



Web API Reference

AWS Glue



API Version 2017-03-31

Copyright © 2024 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

AWS Glue: Web API Reference

Copyright © 2024 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon's trademarks and trade dress may not be used in connection with any product or service that is not Amazon's, in any manner that is likely to cause confusion among customers, or in any manner that disparages or discredits Amazon. All other trademarks not owned by Amazon are the property of their respective owners, who may or may not be affiliated with, connected to, or sponsored by Amazon.

Table of Contents

Welcome to the AWS Glue Web API Reference	1
Actions	2
BatchCreatePartition	10
Request Syntax	10
Request Parameters	11
Response Syntax	12
Response Elements	13
Errors	13
See Also	14
BatchDeleteConnection	15
Request Syntax	15
Request Parameters	15
Response Syntax	16
Response Elements	16
Errors	16
See Also	17
BatchDeletePartition	18
Request Syntax	18
Request Parameters	18
Response Syntax	19
Response Elements	19
Errors	20
See Also	20
BatchDeleteTable	22
Request Syntax	22
Request Parameters	22
Response Syntax	23
Response Elements	24
Errors	24
See Also	25
BatchDeleteTableVersion	26
Request Syntax	26
Request Parameters	26
Response Syntax	27

Response Elements	28
Errors	28
See Also	28
BatchGetBlueprints	30
Request Syntax	30
Request Parameters	30
Response Syntax	31
Response Elements	31
Errors	32
See Also	32
BatchGetCrawlers	34
Request Syntax	34
Request Parameters	34
Response Syntax	34
Response Elements	37
Errors	38
See Also	38
BatchGetCustomEntityTypes	39
Request Syntax	39
Request Parameters	39
Response Syntax	39
Response Elements	40
Errors	40
See Also	41
BatchGetDataQualityResult	42
Request Syntax	42
Request Parameters	42
Response Syntax	42
Response Elements	44
Errors	44
See Also	45
BatchGetDevEndpoints	46
Request Syntax	46
Request Parameters	46
Response Syntax	46
Response Elements	47

Errors	48
See Also	48
BatchGetJobs	50
Request Syntax	50
Request Parameters	50
Response Syntax	50
Response Elements	80
Errors	80
See Also	81
BatchGetPartition	82
Request Syntax	82
Request Parameters	82
Response Syntax	83
Response Elements	85
Errors	85
See Also	86
BatchGetTableOptimizer	88
Request Syntax	88
Request Parameters	88
Response Syntax	88
Response Elements	89
Errors	90
See Also	90
BatchGetTriggers	91
Request Syntax	91
Request Parameters	91
Response Syntax	91
Response Elements	92
Errors	93
See Also	93
BatchGetWorkflows	95
Request Syntax	95
Request Parameters	95
Response Syntax	96
Response Elements	102
Errors	102

See Also	103
BatchStopJobRun	104
Request Syntax	104
Request Parameters	104
Response Syntax	105
Response Elements	105
Errors	105
See Also	106
BatchUpdatePartition	107
Request Syntax	107
Request Parameters	108
Response Syntax	110
Response Elements	110
Errors	110
See Also	111
CancelDataQualityRuleRecommendationRun	112
Request Syntax	112
Request Parameters	112
Response Elements	112
Errors	112
See Also	113
CancelDataQualityRulesetEvaluationRun	114
Request Syntax	114
Request Parameters	114
Response Elements	114
Errors	114
See Also	115
CancelMLTaskRun	116
Request Syntax	116
Request Parameters	116
Response Syntax	117
Response Elements	117
Errors	118
See Also	118
CancelStatement	120
Request Syntax	120

Request Parameters	120
Response Elements	121
Errors	121
See Also	122
CheckSchemaVersionValidity	123
Request Syntax	123
Request Parameters	123
Response Syntax	124
Response Elements	124
Errors	124
See Also	125
CreateBlueprint	126
Request Syntax	126
Request Parameters	126
Response Syntax	127
Response Elements	127
Errors	128
See Also	128
CreateClassifier	130
Request Syntax	130
Request Parameters	130
Response Elements	131
Errors	131
See Also	132
CreateConnection	133
Request Syntax	133
Request Parameters	134
Response Syntax	135
Response Elements	135
Errors	135
See Also	136
CreateCrawler	137
Request Syntax	137
Request Parameters	139
Response Elements	142
Errors	142

See Also	143
CreateCustomEntityType	144
Request Syntax	144
Request Parameters	144
Response Syntax	145
Response Elements	146
Errors	146
See Also	147
CreateDatabase	148
Request Syntax	148
Request Parameters	148
Response Elements	149
Errors	149
See Also	151
CreateDataQualityRuleset	152
Request Syntax	152
Request Parameters	152
Response Syntax	154
Response Elements	154
Errors	154
See Also	155
CreateDevEndpoint	156
Request Syntax	156
Request Parameters	156
Response Syntax	161
Response Elements	161
Errors	164
See Also	166
CreateJob	167
Request Syntax	167
Request Parameters	196
Response Syntax	204
Response Elements	204
Errors	204
See Also	205
CreateMLTransform	206

Request Syntax	206
Request Parameters	207
Response Syntax	211
Response Elements	211
Errors	212
See Also	213
CreatePartition	214
Request Syntax	214
Request Parameters	215
Response Elements	216
Errors	216
See Also	217
CreatePartitionIndex	219
Request Syntax	219
Request Parameters	219
Response Elements	220
Errors	220
See Also	221
CreateRegistry	222
Request Syntax	222
Request Parameters	222
Response Syntax	223
Response Elements	223
Errors	224
See Also	225
CreateSchema	226
Request Syntax	226
Request Parameters	226
Response Syntax	229
Response Elements	230
Errors	232
See Also	233
CreateScript	235
Request Syntax	235
Request Parameters	235
Response Syntax	236

Response Elements	236
Errors	237
See Also	237
CreateSecurityConfiguration	238
Request Syntax	238
Request Parameters	238
Response Syntax	239
Response Elements	239
Errors	240
See Also	240
CreateSession	242
Request Syntax	242
Request Parameters	242
Response Syntax	247
Response Elements	248
Errors	248
See Also	249
CreateTable	250
Request Syntax	250
Request Parameters	252
Response Elements	254
Errors	254
See Also	255
CreateTableOptimizer	257
Request Syntax	257
Request Parameters	257
Response Elements	258
Errors	258
See Also	259
CreateTrigger	261
Request Syntax	261
Request Parameters	262
Response Syntax	264
Response Elements	264
Errors	265
See Also	266

CreateUsageProfile	267
Request Syntax	267
Request Parameters	267
Response Syntax	268
Response Elements	269
Errors	269
See Also	270
CreateUserDefinedFunction	271
Request Syntax	271
Request Parameters	271
Response Elements	272
Errors	272
See Also	273
CreateWorkflow	274
Request Syntax	274
Request Parameters	274
Response Syntax	275
Response Elements	276
Errors	276
See Also	277
DeleteBlueprint	278
Request Syntax	278
Request Parameters	278
Response Syntax	278
Response Elements	278
Errors	279
See Also	279
DeleteClassifier	281
Request Syntax	281
Request Parameters	281
Response Elements	281
Errors	281
See Also	282
DeleteColumnStatisticsForPartition	283
Request Syntax	283
Request Parameters	283

Response Elements	284
Errors	284
See Also	285
DeleteColumnStatisticsForTable	287
Request Syntax	287
Request Parameters	287
Response Elements	288
Errors	288
See Also	289
DeleteConnection	290
Request Syntax	290
Request Parameters	290
Response Elements	291
Errors	291
See Also	291
DeleteCrawler	292
Request Syntax	292
Request Parameters	292
Response Elements	292
Errors	292
See Also	293
DeleteCustomEntityType	294
Request Syntax	294
Request Parameters	294
Response Syntax	294
Response Elements	294
Errors	295
See Also	296
DeleteDatabase	297
Request Syntax	297
Request Parameters	297
Response Elements	298
Errors	298
See Also	299
DeleteDataQualityRuleset	300
Request Syntax	300

Request Parameters	300
Response Elements	300
Errors	300
See Also	301
DeleteDevEndpoint	302
Request Syntax	302
Request Parameters	302
Response Elements	302
Errors	302
See Also	303
DeleteJob	304
Request Syntax	304
Request Parameters	304
Response Syntax	304
Response Elements	304
Errors	305
See Also	305
DeleteMLTransform	307
Request Syntax	307
Request Parameters	307
Response Syntax	307
Response Elements	308
Errors	308
See Also	309
DeletePartition	310
Request Syntax	310
Request Parameters	310
Response Elements	311
Errors	311
See Also	312
DeletePartitionIndex	313
Request Syntax	313
Request Parameters	313
Response Elements	314
Errors	314
See Also	315

DeleteRegistry	316
Request Syntax	316
Request Parameters	316
Response Syntax	316
Response Elements	317
Errors	317
See Also	318
DeleteResourcePolicy	319
Request Syntax	319
Request Parameters	319
Response Elements	320
Errors	320
See Also	320
DeleteSchema	322
Request Syntax	322
Request Parameters	322
Response Syntax	322
Response Elements	323
Errors	323
See Also	324
DeleteSchemaVersions	325
Request Syntax	325
Request Parameters	325
Response Syntax	326
Response Elements	326
Errors	327
See Also	327
DeleteSecurityConfiguration	329
Request Syntax	329
Request Parameters	329
Response Elements	329
Errors	329
See Also	330
DeleteSession	331
Request Syntax	331
Request Parameters	331

Response Syntax	332
Response Elements	332
Errors	332
See Also	333
DeleteTable	334
Request Syntax	334
Request Parameters	334
Response Elements	335
Errors	335
See Also	336
DeleteTableOptimizer	338
Request Syntax	338
Request Parameters	338
Response Elements	339
Errors	339
See Also	340
DeleteTableVersion	341
Request Syntax	341
Request Parameters	341
Response Elements	342
Errors	342
See Also	343
DeleteTrigger	344
Request Syntax	344
Request Parameters	344
Response Syntax	344
Response Elements	344
Errors	345
See Also	345
DeleteUsageProfile	347
Request Syntax	347
Request Parameters	347
Response Elements	347
Errors	347
See Also	348
DeleteUserDefinedFunction	349

Request Syntax	349
Request Parameters	349
Response Elements	350
Errors	350
See Also	351
DeleteWorkflow	352
Request Syntax	352
Request Parameters	352
Response Syntax	352
Response Elements	352
Errors	353
See Also	353
GetBlueprint	355
Request Syntax	355
Request Parameters	355
Response Syntax	356
Response Elements	356
Errors	356
See Also	357
GetBlueprintRun	358
Request Syntax	358
Request Parameters	358
Response Syntax	359
Response Elements	359
Errors	359
See Also	360
GetBlueprintRuns	361
Request Syntax	361
Request Parameters	361
Response Syntax	362
Response Elements	362
Errors	363
See Also	363
GetCatalogImportStatus	365
Request Syntax	365
Request Parameters	365

Response Syntax	365
Response Elements	365
Errors	366
See Also	366
GetClassifier	368
Request Syntax	368
Request Parameters	368
Response Syntax	368
Response Elements	369
Errors	370
See Also	370
GetClassifiers	371
Request Syntax	371
Request Parameters	371
Response Syntax	371
Response Elements	373
Errors	373
See Also	373
GetColumnStatisticsForPartition	375
Request Syntax	375
Request Parameters	375
Response Syntax	376
Response Elements	378
Errors	378
See Also	379
GetColumnStatisticsForTable	380
Request Syntax	380
Request Parameters	380
Response Syntax	381
Response Elements	383
Errors	383
See Also	384
GetColumnStatisticsTaskRun	385
Request Syntax	385
Request Parameters	385
Response Syntax	385

Response Elements	386
Errors	386
See Also	387
GetColumnStatisticsTaskRuns	388
Request Syntax	388
Request Parameters	388
Response Syntax	389
Response Elements	389
Errors	390
See Also	390
GetConnection	392
Request Syntax	392
Request Parameters	392
Response Syntax	393
Response Elements	394
Errors	394
See Also	395
GetConnections	396
Request Syntax	396
Request Parameters	396
Response Syntax	397
Response Elements	398
Errors	399
See Also	399
GetCrawler	401
Request Syntax	401
Request Parameters	401
Response Syntax	401
Response Elements	404
Errors	404
See Also	404
GetCrawlerMetrics	406
Request Syntax	406
Request Parameters	406
Response Syntax	407
Response Elements	407

Errors	408
See Also	408
GetCrawlers	409
Request Syntax	409
Request Parameters	409
Response Syntax	409
Response Elements	412
Errors	412
See Also	413
GetCustomEntityType	414
Request Syntax	414
Request Parameters	414
Response Syntax	414
Response Elements	414
Errors	415
See Also	416
GetDatabase	417
Request Syntax	417
Request Parameters	417
Response Syntax	418
Response Elements	418
Errors	419
See Also	419
GetDatabases	421
Request Syntax	421
Request Parameters	421
Response Syntax	422
Response Elements	423
Errors	423
See Also	424
GetDataCatalogEncryptionSettings	425
Request Syntax	425
Request Parameters	425
Response Syntax	425
Response Elements	426
Errors	426

See Also	426
GetDataflowGraph	428
Request Syntax	428
Request Parameters	428
Response Syntax	428
Response Elements	429
Errors	429
See Also	430
GetDataQualityResult	431
Request Syntax	431
Request Parameters	431
Response Syntax	431
Response Elements	433
Errors	435
See Also	436
GetDataQualityRuleRecommendationRun	437
Request Syntax	437
Request Parameters	437
Response Syntax	437
Response Elements	438
Errors	440
See Also	441
GetDataQualityRuleset	442
Request Syntax	442
Request Parameters	442
Response Syntax	442
Response Elements	443
Errors	444
See Also	445
GetDataQualityRulesetEvaluationRun	446
Request Syntax	446
Request Parameters	446
Response Syntax	446
Response Elements	447
Errors	450
See Also	451

GetDevEndpoint	452
Request Syntax	452
Request Parameters	452
Response Syntax	452
Response Elements	453
Errors	453
See Also	454
GetDevEndpoints	455
Request Syntax	455
Request Parameters	455
Response Syntax	456
Response Elements	456
Errors	457
See Also	457
GetJob	459
Request Syntax	459
Request Parameters	459
Response Syntax	459
Response Elements	488
Errors	489
See Also	489
GetJobBookmark	491
Request Syntax	491
Request Parameters	491
Response Syntax	492
Response Elements	492
Errors	492
See Also	493
GetJobRun	494
Request Syntax	494
Request Parameters	494
Response Syntax	495
Response Elements	496
Errors	496
See Also	497
GetJobRuns	498

Request Syntax	498
Request Parameters	498
Response Syntax	499
Response Elements	500
Errors	500
See Also	501
GetJobs	502
Request Syntax	502
Request Parameters	502
Response Syntax	502
Response Elements	532
Errors	532
See Also	533
GetMapping	534
Request Syntax	534
Request Parameters	535
Response Syntax	535
Response Elements	536
Errors	536
See Also	536
GetMLTaskRun	538
Request Syntax	538
Request Parameters	538
Response Syntax	539
Response Elements	539
Errors	541
See Also	542
GetMLTaskRuns	543
Request Syntax	543
Request Parameters	543
Response Syntax	544
Response Elements	545
Errors	546
See Also	546
GetMLTransform	548
Request Syntax	548

Request Parameters	548
Response Syntax	548
Response Elements	550
Errors	554
See Also	554
GetMLTransforms	556
Request Syntax	556
Request Parameters	556
Response Syntax	557
Response Elements	559
Errors	559
See Also	560
GetPartition	561
Request Syntax	561
Request Parameters	561
Response Syntax	562
Response Elements	564
Errors	564
See Also	565
GetPartitionIndexes	566
Request Syntax	566
Request Parameters	566
Response Syntax	567
Response Elements	568
Errors	568
See Also	569
GetPartitions	570
Request Syntax	570
Request Parameters	570
Response Syntax	574
Response Elements	576
Errors	576
See Also	577
GetPlan	579
Request Syntax	579
Request Parameters	580

Response Syntax	581
Response Elements	581
Errors	582
See Also	582
GetRegistry	584
Request Syntax	584
Request Parameters	584
Response Syntax	584
Response Elements	585
Errors	586
See Also	586
GetResourcePolicies	588
Request Syntax	588
Request Parameters	588
Response Syntax	588
Response Elements	589
Errors	589
See Also	590
GetResourcePolicy	591
Request Syntax	591
Request Parameters	591
Response Syntax	591
Response Elements	592
Errors	592
See Also	593
GetSchema	594
Request Syntax	594
Request Parameters	594
Response Syntax	594
Response Elements	595
Errors	597
See Also	598
GetSchemaByDefinition	599
Request Syntax	599
Request Parameters	599
Response Syntax	600

Response Elements	600
Errors	601
See Also	602
GetSchemaVersion	603
Request Syntax	603
Request Parameters	603
Response Syntax	604
Response Elements	604
Errors	606
See Also	606
GetSchemaVersionsDiff	608
Request Syntax	608
Request Parameters	608
Response Syntax	609
Response Elements	609
Errors	610
See Also	610
GetSecurityConfiguration	612
Request Syntax	612
Request Parameters	612
Response Syntax	612
Response Elements	613
Errors	613
See Also	614
GetSecurityConfigurations	615
Request Syntax	615
Request Parameters	615
Response Syntax	615
Response Elements	616
Errors	616
See Also	617
GetSession	618
Request Syntax	618
Request Parameters	618
Response Syntax	619
Response Elements	619

Errors	620
See Also	620
GetStatement	622
Request Syntax	622
Request Parameters	622
Response Syntax	623
Response Elements	623
Errors	624
See Also	624
GetTable	626
Request Syntax	626
Request Parameters	626
Response Syntax	627
Response Elements	630
Errors	630
See Also	631
GetTableOptimizer	633
Request Syntax	633
Request Parameters	633
Response Syntax	634
Response Elements	635
Errors	636
See Also	636
GetTables	638
Request Syntax	638
Request Parameters	638
Response Syntax	640
Response Elements	642
Errors	643
See Also	644
GetTableVersion	645
Request Syntax	645
Request Parameters	645
Response Syntax	646
Response Elements	649
Errors	649

See Also	650
GetTableVersions	651
Request Syntax	651
Request Parameters	651
Response Syntax	652
Response Elements	655
Errors	655
See Also	656
GetTags	657
Request Syntax	657
Request Parameters	657
Response Syntax	657
Response Elements	657
Errors	658
See Also	658
GetTrigger	660
Request Syntax	660
Request Parameters	660
Response Syntax	660
Response Elements	661
Errors	662
See Also	662
GetTriggers	664
Request Syntax	664
Request Parameters	664
Response Syntax	665
Response Elements	666
Errors	666
See Also	667
GetUnfilteredPartitionMetadata	668
Request Syntax	668
Request Parameters	668
Response Syntax	670
Response Elements	672
Errors	672
See Also	674

GetUnfilteredPartitionsMetadata	675
Request Syntax	675
Request Parameters	675
Response Syntax	680
Response Elements	681
Errors	682
See Also	683
GetUnfilteredTableMetadata	684
Request Syntax	684
Request Parameters	684
Response Syntax	688
Response Elements	691
Errors	693
See Also	694
GetUsageProfile	695
Request Syntax	695
Request Parameters	695
Response Syntax	695
Response Elements	696
Errors	697
See Also	698
GetUserDefinedFunction	699
Request Syntax	699
Request Parameters	699
Response Syntax	700
Response Elements	700
Errors	701
See Also	701
GetUserDefinedFunctions	703
Request Syntax	703
Request Parameters	703
Response Syntax	704
Response Elements	705
Errors	705
See Also	706
GetWorkflow	707

Request Syntax	707
Request Parameters	707
Response Syntax	707
Response Elements	713
Errors	714
See Also	714
GetWorkflowRun	716
Request Syntax	716
Request Parameters	716
Response Syntax	717
Response Elements	720
Errors	720
See Also	721
GetWorkflowRunProperties	722
Request Syntax	722
Request Parameters	722
Response Syntax	723
Response Elements	723
Errors	723
See Also	724
GetWorkflowRuns	725
Request Syntax	725
Request Parameters	725
Response Syntax	726
Response Elements	729
Errors	730
See Also	730
ImportCatalogToGlue	732
Request Syntax	732
Request Parameters	732
Response Elements	732
Errors	732
See Also	733
ListBlueprints	734
Request Syntax	734
Request Parameters	734

Response Syntax	735
Response Elements	735
Errors	735
See Also	736
ListColumnStatisticsTaskRuns	737
Request Syntax	737
Request Parameters	737
Response Syntax	737
Response Elements	738
Errors	738
See Also	738
ListCrawlers	740
Request Syntax	740
Request Parameters	740
Response Syntax	741
Response Elements	741
Errors	742
See Also	742
ListCrawls	743
Request Syntax	743
Request Parameters	743
Response Syntax	744
Response Elements	745
Errors	745
See Also	746
ListCustomEntityTypes	747
Request Syntax	747
Request Parameters	747
Response Syntax	748
Response Elements	748
Errors	748
See Also	749
ListDataQualityResults	750
Request Syntax	750
Request Parameters	750
Response Syntax	751

Response Elements	752
Errors	752
See Also	752
ListDataQualityRuleRecommendationRuns	754
Request Syntax	754
Request Parameters	754
Response Syntax	755
Response Elements	755
Errors	756
See Also	756
ListDataQualityRulesetEvaluationRuns	758
Request Syntax	758
Request Parameters	758
Response Syntax	759
Response Elements	759
Errors	760
See Also	760
ListDataQualityRulesets	762
Request Syntax	762
Request Parameters	762
Response Syntax	763
Response Elements	764
Errors	764
See Also	765
ListDevEndpoints	766
Request Syntax	766
Request Parameters	766
Response Syntax	767
Response Elements	767
Errors	768
See Also	768
ListJobs	770
Request Syntax	770
Request Parameters	770
Response Syntax	771
Response Elements	771

Errors	772
See Also	772
ListMLTransforms	774
Request Syntax	774
Request Parameters	774
Response Syntax	776
Response Elements	776
Errors	776
See Also	777
ListRegistries	778
Request Syntax	778
Request Parameters	778
Response Syntax	778
Response Elements	779
Errors	779
See Also	780
ListSchemas	781
Request Syntax	781
Request Parameters	781
Response Syntax	782
Response Elements	782
Errors	783
See Also	783
ListSchemaVersions	785
Request Syntax	785
Request Parameters	785
Response Syntax	786
Response Elements	786
Errors	787
See Also	787
ListSessions	789
Request Syntax	789
Request Parameters	789
Response Syntax	790
Response Elements	791
Errors	792

See Also	792
ListStatements	794
Request Syntax	794
Request Parameters	794
Response Syntax	795
Response Elements	795
Errors	796
See Also	797
ListTableOptimizerRuns	798
Request Syntax	798
Request Parameters	798
Response Syntax	799
Response Elements	800
Errors	801
See Also	802
ListTriggers	803
Request Syntax	803
Request Parameters	803
Response Syntax	804
Response Elements	804
Errors	805
See Also	805
ListUsageProfiles	807
Request Syntax	807
Request Parameters	807
Response Syntax	807
Response Elements	808
Errors	808
See Also	809
ListWorkflows	810
Request Syntax	810
Request Parameters	810
Response Syntax	810
Response Elements	811
Errors	811
See Also	812

PutDataCatalogEncryptionSettings	813
Request Syntax	813
Request Parameters	813
Response Elements	814
Errors	814
See Also	814
PutResourcePolicy	816
Request Syntax	816
Request Parameters	816
Response Syntax	817
Response Elements	818
Errors	818
See Also	819
PutSchemaVersionMetadata	820
Request Syntax	820
Request Parameters	820
Response Syntax	821
Response Elements	821
Errors	823
See Also	824
PutWorkflowRunProperties	825
Request Syntax	825
Request Parameters	825
Response Elements	826
Errors	826
See Also	827
QuerySchemaVersionMetadata	828
Request Syntax	828
Request Parameters	828
Response Syntax	830
Response Elements	830
Errors	831
See Also	831
RegisterSchemaVersion	833
Request Syntax	833
Request Parameters	833

Response Syntax	834
Response Elements	834
Errors	835
See Also	836
RemoveSchemaVersionMetadata	837
Request Syntax	837
Request Parameters	837
Response Syntax	838
Response Elements	838
Errors	840
See Also	841
ResetJobBookmark	842
Request Syntax	842
Request Parameters	842
Response Syntax	843
Response Elements	843
Errors	843
See Also	844
ResumeWorkflowRun	845
Request Syntax	845
Request Parameters	845
Response Syntax	846
Response Elements	846
Errors	847
See Also	848
RunStatement	849
Request Syntax	849
Request Parameters	849
Response Syntax	850
Response Elements	850
Errors	850
See Also	851
SearchTables	853
Request Syntax	853
Request Parameters	853
Response Syntax	855

Response Elements	858
Errors	859
See Also	859
StartBlueprintRun	860
Request Syntax	860
Request Parameters	860
Response Syntax	861
Response Elements	861
Errors	861
See Also	862
StartColumnStatisticsTaskRun	864
Request Syntax	864
Request Parameters	864
Response Syntax	866
Response Elements	866
Errors	866
See Also	867
StartCrawler	869
Request Syntax	869
Request Parameters	869
Response Elements	869
Errors	869
See Also	870
StartCrawlerSchedule	871
Request Syntax	871
Request Parameters	871
Response Elements	871
Errors	871
See Also	872
StartDataQualityRuleRecommendationRun	873
Request Syntax	873
Request Parameters	873
Response Syntax	875
Response Elements	875
Errors	875
See Also	876

StartDataQualityRulesetEvaluationRun	877
Request Syntax	877
Request Parameters	878
Response Syntax	880
Response Elements	880
Errors	880
See Also	881
StartExportLabelsTaskRun	882
Request Syntax	882
Request Parameters	882
Response Syntax	883
Response Elements	883
Errors	883
See Also	884
StartImportLabelsTaskRun	885
Request Syntax	885
Request Parameters	885
Response Syntax	886
Response Elements	886
Errors	887
See Also	887
StartJobRun	889
Request Syntax	889
Request Parameters	889
Response Syntax	894
Response Elements	894
Errors	894
See Also	895
StartMLEvaluationTaskRun	896
Request Syntax	896
Request Parameters	896
Response Syntax	896
Response Elements	897
Errors	897
See Also	898
StartMLLabelingSetGenerationTaskRun	899

Request Syntax	899
Request Parameters	899
Response Syntax	900
Response Elements	900
Errors	900
See Also	901
StartTrigger	902
Request Syntax	902
Request Parameters	902
Response Syntax	902
Response Elements	902
Errors	903
See Also	904
StartWorkflowRun	905
Request Syntax	905
Request Parameters	905
Response Syntax	906
Response Elements	906
Errors	906
See Also	907
StopColumnStatisticsTaskRun	908
Request Syntax	908
Request Parameters	908
Response Elements	908
Errors	909
See Also	909
StopCrawler	911
Request Syntax	911
Request Parameters	911
Response Elements	911
Errors	911
See Also	912
StopCrawlerSchedule	913
Request Syntax	913
Request Parameters	913
Response Elements	913

Errors	913
See Also	914
StopSession	915
Request Syntax	915
Request Parameters	915
Response Syntax	916
Response Elements	916
Errors	916
See Also	917
StopTrigger	918
Request Syntax	918
Request Parameters	918
Response Syntax	918
Response Elements	918
Errors	919
See Also	920
StopWorkflowRun	921
Request Syntax	921
Request Parameters	921
Response Elements	922
Errors	922
See Also	922
TagResource	924
Request Syntax	924
Request Parameters	924
Response Elements	925
Errors	925
See Also	925
UntagResource	927
Request Syntax	927
Request Parameters	927
Response Elements	928
Errors	928
See Also	928
UpdateBlueprint	930
Request Syntax	930

Request Parameters	930
Response Syntax	931
Response Elements	931
Errors	931
See Also	932
UpdateClassifier	933
Request Syntax	933
Request Parameters	933
Response Elements	934
Errors	934
See Also	935
UpdateColumnStatisticsForPartition	936
Request Syntax	936
Request Parameters	937
Response Syntax	939
Response Elements	940
Errors	940
See Also	941
UpdateColumnStatisticsForTable	942
Request Syntax	942
Request Parameters	943
Response Syntax	944
Response Elements	946
Errors	946
See Also	947
UpdateConnection	948
Request Syntax	948
Request Parameters	949
Response Elements	949
Errors	950
See Also	950
UpdateCrawler	952
Request Syntax	952
Request Parameters	954
Response Elements	957
Errors	957

See Also	958
UpdateCrawlerSchedule	959
Request Syntax	959
Request Parameters	959
Response Elements	959
Errors	960
See Also	960
UpdateDatabase	962
Request Syntax	962
Request Parameters	962
Response Elements	963
Errors	963
See Also	964
UpdateDataQualityRuleset	966
Request Syntax	966
Request Parameters	966
Response Syntax	967
Response Elements	967
Errors	968
See Also	969
UpdateDevEndpoint	970
Request Syntax	970
Request Parameters	970
Response Elements	972
Errors	972
See Also	973
UpdateJob	974
Request Syntax	974
Request Parameters	1003
Response Syntax	1003
Response Elements	1004
Errors	1004
See Also	1005
UpdateJobFromSourceControl	1006
Request Syntax	1006
Request Parameters	1006

Response Syntax	1009
Response Elements	1009
Errors	1009
See Also	1010
UpdateMLTransform	1011
Request Syntax	1011
Request Parameters	1011
Response Syntax	1014
Response Elements	1014
Errors	1015
See Also	1016
UpdatePartition	1017
Request Syntax	1017
Request Parameters	1018
Response Elements	1020
Errors	1020
See Also	1020
UpdateRegistry	1022
Request Syntax	1022
Request Parameters	1022
Response Syntax	1023
Response Elements	1023
Errors	1023
See Also	1024
UpdateSchema	1025
Request Syntax	1025
Request Parameters	1025
Response Syntax	1026
Response Elements	1027
Errors	1027
See Also	1028
UpdateSourceControlFromJob	1030
Request Syntax	1030
Request Parameters	1030
Response Syntax	1033
Response Elements	1033

Errors	1033
See Also	1034
UpdateTable	1035
Request Syntax	1035
Request Parameters	1037
Response Elements	1039
Errors	1039
See Also	1040
UpdateTableOptimizer	1042
Request Syntax	1042
Request Parameters	1042
Response Elements	1043
Errors	1043
See Also	1044
UpdateTrigger	1045
Request Syntax	1045
Request Parameters	1046
Response Syntax	1046
Response Elements	1047
Errors	1047
See Also	1048
UpdateUsageProfile	1050
Request Syntax	1050
Request Parameters	1050
Response Syntax	1051
Response Elements	1051
Errors	1052
See Also	1052
UpdateUserDefinedFunction	1054
Request Syntax	1054
Request Parameters	1054
Response Elements	1055
Errors	1055
See Also	1056
UpdateWorkflow	1057
Request Syntax	1057

Request Parameters	1057
Response Syntax	1058
Response Elements	1058
Errors	1059
See Also	1059
Data Types	1061
Action	1072
Contents	1072
See Also	1073
Aggregate	1075
Contents	1075
See Also	1076
AggregateOperation	1077
Contents	1077
See Also	1077
AmazonRedshiftAdvancedOption	1079
Contents	1079
See Also	1079
AmazonRedshiftNodeData	1080
Contents	1080
See Also	1085
AmazonRedshiftSource	1086
Contents	1086
See Also	1086
AmazonRedshiftTarget	1087
Contents	1087
See Also	1087
ApplyMapping	1089
Contents	1089
See Also	1089
AthenaConnectorSource	1091
Contents	1091
See Also	1092
AuditContext	1094
Contents	1094
See Also	1094

AuthenticationConfiguration	1096
Contents	1096
See Also	1096
AuthenticationConfigurationInput	1098
Contents	1098
See Also	1098
AuthorizationCodeProperties	1100
Contents	1100
See Also	1100
BackfillError	1102
Contents	1102
See Also	1103
BasicCatalogTarget	1104
Contents	1104
See Also	1105
BatchGetTableOptimizerEntry	1106
Contents	1106
See Also	1107
BatchGetTableOptimizerError	1108
Contents	1108
See Also	1109
BatchStopJobRunError	1110
Contents	1110
See Also	1110
BatchStopJobRunSuccessfulSubmission	1112
Contents	1112
See Also	1112
BatchTableOptimizer	1113
Contents	1113
See Also	1114
BatchUpdatePartitionFailureEntry	1115
Contents	1115
See Also	1115
BatchUpdatePartitionRequestEntry	1116
Contents	1116
See Also	1116

BinaryColumnStatisticsData	1117
Contents	1117
See Also	1117
Blueprint	1119
Contents	1119
See Also	1121
BlueprintDetails	1122
Contents	1122
See Also	1122
BlueprintRun	1123
Contents	1123
See Also	1125
BooleanColumnStatisticsData	1126
Contents	1126
See Also	1126
CatalogDeltaSource	1128
Contents	1128
See Also	1129
CatalogEntry	1130
Contents	1130
See Also	1130
CatalogHudiSource	1131
Contents	1131
See Also	1132
CatalogImportStatus	1133
Contents	1133
See Also	1133
CatalogKafkaSource	1135
Contents	1135
See Also	1136
CatalogKinesisSource	1137
Contents	1137
See Also	1138
CatalogSchemaChangePolicy	1139
Contents	1139
See Also	1139

CatalogSource	1140
Contents	1140
See Also	1140
CatalogTarget	1142
Contents	1142
See Also	1143
Classifier	1144
Contents	1144
See Also	1145
CloudWatchEncryption	1146
Contents	1146
See Also	1146
CodeGenConfigurationNode	1147
Contents	1147
See Also	1159
CodeGenEdge	1160
Contents	1160
See Also	1160
CodeGenNode	1162
Contents	1162
See Also	1163
CodeGenNodeArg	1164
Contents	1164
See Also	1164
Column	1165
Contents	1165
See Also	1166
ColumnError	1167
Contents	1167
See Also	1167
ColumnImportance	1168
Contents	1168
See Also	1168
ColumnRowFilter	1169
Contents	1169
See Also	1169

ColumnStatistics	1170
Contents	1170
See Also	1171
ColumnStatisticsData	1172
Contents	1172
See Also	1173
ColumnStatisticsError	1174
Contents	1174
See Also	1174
ColumnStatisticsTaskRun	1175
Contents	1175
See Also	1179
Condition	1180
Contents	1180
See Also	1181
ConfigurationObject	1182
Contents	1182
See Also	1183
ConfusionMatrix	1184
Contents	1184
See Also	1185
Connection	1186
Contents	1186
See Also	1192
ConnectionInput	1193
Contents	1193
See Also	1197
ConnectionPasswordEncryption	1198
Contents	1198
See Also	1199
ConnectionsList	1200
Contents	1200
See Also	1200
ConnectorDataSource	1201
Contents	1201
See Also	1202

ConnectorDataTarget	1203
Contents	1203
See Also	1204
Crawl	1205
Contents	1205
See Also	1206
Crawler	1207
Contents	1207
See Also	1211
CrawlerHistory	1212
Contents	1212
See Also	1214
CrawlerMetrics	1215
Contents	1215
See Also	1216
CrawlerNodeDetails	1218
Contents	1218
See Also	1218
CrawlerTargets	1219
Contents	1219
See Also	1220
CrawlsFilter	1221
Contents	1221
See Also	1222
CreateCsvClassifierRequest	1223
Contents	1223
See Also	1225
CreateGrokClassifierRequest	1226
Contents	1226
See Also	1227
CreateJsonClassifierRequest	1228
Contents	1228
See Also	1228
CreateXMLClassifierRequest	1229
Contents	1229
See Also	1229

CsvClassifier	1231
Contents	1231
See Also	1234
CustomCode	1235
Contents	1235
See Also	1236
CustomEntityType	1237
Contents	1237
See Also	1238
Database	1239
Contents	1239
See Also	1241
Databaselfentifier	1242
Contents	1242
See Also	1243
Databaselfinput	1244
Contents	1244
See Also	1245
DataCatalogEncryptionSettings	1247
Contents	1247
See Also	1247
DataLakePrincipal	1248
Contents	1248
See Also	1248
DataQualityAnalyzerResult	1249
Contents	1249
See Also	1250
DataQualityEvaluationRunAdditionalRunOptions	1251
Contents	1251
See Also	1251
DataQualityMetricValues	1252
Contents	1252
See Also	1252
DataQualityObservation	1254
Contents	1254
See Also	1254

DataQualityResult	1255
Contents	1255
See Also	1258
DataQualityResultDescription	1259
Contents	1259
See Also	1260
DataQualityResultFilterCriteria	1261
Contents	1261
See Also	1262
DataQualityRuleRecommendationRunDescription	1263
Contents	1263
See Also	1264
DataQualityRuleRecommendationRunFilter	1265
Contents	1265
See Also	1265
DataQualityRuleResult	1266
Contents	1266
See Also	1267
DataQualityRulesetEvaluationRunDescription	1268
Contents	1268
See Also	1269
DataQualityRulesetEvaluationRunFilter	1270
Contents	1270
See Also	1270
DataQualityRulesetFilterCriteria	1271
Contents	1271
See Also	1272
DataQualityRulesetListDetails	1273
Contents	1273
See Also	1274
DataQualityTargetTable	1275
Contents	1275
See Also	1276
DataSource	1277
Contents	1277
See Also	1277

Datatype	1278
Contents	1278
See Also	1278
DateColumnStatisticsData	1279
Contents	1279
See Also	1279
DecimalColumnStatisticsData	1281
Contents	1281
See Also	1281
DecimalNumber	1283
Contents	1283
See Also	1283
DeltaTarget	1284
Contents	1284
See Also	1284
DevEndpoint	1286
Contents	1286
See Also	1292
DevEndpointCustomLibraries	1293
Contents	1293
See Also	1293
DirectJDBCSource	1295
Contents	1295
See Also	1296
DirectKafkaSource	1297
Contents	1297
See Also	1298
DirectKinesisSource	1299
Contents	1299
See Also	1300
DirectSchemaChangePolicy	1301
Contents	1301
See Also	1302
DoubleColumnStatisticsData	1303
Contents	1303
See Also	1303

DQResultsPublishingOptions	1305
Contents	1305
See Also	1306
DQStopJobOnFailureOptions	1307
Contents	1307
See Also	1307
DropDuplicates	1308
Contents	1308
See Also	1308
DropFields	1310
Contents	1310
See Also	1310
DropNullFields	1312
Contents	1312
See Also	1313
DynamicTransform	1314
Contents	1314
See Also	1316
DynamoDBCatalogSource	1317
Contents	1317
See Also	1317
DynamoDBTarget	1319
Contents	1319
See Also	1320
Edge	1321
Contents	1321
See Also	1321
EncryptionAtRest	1322
Contents	1322
See Also	1322
EncryptionConfiguration	1324
Contents	1324
See Also	1324
ErrorDetail	1325
Contents	1325
See Also	1325

ErrorDetails	1326
Contents	1326
See Also	1326
EvaluateDataQuality	1327
Contents	1327
See Also	1328
EvaluateDataQualityMultiFrame	1329
Contents	1329
See Also	1330
EvaluationMetrics	1331
Contents	1331
See Also	1331
EventBatchingCondition	1332
Contents	1332
See Also	1332
ExecutionProperty	1333
Contents	1333
See Also	1333
ExportLabelsTaskRunProperties	1334
Contents	1334
See Also	1334
FederatedDatabase	1335
Contents	1335
See Also	1335
FederatedTable	1336
Contents	1336
See Also	1337
FillMissingValues	1338
Contents	1338
See Also	1339
Filter	1340
Contents	1340
See Also	1341
FilterExpression	1342
Contents	1342
See Also	1342

FilterValue	1343
Contents	1343
See Also	1343
FindMatchesMetrics	1344
Contents	1344
See Also	1345
FindMatchesParameters	1347
Contents	1347
See Also	1348
FindMatchesTaskRunProperties	1349
Contents	1349
See Also	1350
GetConnectionsFilter	1351
Contents	1351
See Also	1351
GluePolicy	1352
Contents	1352
See Also	1353
GlueSchema	1354
Contents	1354
See Also	1354
GlueStudioSchemaColumn	1355
Contents	1355
See Also	1355
GlueTable	1356
Contents	1356
See Also	1357
GovernedCatalogSource	1358
Contents	1358
See Also	1359
GovernedCatalogTarget	1360
Contents	1360
See Also	1361
GrokClassifier	1362
Contents	1362
See Also	1363

HudiTarget	1364
Contents	1364
See Also	1365
IcebergInput	1366
Contents	1366
See Also	1366
IcebergTarget	1367
Contents	1367
See Also	1367
ImportLabelsTaskRunProperties	1369
Contents	1369
See Also	1369
JDBCConnectorOptions	1370
Contents	1370
See Also	1372
JDBCConnectorSource	1373
Contents	1373
See Also	1374
JDBCConnectorTarget	1376
Contents	1376
See Also	1378
JdbcTarget	1379
Contents	1379
See Also	1380
Job	1381
Contents	1381
See Also	1389
JobBookmarkEntry	1390
Contents	1390
See Also	1391
JobBookmarksEncryption	1392
Contents	1392
See Also	1392
JobCommand	1393
Contents	1393
See Also	1394

JobNodeDetails	1395
Contents	1395
See Also	1395
JobRun	1396
Contents	1396
See Also	1404
JobUpdate	1405
Contents	1405
See Also	1412
Join	1413
Contents	1413
See Also	1414
JoinColumn	1415
Contents	1415
See Also	1415
JsonClassifier	1416
Contents	1416
See Also	1417
KafkaStreamingSourceOptions	1418
Contents	1418
See Also	1423
KeySchemaElement	1424
Contents	1424
See Also	1424
KinesisStreamingSourceOptions	1425
Contents	1425
See Also	1430
LabelingSetGenerationTaskRunProperties	1431
Contents	1431
See Also	1431
LakeFormationConfiguration	1432
Contents	1432
See Also	1432
LastActiveDefinition	1433
Contents	1433
See Also	1434

LastCrawlInfo	1435
Contents	1435
See Also	1436
LineageConfiguration	1437
Contents	1437
See Also	1437
Location	1438
Contents	1438
See Also	1438
LongColumnStatisticsData	1440
Contents	1440
See Also	1440
Mapping	1442
Contents	1442
See Also	1443
MappingEntry	1445
Contents	1445
See Also	1446
Merge	1447
Contents	1447
See Also	1448
MetadataInfo	1449
Contents	1449
See Also	1449
MetadataKeyValuePair	1451
Contents	1451
See Also	1451
MetricBasedObservation	1452
Contents	1452
See Also	1452
MicrosoftSQLServerCatalogSource	1454
Contents	1454
See Also	1454
MicrosoftSQLServerCatalogTarget	1456
Contents	1456
See Also	1457

MLTransform	1458
Contents	1458
See Also	1463
MLUserDataEncryption	1464
Contents	1464
See Also	1464
MongoDBTarget	1465
Contents	1465
See Also	1465
MySQLCatalogSource	1467
Contents	1467
See Also	1467
MySQLCatalogTarget	1469
Contents	1469
See Also	1470
Node	1471
Contents	1471
See Also	1472
NotificationProperty	1473
Contents	1473
See Also	1473
NullCheckBoxList	1474
Contents	1474
See Also	1474
NullValueField	1475
Contents	1475
See Also	1475
OAuth2ClientApplication	1476
Contents	1476
See Also	1476
OAuth2Properties	1477
Contents	1477
See Also	1478
OAuth2PropertiesInput	1479
Contents	1479
See Also	1480

OpenTableFormatInput	1481
Contents	1481
See Also	1481
Option	1482
Contents	1482
See Also	1482
OracleSQLCatalogSource	1484
Contents	1484
See Also	1484
OracleSQLCatalogTarget	1486
Contents	1486
See Also	1487
Order	1488
Contents	1488
See Also	1488
OtherMetadataValueListItem	1489
Contents	1489
See Also	1489
Partition	1490
Contents	1490
See Also	1492
PartitionError	1493
Contents	1493
See Also	1493
PartitionIndex	1494
Contents	1494
See Also	1494
PartitionIndexDescriptor	1495
Contents	1495
See Also	1496
PartitionInput	1497
Contents	1497
See Also	1498
PartitionValueList	1499
Contents	1499
See Also	1499

PhysicalConnectionRequirements	1500
Contents	1500
See Also	1501
PIIDetection	1502
Contents	1502
See Also	1504
PostgreSQLCatalogSource	1505
Contents	1505
See Also	1505
PostgreSQLCatalogTarget	1507
Contents	1507
See Also	1508
Predecessor	1509
Contents	1509
See Also	1509
Predicate	1510
Contents	1510
See Also	1510
PrincipalPermissions	1511
Contents	1511
See Also	1511
ProfileConfiguration	1512
Contents	1512
See Also	1512
PropertyPredicate	1513
Contents	1513
See Also	1513
QuerySessionContext	1515
Contents	1515
See Also	1516
Recipe	1517
Contents	1517
See Also	1517
RecipeReference	1519
Contents	1519
See Also	1519

RecrawlPolicy	1520
Contents	1520
See Also	1520
RedshiftSource	1521
Contents	1521
See Also	1522
RedshiftTarget	1523
Contents	1523
See Also	1524
RegistryId	1526
Contents	1526
See Also	1526
RegistryListItem	1527
Contents	1527
See Also	1528
RelationalCatalogSource	1529
Contents	1529
See Also	1529
RenameField	1531
Contents	1531
See Also	1532
ResourceUri	1533
Contents	1533
See Also	1533
RunMetrics	1534
Contents	1534
See Also	1534
S3CatalogDeltaSource	1536
Contents	1536
See Also	1537
S3CatalogHudiSource	1538
Contents	1538
See Also	1539
S3CatalogSource	1540
Contents	1540
See Also	1541

S3CatalogTarget	1542
Contents	1542
See Also	1543
S3CsvSource	1544
Contents	1544
See Also	1548
S3DeltaCatalogTarget	1549
Contents	1549
See Also	1550
S3DeltaDirectTarget	1552
Contents	1552
See Also	1554
S3DeltaSource	1555
Contents	1555
See Also	1556
S3DirectSourceAdditionalOptions	1557
Contents	1557
See Also	1557
S3DirectTarget	1559
Contents	1559
See Also	1560
S3Encryption	1561
Contents	1561
See Also	1561
S3GlueParquetTarget	1562
Contents	1562
See Also	1563
S3HudiCatalogTarget	1564
Contents	1564
See Also	1565
S3HudiDirectTarget	1567
Contents	1567
See Also	1569
S3HudiSource	1570
Contents	1570
See Also	1571

S3JsonSource	1572
Contents	1572
See Also	1575
S3ParquetSource	1576
Contents	1576
See Also	1578
S3SourceAdditionalOptions	1579
Contents	1579
See Also	1579
S3Target	1580
Contents	1580
See Also	1581
Schedule	1582
Contents	1582
See Also	1582
SchemaChangePolicy	1583
Contents	1583
See Also	1583
SchemaColumn	1584
Contents	1584
See Also	1584
Schemald	1585
Contents	1585
See Also	1586
SchemaListItem	1587
Contents	1587
See Also	1588
SchemaReference	1589
Contents	1589
See Also	1589
SchemaVersionErrorItem	1591
Contents	1591
See Also	1591
SchemaVersionListItem	1592
Contents	1592
See Also	1593

SchemaVersionNumber	1594
Contents	1594
See Also	1594
SecurityConfiguration	1595
Contents	1595
See Also	1595
Segment	1597
Contents	1597
See Also	1597
SelectFields	1598
Contents	1598
See Also	1598
SelectFromCollection	1600
Contents	1600
See Also	1600
SerDelInfo	1602
Contents	1602
See Also	1603
Session	1604
Contents	1604
See Also	1608
SessionCommand	1609
Contents	1609
See Also	1609
SkewedInfo	1610
Contents	1610
See Also	1610
SnowflakeNodeData	1612
Contents	1612
See Also	1616
SnowflakeSource	1617
Contents	1617
See Also	1617
SnowflakeTarget	1618
Contents	1618
See Also	1618

SortCriterion	1620
Contents	1620
See Also	1620
SourceControlDetails	1621
Contents	1621
See Also	1622
SparkConnectorSource	1624
Contents	1624
See Also	1625
SparkConnectorTarget	1626
Contents	1626
See Also	1627
SparkSQL	1629
Contents	1629
See Also	1630
Spigot	1631
Contents	1631
See Also	1632
SplitFields	1633
Contents	1633
See Also	1634
SqlAlias	1635
Contents	1635
See Also	1635
StartingEventBatchCondition	1636
Contents	1636
See Also	1636
Statement	1637
Contents	1637
See Also	1638
StatementOutput	1639
Contents	1639
See Also	1640
StatementOutputData	1641
Contents	1641
See Also	1641

StorageDescriptor	1642
Contents	1642
See Also	1645
StreamingDataPreviewOptions	1646
Contents	1646
See Also	1646
StringColumnStatisticsData	1647
Contents	1647
See Also	1648
SupportedDialect	1649
Contents	1649
See Also	1649
Table	1650
Contents	1650
See Also	1655
TableError	1656
Contents	1656
See Also	1656
TableIdentifier	1657
Contents	1657
See Also	1658
TableInput	1659
Contents	1659
See Also	1662
TableOptimizer	1663
Contents	1663
See Also	1663
TableOptimizerConfiguration	1665
Contents	1665
See Also	1665
TableOptimizerRun	1666
Contents	1666
See Also	1667
TableVersion	1668
Contents	1668
See Also	1668

TableVersionError	1669
Contents	1669
See Also	1669
TaskRun	1671
Contents	1671
See Also	1673
TaskRunFilterCriteria	1674
Contents	1674
See Also	1675
TaskRunProperties	1676
Contents	1676
See Also	1677
TaskRunSortCriteria	1678
Contents	1678
See Also	1678
TransformConfigParameter	1679
Contents	1679
See Also	1680
TransformEncryption	1681
Contents	1681
See Also	1681
TransformFilterCriteria	1682
Contents	1682
See Also	1684
TransformParameters	1685
Contents	1685
See Also	1685
TransformSortCriteria	1686
Contents	1686
See Also	1686
Trigger	1687
Contents	1687
See Also	1689
TriggerNodeDetails	1690
Contents	1690
See Also	1690

TriggerUpdate	1691
Contents	1691
See Also	1692
UnfilteredPartition	1693
Contents	1693
See Also	1693
Union	1695
Contents	1695
See Also	1696
UpdateCsvClassifierRequest	1697
Contents	1697
See Also	1699
UpdateGrokClassifierRequest	1700
Contents	1700
See Also	1701
UpdateJsonClassifierRequest	1702
Contents	1702
See Also	1702
UpdateXMLClassifierRequest	1703
Contents	1703
See Also	1703
UpsertRedshiftTargetOptions	1705
Contents	1705
See Also	1705
UsageProfileDefinition	1707
Contents	1707
See Also	1708
UserDefinedFunction	1709
Contents	1709
See Also	1711
UserDefinedFunctionInput	1712
Contents	1712
See Also	1713
ViewDefinition	1714
Contents	1714
See Also	1715

ViewDefinitionInput	1716
Contents	1716
See Also	1717
ViewRepresentation	1718
Contents	1718
See Also	1719
ViewRepresentationInput	1720
Contents	1720
See Also	1721
Workflow	1722
Contents	1722
See Also	1724
WorkflowGraph	1725
Contents	1725
See Also	1725
WorkflowRun	1726
Contents	1726
See Also	1728
WorkflowRunStatistics	1729
Contents	1729
See Also	1730
XMLClassifier	1731
Contents	1731
See Also	1732
Common Parameters	1733
Common Errors	1736

Welcome to the AWS Glue Web API Reference

AWS Glue is a fully managed ETL (extract, transform, and load) service that makes it simple and cost-effective to categorize your data, clean it, enrich it, and move it reliably between various data stores. AWS Glue consists of a central metadata repository known as the AWS Glue Data Catalog, an ETL engine that automatically generates Python code, and a flexible scheduler that handles dependency resolution, job monitoring, and retries. AWS Glue is serverless, so there's no infrastructure to set up or manage.

Actions

The following actions are supported:

- [BatchCreatePartition](#)
- [BatchDeleteConnection](#)
- [BatchDeletePartition](#)
- [BatchDeleteTable](#)
- [BatchDeleteTableVersion](#)
- [BatchGetBlueprints](#)
- [BatchGetCrawlers](#)
- [BatchGetCustomEntityTypes](#)
- [BatchGetDataQualityResult](#)
- [BatchGetDevEndpoints](#)
- [BatchGetJobs](#)
- [BatchGetPartition](#)
- [BatchGetTableOptimizer](#)
- [BatchGetTriggers](#)
- [BatchGetWorkflows](#)
- [BatchStopJobRun](#)
- [BatchUpdatePartition](#)
- [CancelDataQualityRuleRecommendationRun](#)
- [CancelDataQualityRulesetEvaluationRun](#)
- [CancelMLTaskRun](#)
- [CancelStatement](#)
- [CheckSchemaVersionValidity](#)
- [CreateBlueprint](#)
- [CreateClassifier](#)
- [CreateConnection](#)
- [CreateCrawler](#)
- [CreateCustomEntityType](#)

- [CreateDatabase](#)
- [CreateDataQualityRuleset](#)
- [CreateDevEndpoint](#)
- [CreateJob](#)
- [CreateMLTransform](#)
- [CreatePartition](#)
- [CreatePartitionIndex](#)
- [CreateRegistry](#)
- [CreateSchema](#)
- [CreateScript](#)
- [CreateSecurityConfiguration](#)
- [CreateSession](#)
- [CreateTable](#)
- [CreateTableOptimizer](#)
- [CreateTrigger](#)
- [CreateUsageProfile](#)
- [CreateUserDefinedFunction](#)
- [CreateWorkflow](#)
- [DeleteBlueprint](#)
- [DeleteClassifier](#)
- [DeleteColumnStatisticsForPartition](#)
- [DeleteColumnStatisticsForTable](#)
- [DeleteConnection](#)
- [DeleteCrawler](#)
- [DeleteCustomEntityType](#)
- [DeleteDatabase](#)
- [DeleteDataQualityRuleset](#)
- [DeleteDevEndpoint](#)
- [DeleteJob](#)
- [DeleteMLTransform](#)

- [DeletePartition](#)
- [DeletePartitionIndex](#)
- [DeleteRegistry](#)
- [DeleteResourcePolicy](#)
- [DeleteSchema](#)
- [DeleteSchemaVersions](#)
- [DeleteSecurityConfiguration](#)
- [DeleteSession](#)
- [DeleteTable](#)
- [DeleteTableOptimizer](#)
- [DeleteTableVersion](#)
- [DeleteTrigger](#)
- [DeleteUsageProfile](#)
- [DeleteUserDefinedFunction](#)
- [DeleteWorkflow](#)
- [GetBlueprint](#)
- [GetBlueprintRun](#)
- [GetBlueprintRuns](#)
- [GetCatalogImportStatus](#)
- [GetClassifier](#)
- [GetClassifiers](#)
- [GetColumnStatisticsForPartition](#)
- [GetColumnStatisticsForTable](#)
- [GetColumnStatisticsTaskRun](#)
- [GetColumnStatisticsTaskRuns](#)
- [GetConnection](#)
- [GetConnections](#)
- [GetCrawler](#)
- [GetCrawlerMetrics](#)
- [GetCrawlers](#)

- [GetCustomEntityType](#)
- [GetDatabase](#)
- [GetDatabases](#)
- [GetDataCatalogEncryptionSettings](#)
- [GetDataflowGraph](#)
- [GetDataQualityResult](#)
- [GetDataQualityRuleRecommendationRun](#)
- [GetDataQualityRuleset](#)
- [GetDataQualityRulesetEvaluationRun](#)
- [GetDevEndpoint](#)
- [GetDevEndpoints](#)
- [GetJob](#)
- [GetJobBookmark](#)
- [GetJobRun](#)
- [GetJobRuns](#)
- [GetJobs](#)
- [GetMapping](#)
- [GetMLTaskRun](#)
- [GetMLTaskRuns](#)
- [GetMLTransform](#)
- [GetMLTransforms](#)
- [GetPartition](#)
- [GetPartitionIndexes](#)
- [GetPartitions](#)
- [GetPlan](#)
- [GetRegistry](#)
- [GetResourcePolicies](#)
- [GetResourcePolicy](#)
- [GetSchema](#)
- [GetSchemaByDefinition](#)

- [GetSchemaVersion](#)
- [GetSchemaVersionsDiff](#)
- [GetSecurityConfiguration](#)
- [GetSecurityConfigurations](#)
- [GetSession](#)
- [GetStatement](#)
- [GetTable](#)
- [GetTableOptimizer](#)
- [GetTables](#)
- [GetTableVersion](#)
- [GetTableVersions](#)
- [GetTags](#)
- [GetTrigger](#)
- [GetTriggers](#)
- [GetUnfilteredPartitionMetadata](#)
- [GetUnfilteredPartitionsMetadata](#)
- [GetUnfilteredTableMetadata](#)
- [GetUsageProfile](#)
- [GetUserDefinedFunction](#)
- [GetUserDefinedFunctions](#)
- [GetWorkflow](#)
- [GetWorkflowRun](#)
- [GetWorkflowRunProperties](#)
- [GetWorkflowRuns](#)
- [ImportCatalogToGlue](#)
- [ListBlueprints](#)
- [ListColumnStatisticsTaskRuns](#)
- [ListCrawlers](#)
- [ListCrawls](#)
- [ListCustomEntityTypes](#)

- [ListDataQualityResults](#)
- [ListDataQualityRuleRecommendationRuns](#)
- [ListDataQualityRulesetEvaluationRuns](#)
- [ListDataQualityRulesets](#)
- [ListDevEndpoints](#)
- [ListJobs](#)
- [ListMLTransforms](#)
- [ListRegistries](#)
- [ListSchemas](#)
- [ListSchemaVersions](#)
- [ListSessions](#)
- [ListStatements](#)
- [ListTableOptimizerRuns](#)
- [ListTriggers](#)
- [ListUsageProfiles](#)
- [ListWorkflows](#)
- [PutDataCatalogEncryptionSettings](#)
- [PutResourcePolicy](#)
- [PutSchemaVersionMetadata](#)
- [PutWorkflowRunProperties](#)
- [QuerySchemaVersionMetadata](#)
- [RegisterSchemaVersion](#)
- [RemoveSchemaVersionMetadata](#)
- [ResetJobBookmark](#)
- [ResumeWorkflowRun](#)
- [RunStatement](#)
- [SearchTables](#)
- [StartBlueprintRun](#)
- [StartColumnStatisticsTaskRun](#)
- [StartCrawler](#)

- [StartCrawlerSchedule](#)
- [StartDataQualityRuleRecommendationRun](#)
- [StartDataQualityRulesetEvaluationRun](#)
- [StartExportLabelsTaskRun](#)
- [StartImportLabelsTaskRun](#)
- [StartJobRun](#)
- [StartMLEvaluationTaskRun](#)
- [StartMLLabelingSetGenerationTaskRun](#)
- [StartTrigger](#)
- [StartWorkflowRun](#)
- [StopColumnStatisticsTaskRun](#)
- [StopCrawler](#)
- [StopCrawlerSchedule](#)
- [StopSession](#)
- [StopTrigger](#)
- [StopWorkflowRun](#)
- [TagResource](#)
- [UntagResource](#)
- [UpdateBlueprint](#)
- [UpdateClassifier](#)
- [UpdateColumnStatisticsForPartition](#)
- [UpdateColumnStatisticsForTable](#)
- [UpdateConnection](#)
- [UpdateCrawler](#)
- [UpdateCrawlerSchedule](#)
- [UpdateDatabase](#)
- [UpdateDataQualityRuleset](#)
- [UpdateDevEndpoint](#)
- [UpdateJob](#)
- [UpdateJobFromSourceControl](#)

- [UpdateMLTransform](#)
- [UpdatePartition](#)
- [UpdateRegistry](#)
- [UpdateSchema](#)
- [UpdateSourceControlFromJob](#)
- [UpdateTable](#)
- [UpdateTableOptimizer](#)
- [UpdateTrigger](#)
- [UpdateUsageProfile](#)
- [UpdateUserDefinedFunction](#)
- [UpdateWorkflow](#)

BatchCreatePartition

Creates one or more partitions in a batch operation.

Request Syntax

```
{
  "CatalogId": "string",
  "DatabaseName": "string",
  "PartitionInputList": [
    {
      "LastAccessTime": number,
      "LastAnalyzedTime": number,
      "Parameters": {
        "string": "string"
      },
      "StorageDescriptor": {
        "AdditionalLocations": [ "string" ],
        "BucketColumns": [ "string" ],
        "Columns": [
          {
            "Comment": "string",
            "Name": "string",
            "Parameters": {
              "string": "string"
            },
            "Type": "string"
          }
        ],
        "Compressed": boolean,
        "InputFormat": "string",
        "Location": "string",
        "NumberOfBuckets": number,
        "OutputFormat": "string",
        "Parameters": {
          "string": "string"
        },
        "SchemaReference": {
          "SchemaId": {
            "RegistryName": "string",
            "SchemaArn": "string",
            "SchemaName": "string"
          }
        }
      }
    }
  ]
}
```

```

    "SchemaVersionId": "string",
    "SchemaVersionNumber": number
  },
  "SerdeInfo": {
    "Name": "string",
    "Parameters": {
      "string" : "string"
    },
    "SerializationLibrary": "string"
  },
  "SkewedInfo": {
    "SkewedColumnNames": [ "string" ],
    "SkewedColumnValueLocationMaps": {
      "string" : "string"
    },
    "SkewedColumnValues": [ "string" ]
  },
  "SortColumns": [
    {
      "Column": "string",
      "SortOrder": number
    }
  ],
  "StoredAsSubDirectories": boolean
},
"Values": [ "string" ]
}
],
"TableName": "string"
}

```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the catalog in which the partition is to be created. Currently, this should be the AWS account ID.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF\\t]*`

Required: No

DatabaseName

The name of the metadata database in which the partition is to be created.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF\\t]*`

Required: Yes

PartitionInputList

A list of `PartitionInput` structures that define the partitions to be created.

Type: Array of [PartitionInput](#) objects

Array Members: Minimum number of 0 items. Maximum number of 100 items.

Required: Yes

TableName

The name of the metadata table in which the partition is to be created.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF\\t]*`

Required: Yes

Response Syntax

```
{  
  "Errors": [  
    ...  
  ]  
}
```

```
{
  "ErrorDetail": {
    "ErrorCode": "string",
    "ErrorMessage": "string"
  },
  "PartitionValues": [ "string" ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Errors

The errors encountered when trying to create the requested partitions.

Type: Array of [PartitionError](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

BatchDeleteConnection

Deletes a list of connection definitions from the Data Catalog.

Request Syntax

```
{  
  "CatalogId": "string",  
  "ConnectionNameList": [ "string" ]  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog in which the connections reside. If none is provided, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

ConnectionNameList

A list of names of the connections to delete.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 25 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "Errors": {
    "string": {
      "ErrorCode": "string",
      "ErrorMessage": "string"
    }
  },
  "Succeeded": [ "string" ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Errors

A map of the names of connections that were not successfully deleted to error details.

Type: String to [ErrorDetail](#) object map

Key Length Constraints: Minimum length of 1. Maximum length of 255.

Key Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Succeeded

A list of names of the connection definitions that were successfully deleted.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

BatchDeletePartition

Deletes one or more partitions in a batch operation.

Request Syntax

```
{
  "CatalogId": "string",
  "DatabaseName": "string",
  "PartitionsToDelete": [
    {
      "Values": [ "string" ]
    }
  ],
  "TableName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog where the partition to be deleted resides. If none is provided, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

DatabaseName

The name of the catalog database in which the table in question resides.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

PartitionsToDelete

A list of `PartitionInput` structures that define the partitions to be deleted.

Type: Array of [PartitionValueList](#) objects

Array Members: Minimum number of 0 items. Maximum number of 25 items.

Required: Yes

TableName

The name of the table that contains the partitions to be deleted.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "Errors": [
    {
      "ErrorDetail": {
        "ErrorCode": "string",
        "ErrorMessage": "string"
      },
      "PartitionValues": [ "string" ]
    }
  ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Errors

The errors encountered when trying to delete the requested partitions.

Type: Array of [PartitionError](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)

- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

BatchDeleteTable

Deletes multiple tables at once.

Note

After completing this operation, you no longer have access to the table versions and partitions that belong to the deleted table. AWS Glue deletes these "orphaned" resources asynchronously in a timely manner, at the discretion of the service.

To ensure the immediate deletion of all related resources, before calling `BatchDeleteTable`, use `DeleteTableVersion` or `BatchDeleteTableVersion`, and `DeletePartition` or `BatchDeletePartition`, to delete any resources that belong to the table.

Request Syntax

```
{
  "CatalogId": "string",
  "DatabaseName": "string",
  "TablesToDelete": [ "string" ],
  "TransactionId": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog where the table resides. If none is provided, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF\\t]*`

Required: No

DatabaseName

The name of the catalog database in which the tables to delete reside. For Hive compatibility, this name is entirely lowercase.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

TablesToDelete

A list of the table to delete.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 100 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

TransactionId

The transaction ID at which to delete the table contents.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\p{L}\p{N}\p{P}]*`

Required: No

Response Syntax

```
{  
  "Errors": [  
    ...  
  ]  
}
```

```
{
  "ErrorDetail": {
    "ErrorCode": "string",
    "ErrorMessage": "string"
  },
  "TableName": "string"
}
]
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Errors

A list of errors encountered in attempting to delete the specified tables.

Type: Array of [TableError](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNotReadyException

A resource was not ready for a transaction.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

Required: Yes

TableName

The name of the table. For Hive compatibility, this name is entirely lowercase.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

VersionIds

A list of the IDs of versions to be deleted. A `VersionId` is a string representation of an integer. Each version is incremented by 1.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 100 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "Errors": [
    {
      "ErrorDetail": {
        "ErrorCode": "string",
        "ErrorMessage": "string"
      },
      "TableName": "string",
      "VersionId": "string"
    }
  ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Errors

A list of errors encountered while trying to delete the specified table versions.

Type: Array of [TableVersionError](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

BatchGetBlueprints

Retrieves information about a list of blueprints.

Request Syntax

```
{  
  "IncludeBlueprint": boolean,  
  "IncludeParameterSpec": boolean,  
  "Names": [ "string" ]  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

[IncludeBlueprint](#)

Specifies whether or not to include the blueprint in the response.

Type: Boolean

Required: No

[IncludeParameterSpec](#)

Specifies whether or not to include the parameters, as a JSON string, for the blueprint in the response.

Type: Boolean

Required: No

[Names](#)

A list of blueprint names.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 25 items.

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[\._\-A-Za-z0-9]+`

Required: Yes

Response Syntax

```
{
  "Blueprints": [
    {
      "BlueprintLocation": "string",
      "BlueprintServiceLocation": "string",
      "CreatedOn": number,
      "Description": "string",
      "ErrorMessage": "string",
      "LastActiveDefinition": {
        "BlueprintLocation": "string",
        "BlueprintServiceLocation": "string",
        "Description": "string",
        "LastModifiedOn": number,
        "ParameterSpec": "string"
      },
      "LastModifiedOn": number,
      "Name": "string",
      "ParameterSpec": "string",
      "Status": "string"
    }
  ],
  "MissingBlueprints": [ "string" ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Blueprints

Returns a list of blueprint as a Blueprints object.

Type: Array of [Blueprint](#) objects

[MissingBlueprints](#)

Returns a list of `BlueprintNames` that were not found.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[\.\-_\A-Za-z0-9]+`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)

- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)


```
"CrawlElapsedTime": number,
"CrawlerSecurityConfiguration": "string",
"CreationTime": number,
"DatabaseName": "string",
"Description": "string",
"LakeFormationConfiguration": {
  "AccountId": "string",
  "UseLakeFormationCredentials": boolean
},
>LastCrawl": {
  "ErrorMessage": "string",
  "LogGroup": "string",
  "LogStream": "string",
  "MessagePrefix": "string",
  "StartTime": number,
  "Status": "string"
},
>LastUpdated": number,
"LineageConfiguration": {
  "CrawlerLineageSettings": "string"
},
>Name": "string",
"RecrawlPolicy": {
  "RecrawlBehavior": "string"
},
>Role": "string",
>Schedule": {
  "ScheduleExpression": "string",
  "State": "string"
},
>SchemaChangePolicy": {
  "DeleteBehavior": "string",
  "UpdateBehavior": "string"
},
>State": "string",
>TablePrefix": "string",
>Targets": {
  "CatalogTargets": [
    {
      "ConnectionName": "string",
      "DatabaseName": "string",
      "DlqEventQueueArn": "string",
      "EventQueueArn": "string",
      "Tables": [ "string" ]
    }
  ]
}
```

```
    }
  ],
  "DeltaTargets": [
    {
      "ConnectionName": "string",
      "CreateNativeDeltaTable": boolean,
      "DeltaTables": [ "string " ],
      "WriteManifest": boolean
    }
  ],
  "DynamoDBTargets": [
    {
      "Path": "string",
      "scanAll": boolean,
      "scanRate": number
    }
  ],
  "HudiTargets": [
    {
      "ConnectionName": "string",
      "Exclusions": [ "string " ],
      "MaximumTraversalDepth": number,
      "Paths": [ "string " ]
    }
  ],
  "IcebergTargets": [
    {
      "ConnectionName": "string",
      "Exclusions": [ "string " ],
      "MaximumTraversalDepth": number,
      "Paths": [ "string " ]
    }
  ],
  "JdbcTargets": [
    {
      "ConnectionName": "string",
      "EnableAdditionalMetadata": [ "string " ],
      "Exclusions": [ "string " ],
      "Path": "string"
    }
  ],
  "MongoDBTargets": [
    {
      "ConnectionName": "string",
```

```
        "Path": "string",
        "ScanAll": boolean
    }
],
"S3Targets": [
    {
        "ConnectionName": "string",
        "DlqEventQueueArn": "string",
        "EventQueueArn": "string",
        "Exclusions": [ "string" ],
        "Path": "string",
        "SampleSize": number
    }
]
},
"Version": number
}
],
"CrawlersNotFound": [ "string" ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Crawlers

A list of crawler definitions.

Type: Array of [Crawler](#) objects

CrawlersNotFound

A list of names of crawlers that were not found.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 100 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

BatchGetCustomEntityTypes

Retrieves the details for the custom patterns specified by a list of names.

Request Syntax

```
{
  "Names": [ "string" ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Names

A list of names of the custom patterns that you want to retrieve.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 50 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "CustomEntityTypes": [
    {
      "ContextWords": [ "string" ],
      "Name": "string",
      "RegexString": "string"
    }
  ],
  "CustomEntityTypesNotFound": [ "string" ]
}
```

```
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

CustomEntityType

A list of `CustomEntityType` objects representing the custom patterns that have been created.

Type: Array of [CustomEntityType](#) objects

CustomEntityTypesNotFound

A list of the names of custom patterns that were not found.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 50 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

BatchGetDataQualityResult

Retrieves a list of data quality results for the specified result IDs.

Request Syntax

```
{
  "ResultIds": [ "string" ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

ResultIds

A list of unique result IDs for the data quality results.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 100 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "Results": [
    {
      "AnalyzerResults": [
        {
          "Description": "string",
          "EvaluatedMetrics": {
            "string": number
          },
          "EvaluationMessage": "string",

```

```
    "Name": "string"
  }
],
"CompletedOn": number,
"DataSource": {
  "GlueTable": {
    "AdditionalOptions": {
      "string" : "string"
    },
    "CatalogId": "string",
    "ConnectionName": "string",
    "DatabaseName": "string",
    "TableName": "string"
  }
},
"EvaluationContext": "string",
"JobName": "string",
"JobRunId": "string",
"Observations": [
  {
    "Description": "string",
    "MetricBasedObservation": {
      "MetricName": "string",
      "MetricValues": {
        "ActualValue": number,
        "ExpectedValue": number,
        "LowerLimit": number,
        "UpperLimit": number
      },
      "NewRules": [ "string" ]
    }
  }
],
"ResultId": "string",
"RuleResults": [
  {
    "Description": "string",
    "EvaluatedMetrics": {
      "string" : number
    },
    "EvaluationMessage": "string",
    "Name": "string",
    "Result": "string"
  }
]
```

```
    ],
    "RulesetEvaluationRunId": "string",
    "RulesetName": "string",
    "Score": number,
    "StartedOn": number
  }
],
"ResultsNotFound": [ "string" ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

[Results](#)

A list of `DataQualityResult` objects representing the data quality results.

Type: Array of [DataQualityResult](#) objects

[ResultsNotFound](#)

A list of result IDs for which results were not found.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 100 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF\\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

BatchGetDevEndpoints

Returns a list of resource metadata for a given list of development endpoint names. After calling the `ListDevEndpoints` operation, you can call this operation to access the data to which you have been granted permissions. This operation supports all IAM permissions, including permission conditions that uses tags.

Request Syntax

```
{
  "DevEndpointNames": [ "string" ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

DevEndpointNames

The list of DevEndpoint names, which might be the names returned from the `ListDevEndpoint` operation.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 25 items.

Required: Yes

Response Syntax

```
{
  "DevEndpoints": [
    {
      "Arguments": {
        "string" : "string"
      },
      "AvailabilityZone": "string",
      "CreatedTimestamp": number,

```



```
"EndpointName": "string",
"ExtraJarsS3Path": "string",
"ExtraPythonLibsS3Path": "string",
"FailureReason": "string",
"GlueVersion": "string",
"LastModifiedTimestamp": number,
"LastUpdateStatus": "string",
"NumberOfNodes": number,
"NumberOfWorkers": number,
"PrivateAddress": "string",
"PublicAddress": "string",
"PublicKey": "string",
"PublicKeys": [ "string" ],
"RoleArn": "string",
"SecurityConfiguration": "string",
"SecurityGroupIds": [ "string" ],
>Status": "string",
"SubnetId": "string",
"VpcId": "string",
"WorkerType": "string",
"YarnEndpointAddress": "string",
"ZeppelinRemoteSparkInterpreterPort": number
}
],
"DevEndpointsNotFound": [ "string" ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

DevEndpoints

A list of DevEndpoint definitions.

Type: Array of [DevEndpoint](#) objects

DevEndpointsNotFound

A list of DevEndpoints not found.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 25 items.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)

- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

BatchGetJobs

Returns a list of resource metadata for a given list of job names. After calling the `ListJobs` operation, you can call this operation to access the data to which you have been granted permissions. This operation supports all IAM permissions, including permission conditions that uses tags.

Request Syntax

```
{  
  "JobNames": [ "string" ]  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

JobNames

A list of job names, which might be the names returned from the `ListJobs` operation.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{  
  "Jobs": [  
    {  
      "AllocatedCapacity": number,  
      "CodeGenConfigurationNodes": {  
        "string" : {  
          "Aggregate": {  
            "Aggs": [  

```

```
    {
      "AggFunc": "string",
      "Column": [ "string" ]
    }
  ],
  "Groups": [
    [ "string" ]
  ],
  "Inputs": [ "string" ],
  "Name": "string"
},
"AmazonRedshiftSource": {
  "Data": {
    "AccessType": "string",
    "Action": "string",
    "AdvancedOptions": [
      {
        "Key": "string",
        "Value": "string"
      }
    ],
    "CatalogDatabase": {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    },
    "CatalogRedshiftSchema": "string",
    "CatalogRedshiftTable": "string",
    "CatalogTable": {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    },
    "Connection": {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    },
    "CrawlerConnection": "string",
    "IamRole": {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    }
  },
}
```

```
"MergeAction": "string",
"MergeClause": "string",
"MergeWhenMatched": "string",
"MergeWhenNotMatched": "string",
"PostAction": "string",
"PreAction": "string",
"SampleQuery": "string",
"Schema": {
  "Description": "string",
  "Label": "string",
  "Value": "string"
},
"SelectedColumns": [
  {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  }
],
"SourceType": "string",
"StagingTable": "string",
"Table": {
  "Description": "string",
  "Label": "string",
  "Value": "string"
},
"TablePrefix": "string",
"TableSchema": [
  {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  }
],
"TempDir": "string",
"Upsert": boolean
},
"Name": "string"
},
"AmazonRedshiftTarget": {
  "Data": {
    "AccessType": "string",
    "Action": "string",
    "AdvancedOptions": [
```

```
    {
      "Key": "string",
      "Value": "string"
    }
  ],
  "CatalogDatabase": {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  },
  "CatalogRedshiftSchema": "string",
  "CatalogRedshiftTable": "string",
  "CatalogTable": {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  },
  "Connection": {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  },
  "CrawlerConnection": "string",
  "IamRole": {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  },
  "MergeAction": "string",
  "MergeClause": "string",
  "MergeWhenMatched": "string",
  "MergeWhenNotMatched": "string",
  "PostAction": "string",
  "PreAction": "string",
  "SampleQuery": "string",
  "Schema": {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  },
  "SelectedColumns": [
    {
      "Description": "string",
      "Label": "string",
```

```

        "Value": "string"
      }
    ],
    "SourceType": "string",
    "StagingTable": "string",
    "Table": {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    },
    "TablePrefix": "string",
    "TableSchema": [
      {
        "Description": "string",
        "Label": "string",
        "Value": "string"
      }
    ],
    "TempDir": "string",
    "Upsert": boolean
  },
  "Inputs": [ "string" ],
  "Name": "string"
},
"ApplyMapping": {
  "Inputs": [ "string" ],
  "Mapping": [
    {
      "Children": [
        "Mapping"
      ],
      "Dropped": boolean,
      "FromPath": [ "string" ],
      "FromType": "string",
      "ToKey": "string",
      "ToType": "string"
    }
  ],
  "Name": "string"
},
"AthenaConnectorSource": {
  "ConnectionName": "string",
  "ConnectionTable": "string",
  "ConnectionType": "string",

```



```
    "ConnectorName": "string",
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
    "SchemaName": "string"
  },
  "CatalogDeltaSource": {
    "AdditionalDeltaOptions": {
      "string": "string"
    },
    "Database": "string",
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
    "Table": "string"
  },
  "CatalogHudiSource": {
    "AdditionalHudiOptions": {
      "string": "string"
    },
    "Database": "string",
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ]
  }
}
```

```
    }
  ]
}
],
"Table": "string"
},
"CatalogKafkaSource": {
  "Database": "string",
  "DataPreviewOptions": {
    "PollingTime": number,
    "RecordPollingLimit": number
  },
  "DetectSchema": boolean,
  "Name": "string",
  "StreamingOptions": {
    "AddRecordTimestamp": "string",
    "Assign": "string",
    "BootstrapServers": "string",
    "Classification": "string",
    "ConnectionName": "string",
    "Delimiter": "string",
    "EmitConsumerLagMetrics": "string",
    "EndingOffsets": "string",
    "IncludeHeaders": boolean,
    "MaxOffsetsPerTrigger": number,
    "MinPartitions": number,
    "NumRetries": number,
    "PollTimeoutMs": number,
    "RetryIntervalMs": number,
    "SecurityProtocol": "string",
    "StartingOffsets": "string",
    "StartingTimestamp": "string",
    "SubscribePattern": "string",
    "TopicName": "string"
  },
  "Table": "string",
  "WindowSize": number
},
"CatalogKinesisSource": {
  "Database": "string",
  "DataPreviewOptions": {
    "PollingTime": number,
    "RecordPollingLimit": number
  },
}
```

```

    "DetectSchema": boolean,
    "Name": "string",
    "StreamingOptions": {
      "AddIdleTimeBetweenReads": boolean,
      "AddRecordTimestamp": "string",
      "AvoidEmptyBatches": boolean,
      "Classification": "string",
      "Delimiter": "string",
      "DescribeShardInterval": number,
      "EmitConsumerLagMetrics": "string",
      "EndpointUrl": "string",
      "IdleTimeBetweenReadsInMs": number,
      "MaxFetchRecordsPerShard": number,
      "MaxFetchTimeInMs": number,
      "MaxRecordPerRead": number,
      "MaxRetryIntervalMs": number,
      "NumRetries": number,
      "RetryIntervalMs": number,
      "RoleArn": "string",
      "RoleSessionName": "string",
      "StartingPosition": "string",
      "StartingTimestamp": "string",
      "StreamArn": "string",
      "StreamName": "string"
    },
    "Table": "string",
    "WindowSize": number
  },
  "CatalogSource": {
    "Database": "string",
    "Name": "string",
    "Table": "string"
  },
  "CatalogTarget": {
    "Database": "string",
    "Inputs": [ "string " ],
    "Name": "string",
    "Table": "string"
  },
  "ConnectorDataSource": {
    "ConnectionType": "string",
    "Data": {
      "string": "string"
    }
  },

```

```
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
  },
  "ConnectorDataTarget": {
    "ConnectionType": "string",
    "Data": {
      "string": "string"
    },
    "Inputs": [ "string" ],
    "Name": "string"
  },
  "CustomCode": {
    "ClassName": "string",
    "Code": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ]
  },
  "DirectJDBCSource": {
    "ConnectionName": "string",
    "ConnectionType": "string",
    "Database": "string",
    "Name": "string",
    "RedshiftTmpDir": "string",
    "Table": "string"
  },
}
```

```
"DirectKafkaSource": {
  "DataPreviewOptions": {
    "PollingTime": number,
    "RecordPollingLimit": number
  },
  "DetectSchema": boolean,
  "Name": "string",
  "StreamingOptions": {
    "AddRecordTimestamp": "string",
    "Assign": "string",
    "BootstrapServers": "string",
    "Classification": "string",
    "ConnectionName": "string",
    "Delimiter": "string",
    "EmitConsumerLagMetrics": "string",
    "EndingOffsets": "string",
    "IncludeHeaders": boolean,
    "MaxOffsetsPerTrigger": number,
    "MinPartitions": number,
    "NumRetries": number,
    "PollTimeoutMs": number,
    "RetryIntervalMs": number,
    "SecurityProtocol": "string",
    "StartingOffsets": "string",
    "StartingTimestamp": "string",
    "SubscribePattern": "string",
    "TopicName": "string"
  },
  "WindowSize": number
},
"DirectKinesisSource": {
  "DataPreviewOptions": {
    "PollingTime": number,
    "RecordPollingLimit": number
  },
  "DetectSchema": boolean,
  "Name": "string",
  "StreamingOptions": {
    "AddIdleTimeBetweenReads": boolean,
    "AddRecordTimestamp": "string",
    "AvoidEmptyBatches": boolean,
    "Classification": "string",
    "Delimiter": "string",
    "DescribeShardInterval": number,
```

```

    "EmitConsumerLagMetrics": "string",
    "EndpointUrl": "string",
    "IdleTimeBetweenReadsInMs": number,
    "MaxFetchRecordsPerShard": number,
    "MaxFetchTimeInMs": number,
    "MaxRecordPerRead": number,
    "MaxRetryIntervalMs": number,
    "NumRetries": number,
    "RetryIntervalMs": number,
    "RoleArn": "string",
    "RoleSessionName": "string",
    "StartingPosition": "string",
    "StartingTimestamp": "string",
    "StreamArn": "string",
    "StreamName": "string"
  },
  "WindowSize": number
},
"DropDuplicates": {
  "Columns": [
    [ "string" ]
  ],
  "Inputs": [ "string" ],
  "Name": "string"
},
"DropFields": {
  "Inputs": [ "string" ],
  "Name": "string",
  "Paths": [
    [ "string" ]
  ]
},
"DropNullFields": {
  "Inputs": [ "string" ],
  "Name": "string",
  "NullCheckBoxList": {
    "IsEmpty": boolean,
    "IsNegOne": boolean,
    "IsNullString": boolean
  }
},
"NullTextList": [
  {
    "Datatype": {
      "Id": "string",

```

```
        "Label": "string"
      },
      "Value": "string"
    }
  ]
},
"DynamicTransform": {
  "FunctionName": "string",
  "Inputs": [ "string" ],
  "Name": "string",
  "OutputSchemas": [
    {
      "Columns": [
        {
          "Name": "string",
          "Type": "string"
        }
      ]
    }
  ]
},
"Parameters": [
  {
    "IsOptional": boolean,
    "ListType": "string",
    "Name": "string",
    "Type": "string",
    "ValidationMessage": "string",
    "ValidationRule": "string",
    "Value": [ "string" ]
  }
],
"Path": "string",
"TransformName": "string",
"Version": "string"
},
"DynamoDBCatalogSource": {
  "Database": "string",
  "Name": "string",
  "Table": "string"
},
"EvaluateDataQuality": {
  "Inputs": [ "string" ],
  "Name": "string",
  "Output": "string",
```

```

    "PublishingOptions": {
      "CloudWatchMetricsEnabled": boolean,
      "EvaluationContext": "string",
      "ResultsPublishingEnabled": boolean,
      "ResultsS3Prefix": "string"
    },
    "Ruleset": "string",
    "StopJobOnFailureOptions": {
      "StopJobOnFailureTiming": "string"
    }
  },
  "EvaluateDataQualityMultiFrame": {
    "AdditionalDataSources": {
      "string" : "string"
    },
    "AdditionalOptions": {
      "string" : "string"
    },
    "Inputs": [ "string" ],
    "Name": "string",
    "PublishingOptions": {
      "CloudWatchMetricsEnabled": boolean,
      "EvaluationContext": "string",
      "ResultsPublishingEnabled": boolean,
      "ResultsS3Prefix": "string"
    },
    "Ruleset": "string",
    "StopJobOnFailureOptions": {
      "StopJobOnFailureTiming": "string"
    }
  },
  "FillMissingValues": {
    "FilledPath": "string",
    "ImputedPath": "string",
    "Inputs": [ "string" ],
    "Name": "string"
  },
  "Filter": {
    "Filters": [
      {
        "Negated": boolean,
        "Operation": "string",
        "Values": [

```



```

        "Type": "string",
        "Value": [ "string" ]
      }
    ]
  },
  "Inputs": [ "string" ],
  "LogicalOperator": "string",
  "Name": "string"
},
"GovernedCatalogSource": {
  "AdditionalOptions": {
    "BoundedFiles": number,
    "BoundedSize": number
  },
  "Database": "string",
  "Name": "string",
  "PartitionPredicate": "string",
  "Table": "string"
},
"GovernedCatalogTarget": {
  "Database": "string",
  "Inputs": [ "string" ],
  "Name": "string",
  "PartitionKeys": [
    [ "string" ]
  ],
  "SchemaChangePolicy": {
    "EnableUpdateCatalog": boolean,
    "UpdateBehavior": "string"
  },
  "Table": "string"
},
"JDBCConnectorSource": {
  "AdditionalOptions": {
    "DataTypeMapping": {
      "string" : "string"
    },
    "FilterPredicate": "string",
    "JobBookmarkKeys": [ "string" ],
    "JobBookmarkKeysSortOrder": "string",
    "LowerBound": number,
    "NumPartitions": number,
    "PartitionColumn": "string",

```

```

        "UpperBound": number
    },
    "ConnectionName": "string",
    "ConnectionTable": "string",
    "ConnectionType": "string",
    "ConnectorName": "string",
    "Name": "string",
    "OutputSchemas": [
        {
            "Columns": [
                {
                    "Name": "string",
                    "Type": "string"
                }
            ]
        }
    ],
    "Query": "string"
},
"JDBCConnectorTarget": {
    "AdditionalOptions": {
        "string" : "string"
    },
    "ConnectionName": "string",
    "ConnectionTable": "string",
    "ConnectionType": "string",
    "ConnectorName": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "OutputSchemas": [
        {
            "Columns": [
                {
                    "Name": "string",
                    "Type": "string"
                }
            ]
        }
    ]
},
"Join": {
    "Columns": [
        {
            "From": "string",

```

```
        "Keys": [
            [ "string" ]
        ]
    },
    ],
    "Inputs": [ "string" ],
    "JoinType": "string",
    "Name": "string"
},
"Merge": {
    "Inputs": [ "string" ],
    "Name": "string",
    "PrimaryKeys": [
        [ "string" ]
    ],
    "Source": "string"
},
"MicrosoftSQLServerCatalogSource": {
    "Database": "string",
    "Name": "string",
    "Table": "string"
},
"MicrosoftSQLServerCatalogTarget": {
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "Table": "string"
},
"MySQLCatalogSource": {
    "Database": "string",
    "Name": "string",
    "Table": "string"
},
"MySQLCatalogTarget": {
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "Table": "string"
},
"OracleSQLCatalogSource": {
    "Database": "string",
    "Name": "string",
    "Table": "string"
},
},
```

```
"OracleSQLCatalogTarget": {
  "Database": "string",
  "Inputs": [ "string" ],
  "Name": "string",
  "Table": "string"
},
"PIIDetection": {
  "EntityTypesToDetect": [ "string" ],
  "Inputs": [ "string" ],
  "MaskValue": "string",
  "Name": "string",
  "OutputColumnName": "string",
  "PiiType": "string",
  "SampleFraction": number,
  "ThresholdFraction": number
},
"PostgreSQLCatalogSource": {
  "Database": "string",
  "Name": "string",
  "Table": "string"
},
"PostgreSQLCatalogTarget": {
  "Database": "string",
  "Inputs": [ "string" ],
  "Name": "string",
  "Table": "string"
},
"Recipe": {
  "Inputs": [ "string" ],
  "Name": "string",
  "RecipeReference": {
    "RecipeArn": "string",
    "RecipeVersion": "string"
  }
},
"RedshiftSource": {
  "Database": "string",
  "Name": "string",
  "RedshiftTmpDir": "string",
  "Table": "string",
  "TmpDirIAMRole": "string"
},
"RedshiftTarget": {
  "Database": "string",
```

```

    "Inputs": [ "string" ],
    "Name": "string",
    "RedshiftTmpDir": "string",
    "Table": "string",
    "TmpDirIAMRole": "string",
    "UpsertRedshiftOptions": {
      "ConnectionName": "string",
      "TableLocation": "string",
      "UpsertKeys": [ "string" ]
    }
  },
  "RelationalCatalogSource": {
    "Database": "string",
    "Name": "string",
    "Table": "string"
  },
  "RenameField": {
    "Inputs": [ "string" ],
    "Name": "string",
    "SourcePath": [ "string" ],
    "TargetPath": [ "string" ]
  },
  "S3CatalogDeltaSource": {
    "AdditionalDeltaOptions": {
      "string" : "string"
    },
    "Database": "string",
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
    "Table": "string"
  },
  "S3CatalogHudiSource": {
    "AdditionalHudiOptions": {
      "string" : "string"
    }
  },

```

```
    "Database": "string",
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
    "Table": "string"
  },
  "S3CatalogSource": {
    "AdditionalOptions": {
      "BoundedFiles": number,
      "BoundedSize": number
    },
    "Database": "string",
    "Name": "string",
    "PartitionPredicate": "string",
    "Table": "string"
  },
  "S3CatalogTarget": {
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
      [ "string" ]
    ],
    "SchemaChangePolicy": {
      "EnableUpdateCatalog": boolean,
      "UpdateBehavior": "string"
    },
    "Table": "string"
  },
  "S3CsvSource": {
    "AdditionalOptions": {
      "BoundedFiles": number,
      "BoundedSize": number,
      "EnableSamplePath": boolean,
      "SamplePath": "string"
    }
  },
```

```

    "CompressionType": "string",
    "Escaper": "string",
    "Exclusions": [ "string" ],
    "GroupFiles": "string",
    "GroupSize": "string",
    "MaxBand": number,
    "MaxFilesInBand": number,
    "Multiline": boolean,
    "Name": "string",
    "OptimizePerformance": boolean,
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
    "Paths": [ "string" ],
    "QuoteChar": "string",
    "Recurse": boolean,
    "Separator": "string",
    "SkipFirst": boolean,
    "WithHeader": boolean,
    "WriteHeader": boolean
  },
  "S3DeltaCatalogTarget": {
    "AdditionalOptions": {
      "string" : "string"
    },
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
      [ "string" ]
    ],
    "SchemaChangePolicy": {
      "EnableUpdateCatalog": boolean,
      "UpdateBehavior": "string"
    },
    "Table": "string"
  },
},

```

```

    "S3DeltaDirectTarget": {
      "AdditionalOptions": {
        "string" : "string"
      },
      "Compression": "string",
      "Format": "string",
      "Inputs": [ "string" ],
      "Name": "string",
      "PartitionKeys": [
        [ "string" ]
      ],
      "Path": "string",
      "SchemaChangePolicy": {
        "Database": "string",
        "EnableUpdateCatalog": boolean,
        "Table": "string",
        "UpdateBehavior": "string"
      }
    },
    "S3DeltaSource": {
      "AdditionalDeltaOptions": {
        "string" : "string"
      },
      "AdditionalOptions": {
        "BoundedFiles": number,
        "BoundedSize": number,
        "EnableSamplePath": boolean,
        "SamplePath": "string"
      },
      "Name": "string",
      "OutputSchemas": [
        {
          "Columns": [
            {
              "Name": "string",
              "Type": "string"
            }
          ]
        }
      ],
      "Paths": [ "string" ]
    },
    "S3DirectTarget": {
      "Compression": "string",

```



```

    "Format": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
      [ "string" ]
    ],
    "Path": "string",
    "SchemaChangePolicy": {
      "Database": "string",
      "EnableUpdateCatalog": boolean,
      "Table": "string",
      "UpdateBehavior": "string"
    }
  },
  "S3GlueParquetTarget": {
    "Compression": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
      [ "string" ]
    ],
    "Path": "string",
    "SchemaChangePolicy": {
      "Database": "string",
      "EnableUpdateCatalog": boolean,
      "Table": "string",
      "UpdateBehavior": "string"
    }
  },
  "S3HudiCatalogTarget": {
    "AdditionalOptions": {
      "string" : "string"
    },
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
      [ "string" ]
    ],
    "SchemaChangePolicy": {
      "EnableUpdateCatalog": boolean,
      "UpdateBehavior": "string"
    },
    "Table": "string"
  }
}

```

```

},
  "S3HudiDirectTarget": {
    "AdditionalOptions": {
      "string" : "string"
    },
    "Compression": "string",
    "Format": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
      [ "string" ]
    ],
    "Path": "string",
    "SchemaChangePolicy": {
      "Database": "string",
      "EnableUpdateCatalog": boolean,
      "Table": "string",
      "UpdateBehavior": "string"
    }
  },
  "S3HudiSource": {
    "AdditionalHudiOptions": {
      "string" : "string"
    },
    "AdditionalOptions": {
      "BoundedFiles": number,
      "BoundedSize": number,
      "EnableSamplePath": boolean,
      "SamplePath": "string"
    },
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
    "Paths": [ "string" ]
  },
  "S3JsonSource": {

```

```
    "AdditionalOptions": {
      "BoundedFiles": number,
      "BoundedSize": number,
      "EnableSamplePath": boolean,
      "SamplePath": "string"
    },
    "CompressionType": "string",
    "Exclusions": [ "string" ],
    "GroupFiles": "string",
    "GroupSize": "string",
    "JsonPath": "string",
    "MaxBand": number,
    "MaxFilesInBand": number,
    "Multiline": boolean,
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
    "Paths": [ "string" ],
    "Recurse": boolean
  },
  "S3ParquetSource": {
    "AdditionalOptions": {
      "BoundedFiles": number,
      "BoundedSize": number,
      "EnableSamplePath": boolean,
      "SamplePath": "string"
    },
    "CompressionType": "string",
    "Exclusions": [ "string" ],
    "GroupFiles": "string",
    "GroupSize": "string",
    "MaxBand": number,
    "MaxFilesInBand": number,
    "Name": "string",
    "OutputSchemas": [
      {
```

```
        "Columns": [
            {
                "Name": "string",
                "Type": "string"
            }
        ]
    },
    ],
    "Paths": [ "string" ],
    "Recurse": boolean
},
"SelectFields": {
    "Inputs": [ "string" ],
    "Name": "string",
    "Paths": [
        [ "string" ]
    ]
},
"SelectFromCollection": {
    "Index": number,
    "Inputs": [ "string" ],
    "Name": "string"
},
"SnowflakeSource": {
    "Data": {
        "Action": "string",
        "AdditionalOptions": {
            "string" : "string"
        },
        "AutoPushdown": boolean,
        "Connection": {
            "Description": "string",
            "Label": "string",
            "Value": "string"
        },
        "Database": "string",
        "IamRole": {
            "Description": "string",
            "Label": "string",
            "Value": "string"
        },
        "MergeAction": "string",
        "MergeClause": "string",
        "MergeWhenMatched": "string",
```

```

    "MergeWhenNotMatched": "string",
    "PostAction": "string",
    "PreAction": "string",
    "SampleQuery": "string",
    "Schema": "string",
    "SelectedColumns": [
      {
        "Description": "string",
        "Label": "string",
        "Value": "string"
      }
    ],
    "SourceType": "string",
    "StagingTable": "string",
    "Table": "string",
    "TableSchema": [
      {
        "Description": "string",
        "Label": "string",
        "Value": "string"
      }
    ],
    "TempDir": "string",
    "Upsert": boolean
  },
  "Name": "string",
  "OutputSchemas": [
    {
      "Columns": [
        {
          "Name": "string",
          "Type": "string"
        }
      ]
    }
  ]
},
"SnowflakeTarget": {
  "Data": {
    "Action": "string",
    "AdditionalOptions": {
      "string": "string"
    },
    "AutoPushdown": boolean,

```

```
"Connection": {
  "Description": "string",
  "Label": "string",
  "Value": "string"
},
"Database": "string",
"IamRole": {
  "Description": "string",
  "Label": "string",
  "Value": "string"
},
"MergeAction": "string",
"MergeClause": "string",
"MergeWhenMatched": "string",
"MergeWhenNotMatched": "string",
"PostAction": "string",
"PreAction": "string",
"SampleQuery": "string",
"Schema": "string",
"SelectedColumns": [
  {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  }
],
"SourceType": "string",
"StagingTable": "string",
"Table": "string",
"TableSchema": [
  {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  }
],
"TempDir": "string",
"Upsert": boolean
},
"Inputs": [ "string" ],
"Name": "string"
},
"SparkConnectorSource": {
  "AdditionalOptions": {
```

```
        "string" : "string"
    },
    "ConnectionName": "string",
    "ConnectionType": "string",
    "ConnectorName": "string",
    "Name": "string",
    "OutputSchemas": [
        {
            "Columns": [
                {
                    "Name": "string",
                    "Type": "string"
                }
            ]
        }
    ]
},
"SparkConnectorTarget": {
    "AdditionalOptions": {
        "string" : "string"
    },
    "ConnectionName": "string",
    "ConnectionType": "string",
    "ConnectorName": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "OutputSchemas": [
        {
            "Columns": [
                {
                    "Name": "string",
                    "Type": "string"
                }
            ]
        }
    ]
},
"SparkSQL": {
    "Inputs": [ "string" ],
    "Name": "string",
    "OutputSchemas": [
        {
            "Columns": [
                {
```

```

        "Name": "string",
        "Type": "string"
    }
]
},
"SqlAliases": [
    {
        "Alias": "string",
        "From": "string"
    }
],
"SqlQuery": "string"
},
"Spigot": {
    "Inputs": [ "string" ],
    "Name": "string",
    "Path": "string",
    "Prob": number,
    "Topk": number
},
"SplitFields": {
    "Inputs": [ "string" ],
    "Name": "string",
    "Paths": [
        [ "string" ]
    ]
},
"Union": {
    "Inputs": [ "string" ],
    "Name": "string",
    "UnionType": "string"
}
}
},
"Command": {
    "Name": "string",
    "PythonVersion": "string",
    "Runtime": "string",
    "ScriptLocation": "string"
},
"Connections": {
    "Connections": [ "string" ]
},

```



```
"CreatedOn": number,
"DefaultArguments": {
  "string" : "string"
},
"Description": "string",
"ExecutionClass": "string",
"ExecutionProperty": {
  "MaxConcurrentRuns": number
},
"GlueVersion": "string",
"JobMode": "string",
"LastModifiedOn": number,
"LogUri": "string",
"MaintenanceWindow": "string",
"MaxCapacity": number,
"MaxRetries": number,
"Name": "string",
"NonOverridableArguments": {
  "string" : "string"
},
"NotificationProperty": {
  "NotifyDelayAfter": number
},
"NumberOfWorkers": number,
"ProfileName": "string",
"Role": "string",
"SecurityConfiguration": "string",
"SourceControlDetails": {
  "AuthStrategy": "string",
  "AuthToken": "string",
  "Branch": "string",
  "Folder": "string",
  "LastCommitId": "string",
  "Owner": "string",
  "Provider": "string",
  "Repository": "string"
},
"Timeout": number,
"WorkerType": "string"
}
],
"JobsNotFound": [ "string" ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Jobs

A list of job definitions.

Type: Array of [Job](#) objects

JobsNotFound

A list of names of jobs not found.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

BatchGetPartition

Retrieves partitions in a batch request.

Request Syntax

```
{
  "CatalogId": "string",
  "DatabaseName": "string",
  "PartitionsToGet": [
    {
      "Values": [ "string" ]
    }
  ],
  "TableName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog where the partitions in question reside. If none is supplied, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

DatabaseName

The name of the catalog database where the partitions reside.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

PartitionsToGet

A list of partition values identifying the partitions to retrieve.

Type: Array of [PartitionValueList](#) objects

Array Members: Minimum number of 0 items. Maximum number of 1000 items.

Required: Yes

TableName

The name of the partitions' table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "Partitions": [
    {
      "CatalogId": "string",
      "CreationTime": number,
      "DatabaseName": "string",
      "LastAccessTime": number,
      "LastAnalyzedTime": number,
      "Parameters": {
        "string" : "string"
      },
      "StorageDescriptor": {
        "AdditionalLocations": [ "string" ],
        "BucketColumns": [ "string" ],
        "Columns": [
          {
            "Comment": "string",
```

```
        "Name": "string",
        "Parameters": {
            "string" : "string"
        },
        "Type": "string"
    }
],
"Compressed": boolean,
"InputFormat": "string",
"Location": "string",
"NumberOfBuckets": number,
"OutputFormat": "string",
"Parameters": {
    "string" : "string"
},
"SchemaReference": {
    "SchemaId": {
        "RegistryName": "string",
        "SchemaArn": "string",
        "SchemaName": "string"
    },
    "SchemaVersionId": "string",
    "SchemaVersionNumber": number
},
"SerdeInfo": {
    "Name": "string",
    "Parameters": {
        "string" : "string"
    },
    "SerializationLibrary": "string"
},
"SkewedInfo": {
    "SkewedColumnNames": [ "string" ],
    "SkewedColumnValueLocationMaps": {
        "string" : "string"
    },
    "SkewedColumnValues": [ "string" ]
},
"SortColumns": [
    {
        "Column": "string",
        "SortOrder": number
    }
],
```

```
    "StoredAsSubDirectories": boolean
  },
  "TableName": "string",
  "Values": [ "string " ]
}
],
"UnprocessedKeys": [
  {
    "Values": [ "string " ]
  }
]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Partitions

A list of the requested partitions.

Type: Array of [Partition](#) objects

UnprocessedKeys

A list of the partition values in the request for which partitions were not returned.

Type: Array of [PartitionValueList](#) objects

Array Members: Minimum number of 0 items. Maximum number of 1000 items.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

FederationSourceException

A federation source failed.

HTTP Status Code: 400

FederationSourceRetryableException

A federation source failed, but the operation may be retried.

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

InvalidStateException

An error that indicates your data is in an invalid state.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

BatchGetTableOptimizer

Returns the configuration for the specified table optimizers.

Request Syntax

```
{
  "Entries": [
    {
      "catalogId": "string",
      "databaseName": "string",
      "tableName": "string",
      "type": "string"
    }
  ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Entries

A list of BatchGetTableOptimizerEntry objects specifying the table optimizers to retrieve.

Type: Array of [BatchGetTableOptimizerEntry](#) objects

Required: Yes

Response Syntax

```
{
  "Failures": [
    {
      "catalogId": "string",
      "databaseName": "string",
      "error": {
        "ErrorCode": "string",

```

```

    "ErrorMessage": "string"
  },
  "tableName": "string",
  "type": "string"
}
],
"TableOptimizers": [
  {
    "catalogId": "string",
    "databaseName": "string",
    "tableName": "string",
    "tableOptimizer": {
      "configuration": {
        "enabled": boolean,
        "roleArn": "string"
      },
      "lastRun": {
        "endTimeStamp": number,
        "error": "string",
        "eventType": "string",
        "metrics": {
          "JobDurationInHour": "string",
          "NumberOfBytesCompacted": "string",
          "NumberOfDpus": "string",
          "NumberOfFilesCompacted": "string"
        },
        "startTimeStamp": number
      },
      "type": "string"
    }
  }
]
}

```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Failures

A list of errors from the operation.

Type: Array of [BatchGetTableOptimizerError](#) objects

[TableOptimizers](#)

A list of `BatchTableOptimizer` objects.

Type: Array of [BatchTableOptimizer](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

BatchGetTriggers

Returns a list of resource metadata for a given list of trigger names. After calling the `ListTriggers` operation, you can call this operation to access the data to which you have been granted permissions. This operation supports all IAM permissions, including permission conditions that uses tags.

Request Syntax

```
{
  "TriggerNames": [ "string" ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

TriggerNames

A list of trigger names, which may be the names returned from the `ListTriggers` operation.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "Triggers": [
    {
      "Actions": [
        {
          "Arguments": {
            "string" : "string"
          }
        }
      ]
    }
  ]
}
```

```

    },
    "CrawlerName": "string",
    "JobName": "string",
    "NotificationProperty": {
      "NotifyDelayAfter": number
    },
    "SecurityConfiguration": "string",
    "Timeout": number
  }
],
"Description": "string",
"EventBatchingCondition": {
  "BatchSize": number,
  "BatchWindow": number
},
"Id": "string",
"Name": "string",
"Predicate": {
  "Conditions": [
    {
      "CrawlerName": "string",
      "CrawlState": "string",
      "JobName": "string",
      "LogicalOperator": "string",
      "State": "string"
    }
  ],
  "Logical": "string"
},
"Schedule": "string",
"State": "string",
"Type": "string",
"WorkflowName": "string"
}
],
"TriggersNotFound": [ "string" ]
}

```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

[Triggers](#)

A list of trigger definitions.

Type: Array of [Trigger](#) objects

[TriggersNotFound](#)

A list of names of triggers not found.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)

- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

BatchGetWorkflows

Returns a list of resource metadata for a given list of workflow names. After calling the `ListWorkflows` operation, you can call this operation to access the data to which you have been granted permissions. This operation supports all IAM permissions, including permission conditions that uses tags.

Request Syntax

```
{
  "IncludeGraph": boolean,
  "Names": [ "string" ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

IncludeGraph

Specifies whether to include a graph when returning the workflow resource metadata.

Type: Boolean

Required: No

Names

A list of workflow names, which may be the names returned from the `ListWorkflows` operation.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 25 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "MissingWorkflows": [ "string" ],
  "Workflows": [
    {
      "BlueprintDetails": {
        "BlueprintName": "string",
        "RunId": "string"
      },
      "CreatedOn": number,
      "DefaultRunProperties": {
        "string" : "string"
      },
      "Description": "string",
      "Graph": {
        "Edges": [
          {
            "DestinationId": "string",
            "SourceId": "string"
          }
        ],
        "Nodes": [
          {
            "CrawlerDetails": {
              "Crawls": [
                {
                  "CompletedOn": number,
                  "ErrorMessage": "string",
                  "LogGroup": "string",
                  "LogStream": "string",
                  "StartedOn": number,
                  "State": "string"
                }
              ]
            },
            "JobDetails": {
              "JobRuns": [
                {
                  "AllocatedCapacity": number,
                  "Arguments": {
                    "string" : "string"
                  }
                }
              ]
            }
          }
        ]
      }
    }
  ]
}
```

```

    "Attempt": number,
    "CompletedOn": number,
    "DPUSecods": number,
    "ErrorMessage": "string",
    "ExecutionClass": "string",
    "ExecutionTime": number,
    "GlueVersion": "string",
    "Id": "string",
    "JobMode": "string",
    "JobName": "string",
    "JobRunState": "string",
    "LastModifiedOn": number,
    "LogGroupName": "string",
    "MaintenanceWindow": "string",
    "MaxCapacity": number,
    "NotificationProperty": {
      "NotifyDelayAfter": number
    },
    "NumberOfWorkers": number,
    "PredecessorRuns": [
      {
        "JobName": "string",
        "RunId": "string"
      }
    ],
    "PreviousRunId": "string",
    "ProfileName": "string",
    "SecurityConfiguration": "string",
    "StartedOn": number,
    "Timeout": number,
    "TriggerName": "string",
    "WorkerType": "string"
  }
]
},
"Name": "string",
"TriggerDetails": {
  "Trigger": {
    "Actions": [
      {
        "Arguments": {
          "string": "string"
        },
        "CrawlerName": "string",

```

```

        "JobName": "string",
        "NotificationProperty": {
            "NotifyDelayAfter": number
        },
        "SecurityConfiguration": "string",
        "Timeout": number
    }
],
"Description": "string",
"EventBatchingCondition": {
    "BatchSize": number,
    "BatchWindow": number
},
"Id": "string",
"Name": "string",
"Predicate": {
    "Conditions": [
        {
            "CrawlerName": "string",
            "CrawlState": "string",
            "JobName": "string",
            "LogicalOperator": "string",
            "State": "string"
        }
    ],
    "Logical": "string"
},
"Schedule": "string",
"State": "string",
"Type": "string",
"WorkflowName": "string"
}
},
"Type": "string",
"UniqueId": "string"
}
]
},
"LastModifiedOn": number,
"LastRun": {
    "CompletedOn": number,
    "ErrorMessage": "string",
    "Graph": {
        "Edges": [

```

```

    {
      "DestinationId": "string",
      "SourceId": "string"
    }
  ],
  "Nodes": [
    {
      "CrawlerDetails": {
        "Crawls": [
          {
            "CompletedOn": number,
            "ErrorMessage": "string",
            "LogGroup": "string",
            "LogStream": "string",
            "StartedOn": number,
            "State": "string"
          }
        ]
      },
      "JobDetails": {
        "JobRuns": [
          {
            "AllocatedCapacity": number,
            "Arguments": {
              "string": "string"
            },
            "Attempt": number,
            "CompletedOn": number,
            "DPUSecods": number,
            "ErrorMessage": "string",
            "ExecutionClass": "string",
            "ExecutionTime": number,
            "GlueVersion": "string",
            "Id": "string",
            "JobMode": "string",
            "JobName": "string",
            "JobRunState": "string",
            "LastModifiedOn": number,
            "LogGroupName": "string",
            "MaintenanceWindow": "string",
            "MaxCapacity": number,
            "NotificationProperty": {
              "NotifyDelayAfter": number
            }
          }
        ]
      }
    }
  ]
}

```

```

        "NumberOfWorkers": number,
        "PredecessorRuns": [
            {
                "JobName": "string",
                "RunId": "string"
            }
        ],
        "PreviousRunId": "string",
        "ProfileName": "string",
        "SecurityConfiguration": "string",
        "StartedOn": number,
        "Timeout": number,
        "TriggerName": "string",
        "WorkerType": "string"
    }
]
},
"Name": "string",
"TriggerDetails": {
    "Trigger": {
        "Actions": [
            {
                "Arguments": {
                    "string": "string"
                },
                "CrawlerName": "string",
                "JobName": "string",
                "NotificationProperty": {
                    "NotifyDelayAfter": number
                },
                "SecurityConfiguration": "string",
                "Timeout": number
            }
        ],
        "Description": "string",
        "EventBatchingCondition": {
            "BatchSize": number,
            "BatchWindow": number
        },
        "Id": "string",
        "Name": "string",
        "Predicate": {
            "Conditions": [
                {

```

```

        "CrawlerName": "string",
        "CrawlState": "string",
        "JobName": "string",
        "LogicalOperator": "string",
        "State": "string"
    }
  ],
  "Logical": "string"
},
"Schedule": "string",
"State": "string",
"Type": "string",
"WorkflowName": "string"
}
},
"Type": "string",
"UniqueId": "string"
}
]
},
"Name": "string",
"PreviousRunId": "string",
"StartedOn": number,
"StartingEventBatchCondition": {
  "BatchSize": number,
  "BatchWindow": number
},
"Statistics": {
  "ErroredActions": number,
  "FailedActions": number,
  "RunningActions": number,
  "StoppedActions": number,
  "SucceededActions": number,
  "TimeoutActions": number,
  "TotalActions": number,
  "WaitingActions": number
},
"Status": "string",
"WorkflowRunId": "string",
"WorkflowRunProperties": {
  "string" : "string"
}
},
"MaxConcurrentRuns": number,

```

```
    "Name": "string"  
  }  
]  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

MissingWorkflows

A list of names of workflows not found.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 25 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Workflows

A list of workflow resource metadata.

Type: Array of [Workflow](#) objects

Array Members: Minimum number of 1 item. Maximum number of 25 items.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

BatchStopJobRun

Stops one or more job runs for a specified job definition.

Request Syntax

```
{
  "JobName": "string",
  "JobRunIds": [ "string" ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

JobName

The name of the job definition for which to stop job runs.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

JobRunIds

A list of the JobRunIds that should be stopped for that job definition.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 25 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "Errors": [
    {
      "ErrorDetail": {
        "ErrorCode": "string",
        "ErrorMessage": "string"
      },
      "JobName": "string",
      "JobRunId": "string"
    }
  ],
  "SuccessfulSubmissions": [
    {
      "JobName": "string",
      "JobRunId": "string"
    }
  ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Errors

A list of the errors that were encountered in trying to stop JobRuns, including the JobRunId for which each error was encountered and details about the error.

Type: Array of [BatchStopJobRunError](#) objects

SuccessfulSubmissions

A list of the JobRuns that were successfully submitted for stopping.

Type: Array of [BatchStopJobRunSuccessfulSubmission](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

BatchUpdatePartition

Updates one or more partitions in a batch operation.

Request Syntax

```
{
  "CatalogId": "string",
  "DatabaseName": "string",
  "Entries": [
    {
      "PartitionInput": {
        "LastAccessTime": number,
        "LastAnalyzedTime": number,
        "Parameters": {
          "string": "string"
        },
      },
      "StorageDescriptor": {
        "AdditionalLocations": [ "string" ],
        "BucketColumns": [ "string" ],
        "Columns": [
          {
            "Comment": "string",
            "Name": "string",
            "Parameters": {
              "string": "string"
            },
            "Type": "string"
          }
        ],
      },
      "Compressed": boolean,
      "InputFormat": "string",
      "Location": "string",
      "NumberOfBuckets": number,
      "OutputFormat": "string",
      "Parameters": {
        "string": "string"
      },
      "SchemaReference": {
        "SchemaId": {
          "RegistryName": "string",
          "SchemaArn": "string",
          "SchemaName": "string"
        }
      }
    }
  ]
}
```

```

    },
    "SchemaVersionId": "string",
    "SchemaVersionNumber": number
  },
  "SerdeInfo": {
    "Name": "string",
    "Parameters": {
      "string" : "string"
    },
    "SerializationLibrary": "string"
  },
  "SkewedInfo": {
    "SkewedColumnNames": [ "string" ],
    "SkewedColumnValueLocationMaps": {
      "string" : "string"
    },
    "SkewedColumnValues": [ "string" ]
  },
  "SortColumns": [
    {
      "Column": "string",
      "SortOrder": number
    }
  ],
  "StoredAsSubDirectories": boolean
},
"Values": [ "string" ]
},
"PartitionValueList": [ "string" ]
}
],
"TableName": "string"
}

```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the catalog in which the partition is to be updated. Currently, this should be the AWS account ID.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

DatabaseName

The name of the metadata database in which the partition is to be updated.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Entries

A list of up to 100 `BatchUpdatePartitionRequestEntry` objects to update.

Type: Array of [BatchUpdatePartitionRequestEntry](#) objects

Array Members: Minimum number of 1 item. Maximum number of 100 items.

Required: Yes

TableName

The name of the metadata table in which the partition is to be updated.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "Errors": [
    {
      "ErrorDetail": {
        "ErrorCode": "string",
        "ErrorMessage": "string"
      },
      "PartitionValueList": [ "string" ]
    }
  ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Errors

The errors encountered when trying to update the requested partitions. A list of `BatchUpdatePartitionFailureEntry` objects.

Type: Array of [BatchUpdatePartitionFailureEntry](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CancelDataQualityRuleRecommendationRun

Cancels the specified recommendation run that was being used to generate rules.

Request Syntax

```
{  
  "RunId": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

[RunId](#)

The unique run identifier associated with this run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\u007F\u00E0-\u00FF\uD800-\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CancelDataQualityRulesetEvaluationRun

Cancels a run where a ruleset is being evaluated against a data source.

Request Syntax

```
{
  "RunId": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

RunId

The unique run identifier associated with this run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CancelMLTaskRun

Cancels (stops) a task run. Machine learning task runs are asynchronous tasks that AWS Glue runs on your behalf as part of various machine learning workflows. You can cancel a machine learning task run at any time by calling `CancelMLTaskRun` with a task run's parent transform's `TransformID` and the task run's `TaskRunId`.

Request Syntax

```
{
  "TaskRunId": "string",
  "TransformId": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

TaskRunId

A unique identifier for the task run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

TransformId

The unique identifier of the machine learning transform.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "Status": "string",
  "TaskRunId": "string",
  "TransformId": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Status

The status for this run.

Type: String

Valid Values: STARTING | RUNNING | STOPPING | STOPPED | SUCCEEDED | FAILED | TIMEOUT

TaskRunId

The unique identifier for the task run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

TransformId

The unique identifier of the machine learning transform.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)

- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CancelStatement

Cancels the statement.

Request Syntax

```
{  
  "Id": number,  
  "RequestOrigin": "string",  
  "SessionId": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Id

The ID of the statement to be cancelled.

Type: Integer

Required: Yes

RequestOrigin

The origin of the request to cancel the statement.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[\.\-_\A-Za-z0-9]+`

Required: No

SessionId

The Session ID of the statement to be cancelled.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

IllegalSessionStateException

The session is in an invalid state to perform a requested operation.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CheckSchemaVersionValidity

Validates the supplied schema. This call has no side effects, it simply validates using the supplied schema using DataFormat as the format. Since it does not take a schema set name, no compatibility checks are performed.

Request Syntax

```
{
  "DataFormat": "string",
  "SchemaDefinition": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

DataFormat

The data format of the schema definition. Currently AVRO, JSON and PROTOBUF are supported.

Type: String

Valid Values: AVRO | JSON | PROTOBUF

Required: Yes

SchemaDefinition

The definition of the schema that has to be validated.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 170000.

Pattern: .*\\S.*

Required: Yes

Response Syntax

```
{  
  "Error": "string",  
  "Valid": boolean  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Error

A validation failure error message.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 5000.

Valid

Return true, if the schema is valid and false otherwise.

Type: Boolean

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreateBlueprint

Registers a blueprint with AWS Glue.

Request Syntax

```
{
  "BlueprintLocation": "string",
  "Description": "string",
  "Name": "string",
  "Tags": {
    "string" : "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

BlueprintLocation

Specifies a path in Amazon S3 where the blueprint is published.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 8192.

Pattern: `^s3://([^\s/]+)/(^[^\s/]+/)*([^\s/]+)$`

Required: Yes

Description

A description of the blueprint.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 512.

Required: No

Name

The name of the blueprint.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[\.\-_\A-Za-z0-9]+`

Required: Yes

Tags

The tags to be applied to this blueprint.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Required: No

Response Syntax

```
{
  "Name": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Name

Returns the name of the blueprint that was registered.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)

- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreateClassifier

Creates a classifier in the user's account. This can be a `GrokClassifier`, an `XMLClassifier`, a `JsonClassifier`, or a `CsvClassifier`, depending on which field of the request is present.

Request Syntax

```
{
  "CsvClassifier": {
    "AllowSingleColumn": boolean,
    "ContainsHeader": "string",
    "CustomDatatypeConfigured": boolean,
    "CustomDatatypes": [ "string" ],
    "Delimiter": "string",
    "DisableValueTrimming": boolean,
    "Header": [ "string" ],
    "Name": "string",
    "QuoteSymbol": "string",
    "Serde": "string"
  },
  "GrokClassifier": {
    "Classification": "string",
    "CustomPatterns": "string",
    "GrokPattern": "string",
    "Name": "string"
  },
  "JsonClassifier": {
    "JsonPath": "string",
    "Name": "string"
  },
  "XMLClassifier": {
    "Classification": "string",
    "Name": "string",
    "RowTag": "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CsvClassifier

A `CsvClassifier` object specifying the classifier to create.

Type: [CreateCsvClassifierRequest](#) object

Required: No

GrokClassifier

A `GrokClassifier` object specifying the classifier to create.

Type: [CreateGrokClassifierRequest](#) object

Required: No

JsonClassifier

A `JsonClassifier` object specifying the classifier to create.

Type: [CreateJsonClassifierRequest](#) object

Required: No

XMLClassifier

An `XMLClassifier` object specifying the classifier to create.

Type: [CreateXMLClassifierRequest](#) object

Required: No

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreateConnection

Creates a connection definition in the Data Catalog.

Connections used for creating federated resources require the IAM `glue:PassConnection` permission.

Request Syntax

```
{
  "CatalogId": "string",
  "ConnectionInput": {
    "AuthenticationConfiguration": {
      "AuthenticationType": "string",
      "OAuth2Properties": {
        "AuthorizationCodeProperties": {
          "AuthorizationCode": "string",
          "RedirectUri": "string"
        },
        "OAuth2ClientApplication": {
          "AWSManagedClientApplicationReference": "string",
          "UserManagedClientApplicationClientId": "string"
        },
        "OAuth2GrantType": "string",
        "TokenUrl": "string",
        "TokenUrlParametersMap": {
          "string": "string"
        }
      },
      "SecretArn": "string"
    },
    "ConnectionProperties": {
      "string": "string"
    },
    "ConnectionType": "string",
    "Description": "string",
    "MatchCriteria": [ "string" ],
    "Name": "string",
    "PhysicalConnectionRequirements": {
      "AvailabilityZone": "string",
      "SecurityGroupIdList": [ "string" ],
      "SubnetId": "string"
    }
  },
}
```

```
    "ValidateCredentials": boolean
  },
  "Tags": {
    "string" : "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog in which to create the connection. If none is provided, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

ConnectionInput

A ConnectionInput object defining the connection to create.

Type: [ConnectionInput](#) object

Required: Yes

Tags

The tags you assign to the connection.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Required: No

Response Syntax

```
{  
  "CreateConnectionStatus": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

[CreateConnectionStatus](#)

The status of the connection creation request. The request can take some time for certain authentication types, for example when creating an OAuth connection with token exchange over VPC.

Type: String

Valid Values: READY | IN_PROGRESS | FAILED

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreateCrawler

Creates a new crawler with specified targets, role, configuration, and optional schedule. At least one crawl target must be specified, in the `s3Targets` field, the `jdbcTargets` field, or the `DynamoDBTargets` field.

Request Syntax

```
{
  "Classifiers": [ "string" ],
  "Configuration": "string",
  "CrawlerSecurityConfiguration": "string",
  "DatabaseName": "string",
  "Description": "string",
  "LakeFormationConfiguration": {
    "AccountId": "string",
    "UseLakeFormationCredentials": boolean
  },
  "LineageConfiguration": {
    "CrawlerLineageSettings": "string"
  },
  "Name": "string",
  "RecrawlPolicy": {
    "RecrawlBehavior": "string"
  },
  "Role": "string",
  "Schedule": "string",
  "SchemaChangePolicy": {
    "DeleteBehavior": "string",
    "UpdateBehavior": "string"
  },
  "TablePrefix": "string",
  "Tags": {
    "string" : "string"
  },
  "Targets": {
    "CatalogTargets": [
      {
        "ConnectionName": "string",
        "DatabaseName": "string",
        "DlqEventQueueArn": "string",
        "EventQueueArn": "string",
```

```
    "Tables": [ "string" ]
  }
],
"DeltaTargets": [
  {
    "ConnectionName": "string",
    "CreateNativeDeltaTable": boolean,
    "DeltaTables": [ "string" ],
    "WriteManifest": boolean
  }
],
"DynamoDBTargets": [
  {
    "Path": "string",
    "scanAll": boolean,
    "scanRate": number
  }
],
"HudiTargets": [
  {
    "ConnectionName": "string",
    "Exclusions": [ "string" ],
    "MaximumTraversalDepth": number,
    "Paths": [ "string" ]
  }
],
"IcebergTargets": [
  {
    "ConnectionName": "string",
    "Exclusions": [ "string" ],
    "MaximumTraversalDepth": number,
    "Paths": [ "string" ]
  }
],
"JdbcTargets": [
  {
    "ConnectionName": "string",
    "EnableAdditionalMetadata": [ "string" ],
    "Exclusions": [ "string" ],
    "Path": "string"
  }
],
"MongoDBTargets": [
  {
```

```

        "ConnectionName": "string",
        "Path": "string",
        "ScanAll": boolean
    }
],
"S3Targets": [
    {
        "ConnectionName": "string",
        "DlqEventQueueArn": "string",
        "EventQueueArn": "string",
        "Exclusions": [ "string" ],
        "Path": "string",
        "SampleSize": number
    }
]
}
}

```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Classifiers

A list of custom classifiers that the user has registered. By default, all built-in classifiers are included in a crawl, but these custom classifiers always override the default classifiers for a given classification.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Configuration

Crawler configuration information. This versioned JSON string allows users to specify aspects of a crawler's behavior. For more information, see [Setting crawler configuration options](#).

Type: String

Required: No

CrawlerSecurityConfiguration

The name of the SecurityConfiguration structure to be used by this crawler.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 128.

Required: No

DatabaseName

The AWS Glue database where results are written, such as: `arn:aws:daylight:us-east-1::database/sometable/*`.

Type: String

Required: No

Description

A description of the new crawler.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

LakeFormationConfiguration

Specifies AWS Lake Formation configuration settings for the crawler.

Type: [LakeFormationConfiguration](#) object

Required: No

LineageConfiguration

Specifies data lineage configuration settings for the crawler.

Type: [LineageConfiguration](#) object

Required: No

Name

Name of the new crawler.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\u007F\u00E0-\u00FF\uD800-\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

RecrawlPolicy

A policy that specifies whether to crawl the entire dataset again, or to crawl only folders that were added since the last crawler run.

Type: [RecrawlPolicy](#) object

Required: No

Role

The IAM role or Amazon Resource Name (ARN) of an IAM role used by the new crawler to access customer resources.

Type: String

Required: Yes

Schedule

A cron expression used to specify the schedule (see [Time-Based Schedules for Jobs and Crawlers](#)). For example, to run something every day at 12:15 UTC, you would specify: `cron(15 12 * * ? *)`.

Type: String

Required: No

SchemaChangePolicy

The policy for the crawler's update and deletion behavior.

Type: [SchemaChangePolicy](#) object

Required: No

[TablePrefix](#)

The table prefix used for catalog tables that are created.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 128.

Required: No

[Tags](#)

The tags to use with this crawler request. You may use tags to limit access to the crawler. For more information about tags in AWS Glue, see [AWS Tags in AWS Glue](#) in the developer guide.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Required: No

[Targets](#)

A list of collection of targets to crawl.

Type: [CrawlerTargets](#) object

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreateCustomEntityType

Creates a custom pattern that is used to detect sensitive data across the columns and rows of your structured data.

Each custom pattern you create specifies a regular expression and an optional list of context words. If no context words are passed only a regular expression is checked.

Request Syntax

```
{
  "ContextWords": [ "string" ],
  "Name": "string",
  "RegexString": "string",
  "Tags": {
    "string" : "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

ContextWords

A list of context words. If none of these context words are found within the vicinity of the regular expression the data will not be detected as sensitive data.

If no context words are passed only a regular expression is checked.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 20 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Name

A name for the custom pattern that allows it to be retrieved or deleted later. This name must be unique per AWS account.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

RegexString

A regular expression string that is used for detecting sensitive data in a custom pattern.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Tags

A list of tags applied to the custom entity type.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Required: No

Response Syntax

```
{  
  "Name": "string"
```

```
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Name

The name of the custom pattern you created.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

IdempotentParameterMismatchException

The same unique identifier was associated with two different records.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreateDatabase

Creates a new database in a Data Catalog.

Request Syntax

```
{
  "CatalogId": "string",
  "DatabaseInput": {
    "CreateTableDefaultPermissions": [
      {
        "Permissions": [ "string" ],
        "Principal": {
          "DataLakePrincipalIdentifier": "string"
        }
      }
    ],
    "Description": "string",
    "FederatedDatabase": {
      "ConnectionName": "string",
      "Identifier": "string"
    },
    "LocationUri": "string",
    "Name": "string",
    "Parameters": {
      "string" : "string"
    },
    "TargetDatabase": {
      "CatalogId": "string",
      "DatabaseName": "string",
      "Region": "string"
    }
  },
  "Tags": {
    "string" : "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog in which to create the database. If none is provided, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

DatabaseInput

The metadata for the database.

Type: [DatabaseInput](#) object

Required: Yes

Tags

The tags you assign to the database.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Required: No

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

FederatedResourceAlreadyExistsException

A federated resource already exists.

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreateDataQualityRuleset

Creates a data quality ruleset with DQDL rules applied to a specified AWS Glue table.

You create the ruleset using the Data Quality Definition Language (DQDL). For more information, see the AWS Glue developer guide.

Request Syntax

```
{
  "ClientToken": "string",
  "Description": "string",
  "Name": "string",
  "Ruleset": "string",
  "Tags": {
    "string" : "string"
  },
  "TargetTable": {
    "CatalogId": "string",
    "DatabaseName": "string",
    "TableName": "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

ClientToken

Used for idempotency and is recommended to be set to a random ID (such as a UUID) to avoid creating or starting multiple instances of the same resource.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF\\t]*`

Required: No

Description

A description of the data quality ruleset.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Name

A unique name for the data quality ruleset.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Ruleset

A Data Quality Definition Language (DQDL) ruleset. For more information, see the AWS Glue developer guide.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 65536.

Required: Yes

Tags

A list of tags applied to the data quality ruleset.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Required: No

TargetTable

A target table associated with the data quality ruleset.

Type: [DataQualityTargetTable](#) object

Required: No

Response Syntax

```
{  
  "Name": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Name

A unique name for the data quality ruleset.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreateDevEndpoint

Creates a new development endpoint.

Request Syntax

```
{
  "Arguments": {
    "string" : "string"
  },
  "EndpointName": "string",
  "ExtraJarsS3Path": "string",
  "ExtraPythonLibsS3Path": "string",
  "GlueVersion": "string",
  "NumberOfNodes": number,
  "NumberOfWorkers": number,
  "PublicKey": "string",
  "PublicKeys": [ "string" ],
  "RoleArn": "string",
  "SecurityConfiguration": "string",
  "SecurityGroupIds": [ "string" ],
  "SubnetId": "string",
  "Tags": {
    "string" : "string"
  },
  "WorkerType": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Arguments

A map of arguments used to configure the DevEndpoint.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 100 items.

Required: No

EndpointName

The name to be assigned to the new DevEndpoint.

Type: String

Required: Yes

ExtraJarsS3Path

The path to one or more Java `.jar` files in an S3 bucket that should be loaded in your DevEndpoint.

Type: String

Required: No

ExtraPythonLibsS3Path

The paths to one or more Python libraries in an Amazon S3 bucket that should be loaded in your DevEndpoint. Multiple values must be complete paths separated by a comma.

Note

You can only use pure Python libraries with a DevEndpoint. Libraries that rely on C extensions, such as the [pandas](#) Python data analysis library, are not yet supported.

Type: String

Required: No

GlueVersion

Glue version determines the versions of Apache Spark and Python that AWS Glue supports. The Python version indicates the version supported for running your ETL scripts on development endpoints.

For more information about the available AWS Glue versions and corresponding Spark and Python versions, see [Glue version](#) in the developer guide.

Development endpoints that are created without specifying a Glue version default to Glue 0.9.

You can specify a version of Python support for development endpoints by using the `Arguments` parameter in the `CreateDevEndpoint` or `UpdateDevEndpoint` APIs. If no arguments are provided, the version defaults to Python 2.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `^\w+\.\w+$`

Required: No

NumberOfNodes

The number of AWS Glue Data Processing Units (DPUs) to allocate to this `DevEndpoint`.

Type: Integer

Required: No

NumberOfWorkers

The number of workers of a defined `workerType` that are allocated to the development endpoint.

The maximum number of workers you can define are 299 for `G.1X`, and 149 for `G.2X`.

Type: Integer

Required: No

PublicKey

The public key to be used by this `DevEndpoint` for authentication. This attribute is provided for backward compatibility because the recommended attribute to use is public keys.

Type: String

Required: No

PublicKeys

A list of public keys to be used by the development endpoints for authentication. The use of this attribute is preferred over a single public key because the public keys allow you to have a different private key per client.

Note

If you previously created an endpoint with a public key, you must remove that key to be able to set a list of public keys. Call the `UpdateDevEndpoint` API with the public key content in the `deletePublicKeys` attribute, and the list of new keys in the `addPublicKeys` attribute.

Type: Array of strings

Array Members: Maximum number of 5 items.

Required: No

RoleArn

The IAM role for the DevEndpoint.

Type: String

Pattern: `arn:aws:iam::\d{12}:role/.*`

Required: Yes

SecurityConfiguration

The name of the SecurityConfiguration structure to be used with this DevEndpoint.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

SecurityGroupIds

Security group IDs for the security groups to be used by the new DevEndpoint.

Type: Array of strings

Required: No

SubnetId

The subnet ID for the new DevEndpoint to use.

Type: String

Required: No

Tags

The tags to use with this DevEndpoint. You may use tags to limit access to the DevEndpoint. For more information about tags in AWS Glue, see [AWS Tags in AWS Glue](#) in the developer guide.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Required: No

WorkerType

The type of predefined worker that is allocated to the development endpoint. Accepts a value of Standard, G.1X, or G.2X.

- For the Standard worker type, each worker provides 4 vCPU, 16 GB of memory and a 50GB disk, and 2 executors per worker.
- For the G.1X worker type, each worker maps to 1 DPU (4 vCPU, 16 GB of memory, 64 GB disk), and provides 1 executor per worker. We recommend this worker type for memory-intensive jobs.
- For the G.2X worker type, each worker maps to 2 DPU (8 vCPU, 32 GB of memory, 128 GB disk), and provides 1 executor per worker. We recommend this worker type for memory-intensive jobs.

Known issue: when a development endpoint is created with the G.2X WorkerType configuration, the Spark drivers for the development endpoint will run on 4 vCPU, 16 GB of memory, and a 64 GB disk.

Type: String

Valid Values: Standard | G.1X | G.2X | G.025X | G.4X | G.8X | Z.2X

Required: No

Response Syntax

```
{
  "Arguments": {
    "string" : "string"
  },
  "AvailabilityZone": "string",
  "CreatedTimestamp": number,
  "EndpointName": "string",
  "ExtraJarsS3Path": "string",
  "ExtraPythonLibsS3Path": "string",
  "FailureReason": "string",
  "GlueVersion": "string",
  "NumberOfNodes": number,
  "NumberOfWorkers": number,
  "RoleArn": "string",
  "SecurityConfiguration": "string",
  "SecurityGroupIds": [ "string" ],
  "Status": "string",
  "SubnetId": "string",
  "VpcId": "string",
  "WorkerType": "string",
  "YarnEndpointAddress": "string",
  "ZeppelinRemoteSparkInterpreterPort": number
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Arguments

The map of arguments used to configure this DevEndpoint.

Valid arguments are:

- `--enable-glue-datacatalog": ""`

You can specify a version of Python support for development endpoints by using the `Arguments` parameter in the `CreateDevEndpoint` or `UpdateDevEndpoint` APIs. If no arguments are provided, the version defaults to Python 2.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 100 items.

AvailabilityZone

The AWS Availability Zone where this `DevEndpoint` is located.

Type: String

CreatedTimestamp

The point in time at which this `DevEndpoint` was created.

Type: Timestamp

EndpointName

The name assigned to the new `DevEndpoint`.

Type: String

ExtraJarsS3Path

Path to one or more Java `.jar` files in an S3 bucket that will be loaded in your `DevEndpoint`.

Type: String

ExtraPythonLibsS3Path

The paths to one or more Python libraries in an S3 bucket that will be loaded in your `DevEndpoint`.

Type: String

FailureReason

The reason for a current failure in this `DevEndpoint`.

Type: String

GlueVersion

Glue version determines the versions of Apache Spark and Python that AWS Glue supports. The Python version indicates the version supported for running your ETL scripts on development endpoints.

For more information about the available AWS Glue versions and corresponding Spark and Python versions, see [Glue version](#) in the developer guide.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `^\w+\.\w+$`

NumberOfNodes

The number of AWS Glue Data Processing Units (DPUs) allocated to this DevEndpoint.

Type: Integer

NumberOfWorkers

The number of workers of a defined `workerType` that are allocated to the development endpoint.

Type: Integer

RoleArn

The Amazon Resource Name (ARN) of the role assigned to the new DevEndpoint.

Type: String

Pattern: `arn:aws:iam::\d{12}:role/.*`

SecurityConfiguration

The name of the `SecurityConfiguration` structure being used with this DevEndpoint.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\u007F\u00E0-\u00FF\uD800-\uDC00-\uDBFF\uDFFF\t]*`

SecurityGroupIds

The security groups assigned to the new DevEndpoint.

Type: Array of strings

Status

The current status of the new DevEndpoint.

Type: String

SubnetId

The subnet ID assigned to the new DevEndpoint.

Type: String

VpcId

The ID of the virtual private cloud (VPC) used by this DevEndpoint.

Type: String

WorkerType

The type of predefined worker that is allocated to the development endpoint. May be a value of Standard, G.1X, or G.2X.

Type: String

Valid Values: Standard | G.1X | G.2X | G.025X | G.4X | G.8X | Z.2X

YarnEndpointAddress

The address of the YARN endpoint used by this DevEndpoint.

Type: String

ZeppelinRemoteSparkInterpreterPort

The Apache Zeppelin port for the remote Apache Spark interpreter.

Type: Integer

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

IdempotentParameterMismatchException

The same unique identifier was associated with two different records.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

ValidationException

A value could not be validated.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreateJob

Creates a new job definition.

Request Syntax

```
{
  "AllocatedCapacity": number,
  "CodeGenConfigurationNodes": {
    "string" : {
      "Aggregate": {
        "Aggs": [
          {
            "AggFunc": "string",
            "Column": [ "string" ]
          }
        ],
        "Groups": [
          [ "string" ]
        ],
        "Inputs": [ "string" ],
        "Name": "string"
      },
      "AmazonRedshiftSource": {
        "Data": {
          "AccessType": "string",
          "Action": "string",
          "AdvancedOptions": [
            {
              "Key": "string",
              "Value": "string"
            }
          ],
          "CatalogDatabase": {
            "Description": "string",
            "Label": "string",
            "Value": "string"
          },
          "CatalogRedshiftSchema": "string",
          "CatalogRedshiftTable": "string",
          "CatalogTable": {
            "Description": "string",
            "Label": "string",
```

```
    "Value": "string"
  },
  "Connection": {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  },
  "CrawlerConnection": "string",
  "IamRole": {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  },
  "MergeAction": "string",
  "MergeClause": "string",
  "MergeWhenMatched": "string",
  "MergeWhenNotMatched": "string",
  "PostAction": "string",
  "PreAction": "string",
  "SampleQuery": "string",
  "Schema": {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  },
  "SelectedColumns": [
    {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    }
  ],
  "SourceType": "string",
  "StagingTable": "string",
  "Table": {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  },
  "TablePrefix": "string",
  "TableSchema": [
    {
      "Description": "string",
      "Label": "string",
```

```

        "Value": "string"
    }
],
"TempDir": "string",
"Upsert": boolean
},
"Name": "string"
},
"AmazonRedshiftTarget": {
    "Data": {
        "AccessType": "string",
        "Action": "string",
        "AdvancedOptions": [
            {
                "Key": "string",
                "Value": "string"
            }
        ],
        "CatalogDatabase": {
            "Description": "string",
            "Label": "string",
            "Value": "string"
        },
        "CatalogRedshiftSchema": "string",
        "CatalogRedshiftTable": "string",
        "CatalogTable": {
            "Description": "string",
            "Label": "string",
            "Value": "string"
        },
        "Connection": {
            "Description": "string",
            "Label": "string",
            "Value": "string"
        },
        "CrawlerConnection": "string",
        "IamRole": {
            "Description": "string",
            "Label": "string",
            "Value": "string"
        },
        "MergeAction": "string",
        "MergeClause": "string",
        "MergeWhenMatched": "string",

```

```
"MergeWhenNotMatched": "string",
"PostAction": "string",
"PreAction": "string",
"SampleQuery": "string",
"Schema": {
  "Description": "string",
  "Label": "string",
  "Value": "string"
},
"SelectedColumns": [
  {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  }
],
"SourceType": "string",
"StagingTable": "string",
"Table": {
  "Description": "string",
  "Label": "string",
  "Value": "string"
},
"TablePrefix": "string",
"TableSchema": [
  {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  }
],
"TempDir": "string",
"Upsert": boolean
},
"Inputs": [ "string" ],
"Name": "string"
},
"ApplyMapping": {
  "Inputs": [ "string" ],
  "Mapping": [
    {
      "Children": [
        "Mapping"
      ]
    }
  ]
},
```

```

        "Dropped": boolean,
        "FromPath": [ "string" ],
        "FromType": "string",
        "ToKey": "string",
        "ToType": "string"
    }
],
    "Name": "string"
},
"AthenaConnectorSource": {
    "ConnectionName": "string",
    "ConnectionTable": "string",
    "ConnectionType": "string",
    "ConnectorName": "string",
    "Name": "string",
    "OutputSchemas": [
        {
            "Columns": [
                {
                    "Name": "string",
                    "Type": "string"
                }
            ]
        }
    ],
    "SchemaName": "string"
},
"CatalogDeltaSource": {
    "AdditionalDeltaOptions": {
        "string" : "string"
    },
    "Database": "string",
    "Name": "string",
    "OutputSchemas": [
        {
            "Columns": [
                {
                    "Name": "string",
                    "Type": "string"
                }
            ]
        }
    ],
    "Table": "string"
}

```

```
},
  "CatalogHudiSource": {
    "AdditionalHudiOptions": {
      "string": "string"
    },
    "Database": "string",
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
    "Table": "string"
  },
  "CatalogKafkaSource": {
    "Database": "string",
    "DataPreviewOptions": {
      "PollingTime": number,
      "RecordPollingLimit": number
    },
    "DetectSchema": boolean,
    "Name": "string",
    "StreamingOptions": {
      "AddRecordTimestamp": "string",
      "Assign": "string",
      "BootstrapServers": "string",
      "Classification": "string",
      "ConnectionName": "string",
      "Delimiter": "string",
      "EmitConsumerLagMetrics": "string",
      "EndingOffsets": "string",
      "IncludeHeaders": boolean,
      "MaxOffsetsPerTrigger": number,
      "MinPartitions": number,
      "NumRetries": number,
      "PollTimeoutMs": number,
      "RetryIntervalMs": number,
      "SecurityProtocol": "string",
      "StartingOffsets": "string",
```

```
    "StartingTimestamp": "string",
    "SubscribePattern": "string",
    "TopicName": "string"
  },
  "Table": "string",
  "WindowSize": number
},
"CatalogKinesisSource": {
  "Database": "string",
  "DataPreviewOptions": {
    "PollingTime": number,
    "RecordPollingLimit": number
  },
  "DetectSchema": boolean,
  "Name": "string",
  "StreamingOptions": {
    "AddIdleTimeBetweenReads": boolean,
    "AddRecordTimestamp": "string",
    "AvoidEmptyBatches": boolean,
    "Classification": "string",
    "Delimiter": "string",
    "DescribeShardInterval": number,
    "EmitConsumerLagMetrics": "string",
    "EndpointUrl": "string",
    "IdleTimeBetweenReadsInMs": number,
    "MaxFetchRecordsPerShard": number,
    "MaxFetchTimeInMs": number,
    "MaxRecordPerRead": number,
    "MaxRetryIntervalMs": number,
    "NumRetries": number,
    "RetryIntervalMs": number,
    "RoleArn": "string",
    "RoleSessionName": "string",
    "StartingPosition": "string",
    "StartingTimestamp": "string",
    "StreamArn": "string",
    "StreamName": "string"
  },
  "Table": "string",
  "WindowSize": number
},
"CatalogSource": {
  "Database": "string",
  "Name": "string",
```

```
    "Table": "string"
  },
  "CatalogTarget": {
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "Table": "string"
  },
  "ConnectorDataSource": {
    "ConnectionType": "string",
    "Data": {
      "string" : "string"
    },
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ]
  },
  "ConnectorDataTarget": {
    "ConnectionType": "string",
    "Data": {
      "string" : "string"
    },
    "Inputs": [ "string" ],
    "Name": "string"
  },
  "CustomCode": {
    "ClassName": "string",
    "Code": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ]
  }
}
```



```

    }
  ]
}
},
"DirectJDBCSource": {
  "ConnectionName": "string",
  "ConnectionType": "string",
  "Database": "string",
  "Name": "string",
  "RedshiftTmpDir": "string",
  "Table": "string"
},
"DirectKafkaSource": {
  "DataPreviewOptions": {
    "PollingTime": number,
    "RecordPollingLimit": number
  },
  "DetectSchema": boolean,
  "Name": "string",
  "StreamingOptions": {
    "AddRecordTimestamp": "string",
    "Assign": "string",
    "BootstrapServers": "string",
    "Classification": "string",
    "ConnectionName": "string",
    "Delimiter": "string",
    "EmitConsumerLagMetrics": "string",
    "EndingOffsets": "string",
    "IncludeHeaders": boolean,
    "MaxOffsetsPerTrigger": number,
    "MinPartitions": number,
    "NumRetries": number,
    "PollTimeoutMs": number,
    "RetryIntervalMs": number,
    "SecurityProtocol": "string",
    "StartingOffsets": "string",
    "StartingTimestamp": "string",
    "SubscribePattern": "string",
    "TopicName": "string"
  },
  "WindowSize": number
},
"DirectKinesisSource": {

```

```
"DataPreviewOptions": {
  "PollingTime": number,
  "RecordPollingLimit": number
},
"DetectSchema": boolean,
"Name": "string",
"StreamingOptions": {
  "AddIdleTimeBetweenReads": boolean,
  "AddRecordTimestamp": "string",
  "AvoidEmptyBatches": boolean,
  "Classification": "string",
  "Delimiter": "string",
  "DescribeShardInterval": number,
  "EmitConsumerLagMetrics": "string",
  "EndpointUrl": "string",
  "IdleTimeBetweenReadsInMs": number,
  "MaxFetchRecordsPerShard": number,
  "MaxFetchTimeInMs": number,
  "MaxRecordPerRead": number,
  "MaxRetryIntervalMs": number,
  "NumRetries": number,
  "RetryIntervalMs": number,
  "RoleArn": "string",
  "RoleSessionName": "string",
  "StartingPosition": "string",
  "StartingTimestamp": "string",
  "StreamArn": "string",
  "StreamName": "string"
},
"WindowSize": number
},
"DropDuplicates": {
  "Columns": [
    [ "string" ]
  ],
  "Inputs": [ "string" ],
  "Name": "string"
},
"DropFields": {
  "Inputs": [ "string" ],
  "Name": "string",
  "Paths": [
    [ "string" ]
  ]
}
```

```
},
  "DropNullFields": {
    "Inputs": [ "string" ],
    "Name": "string",
    "NullCheckBoxList": {
      "IsEmpty": boolean,
      "IsNegOne": boolean,
      "IsNullString": boolean
    },
    "NullTextList": [
      {
        "Datatype": {
          "Id": "string",
          "Label": "string"
        },
        "Value": "string"
      }
    ]
  },
  "DynamicTransform": {
    "FunctionName": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ]
  },
  "Parameters": [
    {
      "IsOptional": boolean,
      "ListType": "string",
      "Name": "string",
      "Type": "string",
      "ValidationMessage": "string",
      "ValidationRule": "string",
      "Value": [ "string" ]
    }
  ],
}
```

```
    "Path": "string",
    "TransformName": "string",
    "Version": "string"
  },
  "DynamoDBCatalogSource": {
    "Database": "string",
    "Name": "string",
    "Table": "string"
  },
  "EvaluateDataQuality": {
    "Inputs": [ "string" ],
    "Name": "string",
    "Output": "string",
    "PublishingOptions": {
      "CloudWatchMetricsEnabled": boolean,
      "EvaluationContext": "string",
      "ResultsPublishingEnabled": boolean,
      "ResultsS3Prefix": "string"
    },
    "Ruleset": "string",
    "StopJobOnFailureOptions": {
      "StopJobOnFailureTiming": "string"
    }
  },
  "EvaluateDataQualityMultiFrame": {
    "AdditionalDataSources": {
      "string" : "string"
    },
    "AdditionalOptions": {
      "string" : "string"
    },
    "Inputs": [ "string" ],
    "Name": "string",
    "PublishingOptions": {
      "CloudWatchMetricsEnabled": boolean,
      "EvaluationContext": "string",
      "ResultsPublishingEnabled": boolean,
      "ResultsS3Prefix": "string"
    },
    "Ruleset": "string",
    "StopJobOnFailureOptions": {
      "StopJobOnFailureTiming": "string"
    }
  },
}
```

```
"FillMissingValues": {
  "FilledPath": "string",
  "ImputedPath": "string",
  "Inputs": [ "string" ],
  "Name": "string"
},
"Filter": {
  "Filters": [
    {
      "Negated": boolean,
      "Operation": "string",
      "Values": [
        {
          "Type": "string",
          "Value": [ "string" ]
        }
      ]
    }
  ],
  "Inputs": [ "string" ],
  "LogicalOperator": "string",
  "Name": "string"
},
"GovernedCatalogSource": {
  "AdditionalOptions": {
    "BoundedFiles": number,
    "BoundedSize": number
  },
  "Database": "string",
  "Name": "string",
  "PartitionPredicate": "string",
  "Table": "string"
},
"GovernedCatalogTarget": {
  "Database": "string",
  "Inputs": [ "string" ],
  "Name": "string",
  "PartitionKeys": [
    [ "string" ]
  ],
  "SchemaChangePolicy": {
    "EnableUpdateCatalog": boolean,
    "UpdateBehavior": "string"
  }
},
```

```
    "Table": "string"
  },
  "JDBCConnectorSource": {
    "AdditionalOptions": {
      "DataTypeMapping": {
        "string": "string"
      },
      "FilterPredicate": "string",
      "JobBookmarkKeys": [ "string" ],
      "JobBookmarkKeysSortOrder": "string",
      "LowerBound": number,
      "NumPartitions": number,
      "PartitionColumn": "string",
      "UpperBound": number
    },
    "ConnectionName": "string",
    "ConnectionTable": "string",
    "ConnectionType": "string",
    "ConnectorName": "string",
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
    "Query": "string"
  },
  "JDBCConnectorTarget": {
    "AdditionalOptions": {
      "string": "string"
    },
    "ConnectionName": "string",
    "ConnectionTable": "string",
    "ConnectionType": "string",
    "ConnectorName": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "OutputSchemas": [
      {
```

```
        "Columns": [
            {
                "Name": "string",
                "Type": "string"
            }
        ]
    },
    ],
    "Join": {
        "Columns": [
            {
                "From": "string",
                "Keys": [
                    [ "string" ]
                ]
            }
        ],
        "Inputs": [ "string" ],
        "JoinType": "string",
        "Name": "string"
    },
    "Merge": {
        "Inputs": [ "string" ],
        "Name": "string",
        "PrimaryKeys": [
            [ "string" ]
        ],
        "Source": "string"
    },
    "MicrosoftSQLServerCatalogSource": {
        "Database": "string",
        "Name": "string",
        "Table": "string"
    },
    "MicrosoftSQLServerCatalogTarget": {
        "Database": "string",
        "Inputs": [ "string" ],
        "Name": "string",
        "Table": "string"
    },
    "MySQLCatalogSource": {
        "Database": "string",
        "Name": "string",
```

```
    "Table": "string"
  },
  "MySQLCatalogTarget": {
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "Table": "string"
  },
  "OracleSQLCatalogSource": {
    "Database": "string",
    "Name": "string",
    "Table": "string"
  },
  "OracleSQLCatalogTarget": {
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "Table": "string"
  },
  "PIIDetection": {
    "EntityTypesToDetect": [ "string" ],
    "Inputs": [ "string" ],
    "MaskValue": "string",
    "Name": "string",
    "OutputColumnName": "string",
    "PiiType": "string",
    "SampleFraction": number,
    "ThresholdFraction": number
  },
  "PostgreSQLCatalogSource": {
    "Database": "string",
    "Name": "string",
    "Table": "string"
  },
  "PostgreSQLCatalogTarget": {
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "Table": "string"
  },
  "Recipe": {
    "Inputs": [ "string" ],
    "Name": "string",
    "RecipeReference": {
```



```

        "RecipeArn": "string",
        "RecipeVersion": "string"
    }
},
"RedshiftSource": {
    "Database": "string",
    "Name": "string",
    "RedshiftTmpDir": "string",
    "Table": "string",
    "TmpDirIAMRole": "string"
},
"RedshiftTarget": {
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "RedshiftTmpDir": "string",
    "Table": "string",
    "TmpDirIAMRole": "string",
    "UpsertRedshiftOptions": {
        "ConnectionName": "string",
        "TableLocation": "string",
        "UpsertKeys": [ "string" ]
    }
},
"RelationalCatalogSource": {
    "Database": "string",
    "Name": "string",
    "Table": "string"
},
"RenameField": {
    "Inputs": [ "string" ],
    "Name": "string",
    "SourcePath": [ "string" ],
    "TargetPath": [ "string" ]
},
"S3CatalogDeltaSource": {
    "AdditionalDeltaOptions": {
        "string" : "string"
    },
    "Database": "string",
    "Name": "string",
    "OutputSchemas": [
        {
            "Columns": [

```

```

        {
            "Name": "string",
            "Type": "string"
        }
    ]
}
],
"Table": "string"
},
"S3CatalogHudiSource": {
    "AdditionalHudiOptions": {
        "string" : "string"
    },
    "Database": "string",
    "Name": "string",
    "OutputSchemas": [
        {
            "Columns": [
                {
                    "Name": "string",
                    "Type": "string"
                }
            ]
        }
    ],
    "Table": "string"
},
"S3CatalogSource": {
    "AdditionalOptions": {
        "BoundedFiles": number,
        "BoundedSize": number
    },
    "Database": "string",
    "Name": "string",
    "PartitionPredicate": "string",
    "Table": "string"
},
"S3CatalogTarget": {
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
        [ "string" ]
    ],
}

```

```
    "SchemaChangePolicy": {
      "EnableUpdateCatalog": boolean,
      "UpdateBehavior": "string"
    },
    "Table": "string"
  },
  "S3CsvSource": {
    "AdditionalOptions": {
      "BoundedFiles": number,
      "BoundedSize": number,
      "EnableSamplePath": boolean,
      "SamplePath": "string"
    },
    "CompressionType": "string",
    "Escaper": "string",
    "Exclusions": [ "string " ],
    "GroupFiles": "string",
    "GroupSize": "string",
    "MaxBand": number,
    "MaxFilesInBand": number,
    "Multiline": boolean,
    "Name": "string",
    "OptimizePerformance": boolean,
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
    "Paths": [ "string " ],
    "QuoteChar": "string",
    "Recurse": boolean,
    "Separator": "string",
    "SkipFirst": boolean,
    "WithHeader": boolean,
    "WriteHeader": boolean
  },
  "S3DeltaCatalogTarget": {
    "AdditionalOptions": {
      "string": "string"
    }
  }
}
```

```
    },
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
      [ "string" ]
    ],
    "SchemaChangePolicy": {
      "EnableUpdateCatalog": boolean,
      "UpdateBehavior": "string"
    },
    "Table": "string"
  },
  "S3DeltaDirectTarget": {
    "AdditionalOptions": {
      "string" : "string"
    },
    "Compression": "string",
    "Format": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
      [ "string" ]
    ],
    "Path": "string",
    "SchemaChangePolicy": {
      "Database": "string",
      "EnableUpdateCatalog": boolean,
      "Table": "string",
      "UpdateBehavior": "string"
    }
  },
  "S3DeltaSource": {
    "AdditionalDeltaOptions": {
      "string" : "string"
    },
    "AdditionalOptions": {
      "BoundedFiles": number,
      "BoundedSize": number,
      "EnableSamplePath": boolean,
      "SamplePath": "string"
    },
    "Name": "string",
    "OutputSchemas": [
```

```

    {
      "Columns": [
        {
          "Name": "string",
          "Type": "string"
        }
      ]
    }
  ],
  "Paths": [ "string" ]
},
"S3DirectTarget": {
  "Compression": "string",
  "Format": "string",
  "Inputs": [ "string" ],
  "Name": "string",
  "PartitionKeys": [
    [ "string" ]
  ],
  "Path": "string",
  "SchemaChangePolicy": {
    "Database": "string",
    "EnableUpdateCatalog": boolean,
    "Table": "string",
    "UpdateBehavior": "string"
  }
},
"S3GlueParquetTarget": {
  "Compression": "string",
  "Inputs": [ "string" ],
  "Name": "string",
  "PartitionKeys": [
    [ "string" ]
  ],
  "Path": "string",
  "SchemaChangePolicy": {
    "Database": "string",
    "EnableUpdateCatalog": boolean,
    "Table": "string",
    "UpdateBehavior": "string"
  }
},
"S3HudiCatalogTarget": {
  "AdditionalOptions": {

```

```

        "string" : "string"
    },
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
        [ "string" ]
    ],
    "SchemaChangePolicy": {
        "EnableUpdateCatalog": boolean,
        "UpdateBehavior": "string"
    },
    "Table": "string"
},
"S3HudiDirectTarget": {
    "AdditionalOptions": {
        "string" : "string"
    },
    "Compression": "string",
    "Format": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
        [ "string" ]
    ],
    "Path": "string",
    "SchemaChangePolicy": {
        "Database": "string",
        "EnableUpdateCatalog": boolean,
        "Table": "string",
        "UpdateBehavior": "string"
    }
},
"S3HudiSource": {
    "AdditionalHudiOptions": {
        "string" : "string"
    },
    "AdditionalOptions": {
        "BoundedFiles": number,
        "BoundedSize": number,
        "EnableSamplePath": boolean,
        "SamplePath": "string"
    },
    "Name": "string",

```

```

    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
    "Paths": [ "string" ]
  },
  "S3JsonSource": {
    "AdditionalOptions": {
      "BoundedFiles": number,
      "BoundedSize": number,
      "EnableSamplePath": boolean,
      "SamplePath": "string"
    },
    "CompressionType": "string",
    "Exclusions": [ "string" ],
    "GroupFiles": "string",
    "GroupSize": "string",
    "JsonPath": "string",
    "MaxBand": number,
    "MaxFilesInBand": number,
    "Multiline": boolean,
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
    "Paths": [ "string" ],
    "Recurse": boolean
  },
  "S3ParquetSource": {
    "AdditionalOptions": {
      "BoundedFiles": number,

```

```

    "BoundedSize": number,
    "EnableSamplePath": boolean,
    "SamplePath": "string"
  },
  "CompressionType": "string",
  "Exclusions": [ "string" ],
  "GroupFiles": "string",
  "GroupSize": "string",
  "MaxBand": number,
  "MaxFilesInBand": number,
  "Name": "string",
  "OutputSchemas": [
    {
      "Columns": [
        {
          "Name": "string",
          "Type": "string"
        }
      ]
    }
  ],
  "Paths": [ "string" ],
  "Recurse": boolean
},
"SelectFields": {
  "Inputs": [ "string" ],
  "Name": "string",
  "Paths": [
    [ "string" ]
  ]
},
"SelectFromCollection": {
  "Index": number,
  "Inputs": [ "string" ],
  "Name": "string"
},
"SnowflakeSource": {
  "Data": {
    "Action": "string",
    "AdditionalOptions": {
      "string" : "string"
    },
  },
  "AutoPushdown": boolean,
  "Connection": {

```



```
    "Description": "string",
    "Label": "string",
    "Value": "string"
  },
  "Database": "string",
  "IamRole": {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  },
  "MergeAction": "string",
  "MergeClause": "string",
  "MergeWhenMatched": "string",
  "MergeWhenNotMatched": "string",
  "PostAction": "string",
  "PreAction": "string",
  "SampleQuery": "string",
  "Schema": "string",
  "SelectedColumns": [
    {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    }
  ],
  "SourceType": "string",
  "StagingTable": "string",
  "Table": "string",
  "TableSchema": [
    {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    }
  ],
  "TempDir": "string",
  "Upsert": boolean
},
"Name": "string",
"OutputSchemas": [
  {
    "Columns": [
      {
        "Name": "string",
```

```

        "Type": "string"
      }
    ]
  },
  "SnowflakeTarget": {
    "Data": {
      "Action": "string",
      "AdditionalOptions": {
        "string": "string"
      },
      "AutoPushdown": boolean,
      "Connection": {
        "Description": "string",
        "Label": "string",
        "Value": "string"
      },
      "Database": "string",
      "IamRole": {
        "Description": "string",
        "Label": "string",
        "Value": "string"
      },
      "MergeAction": "string",
      "MergeClause": "string",
      "MergeWhenMatched": "string",
      "MergeWhenNotMatched": "string",
      "PostAction": "string",
      "PreAction": "string",
      "SampleQuery": "string",
      "Schema": "string",
      "SelectedColumns": [
        {
          "Description": "string",
          "Label": "string",
          "Value": "string"
        }
      ],
      "SourceType": "string",
      "StagingTable": "string",
      "Table": "string",
      "TableSchema": [
        {

```

```
        "Description": "string",
        "Label": "string",
        "Value": "string"
    }
],
"TempDir": "string",
"Upsert": boolean
},
"Inputs": [ "string" ],
"Name": "string"
},
"SparkConnectorSource": {
    "AdditionalOptions": {
        "string" : "string"
    },
    "ConnectionName": "string",
    "ConnectionType": "string",
    "ConnectorName": "string",
    "Name": "string",
    "OutputSchemas": [
        {
            "Columns": [
                {
                    "Name": "string",
                    "Type": "string"
                }
            ]
        }
    ]
},
"SparkConnectorTarget": {
    "AdditionalOptions": {
        "string" : "string"
    },
    "ConnectionName": "string",
    "ConnectionType": "string",
    "ConnectorName": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "OutputSchemas": [
        {
            "Columns": [
                {
                    "Name": "string",
```

```

        "Type": "string"
      }
    ]
  },
  "SparkSQL": {
    "Inputs": [ "string" ],
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
    "SqlAliases": [
      {
        "Alias": "string",
        "From": "string"
      }
    ],
    "SqlQuery": "string"
  },
  "Spigot": {
    "Inputs": [ "string" ],
    "Name": "string",
    "Path": "string",
    "Prob": number,
    "Topk": number
  },
  "SplitFields": {
    "Inputs": [ "string" ],
    "Name": "string",
    "Paths": [
      [ "string" ]
    ]
  },
  "Union": {
    "Inputs": [ "string" ],
    "Name": "string",

```

```
        "UnionType": "string"
      }
    }
  },
  "Command": {
    "Name": "string",
    "PythonVersion": "string",
    "Runtime": "string",
    "ScriptLocation": "string"
  },
  "Connections": {
    "Connections": [ "string" ]
  },
  "DefaultArguments": {
    "string" : "string"
  },
  "Description": "string",
  "ExecutionClass": "string",
  "ExecutionProperty": {
    "MaxConcurrentRuns": number
  },
  "GlueVersion": "string",
  "JobMode": "string",
  "LogUri": "string",
  "MaintenanceWindow": "string",
  "MaxCapacity": number,
  "MaxRetries": number,
  "Name": "string",
  "NonOverridableArguments": {
    "string" : "string"
  },
  "NotificationProperty": {
    "NotifyDelayAfter": number
  },
  "NumberOfWorkers": number,
  "Role": "string",
  "SecurityConfiguration": "string",
  "SourceControlDetails": {
    "AuthStrategy": "string",
    "AuthToken": "string",
    "Branch": "string",
    "Folder": "string",
    "LastCommitId": "string",
    "Owner": "string",
```

```
    "Provider": "string",
    "Repository": "string"
  },
  "Tags": {
    "string" : "string"
  },
  "Timeout": number,
  "WorkerType": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

AllocatedCapacity

This parameter is deprecated. Use `MaxCapacity` instead.

The number of AWS Glue data processing units (DPUs) to allocate to this Job. You can allocate a minimum of 2 DPUs; the default is 10. A DPU is a relative measure of processing power that consists of 4 vCPUs of compute capacity and 16 GB of memory. For more information, see the [AWS Glue pricing page](#).

Type: Integer

Required: No

CodeGenConfigurationNodes

The representation of a directed acyclic graph on which both the Glue Studio visual component and Glue Studio code generation is based.

Type: String to [CodeGenConfigurationNode](#) object map

Key Pattern: `[A-Za-z0-9_-]*`

Required: No

Command

The `JobCommand` that runs this job.

Type: [JobCommand](#) object

Required: Yes

[Connections](#)

The connections used for this job.

Type: [ConnectionsList](#) object

Required: No

[DefaultArguments](#)

The default arguments for every run of this job, specified as name-value pairs.

You can specify arguments here that your own job-execution script consumes, as well as arguments that AWS Glue itself consumes.

Job arguments may be logged. Do not pass plaintext secrets as arguments. Retrieve secrets from a AWS Glue Connection, AWS Secrets Manager or other secret management mechanism if you intend to keep them within the Job.

For information about how to specify and consume your own Job arguments, see the [Calling AWS Glue APIs in Python](#) topic in the developer guide.

For information about the arguments you can provide to this field when configuring Spark jobs, see the [Special Parameters Used by AWS Glue](#) topic in the developer guide.

For information about the arguments you can provide to this field when configuring Ray jobs, see [Using job parameters in Ray jobs](#) in the developer guide.

Type: String to string map

Required: No

[Description](#)

Description of the job being defined.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

ExecutionClass

Indicates whether the job is run with a standard or flexible execution class. The standard execution-class is ideal for time-sensitive workloads that require fast job startup and dedicated resources.

The flexible execution class is appropriate for time-insensitive jobs whose start and completion times may vary.

Only jobs with AWS Glue version 3.0 and above and command type `glueetl` will be allowed to set `ExecutionClass` to FLEX. The flexible execution class is available for Spark jobs.

Type: String

Length Constraints: Maximum length of 16.

Valid Values: FLEX | STANDARD

Required: No

ExecutionProperty

An `ExecutionProperty` specifying the maximum number of concurrent runs allowed for this job.

Type: [ExecutionProperty](#) object

Required: No

GlueVersion

In Spark jobs, `GlueVersion` determines the versions of Apache Spark and Python that AWS Glue available in a job. The Python version indicates the version supported for jobs of type Spark.

Ray jobs should set `GlueVersion` to `4.0` or greater. However, the versions of Ray, Python and additional libraries available in your Ray job are determined by the `Runtime` parameter of the Job command.

For more information about the available AWS Glue versions and corresponding Spark and Python versions, see [Glue version](#) in the developer guide.

Jobs that are created without specifying a Glue version default to Glue 0.9.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `^\w+\.\w+$`

Required: No

JobMode

A mode that describes how a job was created. Valid values are:

- SCRIPT - The job was created using the AWS Glue Studio script editor.
- VISUAL - The job was created using the AWS Glue Studio visual editor.
- NOTEBOOK - The job was created using an interactive sessions notebook.

When the JobMode field is missing or null, SCRIPT is assigned as the default value.

Type: String

Valid Values: SCRIPT | VISUAL | NOTEBOOK

Required: No

LogUri

This field is reserved for future use.

Type: String

Required: No

MaintenanceWindow

This field specifies a day of the week and hour for a maintenance window for streaming jobs. AWS Glue periodically performs maintenance activities. During these maintenance windows, AWS Glue will need to restart your streaming jobs.

AWS Glue will restart the job within 3 hours of the specified maintenance window. For instance, if you set up the maintenance window for Monday at 10:00AM GMT, your jobs will be restarted between 10:00AM GMT to 1:00PM GMT.

Type: String

Pattern: `^(Sun|Mon|Tue|Wed|Thu|Fri|Sat):([01]?[0-9]|2[0-3])$`

Required: No

MaxCapacity

For Glue version 1.0 or earlier jobs, using the standard worker type, the number of AWS Glue data processing units (DPUs) that can be allocated when this job runs. A DPU is a relative measure of processing power that consists of 4 vCPUs of compute capacity and 16 GB of memory. For more information, see the [AWS Glue pricing page](#).

For Glue version 2.0+ jobs, you cannot specify a `MaximumCapacity`. Instead, you should specify a `WorkerType` and the `NumberOfWorkers`.

Do not set `MaxCapacity` if using `WorkerType` and `NumberOfWorkers`.

The value that can be allocated for `MaxCapacity` depends on whether you are running a Python shell job, an Apache Spark ETL job, or an Apache Spark streaming ETL job:

- When you specify a Python shell job (`JobCommand.Name="pythonshell"`), you can allocate either 0.0625 or 1 DPU. The default is 0.0625 DPU.
- When you specify an Apache Spark ETL job (`JobCommand.Name="glueetl"`) or Apache Spark streaming ETL job (`JobCommand.Name="gluestreaming"`), you can allocate from 2 to 100 DPUs. The default is 10 DPUs. This job type cannot have a fractional DPU allocation.

Type: Double

Required: No

MaxRetries

The maximum number of times to retry this job if it fails.

Type: Integer

Required: No

Name

The name you assign to this job definition. It must be unique in your account.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF\\t]*`

Required: Yes

NonOverridableArguments

Arguments for this job that are not overridden when providing job arguments in a job run, specified as name-value pairs.

Type: String to string map

Required: No

NotificationProperty

Specifies configuration properties of a job notification.

Type: [NotificationProperty](#) object

Required: No

NumberOfWorkers

The number of workers of a defined `workerType` that are allocated when a job runs.

Type: Integer

Required: No

Role

The name or Amazon Resource Name (ARN) of the IAM role associated with this job.

Type: String

Required: Yes

SecurityConfiguration

The name of the `SecurityConfiguration` structure to be used with this job.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

SourceControlDetails

The details for a source control configuration for a job, allowing synchronization of job artifacts to or from a remote repository.

Type: [SourceControlDetails](#) object

Required: No

Tags

The tags to use with this job. You may use tags to limit access to the job. For more information about tags in AWS Glue, see [AWS Tags in AWS Glue](#) in the developer guide.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Required: No

Timeout

The job timeout in minutes. This is the maximum time that a job run can consume resources before it is terminated and enters TIMEOUT status. The default is 2,880 minutes (48 hours) for batch jobs.

Streaming jobs must have timeout values less than 7 days or 10080 minutes. When the value is left blank, the job will be restarted after 7 days based if you have not setup a maintenance window. If you have setup maintenance window, it will be restarted during the maintenance window after 7 days.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

WorkerType

The type of predefined worker that is allocated when a job runs. Accepts a value of G.1X, G.2X, G.4X, G.8X or G.025X for Spark jobs. Accepts the value Z.2X for Ray jobs.

- For the G.1X worker type, each worker maps to 1 DPU (4 vCPUs, 16 GB of memory) with 84GB disk (approximately 34GB free), and provides 1 executor per worker. We recommend this worker type for workloads such as data transforms, joins, and queries, to offers a scalable and cost effective way to run most jobs.
- For the G.2X worker type, each worker maps to 2 DPU (8 vCPUs, 32 GB of memory) with 128GB disk (approximately 77GB free), and provides 1 executor per worker. We recommend this worker type for workloads such as data transforms, joins, and queries, to offers a scalable and cost effective way to run most jobs.
- For the G.4X worker type, each worker maps to 4 DPU (16 vCPUs, 64 GB of memory) with 256GB disk (approximately 235GB free), and provides 1 executor per worker. We recommend this worker type for jobs whose workloads contain your most demanding transforms, aggregations, joins, and queries. This worker type is available only for AWS Glue version 3.0 or later Spark ETL jobs in the following AWS Regions: US East (Ohio), US East (N. Virginia), US West (Oregon), Asia Pacific (Singapore), Asia Pacific (Sydney), Asia Pacific (Tokyo), Canada (Central), Europe (Frankfurt), Europe (Ireland), and Europe (Stockholm).
- For the G.8X worker type, each worker maps to 8 DPU (32 vCPUs, 128 GB of memory) with 512GB disk (approximately 487GB free), and provides 1 executor per worker. We recommend this worker type for jobs whose workloads contain your most demanding transforms, aggregations, joins, and queries. This worker type is available only for AWS Glue version 3.0 or later Spark ETL jobs, in the same AWS Regions as supported for the G.4X worker type.
- For the G.025X worker type, each worker maps to 0.25 DPU (2 vCPUs, 4 GB of memory) with 84GB disk (approximately 34GB free), and provides 1 executor per worker. We recommend this worker type for low volume streaming jobs. This worker type is only available for AWS Glue version 3.0 streaming jobs.
- For the Z.2X worker type, each worker maps to 2 M-DPU (8vCPUs, 64 GB of memory) with 128 GB disk (approximately 120GB free), and provides up to 8 Ray workers based on the autoscaler.

Type: String

Valid Values: Standard | G.1X | G.2X | G.025X | G.4X | G.8X | Z.2X

Required: No

Response Syntax

```
{  
  "Name": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Name

The unique name that was provided for this job definition.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

IdempotentParameterMismatchException

The same unique identifier was associated with two different records.

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreateMLTransform

Creates an AWS Glue machine learning transform. This operation creates the transform and all the necessary parameters to train it.

Call this operation as the first step in the process of using a machine learning transform (such as the FindMatches transform) for deduplicating data. You can provide an optional Description, in addition to the parameters that you want to use for your algorithm.

You must also specify certain parameters for the tasks that AWS Glue runs on your behalf as part of learning from your data and creating a high-quality machine learning transform. These parameters include Role, and optionally, AllocatedCapacity, Timeout, and MaxRetries. For more information, see [Jobs](#).

Request Syntax

```
{
  "Description": "string",
  "GlueVersion": "string",
  "InputRecordTables": [
    {
      "AdditionalOptions": {
        "string" : "string"
      },
      "CatalogId": "string",
      "ConnectionName": "string",
      "DatabaseName": "string",
      "TableName": "string"
    }
  ],
  "MaxCapacity": number,
  "MaxRetries": number,
  "Name": "string",
  "NumberOfWorkers": number,
  "Parameters": {
    "FindMatchesParameters": {
      "AccuracyCostTradeoff": number,
      "EnforceProvidedLabels": boolean,
      "PrecisionRecallTradeoff": number,
      "PrimaryKeyColumnName": "string"
    }
  },
  "TransformType": "string"
```



```
  },  
  "Role": "string",  
  "Tags": {  
    "string" : "string"  
  },  
  "Timeout": number,  
  "TransformEncryption": {  
    "MLUserDataEncryption": {  
      "KmsKeyId": "string",  
      "MLUserDataEncryptionMode": "string"  
    },  
    "TaskRunSecurityConfigurationName": "string"  
  },  
  "WorkerType": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Description

A description of the machine learning transform that is being defined. The default is an empty string.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF\\r\\n\\t]*`

Required: No

GlueVersion

This value determines which version of AWS Glue this machine learning transform is compatible with. Glue 1.0 is recommended for most customers. If the value is not set, the Glue compatibility defaults to Glue 0.9. For more information, see [AWS Glue Versions](#) in the developer guide.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `^\w+\.\w+$`

Required: No

InputRecordTables

A list of AWS Glue table definitions used by the transform.

Type: Array of [GlueTable](#) objects

Array Members: Minimum number of 0 items. Maximum number of 10 items.

Required: Yes

MaxCapacity

The number of AWS Glue data processing units (DPUs) that are allocated to task runs for this transform. You can allocate from 2 to 100 DPUs; the default is 10. A DPU is a relative measure of processing power that consists of 4 vCPUs of compute capacity and 16 GB of memory. For more information, see the [AWS Glue pricing page](#).

MaxCapacity is a mutually exclusive option with NumberOfWorkers and WorkerType.

- If either NumberOfWorkers or WorkerType is set, then MaxCapacity cannot be set.
- If MaxCapacity is set then neither NumberOfWorkers or WorkerType can be set.
- If WorkerType is set, then NumberOfWorkers is required (and vice versa).
- MaxCapacity and NumberOfWorkers must both be at least 1.

When the WorkerType field is set to a value other than Standard, the MaxCapacity field is set automatically and becomes read-only.

When the WorkerType field is set to a value other than Standard, the MaxCapacity field is set automatically and becomes read-only.

Type: Double

Required: No

MaxRetries

The maximum number of times to retry a task for this transform after a task run fails.

Type: Integer

Required: No

Name

The unique name that you give the transform when you create it.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

NumberOfWorkers

The number of workers of a defined `workerType` that are allocated when this task runs.

If `WorkerType` is set, then `NumberOfWorkers` is required (and vice versa).

Type: Integer

Required: No

Parameters

The algorithmic parameters that are specific to the transform type used. Conditionally dependent on the transform type.

Type: [TransformParameters](#) object

Required: Yes

Role

The name or Amazon Resource Name (ARN) of the IAM role with the required permissions. The required permissions include both AWS Glue service role permissions to AWS Glue resources, and Amazon S3 permissions required by the transform.

- This role needs AWS Glue service role permissions to allow access to resources in AWS Glue. See [Attach a Policy to IAM Users That Access AWS Glue](#).
- This role needs permission to your Amazon Simple Storage Service (Amazon S3) sources, targets, temporary directory, scripts, and any libraries used by the task run for this transform.

Type: String

Required: Yes

Tags

The tags to use with this machine learning transform. You may use tags to limit access to the machine learning transform. For more information about tags in AWS Glue, see [AWS Tags in AWS Glue](#) in the developer guide.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Required: No

Timeout

The timeout of the task run for this transform in minutes. This is the maximum time that a task run for this transform can consume resources before it is terminated and enters TIMEOUT status. The default is 2,880 minutes (48 hours).

Type: Integer

Valid Range: Minimum value of 1.

Required: No

TransformEncryption

The encryption-at-rest settings of the transform that apply to accessing user data. Machine learning transforms can access user data encrypted in Amazon S3 using KMS.

Type: [TransformEncryption](#) object

Required: No

WorkerType

The type of predefined worker that is allocated when this task runs. Accepts a value of Standard, G.1X, or G.2X.

- For the Standard worker type, each worker provides 4 vCPU, 16 GB of memory and a 50GB disk, and 2 executors per worker.
- For the G.1X worker type, each worker provides 4 vCPU, 16 GB of memory and a 64GB disk, and 1 executor per worker.
- For the G.2X worker type, each worker provides 8 vCPU, 32 GB of memory and a 128GB disk, and 1 executor per worker.

MaxCapacity is a mutually exclusive option with NumberOfWorkers and WorkerType.

- If either NumberOfWorkers or WorkerType is set, then MaxCapacity cannot be set.
- If MaxCapacity is set then neither NumberOfWorkers or WorkerType can be set.
- If WorkerType is set, then NumberOfWorkers is required (and vice versa).
- MaxCapacity and NumberOfWorkers must both be at least 1.

Type: String

Valid Values: Standard | G.1X | G.2X | G.025X | G.4X | G.8X | Z.2X

Required: No

Response Syntax

```
{  
  "TransformId": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

TransformId

A unique identifier that is generated for the transform.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

IdempotentParameterMismatchException

The same unique identifier was associated with two different records.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreatePartition

Creates a new partition.

Request Syntax

```
{
  "CatalogId": "string",
  "DatabaseName": "string",
  "PartitionInput": {
    "LastAccessTime": number,
    "LastAnalyzedTime": number,
    "Parameters": {
      "string" : "string"
    },
    "StorageDescriptor": {
      "AdditionalLocations": [ "string" ],
      "BucketColumns": [ "string" ],
      "Columns": [
        {
          "Comment": "string",
          "Name": "string",
          "Parameters": {
            "string" : "string"
          },
          "Type": "string"
        }
      ],
      "Compressed": boolean,
      "InputFormat": "string",
      "Location": "string",
      "NumberOfBuckets": number,
      "OutputFormat": "string",
      "Parameters": {
        "string" : "string"
      },
      "SchemaReference": {
        "SchemaId": {
          "RegistryName": "string",
          "SchemaArn": "string",
          "SchemaName": "string"
        },
        "SchemaVersionId": "string",
```



```

    "SchemaVersionNumber": number
  },
  "SerdeInfo": {
    "Name": "string",
    "Parameters": {
      "string" : "string"
    },
    "SerializationLibrary": "string"
  },
  "SkewedInfo": {
    "SkewedColumnNames": [ "string" ],
    "SkewedColumnValueLocationMaps": {
      "string" : "string"
    },
    "SkewedColumnValues": [ "string" ]
  },
  "SortColumns": [
    {
      "Column": "string",
      "SortOrder": number
    }
  ],
  "StoredAsSubDirectories": boolean
},
"Values": [ "string" ]
},
"TableName": "string"
}

```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The AWS account ID of the catalog in which the partition is to be created.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF\\t]*`

Required: No

DatabaseName

The name of the metadata database in which the partition is to be created.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

PartitionInput

A `PartitionInput` structure defining the partition to be created.

Type: [PartitionInput](#) object

Required: Yes

TableName

The name of the metadata table in which the partition is to be created.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)

- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreatePartitionIndex

Creates a specified partition index in an existing table.

Request Syntax

```
{
  "CatalogId": "string",
  "DatabaseName": "string",
  "PartitionIndex": {
    "IndexName": "string",
    "Keys": [ "string" ]
  },
  "TableName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The catalog ID where the table resides.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

DatabaseName

Specifies the name of a database in which you want to create a partition index.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

PartitionIndex

Specifies a `PartitionIndex` structure to create a partition index in an existing table.

Type: [PartitionIndex](#) object

Required: Yes

TableName

Specifies the name of a table in which you want to create a partition index.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreateRegistry

Creates a new registry which may be used to hold a collection of schemas.

Request Syntax

```
{
  "Description": "string",
  "RegistryName": "string",
  "Tags": {
    "string" : "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Description

A description of the registry. If description is not provided, there will not be any default value for this.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

RegistryName

Name of the registry to be created of max length of 255, and may only contain letters, numbers, hyphen, underscore, dollar sign, or hash mark. No whitespace.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[a-zA-Z0-9-_$#.]+`

Required: Yes

Tags

AWS tags that contain a key value pair and may be searched by console, command line, or API.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Required: No

Response Syntax

```
{
  "Description": "string",
  "RegistryArn": "string",
  "RegistryName": "string",
  "Tags": {
    "string" : "string"
  }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Description

A description of the registry.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF\\r\\n\\t]*`

RegistryArn

The Amazon Resource Name (ARN) of the newly created registry.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 10240.

Pattern: `arn:(aws|aws-us-gov|aws-cn):glue:.*`

RegistryName

The name of the registry.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[a-zA-Z0-9-_$#.]+`

Tags

The tags for the registry.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreateSchema

Creates a new schema set and registers the schema definition. Returns an error if the schema set already exists without actually registering the version.

When the schema set is created, a version checkpoint will be set to the first version. Compatibility mode "DISABLED" restricts any additional schema versions from being added after the first schema version. For all other compatibility modes, validation of compatibility settings will be applied only from the second version onwards when the RegisterSchemaVersion API is used.

When this API is called without a RegistryId, this will create an entry for a "default-registry" in the registry database tables, if it is not already present.

Request Syntax

```
{
  "Compatibility": "string",
  "DataFormat": "string",
  "Description": "string",
  "RegistryId": {
    "RegistryArn": "string",
    "RegistryName": "string"
  },
  "SchemaDefinition": "string",
  "SchemaName": "string",
  "Tags": {
    "string" : "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Compatibility

The compatibility mode of the schema. The possible values are:

- **NONE**: No compatibility mode applies. You can use this choice in development scenarios or if you do not know the compatibility mode that you want to apply to schemas. Any new version added will be accepted without undergoing a compatibility check.
- **DISABLED**: This compatibility choice prevents versioning for a particular schema. You can use this choice to prevent future versioning of a schema.
- **BACKWARD**: This compatibility choice is recommended as it allows data receivers to read both the current and one previous schema version. This means that for instance, a new schema version cannot drop data fields or change the type of these fields, so they can't be read by readers using the previous version.
- **BACKWARD_ALL**: This compatibility choice allows data receivers to read both the current and all previous schema versions. You can use this choice when you need to delete fields or add optional fields, and check compatibility against all previous schema versions.
- **FORWARD**: This compatibility choice allows data receivers to read both the current and one next schema version, but not necessarily later versions. You can use this choice when you need to add fields or delete optional fields, but only check compatibility against the last schema version.
- **FORWARD_ALL**: This compatibility choice allows data receivers to read written by producers of any new registered schema. You can use this choice when you need to add fields or delete optional fields, and check compatibility against all previous schema versions.
- **FULL**: This compatibility choice allows data receivers to read data written by producers using the previous or next version of the schema, but not necessarily earlier or later versions. You can use this choice when you need to add or remove optional fields, but only check compatibility against the last schema version.
- **FULL_ALL**: This compatibility choice allows data receivers to read data written by producers using all previous schema versions. You can use this choice when you need to add or remove optional fields, and check compatibility against all previous schema versions.

Type: String

Valid Values: NONE | DISABLED | BACKWARD | BACKWARD_ALL | FORWARD | FORWARD_ALL | FULL | FULL_ALL

Required: No

DataFormat

The data format of the schema definition. Currently AVRO, JSON and PROTOBUF are supported.

Type: String

Valid Values: AVRO | JSON | PROTOBUF

Required: Yes

Description

An optional description of the schema. If description is not provided, there will not be any automatic default value for this.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

RegistryId

This is a wrapper shape to contain the registry identity fields. If this is not provided, the default registry will be used. The ARN format for the same will be: `arn:aws:glue:us-east-2:<customer id>:registry/default-registry:random-5-letter-id`.

Type: [RegistryId](#) object

Required: No

SchemaDefinition

The schema definition using the DataFormat setting for SchemaName.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 170000.

Pattern: `.*\S.*`

Required: No

SchemaName

Name of the schema to be created of max length of 255, and may only contain letters, numbers, hyphen, underscore, dollar sign, or hash mark. No whitespace.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: [a-zA-Z0-9-_\$#.]+

Required: Yes

Tags

AWS tags that contain a key value pair and may be searched by console, command line, or API. If specified, follows the AWS tags-on-create pattern.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Required: No

Response Syntax

```
{
  "Compatibility": "string",
  "DataFormat": "string",
  "Description": "string",
  "LatestSchemaVersion": number,
  "NextSchemaVersion": number,
  "RegistryArn": "string",
  "RegistryName": "string",
  "SchemaArn": "string",
  "SchemaCheckpoint": number,
  "SchemaName": "string",
  "SchemaStatus": "string",
  "SchemaVersionId": "string",
  "SchemaVersionStatus": "string",
  "Tags": {
    "string" : "string"
  }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Compatibility

The schema compatibility mode.

Type: String

Valid Values: NONE | DISABLED | BACKWARD | BACKWARD_ALL | FORWARD | FORWARD_ALL | FULL | FULL_ALL

DataFormat

The data format of the schema definition. Currently AVRO, JSON and PROTOBUF are supported.

Type: String

Valid Values: AVRO | JSON | PROTOBUF

Description

A description of the schema if specified when created.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

LatestSchemaVersion

The latest version of the schema associated with the returned schema definition.

Type: Long

Valid Range: Minimum value of 1. Maximum value of 100000.

NextSchemaVersion

The next version of the schema associated with the returned schema definition.

Type: Long

Valid Range: Minimum value of 1. Maximum value of 100000.

RegistryArn

The Amazon Resource Name (ARN) of the registry.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 10240.

Pattern: `arn:(aws|aws-us-gov|aws-cn):glue:.*`

RegistryName

The name of the registry.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[a-zA-Z0-9-_$#.]+`

SchemaArn

The Amazon Resource Name (ARN) of the schema.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 10240.

Pattern: `arn:(aws|aws-us-gov|aws-cn):glue:.*`

SchemaCheckpoint

The version number of the checkpoint (the last time the compatibility mode was changed).

Type: Long

Valid Range: Minimum value of 1. Maximum value of 100000.

SchemaName

The name of the schema.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: [a-zA-Z0-9-_\$#.]+

SchemaStatus

The status of the schema.

Type: String

Valid Values: AVAILABLE | PENDING | DELETING

SchemaVersionId

The unique identifier of the first schema version.

Type: String

Length Constraints: Fixed length of 36.

Pattern: [a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}

SchemaVersionStatus

The status of the first schema version created.

Type: String

Valid Values: AVAILABLE | PENDING | FAILURE | DELETING

Tags

The tags for the schema.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreateScript

Transforms a directed acyclic graph (DAG) into code.

Request Syntax

```
{
  "DagEdges": [
    {
      "Source": "string",
      "Target": "string",
      "TargetParameter": "string"
    }
  ],
  "DagNodes": [
    {
      "Args": [
        {
          "Name": "string",
          "Param": boolean,
          "Value": "string"
        }
      ],
      "Id": "string",
      "LineNumber": number,
      "NodeType": "string"
    }
  ],
  "Language": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

DagEdges

A list of the edges in the DAG.

Type: Array of [CodeGenEdge](#) objects

Required: No

DagNodes

A list of the nodes in the DAG.

Type: Array of [CodeGenNode](#) objects

Required: No

Language

The programming language of the resulting code from the DAG.

Type: String

Valid Values: PYTHON | SCALA

Required: No

Response Syntax

```
{
  "PythonScript": "string",
  "ScalaCode": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

PythonScript

The Python script generated from the DAG.

Type: String

ScalaCode

The Scala code generated from the DAG.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreateSecurityConfiguration

Creates a new security configuration. A security configuration is a set of security properties that can be used by AWS Glue. You can use a security configuration to encrypt data at rest. For information about using security configurations in AWS Glue, see [Encrypting Data Written by Crawlers, Jobs, and Development Endpoints](#).

Request Syntax

```
{
  "EncryptionConfiguration": {
    "CloudWatchEncryption": {
      "CloudWatchEncryptionMode": "string",
      "KmsKeyArn": "string"
    },
    "JobBookmarksEncryption": {
      "JobBookmarksEncryptionMode": "string",
      "KmsKeyArn": "string"
    },
    "S3Encryption": [
      {
        "KmsKeyArn": "string",
        "S3EncryptionMode": "string"
      }
    ]
  },
  "Name": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

[EncryptionConfiguration](#)

The encryption configuration for the new security configuration.

Type: [EncryptionConfiguration](#) object

Required: Yes

Name

The name for the new security configuration.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "CreatedTimestamp": number,
  "Name": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

CreatedTimestamp

The time at which the new security configuration was created.

Type: Timestamp

Name

The name assigned to the new security configuration.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)

- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreateSession

Creates a new session.

Request Syntax

```
{
  "Command": {
    "Name": "string",
    "PythonVersion": "string"
  },
  "Connections": {
    "Connections": [ "string" ]
  },
  "DefaultArguments": {
    "string" : "string"
  },
  "Description": "string",
  "GlueVersion": "string",
  "Id": "string",
  "IdleTimeout": number,
  "MaxCapacity": number,
  "NumberOfWorkers": number,
  "RequestOrigin": "string",
  "Role": "string",
  "SecurityConfiguration": "string",
  "Tags": {
    "string" : "string"
  },
  "Timeout": number,
  "WorkerType": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Command

The SessionCommand that runs the job.

Type: [SessionCommand](#) object

Required: Yes

Connections

The number of connections to use for the session.

Type: [ConnectionsList](#) object

Required: No

DefaultArguments

A map array of key-value pairs. Max is 75 pairs.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 75 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Key Pattern: `[\.\- _A-Za-z0-9]+`

Value Length Constraints: Minimum length of 0. Maximum length of 4096.

Value Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Description

The description of the session.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

GlueVersion

The AWS Glue version determines the versions of Apache Spark and Python that AWS Glue supports. The GlueVersion must be greater than 2.0.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `^\w+\.\w+$`

Required: No

Id

The ID of the session request.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

IdleTimeout

The number of minutes when idle before session times out. Default for Spark ETL jobs is value of Timeout. Consult the documentation for other job types.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

MaxCapacity

The number of AWS Glue data processing units (DPUs) that can be allocated when the job runs. A DPU is a relative measure of processing power that consists of 4 vCPUs of compute capacity and 16 GB memory.

Type: Double

Required: No

NumberOfWorkers

The number of workers of a defined `WorkerType` to use for the session.

Type: Integer

Required: No

RequestOrigin

The origin of the request.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[\.\- _A-Za-z0-9]+`

Required: No

Role

The IAM Role ARN

Type: String

Length Constraints: Minimum length of 20. Maximum length of 2048.

Pattern: `arn:aws[^:]*:iam:[0-9]*:role/.+`

Required: Yes

SecurityConfiguration

The name of the SecurityConfiguration structure to be used with the session

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Tags

The map of key value pairs (tags) belonging to the session.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Required: No

Timeout

The number of minutes before session times out. Default for Spark ETL jobs is 48 hours (2880 minutes), the maximum session lifetime for this job type. Consult the documentation for other job types.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

WorkerType

The type of predefined worker that is allocated when a job runs. Accepts a value of G.1X, G.2X, G.4X, or G.8X for Spark jobs. Accepts the value Z.2X for Ray notebooks.

- For the G.1X worker type, each worker maps to 1 DPU (4 vCPUs, 16 GB of memory) with 84GB disk (approximately 34GB free), and provides 1 executor per worker. We recommend this worker type for workloads such as data transforms, joins, and queries, to offers a scalable and cost effective way to run most jobs.
- For the G.2X worker type, each worker maps to 2 DPU (8 vCPUs, 32 GB of memory) with 128GB disk (approximately 77GB free), and provides 1 executor per worker. We recommend this worker type for workloads such as data transforms, joins, and queries, to offers a scalable and cost effective way to run most jobs.
- For the G.4X worker type, each worker maps to 4 DPU (16 vCPUs, 64 GB of memory) with 256GB disk (approximately 235GB free), and provides 1 executor per worker. We recommend this worker type for jobs whose workloads contain your most demanding transforms, aggregations, joins, and queries. This worker type is available only for AWS Glue version 3.0 or later Spark ETL jobs in the following AWS Regions: US East (Ohio), US East (N. Virginia), US West (Oregon), Asia Pacific (Singapore), Asia Pacific (Sydney), Asia Pacific (Tokyo), Canada (Central), Europe (Frankfurt), Europe (Ireland), and Europe (Stockholm).
- For the G.8X worker type, each worker maps to 8 DPU (32 vCPUs, 128 GB of memory) with 512GB disk (approximately 487GB free), and provides 1 executor per worker. We recommend this worker type for jobs whose workloads contain your most demanding transforms, aggregations, joins, and queries. This worker type is available only for AWS Glue version 3.0 or later Spark ETL jobs, in the same AWS Regions as supported for the G.4X worker type.

- For the Z.2X worker type, each worker maps to 2 M-DPU (8vCPUs, 64 GB of memory) with 128 GB disk (approximately 120GB free), and provides up to 8 Ray workers based on the autoscaler.

Type: String

Valid Values: Standard | G.1X | G.2X | G.025X | G.4X | G.8X | Z.2X

Required: No

Response Syntax

```
{
  "Session": {
    "Command": {
      "Name": "string",
      "PythonVersion": "string"
    },
    "CompletedOn": number,
    "Connections": {
      "Connections": [ "string" ]
    },
    "CreatedOn": number,
    "DefaultArguments": {
      "string" : "string"
    },
    "Description": "string",
    "DPUSecods": number,
    "ErrorMessage": "string",
    "ExecutionTime": number,
    "GlueVersion": "string",
    "Id": "string",
    "IdleTimeout": number,
    "MaxCapacity": number,
    "NumberOfWorkers": number,
    "ProfileName": "string",
    "Progress": number,
    "Role": "string",
    "SecurityConfiguration": "string",
    "Status": "string",
    "WorkerType": "string"
  }
}
```

```
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Session

Returns the session object in the response.

Type: [Session](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

IdempotentParameterMismatchException

The same unique identifier was associated with two different records.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

ValidationException

A value could not be validated.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreateTable

Creates a new table definition in the Data Catalog.

Request Syntax

```
{
  "CatalogId": "string",
  "DatabaseName": "string",
  "OpenTableFormatInput": {
    "IcebergInput": {
      "MetadataOperation": "string",
      "Version": "string"
    }
  },
  "PartitionIndexes": [
    {
      "IndexName": "string",
      "Keys": [ "string" ]
    }
  ],
  "TableInput": {
    "Description": "string",
    "LastAccessTime": number,
    "LastAnalyzedTime": number,
    "Name": "string",
    "Owner": "string",
    "Parameters": {
      "string" : "string"
    },
    "PartitionKeys": [
      {
        "Comment": "string",
        "Name": "string",
        "Parameters": {
          "string" : "string"
        },
        "Type": "string"
      }
    ],
    "Retention": number,
    "StorageDescriptor": {
      "AdditionalLocations": [ "string" ],
```

```
"BucketColumns": [ "string" ],
"Columns": [
  {
    "Comment": "string",
    "Name": "string",
    "Parameters": {
      "string" : "string"
    },
    "Type": "string"
  }
],
"Compressed": boolean,
"InputFormat": "string",
"Location": "string",
"NumberOfBuckets": number,
"OutputFormat": "string",
"Parameters": {
  "string" : "string"
},
"SchemaReference": {
  "SchemaId": {
    "RegistryName": "string",
    "SchemaArn": "string",
    "SchemaName": "string"
  },
  "SchemaVersionId": "string",
  "SchemaVersionNumber": number
},
"SerdeInfo": {
  "Name": "string",
  "Parameters": {
    "string" : "string"
  },
  "SerializationLibrary": "string"
},
"SkewedInfo": {
  "SkewedColumnNames": [ "string" ],
  "SkewedColumnValueLocationMaps": {
    "string" : "string"
  },
  "SkewedColumnValues": [ "string" ]
},
"SortColumns": [
  {
```

```

        "Column": "string",
        "SortOrder": number
    }
],
"StoredAsSubDirectories": boolean
},
"TableType": "string",
"TargetTable": {
    "CatalogId": "string",
    "DatabaseName": "string",
    "Name": "string",
    "Region": "string"
},
"ViewDefinition": {
    "Definer": "string",
    "IsProtected": boolean,
    "Representations": [
        {
            "Dialect": "string",
            "DialectVersion": "string",
            "ValidationConnection": "string",
            "ViewExpandedText": "string",
            "ViewOriginalText": "string"
        }
    ],
    "SubObjects": [ "string" ]
},
"ViewExpandedText": "string",
"ViewOriginalText": "string"
},
"TransactionId": "string"
}

```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog in which to create the Table. If none is supplied, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

DatabaseName

The catalog database in which to create the new table. For Hive compatibility, this name is entirely lowercase.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

OpenTableFormatInput

Specifies an `OpenTableFormatInput` structure when creating an open format table.

Type: [OpenTableFormatInput](#) object

Required: No

PartitionIndexes

A list of partition indexes, `PartitionIndex` structures, to create in the table.

Type: Array of [PartitionIndex](#) objects

Array Members: Maximum number of 3 items.

Required: No

TableInput

The `TableInput` object that defines the metadata table to create in the catalog.

Type: [TableInput](#) object

Required: Yes

TransactionId

The ID of the transaction.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\p{L}\p{N}\p{P}]*`

Required: No

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNotReadyException

A resource was not ready for a transaction.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)

- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreateTableOptimizer

Creates a new table optimizer for a specific function. `compaction` is the only currently supported optimizer type.

Request Syntax

```

{
  "CatalogId": "string",
  "DatabaseName": "string",
  "TableName": "string",
  "TableOptimizerConfiguration": {
    "enabled": boolean,
    "roleArn": "string"
  },
  "Type": "string"
}

```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

[CatalogId](#)

The Catalog ID of the table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF\\t]*`

Required: Yes

[DatabaseName](#)

The name of the database in the catalog in which the table resides.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

TableName

The name of the table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

TableOptimizerConfiguration

A `TableOptimizerConfiguration` object representing the configuration of a table optimizer.

Type: [TableOptimizerConfiguration](#) object

Required: Yes

Type

The type of table optimizer. Currently, the only valid value is `compaction`.

Type: String

Valid Values: `compaction`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)

- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreateTrigger

Creates a new trigger.

Request Syntax

```
{
  "Actions": [
    {
      "Arguments": {
        "string" : "string"
      },
      "CrawlerName": "string",
      "JobName": "string",
      "NotificationProperty": {
        "NotifyDelayAfter": number
      },
      "SecurityConfiguration": "string",
      "Timeout": number
    }
  ],
  "Description": "string",
  "EventBatchingCondition": {
    "BatchSize": number,
    "BatchWindow": number
  },
  "Name": "string",
  "Predicate": {
    "Conditions": [
      {
        "CrawlerName": "string",
        "CrawlState": "string",
        "JobName": "string",
        "LogicalOperator": "string",
        "State": "string"
      }
    ],
    "Logical": "string"
  },
  "Schedule": "string",
  "StartOnCreation": boolean,
  "Tags": {
    "string" : "string"
  }
}
```

```
  },  
  "Type": "string",  
  "WorkflowName": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Actions

The actions initiated by this trigger when it fires.

Type: Array of [Action](#) objects

Required: Yes

Description

A description of the new trigger.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

EventBatchingCondition

Batch condition that must be met (specified number of events received or batch time window expired) before EventBridge event trigger fires.

Type: [EventBatchingCondition](#) object

Required: No

Name

The name of the trigger.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Predicate

A predicate to specify when the new trigger should fire.

This field is required when the trigger type is `CONDITIONAL`.

Type: [Predicate](#) object

Required: No

Schedule

A cron expression used to specify the schedule (see [Time-Based Schedules for Jobs and Crawlers](#)). For example, to run something every day at 12:15 UTC, you would specify: `cron(15 12 * * ? *)`.

This field is required when the trigger type is `SCHEDULED`.

Type: String

Required: No

StartOnCreation

Set to `true` to start `SCHEDULED` and `CONDITIONAL` triggers when created. `True` is not supported for `ON_DEMAND` triggers.

Type: Boolean

Required: No

Tags

The tags to use with this trigger. You may use tags to limit access to the trigger. For more information about tags in AWS Glue, see [AWS Tags in AWS Glue](#) in the developer guide.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Required: No

Type

The type of the new trigger.

Type: String

Valid Values: SCHEDULED | CONDITIONAL | ON_DEMAND | EVENT

Required: Yes

WorkflowName

The name of the workflow associated with the trigger.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Response Syntax

```
{
  "Name": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Name

The name of the trigger.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

IdempotentParameterMismatchException

The same unique identifier was associated with two different records.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreateUsageProfile

Creates an AWS Glue usage profile.

Request Syntax

```
{
  "Configuration": {
    "JobConfiguration": {
      "string": {
        "AllowedValues": [ "string" ],
        "DefaultValue": "string",
        "MaxValue": "string",
        "MinValue": "string"
      }
    },
    "SessionConfiguration": {
      "string": {
        "AllowedValues": [ "string" ],
        "DefaultValue": "string",
        "MaxValue": "string",
        "MinValue": "string"
      }
    }
  },
  "Description": "string",
  "Name": "string",
  "Tags": {
    "string": "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Configuration

A ProfileConfiguration object specifying the job and session values for the profile.

Type: [ProfileConfiguration](#) object

Required: Yes

Description

A description of the usage profile.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Name

The name of the usage profile.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Tags

A list of tags applied to the usage profile.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Required: No

Response Syntax

```
{  
  "Name": "string"
```

```
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Name

The name of the usage profile that was created.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationNotSupportedException

The operation is not available in the region.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreateUserDefinedFunction

Creates a new function definition in the Data Catalog.

Request Syntax

```
{
  "CatalogId": "string",
  "DatabaseName": "string",
  "FunctionInput": {
    "ClassName": "string",
    "FunctionName": "string",
    "OwnerName": "string",
    "OwnerType": "string",
    "ResourceUris": [
      {
        "ResourceType": "string",
        "Uri": "string"
      }
    ]
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog in which to create the function. If none is provided, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

DatabaseName

The name of the catalog database in which to create the function.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

FunctionInput

A `FunctionInput` object that defines the function to create in the Data Catalog.

Type: [UserDefinedFunctionInput](#) object

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreateWorkflow

Creates a new workflow.

Request Syntax

```
{
  "DefaultRunProperties": {
    "string" : "string"
  },
  "Description": "string",
  "MaxConcurrentRuns": number,
  "Name": "string",
  "Tags": {
    "string" : "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

DefaultRunProperties

A collection of properties to be used as part of each execution of the workflow.

Type: String to string map

Key Length Constraints: Minimum length of 1. Maximum length of 255.

Key Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Description

A description of the workflow.

Type: String

Required: No

MaxConcurrentRuns

You can use this parameter to prevent unwanted multiple updates to data, to control costs, or in some cases, to prevent exceeding the maximum number of concurrent runs of any of the component jobs. If you leave this parameter blank, there is no limit to the number of concurrent workflow runs.

Type: Integer

Required: No

Name

The name to be assigned to the workflow. It should be unique within your account.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Tags

The tags to be used with this workflow.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Required: No

Response Syntax

```
{
  "Name": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Name

The name of the workflow which was provided as part of the request.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteBlueprint

Deletes an existing blueprint.

Request Syntax

```
{  
  "Name": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Name

The name of the blueprint to delete.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{  
  "Name": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Name

Returns the name of the blueprint that was deleted.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)

- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteClassifier

Removes a classifier from the Data Catalog.

Request Syntax

```
{  
  "Name": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Name

Name of the classifier to remove.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteColumnStatisticsForPartition

Delete the partition column statistics of a column.

The Identity and Access Management (IAM) permission required for this operation is `DeletePartition`.

Request Syntax

```
{
  "CatalogId": "string",
  "ColumnName": "string",
  "DatabaseName": "string",
  "PartitionValues": [ "string" ],
  "TableName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog where the partitions in question reside. If none is supplied, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

ColumnName

Name of the column.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

DatabaseName

The name of the catalog database where the partitions reside.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

PartitionValues

A list of partition values identifying the partition.

Type: Array of strings

Length Constraints: Maximum length of 1024.

Required: Yes

TableName

The name of the partitions' table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)

- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteColumnStatisticsForTable

Retrieves table statistics of columns.

The Identity and Access Management (IAM) permission required for this operation is DeleteTable.

Request Syntax

```
{  
  "CatalogId": "string",  
  "ColumnName": "string",  
  "DatabaseName": "string",  
  "TableName": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog where the partitions in question reside. If none is supplied, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\u007F\u00E0-\u00FF\u0100-\u017F\u0180-\u01FF\u0200-\u02FF\u0300-\u037F\u0380-\u03FF\u0400-\u047F\u0480-\u04FF\u0500-\u057F\u0580-\u05FF\u0600-\u06FF\u0700-\u077F\u0780-\u07FF\u0800-\u087F\u0880-\u08FF\u0900-\u097F\u0980-\u09FF\u0A00-\u0A7F\u0A80-\u0AFF\u0B00-\u0B7F\u0B80-\u0BFF\u0C00-\u0C7F\u0C80-\u0CFF\u0D00-\u0D7F\u0D80-\u0DBF\u0DC0-\u0DBF\u0E00-\u0E7F\u0E80-\u0EFF\u0F00-\u0F7F\u0F80-\u0FFF\u1000-\u107F\u1080-\u10FF\u1100-\u117F\u1180-\u11FF\u1200-\u127F\u1280-\u12FF\u1300-\u137F\u1380-\u13FF\u1400-\u147F\u1480-\u14FF\u1500-\u157F\u1580-\u15FF\u1600-\u167F\u1680-\u16FF\u1700-\u177F\u1780-\u17FF\u1800-\u187F\u1880-\u18FF\u1900-\u197F\u1980-\u19FF\u1A00-\u1A7F\u1A80-\u1AFF\u1B00-\u1B7F\u1B80-\u1BFF\u1C00-\u1C7F\u1C80-\u1CFF\u1D00-\u1D7F\u1D80-\u1DBF\u1E00-\u1E7F\u1E80-\u1EFF\u1F00-\u1F7F\u1F80-\u1FFF\u2000-\u207F\u2080-\u20FF\u2100-\u217F\u2180-\u21FF\u2200-\u227F\u2280-\u22FF\u2300-\u237F\u2380-\u23FF\u2400-\u247F\u2480-\u24FF\u2500-\u257F\u2580-\u25FF\u2600-\u267F\u2680-\u26FF\u2700-\u277F\u2780-\u27FF\u2800-\u287F\u2880-\u28FF\u2900-\u297F\u2980-\u29FF\u2A00-\u2A7F\u2A80-\u2AFF\u2B00-\u2B7F\u2B80-\u2BFF\u2C00-\u2C7F\u2C80-\u2CFF\u2D00-\u2D7F\u2D80-\u2DBF\u2E00-\u2E7F\u2E80-\u2EFF\u2F00-\u2F7F\u2F80-\u2FFF\u3000-\u307F\u3080-\u30FF\u3100-\u317F\u3180-\u31FF\u3200-\u327F\u3280-\u32FF\u3300-\u337F\u3380-\u33FF\u3400-\u347F\u3480-\u34FF\u3500-\u357F\u3580-\u35FF\u3600-\u367F\u3680-\u36FF\u3700-\u377F\u3780-\u37FF\u3800-\u387F\u3880-\u38FF\u3900-\u397F\u3980-\u39FF\u3A00-\u3A7F\u3A80-\u3AFF\u3B00-\u3B7F\u3B80-\u3BFF\u3C00-\u3C7F\u3C80-\u3CFF\u3D00-\u3D7F\u3D80-\u3DBF\u3E00-\u3E7F\u3E80-\u3EFF\u3F00-\u3F7F\u3F80-\u3FFF\u4000-\u407F\u4080-\u40FF\u4100-\u417F\u4180-\u41FF\u4200-\u427F\u4280-\u42FF\u4300-\u437F\u4380-\u43FF\u4400-\u447F\u4480-\u44FF\u4500-\u457F\u4580-\u45FF\u4600-\u467F\u4680-\u46FF\u4700-\u477F\u4780-\u47FF\u4800-\u487F\u4880-\u48FF\u4900-\u497F\u4980-\u49FF\u4A00-\u4A7F\u4A80-\u4AFF\u4B00-\u4B7F\u4B80-\u4BFF\u4C00-\u4C7F\u4C80-\u4CFF\u4D00-\u4D7F\u4D80-\u4DBF\u4E00-\u4E7F\u4E80-\u4EFF\u4F00-\u4F7F\u4F80-\u4FFF\u5000-\u507F\u5080-\u50FF\u5100-\u517F\u5180-\u51FF\u5200-\u527F\u5280-\u52FF\u5300-\u537F\u5380-\u53FF\u5400-\u547F\u5480-\u54FF\u5500-\u557F\u5580-\u55FF\u5600-\u567F\u5680-\u56FF\u5700-\u577F\u5780-\u57FF\u5800-\u587F\u5880-\u58FF\u5900-\u597F\u5980-\u59FF\u5A00-\u5A7F\u5A80-\u5AFF\u5B00-\u5B7F\u5B80-\u5BFF\u5C00-\u5C7F\u5C80-\u5CFF\u5D00-\u5D7F\u5D80-\u5DBF\u5E00-\u5E7F\u5E80-\u5EFF\u5F00-\u5F7F\u5F80-\u5FFF\u6000-\u607F\u6080-\u60FF\u6100-\u617F\u6180-\u61FF\u6200-\u627F\u6280-\u62FF\u6300-\u637F\u6380-\u63FF\u6400-\u647F\u6480-\u64FF\u6500-\u657F\u6580-\u65FF\u6600-\u667F\u6680-\u66FF\u6700-\u677F\u6780-\u67FF\u6800-\u687F\u6880-\u68FF\u6900-\u697F\u6980-\u69FF\u6A00-\u6A7F\u6A80-\u6AFF\u6B00-\u6B7F\u6B80-\u6BFF\u6C00-\u6C7F\u6C80-\u6CFF\u6D00-\u6D7F\u6D80-\u6DBF\u6E00-\u6E7F\u6E80-\u6EFF\u6F00-\u6F7F\u6F80-\u6FFF\u7000-\u707F\u7080-\u70FF\u7100-\u717F\u7180-\u71FF\u7200-\u727F\u7280-\u72FF\u7300-\u737F\u7380-\u73FF\u7400-\u747F\u7480-\u74FF\u7500-\u757F\u7580-\u75FF\u7600-\u767F\u7680-\u76FF\u7700-\u777F\u7780-\u77FF\u7800-\u787F\u7880-\u78FF\u7900-\u797F\u7980-\u79FF\u7A00-\u7A7F\u7A80-\u7AFF\u7B00-\u7B7F\u7B80-\u7BFF\u7C00-\u7C7F\u7C80-\u7CFF\u7D00-\u7D7F\u7D80-\u7DBF\u7E00-\u7E7F\u7E80-\u7EFF\u7F00-\u7F7F\u7F80-\u7FFF\u8000-\u807F\u8080-\u80FF\u8100-\u817F\u8180-\u81FF\u8200-\u827F\u8280-\u82FF\u8300-\u837F\u8380-\u83FF\u8400-\u847F\u8480-\u84FF\u8500-\u857F\u8580-\u85FF\u8600-\u867F\u8680-\u86FF\u8700-\u877F\u8780-\u87FF\u8800-\u887F\u8880-\u88FF\u8900-\u897F\u8980-\u89FF\u8A00-\u8A7F\u8A80-\u8AFF\u8B00-\u8B7F\u8B80-\u8BFF\u8C00-\u8C7F\u8C80-\u8CFF\u8D00-\u8D7F\u8D80-\u8DBF\u8E00-\u8E7F\u8E80-\u8EFF\u8F00-\u8F7F\u8F80-\u8FFF\u9000-\u907F\u9080-\u90FF\u9100-\u917F\u9180-\u91FF\u9200-\u927F\u9280-\u92FF\u9300-\u937F\u9380-\u93FF\u9400-\u947F\u9480-\u94FF\u9500-\u957F\u9580-\u95FF\u9600-\u967F\u9680-\u96FF\u9700-\u977F\u9780-\u97FF\u9800-\u987F\u9880-\u98FF\u9900-\u997F\u9980-\u99FF\u9A00-\u9A7F\u9A80-\u9AFF\u9B00-\u9B7F\u9B80-\u9BFF\u9C00-\u9C7F\u9C80-\u9CFF\u9D00-\u9D7F\u9D80-\u9DBF\u9E00-\u9E7F\u9E80-\u9EFF\u9F00-\u9F7F\u9F80-\u9FFF\uA000-\uA07F\uA080-\uA0FF\uA100-\uA17F\uA180-\uA1FF\uA200-\uA27F\uA280-\uA2FF\uA300-\uA37F\uA380-\uA3FF\uA400-\uA47F\uA480-\uA4FF\uA500-\uA57F\uA580-\uA5FF\uA600-\uA67F\uA680-\uA6FF\uA700-\uA77F\uA780-\uA7FF\uA800-\uA87F\uA880-\uA8FF\uA900-\uA97F\uA980-\uA9FF\uAA00-\uAA7F\uAA80-\uAAFF\uAB00-\uAB7F\uAB80-\uABFF\uAC00-\uAC7F\uAC80-\uACFF\uAD00-\uAD7F\uAD80-\uADBFF\uAE00-\uAE7F\uAE80-\uAEFF\uAF00-\uAF7F\uAF80-\uAFFF\uB000-\uB07F\uB080-\uB0FF\uB100-\uB17F\uB180-\uB1FF\uB200-\uB27F\uB280-\uB2FF\uB300-\uB37F\uB380-\uB3FF\uB400-\uB47F\uB480-\uB4FF\uB500-\uB57F\uB580-\uB5FF\uB600-\uB67F\uB680-\uB6FF\uB700-\uB77F\uB780-\uB7FF\uB800-\uB87F\uB880-\uB8FF\uB900-\uB97F\uB980-\uB9FF\uBA00-\uBA7F\uBA80-\uBAFF\uBB00-\uBB7F\uBB80-\uBBFF\uBC00-\uBC7F\uBC80-\uBCFF\uBD00-\uBD7F\uBD80-\uBDBFF\uBE00-\uBE7F\uBE80-\uBEFF\uBF00-\uBF7F\uBF80-\uBFFF\uC000-\uC07F\uC080-\uC0FF\uC100-\uC17F\uC180-\uC1FF\uC200-\uC27F\uC280-\uC2FF\uC300-\uC37F\uC380-\uC3FF\uC400-\uC47F\uC480-\uC4FF\uC500-\uC57F\uC580-\uC5FF\uC600-\uC67F\uC680-\uC6FF\uC700-\uC77F\uC780-\uC7FF\uC800-\uC87F\uC880-\uC8FF\uC900-\uC97F\uC980-\uC9FF\uCA00-\uCA7F\uCA80-\uCAFF\uCB00-\uCB7F\uCB80-\uCBFF\uCC00-\uCC7F\uCC80-\uCCFF\uCD00-\uCD7F\uCD80-\uCDFF\uCE00-\uCE7F\uCE80-\uCEFF\uCF00-\uCF7F\uCF80-\uCFFF\uD000-\uD07F\uD080-\uD0FF\uD100-\uD17F\uD180-\uD1FF\uD200-\uD27F\uD280-\uD2FF\uD300-\uD37F\uD380-\uD3FF\uD400-\uD47F\uD480-\uD4FF\uD500-\uD57F\uD580-\uD5FF\uD600-\uD67F\uD680-\uD6FF\uD700-\uD77F\uD780-\uD7FF\uD800-\uD87F\uD880-\uD8FF\uD900-\uD97F\uD980-\uD9FF\uDA00-\uDA7F\uDA80-\uDAFF\uDB00-\uDB7F\uDB80-\uDBFF\uDC00-\uDC7F\uDC80-\uDCFF\uDD00-\uDD7F\uDD80-\uDDFF\uDE00-\uDE7F\uDE80-\uDEFF\uDF00-\uDF7F\uDF80-\uDFFF\uE000-\uE07F\uE080-\uE0FF\uE100-\uE17F\uE180-\uE1FF\uE200-\uE27F\uE280-\uE2FF\uE300-\uE37F\uE380-\uE3FF\uE400-\uE47F\uE480-\uE4FF\uE500-\uE57F\uE580-\uE5FF\uE600-\uE67F\uE680-\uE6FF\uE700-\uE77F\uE780-\uE7FF\uE800-\uE87F\uE880-\uE8FF\uE900-\uE97F\uE980-\uE9FF\uEA00-\uEA7F\uEA80-\uEAFF\uEB00-\uEB7F\uEB80-\uEBFF\uEC00-\uEC7F\uEC80-\uECFF\uED00-\uED7F\uED80-\uEDFF\uEE00-\uEE7F\uEE80-\uEEFF\uEF00-\uEF7F\uEF80-\uEFFF\uF000-\uF07F\uF080-\uF0FF\uF100-\uF17F\uF180-\uF1FF\uF200-\uF27F\uF280-\uF2FF\uF300-\uF37F\uF380-\uF3FF\uF400-\uF47F\uF480-\uF4FF\uF500-\uF57F\uF580-\uF5FF\uF600-\uF67F\uF680-\uF6FF\uF700-\uF77F\uF780-\uF7FF\uF800-\uF87F\uF880-\uF8FF\uF900-\uF97F\uF980-\uF9FF\uFA00-\uFA7F\uFA80-\uFAFF\uFB00-\uFB7F\uFB80-\uFBFF\uFC00-\uFC7F\uFC80-\uFCFF\uFD00-\uFD7F\uFD80-\uFDFF\uFE00-\uFE7F\uFE80-\uFEFF\uFF00-\uFF7F\uFF80-\uFFFF`

Required: No

ColumnName

The name of the column.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

DatabaseName

The name of the catalog database where the partitions reside.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

TableName

The name of the partitions' table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteConnection

Deletes a connection from the Data Catalog.

Request Syntax

```
{  
  "CatalogId": "string",  
  "ConnectionName": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog in which the connection resides. If none is provided, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

ConnectionName

The name of the connection to delete.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteCrawler

Removes a specified crawler from the AWS Glue Data Catalog, unless the crawler state is RUNNING.

Request Syntax

```
{
  "Name": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Name

The name of the crawler to remove.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

CrawlerRunningException

The operation cannot be performed because the crawler is already running.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

SchedulerTransitioningException

The specified scheduler is transitioning.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteCustomEntityType

Deletes a custom pattern by specifying its name.

Request Syntax

```
{  
  "Name": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Name

The name of the custom pattern that you want to delete.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{  
  "Name": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Name

The name of the custom pattern you deleted.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteDatabase

Removes a specified database from a Data Catalog.

Note

After completing this operation, you no longer have access to the tables (and all table versions and partitions that might belong to the tables) and the user-defined functions in the deleted database. AWS Glue deletes these "orphaned" resources asynchronously in a timely manner, at the discretion of the service.

To ensure the immediate deletion of all related resources, before calling `DeleteDatabase`, use `DeleteTableVersion` or `BatchDeleteTableVersion`, `DeletePartition` or `BatchDeletePartition`, `DeleteUserDefinedFunction`, and `DeleteTable` or `BatchDeleteTable`, to delete any resources that belong to the database.

Request Syntax

```
{
  "CatalogId": "string",
  "Name": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog in which the database resides. If none is provided, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF\\t]*`

Required: No

Name

The name of the database to delete. For Hive compatibility, this must be all lowercase.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteDataQualityRuleset

Deletes a data quality ruleset.

Request Syntax

```
{
  "Name": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Name

A name for the data quality ruleset.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteDevEndpoint

Deletes a specified development endpoint.

Request Syntax

```
{  
  "EndpointName": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

EndpointName

The name of the DevEndpoint.

Type: String

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteJob

Deletes a specified job definition. If the job definition is not found, no exception is thrown.

Request Syntax

```
{
  "JobName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

JobName

The name of the job definition to delete.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "JobName": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

JobName

The name of the job definition that was deleted.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)

- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteMLTransform

Deletes an AWS Glue machine learning transform. Machine learning transforms are a special type of transform that use machine learning to learn the details of the transformation to be performed by learning from examples provided by humans. These transformations are then saved by AWS Glue. If you no longer need a transform, you can delete it by calling `DeleteMLTransforms`. However, any AWS Glue jobs that still reference the deleted transform will no longer succeed.

Request Syntax

```
{
  "TransformId": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

TransformId

The unique identifier of the transform to delete.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "TransformId": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

TransformId

The unique identifier of the transform that was deleted.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeletePartition

Deletes a specified partition.

Request Syntax

```
{
  "CatalogId": "string",
  "DatabaseName": "string",
  "PartitionValues": [ "string" ],
  "TableName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog where the partition to be deleted resides. If none is provided, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

DatabaseName

The name of the catalog database in which the table in question resides.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

PartitionValues

The values that define the partition.

Type: Array of strings

Length Constraints: Maximum length of 1024.

Required: Yes

TableName

The name of the table that contains the partition to be deleted.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeletePartitionIndex

Deletes a specified partition index from an existing table.

Request Syntax

```
{
  "CatalogId": "string",
  "DatabaseName": "string",
  "IndexName": "string",
  "TableName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The catalog ID where the table resides.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

DatabaseName

Specifies the name of a database from which you want to delete a partition index.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

IndexName

The name of the partition index to be deleted.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

TableName

Specifies the name of a table from which you want to delete a partition index.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ConflictException

The `CreatePartitions` API was called on a table that has indexes enabled.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteRegistry

Delete the entire registry including schema and all of its versions. To get the status of the delete operation, you can call the GetRegistry API after the asynchronous call. Deleting a registry will deactivate all online operations for the registry such as the UpdateRegistry, CreateSchema, UpdateSchema, and RegisterSchemaVersion APIs.

Request Syntax

```
{
  "RegistryId": {
    "RegistryArn": "string",
    "RegistryName": "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

[RegistryId](#)

This is a wrapper structure that may contain the registry name and Amazon Resource Name (ARN).

Type: [RegistryId](#) object

Required: Yes

Response Syntax

```
{
  "RegistryArn": "string",
  "RegistryName": "string",
  "Status": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

RegistryArn

The Amazon Resource Name (ARN) of the registry being deleted.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 10240.

Pattern: `arn:(aws|aws-us-gov|aws-cn):glue:.*`

RegistryName

The name of the registry being deleted.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[a-zA-Z0-9-_$#.]+`

Status

The status of the registry. A successful operation will return the `Deleting` status.

Type: String

Valid Values: `AVAILABLE | DELETING`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteResourcePolicy

Deletes a specified policy.

Request Syntax

```
{
  "PolicyHashCondition": "string",
  "ResourceArn": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

PolicyHashCondition

The hash value returned when this policy was set.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

ResourceArn

The ARN of the AWS Glue resource for the resource policy to be deleted.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 10240.

Pattern: `arn:(aws|aws-us-gov|aws-cn):glue:.*`

Required: No

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ConditionCheckFailureException

A specified condition was not satisfied.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)

- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteSchema

Deletes the entire schema set, including the schema set and all of its versions. To get the status of the delete operation, you can call GetSchema API after the asynchronous call. Deleting a registry will deactivate all online operations for the schema, such as the GetSchemaByDefinition, and RegisterSchemaVersion APIs.

Request Syntax

```
{
  "SchemaId": {
    "RegistryName": "string",
    "SchemaArn": "string",
    "SchemaName": "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

SchemaId

This is a wrapper structure that may contain the schema name and Amazon Resource Name (ARN).

Type: [SchemaId](#) object

Required: Yes

Response Syntax

```
{
  "SchemaArn": "string",
  "SchemaName": "string",
  "Status": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

SchemaArn

The Amazon Resource Name (ARN) of the schema being deleted.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 10240.

Pattern: `arn:(aws|aws-us-gov|aws-cn):glue:.*`

SchemaName

The name of the schema being deleted.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[a-zA-Z0-9-_$#.]+`

Status

The status of the schema.

Type: String

Valid Values: AVAILABLE | PENDING | DELETING

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteSchemaVersions

Remove versions from the specified schema. A version number or range may be supplied. If the compatibility mode forbids deleting of a version that is necessary, such as `BACKWARDS_FULL`, an error is returned. Calling the `GetSchemaVersions` API after this call will list the status of the deleted versions.

When the range of version numbers contain check pointed version, the API will return a 409 conflict and will not proceed with the deletion. You have to remove the checkpoint first using the `DeleteSchemaCheckpoint` API before using this API.

You cannot use the `DeleteSchemaVersions` API to delete the first schema version in the schema set. The first schema version can only be deleted by the `DeleteSchema` API. This operation will also delete the attached `SchemaVersionMetadata` under the schema versions. Hard deletes will be enforced on the database.

If the compatibility mode forbids deleting of a version that is necessary, such as `BACKWARDS_FULL`, an error is returned.

Request Syntax

```
{
  "SchemaId": {
    "RegistryName": "string",
    "SchemaArn": "string",
    "SchemaName": "string"
  },
  "Versions": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

SchemaId

This is a wrapper structure that may contain the schema name and Amazon Resource Name (ARN).

Type: [Schemald](#) object

Required: Yes

[Versions](#)

A version range may be supplied which may be of the format:

- a single version number, 5
- a range, 5-8 : deletes versions 5, 6, 7, 8

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100000.

Pattern: `[1-9][0-9]* | [1-9][0-9]* - [1-9][0-9]*`

Required: Yes

Response Syntax

```
{
  "SchemaVersionErrors": [
    {
      "ErrorDetails": {
        "ErrorCode": "string",
        "ErrorMessage": "string"
      },
      "VersionNumber": number
    }
  ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

[SchemaVersionErrors](#)

A list of `SchemaVersionErrorItem` objects, each containing an error and schema version.

Type: Array of [SchemaVersionErrorItem](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)

- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteSecurityConfiguration

Deletes a specified security configuration.

Request Syntax

```
{  
  "Name": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Name

The name of the security configuration to delete.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteSession

Deletes the session.

Request Syntax

```
{  
  "Id": "string",  
  "RequestOrigin": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Id

The ID of the session to be deleted.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF\\t]*`

Required: Yes

RequestOrigin

The name of the origin of the delete session request.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[\\._-A-Za-z0-9]+`

Required: No

Response Syntax

```
{  
  "Id": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Id

Returns the ID of the deleted session.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

IllegalSessionStateException

The session is in an invalid state to perform a requested operation.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteTable

Removes a table definition from the Data Catalog.

Note

After completing this operation, you no longer have access to the table versions and partitions that belong to the deleted table. AWS Glue deletes these "orphaned" resources asynchronously in a timely manner, at the discretion of the service.

To ensure the immediate deletion of all related resources, before calling `DeleteTable`, use `DeleteTableVersion` or `BatchDeleteTableVersion`, and `DeletePartition` or `BatchDeletePartition`, to delete any resources that belong to the table.

Request Syntax

```
{  
  "CatalogId": "string",  
  "DatabaseName": "string",  
  "Name": "string",  
  "TransactionId": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog where the table resides. If none is provided, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF\\t]*`

Required: No

DatabaseName

The name of the catalog database in which the table resides. For Hive compatibility, this name is entirely lowercase.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Name

The name of the table to be deleted. For Hive compatibility, this name is entirely lowercase.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

TransactionId

The transaction ID at which to delete the table contents.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\p{L}\p{N}\p{P}]*`

Required: No

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNotReadyException

A resource was not ready for a transaction.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)

- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteTableOptimizer

Deletes an optimizer and all associated metadata for a table. The optimization will no longer be performed on the table.

Request Syntax

```
{
  "CatalogId": "string",
  "DatabaseName": "string",
  "TableName": "string",
  "Type": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The Catalog ID of the table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

DatabaseName

The name of the database in the catalog in which the table resides.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

TableName

The name of the table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Type

The type of table optimizer.

Type: String

Valid Values: `compaction`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteTableVersion

Deletes a specified version of a table.

Request Syntax

```
{  
  "CatalogId": "string",  
  "DatabaseName": "string",  
  "TableName": "string",  
  "VersionId": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog where the tables reside. If none is provided, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF\\t]*`

Required: No

DatabaseName

The database in the catalog in which the table resides. For Hive compatibility, this name is entirely lowercase.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF\\t]*`

Required: Yes

TableName

The name of the table. For Hive compatibility, this name is entirely lowercase.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

VersionId

The ID of the table version to be deleted. A `VersionID` is a string representation of an integer. Each version is incremented by 1.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteTrigger

Deletes a specified trigger. If the trigger is not found, no exception is thrown.

Request Syntax

```
{  
  "Name": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Name

The name of the trigger to delete.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{  
  "Name": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Name

The name of the trigger that was deleted.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)

- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteUsageProfile

Deletes the AWS Glue specified usage profile.

Request Syntax

```
{
  "Name": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Name

The name of the usage profile to delete.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationNotSupportedException

The operation is not available in the region.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteUserDefinedFunction

Deletes an existing function definition from the Data Catalog.

Request Syntax

```
{
  "CatalogId": "string",
  "DatabaseName": "string",
  "FunctionName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog where the function to be deleted is located. If none is supplied, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

DatabaseName

The name of the catalog database where the function is located.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

FunctionName

The name of the function definition to be deleted.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServerError

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteWorkflow

Deletes a workflow.

Request Syntax

```
{  
  "Name": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Name

Name of the workflow to be deleted.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{  
  "Name": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Name

Name of the workflow specified in input.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)

- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetBlueprint

Retrieves the details of a blueprint.

Request Syntax

```
{  
  "IncludeBlueprint": boolean,  
  "IncludeParameterSpec": boolean,  
  "Name": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

[IncludeBlueprint](#)

Specifies whether or not to include the blueprint in the response.

Type: Boolean

Required: No

[IncludeParameterSpec](#)

Specifies whether or not to include the parameter specification.

Type: Boolean

Required: No

[Name](#)

The name of the blueprint.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "Blueprint": {
    "BlueprintLocation": "string",
    "BlueprintServiceLocation": "string",
    "CreatedOn": number,
    "Description": "string",
    "ErrorMessage": "string",
    "LastActiveDefinition": {
      "BlueprintLocation": "string",
      "BlueprintServiceLocation": "string",
      "Description": "string",
      "LastModifiedOn": number,
      "ParameterSpec": "string"
    },
    "LastModifiedOn": number,
    "Name": "string",
    "ParameterSpec": "string",
    "Status": "string"
  }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Blueprint

Returns a Blueprint object.

Type: [Blueprint](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetBlueprintRun

Retrieves the details of a blueprint run.

Request Syntax

```
{
  "BlueprintName": "string",
  "RunId": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

BlueprintName

The name of the blueprint.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[\.\-_\A-Za-z0-9]+`

Required: Yes

RunId

The run ID for the blueprint run you want to retrieve.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "BlueprintRun": {
    "BlueprintName": "string",
    "CompletedOn": number,
    "ErrorMessage": "string",
    "Parameters": "string",
    "RoleArn": "string",
    "RollbackErrorMessage": "string",
    "RunId": "string",
    "StartedOn": number,
    "State": "string",
    "WorkflowName": "string"
  }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

BlueprintRun

Returns a `BlueprintRun` object.

Type: [BlueprintRun](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetBlueprintRuns

Retrieves the details of blueprint runs for a specified blueprint.

Request Syntax

```
{  
  "BlueprintName": string,  
  "MaxResults": number,  
  "NextToken": string  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

[BlueprintName](#)

The name of the blueprint.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

[MaxResults](#)

The maximum size of a list to return.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

[NextToken](#)

A continuation token, if this is a continuation request.

Type: String

Required: No

Response Syntax

```
{
  "BlueprintRuns": [
    {
      "BlueprintName": "string",
      "CompletedOn": number,
      "ErrorMessage": "string",
      "Parameters": "string",
      "RoleArn": "string",
      "RollbackErrorMessage": "string",
      "RunId": "string",
      "StartedOn": number,
      "State": "string",
      "WorkflowName": "string"
    }
  ],
  "NextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

BlueprintRuns

Returns a list of `BlueprintRun` objects.

Type: Array of [BlueprintRun](#) objects

NextToken

A continuation token, if not all blueprint runs have been returned.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)

- [AWS SDK for Ruby V3](#)

GetCatalogImportStatus

Retrieves the status of a migration operation.

Request Syntax

```
{  
  "CatalogId": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the catalog to migrate. Currently, this should be the AWS account ID.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Response Syntax

```
{  
  "ImportStatus": {  
    "ImportCompleted": boolean,  
    "ImportedBy": "string",  
    "ImportTime": number  
  }  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

ImportStatus

The status of the specified catalog migration.

Type: [CatalogImportStatus](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetClassifier

Retrieve a classifier by name.

Request Syntax

```
{
  "Name": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Name

Name of the classifier to retrieve.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "Classifier": {
    "CsvClassifier": {
      "AllowSingleColumn": boolean,
      "ContainsHeader": "string",
      "CreationTime": number,
      "CustomDatatypeConfigured": boolean,
      "CustomDatatypes": [ "string" ],
      "Delimiter": "string",
      "DisableValueTrimming": boolean,

```

```
    "Header": [ "string" ],
    "LastUpdated": number,
    "Name": "string",
    "QuoteSymbol": "string",
    "Serde": "string",
    "Version": number
  },
  "GrokClassifier": {
    "Classification": "string",
    "CreationTime": number,
    "CustomPatterns": "string",
    "GrokPattern": "string",
    "LastUpdated": number,
    "Name": "string",
    "Version": number
  },
  "JsonClassifier": {
    "CreationTime": number,
    "JsonPath": "string",
    "LastUpdated": number,
    "Name": "string",
    "Version": number
  },
  "XMLClassifier": {
    "Classification": "string",
    "CreationTime": number,
    "LastUpdated": number,
    "Name": "string",
    "RowTag": "string",
    "Version": number
  }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Classifier

The requested classifier.

Type: [Classifier](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetClassifiers

Lists all classifier objects in the Data Catalog.

Request Syntax

```
{
  "MaxResults": number,
  "NextToken": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

MaxResults

The size of the list to return (optional).

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

NextToken

An optional continuation token.

Type: String

Required: No

Response Syntax

```
{
  "Classifiers": [
    {
      "CsvClassifier": {
        "AllowSingleColumn": boolean,
        "ContainsHeader": "string",
```

```
    "CreationTime": number,
    "CustomDatatypeConfigured": boolean,
    "CustomDatatypes": [ "string" ],
    "Delimiter": "string",
    "DisableValueTrimming": boolean,
    "Header": [ "string" ],
    "LastUpdated": number,
    "Name": "string",
    "QuoteSymbol": "string",
    "Serde": "string",
    "Version": number
  },
  "GrokClassifier": {
    "Classification": "string",
    "CreationTime": number,
    "CustomPatterns": "string",
    "GrokPattern": "string",
    "LastUpdated": number,
    "Name": "string",
    "Version": number
  },
  "JsonClassifier": {
    "CreationTime": number,
    "JsonPath": "string",
    "LastUpdated": number,
    "Name": "string",
    "Version": number
  },
  "XMLClassifier": {
    "Classification": "string",
    "CreationTime": number,
    "LastUpdated": number,
    "Name": "string",
    "RowTag": "string",
    "Version": number
  }
}
],
"NextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Classifiers

The requested list of classifier objects.

Type: Array of [Classifier](#) objects

NextToken

A continuation token.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)

- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetColumnStatisticsForPartition

Retrieves partition statistics of columns.

The Identity and Access Management (IAM) permission required for this operation is `GetPartition`.

Request Syntax

```
{
  "CatalogId": "string",
  "ColumnNames": [ "string" ],
  "DatabaseName": "string",
  "PartitionValues": [ "string" ],
  "TableName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog where the partitions in question reside. If none is supplied, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

ColumnNames

A list of the column names.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 100 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

DatabaseName

The name of the catalog database where the partitions reside.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

PartitionValues

A list of partition values identifying the partition.

Type: Array of strings

Length Constraints: Maximum length of 1024.

Required: Yes

TableName

The name of the partitions' table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "ColumnStatisticsList": [
    {
      "AnalyzedTime": number,
      "ColumnName": "string",
```

```
"ColumnType": "string",
"StatisticsData": {
  "BinaryColumnStatisticsData": {
    "AverageLength": number,
    "MaximumLength": number,
    "NumberOfNulls": number
  },
  "BooleanColumnStatisticsData": {
    "NumberOfFalses": number,
    "NumberOfNulls": number,
    "NumberOfTrues": number
  },
  "DateColumnStatisticsData": {
    "MaximumValue": number,
    "MinimumValue": number,
    "NumberOfDistinctValues": number,
    "NumberOfNulls": number
  },
  "DecimalColumnStatisticsData": {
    "MaximumValue": {
      "Scale": number,
      "UnscaledValue": blob
    },
    "MinimumValue": {
      "Scale": number,
      "UnscaledValue": blob
    },
    "NumberOfDistinctValues": number,
    "NumberOfNulls": number
  },
  "DoubleColumnStatisticsData": {
    "MaximumValue": number,
    "MinimumValue": number,
    "NumberOfDistinctValues": number,
    "NumberOfNulls": number
  },
  "LongColumnStatisticsData": {
    "MaximumValue": number,
    "MinimumValue": number,
    "NumberOfDistinctValues": number,
    "NumberOfNulls": number
  },
  "StringColumnStatisticsData": {
    "AverageLength": number,
```

```
        "MaxLength": number,
        "NumberOfDistinctValues": number,
        "NumberOfNulls": number
    },
    "Type": "string"
}
],
"Errors": [
{
    "ColumnName": "string",
    "Error": {
        "ErrorCode": "string",
        "ErrorMessage": "string"
    }
}
]
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

ColumnStatisticsList

List of ColumnStatistics that failed to be retrieved.

Type: Array of [ColumnStatistics](#) objects

Errors

Error occurred during retrieving column statistics data.

Type: Array of [ColumnError](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetColumnStatisticsForTable

Retrieves table statistics of columns.

The Identity and Access Management (IAM) permission required for this operation is `GetTable`.

Request Syntax

```
{
  "CatalogId": "string",
  "ColumnNames": [ "string" ],
  "DatabaseName": "string",
  "TableName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog where the partitions in question reside. If none is supplied, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

ColumnNames

A list of the column names.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 100 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

DatabaseName

The name of the catalog database where the partitions reside.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

TableName

The name of the partitions' table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "ColumnStatisticsList": [
    {
      "AnalyzedTime": number,
      "ColumnName": "string",
      "ColumnType": "string",
      "StatisticsData": {
        "BinaryColumnStatisticsData": {
          "AverageLength": number,
          "MaximumLength": number,
          "NumberOfNulls": number
        },
        "BooleanColumnStatisticsData": {
          "NumberOfFalses": number,
          "NumberOfNulls": number,
          "NumberOfTrues": number
        }
      }
    }
  ]
}
```

```

    "DateColumnStatisticsData": {
      "MaximumValue": number,
      "MinimumValue": number,
      "NumberOfDistinctValues": number,
      "NumberOfNulls": number
    },
    "DecimalColumnStatisticsData": {
      "MaximumValue": {
        "Scale": number,
        "UnscaledValue": blob
      },
      "MinimumValue": {
        "Scale": number,
        "UnscaledValue": blob
      },
      "NumberOfDistinctValues": number,
      "NumberOfNulls": number
    },
    "DoubleColumnStatisticsData": {
      "MaximumValue": number,
      "MinimumValue": number,
      "NumberOfDistinctValues": number,
      "NumberOfNulls": number
    },
    "LongColumnStatisticsData": {
      "MaximumValue": number,
      "MinimumValue": number,
      "NumberOfDistinctValues": number,
      "NumberOfNulls": number
    },
    "StringColumnStatisticsData": {
      "AverageLength": number,
      "MaximumLength": number,
      "NumberOfDistinctValues": number,
      "NumberOfNulls": number
    },
    "Type": "string"
  }
}
],
"Errors": [
  {
    "ColumnName": "string",
    "Error": {

```



```
        "ErrorCode": "string",
        "ErrorMessage": "string"
    }
}
]
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

ColumnStatisticsList

List of ColumnStatistics.

Type: Array of [ColumnStatistics](#) objects

Errors

List of ColumnStatistics that failed to be retrieved.

Type: Array of [ColumnError](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)


```
"LastUpdated": number,  
"NumberOfWorkers": number,  
"Role": "string",  
"SampleSize": number,  
"SecurityConfiguration": "string",  
"StartTime": number,  
"Status": "string",  
"TableName": "string",  
"WorkerType": "string"  
}  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

ColumnStatisticsTaskRun

A ColumnStatisticsTaskRun object representing the details of the column stats run.

Type: [ColumnStatisticsTaskRun](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetColumnStatisticsTaskRuns

Retrieves information about all runs associated with the specified table.

Request Syntax

```
{  
  "DatabaseName": "string",  
  "MaxResults": number,  
  "NextToken": "string",  
  "TableName": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

DatabaseName

The name of the database where the table resides.

Type: String

Required: Yes

MaxResults

The maximum size of the response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

NextToken

A continuation token, if this is a continuation call.

Type: String

Required: No

TableName

The name of the table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "ColumnStatisticsTaskRuns": [
    {
      "CatalogID": "string",
      "ColumnNameList": [ "string" ],
      "ColumnStatisticsTaskRunId": "string",
      "CreationTime": number,
      "CustomerId": "string",
      "DatabaseName": "string",
      "DPUSeconds": number,
      "EndTime": number,
      "ErrorMessage": "string",
      "LastUpdated": number,
      "NumberOfWorkers": number,
      "Role": "string",
      "SampleSize": number,
      "SecurityConfiguration": "string",
      "StartTime": number,
      "Status": "string",
      "TableName": "string",
      "WorkerType": "string"
    }
  ],
  "NextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

ColumnStatisticsTaskRuns

A list of column statistics task runs.

Type: Array of [ColumnStatisticsTaskRun](#) objects

NextToken

A continuation token, if not all task runs have yet been returned.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

Name

The name of the connection definition to retrieve.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```

{
  "Connection": {
    "AuthenticationConfiguration": {
      "AuthenticationType": "string",
      "OAuth2Properties": {
        "OAuth2ClientApplication": {
          "AWSManagedClientApplicationReference": "string",
          "UserManagedClientApplicationClientId": "string"
        },
        "OAuth2GrantType": "string",
        "TokenUrl": "string",
        "TokenUrlParametersMap": {
          "string": "string"
        }
      },
      "SecretArn": "string"
    },
    "ConnectionProperties": {
      "string": "string"
    },
    "ConnectionType": "string",
    "CreationTime": number,
    "Description": "string",
    "LastConnectionValidationTime": number,
    "LastUpdatedBy": "string",
    "LastUpdateTime": number,
    "MatchCriteria": [ "string" ],
    "Name": "string",
    "PhysicalConnectionRequirements": {

```

```
    "AvailabilityZone": "string",
    "SecurityGroupIdList": [ "string" ],
    "SubnetId": "string"
  },
  "Status": "string",
  "StatusReason": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Connection

The requested connection definition.

Type: [Connection](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetConnections

Retrieves a list of connection definitions from the Data Catalog.

Request Syntax

```
{
  "CatalogId": "string",
  "Filter": {
    "ConnectionType": "string",
    "MatchCriteria": [ "string" ]
  },
  "HidePassword": boolean,
  "MaxResults": number,
  "NextToken": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog in which the connections reside. If none is provided, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Filter

A filter that controls which connections are returned.

Type: [GetConnectionsFilter](#) object

Required: No

HidePassword

Allows you to retrieve the connection metadata without returning the password. For instance, the AWS Glue console uses this flag to retrieve the connection, and does not display the password. Set this parameter when the caller might not have permission to use the AWS KMS key to decrypt the password, but it does have permission to access the rest of the connection properties.

Type: Boolean

Required: No

MaxResults

The maximum number of connections to return in one response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

NextToken

A continuation token, if this is a continuation call.

Type: String

Required: No

Response Syntax

```
{
  "ConnectionList": [
    {
      "AuthenticationConfiguration": {
        "AuthenticationType": "string",
        "OAuth2Properties": {
          "OAuth2ClientApplication": {
            "AWSManagedClientApplicationReference": "string",
            "UserManagedClientApplicationClientId": "string"
          }
        }
      }
    }
  ]
}
```

```

    },
    "OAuth2GrantType": "string",
    "TokenUrl": "string",
    "TokenUrlParametersMap": {
      "string" : "string"
    }
  },
  "SecretArn": "string"
},
"ConnectionProperties": {
  "string" : "string"
},
"ConnectionType": "string",
"CreationTime": number,
"Description": "string",
"LastConnectionValidationTime": number,
"LastUpdatedBy": "string",
"LastUpdatedTime": number,
"MatchCriteria": [ "string" ],
"Name": "string",
"PhysicalConnectionRequirements": {
  "AvailabilityZone": "string",
  "SecurityGroupIdList": [ "string" ],
  "SubnetId": "string"
},
"Status": "string",
"StatusReason": "string"
}
],
"NextToken": "string"
}

```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

ConnectionList

A list of requested connection definitions.

Type: Array of [Connection](#) objects

NextToken

A continuation token, if the list of connections returned does not include the last of the filtered connections.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)

- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)


```
},
  "LastCrawl": {
    "ErrorMessage": "string",
    "LogGroup": "string",
    "LogStream": "string",
    "MessagePrefix": "string",
    "StartTime": number,
    "Status": "string"
  },
  "LastUpdated": number,
  "LineageConfiguration": {
    "CrawlerLineageSettings": "string"
  },
  "Name": "string",
  "RecrawlPolicy": {
    "RecrawlBehavior": "string"
  },
  "Role": "string",
  "Schedule": {
    "ScheduleExpression": "string",
    "State": "string"
  },
  "SchemaChangePolicy": {
    "DeleteBehavior": "string",
    "UpdateBehavior": "string"
  },
  "State": "string",
  "TablePrefix": "string",
  "Targets": {
    "CatalogTargets": [
      {
        "ConnectionName": "string",
        "DatabaseName": "string",
        "DlqEventQueueArn": "string",
        "EventQueueArn": "string",
        "Tables": [ "string" ]
      }
    ],
    "DeltaTargets": [
      {
        "ConnectionName": "string",
        "CreateNativeDeltaTable": boolean,
        "DeltaTables": [ "string" ],
        "WriteManifest": boolean
      }
    ]
  }
}
```

```
    }
  ],
  "DynamoDBTargets": [
    {
      "Path": "string",
      "scanAll": boolean,
      "scanRate": number
    }
  ],
  "HudiTargets": [
    {
      "ConnectionName": "string",
      "Exclusions": [ "string" ],
      "MaximumTraversalDepth": number,
      "Paths": [ "string" ]
    }
  ],
  "IcebergTargets": [
    {
      "ConnectionName": "string",
      "Exclusions": [ "string" ],
      "MaximumTraversalDepth": number,
      "Paths": [ "string" ]
    }
  ],
  "JdbcTargets": [
    {
      "ConnectionName": "string",
      "EnableAdditionalMetadata": [ "string" ],
      "Exclusions": [ "string" ],
      "Path": "string"
    }
  ],
  "MongoDBTargets": [
    {
      "ConnectionName": "string",
      "Path": "string",
      "ScanAll": boolean
    }
  ],
  "S3Targets": [
    {
      "ConnectionName": "string",
      "DlqEventQueueArn": "string",
```

```
        "EventQueueArn": "string",
        "Exclusions": [ "string" ],
        "Path": "string",
        "SampleSize": number
    }
]
},
"Version": number
}
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Crawler

The metadata for the specified crawler.

Type: [Crawler](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetCrawlerMetrics

Retrieves metrics about specified crawlers.

Request Syntax

```
{
  "CrawlerNameList": [ string ],
  "MaxResults": number,
  "NextToken": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CrawlerNameList

A list of the names of crawlers about which to retrieve metrics.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 100 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\u007F\u00E0-\u00FF\u0100-\u017F\u0180-\u01FF\u0200-\u02FF\u0300-\u037F\u0380-\u03FF\u0400-\u047F\u0480-\u04FF\u0500-\u057F\u0580-\u05FF\u0600-\u06FF\u0700-\u07FF\u0800-\u08FF\u0900-\u097F\u0980-\u09FF\u0A00-\u0A7F\u0A80-\u0AFF\u0B00-\u0B7F\u0B80-\u0BFF\u0C00-\u0C7F\u0C80-\u0CFF\u0D00-\u0D7F\u0D80-\u0DBF\u0DC0-\u0DBF\u0E00-\u0E7F\u0E80-\u0EFF\u0F00-\u0F7F\u0F80-\u0FFF\u1000-\u107F\u1080-\u10FF\u1100-\u117F\u1180-\u11FF\u1200-\u127F\u1280-\u12FF\u1300-\u137F\u1380-\u13FF\u1400-\u147F\u1480-\u14FF\u1500-\u157F\u1580-\u15FF\u1600-\u167F\u1680-\u16FF\u1700-\u177F\u1780-\u17FF\u1800-\u187F\u1880-\u18FF\u1900-\u197F\u1980-\u19FF\u1A00-\u1A7F\u1A80-\u1AFF\u1B00-\u1B7F\u1B80-\u1BFF\u1C00-\u1C7F\u1C80-\u1CFF\u1D00-\u1D7F\u1D80-\u1DBF\u1E00-\u1E7F\u1E80-\u1EFF\u1F00-\u1F7F\u1F80-\u1FFF\u2000-\u207F\u2080-\u20FF\u2100-\u217F\u2180-\u21FF\u2200-\u227F\u2280-\u22FF\u2300-\u237F\u2380-\u23FF\u2400-\u247F\u2480-\u24FF\u2500-\u257F\u2580-\u25FF\u2600-\u267F\u2680-\u26FF\u2700-\u277F\u2780-\u27FF\u2800-\u287F\u2880-\u28FF\u2900-\u297F\u2980-\u29FF\u2A00-\u2A7F\u2A80-\u2AFF\u2B00-\u2B7F\u2B80-\u2BFF\u2C00-\u2C7F\u2C80-\u2CFF\u2D00-\u2D7F\u2D80-\u2DBF\u2E00-\u2E7F\u2E80-\u2EFF\u2F00-\u2F7F\u2F80-\u2FFF\u3000-\u307F\u3080-\u30FF\u3100-\u317F\u3180-\u31FF\u3200-\u327F\u3280-\u32FF\u3300-\u337F\u3380-\u33FF\u3400-\u347F\u3480-\u34FF\u3500-\u357F\u3580-\u35FF\u3600-\u367F\u3680-\u36FF\u3700-\u377F\u3780-\u37FF\u3800-\u387F\u3880-\u38FF\u3900-\u397F\u3980-\u39FF\u3A00-\u3A7F\u3A80-\u3AFF\u3B00-\u3B7F\u3B80-\u3BFF\u3C00-\u3C7F\u3C80-\u3CFF\u3D00-\u3D7F\u3D80-\u3DBF\u3E00-\u3E7F\u3E80-\u3EFF\u3F00-\u3F7F\u3F80-\u3FFF\u4000-\u407F\u4080-\u40FF\u4100-\u417F\u4180-\u41FF\u4200-\u427F\u4280-\u42FF\u4300-\u437F\u4380-\u43FF\u4400-\u447F\u4480-\u44FF\u4500-\u457F\u4580-\u45FF\u4600-\u467F\u4680-\u46FF\u4700-\u477F\u4780-\u47FF\u4800-\u487F\u4880-\u48FF\u4900-\u497F\u4980-\u49FF\u4A00-\u4A7F\u4A80-\u4AFF\u4B00-\u4B7F\u4B80-\u4BFF\u4C00-\u4C7F\u4C80-\u4CFF\u4D00-\u4D7F\u4D80-\u4DBF\u4E00-\u4E7F\u4E80-\u4EFF\u4F00-\u4F7F\u4F80-\u4FFF\u5000-\u507F\u5080-\u50FF\u5100-\u517F\u5180-\u51FF\u5200-\u527F\u5280-\u52FF\u5300-\u537F\u5380-\u53FF\u5400-\u547F\u5480-\u54FF\u5500-\u557F\u5580-\u55FF\u5600-\u567F\u5680-\u56FF\u5700-\u577F\u5780-\u57FF\u5800-\u587F\u5880-\u58FF\u5900-\u597F\u5980-\u59FF\u5A00-\u5A7F\u5A80-\u5AFF\u5B00-\u5B7F\u5B80-\u5BFF\u5C00-\u5C7F\u5C80-\u5CFF\u5D00-\u5D7F\u5D80-\u5DBF\u5E00-\u5E7F\u5E80-\u5EFF\u5F00-\u5F7F\u5F80-\u5FFF\u6000-\u607F\u6080-\u60FF\u6100-\u617F\u6180-\u61FF\u6200-\u627F\u6280-\u62FF\u6300-\u637F\u6380-\u63FF\u6400-\u647F\u6480-\u64FF\u6500-\u657F\u6580-\u65FF\u6600-\u667F\u6680-\u66FF\u6700-\u677F\u6780-\u67FF\u6800-\u687F\u6880-\u68FF\u6900-\u697F\u6980-\u69FF\u6A00-\u6A7F\u6A80-\u6AFF\u6B00-\u6B7F\u6B80-\u6BFF\u6C00-\u6C7F\u6C80-\u6CFF\u6D00-\u6D7F\u6D80-\u6DBF\u6E00-\u6E7F\u6E80-\u6EFF\u6F00-\u6F7F\u6F80-\u6FFF\u7000-\u707F\u7080-\u70FF\u7100-\u717F\u7180-\u71FF\u7200-\u727F\u7280-\u72FF\u7300-\u737F\u7380-\u73FF\u7400-\u747F\u7480-\u74FF\u7500-\u757F\u7580-\u75FF\u7600-\u767F\u7680-\u76FF\u7700-\u777F\u7780-\u77FF\u7800-\u787F\u7880-\u78FF\u7900-\u797F\u7980-\u79FF\u7A00-\u7A7F\u7A80-\u7AFF\u7B00-\u7B7F\u7B80-\u7BFF\u7C00-\u7C7F\u7C80-\u7CFF\u7D00-\u7D7F\u7D80-\u7DBF\u7E00-\u7E7F\u7E80-\u7EFF\u7F00-\u7F7F\u7F80-\u7FFF\u8000-\u807F\u8080-\u80FF\u8100-\u817F\u8180-\u81FF\u8200-\u827F\u8280-\u82FF\u8300-\u837F\u8380-\u83FF\u8400-\u847F\u8480-\u84FF\u8500-\u857F\u8580-\u85FF\u8600-\u867F\u8680-\u86FF\u8700-\u877F\u8780-\u87FF\u8800-\u887F\u8880-\u88FF\u8900-\u897F\u8980-\u89FF\u8A00-\u8A7F\u8A80-\u8AFF\u8B00-\u8B7F\u8B80-\u8BFF\u8C00-\u8C7F\u8C80-\u8CFF\u8D00-\u8D7F\u8D80-\u8DBF\u8E00-\u8E7F\u8E80-\u8EFF\u8F00-\u8F7F\u8F80-\u8FFF\u9000-\u907F\u9080-\u90FF\u9100-\u917F\u9180-\u91FF\u9200-\u927F\u9280-\u92FF\u9300-\u937F\u9380-\u93FF\u9400-\u947F\u9480-\u94FF\u9500-\u957F\u9580-\u95FF\u9600-\u967F\u9680-\u96FF\u9700-\u977F\u9780-\u97FF\u9800-\u987F\u9880-\u98FF\u9900-\u997F\u9980-\u99FF\u9A00-\u9A7F\u9A80-\u9AFF\u9B00-\u9B7F\u9B80-\u9BFF\u9C00-\u9C7F\u9C80-\u9CFF\u9D00-\u9D7F\u9D80-\u9DBF\u9E00-\u9E7F\u9E80-\u9EFF\u9F00-\u9F7F\u9F80-\u9FFF\uA000-\uA07F\uA080-\uA0FF\uA100-\uA17F\uA180-\uA1FF\uA200-\uA27F\uA280-\uA2FF\uA300-\uA37F\uA380-\uA3FF\uA400-\uA47F\uA480-\uA4FF\uA500-\uA57F\uA580-\uA5FF\uA600-\uA67F\uA680-\uA6FF\uA700-\uA77F\uA780-\uA7FF\uA800-\uA87F\uA880-\uA8FF\uA900-\uA97F\uA980-\uA9FF\uAA00-\uAA7F\uAA80-\uAAFF\uAB00-\uAB7F\uAB80-\uABFF\uAC00-\uAC7F\uAC80-\uACFF\uAD00-\uAD7F\uAD80-\uADBFF\uAE00-\uAE7F\uAE80-\uAEFF\uAF00-\uAF7F\uAF80-\uAFFF\uB000-\uB07F\uB080-\uB0FF\uB100-\uB17F\uB180-\uB1FF\uB200-\uB27F\uB280-\uB2FF\uB300-\uB37F\uB380-\uB3FF\uB400-\uB47F\uB480-\uB4FF\uB500-\uB57F\uB580-\uB5FF\uB600-\uB67F\uB680-\uB6FF\uB700-\uB77F\uB780-\uB7FF\uB800-\uB87F\uB880-\uB8FF\uB900-\uB97F\uB980-\uB9FF\uBA00-\uBA7F\uBA80-\uBAFF\uBB00-\uBB7F\uBB80-\uBBFF\uBC00-\uBC7F\uBC80-\uBCFF\uBD00-\uBD7F\uBD80-\uBDBF\uBE00-\uBE7F\uBE80-\uBEFF\uBF00-\uBF7F\uBF80-\uBFFF\uC000-\u007F\uC080-\u00FF\uC100-\u017F\uC180-\u01FF\uC200-\u027F\uC280-\u02FF\uC300-\u037F\uC380-\u03FF\uC400-\u047F\uC480-\u04FF\uC500-\u057F\uC580-\u05FF\uC600-\u067F\uC680-\u06FF\uC700-\u077F\uC780-\u07FF\uC800-\u087F\uC880-\u08FF\uC900-\u097F\uC980-\u09FF\uCA00-\u0A7F\uCA80-\u0AFF\uCB00-\u0B7F\uCB80-\u0BFF\uCC00-\u0C7F\uCC80-\u0CFF\uCD00-\u0D7F\uCD80-\u0DBF\uCE00-\u0E7F\uCE80-\u0EFF\uCF00-\u0F7F\uCF80-\u0FFF\uD000-\u107F\uD080-\u10FF\uD100-\u117F\uD180-\u11FF\uD200-\u127F\uD280-\u12FF\uD300-\u137F\uD380-\u13FF\uD400-\u147F\uD480-\u14FF\uD500-\u157F\uD580-\u15FF\uD600-\u167F\uD680-\u16FF\uD700-\u177F\uD780-\u17FF\uD800-\u187F\uD880-\u18FF\uD900-\u197F\uD980-\u19FF\uDA00-\u1A7F\uDA80-\u1AFF\uDB00-\u1B7F\uDB80-\u1BFF\uDC00-\u1C7F\uDC80-\u1CFF\uDD00-\u1D7F\uDD80-\u1DBF\uDE00-\u1E7F\uDE80-\u1EFF\uDF00-\u1F7F\uDF80-\u1FFF\uE000-\u207F\uE080-\u20FF\uE100-\u217F\uE180-\u21FF\uE200-\u227F\uE280-\u22FF\uE300-\u237F\uE380-\u23FF\uE400-\u247F\uE480-\u24FF\uE500-\u257F\uE580-\u25FF\uE600-\u267F\uE680-\u26FF\uE700-\u277F\uE780-\u27FF\uE800-\u287F\uE880-\u28FF\uE900-\u297F\uE980-\u29FF\uEA00-\u2A7F\uEA80-\u2AFF\uEB00-\u2B7F\uEB80-\u2BFF\uEC00-\u2C7F\uEC80-\u2CFF\uED00-\u2D7F\uED80-\u2DBF\uEE00-\u2E7F\uEE80-\u2EFF\uEF00-\u2F7F\uEF80-\u2FFF\uF000-\u307F\uF080-\u30FF\uF100-\u317F\uF180-\u31FF\uF200-\u327F\uF280-\u32FF\uF300-\u337F\uF380-\u33FF\uF400-\u347F\uF480-\u34FF\uF500-\u357F\uF580-\u35FF\uF600-\u367F\uF680-\u36FF\uF700-\u377F\uF780-\u37FF\uF800-\u387F\uF880-\u38FF\uF900-\u397F\uF980-\u39FF\uFA00-\u3A7F\uFA80-\u3AFF\uFB00-\u3B7F\uFB80-\u3BFF\uFC00-\u3C7F\uFC80-\u3CFF\uFD00-\u3D7F\uFD80-\u3DBF\uFE00-\u3E7F\uFE80-\u3EFF\uFF00-\u3F7F\uFF80-\u3FFF`

Required: No

MaxResults

The maximum size of a list to return.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

NextToken

A continuation token, if this is a continuation call.

Type: String

Required: No

Response Syntax

```
{
  "CrawlerMetricsList": [
    {
      "CrawlerName": "string",
      "LastRuntimeSeconds": number,
      "MedianRuntimeSeconds": number,
      "StillEstimating": boolean,
      "TablesCreated": number,
      "TablesDeleted": number,
      "TablesUpdated": number,
      "TimeLeftSeconds": number
    }
  ],
  "NextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

CrawlerMetricsList

A list of metrics for the specified crawler.

Type: Array of [CrawlerMetrics](#) objects

NextToken

A continuation token, if the returned list does not contain the last metric available.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetCrawlers

Retrieves metadata for all crawlers defined in the customer account.

Request Syntax

```
{  
  "MaxResults": number,  
  "NextToken": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

MaxResults

The number of crawlers to return on each call.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

NextToken

A continuation token, if this is a continuation request.

Type: String

Required: No

Response Syntax

```
{  
  "Crawlers": [  
    {  
      "Classifiers": [ "string" ],  
      "Configuration": "string",  
      "CrawlElapsedTime": number,  
    }  
  ]  
}
```

```
"CrawlerSecurityConfiguration": "string",
"CreationTime": number,
"DatabaseName": "string",
"Description": "string",
"LakeFormationConfiguration": {
  "AccountId": "string",
  "UseLakeFormationCredentials": boolean
},
>LastCrawl": {
  "ErrorMessage": "string",
  "LogGroup": "string",
  "LogStream": "string",
  "MessagePrefix": "string",
  "StartTime": number,
  "Status": "string"
},
>LastUpdated": number,
"LineageConfiguration": {
  "CrawlerLineageSettings": "string"
},
>Name": "string",
"RecrawlPolicy": {
  "RecrawlBehavior": "string"
},
>Role": "string",
"Schedule": {
  "ScheduleExpression": "string",
  "State": "string"
},
>SchemaChangePolicy": {
  "DeleteBehavior": "string",
  "UpdateBehavior": "string"
},
>State": "string",
"TablePrefix": "string",
"Targets": {
  "CatalogTargets": [
    {
      "ConnectionName": "string",
      "DatabaseName": "string",
      "DlqEventQueueArn": "string",
      "EventQueueArn": "string",
      "Tables": [ "string" ]
    }
  ]
}
```

```
],
  "DeltaTargets": [
    {
      "ConnectionName": "string",
      "CreateNativeDeltaTable": boolean,
      "DeltaTables": [ "string " ],
      "WriteManifest": boolean
    }
  ],
  "DynamoDBTargets": [
    {
      "Path": "string",
      "scanAll": boolean,
      "scanRate": number
    }
  ],
  "HudiTargets": [
    {
      "ConnectionName": "string",
      "Exclusions": [ "string " ],
      "MaximumTraversalDepth": number,
      "Paths": [ "string " ]
    }
  ],
  "IcebergTargets": [
    {
      "ConnectionName": "string",
      "Exclusions": [ "string " ],
      "MaximumTraversalDepth": number,
      "Paths": [ "string " ]
    }
  ],
  "JdbcTargets": [
    {
      "ConnectionName": "string",
      "EnableAdditionalMetadada": [ "string " ],
      "Exclusions": [ "string " ],
      "Path": "string"
    }
  ],
  "MongoDBTargets": [
    {
      "ConnectionName": "string",
      "Path": "string",
```

```
        "ScanAll": boolean
      }
    ],
    "S3Targets": [
      {
        "ConnectionName": "string",
        "DlqEventQueueArn": "string",
        "EventQueueArn": "string",
        "Exclusions": [ "string " ],
        "Path": "string",
        "SampleSize": number
      }
    ]
  },
  "Version": number
}
],
"NextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Crawlers

A list of crawler metadata.

Type: Array of [Crawler](#) objects

NextToken

A continuation token, if the returned list has not reached the end of those defined in this customer account.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetCustomEntityType

Retrieves the details of a custom pattern by specifying its name.

Request Syntax

```
{  
  "Name": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Name

The name of the custom pattern that you want to retrieve.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{  
  "ContextWords": [ "string" ],  
  "Name": "string",  
  "RegexString": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

ContextWords

A list of context words if specified when you created the custom pattern. If none of these context words are found within the vicinity of the regular expression the data will not be detected as sensitive data.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 20 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Name

The name of the custom pattern that you retrieved.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

RegexString

A regular expression string that is used for detecting sensitive data in a custom pattern.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetDatabase

Retrieves the definition of a specified database.

Request Syntax

```
{  
  "CatalogId": "string",  
  "Name": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog in which the database resides. If none is provided, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Name

The name of the database to retrieve. For Hive compatibility, this should be all lowercase.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "Database": {
    "CatalogId": "string",
    "CreateTableDefaultPermissions": [
      {
        "Permissions": [ "string" ],
        "Principal": {
          "DataLakePrincipalIdentifier": "string"
        }
      }
    ],
    "CreateTime": number,
    "Description": "string",
    "FederatedDatabase": {
      "ConnectionName": "string",
      "Identifier": "string"
    },
    "LocationUri": "string",
    "Name": "string",
    "Parameters": {
      "string" : "string"
    },
    "TargetDatabase": {
      "CatalogId": "string",
      "DatabaseName": "string",
      "Region": "string"
    }
  }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Database

The definition of the specified database in the Data Catalog.

Type: [Database](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

FederationSourceException

A federation source failed.

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)

- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetDatabases

Retrieves all databases defined in a given Data Catalog.

Request Syntax

```
{
  "CatalogId": "string",
  "MaxResults": number,
  "NextToken": "string",
  "ResourceShareType": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog from which to retrieve Databases. If none is provided, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

MaxResults

The maximum number of databases to return in one response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

NextToken

A continuation token, if this is a continuation call.

Type: String

Required: No

ResourceShareType

Allows you to specify that you want to list the databases shared with your account. The allowable values are FEDERATED, FOREIGN or ALL.

- If set to FEDERATED, will list the federated databases (referencing an external entity) shared with your account.
- If set to FOREIGN, will list the databases shared with your account.
- If set to ALL, will list the databases shared with your account, as well as the databases in your local account.

Type: String

Valid Values: FOREIGN | ALL | FEDERATED

Required: No

Response Syntax

```
{
  "DatabaseList": [
    {
      "CatalogId": "string",
      "CreateTableDefaultPermissions": [
        {
          "Permissions": [ "string" ],
          "Principal": {
            "DataLakePrincipalIdentifier": "string"
          }
        }
      ],
      "CreateTime": number,
      "Description": "string",
      "FederatedDatabase": {
        "ConnectionName": "string",
```



```
    "Identifier": "string"
  },
  "LocationUri": "string",
  "Name": "string",
  "Parameters": {
    "string" : "string"
  },
  "TargetDatabase": {
    "CatalogId": "string",
    "DatabaseName": "string",
    "Region": "string"
  }
}
],
"NextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

DatabaseList

A list of Database objects from the specified catalog.

Type: Array of [Database](#) objects

NextToken

A continuation token for paginating the returned list of tokens, returned if the current segment of the list is not the last.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetDataCatalogEncryptionSettings

Retrieves the security configuration for a specified catalog.

Request Syntax

```
{
  "CatalogId": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog to retrieve the security configuration for. If none is provided, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Response Syntax

```
{
  "DataCatalogEncryptionSettings": {
    "ConnectionPasswordEncryption": {
      "AwsKmsKeyId": "string",
      "ReturnConnectionPasswordEncrypted": boolean
    },
    "EncryptionAtRest": {
      "CatalogEncryptionMode": "string",
      "CatalogEncryptionServiceRole": "string",
      "SseAwsKmsKeyId": "string"
    }
  }
}
```

```
    }  
  }  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

[DataCatalogEncryptionSettings](#)

The requested security configuration.

Type: [DataCatalogEncryptionSettings](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetDataflowGraph

Transforms a Python script into a directed acyclic graph (DAG).

Request Syntax

```
{  
  "PythonScript": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

PythonScript

The Python script to transform.

Type: String

Required: No

Response Syntax

```
{  
  "DagEdges": [  
    {  
      "Source": "string",  
      "Target": "string",  
      "TargetParameter": "string"  
    }  
  ],  
  "DagNodes": [  
    {  
      "Args": [  
        {  
          "Name": "string",  
          "Param": boolean,  
        }  
      ]  
    }  
  ]  
}
```

```
        "Value": "string"
      }
    ],
    "Id": "string",
    "LineNumber": number,
    "NodeType": "string"
  }
]
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

DagEdges

A list of the edges in the resulting DAG.

Type: Array of [CodeGenEdge](#) objects

DagNodes

A list of the nodes in the resulting DAG.

Type: Array of [CodeGenNode](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetDataQualityResult

Retrieves the result of a data quality rule evaluation.

Request Syntax

```
{
  "ResultId": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

ResultId

A unique result ID for the data quality result.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\u0D7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "AnalyzerResults": [
    {
      "Description": "string",
      "EvaluatedMetrics": {
        "string": number
      },
      "EvaluationMessage": "string",
      "Name": "string"
    }
  ],
  "CompletedOn": number,
```

```

"DataSource": {
  "GlueTable": {
    "AdditionalOptions": {
      "string" : "string"
    },
    "CatalogId": "string",
    "ConnectionName": "string",
    "DatabaseName": "string",
    "TableName": "string"
  }
},
"EvaluationContext": "string",
"JobName": "string",
"JobRunId": "string",
"Observations": [
  {
    "Description": "string",
    "MetricBasedObservation": {
      "MetricName": "string",
      "MetricValues": {
        "ActualValue": number,
        "ExpectedValue": number,
        "LowerLimit": number,
        "UpperLimit": number
      },
      "NewRules": [ "string" ]
    }
  }
],
"ResultId": "string",
"RuleResults": [
  {
    "Description": "string",
    "EvaluatedMetrics": {
      "string" : number
    },
    "EvaluationMessage": "string",
    "Name": "string",
    "Result": "string"
  }
],
"RulesetEvaluationRunId": "string",
"RulesetName": "string",
"Score": number,

```

```
"StartedOn": number
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

AnalyzerResults

A list of `DataQualityAnalyzerResult` objects representing the results for each analyzer.

Type: Array of [DataQualityAnalyzerResult](#) objects

Array Members: Minimum number of 0 items. Maximum number of 2000 items.

CompletedOn

The date and time when the run for this data quality result was completed.

Type: Timestamp

DataSource

The table associated with the data quality result, if any.

Type: [DataSource](#) object

EvaluationContext

In the context of a job in AWS Glue Studio, each node in the canvas is typically assigned some sort of name and data quality nodes will have names. In the case of multiple nodes, the `evaluationContext` can differentiate the nodes.

Type: String

JobName

The job name associated with the data quality result, if any.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

JobRunId

The job run ID associated with the data quality result, if any.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Observations

A list of `DataQualityObservation` objects representing the observations generated after evaluating the rules and analyzers.

Type: Array of [DataQualityObservation](#) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

ResultId

A unique result ID for the data quality result.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

RuleResults

A list of `DataQualityRuleResult` objects representing the results for each rule.

Type: Array of [DataQualityRuleResult](#) objects

Array Members: Minimum number of 0 items. Maximum number of 2000 items.

RulesetEvaluationRunId

The unique run ID associated with the ruleset evaluation.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

RulesetName

The name of the ruleset associated with the data quality result.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Score

An aggregate data quality score. Represents the ratio of rules that passed to the total number of rules.

Type: Double

Valid Range: Minimum value of 0.0. Maximum value of 1.0.

StartedOn

The date and time when the run for this data quality result started.

Type: Timestamp

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

Type: String

ExecutionTime

The amount of time (in seconds) that the run consumed resources.

Type: Integer

LastModifiedOn

A timestamp. The last point in time when this data quality rule recommendation run was modified.

Type: Timestamp

NumberOfWorkers

The number of G.1X workers to be used in the run. The default is 5.

Type: Integer

RecommendedRuleset

When a start rule recommendation run completes, it creates a recommended ruleset (a set of rules). This member has those rules in Data Quality Definition Language (DQDL) format.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 65536.

Role

An IAM role supplied to encrypt the results of the run.

Type: String

RunId

The unique run identifier associated with this run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

StartedOn

The date and time when this run started.

Type: Timestamp

Status

The status for this run.

Type: String

Valid Values: STARTING | RUNNING | STOPPING | STOPPED | SUCCEEDED | FAILED | TIMEOUT

Timeout

The timeout for a run in minutes. This is the maximum time that a run can consume resources before it is terminated and enters TIMEOUT status. The default is 2,880 minutes (48 hours).

Type: Integer

Valid Range: Minimum value of 1.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetDataQualityRuleset

Returns an existing ruleset by identifier or name.

Request Syntax

```
{
  "Name": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Name

The name of the ruleset.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: [0020-0D7FF0E000-0FFFFD800DC00-0DBFF0DFFFt]*

Required: Yes

Response Syntax

```
{
  "CreatedOn": number,
  "Description": "string",
  "LastModifiedOn": number,
  "Name": "string",
  "RecommendationRunId": "string",
  "Ruleset": "string",
  "TargetTable": {
    "CatalogId": "string",
    "DatabaseName": "string",
    "TableName": "string"
  }
}
```

```
}  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

CreatedOn

A timestamp. The time and date that this data quality ruleset was created.

Type: Timestamp

Description

A description of the ruleset.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\i\n\t]*`

LastModifiedOn

A timestamp. The last point in time when this data quality ruleset was modified.

Type: Timestamp

Name

The name of the ruleset.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

RecommendationRunId

When a ruleset was created from a recommendation run, this run ID is generated to link the two together.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Ruleset

A Data Quality Definition Language (DQDL) ruleset. For more information, see the AWS Glue developer guide.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 65536.

TargetTable

The name and database name of the target table.

Type: [DataQualityTargetTable](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetDataQualityRulesetEvaluationRun

Retrieves a specific run where a ruleset is evaluated against a data source.

Request Syntax

```
{  
  "RunId": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

RunId

The unique run identifier associated with this run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{  
  "AdditionalDataSources": {  
    "string" : {  
      "GlueTable": {  
        "AdditionalOptions": {  
          "string" : "string"  
        },  
        "CatalogId": "string",  
        "ConnectionName": "string",  
        "DatabaseName": "string",
```



```

        "TableName": "string"
    }
}
},
"AdditionalRunOptions": {
    "CloudWatchMetricsEnabled": boolean,
    "CompositeRuleEvaluationMethod": "string",
    "ResultsS3Prefix": "string"
},
"CompletedOn": number,
"DataSource": {
    "GlueTable": {
        "AdditionalOptions": {
            "string" : "string"
        },
        "CatalogId": "string",
        "ConnectionName": "string",
        "DatabaseName": "string",
        "TableName": "string"
    }
},
"ErrorString": "string",
"ExecutionTime": number,
"LastModifiedOn": number,
"NumberOfWorkers": number,
"ResultIds": [ "string" ],
"Role": "string",
"RulesetNames": [ "string" ],
"RunId": "string",
"StartedOn": number,
>Status": "string",
"Timeout": number
}

```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

AdditionalDataSources

A map of reference strings to additional data sources you can specify for an evaluation run.

Type: String to [DataSource](#) object map

Key Length Constraints: Minimum length of 1. Maximum length of 255.

Key Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

[AdditionalRunOptions](#)

Additional run options you can specify for an evaluation run.

Type: [DataQualityEvaluationRunAdditionalRunOptions](#) object

[CompletedOn](#)

The date and time when this run was completed.

Type: Timestamp

[DataSource](#)

The data source (an AWS Glue table) associated with this evaluation run.

Type: [DataSource](#) object

[ErrorString](#)

The error strings that are associated with the run.

Type: String

[ExecutionTime](#)

The amount of time (in seconds) that the run consumed resources.

Type: Integer

[LastModifiedOn](#)

A timestamp. The last point in time when this data quality rule recommendation run was modified.

Type: Timestamp

[NumberOfWorkers](#)

The number of G.1X workers to be used in the run. The default is 5.

Type: Integer

ResultIds

A list of result IDs for the data quality results for the run.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 10 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Role

An IAM role supplied to encrypt the results of the run.

Type: String

RulesetNames

A list of ruleset names for the run. Currently, this parameter takes only one Ruleset name.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 10 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

RunId

The unique run identifier associated with this run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

StartedOn

The date and time when this run started.

Type: Timestamp

Status

The status for this run.

Type: String

Valid Values: STARTING | RUNNING | STOPPING | STOPPED | SUCCEEDED | FAILED | TIMEOUT

Timeout

The timeout for a run in minutes. This is the maximum time that a run can consume resources before it is terminated and enters TIMEOUT status. The default is 2,880 minutes (48 hours).

Type: Integer

Valid Range: Minimum value of 1.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetDevEndpoint

Retrieves information about a specified development endpoint.

Note

When you create a development endpoint in a virtual private cloud (VPC), AWS Glue returns only a private IP address, and the public IP address field is not populated. When you create a non-VPC development endpoint, AWS Glue returns only a public IP address.

Request Syntax

```
{  
  "EndpointName": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

EndpointName

Name of the DevEndpoint to retrieve information for.

Type: String

Required: Yes

Response Syntax

```
{  
  "DevEndpoint": {  
    "Arguments": {  
      "string" : "string"  
    },  
  },  
}
```

```
"AvailabilityZone": "string",
"CreatedTimestamp": number,
"EndpointName": "string",
"ExtraJarsS3Path": "string",
"ExtraPythonLibsS3Path": "string",
"FailureReason": "string",
"GlueVersion": "string",
"LastModifiedTimestamp": number,
"LastUpdateStatus": "string",
"NumberOfNodes": number,
"NumberOfWorkers": number,
"PrivateAddress": "string",
"PublicAddress": "string",
"PublicKey": "string",
"PublicKeys": [ "string" ],
"RoleArn": "string",
"SecurityConfiguration": "string",
"SecurityGroupIds": [ "string" ],
"Status": "string",
"SubnetId": "string",
"VpcId": "string",
"WorkerType": "string",
"YarnEndpointAddress": "string",
"ZeppelinRemoteSparkInterpreterPort": number
}
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

DevEndpoint

A DevEndpoint definition.

Type: [DevEndpoint](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetDevEndpoints

Retrieves all the development endpoints in this AWS account.

Note

When you create a development endpoint in a virtual private cloud (VPC), AWS Glue returns only a private IP address and the public IP address field is not populated. When you create a non-VPC development endpoint, AWS Glue returns only a public IP address.

Request Syntax

```
{
  "MaxResults": number,
  "NextToken": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

MaxResults

The maximum size of information to return.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

NextToken

A continuation token, if this is a continuation call.

Type: String

Required: No

Response Syntax

```
{
  "DevEndpoints": [
    {
      "Arguments": {
        "string": "string"
      },
      "AvailabilityZone": "string",
      "CreatedTimestamp": number,
      "EndpointName": "string",
      "ExtraJarsS3Path": "string",
      "ExtraPythonLibsS3Path": "string",
      "FailureReason": "string",
      "GlueVersion": "string",
      "LastModifiedTimestamp": number,
      "LastUpdateStatus": "string",
      "NumberOfNodes": number,
      "NumberOfWorkers": number,
      "PrivateAddress": "string",
      "PublicAddress": "string",
      "PublicKey": "string",
      "PublicKeys": [ "string" ],
      "RoleArn": "string",
      "SecurityConfiguration": "string",
      "SecurityGroupIds": [ "string" ],
      "Status": "string",
      "SubnetId": "string",
      "VpcId": "string",
      "WorkerType": "string",
      "YarnEndpointAddress": "string",
      "ZeppelinRemoteSparkInterpreterPort": number
    }
  ],
  "NextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

DevEndpoints

A list of DevEndpoint definitions.

Type: Array of [DevEndpoint](#) objects

NextToken

A continuation token, if not all DevEndpoint definitions have yet been returned.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetJob

Retrieves an existing job definition.

Request Syntax

```
{
  "JobName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

JobName

The name of the job definition to retrieve.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "Job": {
    "AllocatedCapacity": number,
    "CodeGenConfigurationNodes": {
      "string" : {
        "Aggregate": {
          "Aggs": [
            {
              "AggFunc": "string",
              "Column": [ "string" ]
            }
          ]
        }
      }
    ],
  },
}
```

```
"Groups": [
  [ "string" ]
],
"Inputs": [ "string" ],
"Name": "string"
},
"AmazonRedshiftSource": {
  "Data": {
    "AccessType": "string",
    "Action": "string",
    "AdvancedOptions": [
      {
        "Key": "string",
        "Value": "string"
      }
    ],
  },
  "CatalogDatabase": {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  },
  "CatalogRedshiftSchema": "string",
  "CatalogRedshiftTable": "string",
  "CatalogTable": {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  },
  "Connection": {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  },
  "CrawlerConnection": "string",
  "IamRole": {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  },
  "MergeAction": "string",
  "MergeClause": "string",
  "MergeWhenMatched": "string",
  "MergeWhenNotMatched": "string",
  "PostAction": "string",
```

```
    "PreAction": "string",
    "SampleQuery": "string",
    "Schema": {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    },
    "SelectedColumns": [
      {
        "Description": "string",
        "Label": "string",
        "Value": "string"
      }
    ],
    "SourceType": "string",
    "StagingTable": "string",
    "Table": {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    },
    "TablePrefix": "string",
    "TableSchema": [
      {
        "Description": "string",
        "Label": "string",
        "Value": "string"
      }
    ],
    "TempDir": "string",
    "Upsert": boolean
  },
  "Name": "string"
},
"AmazonRedshiftTarget": {
  "Data": {
    "AccessType": "string",
    "Action": "string",
    "AdvancedOptions": [
      {
        "Key": "string",
        "Value": "string"
      }
    ]
  }
},
```

```
"CatalogDatabase": {
  "Description": "string",
  "Label": "string",
  "Value": "string"
},
"CatalogRedshiftSchema": "string",
"CatalogRedshiftTable": "string",
"CatalogTable": {
  "Description": "string",
  "Label": "string",
  "Value": "string"
},
"Connection": {
  "Description": "string",
  "Label": "string",
  "Value": "string"
},
"CrawlerConnection": "string",
"IamRole": {
  "Description": "string",
  "Label": "string",
  "Value": "string"
},
"MergeAction": "string",
"MergeClause": "string",
"MergeWhenMatched": "string",
"MergeWhenNotMatched": "string",
"PostAction": "string",
"PreAction": "string",
"SampleQuery": "string",
"Schema": {
  "Description": "string",
  "Label": "string",
  "Value": "string"
},
"SelectedColumns": [
  {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  }
],
"SourceType": "string",
"StagingTable": "string",
```



```
    "Table": {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    },
    "TablePrefix": "string",
    "TableSchema": [
      {
        "Description": "string",
        "Label": "string",
        "Value": "string"
      }
    ],
    "TempDir": "string",
    "Upsert": boolean
  },
  "Inputs": [ "string" ],
  "Name": "string"
},
"ApplyMapping": {
  "Inputs": [ "string" ],
  "Mapping": [
    {
      "Children": [
        "Mapping"
      ],
      "Dropped": boolean,
      "FromPath": [ "string" ],
      "FromType": "string",
      "ToKey": "string",
      "ToType": "string"
    }
  ],
  "Name": "string"
},
"AthenaConnectorSource": {
  "ConnectionName": "string",
  "ConnectionTable": "string",
  "ConnectionType": "string",
  "ConnectorName": "string",
  "Name": "string",
  "OutputSchemas": [
    {
      "Columns": [
```

```
        {
            "Name": "string",
            "Type": "string"
        }
    ]
}
],
"SchemaName": "string"
},
"CatalogDeltaSource": {
    "AdditionalDeltaOptions": {
        "string" : "string"
    },
    "Database": "string",
    "Name": "string",
    "OutputSchemas": [
        {
            "Columns": [
                {
                    "Name": "string",
                    "Type": "string"
                }
            ]
        }
    ],
    "Table": "string"
},
"CatalogHudiSource": {
    "AdditionalHudiOptions": {
        "string" : "string"
    },
    "Database": "string",
    "Name": "string",
    "OutputSchemas": [
        {
            "Columns": [
                {
                    "Name": "string",
                    "Type": "string"
                }
            ]
        }
    ],
    "Table": "string"
}
```

```
},
  "CatalogKafkaSource": {
    "Database": "string",
    "DataPreviewOptions": {
      "PollingTime": number,
      "RecordPollingLimit": number
    },
    "DetectSchema": boolean,
    "Name": "string",
    "StreamingOptions": {
      "AddRecordTimestamp": "string",
      "Assign": "string",
      "BootstrapServers": "string",
      "Classification": "string",
      "ConnectionName": "string",
      "Delimiter": "string",
      "EmitConsumerLagMetrics": "string",
      "EndingOffsets": "string",
      "IncludeHeaders": boolean,
      "MaxOffsetsPerTrigger": number,
      "MinPartitions": number,
      "NumRetries": number,
      "PollTimeoutMs": number,
      "RetryIntervalMs": number,
      "SecurityProtocol": "string",
      "StartingOffsets": "string",
      "StartingTimestamp": "string",
      "SubscribePattern": "string",
      "TopicName": "string"
    },
    "Table": "string",
    "WindowSize": number
  },
  "CatalogKinesisSource": {
    "Database": "string",
    "DataPreviewOptions": {
      "PollingTime": number,
      "RecordPollingLimit": number
    },
    "DetectSchema": boolean,
    "Name": "string",
    "StreamingOptions": {
      "AddIdleTimeBetweenReads": boolean,
      "AddRecordTimestamp": "string",
```

```

        "AvoidEmptyBatches": boolean,
        "Classification": "string",
        "Delimiter": "string",
        "DescribeShardInterval": number,
        "EmitConsumerLagMetrics": "string",
        "EndpointUrl": "string",
        "IdleTimeBetweenReadsInMs": number,
        "MaxFetchRecordsPerShard": number,
        "MaxFetchTimeInMs": number,
        "MaxRecordPerRead": number,
        "MaxRetryIntervalMs": number,
        "NumRetries": number,
        "RetryIntervalMs": number,
        "RoleArn": "string",
        "RoleSessionName": "string",
        "StartingPosition": "string",
        "StartingTimestamp": "string",
        "StreamArn": "string",
        "StreamName": "string"
    },
    "Table": "string",
    "WindowSize": number
},
"CatalogSource": {
    "Database": "string",
    "Name": "string",
    "Table": "string"
},
"CatalogTarget": {
    "Database": "string",
    "Inputs": [ "string ],
    "Name": "string",
    "Table": "string"
},
"ConnectorDataSource": {
    "ConnectionType": "string",
    "Data": {
        "string": "string"
    },
    "Name": "string",
    "OutputSchemas": [
        {
            "Columns": [
                {

```

```

        "Name": "string",
        "Type": "string"
      }
    ]
  }
]
},
"ConnectorDataTarget": {
  "ConnectionType": "string",
  "Data": {
    "string" : "string"
  },
  "Inputs": [ "string" ],
  "Name": "string"
},
"CustomCode": {
  "ClassName": "string",
  "Code": "string",
  "Inputs": [ "string" ],
  "Name": "string",
  "OutputSchemas": [
    {
      "Columns": [
        {
          "Name": "string",
          "Type": "string"
        }
      ]
    }
  ]
},
"DirectJDBCSource": {
  "ConnectionName": "string",
  "ConnectionType": "string",
  "Database": "string",
  "Name": "string",
  "RedshiftTmpDir": "string",
  "Table": "string"
},
"DirectKafkaSource": {
  "DataPreviewOptions": {
    "PollingTime": number,
    "RecordPollingLimit": number
  }
},

```

```
"DetectSchema": boolean,
"Name": "string",
"StreamingOptions": {
  "AddRecordTimestamp": "string",
  "Assign": "string",
  "BootstrapServers": "string",
  "Classification": "string",
  "ConnectionName": "string",
  "Delimiter": "string",
  "EmitConsumerLagMetrics": "string",
  "EndingOffsets": "string",
  "IncludeHeaders": boolean,
  "MaxOffsetsPerTrigger": number,
  "MinPartitions": number,
  "NumRetries": number,
  "PollTimeoutMs": number,
  "RetryIntervalMs": number,
  "SecurityProtocol": "string",
  "StartingOffsets": "string",
  "StartingTimestamp": "string",
  "SubscribePattern": "string",
  "TopicName": "string"
},
"WindowSize": number
},
"DirectKinesisSource": {
  "DataPreviewOptions": {
    "PollingTime": number,
    "RecordPollingLimit": number
  },
  "DetectSchema": boolean,
  "Name": "string",
  "StreamingOptions": {
    "AddIdleTimeBetweenReads": boolean,
    "AddRecordTimestamp": "string",
    "AvoidEmptyBatches": boolean,
    "Classification": "string",
    "Delimiter": "string",
    "DescribeShardInterval": number,
    "EmitConsumerLagMetrics": "string",
    "EndpointUrl": "string",
    "IdleTimeBetweenReadsInMs": number,
    "MaxFetchRecordsPerShard": number,
    "MaxFetchTimeInMs": number,
```

```

        "MaxRecordPerRead": number,
        "MaxRetryIntervalMs": number,
        "NumRetries": number,
        "RetryIntervalMs": number,
        "RoleArn": "string",
        "RoleSessionName": "string",
        "StartingPosition": "string",
        "StartingTimestamp": "string",
        "StreamArn": "string",
        "StreamName": "string"
    },
    "WindowSize": number
},
"DropDuplicates": {
    "Columns": [
        [ "string" ]
    ],
    "Inputs": [ "string" ],
    "Name": "string"
},
"DropFields": {
    "Inputs": [ "string" ],
    "Name": "string",
    "Paths": [
        [ "string" ]
    ]
},
"DropNullFields": {
    "Inputs": [ "string" ],
    "Name": "string",
    "NullCheckBoxList": {
        "IsEmpty": boolean,
        "IsNegOne": boolean,
        "IsNullString": boolean
    },
    "NullTextList": [
        {
            "Datatype": {
                "Id": "string",
                "Label": "string"
            },
            "Value": "string"
        }
    ]
}
]

```

```
},
  "DynamicTransform": {
    "FunctionName": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ]
  },
  "Parameters": [
    {
      "IsOptional": boolean,
      "ListType": "string",
      "Name": "string",
      "Type": "string",
      "ValidationMessage": "string",
      "ValidationRule": "string",
      "Value": [ "string" ]
    }
  ],
  "Path": "string",
  "TransformName": "string",
  "Version": "string"
},
  "DynamoDBCatalogSource": {
    "Database": "string",
    "Name": "string",
    "Table": "string"
  },
  "EvaluateDataQuality": {
    "Inputs": [ "string" ],
    "Name": "string",
    "Output": "string",
    "PublishingOptions": {
      "CloudWatchMetricsEnabled": boolean,
      "EvaluationContext": "string",
      "ResultsPublishingEnabled": boolean,
      "ResultsS3Prefix": "string"
    }
  }
}
```



```

    },
    "Ruleset": "string",
    "StopJobOnFailureOptions": {
      "StopJobOnFailureTiming": "string"
    }
  },
  "EvaluateDataQualityMultiFrame": {
    "AdditionalDataSources": {
      "string" : "string"
    },
    "AdditionalOptions": {
      "string" : "string"
    },
    "Inputs": [ "string" ],
    "Name": "string",
    "PublishingOptions": {
      "CloudWatchMetricsEnabled": boolean,
      "EvaluationContext": "string",
      "ResultsPublishingEnabled": boolean,
      "ResultsS3Prefix": "string"
    },
    "Ruleset": "string",
    "StopJobOnFailureOptions": {
      "StopJobOnFailureTiming": "string"
    }
  },
  "FillMissingValues": {
    "FilledPath": "string",
    "ImputedPath": "string",
    "Inputs": [ "string" ],
    "Name": "string"
  },
  "Filter": {
    "Filters": [
      {
        "Negated": boolean,
        "Operation": "string",
        "Values": [
          {
            "Type": "string",
            "Value": [ "string" ]
          }
        ]
      }
    ]
  }
}

```

```
    ],
    "Inputs": [ "string" ],
    "LogicalOperator": "string",
    "Name": "string"
  },
  "GovernedCatalogSource": {
    "AdditionalOptions": {
      "BoundedFiles": number,
      "BoundedSize": number
    },
    "Database": "string",
    "Name": "string",
    "PartitionPredicate": "string",
    "Table": "string"
  },
  "GovernedCatalogTarget": {
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
      [ "string" ]
    ],
    "SchemaChangePolicy": {
      "EnableUpdateCatalog": boolean,
      "UpdateBehavior": "string"
    },
    "Table": "string"
  },
  "JDBCConnectorSource": {
    "AdditionalOptions": {
      "DataTypeMapping": {
        "string" : "string"
      },
      "FilterPredicate": "string",
      "JobBookmarkKeys": [ "string" ],
      "JobBookmarkKeysSortOrder": "string",
      "LowerBound": number,
      "NumPartitions": number,
      "PartitionColumn": "string",
      "UpperBound": number
    },
    "ConnectionName": "string",
    "ConnectionTable": "string",
    "ConnectionType": "string",
```

```
    "ConnectorName": "string",
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
    "Query": "string"
  },
  "JDBCConnectorTarget": {
    "AdditionalOptions": {
      "string" : "string"
    },
    "ConnectionName": "string",
    "ConnectionTable": "string",
    "ConnectionType": "string",
    "ConnectorName": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ]
  },
  "Join": {
    "Columns": [
      {
        "From": "string",
        "Keys": [
          [ "string" ]
        ]
      }
    ]
  }
],
```

```
    "Inputs": [ "string" ],
    "JoinType": "string",
    "Name": "string"
  },
  "Merge": {
    "Inputs": [ "string" ],
    "Name": "string",
    "PrimaryKeys": [
      [ "string" ]
    ],
    "Source": "string"
  },
  "MicrosoftSQLServerCatalogSource": {
    "Database": "string",
    "Name": "string",
    "Table": "string"
  },
  "MicrosoftSQLServerCatalogTarget": {
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "Table": "string"
  },
  "MySQLCatalogSource": {
    "Database": "string",
    "Name": "string",
    "Table": "string"
  },
  "MySQLCatalogTarget": {
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "Table": "string"
  },
  "OracleSQLCatalogSource": {
    "Database": "string",
    "Name": "string",
    "Table": "string"
  },
  "OracleSQLCatalogTarget": {
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "Table": "string"
  }
}
```

```
},
  "PIIDetection": {
    "EntityTypesToDetect": [ "string" ],
    "Inputs": [ "string" ],
    "MaskValue": "string",
    "Name": "string",
    "OutputColumnName": "string",
    "PiiType": "string",
    "SampleFraction": number,
    "ThresholdFraction": number
  },
  "PostgreSQLCatalogSource": {
    "Database": "string",
    "Name": "string",
    "Table": "string"
  },
  "PostgreSQLCatalogTarget": {
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "Table": "string"
  },
  "Recipe": {
    "Inputs": [ "string" ],
    "Name": "string",
    "RecipeReference": {
      "RecipeArn": "string",
      "RecipeVersion": "string"
    }
  },
  "RedshiftSource": {
    "Database": "string",
    "Name": "string",
    "RedshiftTmpDir": "string",
    "Table": "string",
    "TmpDirIAMRole": "string"
  },
  "RedshiftTarget": {
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "RedshiftTmpDir": "string",
    "Table": "string",
    "TmpDirIAMRole": "string",
```

```
    "UpsertRedshiftOptions": {
      "ConnectionName": "string",
      "TableLocation": "string",
      "UpsertKeys": [ "string" ]
    }
  },
  "RelationalCatalogSource": {
    "Database": "string",
    "Name": "string",
    "Table": "string"
  },
  "RenameField": {
    "Inputs": [ "string" ],
    "Name": "string",
    "SourcePath": [ "string" ],
    "TargetPath": [ "string" ]
  },
  "S3CatalogDeltaSource": {
    "AdditionalDeltaOptions": {
      "string" : "string"
    },
    "Database": "string",
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
    "Table": "string"
  },
  "S3CatalogHudiSource": {
    "AdditionalHudiOptions": {
      "string" : "string"
    },
    "Database": "string",
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
```

```

        {
            "Name": "string",
            "Type": "string"
        }
    ]
}
],
"Table": "string"
},
"S3CatalogSource": {
    "AdditionalOptions": {
        "BoundedFiles": number,
        "BoundedSize": number
    },
    "Database": "string",
    "Name": "string",
    "PartitionPredicate": "string",
    "Table": "string"
},
"S3CatalogTarget": {
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
        [ "string" ]
    ],
    "SchemaChangePolicy": {
        "EnableUpdateCatalog": boolean,
        "UpdateBehavior": "string"
    },
    "Table": "string"
},
"S3CsvSource": {
    "AdditionalOptions": {
        "BoundedFiles": number,
        "BoundedSize": number,
        "EnableSamplePath": boolean,
        "SamplePath": "string"
    },
    "CompressionType": "string",
    "Escaper": "string",
    "Exclusions": [ "string" ],
    "GroupFiles": "string",
    "GroupSize": "string",

```

```

    "MaxBand": number,
    "MaxFilesInBand": number,
    "Multiline": boolean,
    "Name": "string",
    "OptimizePerformance": boolean,
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
    "Paths": [ "string" ],
    "QuoteChar": "string",
    "Recurse": boolean,
    "Separator": "string",
    "SkipFirst": boolean,
    "WithHeader": boolean,
    "WriteHeader": boolean
  },
  "S3DeltaCatalogTarget": {
    "AdditionalOptions": {
      "string" : "string"
    },
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
      [ "string" ]
    ],
    "SchemaChangePolicy": {
      "EnableUpdateCatalog": boolean,
      "UpdateBehavior": "string"
    },
    "Table": "string"
  },
  "S3DeltaDirectTarget": {
    "AdditionalOptions": {
      "string" : "string"
    },
    "Compression": "string",

```



```

    "Format": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
      [ "string" ]
    ],
    "Path": "string",
    "SchemaChangePolicy": {
      "Database": "string",
      "EnableUpdateCatalog": boolean,
      "Table": "string",
      "UpdateBehavior": "string"
    }
  },
  "S3DeltaSource": {
    "AdditionalDeltaOptions": {
      "string" : "string"
    },
    "AdditionalOptions": {
      "BoundedFiles": number,
      "BoundedSize": number,
      "EnableSamplePath": boolean,
      "SamplePath": "string"
    },
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
    "Paths": [ "string" ]
  },
  "S3DirectTarget": {
    "Compression": "string",
    "Format": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
      [ "string" ]
    ]
  }
}

```

```

    ],
    "Path": "string",
    "SchemaChangePolicy": {
      "Database": "string",
      "EnableUpdateCatalog": boolean,
      "Table": "string",
      "UpdateBehavior": "string"
    }
  },
  "S3GlueParquetTarget": {
    "Compression": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
      [ "string" ]
    ],
    "Path": "string",
    "SchemaChangePolicy": {
      "Database": "string",
      "EnableUpdateCatalog": boolean,
      "Table": "string",
      "UpdateBehavior": "string"
    }
  },
  "S3HudiCatalogTarget": {
    "AdditionalOptions": {
      "string" : "string"
    },
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
      [ "string" ]
    ],
    "SchemaChangePolicy": {
      "EnableUpdateCatalog": boolean,
      "UpdateBehavior": "string"
    },
    "Table": "string"
  },
  "S3HudiDirectTarget": {
    "AdditionalOptions": {
      "string" : "string"
    },

```

```
"Compression": "string",
"Format": "string",
"Inputs": [ "string" ],
"Name": "string",
"PartitionKeys": [
  [ "string" ]
],
"Path": "string",
"SchemaChangePolicy": {
  "Database": "string",
  "EnableUpdateCatalog": boolean,
  "Table": "string",
  "UpdateBehavior": "string"
}
},
"S3HudiSource": {
  "AdditionalHudiOptions": {
    "string" : "string"
  },
  "AdditionalOptions": {
    "BoundedFiles": number,
    "BoundedSize": number,
    "EnableSamplePath": boolean,
    "SamplePath": "string"
  },
  "Name": "string",
  "OutputSchemas": [
    {
      "Columns": [
        {
          "Name": "string",
          "Type": "string"
        }
      ]
    }
  ],
  "Paths": [ "string" ]
},
"S3JsonSource": {
  "AdditionalOptions": {
    "BoundedFiles": number,
    "BoundedSize": number,
    "EnableSamplePath": boolean,
    "SamplePath": "string"
  }
}
```

```

    },
    "CompressionType": "string",
    "Exclusions": [ "string" ],
    "GroupFiles": "string",
    "GroupSize": "string",
    "JsonPath": "string",
    "MaxBand": number,
    "MaxFilesInBand": number,
    "Multiline": boolean,
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
    "Paths": [ "string" ],
    "Recurse": boolean
  },
  "S3ParquetSource": {
    "AdditionalOptions": {
      "BoundedFiles": number,
      "BoundedSize": number,
      "EnableSamplePath": boolean,
      "SamplePath": "string"
    },
    "CompressionType": "string",
    "Exclusions": [ "string" ],
    "GroupFiles": "string",
    "GroupSize": "string",
    "MaxBand": number,
    "MaxFilesInBand": number,
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ]
  }
}

```

```

    ]
  }
],
"Paths": [ "string" ],
"Recurse": boolean
},
"SelectFields": {
  "Inputs": [ "string" ],
  "Name": "string",
  "Paths": [
    [ "string" ]
  ]
},
"SelectFromCollection": {
  "Index": number,
  "Inputs": [ "string" ],
  "Name": "string"
},
"SnowflakeSource": {
  "Data": {
    "Action": "string",
    "AdditionalOptions": {
      "string" : "string"
    },
    "AutoPushdown": boolean,
    "Connection": {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    },
    "Database": "string",
    "IamRole": {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    },
    "MergeAction": "string",
    "MergeClause": "string",
    "MergeWhenMatched": "string",
    "MergeWhenNotMatched": "string",
    "PostAction": "string",
    "PreAction": "string",
    "SampleQuery": "string",
    "Schema": "string",

```

```
    "SelectedColumns": [
      {
        "Description": "string",
        "Label": "string",
        "Value": "string"
      }
    ],
    "SourceType": "string",
    "StagingTable": "string",
    "Table": "string",
    "TableSchema": [
      {
        "Description": "string",
        "Label": "string",
        "Value": "string"
      }
    ],
    "TempDir": "string",
    "Upsert": boolean
  },
  "Name": "string",
  "OutputSchemas": [
    {
      "Columns": [
        {
          "Name": "string",
          "Type": "string"
        }
      ]
    }
  ]
},
"SnowflakeTarget": {
  "Data": {
    "Action": "string",
    "AdditionalOptions": {
      "string" : "string"
    },
    "AutoPushdown": boolean,
    "Connection": {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    }
  },
}
```

```
    "Database": "string",
    "IamRole": {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    },
    "MergeAction": "string",
    "MergeClause": "string",
    "MergeWhenMatched": "string",
    "MergeWhenNotMatched": "string",
    "PostAction": "string",
    "PreAction": "string",
    "SampleQuery": "string",
    "Schema": "string",
    "SelectedColumns": [
      {
        "Description": "string",
        "Label": "string",
        "Value": "string"
      }
    ],
    "SourceType": "string",
    "StagingTable": "string",
    "Table": "string",
    "TableSchema": [
      {
        "Description": "string",
        "Label": "string",
        "Value": "string"
      }
    ],
    "TempDir": "string",
    "Upsert": boolean
  },
  "Inputs": [ "string" ],
  "Name": "string"
},
"SparkConnectorSource": {
  "AdditionalOptions": {
    "string" : "string"
  }
},
"ConnectionName": "string",
"ConnectionType": "string",
"ConnectorName": "string",
```

```
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
  },
  "SparkConnectorTarget": {
    "AdditionalOptions": {
      "string" : "string"
    },
    "ConnectionName": "string",
    "ConnectionType": "string",
    "ConnectorName": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ]
  },
  "SparkSQL": {
    "Inputs": [ "string" ],
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ]
  }
}
```



```
    ],
    "SqlAliases": [
      {
        "Alias": "string",
        "From": "string"
      }
    ],
    "SqlQuery": "string"
  },
  "Spigot": {
    "Inputs": [ "string" ],
    "Name": "string",
    "Path": "string",
    "Prob": number,
    "Topk": number
  },
  "SplitFields": {
    "Inputs": [ "string" ],
    "Name": "string",
    "Paths": [
      [ "string" ]
    ]
  },
  "Union": {
    "Inputs": [ "string" ],
    "Name": "string",
    "UnionType": "string"
  }
}
},
"Command": {
  "Name": "string",
  "PythonVersion": "string",
  "Runtime": "string",
  "ScriptLocation": "string"
},
"Connections": {
  "Connections": [ "string" ]
},
"CreatedOn": number,
"DefaultArguments": {
  "string": "string"
},
"Description": "string",
```

```
"ExecutionClass": "string",
"ExecutionProperty": {
  "MaxConcurrentRuns": number
},
"GlueVersion": "string",
"JobMode": "string",
"LastModifiedOn": number,
"LogUri": "string",
"MaintenanceWindow": "string",
"MaxCapacity": number,
"MaxRetries": number,
"Name": "string",
"NonOverridableArguments": {
  "string" : "string"
},
"NotificationProperty": {
  "NotifyDelayAfter": number
},
"NumberOfWorkers": number,
"ProfileName": "string",
"Role": "string",
"SecurityConfiguration": "string",
"SourceControlDetails": {
  "AuthStrategy": "string",
  "AuthToken": "string",
  "Branch": "string",
  "Folder": "string",
  "LastCommitId": "string",
  "Owner": "string",
  "Provider": "string",
  "Repository": "string"
},
"Timeout": number,
"WorkerType": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Job

The requested job definition.

Type: [Job](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)

- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetJobBookmark

Returns information on a job bookmark entry.

For more information about enabling and using job bookmarks, see:

- [Tracking processed data using job bookmarks](#)
- [Job parameters used by AWS Glue](#)
- [Job structure](#)

Request Syntax

```
{
  "JobName": "string",
  "RunId": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

[JobName](#)

The name of the job in question.

Type: String

Required: Yes

[RunId](#)

The unique run identifier associated with this job run.

Type: String

Required: No

Response Syntax

```
{
  "JobBookmarkEntry": {
    "Attempt": number,
    "JobBookmark": "string",
    "JobName": "string",
    "PreviousRunId": "string",
    "Run": number,
    "RunId": "string",
    "Version": number
  }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

JobBookmarkEntry

A structure that defines a point that a job can resume processing.

Type: [JobBookmarkEntry](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ValidationException

A value could not be validated.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetJobRun

Retrieves the metadata for a given job run. Job run history is accessible for 90 days for your workflow and job run.

Request Syntax

```
{  
  "JobName": "string",  
  "PredecessorsIncluded": boolean,  
  "RunId": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

JobName

Name of the job definition being run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

PredecessorsIncluded

True if a list of predecessor runs should be returned.

Type: Boolean

Required: No

RunId

The ID of the job run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "JobRun": {
    "AllocatedCapacity": number,
    "Arguments": {
      "string" : "string"
    },
    "Attempt": number,
    "CompletedOn": number,
    "DPUSeconds": number,
    "ErrorMessage": "string",
    "ExecutionClass": "string",
    "ExecutionTime": number,
    "GlueVersion": "string",
    "Id": "string",
    "JobMode": "string",
    "JobName": "string",
    "JobRunState": "string",
    "LastModifiedOn": number,
    "LogGroupName": "string",
    "MaintenanceWindow": "string",
    "MaxCapacity": number,
    "NotificationProperty": {
      "NotifyDelayAfter": number
    },
    "NumberOfWorkers": number,
    "PredecessorRuns": [
      {
        "JobName": "string",
        "RunId": "string"
      }
    ],
    "PreviousRunId": "string",
    "ProfileName": "string",
```

```
"SecurityConfiguration": "string",  
"StartedOn": number,  
"Timeout": number,  
"TriggerName": "string",  
"WorkerType": "string"  
}  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

JobRun

The requested job-run metadata.

Type: [JobRun](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

Type: String

Required: No

Response Syntax

```
{
  "JobRuns": [
    {
      "AllocatedCapacity": number,
      "Arguments": {
        "string" : "string"
      },
      "Attempt": number,
      "CompletedOn": number,
      "DPUSeconds": number,
      "ErrorMessage": "string",
      "ExecutionClass": "string",
      "ExecutionTime": number,
      "GlueVersion": "string",
      "Id": "string",
      "JobMode": "string",
      "JobName": "string",
      "JobRunState": "string",
      "LastModifiedOn": number,
      "LogGroupName": "string",
      "MaintenanceWindow": "string",
      "MaxCapacity": number,
      "NotificationProperty": {
        "NotifyDelayAfter": number
      },
      "NumberOfWorkers": number,
      "PredecessorRuns": [
        {
          "JobName": "string",
          "RunId": "string"
        }
      ],
      "PreviousRunId": "string",
      "ProfileName": "string",
      "SecurityConfiguration": "string",
      "StartedOn": number,
      "Timeout": number,
    }
  ]
}
```

```
    "TriggerName": "string",  
    "WorkerType": "string"  
  },  
],  
"NextToken": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

JobRuns

A list of job-run metadata objects.

Type: Array of [JobRun](#) objects

NextToken

A continuation token, if not all requested job runs have been returned.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetJobs

Retrieves all current job definitions.

Request Syntax

```
{
  "MaxResults": number,
  "NextToken": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

MaxResults

The maximum size of the response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

NextToken

A continuation token, if this is a continuation call.

Type: String

Required: No

Response Syntax

```
{
  "Jobs": [
    {
      "AllocatedCapacity": number,
      "CodeGenConfigurationNodes": {
        "string" : {
```



```
"Aggregate": {
  "Aggs": [
    {
      "AggFunc": "string",
      "Column": [ "string" ]
    }
  ],
  "Groups": [
    [ "string" ]
  ],
  "Inputs": [ "string" ],
  "Name": "string"
},
"AmazonRedshiftSource": {
  "Data": {
    "AccessType": "string",
    "Action": "string",
    "AdvancedOptions": [
      {
        "Key": "string",
        "Value": "string"
      }
    ],
    "CatalogDatabase": {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    },
    "CatalogRedshiftSchema": "string",
    "CatalogRedshiftTable": "string",
    "CatalogTable": {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    },
    "Connection": {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    },
    "CrawlerConnection": "string",
    "IamRole": {
      "Description": "string",
      "Label": "string",
```

```
    "Value": "string"
  },
  "MergeAction": "string",
  "MergeClause": "string",
  "MergeWhenMatched": "string",
  "MergeWhenNotMatched": "string",
  "PostAction": "string",
  "PreAction": "string",
  "SampleQuery": "string",
  "Schema": {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  },
  "SelectedColumns": [
    {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    }
  ],
  "SourceType": "string",
  "StagingTable": "string",
  "Table": {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  },
  "TablePrefix": "string",
  "TableSchema": [
    {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    }
  ],
  "TempDir": "string",
  "Upsert": boolean
},
"Name": "string"
},
"AmazonRedshiftTarget": {
  "Data": {
    "AccessType": "string",
```

```
"Action": "string",
"AdvancedOptions": [
  {
    "Key": "string",
    "Value": "string"
  }
],
"CatalogDatabase": {
  "Description": "string",
  "Label": "string",
  "Value": "string"
},
"CatalogRedshiftSchema": "string",
"CatalogRedshiftTable": "string",
"CatalogTable": {
  "Description": "string",
  "Label": "string",
  "Value": "string"
},
"Connection": {
  "Description": "string",
  "Label": "string",
  "Value": "string"
},
"CrawlerConnection": "string",
"IamRole": {
  "Description": "string",
  "Label": "string",
  "Value": "string"
},
"MergeAction": "string",
"MergeClause": "string",
"MergeWhenMatched": "string",
"MergeWhenNotMatched": "string",
"PostAction": "string",
"PreAction": "string",
"SampleQuery": "string",
"Schema": {
  "Description": "string",
  "Label": "string",
  "Value": "string"
},
"SelectedColumns": [
  {
```

```

        "Description": "string",
        "Label": "string",
        "Value": "string"
    }
],
"SourceType": "string",
"StagingTable": "string",
"Table": {
    "Description": "string",
    "Label": "string",
    "Value": "string"
},
"TablePrefix": "string",
"TableSchema": [
    {
        "Description": "string",
        "Label": "string",
        "Value": "string"
    }
],
"TempDir": "string",
"Upsert": boolean
},
"Inputs": [ "string" ],
"Name": "string"
},
"ApplyMapping": {
    "Inputs": [ "string" ],
    "Mapping": [
        {
            "Children": [
                "Mapping"
            ],
            "Dropped": boolean,
            "FromPath": [ "string" ],
            "FromType": "string",
            "ToKey": "string",
            "ToType": "string"
        }
    ],
    "Name": "string"
},
"AthenaConnectorSource": {
    "ConnectionName": "string",

```

```
"ConnectionTable": "string",
"ConnectionType": "string",
"ConnectorName": "string",
"Name": "string",
"OutputSchemas": [
  {
    "Columns": [
      {
        "Name": "string",
        "Type": "string"
      }
    ]
  }
],
"SchemaName": "string"
},
"CatalogDeltaSource": {
  "AdditionalDeltaOptions": {
    "string" : "string"
  },
  "Database": "string",
  "Name": "string",
  "OutputSchemas": [
    {
      "Columns": [
        {
          "Name": "string",
          "Type": "string"
        }
      ]
    }
  ],
  "Table": "string"
},
"CatalogHudiSource": {
  "AdditionalHudiOptions": {
    "string" : "string"
  },
  "Database": "string",
  "Name": "string",
  "OutputSchemas": [
    {
      "Columns": [
        {
```

```

        "Name": "string",
        "Type": "string"
    }
]
},
"Table": "string"
},
"CatalogKafkaSource": {
    "Database": "string",
    "DataPreviewOptions": {
        "PollingTime": number,
        "RecordPollingLimit": number
    },
    "DetectSchema": boolean,
    "Name": "string",
    "StreamingOptions": {
        "AddRecordTimestamp": "string",
        "Assign": "string",
        "BootstrapServers": "string",
        "Classification": "string",
        "ConnectionName": "string",
        "Delimiter": "string",
        "EmitConsumerLagMetrics": "string",
        "EndingOffsets": "string",
        "IncludeHeaders": boolean,
        "MaxOffsetsPerTrigger": number,
        "MinPartitions": number,
        "NumRetries": number,
        "PollTimeoutMs": number,
        "RetryIntervalMs": number,
        "SecurityProtocol": "string",
        "StartingOffsets": "string",
        "StartingTimestamp": "string",
        "SubscribePattern": "string",
        "TopicName": "string"
    },
    "Table": "string",
    "WindowSize": number
},
"CatalogKinesisSource": {
    "Database": "string",
    "DataPreviewOptions": {
        "PollingTime": number,

```

```
    "RecordPollingLimit": number
  },
  "DetectSchema": boolean,
  "Name": "string",
  "StreamingOptions": {
    "AddIdleTimeBetweenReads": boolean,
    "AddRecordTimestamp": "string",
    "AvoidEmptyBatches": boolean,
    "Classification": "string",
    "Delimiter": "string",
    "DescribeShardInterval": number,
    "EmitConsumerLagMetrics": "string",
    "EndpointUrl": "string",
    "IdleTimeBetweenReadsInMs": number,
    "MaxFetchRecordsPerShard": number,
    "MaxFetchTimeInMs": number,
    "MaxRecordPerRead": number,
    "MaxRetryIntervalMs": number,
    "NumRetries": number,
    "RetryIntervalMs": number,
    "RoleArn": "string",
    "RoleSessionName": "string",
    "StartingPosition": "string",
    "StartingTimestamp": "string",
    "StreamArn": "string",
    "StreamName": "string"
  },
  "Table": "string",
  "WindowSize": number
},
"CatalogSource": {
  "Database": "string",
  "Name": "string",
  "Table": "string"
},
"CatalogTarget": {
  "Database": "string",
  "Inputs": [ "string" ],
  "Name": "string",
  "Table": "string"
},
"ConnectorDataSource": {
  "ConnectionType": "string",
  "Data": {
```

```
        "string" : "string"
    },
    "Name": "string",
    "OutputSchemas": [
        {
            "Columns": [
                {
                    "Name": "string",
                    "Type": "string"
                }
            ]
        }
    ]
},
"ConnectorDataTarget": {
    "ConnectionType": "string",
    "Data": {
        "string" : "string"
    },
    "Inputs": [ "string" ],
    "Name": "string"
},
"CustomCode": {
    "ClassName": "string",
    "Code": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "OutputSchemas": [
        {
            "Columns": [
                {
                    "Name": "string",
                    "Type": "string"
                }
            ]
        }
    ]
},
"DirectJDBCSource": {
    "ConnectionName": "string",
    "ConnectionType": "string",
    "Database": "string",
    "Name": "string",
    "RedshiftTmpDir": "string",
```



```
    "Table": "string"
  },
  "DirectKafkaSource": {
    "DataPreviewOptions": {
      "PollingTime": number,
      "RecordPollingLimit": number
    },
    "DetectSchema": boolean,
    "Name": "string",
    "StreamingOptions": {
      "AddRecordTimestamp": "string",
      "Assign": "string",
      "BootstrapServers": "string",
      "Classification": "string",
      "ConnectionName": "string",
      "Delimiter": "string",
      "EmitConsumerLagMetrics": "string",
      "EndingOffsets": "string",
      "IncludeHeaders": boolean,
      "MaxOffsetsPerTrigger": number,
      "MinPartitions": number,
      "NumRetries": number,
      "PollTimeoutMs": number,
      "RetryIntervalMs": number,
      "SecurityProtocol": "string",
      "StartingOffsets": "string",
      "StartingTimestamp": "string",
      "SubscribePattern": "string",
      "TopicName": "string"
    },
    "WindowSize": number
  },
  "DirectKinesisSource": {
    "DataPreviewOptions": {
      "PollingTime": number,
      "RecordPollingLimit": number
    },
    "DetectSchema": boolean,
    "Name": "string",
    "StreamingOptions": {
      "AddIdleTimeBetweenReads": boolean,
      "AddRecordTimestamp": "string",
      "AvoidEmptyBatches": boolean,
      "Classification": "string",
```

```

    "Delimiter": "string",
    "DescribeShardInterval": number,
    "EmitConsumerLagMetrics": "string",
    "EndpointUrl": "string",
    "IdleTimeBetweenReadsInMs": number,
    "MaxFetchRecordsPerShard": number,
    "MaxFetchTimeInMs": number,
    "MaxRecordPerRead": number,
    "MaxRetryIntervalMs": number,
    "NumRetries": number,
    "RetryIntervalMs": number,
    "RoleArn": "string",
    "RoleSessionName": "string",
    "StartingPosition": "string",
    "StartingTimestamp": "string",
    "StreamArn": "string",
    "StreamName": "string"
  },
  "WindowSize": number
},
"DropDuplicates": {
  "Columns": [
    [ "string" ]
  ],
  "Inputs": [ "string" ],
  "Name": "string"
},
"DropFields": {
  "Inputs": [ "string" ],
  "Name": "string",
  "Paths": [
    [ "string" ]
  ]
},
"DropNullFields": {
  "Inputs": [ "string" ],
  "Name": "string",
  "NullCheckBoxList": {
    "IsEmpty": boolean,
    "IsNegOne": boolean,
    "IsNullString": boolean
  },
  "NullTextList": [
    {

```

```
        "Datatype": {
            "Id": "string",
            "Label": "string"
        },
        "Value": "string"
    }
]
},
"DynamicTransform": {
    "FunctionName": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "OutputSchemas": [
        {
            "Columns": [
                {
                    "Name": "string",
                    "Type": "string"
                }
            ]
        }
    ],
},
"Parameters": [
    {
        "IsOptional": boolean,
        "ListType": "string",
        "Name": "string",
        "Type": "string",
        "ValidationMessage": "string",
        "ValidationRule": "string",
        "Value": [ "string" ]
    }
],
"Path": "string",
"TransformName": "string",
"Version": "string"
},
"DynamoDBCatalogSource": {
    "Database": "string",
    "Name": "string",
    "Table": "string"
},
"EvaluateDataQuality": {
    "Inputs": [ "string" ],
```

```

    "Name": "string",
    "Output": "string",
    "PublishingOptions": {
      "CloudWatchMetricsEnabled": boolean,
      "EvaluationContext": "string",
      "ResultsPublishingEnabled": boolean,
      "ResultsS3Prefix": "string"
    },
    "Ruleset": "string",
    "StopJobOnFailureOptions": {
      "StopJobOnFailureTiming": "string"
    }
  },
  "EvaluateDataQualityMultiFrame": {
    "AdditionalDataSources": {
      "string" : "string"
    },
    "AdditionalOptions": {
      "string" : "string"
    },
    "Inputs": [ "string" ],
    "Name": "string",
    "PublishingOptions": {
      "CloudWatchMetricsEnabled": boolean,
      "EvaluationContext": "string",
      "ResultsPublishingEnabled": boolean,
      "ResultsS3Prefix": "string"
    },
    "Ruleset": "string",
    "StopJobOnFailureOptions": {
      "StopJobOnFailureTiming": "string"
    }
  },
  "FillMissingValues": {
    "FilledPath": "string",
    "ImputedPath": "string",
    "Inputs": [ "string" ],
    "Name": "string"
  },
  "Filter": {
    "Filters": [
      {
        "Negated": boolean,
        "Operation": "string",

```

```

        "Values": [
            {
                "Type": "string",
                "Value": [ "string" ]
            }
        ]
    },
    "Inputs": [ "string" ],
    "LogicalOperator": "string",
    "Name": "string"
},
"GovernedCatalogSource": {
    "AdditionalOptions": {
        "BoundedFiles": number,
        "BoundedSize": number
    },
    "Database": "string",
    "Name": "string",
    "PartitionPredicate": "string",
    "Table": "string"
},
"GovernedCatalogTarget": {
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
        [ "string" ]
    ],
    "SchemaChangePolicy": {
        "EnableUpdateCatalog": boolean,
        "UpdateBehavior": "string"
    },
    "Table": "string"
},
"JDBCConnectorSource": {
    "AdditionalOptions": {
        "DataTypeMapping": {
            "string" : "string"
        },
        "FilterPredicate": "string",
        "JobBookmarkKeys": [ "string" ],
        "JobBookmarkKeysSortOrder": "string",
        "LowerBound": number,

```

```
        "NumPartitions": number,
        "PartitionColumn": "string",
        "UpperBound": number
    },
    "ConnectionName": "string",
    "ConnectionTable": "string",
    "ConnectionType": "string",
    "ConnectorName": "string",
    "Name": "string",
    "OutputSchemas": [
        {
            "Columns": [
                {
                    "Name": "string",
                    "Type": "string"
                }
            ]
        }
    ],
    "Query": "string"
},
"JDBCConnectorTarget": {
    "AdditionalOptions": {
        "string": "string"
    },
    "ConnectionName": "string",
    "ConnectionTable": "string",
    "ConnectionType": "string",
    "ConnectorName": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "OutputSchemas": [
        {
            "Columns": [
                {
                    "Name": "string",
                    "Type": "string"
                }
            ]
        }
    ]
},
"Join": {
    "Columns": [
```

```
    {
      "From": "string",
      "Keys": [
        [ "string" ]
      ]
    }
  ],
  "Inputs": [ "string" ],
  "JoinType": "string",
  "Name": "string"
},
"Merge": {
  "Inputs": [ "string" ],
  "Name": "string",
  "PrimaryKeys": [
    [ "string" ]
  ],
  "Source": "string"
},
"MicrosoftSQLServerCatalogSource": {
  "Database": "string",
  "Name": "string",
  "Table": "string"
},
"MicrosoftSQLServerCatalogTarget": {
  "Database": "string",
  "Inputs": [ "string" ],
  "Name": "string",
  "Table": "string"
},
"MySQLCatalogSource": {
  "Database": "string",
  "Name": "string",
  "Table": "string"
},
"MySQLCatalogTarget": {
  "Database": "string",
  "Inputs": [ "string" ],
  "Name": "string",
  "Table": "string"
},
"OracleSQLCatalogSource": {
  "Database": "string",
  "Name": "string",
```

```
    "Table": "string"
  },
  "OracleSQLCatalogTarget": {
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "Table": "string"
  },
  "PIIDetection": {
    "EntityTypesToDetect": [ "string" ],
    "Inputs": [ "string" ],
    "MaskValue": "string",
    "Name": "string",
    "OutputColumnName": "string",
    "PiiType": "string",
    "SampleFraction": number,
    "ThresholdFraction": number
  },
  "PostgreSQLCatalogSource": {
    "Database": "string",
    "Name": "string",
    "Table": "string"
  },
  "PostgreSQLCatalogTarget": {
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "Table": "string"
  },
  "Recipe": {
    "Inputs": [ "string" ],
    "Name": "string",
    "RecipeReference": {
      "RecipeArn": "string",
      "RecipeVersion": "string"
    }
  },
  "RedshiftSource": {
    "Database": "string",
    "Name": "string",
    "RedshiftTmpDir": "string",
    "Table": "string",
    "TmpDirIAMRole": "string"
  },
}
```



```
"RedshiftTarget": {
  "Database": "string",
  "Inputs": [ "string" ],
  "Name": "string",
  "RedshiftTmpDir": "string",
  "Table": "string",
  "TmpDirIAMRole": "string",
  "UpsertRedshiftOptions": {
    "ConnectionName": "string",
    "TableLocation": "string",
    "UpsertKeys": [ "string" ]
  }
},
"RelationalCatalogSource": {
  "Database": "string",
  "Name": "string",
  "Table": "string"
},
"RenameField": {
  "Inputs": [ "string" ],
  "Name": "string",
  "SourcePath": [ "string" ],
  "TargetPath": [ "string" ]
},
"S3CatalogDeltaSource": {
  "AdditionalDeltaOptions": {
    "string" : "string"
  },
  "Database": "string",
  "Name": "string",
  "OutputSchemas": [
    {
      "Columns": [
        {
          "Name": "string",
          "Type": "string"
        }
      ]
    }
  ],
  "Table": "string"
},
"S3CatalogHudiSource": {
  "AdditionalHudiOptions": {
```

```
        "string" : "string"
    },
    "Database": "string",
    "Name": "string",
    "OutputSchemas": [
        {
            "Columns": [
                {
                    "Name": "string",
                    "Type": "string"
                }
            ]
        }
    ],
    "Table": "string"
},
"S3CatalogSource": {
    "AdditionalOptions": {
        "BoundedFiles": number,
        "BoundedSize": number
    },
    "Database": "string",
    "Name": "string",
    "PartitionPredicate": "string",
    "Table": "string"
},
"S3CatalogTarget": {
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
        [ "string" ]
    ],
    "SchemaChangePolicy": {
        "EnableUpdateCatalog": boolean,
        "UpdateBehavior": "string"
    },
    "Table": "string"
},
"S3CsvSource": {
    "AdditionalOptions": {
        "BoundedFiles": number,
        "BoundedSize": number,
        "EnableSamplePath": boolean,
```

```

        "SamplePath": "string"
    },
    "CompressionType": "string",
    "Escaper": "string",
    "Exclusions": [ "string" ],
    "GroupFiles": "string",
    "GroupSize": "string",
    "MaxBand": number,
    "MaxFilesInBand": number,
    "Multiline": boolean,
    "Name": "string",
    "OptimizePerformance": boolean,
    "OutputSchemas": [
        {
            "Columns": [
                {
                    "Name": "string",
                    "Type": "string"
                }
            ]
        }
    ],
    "Paths": [ "string" ],
    "QuoteChar": "string",
    "Recurse": boolean,
    "Separator": "string",
    "SkipFirst": boolean,
    "WithHeader": boolean,
    "WriteHeader": boolean
},
"S3DeltaCatalogTarget": {
    "AdditionalOptions": {
        "string" : "string"
    },
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
        [ "string" ]
    ],
    "SchemaChangePolicy": {
        "EnableUpdateCatalog": boolean,
        "UpdateBehavior": "string"
    }
},

```

```

    "Table": "string"
  },
  "S3DeltaDirectTarget": {
    "AdditionalOptions": {
      "string" : "string"
    },
    "Compression": "string",
    "Format": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
      [ "string" ]
    ],
    "Path": "string",
    "SchemaChangePolicy": {
      "Database": "string",
      "EnableUpdateCatalog": boolean,
      "Table": "string",
      "UpdateBehavior": "string"
    }
  },
  "S3DeltaSource": {
    "AdditionalDeltaOptions": {
      "string" : "string"
    },
    "AdditionalOptions": {
      "BoundedFiles": number,
      "BoundedSize": number,
      "EnableSamplePath": boolean,
      "SamplePath": "string"
    },
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
    "Paths": [ "string" ]
  },

```

```
"S3DirectTarget": {
  "Compression": "string",
  "Format": "string",
  "Inputs": [ "string" ],
  "Name": "string",
  "PartitionKeys": [
    [ "string" ]
  ],
  "Path": "string",
  "SchemaChangePolicy": {
    "Database": "string",
    "EnableUpdateCatalog": boolean,
    "Table": "string",
    "UpdateBehavior": "string"
  }
},
"S3GlueParquetTarget": {
  "Compression": "string",
  "Inputs": [ "string" ],
  "Name": "string",
  "PartitionKeys": [
    [ "string" ]
  ],
  "Path": "string",
  "SchemaChangePolicy": {
    "Database": "string",
    "EnableUpdateCatalog": boolean,
    "Table": "string",
    "UpdateBehavior": "string"
  }
},
"S3HudiCatalogTarget": {
  "AdditionalOptions": {
    "string" : "string"
  },
  "Database": "string",
  "Inputs": [ "string" ],
  "Name": "string",
  "PartitionKeys": [
    [ "string" ]
  ],
  "SchemaChangePolicy": {
    "EnableUpdateCatalog": boolean,
    "UpdateBehavior": "string"
  }
}
```

```

    },
    "Table": "string"
  },
  "S3HudiDirectTarget": {
    "AdditionalOptions": {
      "string": "string"
    },
    "Compression": "string",
    "Format": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
      [ "string" ]
    ],
    "Path": "string",
    "SchemaChangePolicy": {
      "Database": "string",
      "EnableUpdateCatalog": boolean,
      "Table": "string",
      "UpdateBehavior": "string"
    }
  },
  "S3HudiSource": {
    "AdditionalHudiOptions": {
      "string": "string"
    },
    "AdditionalOptions": {
      "BoundedFiles": number,
      "BoundedSize": number,
      "EnableSamplePath": boolean,
      "SamplePath": "string"
    },
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
    "Paths": [ "string" ]
  }
}

```

```
},
  "S3JsonSource": {
    "AdditionalOptions": {
      "BoundedFiles": number,
      "BoundedSize": number,
      "EnableSamplePath": boolean,
      "SamplePath": "string"
    },
    "CompressionType": "string",
    "Exclusions": [ "string" ],
    "GroupFiles": "string",
    "GroupSize": "string",
    "JsonPath": "string",
    "MaxBand": number,
    "MaxFilesInBand": number,
    "Multiline": boolean,
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
    "Paths": [ "string" ],
    "Recurse": boolean
  },
  "S3ParquetSource": {
    "AdditionalOptions": {
      "BoundedFiles": number,
      "BoundedSize": number,
      "EnableSamplePath": boolean,
      "SamplePath": "string"
    },
    "CompressionType": "string",
    "Exclusions": [ "string" ],
    "GroupFiles": "string",
    "GroupSize": "string",
    "MaxBand": number,
    "MaxFilesInBand": number,
    "Name": "string",
```

```
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
    "Paths": [ "string" ],
    "Recurse": boolean
  },
  "SelectFields": {
    "Inputs": [ "string" ],
    "Name": "string",
    "Paths": [
      [ "string" ]
    ]
  },
  "SelectFromCollection": {
    "Index": number,
    "Inputs": [ "string" ],
    "Name": "string"
  },
  "SnowflakeSource": {
    "Data": {
      "Action": "string",
      "AdditionalOptions": {
        "string" : "string"
      },
      "AutoPushdown": boolean,
      "Connection": {
        "Description": "string",
        "Label": "string",
        "Value": "string"
      },
      "Database": "string",
      "IamRole": {
        "Description": "string",
        "Label": "string",
        "Value": "string"
      },
      "MergeAction": "string",
```



```
"MergeClause": "string",
"MergeWhenMatched": "string",
"MergeWhenNotMatched": "string",
"PostAction": "string",
"PreAction": "string",
"SampleQuery": "string",
"Schema": "string",
"SelectedColumns": [
  {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  }
],
"SourceType": "string",
"StagingTable": "string",
"Table": "string",
"TableSchema": [
  {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  }
],
"TempDir": "string",
"Upsert": boolean
},
"Name": "string",
"OutputSchemas": [
  {
    "Columns": [
      {
        "Name": "string",
        "Type": "string"
      }
    ]
  }
]
},
"SnowflakeTarget": {
  "Data": {
    "Action": "string",
    "AdditionalOptions": {
      "string" : "string"
    }
  }
}
```

```
    },
    "AutoPushdown": boolean,
    "Connection": {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    },
    "Database": "string",
    "IamRole": {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    },
    },
    "MergeAction": "string",
    "MergeClause": "string",
    "MergeWhenMatched": "string",
    "MergeWhenNotMatched": "string",
    "PostAction": "string",
    "PreAction": "string",
    "SampleQuery": "string",
    "Schema": "string",
    "SelectedColumns": [
      {
        "Description": "string",
        "Label": "string",
        "Value": "string"
      }
    ],
    "SourceType": "string",
    "StagingTable": "string",
    "Table": "string",
    "TableSchema": [
      {
        "Description": "string",
        "Label": "string",
        "Value": "string"
      }
    ],
    "TempDir": "string",
    "Upsert": boolean
  },
  "Inputs": [ "string ],
  "Name": "string"
},
```

```
"SparkConnectorSource": {
  "AdditionalOptions": {
    "string" : "string"
  },
  "ConnectionName": "string",
  "ConnectionType": "string",
  "ConnectorName": "string",
  "Name": "string",
  "OutputSchemas": [
    {
      "Columns": [
        {
          "Name": "string",
          "Type": "string"
        }
      ]
    }
  ]
},
"SparkConnectorTarget": {
  "AdditionalOptions": {
    "string" : "string"
  },
  "ConnectionName": "string",
  "ConnectionType": "string",
  "ConnectorName": "string",
  "Inputs": [ "string" ],
  "Name": "string",
  "OutputSchemas": [
    {
      "Columns": [
        {
          "Name": "string",
          "Type": "string"
        }
      ]
    }
  ]
},
"SparkSQL": {
  "Inputs": [ "string" ],
  "Name": "string",
  "OutputSchemas": [
    {
```

```
        "Columns": [
            {
                "Name": "string",
                "Type": "string"
            }
        ]
    },
    ],
    "SqlAliases": [
        {
            "Alias": "string",
            "From": "string"
        }
    ],
    "SqlQuery": "string"
},
"Spigot": {
    "Inputs": [ "string" ],
    "Name": "string",
    "Path": "string",
    "Prob": number,
    "Topk": number
},
"SplitFields": {
    "Inputs": [ "string" ],
    "Name": "string",
    "Paths": [
        [ "string" ]
    ]
},
"Union": {
    "Inputs": [ "string" ],
    "Name": "string",
    "UnionType": "string"
}
}
},
"Command": {
    "Name": "string",
    "PythonVersion": "string",
    "Runtime": "string",
    "ScriptLocation": "string"
},
"Connections": {
```

```
    "Connections": [ "string" ]
  },
  "CreatedOn": number,
  "DefaultArguments": {
    "string" : "string"
  },
  "Description": "string",
  "ExecutionClass": "string",
  "ExecutionProperty": {
    "MaxConcurrentRuns": number
  },
  "GlueVersion": "string",
  "JobMode": "string",
  "LastModifiedOn": number,
  "LogUri": "string",
  "MaintenanceWindow": "string",
  "MaxCapacity": number,
  "MaxRetries": number,
  "Name": "string",
  "NonOverridableArguments": {
    "string" : "string"
  },
  "NotificationProperty": {
    "NotifyDelayAfter": number
  },
  "NumberOfWorkers": number,
  "ProfileName": "string",
  "Role": "string",
  "SecurityConfiguration": "string",
  "SourceControlDetails": {
    "AuthStrategy": "string",
    "AuthToken": "string",
    "Branch": "string",
    "Folder": "string",
    "LastCommitId": "string",
    "Owner": "string",
    "Provider": "string",
    "Repository": "string"
  },
  "Timeout": number,
  "WorkerType": "string"
}
],
"NextToken": "string"
```

```
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Jobs

A list of job definitions.

Type: Array of [Job](#) objects

NextToken

A continuation token, if not all job definitions have yet been returned.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetMapping

Creates mappings.

Request Syntax

```
{
  "Location": [
    {
      "DynamoDB": [
        {
          "Name": "string",
          "Param": boolean,
          "Value": "string"
        }
      ],
      "Jdbc": [
        {
          "Name": "string",
          "Param": boolean,
          "Value": "string"
        }
      ],
      "S3": [
        {
          "Name": "string",
          "Param": boolean,
          "Value": "string"
        }
      ]
    },
    {
      "Sinks": [
        {
          "DatabaseName": "string",
          "TableName": "string"
        }
      ],
      "Source": {
        "DatabaseName": "string",
        "TableName": "string"
      }
    }
  ]
}
```


Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Location

Parameters for the mapping.

Type: [Location](#) object

Required: No

Sinks

A list of target tables.

Type: Array of [CatalogEntry](#) objects

Required: No

Source

Specifies the source table.

Type: [CatalogEntry](#) object

Required: Yes

Response Syntax

```
{
  "Mapping": [
    {
      "SourcePath": "string",
      "SourceTable": "string",
      "SourceType": "string",
      "TargetPath": "string",
      "TargetTable": "string",
      "TargetType": "string"
    }
  ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Mapping

A list of mappings to the specified targets.

Type: Array of [MappingEntry](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetMLTaskRun

Gets details for a specific task run on a machine learning transform. Machine learning task runs are asynchronous tasks that AWS Glue runs on your behalf as part of various machine learning workflows. You can check the stats of any task run by calling `GetMLTaskRun` with the `TaskRunID` and its parent transform's `TransformID`.

Request Syntax

```
{
  "TaskRunId": "string",
  "TransformId": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

TaskRunId

The unique identifier of the task run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

TransformId

The unique identifier of the machine learning transform.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "CompletedOn": number,
  "ErrorString": "string",
  "ExecutionTime": number,
  "LastModifiedOn": number,
  "LogGroupName": "string",
  "Properties": {
    "ExportLabelsTaskRunProperties": {
      "OutputS3Path": "string"
    },
    "FindMatchesTaskRunProperties": {
      "JobId": "string",
      "JobName": "string",
      "JobRunId": "string"
    },
    "ImportLabelsTaskRunProperties": {
      "InputS3Path": "string",
      "Replace": boolean
    },
    "LabelingSetGenerationTaskRunProperties": {
      "OutputS3Path": "string"
    },
    "TaskType": "string"
  },
  "StartedOn": number,
  "Status": "string",
  "TaskRunId": "string",
  "TransformId": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

CompletedOn

The date and time when this task run was completed.

Type: Timestamp

ErrorString

The error strings that are associated with the task run.

Type: String

ExecutionTime

The amount of time (in seconds) that the task run consumed resources.

Type: Integer

LastModifiedOn

The date and time when this task run was last modified.

Type: Timestamp

LogGroupName

The names of the log groups that are associated with the task run.

Type: String

Properties

The list of properties that are associated with the task run.

Type: [TaskRunProperties](#) object

StartedOn

The date and time when this task run started.

Type: Timestamp

Status

The status for this task run.

Type: String

Valid Values: STARTING | RUNNING | STOPPING | STOPPED | SUCCEEDED | FAILED | TIMEOUT

TaskRunId

The unique run identifier associated with this run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

TransformId

The unique identifier of the task run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetMLTaskRuns

Gets a list of runs for a machine learning transform. Machine learning task runs are asynchronous tasks that AWS Glue runs on your behalf as part of various machine learning workflows. You can get a sortable, filterable list of machine learning task runs by calling `GetMLTaskRuns` with their parent transform's `TransformID` and other optional parameters as documented in this section.

This operation returns a list of historic runs and must be paginated.

Request Syntax

```
{
  "Filter": {
    "StartedAfter": number,
    "StartedBefore": number,
    "Status": "string",
    "TaskRunType": "string"
  },
  "MaxResults": number,
  "NextToken": "string",
  "Sort": {
    "Column": "string",
    "SortDirection": "string"
  },
  "TransformId": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Filter

The filter criteria, in the `TaskRunFilterCriteria` structure, for the task run.

Type: [TaskRunFilterCriteria](#) object

Required: No


```
{
  "CompletedOn": number,
  "ErrorString": "string",
  "ExecutionTime": number,
  "LastModifiedOn": number,
  "LogGroupName": "string",
  "Properties": {
    "ExportLabelsTaskRunProperties": {
      "OutputS3Path": "string"
    },
    "FindMatchesTaskRunProperties": {
      "JobId": "string",
      "JobName": "string",
      "JobRunId": "string"
    },
    "ImportLabelsTaskRunProperties": {
      "InputS3Path": "string",
      "Replace": boolean
    },
    "LabelingSetGenerationTaskRunProperties": {
      "OutputS3Path": "string"
    },
    "TaskType": "string"
  },
  "StartedOn": number,
  "Status": "string",
  "TaskRunId": "string",
  "TransformId": "string"
}
]
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken

A pagination token, if more results are available.

Type: String

[TaskRuns](#)

A list of task runs that are associated with the transform.

Type: Array of [TaskRun](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)

- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetMLTransform

Gets an AWS Glue machine learning transform artifact and all its corresponding metadata. Machine learning transforms are a special type of transform that use machine learning to learn the details of the transformation to be performed by learning from examples provided by humans. These transformations are then saved by AWS Glue. You can retrieve their metadata by calling `GetMLTransform`.

Request Syntax

```
{
  "TransformId": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

TransformId

The unique identifier of the transform, generated at the time that the transform was created.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "CreatedOn": number,
  "Description": "string",
  "EvaluationMetrics": {
    "FindMatchesMetrics": {
      "AreaUnderPRCurve": number,
      "ColumnImportances": [
```

```

    {
      "ColumnName": "string",
      "Importance": number
    }
  ],
  "ConfusionMatrix": {
    "NumFalseNegatives": number,
    "NumFalsePositives": number,
    "NumTrueNegatives": number,
    "NumTruePositives": number
  },
  "F1": number,
  "Precision": number,
  "Recall": number
},
"TransformType": "string"
},
"GlueVersion": "string",
"InputRecordTables": [
  {
    "AdditionalOptions": {
      "string" : "string"
    },
    "CatalogId": "string",
    "ConnectionName": "string",
    "DatabaseName": "string",
    "TableName": "string"
  }
],
"LabelCount": number,
"LastModifiedOn": number,
"MaxCapacity": number,
"MaxRetries": number,
"Name": "string",
"NumberOfWorkers": number,
"Parameters": {
  "FindMatchesParameters": {
    "AccuracyCostTradeoff": number,
    "EnforceProvidedLabels": boolean,
    "PrecisionRecallTradeoff": number,
    "PrimaryKeyColumnName": "string"
  },
  "TransformType": "string"
},
},

```

```
"Role": "string",
"Schema": [
  {
    "DataType": "string",
    "Name": "string"
  }
],
"Status": "string",
"Timeout": number,
"TransformEncryption": {
  "MLUserDataEncryption": {
    "KmsKeyId": "string",
    "MLUserDataEncryptionMode": "string"
  },
  "TaskRunSecurityConfigurationName": "string"
},
"TransformId": "string",
"WorkerType": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

CreatedOn

The date and time when the transform was created.

Type: Timestamp

Description

A description of the transform.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

EvaluationMetrics

The latest evaluation metrics.

Type: [EvaluationMetrics](#) object

GlueVersion

This value determines which version of AWS Glue this machine learning transform is compatible with. Glue 1.0 is recommended for most customers. If the value is not set, the Glue compatibility defaults to Glue 0.9. For more information, see [AWS Glue Versions](#) in the developer guide.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `^\w+\.\w+$`

InputRecordTables

A list of AWS Glue table definitions used by the transform.

Type: Array of [GlueTable](#) objects

Array Members: Minimum number of 0 items. Maximum number of 10 items.

LabelCount

The number of labels available for this transform.

Type: Integer

LastModifiedOn

The date and time when the transform was last modified.

Type: Timestamp

MaxCapacity

The number of AWS Glue data processing units (DPUs) that are allocated to task runs for this transform. You can allocate from 2 to 100 DPUs; the default is 10. A DPU is a relative measure of processing power that consists of 4 vCPUs of compute capacity and 16 GB of memory. For more information, see the [AWS Glue pricing page](#).

When the `WorkerType` field is set to a value other than `Standard`, the `MaxCapacity` field is set automatically and becomes read-only.

Type: Double

MaxRetries

The maximum number of times to retry a task for this transform after a task run fails.

Type: Integer

Name

The unique name given to the transform when it was created.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

NumberOfWorkers

The number of workers of a defined `workerType` that are allocated when this task runs.

Type: Integer

Parameters

The configuration parameters that are specific to the algorithm used.

Type: [TransformParameters](#) object

Role

The name or Amazon Resource Name (ARN) of the IAM role with the required permissions.

Type: String

Schema

The `Map<Column, Type>` object that represents the schema that this transform accepts. Has an upper bound of 100 columns.

Type: Array of [SchemaColumn](#) objects

Array Members: Maximum number of 100 items.

Status

The last known status of the transform (to indicate whether it can be used or not). One of "NOT_READY", "READY", or "DELETING".

Type: String

Valid Values: NOT_READY | READY | DELETING

Timeout

The timeout for a task run for this transform in minutes. This is the maximum time that a task run for this transform can consume resources before it is terminated and enters TIMEOUT status. The default is 2,880 minutes (48 hours).

Type: Integer

Valid Range: Minimum value of 1.

TransformEncryption

The encryption-at-rest settings of the transform that apply to accessing user data. Machine learning transforms can access user data encrypted in Amazon S3 using KMS.

Type: [TransformEncryption](#) object

TransformId

The unique identifier of the transform, generated at the time that the transform was created.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\u007F\u00E0-\u00FF\u0080-\u00FF\u00DC-\u00BF\u00DF\u00t]*`

WorkerType

The type of predefined worker that is allocated when this task runs. Accepts a value of Standard, G.1X, or G.2X.

- For the Standard worker type, each worker provides 4 vCPU, 16 GB of memory and a 50GB disk, and 2 executors per worker.
- For the G.1X worker type, each worker provides 4 vCPU, 16 GB of memory and a 64GB disk, and 1 executor per worker.
- For the G.2X worker type, each worker provides 8 vCPU, 32 GB of memory and a 128GB disk, and 1 executor per worker.

Type: String

Valid Values: Standard | G.1X | G.2X | G.025X | G.4X | G.8X | Z.2X

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)

- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetMLTransforms

Gets a sortable, filterable list of existing AWS Glue machine learning transforms. Machine learning transforms are a special type of transform that use machine learning to learn the details of the transformation to be performed by learning from examples provided by humans. These transformations are then saved by AWS Glue, and you can retrieve their metadata by calling `GetMLTransforms`.

Request Syntax

```
{
  "Filter": {
    "CreatedAfter": number,
    "CreatedBefore": number,
    "GlueVersion": "string",
    "LastModifiedAfter": number,
    "LastModifiedBefore": number,
    "Name": "string",
    "Schema": [
      {
        "DataType": "string",
        "Name": "string"
      }
    ],
    "Status": "string",
    "TransformType": "string"
  },
  "MaxResults": number,
  "NextToken": "string",
  "Sort": {
    "Column": "string",
    "SortDirection": "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Filter

The filter transformation criteria.

Type: [TransformFilterCriteria](#) object

Required: No

MaxResults

The maximum number of results to return.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

NextToken

A paginated token to offset the results.

Type: String

Required: No

Sort

The sorting criteria.

Type: [TransformSortCriteria](#) object

Required: No

Response Syntax

```
{
  "NextToken": "string",
  "Transforms": [
    {
      "CreatedOn": number,
      "Description": "string",
      "EvaluationMetrics": {
        "FindMatchesMetrics": {
          "AreaUnderPRCurve": number,
          "ColumnImportances": [
```

```
    {
      "ColumnName": "string",
      "Importance": number
    }
  ],
  "ConfusionMatrix": {
    "NumFalseNegatives": number,
    "NumFalsePositives": number,
    "NumTrueNegatives": number,
    "NumTruePositives": number
  },
  "F1": number,
  "Precision": number,
  "Recall": number
},
"TransformType": "string"
},
"GlueVersion": "string",
"InputRecordTables": [
  {
    "AdditionalOptions": {
      "string": "string"
    },
    "CatalogId": "string",
    "ConnectionName": "string",
    "DatabaseName": "string",
    "TableName": "string"
  }
],
"LabelCount": number,
"LastModifiedOn": number,
"MaxCapacity": number,
"MaxRetries": number,
"Name": "string",
"NumberOfWorkers": number,
"Parameters": {
  "FindMatchesParameters": {
    "AccuracyCostTradeoff": number,
    "EnforceProvidedLabels": boolean,
    "PrecisionRecallTradeoff": number,
    "PrimaryKeyColumnName": "string"
  },
  "TransformType": "string"
},
},
```



```
"Role": "string",
"Schema": [
  {
    "DataType": "string",
    "Name": "string"
  }
],
"Status": "string",
"Timeout": number,
"TransformEncryption": {
  "MLUserDataEncryption": {
    "KmsKeyId": "string",
    "MLUserDataEncryptionMode": "string"
  },
  "TaskRunSecurityConfigurationName": "string"
},
"TransformId": "string",
"WorkerType": "string"
}
]
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken

A pagination token, if more results are available.

Type: String

Transforms

A list of machine learning transforms.

Type: Array of [MLTransform](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetPartition

Retrieves information about a specified partition.

Request Syntax

```
{  
  "CatalogId": "string",  
  "DatabaseName": "string",  
  "PartitionValues": [ "string" ],  
  "TableName": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog where the partition in question resides. If none is provided, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

DatabaseName

The name of the catalog database where the partition resides.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

PartitionValues

The values that define the partition.

Type: Array of strings

Length Constraints: Maximum length of 1024.

Required: Yes

TableName

The name of the partition's table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "Partition": {
    "CatalogId": "string",
    "CreationTime": number,
    "DatabaseName": "string",
    "LastAccessTime": number,
    "LastAnalyzedTime": number,
    "Parameters": {
      "string" : "string"
    },
    "StorageDescriptor": {
      "AdditionalLocations": [ "string" ],
      "BucketColumns": [ "string" ],
      "Columns": [
        {
          "Comment": "string",
          "Name": "string",
          "Parameters": {
            "string" : "string"
          }
        }
      ],
    },
  },
}
```

```
    "Type": "string"
  }
],
"Compressed": boolean,
"InputFormat": "string",
"Location": "string",
"NumberOfBuckets": number,
"OutputFormat": "string",
"Parameters": {
  "string" : "string"
},
"SchemaReference": {
  "SchemaId": {
    "RegistryName": "string",
    "SchemaArn": "string",
    "SchemaName": "string"
  },
  "SchemaVersionId": "string",
  "SchemaVersionNumber": number
},
"SerdeInfo": {
  "Name": "string",
  "Parameters": {
    "string" : "string"
  },
  "SerializationLibrary": "string"
},
"SkewedInfo": {
  "SkewedColumnNames": [ "string" ],
  "SkewedColumnValueLocationMaps": {
    "string" : "string"
  },
  "SkewedColumnValues": [ "string" ]
},
"SortColumns": [
  {
    "Column": "string",
    "SortOrder": number
  }
],
"StoredAsSubDirectories": boolean
},
"TableName": "string",
"Values": [ "string" ]
```

```
}  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Partition

The requested information, in the form of a `Partition` object.

Type: [Partition](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

FederationSourceException

A federation source failed.

HTTP Status Code: 400

FederationSourceRetryableException

A federation source failed, but the operation may be retried.

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetPartitionIndexes

Retrieves the partition indexes associated with a table.

Request Syntax

```
{  
  "CatalogId": "string",  
  "DatabaseName": "string",  
  "NextToken": "string",  
  "TableName": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The catalog ID where the table resides.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

DatabaseName

Specifies the name of a database from which you want to retrieve partition indexes.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

NextToken

A continuation token, included if this is a continuation call.

Type: String

Required: No

TableName

Specifies the name of a table for which you want to retrieve the partition indexes.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "NextToken": "string",
  "PartitionIndexDescriptorList": [
    {
      "BackfillErrors": [
        {
          "Code": "string",
          "Partitions": [
            {
              "Values": [ "string" ]
            }
          ]
        }
      ],
      "IndexName": "string",
      "IndexStatus": "string",
      "Keys": [
        {
          "Name": "string",
          "Type": "string"
        }
      ]
    }
  ]
}
```

```
    ]  
  }  
]  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken

A continuation token, present if the current list segment is not the last.

Type: String

PartitionIndexDescriptorList

A list of index descriptors.

Type: Array of [PartitionIndexDescriptor](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ConflictException

The `CreatePartitions` API was called on a table that has indexes enabled.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetPartitions

Retrieves information about the partitions in a table.

Request Syntax

```
{
  "CatalogId": "string",
  "DatabaseName": "string",
  "ExcludeColumnSchema": boolean,
  "Expression": "string",
  "MaxResults": number,
  "NextToken": "string",
  "QueryAsOfTime": number,
  "Segment": {
    "SegmentNumber": number,
    "TotalSegments": number
  },
  "TableName": "string",
  "TransactionId": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog where the partitions in question reside. If none is provided, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\u007F\u00E0-\u00FF\u0100-\u017F\u0180-\u01FF\u0200-\u02FF\u0300-\u037F\u0380-\u03FF\u0400-\u047F\u0480-\u04FF\u0500-\u057F\u0580-\u05FF\u0600-\u06FF\u0700-\u07FF\u0800-\u08FF\u0900-\u097F\u0980-\u09FF\u0A00-\u0A7F\u0A80-\u0AFF\u0B00-\u0B7F\u0B80-\u0BFF\u0C00-\u0C7F\u0C80-\u0CFF\u0D00-\u0D7F\u0D80-\u0DBF\u0DC0-\u0DBF\u0E00-\u0E7F\u0E80-\u0EFF\u0F00-\u0F7F\u0F80-\u0FFF\u1000-\u107F\u1080-\u10FF\u1100-\u117F\u1180-\u11FF\u1200-\u127F\u1280-\u12FF\u1300-\u137F\u1380-\u13FF\u1400-\u147F\u1480-\u14FF\u1500-\u157F\u1580-\u15FF\u1600-\u167F\u1680-\u16FF\u1700-\u177F\u1780-\u17FF\u1800-\u187F\u1880-\u18FF\u1900-\u197F\u1980-\u19FF\u1A00-\u1A7F\u1A80-\u1AFF\u1B00-\u1B7F\u1B80-\u1BFF\u1C00-\u1C7F\u1C80-\u1CFF\u1D00-\u1D7F\u1D80-\u1DBF\u1E00-\u1E7F\u1E80-\u1EFF\u1F00-\u1F7F\u1F80-\u1FFF\u2000-\u207F\u2080-\u20FF\u2100-\u217F\u2180-\u21FF\u2200-\u227F\u2280-\u22FF\u2300-\u237F\u2380-\u23FF\u2400-\u247F\u2480-\u24FF\u2500-\u257F\u2580-\u25FF\u2600-\u267F\u2680-\u26FF\u2700-\u277F\u2780-\u27FF\u2800-\u287F\u2880-\u28FF\u2900-\u297F\u2980-\u29FF\u2A00-\u2A7F\u2A80-\u2AFF\u2B00-\u2B7F\u2B80-\u2BFF\u2C00-\u2C7F\u2C80-\u2CFF\u2D00-\u2D7F\u2D80-\u2DBF\u2E00-\u2E7F\u2E80-\u2EFF\u2F00-\u2F7F\u2F80-\u2FFF\u3000-\u307F\u3080-\u30FF\u3100-\u317F\u3180-\u31FF\u3200-\u327F\u3280-\u32FF\u3300-\u337F\u3380-\u33FF\u3400-\u347F\u3480-\u34FF\u3500-\u357F\u3580-\u35FF\u3600-\u367F\u3680-\u36FF\u3700-\u377F\u3780-\u37FF\u3800-\u387F\u3880-\u38FF\u3900-\u397F\u3980-\u39FF\u3A00-\u3A7F\u3A80-\u3AFF\u3B00-\u3B7F\u3B80-\u3BFF\u3C00-\u3C7F\u3C80-\u3CFF\u3D00-\u3D7F\u3D80-\u3DBF\u3E00-\u3E7F\u3E80-\u3EFF\u3F00-\u3F7F\u3F80-\u3FFF\u4000-\u407F\u4080-\u40FF\u4100-\u417F\u4180-\u41FF\u4200-\u427F\u4280-\u42FF\u4300-\u437F\u4380-\u43FF\u4400-\u447F\u4480-\u44FF\u4500-\u457F\u4580-\u45FF\u4600-\u467F\u4680-\u46FF\u4700-\u477F\u4780-\u47FF\u4800-\u487F\u4880-\u48FF\u4900-\u497F\u4980-\u49FF\u4A00-\u4A7F\u4A80-\u4AFF\u4B00-\u4B7F\u4B80-\u4BFF\u4C00-\u4C7F\u4C80-\u4CFF\u4D00-\u4D7F\u4D80-\u4DBF\u4E00-\u4E7F\u4E80-\u4EFF\u4F00-\u4F7F\u4F80-\u4FFF\u5000-\u507F\u5080-\u50FF\u5100-\u517F\u5180-\u51FF\u5200-\u527F\u5280-\u52FF\u5300-\u537F\u5380-\u53FF\u5400-\u547F\u5480-\u54FF\u5500-\u557F\u5580-\u55FF\u5600-\u567F\u5680-\u56FF\u5700-\u577F\u5780-\u57FF\u5800-\u587F\u5880-\u58FF\u5900-\u597F\u5980-\u59FF\u5A00-\u5A7F\u5A80-\u5AFF\u5B00-\u5B7F\u5B80-\u5BFF\u5C00-\u5C7F\u5C80-\u5CFF\u5D00-\u5D7F\u5D80-\u5DBF\u5E00-\u5E7F\u5E80-\u5EFF\u5F00-\u5F7F\u5F80-\u5FFF\u6000-\u607F\u6080-\u60FF\u6100-\u617F\u6180-\u61FF\u6200-\u627F\u6280-\u62FF\u6300-\u637F\u6380-\u63FF\u6400-\u647F\u6480-\u64FF\u6500-\u657F\u6580-\u65FF\u6600-\u667F\u6680-\u66FF\u6700-\u677F\u6780-\u67FF\u6800-\u687F\u6880-\u68FF\u6900-\u697F\u6980-\u69FF\u6A00-\u6A7F\u6A80-\u6AFF\u6B00-\u6B7F\u6B80-\u6BFF\u6C00-\u6C7F\u6C80-\u6CFF\u6D00-\u6D7F\u6D80-\u6DBF\u6E00-\u6E7F\u6E80-\u6EFF\u6F00-\u6F7F\u6F80-\u6FFF\u7000-\u707F\u7080-\u70FF\u7100-\u717F\u7180-\u71FF\u7200-\u727F\u7280-\u72FF\u7300-\u737F\u7380-\u73FF\u7400-\u747F\u7480-\u74FF\u7500-\u757F\u7580-\u75FF\u7600-\u767F\u7680-\u76FF\u7700-\u777F\u7780-\u77FF\u7800-\u787F\u7880-\u78FF\u7900-\u797F\u7980-\u79FF\u7A00-\u7A7F\u7A80-\u7AFF\u7B00-\u7B7F\u7B80-\u7BFF\u7C00-\u7C7F\u7C80-\u7CFF\u7D00-\u7D7F\u7D80-\u7DBF\u7E00-\u7E7F\u7E80-\u7EFF\u7F00-\u7F7F\u7F80-\u7FFF\u8000-\u807F\u8080-\u80FF\u8100-\u817F\u8180-\u81FF\u8200-\u827F\u8280-\u82FF\u8300-\u837F\u8380-\u83FF\u8400-\u847F\u8480-\u84FF\u8500-\u857F\u8580-\u85FF\u8600-\u867F\u8680-\u86FF\u8700-\u877F\u8780-\u87FF\u8800-\u887F\u8880-\u88FF\u8900-\u897F\u8980-\u89FF\u8A00-\u8A7F\u8A80-\u8AFF\u8B00-\u8B7F\u8B80-\u8BFF\u8C00-\u8C7F\u8C80-\u8CFF\u8D00-\u8D7F\u8D80-\u8DBF\u8E00-\u8E7F\u8E80-\u8EFF\u8F00-\u8F7F\u8F80-\u8FFF\u9000-\u907F\u9080-\u90FF\u9100-\u917F\u9180-\u91FF\u9200-\u927F\u9280-\u92FF\u9300-\u937F\u9380-\u93FF\u9400-\u947F\u9480-\u94FF\u9500-\u957F\u9580-\u95FF\u9600-\u967F\u9680-\u96FF\u9700-\u977F\u9780-\u97FF\u9800-\u987F\u9880-\u98FF\u9900-\u997F\u9980-\u99FF\u9A00-\u9A7F\u9A80-\u9AFF\u9B00-\u9B7F\u9B80-\u9BFF\u9C00-\u9C7F\u9C80-\u9CFF\u9D00-\u9D7F\u9D80-\u9DBF\u9E00-\u9E7F\u9E80-\u9EFF\u9F00-\u9F7F\u9F80-\u9FFF\uA000-\uA07F\uA080-\uA0FF\uA100-\uA17F\uA180-\uA1FF\uA200-\uA27F\uA280-\uA2FF\uA300-\uA37F\uA380-\uA3FF\uA400-\uA47F\uA480-\uA4FF\uA500-\uA57F\uA580-\uA5FF\uA600-\uA67F\uA680-\uA6FF\uA700-\uA77F\uA780-\uA7FF\uA800-\uA87F\uA880-\uA8FF\uA900-\uA97F\uA980-\uA9FF\uAA00-\uAA7F\uAA80-\uAAFF\uAB00-\uAB7F\uAB80-\uABFF\uAC00-\uAC7F\uAC80-\uACFF\uAD00-\uAD7F\uAD80-\uADBFF\uAE00-\uAE7F\uAE80-\uAEFF\uAF00-\uAF7F\uAF80-\uAFFF\uB000-\uB07F\uB080-\uB0FF\uB100-\uB17F\uB180-\uB1FF\uB200-\uB27F\uB280-\uB2FF\uB300-\uB37F\uB380-\uB3FF\uB400-\uB47F\uB480-\uB4FF\uB500-\uB57F\uB580-\uB5FF\uB600-\uB67F\uB680-\uB6FF\uB700-\uB77F\uB780-\uB7FF\uB800-\uB87F\uB880-\uB8FF\uB900-\uB97F\uB980-\uB9FF\uBA00-\uBA7F\uBA80-\uBAFF\uBB00-\uBB7F\uBB80-\uBBFF\uBC00-\uBC7F\uBC80-\uBCFF\uBD00-\uBD7F\uBD80-\uBDBF\uBE00-\uBE7F\uBE80-\uBEFF\uBF00-\uBF7F\uBF80-\uBFFF\uC000-\uC07F\uC080-\uC0FF\uC100-\uC17F\uC180-\uC1FF\uC200-\uC27F\uC280-\uC2FF\uC300-\uC37F\uC380-\uC3FF\uC400-\uC47F\uC480-\uC4FF\uC500-\uC57F\uC580-\uC5FF\uC600-\uC67F\uC680-\uC6FF\uC700-\uC77F\uC780-\uC7FF\uC800-\uC87F\uC880-\uC8FF\uC900-\uC97F\uC980-\uC9FF\uCA00-\uCA7F\uCA80-\uCAFF\uCB00-\uCB7F\uCB80-\uCBFF\uCC00-\uCC7F\uCC80-\uCCFF\uCD00-\uCD7F\uCD80-\uCDFF\uCE00-\uCE7F\uCE80-\uCEFF\uCF00-\uCF7F\uCF80-\uCFFF\uD000-\uD07F\uD080-\uD0FF\uD100-\uD17F\uD180-\uD1FF\uD200-\uD27F\uD280-\uD2FF\uD300-\uD37F\uD380-\uD3FF\uD400-\uD47F\uD480-\uD4FF\uD500-\uD57F\uD580-\uD5FF\uD600-\uD67F\uD680-\uD6FF\uD700-\uDBFF\uD780-\uDBFF\uD800-\uDBFF\uD880-\uDBFF\uD900-\uDBFF\uD980-\uDBFF\uDA00-\uDBFF\uDA80-\uDBFF\uDB00-\uDBFF\uDB80-\uDBFF\uDC00-\uDBFF\uDC80-\uDBFF\uDD00-\uDBFF\uDD80-\uDBFF\uDE00-\uDBFF\uDE80-\uDBFF\uDF00-\uDBFF\uDF80-\uDBFF\uE000-\uE07F\uE080-\uE0FF\uE100-\uE17F\uE180-\uE1FF\uE200-\uE27F\uE280-\uE2FF\uE300-\uE37F\uE380-\uE3FF\uE400-\uE47F\uE480-\uE4FF\uE500-\uE57F\uE580-\uE5FF\uE600-\uE67F\uE680-\uE6FF\uE700-\uE77F\uE780-\uE7FF\uE800-\uE87F\uE880-\uE8FF\uE900-\uE97F\uE980-\uE9FF\uEA00-\uEA7F\uEA80-\uEAFF\uEB00-\uEB7F\uEB80-\uEBFF\uEC00-\uEC7F\uEC80-\uECFF\uED00-\uED7F\uED80-\uEDFF\uEE00-\uEE7F\uEE80-\uEEFF\uEF00-\uEF7F\uEF80-\uEFFF\uF000-\uF07F\uF080-\uF0FF\uF100-\uF17F\uF180-\uF1FF\uF200-\uF27F\uF280-\uF2FF\uF300-\uF37F\uF380-\uF3FF\uF400-\uF47F\uF480-\uF4FF\uF500-\uF57F\uF580-\uF5FF\uF600-\uF67F\uF680-\uF6FF\uF700-\uF77F\uF780-\uF7FF\uF800-\uF87F\uF880-\uF8FF\uF900-\uF97F\uF980-\uF9FF\uFA00-\uFA7F\uFA80-\uFAFF\uFB00-\uFB7F\uFB80-\uFBFF\uFC00-\uFC7F\uFC80-\uFCFF\uFD00-\uFD7F\uFD80-\uFDFF\uFE00-\uFE7F\uFE80-\uFEFF\uFF00-\uFF7F\uFF80-\uFFFF`

Required: No

DatabaseName

The name of the catalog database where the partitions reside.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

ExcludeColumnSchema

When true, specifies not returning the partition column schema. Useful when you are interested only in other partition attributes such as partition values or location. This approach avoids the problem of a large response by not returning duplicate data.

Type: Boolean

Required: No

Expression

An expression that filters the partitions to be returned.

The expression uses SQL syntax similar to the SQL WHERE filter clause. The SQL statement parser [JSQLParser](#) parses the expression.

Operators: The following are the operators that you can use in the Expression API call:

=

Checks whether the values of the two operands are equal; if yes, then the condition becomes true.

Example: Assume 'variable a' holds 10 and 'variable b' holds 20.

(a = b) is not true.

< >

Checks whether the values of two operands are equal; if the values are not equal, then the condition becomes true.

Example: (a < > b) is true.

>

Checks whether the value of the left operand is greater than the value of the right operand; if yes, then the condition becomes true.

Example: (a > b) is not true.

<

Checks whether the value of the left operand is less than the value of the right operand; if yes, then the condition becomes true.

Example: (a < b) is true.

>=

Checks whether the value of the left operand is greater than or equal to the value of the right operand; if yes, then the condition becomes true.

Example: (a >= b) is not true.

<=

Checks whether the value of the left operand is less than or equal to the value of the right operand; if yes, then the condition becomes true.

Example: (a <= b) is true.

AND, OR, IN, BETWEEN, LIKE, NOT, IS NULL

Logical operators.

Supported Partition Key Types: The following are the supported partition keys.

- string
- date
- timestamp
- int
- bigint
- long
- tinyint
- smallint
- decimal

If an type is encountered that is not valid, an exception is thrown.

The following list shows the valid operators on each type. When you define a crawler, the `partitionKey` type is created as a `STRING`, to be compatible with the catalog partitions.

Sample API Call:

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

MaxResults

The maximum number of partitions to return in a single response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

NextToken

A continuation token, if this is not the first call to retrieve these partitions.

Type: String

Required: No

QueryAsOfTime

The time as of when to read the partition contents. If not set, the most recent transaction commit time will be used. Cannot be specified along with `TransactionId`.

Type: Timestamp

Required: No

Segment

The segment of the table's partitions to scan in this request.

Type: [Segment](#) object

Required: No

TableName

The name of the partitions' table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

TransactionId

The transaction ID at which to read the partition contents.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\p{L}\p{N}\p{P}]*`

Required: No

Response Syntax

```
{
  "NextToken": "string",
  "Partitions": [
    {
      "CatalogId": "string",
      "CreationTime": number,
      "DatabaseName": "string",
      "LastAccessTime": number,
      "LastAnalyzedTime": number,
      "Parameters": {
        "string" : "string"
      },
      "StorageDescriptor": {
        "AdditionalLocations": [ "string" ],
        "BucketColumns": [ "string" ],
        "Columns": [
          {
            "Comment": "string",
            "Name": "string",
            "Parameters": {
              "string" : "string"
            }
          }
        ],
      },
    }
  ]
}
```



```

        "Type": "string"
    }
],
"Compressed": boolean,
"InputFormat": "string",
"Location": "string",
"NumberOfBuckets": number,
"OutputFormat": "string",
"Parameters": {
    "string" : "string"
},
"SchemaReference": {
    "SchemaId": {
        "RegistryName": "string",
        "SchemaArn": "string",
        "SchemaName": "string"
    },
    "SchemaVersionId": "string",
    "SchemaVersionNumber": number
},
"SerdeInfo": {
    "Name": "string",
    "Parameters": {
        "string" : "string"
    },
    "SerializationLibrary": "string"
},
"SkewedInfo": {
    "SkewedColumnNames": [ "string" ],
    "SkewedColumnValueLocationMaps": {
        "string" : "string"
    },
    "SkewedColumnValues": [ "string" ]
},
"SortColumns": [
    {
        "Column": "string",
        "SortOrder": number
    }
],
"StoredAsSubDirectories": boolean
},
"TableName": "string",
"Values": [ "string" ]

```

```
    }  
  ]  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken

A continuation token, if the returned list of partitions does not include the last one.

Type: String

Partitions

A list of requested partitions.

Type: Array of [Partition](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

FederationSourceException

A federation source failed.

HTTP Status Code: 400

FederationSourceRetryableException

A federation source failed, but the operation may be retried.

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

InvalidStateException

An error that indicates your data is in an invalid state.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNotReadyException

A resource was not ready for a transaction.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)

- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetPlan

Gets code to perform a specified mapping.

Request Syntax

```
{
  "AdditionalPlanOptionsMap": {
    "string" : "string"
  },
  "Language": "string",
  "Location": {
    "DynamoDB": [
      {
        "Name": "string",
        "Param": boolean,
        "Value": "string"
      }
    ],
    "Jdbc": [
      {
        "Name": "string",
        "Param": boolean,
        "Value": "string"
      }
    ],
    "S3": [
      {
        "Name": "string",
        "Param": boolean,
        "Value": "string"
      }
    ]
  },
  "Mapping": [
    {
      "SourcePath": "string",
      "SourceTable": "string",
      "SourceType": "string",
      "TargetPath": "string",
      "TargetTable": "string",
      "TargetType": "string"
    }
  ]
}
```

```
],
  "Sinks": [
    {
      "DatabaseName": "string",
      "TableName": "string"
    }
  ],
  "Source": {
    "DatabaseName": "string",
    "TableName": "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

[AdditionalPlanOptionsMap](#)

A map to hold additional optional key-value parameters.

Currently, these key-value pairs are supported:

- `inferSchema` — Specifies whether to set `inferSchema` to true or false for the default script generated by an AWS Glue job. For example, to set `inferSchema` to true, pass the following key value pair:

```
--additional-plan-options-map '{"inferSchema":"true"}
```

Type: String to string map

Required: No

[Language](#)

The programming language of the code to perform the mapping.

Type: String

Valid Values: PYTHON | SCALA

Required: No

Location

The parameters for the mapping.

Type: [Location](#) object

Required: No

Mapping

The list of mappings from a source table to target tables.

Type: Array of [MappingEntry](#) objects

Required: Yes

Sinks

The target tables.

Type: Array of [CatalogEntry](#) objects

Required: No

Source

The source table.

Type: [CatalogEntry](#) object

Required: Yes

Response Syntax

```
{  
  "PythonScript": "string",  
  "ScalaCode": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

PythonScript

A Python script to perform the mapping.

Type: String

ScalaCode

The Scala code to perform the mapping.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)

- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetRegistry

Describes the specified registry in detail.

Request Syntax

```
{
  "RegistryId": {
    "RegistryArn": "string",
    "RegistryName": "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

[RegistryId](#)

This is a wrapper structure that may contain the registry name and Amazon Resource Name (ARN).

Type: [RegistryId](#) object

Required: Yes

Response Syntax

```
{
  "CreatedTime": "string",
  "Description": "string",
  "RegistryArn": "string",
  "RegistryName": "string",
  "Status": "string",
  "UpdatedTime": "string"
}
```


Type: String

Valid Values: AVAILABLE | DELETING

UpdatedTime

The date and time the registry was updated.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)

- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetResourcePolicies

Retrieves the resource policies set on individual resources by AWS Resource Access Manager during cross-account permission grants. Also retrieves the Data Catalog resource policy.

If you enabled metadata encryption in Data Catalog settings, and you do not have permission on the AWS KMS key, the operation can't return the Data Catalog resource policy.

Request Syntax

```
{  
  "MaxResults": number,  
  "NextToken": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

MaxResults

The maximum size of a list to return.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

NextToken

A continuation token, if this is a continuation request.

Type: String

Required: No

Response Syntax

```
{
```

```
"GetResourcePoliciesResponseList": [  
  {  
    "CreateTime": number,  
    "PolicyHash": "string",  
    "PolicyInJson": "string",  
    "UpdateTime": number  
  }  
],  
"NextToken": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

GetResourcePoliciesResponseList

A list of the individual resource policies and the account-level resource policy.

Type: Array of [GluePolicy](#) objects

NextToken

A continuation token, if the returned list does not contain the last resource policy available.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetResourcePolicy

Retrieves a specified resource policy.

Request Syntax

```
{  
  "ResourceArn": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

ResourceArn

The ARN of the AWS Glue resource for which to retrieve the resource policy. If not supplied, the Data Catalog resource policy is returned. Use `GetResourcePolicies` to view all existing resource policies. For more information see [Specifying AWS Glue Resource ARNs](#).

Type: String

Length Constraints: Minimum length of 1. Maximum length of 10240.

Pattern: `arn:(aws|aws-us-gov|aws-cn):glue:.*`

Required: No

Response Syntax

```
{  
  "CreateTime": number,  
  "PolicyHash": "string",  
  "PolicyInJson": "string",  
  "UpdateTime": number  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

CreateTime

The date and time at which the policy was created.

Type: Timestamp

PolicyHash

Contains the hash value associated with this policy.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

PolicyInJson

Contains the requested policy document, in JSON format.

Type: String

Length Constraints: Minimum length of 2.

UpdateTime

The date and time at which the policy was last updated.

Type: Timestamp

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetSchema

Describes the specified schema in detail.

Request Syntax

```
{
  "SchemaId": {
    "RegistryName": "string",
    "SchemaArn": "string",
    "SchemaName": "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

SchemaId

This is a wrapper structure to contain schema identity fields. The structure contains:

- `SchemaId$SchemaArn`: The Amazon Resource Name (ARN) of the schema. Either `SchemaArn` or `SchemaName` and `RegistryName` has to be provided.
- `SchemaId$SchemaName`: The name of the schema. Either `SchemaArn` or `SchemaName` and `RegistryName` has to be provided.

Type: [SchemaId](#) object

Required: Yes

Response Syntax

```
{
  "Compatibility": "string",
  "CreatedTime": "string",
  "DataFormat": "string",
  "Description": "string",
```

```
"LatestSchemaVersion": number,
"NextSchemaVersion": number,
"RegistryArn": "string",
"RegistryName": "string",
"SchemaArn": "string",
"SchemaCheckpoint": number,
"SchemaName": "string",
"SchemaStatus": "string",
"UpdateTime": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Compatibility

The compatibility mode of the schema.

Type: String

Valid Values: NONE | DISABLED | BACKWARD | BACKWARD_ALL | FORWARD | FORWARD_ALL | FULL | FULL_ALL

CreatedTime

The date and time the schema was created.

Type: String

DataFormat

The data format of the schema definition. Currently AVRO, JSON and PROTOBUF are supported.

Type: String

Valid Values: AVRO | JSON | PROTOBUF

Description

A description of schema if specified when created

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

LatestSchemaVersion

The latest version of the schema associated with the returned schema definition.

Type: Long

Valid Range: Minimum value of 1. Maximum value of 100000.

NextSchemaVersion

The next version of the schema associated with the returned schema definition.

Type: Long

Valid Range: Minimum value of 1. Maximum value of 100000.

RegistryArn

The Amazon Resource Name (ARN) of the registry.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 10240.

Pattern: `arn:(aws|aws-us-gov|aws-cn):glue:.*`

RegistryName

The name of the registry.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[a-zA-Z0-9-_$#.]+`

SchemaArn

The Amazon Resource Name (ARN) of the schema.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 10240.

Pattern: `arn:(aws|aws-us-gov|aws-cn):glue:.*`

SchemaCheckpoint

The version number of the checkpoint (the last time the compatibility mode was changed).

Type: Long

Valid Range: Minimum value of 1. Maximum value of 100000.

SchemaName

The name of the schema.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[a-zA-Z0-9-_$#.]+`

SchemaStatus

The status of the schema.

Type: String

Valid Values: AVAILABLE | PENDING | DELETING

UpdateTime

The date and time the schema was updated.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetSchemaByDefinition

Retrieves a schema by the SchemaDefinition. The schema definition is sent to the Schema Registry, canonicalized, and hashed. If the hash is matched within the scope of the SchemaName or ARN (or the default registry, if none is supplied), that schema's metadata is returned. Otherwise, a 404 or NotFound error is returned. Schema versions in Deleted statuses will not be included in the results.

Request Syntax

```
{
  "SchemaDefinition": "string",
  "SchemaId": {
    "RegistryName": "string",
    "SchemaArn": "string",
    "SchemaName": "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

SchemaDefinition

The definition of the schema for which schema details are required.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 170000.

Pattern: .*\\S.*

Required: Yes

SchemaId

This is a wrapper structure to contain schema identity fields. The structure contains:

- SchemaId\$SchemaArn: The Amazon Resource Name (ARN) of the schema. One of SchemaArn or SchemaName has to be provided.

- `SchemaId$SchemaName`: The name of the schema. One of `SchemaArn` or `SchemaName` has to be provided.

Type: [SchemaId](#) object

Required: Yes

Response Syntax

```
{
  "CreatedTime": "string",
  "DataFormat": "string",
  "SchemaArn": "string",
  "SchemaVersionId": "string",
  "Status": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

[CreatedTime](#)

The date and time the schema was created.

Type: String

[DataFormat](#)

The data format of the schema definition. Currently AVRO, JSON and PROTOBUF are supported.

Type: String

Valid Values: AVRO | JSON | PROTOBUF

[SchemaArn](#)

The Amazon Resource Name (ARN) of the schema.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 10240.

Pattern: `arn:(aws|aws-us-gov|aws-cn):glue:.*`

SchemaVersionId

The schema ID of the schema version.

Type: String

Length Constraints: Fixed length of 36.

Pattern: `[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

Status

The status of the schema version.

Type: String

Valid Values: AVAILABLE | PENDING | FAILURE | DELETING

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetSchemaVersion

Get the specified schema by its unique ID assigned when a version of the schema is created or registered. Schema versions in Deleted status will not be included in the results.

Request Syntax

```
{
  "SchemaId": {
    "RegistryName": "string",
    "SchemaArn": "string",
    "SchemaName": "string"
  },
  "SchemaVersionId": "string",
  "SchemaVersionNumber": {
    "LatestVersion": boolean,
    "VersionNumber": number
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

SchemaId

This is a wrapper structure to contain schema identity fields. The structure contains:

- `SchemaId$SchemaArn`: The Amazon Resource Name (ARN) of the schema. Either `SchemaArn` or `SchemaName` and `RegistryName` has to be provided.
- `SchemaId$SchemaName`: The name of the schema. Either `SchemaArn` or `SchemaName` and `RegistryName` has to be provided.

Type: [SchemaId](#) object

Required: No

SchemaVersionId

The `SchemaVersionId` of the schema version. This field is required for fetching by schema ID. Either this or the `SchemaId` wrapper has to be provided.

Type: String

Length Constraints: Fixed length of 36.

Pattern: `[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

Required: No

SchemaVersionNumber

The version number of the schema.

Type: [SchemaVersionNumber](#) object

Required: No

Response Syntax

```
{
  "CreatedTime": "string",
  "DataFormat": "string",
  "SchemaArn": "string",
  "SchemaDefinition": "string",
  "SchemaVersionId": "string",
  "Status": "string",
  "VersionNumber": number
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

CreatedTime

The date and time the schema version was created.

Type: String

DataFormat

The data format of the schema definition. Currently AVRO, JSON and PROTOBUF are supported.

Type: String

Valid Values: AVRO | JSON | PROTOBUF

SchemaArn

The Amazon Resource Name (ARN) of the schema.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 10240.

Pattern: `arn:(aws|aws-us-gov|aws-cn):glue:.*`

SchemaDefinition

The schema definition for the schema ID.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 170000.

Pattern: `.*\S.*`

SchemaVersionId

The SchemaVersionId of the schema version.

Type: String

Length Constraints: Fixed length of 36.

Pattern: `[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

Status

The status of the schema version.

Type: String

Valid Values: AVAILABLE | PENDING | FAILURE | DELETING

VersionNumber

The version number of the schema.

Type: Long

Valid Range: Minimum value of 1. Maximum value of 100000.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)

- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetSchemaVersionsDiff

Fetches the schema version difference in the specified difference type between two stored schema versions in the Schema Registry.

This API allows you to compare two schema versions between two schema definitions under the same schema.

Request Syntax

```
{
  "FirstSchemaVersionNumber": {
    "LatestVersion": boolean,
    "VersionNumber": number
  },
  "SchemaDiffType": "string",
  "SchemaId": {
    "RegistryName": "string",
    "SchemaArn": "string",
    "SchemaName": "string"
  },
  "SecondSchemaVersionNumber": {
    "LatestVersion": boolean,
    "VersionNumber": number
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

FirstSchemaVersionNumber

The first of the two schema versions to be compared.

Type: [SchemaVersionNumber](#) object

Required: Yes

SchemaDiffType

Refers to SYNTAX_DIFF, which is the currently supported diff type.

Type: String

Valid Values: SYNTAX_DIFF

Required: Yes

Schemald

This is a wrapper structure to contain schema identity fields. The structure contains:

- Schemald\$SchemaArn: The Amazon Resource Name (ARN) of the schema. One of SchemaArn or SchemaName has to be provided.
- Schemald\$SchemaName: The name of the schema. One of SchemaArn or SchemaName has to be provided.

Type: [Schemald](#) object

Required: Yes

SecondSchemaVersionNumber

The second of the two schema versions to be compared.

Type: [SchemaVersionNumber](#) object

Required: Yes

Response Syntax

```
{
  "Diff": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Diff

The difference between schemas as a string in JsonPatch format.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 340000.

Pattern: .*\\S.*

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)

- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)


```
    },
    "S3Encryption": [
      {
        "KmsKeyArn": "string",
        "S3EncryptionMode": "string"
      }
    ]
  },
  "Name": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

SecurityConfiguration

The requested security configuration.

Type: [SecurityConfiguration](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetSecurityConfigurations

Retrieves a list of all security configurations.

Request Syntax

```
{  
  "MaxResults": number,  
  "NextToken": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

MaxResults

The maximum number of results to return.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

NextToken

A continuation token, if this is a continuation call.

Type: String

Required: No

Response Syntax

```
{  
  "NextToken": "string",  
  "SecurityConfigurations": [  
    {
```

```
"CreatedTimeStamp": number,
"EncryptionConfiguration": {
  "CloudWatchEncryption": {
    "CloudWatchEncryptionMode": "string",
    "KmsKeyArn": "string"
  },
  "JobBookmarksEncryption": {
    "JobBookmarksEncryptionMode": "string",
    "KmsKeyArn": "string"
  },
  "S3Encryption": [
    {
      "KmsKeyArn": "string",
      "S3EncryptionMode": "string"
    }
  ]
},
"Name": "string"
}
]
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken

A continuation token, if there are more security configurations to return.

Type: String

SecurityConfigurations

A list of security configurations.

Type: Array of [SecurityConfiguration](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

Response Syntax

```
{
  "Session": {
    "Command": {
      "Name": "string",
      "PythonVersion": "string"
    },
    "CompletedOn": number,
    "Connections": {
      "Connections": [ "string" ]
    },
    "CreatedOn": number,
    "DefaultArguments": {
      "string": "string"
    },
    "Description": "string",
    "DPUSecods": number,
    "ErrorMessage": "string",
    "ExecutionTime": number,
    "GlueVersion": "string",
    "Id": "string",
    "IdleTimeout": number,
    "MaxCapacity": number,
    "NumberOfWorkers": number,
    "ProfileName": "string",
    "Progress": number,
    "Role": "string",
    "SecurityConfiguration": "string",
    "Status": "string",
    "WorkerType": "string"
  }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Session

The session object is returned in the response.

Type: [Session](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)

- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetStatement

Retrieves the statement.

Request Syntax

```
{  
  "Id": number,  
  "RequestOrigin": "string",  
  "SessionId": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Id

The Id of the statement.

Type: Integer

Required: Yes

RequestOrigin

The origin of the request.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[\._\-A-Za-z0-9]+`

Required: No

SessionId

The Session ID of the statement.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "Statement": {
    "Code": "string",
    "CompletedOn": number,
    "Id": number,
    "Output": {
      "Data": {
        "TextPlain": "string"
      },
      "ErrorName": "string",
      "ErrorValue": "string",
      "ExecutionCount": number,
      "Status": "string",
      "Traceback": [ "string" ]
    },
    "Progress": number,
    "StartedOn": number,
    "State": "string"
  }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Statement

Returns the statement.

Type: [Statement](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

IllegalSessionStateException

The session is in an invalid state to perform a requested operation.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)

- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetTable

Retrieves the Table definition in a Data Catalog for a specified table.

Request Syntax

```
{
  "CatalogId": "string",
  "DatabaseName": "string",
  "Name": "string",
  "QueryAsOfTime": number,
  "TransactionId": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog where the table resides. If none is provided, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

DatabaseName

The name of the database in the catalog in which the table resides. For Hive compatibility, this name is entirely lowercase.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Name

The name of the table for which to retrieve the definition. For Hive compatibility, this name is entirely lowercase.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

QueryAsOfTime

The time as of when to read the table contents. If not set, the most recent transaction commit time will be used. Cannot be specified along with `TransactionId`.

Type: Timestamp

Required: No

TransactionId

The transaction ID at which to read the table contents.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\p{L}\p{N}\p{P}]*`

Required: No

Response Syntax

```
{
  "Table": {
    "CatalogId": "string",
    "CreatedBy": "string",
    "CreateTime": number,
    "DatabaseName": "string",
    "Description": "string",
    "FederatedTable": {
```

```
    "ConnectionName": "string",
    "DatabaseIdentifier": "string",
    "Identifier": "string"
  },
  "IsMultiDialectView": boolean,
  "IsRegisteredWithLakeFormation": boolean,
  "LastAccessTime": number,
  "LastAnalyzedTime": number,
  "Name": "string",
  "Owner": "string",
  "Parameters": {
    "string" : "string"
  },
  "PartitionKeys": [
    {
      "Comment": "string",
      "Name": "string",
      "Parameters": {
        "string" : "string"
      },
      "Type": "string"
    }
  ],
  "Retention": number,
  "StorageDescriptor": {
    "AdditionalLocations": [ "string" ],
    "BucketColumns": [ "string" ],
    "Columns": [
      {
        "Comment": "string",
        "Name": "string",
        "Parameters": {
          "string" : "string"
        },
        "Type": "string"
      }
    ],
    "Compressed": boolean,
    "InputFormat": "string",
    "Location": "string",
    "NumberOfBuckets": number,
    "OutputFormat": "string",
    "Parameters": {
      "string" : "string"
    }
  }
}
```

```
    },
    "SchemaReference": {
      "SchemaId": {
        "RegistryName": "string",
        "SchemaArn": "string",
        "SchemaName": "string"
      },
      "SchemaVersionId": "string",
      "SchemaVersionNumber": number
    },
    "SerdeInfo": {
      "Name": "string",
      "Parameters": {
        "string": "string"
      },
      "SerializationLibrary": "string"
    },
    "SkewedInfo": {
      "SkewedColumnNames": [ "string" ],
      "SkewedColumnValueLocationMaps": {
        "string": "string"
      },
      "SkewedColumnValues": [ "string" ]
    },
    "SortColumns": [
      {
        "Column": "string",
        "SortOrder": number
      }
    ],
    "StoredAsSubDirectories": boolean
  },
  "TableType": "string",
  "TargetTable": {
    "CatalogId": "string",
    "DatabaseName": "string",
    "Name": "string",
    "Region": "string"
  },
  "UpdateTime": number,
  "VersionId": "string",
  "ViewDefinition": {
    "Definer": "string",
    "IsProtected": boolean,
```

```
    "Representations": [
      {
        "Dialect": "string",
        "DialectVersion": "string",
        "IsStale": boolean,
        "ValidationConnection": "string",
        "ViewExpandedText": "string",
        "ViewOriginalText": "string"
      }
    ],
    "SubObjects": [ "string" ]
  },
  "ViewExpandedText": "string",
  "ViewOriginalText": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Table

The `Table` object that defines the specified table.

Type: [Table](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

FederationSourceException

A federation source failed.

HTTP Status Code: 400

FederationSourceRetryableException

A federation source failed, but the operation may be retried.

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNotReadyException

A resource was not ready for a transaction.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)

- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetTableOptimizer

Returns the configuration of all optimizers associated with a specified table.

Request Syntax

```
{
  "CatalogId": "string",
  "DatabaseName": "string",
  "TableName": "string",
  "Type": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The Catalog ID of the table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

DatabaseName

The name of the database in the catalog in which the table resides.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

TableName

The name of the table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Type

The type of table optimizer.

Type: String

Valid Values: `compaction`

Required: Yes

Response Syntax

```
{
  "CatalogId": "string",
  "DatabaseName": "string",
  "TableName": "string",
  "TableOptimizer": {
    "configuration": {
      "enabled": boolean,
      "roleArn": "string"
    },
    "lastRun": {
      "endTimeStamp": number,
      "error": "string",
      "eventType": "string",
      "metrics": {
        "JobDurationInHour": "string",
        "NumberOfBytesCompacted": "string",
        "NumberOfDpus": "string",
        "NumberOfFilesCompacted": "string"
      }
    }
  },
}
```

```
    "startTimeStamp": number
  },
  "type": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

CatalogId

The Catalog ID of the table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

DatabaseName

The name of the database in the catalog in which the table resides.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

TableName

The name of the table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

TableOptimizer

The optimizer associated with the specified table.

Type: [TableOptimizer](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)

- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Expression

A regular expression pattern. If present, only those tables whose names match the pattern are returned.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

MaxResults

The maximum number of tables to return in a single response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

NextToken

A continuation token, included if this is a continuation call.

Type: String

Required: No

QueryAsOfTime

The time as of when to read the table contents. If not set, the most recent transaction commit time will be used. Cannot be specified along with `TransactionId`.

Type: Timestamp

Required: No

TransactionId

The transaction ID at which to read the table contents.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\p{L}\p{N}\p{P}]*`

Required: No

Response Syntax

```
{
  "NextToken": "string",
  "TableList": [
    {
      "CatalogId": "string",
      "CreatedBy": "string",
      "CreateTime": number,
      "DatabaseName": "string",
      "Description": "string",
      "FederatedTable": {
        "ConnectionName": "string",
        "DatabaseIdentifier": "string",
        "Identifier": "string"
      },
      "IsMultiDialectView": boolean,
      "IsRegisteredWithLakeFormation": boolean,
      "LastAccessTime": number,
      "LastAnalyzedTime": number,
      "Name": "string",
      "Owner": "string",
      "Parameters": {
        "string" : "string"
      },
      "PartitionKeys": [
        {
          "Comment": "string",
          "Name": "string",
          "Parameters": {
            "string" : "string"
          },
          "Type": "string"
        }
      ],
      "Retention": number,
      "StorageDescriptor": {
```

```
"AdditionalLocations": [ "string" ],
"BucketColumns": [ "string" ],
"Columns": [
  {
    "Comment": "string",
    "Name": "string",
    "Parameters": {
      "string" : "string"
    },
    "Type": "string"
  }
],
"Compressed": boolean,
"InputFormat": "string",
"Location": "string",
"NumberOfBuckets": number,
"OutputFormat": "string",
"Parameters": {
  "string" : "string"
},
"SchemaReference": {
  "SchemaId": {
    "RegistryName": "string",
    "SchemaArn": "string",
    "SchemaName": "string"
  },
  "SchemaVersionId": "string",
  "SchemaVersionNumber": number
},
"SerdeInfo": {
  "Name": "string",
  "Parameters": {
    "string" : "string"
  },
  "SerializationLibrary": "string"
},
"SkewedInfo": {
  "SkewedColumnNames": [ "string" ],
  "SkewedColumnValueLocationMaps": {
    "string" : "string"
  },
  "SkewedColumnValues": [ "string" ]
},
"SortColumns": [
```

```

    {
      "Column": "string",
      "SortOrder": number
    }
  ],
  "StoredAsSubDirectories": boolean
},
"TableType": "string",
"TargetTable": {
  "CatalogId": "string",
  "DatabaseName": "string",
  "Name": "string",
  "Region": "string"
},
"UpdateTime": number,
"VersionId": "string",
"ViewDefinition": {
  "Definer": "string",
  "IsProtected": boolean,
  "Representations": [
    {
      "Dialect": "string",
      "DialectVersion": "string",
      "IsStale": boolean,
      "ValidationConnection": "string",
      "ViewExpandedText": "string",
      "ViewOriginalText": "string"
    }
  ],
  "SubObjects": [ "string" ]
},
"ViewExpandedText": "string",
"ViewOriginalText": "string"
}
]
}

```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken

A continuation token, present if the current list segment is not the last.

Type: String

TableList

A list of the requested Table objects.

Type: Array of [Table](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

FederationSourceException

A federation source failed.

HTTP Status Code: 400

FederationSourceRetryableException

A federation source failed, but the operation may be retried.

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetTableVersion

Retrieves a specified version of a table.

Request Syntax

```
{  
  "CatalogId": "string",  
  "DatabaseName": "string",  
  "TableName": "string",  
  "VersionId": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog where the tables reside. If none is provided, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

DatabaseName

The database in the catalog in which the table resides. For Hive compatibility, this name is entirely lowercase.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

TableName

The name of the table. For Hive compatibility, this name is entirely lowercase.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

VersionId

The ID value of the table version to be retrieved. A VersionID is a string representation of an integer. Each version is incremented by 1.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Response Syntax

```
{
  "TableVersion": {
    "Table": {
      "CatalogId": "string",
      "CreatedBy": "string",
      "CreateTime": number,
      "DatabaseName": "string",
      "Description": "string",
      "FederatedTable": {
        "ConnectionName": "string",
        "DatabaseIdentifier": "string",
        "Identifier": "string"
      },
      "IsMultiDialectView": boolean,
      "IsRegisteredWithLakeFormation": boolean,
      "LastAccessTime": number,
      "LastAnalyzedTime": number,

```



```
"Name": "string",
"Owner": "string",
"Parameters": {
  "string" : "string"
},
"PartitionKeys": [
  {
    "Comment": "string",
    "Name": "string",
    "Parameters": {
      "string" : "string"
    },
    "Type": "string"
  }
],
"Retention": number,
"StorageDescriptor": {
  "AdditionalLocations": [ "string" ],
  "BucketColumns": [ "string" ],
  "Columns": [
    {
      "Comment": "string",
      "Name": "string",
      "Parameters": {
        "string" : "string"
      },
      "Type": "string"
    }
  ],
  "Compressed": boolean,
  "InputFormat": "string",
  "Location": "string",
  "NumberOfBuckets": number,
  "OutputFormat": "string",
  "Parameters": {
    "string" : "string"
  },
  "SchemaReference": {
    "SchemaId": {
      "RegistryName": "string",
      "SchemaArn": "string",
      "SchemaName": "string"
    },
    "SchemaVersionId": "string",
```

```

    "SchemaVersionNumber": number
  },
  "SerdeInfo": {
    "Name": "string",
    "Parameters": {
      "string" : "string"
    },
    "SerializationLibrary": "string"
  },
  "SkewedInfo": {
    "SkewedColumnNames": [ "string" ],
    "SkewedColumnValueLocationMaps": {
      "string" : "string"
    },
    "SkewedColumnValues": [ "string" ]
  },
  "SortColumns": [
    {
      "Column": "string",
      "SortOrder": number
    }
  ],
  "StoredAsSubDirectories": boolean
},
"TableType": "string",
"TargetTable": {
  "CatalogId": "string",
  "DatabaseName": "string",
  "Name": "string",
  "Region": "string"
},
"UpdateTime": number,
"VersionId": "string",
"ViewDefinition": {
  "Definer": "string",
  "IsProtected": boolean,
  "Representations": [
    {
      "Dialect": "string",
      "DialectVersion": "string",
      "IsStale": boolean,
      "ValidationConnection": "string",
      "ViewExpandedText": "string",
      "ViewOriginalText": "string"
    }
  ]
}

```

```
    }  
    ],  
    "SubObjects": [ "string" ]  
  },  
  "ViewExpandedText": "string",  
  "ViewOriginalText": "string"  
},  
"VersionId": "string"  
}  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

TableVersion

The requested table version.

Type: [TableVersion](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)


```

    "DatabaseIdentifier": "string",
    "Identifier": "string"
  },
  "IsMultiDialectView": boolean,
  "IsRegisteredWithLakeFormation": boolean,
  "LastAccessTime": number,
  "LastAnalyzedTime": number,
  "Name": "string",
  "Owner": "string",
  "Parameters": {
    "string" : "string"
  },
  "PartitionKeys": [
    {
      "Comment": "string",
      "Name": "string",
      "Parameters": {
        "string" : "string"
      },
      "Type": "string"
    }
  ],
  "Retention": number,
  "StorageDescriptor": {
    "AdditionalLocations": [ "string" ],
    "BucketColumns": [ "string" ],
    "Columns": [
      {
        "Comment": "string",
        "Name": "string",
        "Parameters": {
          "string" : "string"
        },
        "Type": "string"
      }
    ],
    "Compressed": boolean,
    "InputFormat": "string",
    "Location": "string",
    "NumberOfBuckets": number,
    "OutputFormat": "string",
    "Parameters": {
      "string" : "string"
    }
  },

```

```
"SchemaReference": {
  "SchemaId": {
    "RegistryName": "string",
    "SchemaArn": "string",
    "SchemaName": "string"
  },
  "SchemaVersionId": "string",
  "SchemaVersionNumber": number
},
"SerdeInfo": {
  "Name": "string",
  "Parameters": {
    "string" : "string"
  },
  "SerializationLibrary": "string"
},
"SkewedInfo": {
  "SkewedColumnNames": [ "string" ],
  "SkewedColumnValueLocationMaps": {
    "string" : "string"
  },
  "SkewedColumnValues": [ "string" ]
},
"SortColumns": [
  {
    "Column": "string",
    "SortOrder": number
  }
],
"StoredAsSubDirectories": boolean
},
"TableType": "string",
"TargetTable": {
  "CatalogId": "string",
  "DatabaseName": "string",
  "Name": "string",
  "Region": "string"
},
"UpdateTime": number,
"VersionId": "string",
"ViewDefinition": {
  "Definer": "string",
  "IsProtected": boolean,
  "Representations": [
```



```
    {
      "Dialect": "string",
      "DialectVersion": "string",
      "IsStale": boolean,
      "ValidationConnection": "string",
      "ViewExpandedText": "string",
      "ViewOriginalText": "string"
    },
    "SubObjects": [ "string" ]
  },
  "ViewExpandedText": "string",
  "ViewOriginalText": "string"
},
"VersionId": "string"
}
]
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken

A continuation token, if the list of available versions does not include the last one.

Type: String

TableVersions

A list of strings identifying available versions of the specified table.

Type: Array of [TableVersion](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetTags

Retrieves a list of tags associated with a resource.

Request Syntax

```
{  
  "ResourceArn": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

ResourceArn

The Amazon Resource Name (ARN) of the resource for which to retrieve tags.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 10240.

Pattern: `arn:(aws|aws-us-gov|aws-cn):glue:.*`

Required: Yes

Response Syntax

```
{  
  "Tags": {  
    "string" : "string"  
  }  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Tags

The requested tags.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetTrigger

Retrieves the definition of a trigger.

Request Syntax

```
{  
  "Name": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Name

The name of the trigger to retrieve.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{  
  "Trigger": {  
    "Actions": [  
      {  
        "Arguments": {  
          "string": "string"  
        },  
        "CrawlerName": "string",  
        "JobName": "string",  
        "NotificationProperty": {
```

```
        "NotifyDelayAfter": number
      },
      "SecurityConfiguration": "string",
      "Timeout": number
    }
  ],
  "Description": "string",
  "EventBatchingCondition": {
    "BatchSize": number,
    "BatchWindow": number
  },
  "Id": "string",
  "Name": "string",
  "Predicate": {
    "Conditions": [
      {
        "CrawlerName": "string",
        "CrawlState": "string",
        "JobName": "string",
        "LogicalOperator": "string",
        "State": "string"
      }
    ]
  },
  "Logical": "string"
},
"Schedule": "string",
"State": "string",
"Type": "string",
"WorkflowName": "string"
}
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Trigger

The requested trigger definition.

Type: [Trigger](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)

- [AWS SDK for Ruby V3](#)

GetTriggers

Gets all the triggers associated with a job.

Request Syntax

```
{
  "DependentJobName": "string",
  "MaxResults": number,
  "NextToken": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

DependentJobName

The name of the job to retrieve triggers for. The trigger that can start this job is returned, and if there is no such trigger, all triggers are returned.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

MaxResults

The maximum size of the response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 200.

Required: No

NextToken

A continuation token, if this is a continuation call.

Type: String

Required: No

Response Syntax

```
{
  "NextToken": "string",
  "Triggers": [
    {
      "Actions": [
        {
          "Arguments": {
            "string": "string"
          },
          "CrawlerName": "string",
          "JobName": "string",
          "NotificationProperty": {
            "NotifyDelayAfter": number
          },
          "SecurityConfiguration": "string",
          "Timeout": number
        }
      ],
      "Description": "string",
      "EventBatchingCondition": {
        "BatchSize": number,
        "BatchWindow": number
      },
      "Id": "string",
      "Name": "string",
      "Predicate": {
        "Conditions": [
          {
            "CrawlerName": "string",
            "CrawlState": "string",
            "JobName": "string",
            "LogicalOperator": "string",
            "State": "string"
          }
        ]
      },
      "Logical": "string"
    }
  ],
}
```

```
    "Schedule": "string",
    "State": "string",
    "Type": "string",
    "WorkflowName": "string"
  }
]
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

[NextToken](#)

A continuation token, if not all the requested triggers have yet been returned.

Type: String

[Triggers](#)

A list of triggers for the specified job.

Type: Array of [Trigger](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetUnfilteredPartitionMetadata

Retrieves partition metadata from the Data Catalog that contains unfiltered metadata.

For IAM authorization, the public IAM action associated with this API is `glue:GetPartition`.

Request Syntax

```
{
  "AuditContext": {
    "AdditionalAuditContext": "string",
    "AllColumnsRequested": boolean,
    "RequestedColumns": [ "string" ]
  },
  "CatalogId": "string",
  "DatabaseName": "string",
  "PartitionValues": [ "string" ],
  "QuerySessionContext": {
    "AdditionalContext": {
      "string" : "string"
    },
    "ClusterId": "string",
    "QueryAuthorizationId": "string",
    "QueryId": "string",
    "QueryStartTime": number
  },
  "Region": "string",
  "SupportedPermissionTypes": [ "string" ],
  "TableName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

AuditContext

A structure containing Lake Formation audit context information.

Type: [AuditContext](#) object

Required: No

CatalogId

The catalog ID where the partition resides.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

DatabaseName

(Required) Specifies the name of a database that contains the partition.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

PartitionValues

(Required) A list of partition key values.

Type: Array of strings

Length Constraints: Maximum length of 1024.

Required: Yes

QuerySessionContext

A structure used as a protocol between query engines and Lake Formation or AWS Glue. Contains both a Lake Formation generated authorization identifier and information from the request's authorization context.

Type: [QuerySessionContext](#) object

Required: No

Region

Specified only if the base tables belong to a different AWS Region.


```
"StorageDescriptor": {
  "AdditionalLocations": [ "string" ],
  "BucketColumns": [ "string" ],
  "Columns": [
    {
      "Comment": "string",
      "Name": "string",
      "Parameters": {
        "string" : "string"
      },
      "Type": "string"
    }
  ],
  "Compressed": boolean,
  "InputFormat": "string",
  "Location": "string",
  "NumberOfBuckets": number,
  "OutputFormat": "string",
  "Parameters": {
    "string" : "string"
  },
  "SchemaReference": {
    "SchemaId": {
      "RegistryName": "string",
      "SchemaArn": "string",
      "SchemaName": "string"
    },
    "SchemaVersionId": "string",
    "SchemaVersionNumber": number
  },
  "SerdeInfo": {
    "Name": "string",
    "Parameters": {
      "string" : "string"
    },
    "SerializationLibrary": "string"
  },
  "SkewedInfo": {
    "SkewedColumnNames": [ "string" ],
    "SkewedColumnValueLocationMaps": {
      "string" : "string"
    },
    "SkewedColumnValues": [ "string" ]
  },
}
```


EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

FederationSourceException

A federation source failed.

HTTP Status Code: 400

FederationSourceRetryableException

A federation source failed, but the operation may be retried.

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

PermissionTypeMismatchException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetUnfilteredPartitionsMetadata

Retrieves partition metadata from the Data Catalog that contains unfiltered metadata.

For IAM authorization, the public IAM action associated with this API is `glue:GetPartitions`.

Request Syntax

```
{
  "AuditContext": {
    "AdditionalAuditContext": "string",
    "AllColumnsRequested": boolean,
    "RequestedColumns": [ "string" ]
  },
  "CatalogId": "string",
  "DatabaseName": "string",
  "Expression": "string",
  "MaxResults": number,
  "NextToken": "string",
  "QuerySessionContext": {
    "AdditionalContext": {
      "string" : "string"
    },
    "ClusterId": "string",
    "QueryAuthorizationId": "string",
    "QueryId": "string",
    "QueryStartTime": number
  },
  "Region": "string",
  "Segment": {
    "SegmentNumber": number,
    "TotalSegments": number
  },
  "SupportedPermissionTypes": [ "string" ],
  "TableName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

AuditContext

A structure containing Lake Formation audit context information.

Type: [AuditContext](#) object

Required: No

CatalogId

The ID of the Data Catalog where the partitions in question reside. If none is provided, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

DatabaseName

The name of the catalog database where the partitions reside.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Expression

An expression that filters the partitions to be returned.

The expression uses SQL syntax similar to the SQL WHERE filter clause. The SQL statement parser [JSQLParser](#) parses the expression.

Operators: The following are the operators that you can use in the Expression API call:

=

Checks whether the values of the two operands are equal; if yes, then the condition becomes true.

Example: Assume 'variable a' holds 10 and 'variable b' holds 20.

(a = b) is not true.

< >

Checks whether the values of two operands are equal; if the values are not equal, then the condition becomes true.

Example: (a < > b) is true.

>

Checks whether the value of the left operand is greater than the value of the right operand; if yes, then the condition becomes true.

Example: (a > b) is not true.

<

Checks whether the value of the left operand is less than the value of the right operand; if yes, then the condition becomes true.

Example: (a < b) is true.

>=

Checks whether the value of the left operand is greater than or equal to the value of the right operand; if yes, then the condition becomes true.

Example: (a >= b) is not true.

<=

Checks whether the value of the left operand is less than or equal to the value of the right operand; if yes, then the condition becomes true.

Example: (a <= b) is true.

AND, OR, IN, BETWEEN, LIKE, NOT, IS NULL

Logical operators.

Supported Partition Key Types: The following are the supported partition keys.

- `string`

- date
- timestamp
- int
- bigint
- long
- tinyint
- smallint
- decimal

If an type is encountered that is not valid, an exception is thrown.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

MaxResults

The maximum number of partitions to return in a single response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

NextToken

A continuation token, if this is not the first call to retrieve these partitions.

Type: String

Required: No

QuerySessionContext

A structure used as a protocol between query engines and Lake Formation or AWS Glue. Contains both a Lake Formation generated authorization identifier and information from the request's authorization context.

Type: [QuerySessionContext](#) object

Required: No

Region

Specified only if the base tables belong to a different AWS Region.

Type: String

Length Constraints: Maximum length of 1024.

Required: No

Segment

The segment of the table's partitions to scan in this request.

Type: [Segment](#) object

Required: No

SupportedPermissionTypes

A list of supported permission types.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 255 items.

Valid Values: COLUMN_PERMISSION | CELL_FILTER_PERMISSION | NESTED_PERMISSION
| NESTED_CELL_PERMISSION

Required: Yes

TableName

The name of the table that contains the partition.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "NextToken": "string",
  "UnfilteredPartitions": [
    {
      "AuthorizedColumns": [ "string" ],
      "IsRegisteredWithLakeFormation": boolean,
      "Partition": {
        "CatalogId": "string",
        "CreationTime": number,
        "DatabaseName": "string",
        "LastAccessTime": number,
        "LastAnalyzedTime": number,
        "Parameters": {
          "string" : "string"
        },
        "StorageDescriptor": {
          "AdditionalLocations": [ "string" ],
          "BucketColumns": [ "string" ],
          "Columns": [
            {
              "Comment": "string",
              "Name": "string",
              "Parameters": {
                "string" : "string"
              },
              "Type": "string"
            }
          ],
          "Compressed": boolean,
          "InputFormat": "string",
          "Location": "string",
          "NumberOfBuckets": number,
          "OutputFormat": "string",
          "Parameters": {
            "string" : "string"
          },
          "SchemaReference": {
            "SchemaId": {
              "RegistryName": "string",
              "SchemaArn": "string",
              "SchemaName": "string"
            }
          }
        }
      }
    }
  ]
}
```

```

    },
    "SchemaVersionId": "string",
    "SchemaVersionNumber": number
  },
  "SerdeInfo": {
    "Name": "string",
    "Parameters": {
      "string" : "string"
    },
    "SerializationLibrary": "string"
  },
  "SkewedInfo": {
    "SkewedColumnNames": [ "string" ],
    "SkewedColumnValueLocationMaps": {
      "string" : "string"
    },
    "SkewedColumnValues": [ "string" ]
  },
  "SortColumns": [
    {
      "Column": "string",
      "SortOrder": number
    }
  ],
  "StoredAsSubDirectories": boolean
},
"TableName": "string",
"Values": [ "string" ]
}
}
]
}

```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken

A continuation token, if the returned list of partitions does not include the last one.

Type: String

UnfilteredPartitions

A list of requested partitions.

Type: Array of [UnfilteredPartition](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

FederationSourceException

A federation source failed.

HTTP Status Code: 400

FederationSourceRetryableException

A federation source failed, but the operation may be retried.

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

PermissionTypeMismatchException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetUnfilteredTableMetadata

Allows a third-party analytical engine to retrieve unfiltered table metadata from the Data Catalog.

For IAM authorization, the public IAM action associated with this API is `glue:GetTable`.

Request Syntax

```
{
  "AuditContext": {
    "AdditionalAuditContext": "string",
    "AllColumnsRequested": boolean,
    "RequestedColumns": [ "string" ]
  },
  "CatalogId": "string",
  "DatabaseName": "string",
  "Name": "string",
  "ParentResourceArn": "string",
  "Permissions": [ "string" ],
  "QuerySessionContext": {
    "AdditionalContext": {
      "string" : "string"
    },
    "ClusterId": "string",
    "QueryAuthorizationId": "string",
    "QueryId": "string",
    "QueryStartTime": number
  },
  "Region": "string",
  "RootResourceArn": "string",
  "SupportedDialect": {
    "Dialect": "string",
    "DialectVersion": "string"
  },
  "SupportedPermissionTypes": [ "string" ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

AuditContext

A structure containing Lake Formation audit context information.

Type: [AuditContext](#) object

Required: No

CatalogId

The catalog ID where the table resides.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

DatabaseName

(Required) Specifies the name of a database that contains the table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Name

(Required) Specifies the name of a table for which you are requesting metadata.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

ParentResourceArn

The resource ARN of the view.

Type: String

Length Constraints: Minimum length of 20. Maximum length of 2048.

Required: No

Permissions

The Lake Formation data permissions of the caller on the table. Used to authorize the call when no view context is found.

Type: Array of strings

Valid Values: ALL | SELECT | ALTER | DROP | DELETE | INSERT |
CREATE_DATABASE | CREATE_TABLE | DATA_LOCATION_ACCESS

Required: No

QuerySessionContext

A structure used as a protocol between query engines and Lake Formation or AWS Glue. Contains both a Lake Formation generated authorization identifier and information from the request's authorization context.

Type: [QuerySessionContext](#) object

Required: No

Region

Specified only if the base tables belong to a different AWS Region.

Type: String

Length Constraints: Maximum length of 1024.

Required: No

RootResourceArn

The resource ARN of the root view in a chain of nested views.

Type: String

Length Constraints: Minimum length of 20. Maximum length of 2048.

Required: No

SupportedDialect

A structure specifying the dialect and dialect version used by the query engine.

Type: [SupportedDialect](#) object

Required: No

SupportedPermissionTypes

Indicates the level of filtering a third-party analytical engine is capable of enforcing when calling the `GetUnfilteredTableMetadata` API operation. Accepted values are:

- `COLUMN_PERMISSION` - Column permissions ensure that users can access only specific columns in the table. If there are particular columns contain sensitive data, data lake administrators can define column filters that exclude access to specific columns.
- `CELL_FILTER_PERMISSION` - Cell-level filtering combines column filtering (include or exclude columns) and row filter expressions to restrict access to individual elements in the table.
- `NESTED_PERMISSION` - Nested permissions combines cell-level filtering and nested column filtering to restrict access to columns and/or nested columns in specific rows based on row filter expressions.
- `NESTED_CELL_PERMISSION` - Nested cell permissions combines nested permission with nested cell-level filtering. This allows different subsets of nested columns to be restricted based on an array of row filter expressions.

Note: Each of these permission types follows a hierarchical order where each subsequent permission type includes all permission of the previous type.

Important: If you provide a supported permission type that doesn't match the user's level of permissions on the table, then Lake Formation raises an exception. For example, if the third-party engine calling the `GetUnfilteredTableMetadata` operation can enforce only column-level filtering, and the user has nested cell filtering applied on the table, Lake Formation throws an exception, and will not return unfiltered table metadata and data access credentials.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 255 items.

Valid Values: COLUMN_PERMISSION | CELL_FILTER_PERMISSION | NESTED_PERMISSION
| NESTED_CELL_PERMISSION

Required: Yes

Response Syntax

```
{
  "AuthorizedColumns": [ "string" ],
  "CellFilters": [
    {
      "ColumnName": "string",
      "RowFilterExpression": "string"
    }
  ],
  "IsMultiDialectView": boolean,
  "IsProtected": boolean,
  "IsRegisteredWithLakeFormation": boolean,
  "Permissions": [ "string" ],
  "QueryAuthorizationId": "string",
  "ResourceArn": "string",
  "RowFilter": "string",
  "Table": {
    "CatalogId": "string",
    "CreatedBy": "string",
    "CreateTime": number,
    "DatabaseName": "string",
    "Description": "string",
    "FederatedTable": {
      "ConnectionName": "string",
      "DatabaseIdentifier": "string",
      "Identifier": "string"
    }
  },
  "IsMultiDialectView": boolean,
  "IsRegisteredWithLakeFormation": boolean,
  "LastAccessTime": number,
  "LastAnalyzedTime": number,
  "Name": "string",
  "Owner": "string",
  "Parameters": {
```

```
    "string" : "string"
  },
  "PartitionKeys": [
    {
      "Comment": "string",
      "Name": "string",
      "Parameters": {
        "string" : "string"
      },
      "Type": "string"
    }
  ],
  "Retention": number,
  "StorageDescriptor": {
    "AdditionalLocations": [ "string" ],
    "BucketColumns": [ "string" ],
    "Columns": [
      {
        "Comment": "string",
        "Name": "string",
        "Parameters": {
          "string" : "string"
        },
        "Type": "string"
      }
    ],
    "Compressed": boolean,
    "InputFormat": "string",
    "Location": "string",
    "NumberOfBuckets": number,
    "OutputFormat": "string",
    "Parameters": {
      "string" : "string"
    },
    "SchemaReference": {
      "SchemaId": {
        "RegistryName": "string",
        "SchemaArn": "string",
        "SchemaName": "string"
      },
      "SchemaVersionId": "string",
      "SchemaVersionNumber": number
    },
    "SerdeInfo": {
```

```
    "Name": "string",
    "Parameters": {
      "string": "string"
    },
    "SerializationLibrary": "string"
  },
  "SkewedInfo": {
    "SkewedColumnNames": [ "string" ],
    "SkewedColumnValueLocationMaps": {
      "string": "string"
    },
    "SkewedColumnValues": [ "string" ]
  },
  "SortColumns": [
    {
      "Column": "string",
      "SortOrder": number
    }
  ],
  "StoredAsSubDirectories": boolean
},
"TableType": "string",
"TargetTable": {
  "CatalogId": "string",
  "DatabaseName": "string",
  "Name": "string",
  "Region": "string"
},
"UpdateTime": number,
"VersionId": "string",
"ViewDefinition": {
  "Definer": "string",
  "IsProtected": boolean,
  "Representations": [
    {
      "Dialect": "string",
      "DialectVersion": "string",
      "IsStale": boolean,
      "ValidationConnection": "string",
      "ViewExpandedText": "string",
      "ViewOriginalText": "string"
    }
  ]
},
"SubObjects": [ "string" ]
```


Type: Boolean

Permissions

The Lake Formation data permissions of the caller on the table. Used to authorize the call when no view context is found.

Type: Array of strings

Valid Values: ALL | SELECT | ALTER | DROP | DELETE | INSERT |
CREATE_DATABASE | CREATE_TABLE | DATA_LOCATION_ACCESS

QueryAuthorizationId

A cryptographically generated query identifier generated by AWS Glue or Lake Formation.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

ResourceArn

The resource ARN of the parent resource extracted from the request.

Type: String

Length Constraints: Minimum length of 20. Maximum length of 2048.

RowFilter

The filter that applies to the table. For example when applying the filter in SQL, it would go in the WHERE clause and can be evaluated by using an AND operator with any other predicates applied by the user querying the table.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Table

A Table object containing the table metadata.

Type: [Table](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

FederationSourceException

A federation source failed.

HTTP Status Code: 400

FederationSourceRetryableException

A federation source failed, but the operation may be retried.

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

PermissionTypeMismatchException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)


```
    "string" : {
      "AllowedValues": [ "string" ],
      "DefaultValue": "string",
      "MaxValue": "string",
      "MinValue": "string"
    }
  },
  "CreatedOn": number,
  "Description": "string",
  "LastModifiedOn": number,
  "Name": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Configuration

A ProfileConfiguration object specifying the job and session values for the profile.

Type: [ProfileConfiguration](#) object

CreatedOn

The date and time when the usage profile was created.

Type: Timestamp

Description

A description of the usage profile.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

LastModifiedOn

The date and time when the usage profile was last modified.

Type: Timestamp

Name

The name of the usage profile.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationNotSupportedException

The operation is not available in the region.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

FunctionName

The name of the function.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "UserDefinedFunction": {
    "CatalogId": "string",
    "ClassName": "string",
    "CreateTime": number,
    "DatabaseName": "string",
    "FunctionName": "string",
    "OwnerName": "string",
    "OwnerType": "string",
    "ResourceUri": [
      {
        "ResourceType": "string",
        "Uri": "string"
      }
    ]
  }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

UserDefinedFunction

The requested function definition.

Type: [UserDefinedFunction](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)

- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetUserDefinedFunctions

Retrieves multiple function definitions from the Data Catalog.

Request Syntax

```
{
  "CatalogId": "string",
  "DatabaseName": "string",
  "MaxResults": number,
  "NextToken": "string",
  "Pattern": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog where the functions to be retrieved are located. If none is provided, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

DatabaseName

The name of the catalog database where the functions are located. If none is provided, functions from all the databases across the catalog will be returned.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

MaxResults

The maximum number of functions to return in one response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

NextToken

A continuation token, if this is a continuation call.

Type: String

Required: No

Pattern

An optional function-name pattern string that filters the function definitions returned.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "NextToken": "string",
  "UserDefinedFunctions": [
    {
      "CatalogId": "string",
      "ClassName": "string",
      "CreateTime": number,
      "DatabaseName": "string",
      "FunctionName": "string",
      "OwnerName": "string",
    }
  ]
}
```

```
    "OwnerType": "string",
    "ResourceUris": [
      {
        "ResourceType": "string",
        "Uri": "string"
      }
    ]
  }
]
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken

A continuation token, if the list of functions returned does not include the last requested function.

Type: String

UserDefinedFunctions

A list of requested function definitions.

Type: Array of [UserDefinedFunction](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetWorkflow

Retrieves resource metadata for a workflow.

Request Syntax

```
{
  "IncludeGraph": boolean,
  "Name": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

IncludeGraph

Specifies whether to include a graph when returning the workflow resource metadata.

Type: Boolean

Required: No

Name

The name of the workflow to retrieve.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "Workflow": {
    "BlueprintDetails": {
      "BlueprintName": "string",
    }
  }
}
```

```

    "RunId": "string"
  },
  "CreatedOn": number,
  "DefaultRunProperties": {
    "string": "string"
  },
  "Description": "string",
  "Graph": {
    "Edges": [
      {
        "DestinationId": "string",
        "SourceId": "string"
      }
    ],
    "Nodes": [
      {
        "CrawlerDetails": {
          "Crawls": [
            {
              "CompletedOn": number,
              "ErrorMessage": "string",
              "LogGroup": "string",
              "LogStream": "string",
              "StartedOn": number,
              "State": "string"
            }
          ]
        },
        "JobDetails": {
          "JobRuns": [
            {
              "AllocatedCapacity": number,
              "Arguments": {
                "string": "string"
              },
              "Attempt": number,
              "CompletedOn": number,
              "DPUSeconds": number,
              "ErrorMessage": "string",
              "ExecutionClass": "string",
              "ExecutionTime": number,
              "GlueVersion": "string",
              "Id": "string",
              "JobMode": "string",

```

```

    "JobName": "string",
    "JobRunState": "string",
    "LastModifiedOn": number,
    "LogGroupName": "string",
    "MaintenanceWindow": "string",
    "MaxCapacity": number,
    "NotificationProperty": {
      "NotifyDelayAfter": number
    },
    "NumberOfWorkers": number,
    "PredecessorRuns": [
      {
        "JobName": "string",
        "RunId": "string"
      }
    ],
    "PreviousRunId": "string",
    "ProfileName": "string",
    "SecurityConfiguration": "string",
    "StartedOn": number,
    "Timeout": number,
    "TriggerName": "string",
    "WorkerType": "string"
  }
]
},
"Name": "string",
"TriggerDetails": {
  "Trigger": {
    "Actions": [
      {
        "Arguments": {
          "string": "string"
        },
        "CrawlerName": "string",
        "JobName": "string",
        "NotificationProperty": {
          "NotifyDelayAfter": number
        },
        "SecurityConfiguration": "string",
        "Timeout": number
      }
    ],
    "Description": "string",

```

```

    "EventBatchingCondition": {
      "BatchSize": number,
      "BatchWindow": number
    },
    "Id": "string",
    "Name": "string",
    "Predicate": {
      "Conditions": [
        {
          "CrawlerName": "string",
          "CrawlState": "string",
          "JobName": "string",
          "LogicalOperator": "string",
          "State": "string"
        }
      ],
      "Logical": "string"
    },
    "Schedule": "string",
    "State": "string",
    "Type": "string",
    "WorkflowName": "string"
  }
},
"Type": "string",
"UniqueId": "string"
}
]
},
"LastModifiedOn": number,
"LastRun": {
  "CompletedOn": number,
  "ErrorMessage": "string",
  "Graph": {
    "Edges": [
      {
        "DestinationId": "string",
        "SourceId": "string"
      }
    ],
    "Nodes": [
      {
        "CrawlerDetails": {
          "Crawls": [

```



```

        {
            "CompletedOn": number,
            "ErrorMessage": "string",
            "LogGroup": "string",
            "LogStream": "string",
            "StartedOn": number,
            "State": "string"
        }
    ],
},
"JobDetails": {
    "JobRuns": [
        {
            "AllocatedCapacity": number,
            "Arguments": {
                "string" : "string"
            },
            "Attempt": number,
            "CompletedOn": number,
            "DPUSecods": number,
            "ErrorMessage": "string",
            "ExecutionClass": "string",
            "ExecutionTime": number,
            "GlueVersion": "string",
            "Id": "string",
            "JobMode": "string",
            "JobName": "string",
            "JobRunState": "string",
            "LastModifiedOn": number,
            "LogGroupName": "string",
            "MaintenanceWindow": "string",
            "MaxCapacity": number,
            "NotificationProperty": {
                "NotifyDelayAfter": number
            },
            "NumberOfWorkers": number,
            "PredecessorRuns": [
                {
                    "JobName": "string",
                    "RunId": "string"
                }
            ],
            "PreviousRunId": "string",
            "ProfileName": "string",

```

```

        "SecurityConfiguration": "string",
        "StartedOn": number,
        "Timeout": number,
        "TriggerName": "string",
        "WorkerType": "string"
    }
]
},
"Name": "string",
"TriggerDetails": {
    "Trigger": {
        "Actions": [
            {
                "Arguments": {
                    "string": "string"
                },
                "CrawlerName": "string",
                "JobName": "string",
                "NotificationProperty": {
                    "NotifyDelayAfter": number
                },
                "SecurityConfiguration": "string",
                "Timeout": number
            }
        ],
        "Description": "string",
        "EventBatchingCondition": {
            "BatchSize": number,
            "BatchWindow": number
        },
        "Id": "string",
        "Name": "string",
        "Predicate": {
            "Conditions": [
                {
                    "CrawlerName": "string",
                    "CrawlState": "string",
                    "JobName": "string",
                    "LogicalOperator": "string",
                    "State": "string"
                }
            ],
            "Logical": "string"
        }
    }
},

```

```

        "Schedule": "string",
        "State": "string",
        "Type": "string",
        "WorkflowName": "string"
    }
},
"Type": "string",
"UniqueId": "string"
}
]
},
"Name": "string",
"PreviousRunId": "string",
"StartedOn": number,
"StartingEventBatchCondition": {
    "BatchSize": number,
    "BatchWindow": number
},
"Statistics": {
    "ErroredActions": number,
    "FailedActions": number,
    "RunningActions": number,
    "StoppedActions": number,
    "SucceededActions": number,
    "TimeoutActions": number,
    "TotalActions": number,
    "WaitingActions": number
},
"Status": "string",
"WorkflowRunId": "string",
"WorkflowRunProperties": {
    "string" : "string"
}
},
"MaxConcurrentRuns": number,
"Name": "string"
}
}

```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Workflow

The resource metadata for the workflow.

Type: [Workflow](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)

- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "Run": {
    "CompletedOn": number,
    "ErrorMessage": "string",
    "Graph": {
      "Edges": [
        {
          "DestinationId": "string",
          "SourceId": "string"
        }
      ],
      "Nodes": [
        {
          "CrawlerDetails": {
            "Crawls": [
              {
                "CompletedOn": number,
                "ErrorMessage": "string",
                "LogGroup": "string",
                "LogStream": "string",
                "StartedOn": number,
                "State": "string"
              }
            ]
          },
          "JobDetails": {
            "JobRuns": [
              {
                "AllocatedCapacity": number,
                "Arguments": {
                  "string": "string"
                },
                "Attempt": number,
                "CompletedOn": number,
                "DPUSecods": number,

```

```

    "ErrorMessage": "string",
    "ExecutionClass": "string",
    "ExecutionTime": number,
    "GlueVersion": "string",
    "Id": "string",
    "JobMode": "string",
    "JobName": "string",
    "JobRunState": "string",
    "LastModifiedOn": number,
    "LogGroupName": "string",
    "MaintenanceWindow": "string",
    "MaxCapacity": number,
    "NotificationProperty": {
      "NotifyDelayAfter": number
    },
    "NumberOfWorkers": number,
    "PredecessorRuns": [
      {
        "JobName": "string",
        "RunId": "string"
      }
    ],
    "PreviousRunId": "string",
    "ProfileName": "string",
    "SecurityConfiguration": "string",
    "StartedOn": number,
    "Timeout": number,
    "TriggerName": "string",
    "WorkerType": "string"
  }
]
},
"Name": "string",
"TriggerDetails": {
  "Trigger": {
    "Actions": [
      {
        "Arguments": {
          "string": "string"
        },
        "CrawlerName": "string",
        "JobName": "string",
        "NotificationProperty": {
          "NotifyDelayAfter": number
        }
      }
    ]
  }
}

```



```

        },
        "SecurityConfiguration": "string",
        "Timeout": number
    }
],
"Description": "string",
"EventBatchingCondition": {
    "BatchSize": number,
    "BatchWindow": number
},
"Id": "string",
"Name": "string",
"Predicate": {
    "Conditions": [
        {
            "CrawlerName": "string",
            "CrawlState": "string",
            "JobName": "string",
            "LogicalOperator": "string",
            "State": "string"
        }
    ],
    "Logical": "string"
},
"Schedule": "string",
"State": "string",
"Type": "string",
"WorkflowName": "string"
}
},
"Type": "string",
"UniqueId": "string"
}
]
},
"Name": "string",
"PreviousRunId": "string",
"StartedOn": number,
"StartingEventBatchCondition": {
    "BatchSize": number,
    "BatchWindow": number
},
"Statistics": {
    "ErroredActions": number,

```

```
    "FailedActions": number,
    "RunningActions": number,
    "StoppedActions": number,
    "SucceededActions": number,
    "TimeoutActions": number,
    "TotalActions": number,
    "WaitingActions": number
  },
  "Status": "string",
  "WorkflowRunId": "string",
  "WorkflowRunProperties": {
    "string" : "string"
  }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Run

The requested workflow run metadata.

Type: [WorkflowRun](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetWorkflowRunProperties

Retrieves the workflow run properties which were set during the run.

Request Syntax

```
{  
  "Name": "string",  
  "RunId": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Name

Name of the workflow which was run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

RunId

The ID of the workflow run whose run properties should be returned.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "RunProperties": {
    "string" : "string"
  }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

RunProperties

The workflow run properties which were set during the specified run.

Type: String to string map

Key Length Constraints: Minimum length of 1. Maximum length of 255.

Key Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetWorkflowRuns

Retrieves metadata for all runs of a given workflow.

Request Syntax

```
{  
  "IncludeGraph": boolean,  
  "MaxResults": number,  
  "Name": "string",  
  "NextToken": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

IncludeGraph

Specifies whether to include the workflow graph in response or not.

Type: Boolean

Required: No

MaxResults

The maximum number of workflow runs to be included in the response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

Name

Name of the workflow whose metadata of runs should be returned.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.


```

    "AllocatedCapacity": number,
    "Arguments": {
      "string" : "string"
    },
    "Attempt": number,
    "CompletedOn": number,
    "DPUSecods": number,
    "ErrorMessage": "string",
    "ExecutionClass": "string",
    "ExecutionTime": number,
    "GlueVersion": "string",
    "Id": "string",
    "JobMode": "string",
    "JobName": "string",
    "JobRunState": "string",
    "LastModifiedOn": number,
    "LogGroupName": "string",
    "MaintenanceWindow": "string",
    "MaxCapacity": number,
    "NotificationProperty": {
      "NotifyDelayAfter": number
    },
    "NumberOfWorkers": number,
    "PredecessorRuns": [
      {
        "JobName": "string",
        "RunId": "string"
      }
    ],
    "PreviousRunId": "string",
    "ProfileName": "string",
    "SecurityConfiguration": "string",
    "StartedOn": number,
    "Timeout": number,
    "TriggerName": "string",
    "WorkerType": "string"
  }
]
},
"Name": "string",
"TriggerDetails": {
  "Trigger": {
    "Actions": [
      {

```

```

        "Arguments": {
            "string": "string"
        },
        "CrawlerName": "string",
        "JobName": "string",
        "NotificationProperty": {
            "NotifyDelayAfter": number
        },
        "SecurityConfiguration": "string",
        "Timeout": number
    }
],
"Description": "string",
"EventBatchingCondition": {
    "BatchSize": number,
    "BatchWindow": number
},
"Id": "string",
"Name": "string",
"Predicate": {
    "Conditions": [
        {
            "CrawlerName": "string",
            "CrawlState": "string",
            "JobName": "string",
            "LogicalOperator": "string",
            "State": "string"
        }
    ],
    "Logical": "string"
},
"Schedule": "string",
"State": "string",
"Type": "string",
"WorkflowName": "string"
}
},
"Type": "string",
"UniqueId": "string"
}
],
},
"Name": "string",
"PreviousRunId": "string",

```

```
"StartedOn": number,
"StartingEventBatchCondition": {
  "BatchSize": number,
  "BatchWindow": number
},
"Statistics": {
  "ErroredActions": number,
  "FailedActions": number,
  "RunningActions": number,
  "StoppedActions": number,
  "SucceededActions": number,
  "TimeoutActions": number,
  "TotalActions": number,
  "WaitingActions": number
},
"Status": "string",
"WorkflowRunId": "string",
"WorkflowRunProperties": {
  "string" : "string"
}
}
]
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken

A continuation token, if not all requested workflow runs have been returned.

Type: String

Runs

A list of workflow run metadata objects.

Type: Array of [WorkflowRun](#) objects

Array Members: Minimum number of 1 item. Maximum number of 1000 items.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)

- [AWS SDK for Ruby V3](#)

ImportCatalogToGlue

Imports an existing Amazon Athena Data Catalog to AWS Glue.

Request Syntax

```
{  
  "CatalogId": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the catalog to import. Currently, this should be the AWS account ID.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListBlueprints

Lists all the blueprint names in an account.

Request Syntax

```
{
  "MaxResults": number,
  "NextToken": "string",
  "Tags": {
    "string" : "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

MaxResults

The maximum size of a list to return.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 25.

Required: No

NextToken

A continuation token, if this is a continuation request.

Type: String

Required: No

Tags

Filters the list by an AWS resource tag.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Required: No

Response Syntax

```
{
  "Blueprints": [ "string" ],
  "NextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Blueprints

List of names of blueprints in the account.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[\.\-_\A-Za-z0-9]+`

NextToken

A continuation token, if not all blueprint names have been returned.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListColumnStatisticsTaskRuns

List all task runs for a particular account.

Request Syntax

```
{  
  "MaxResults": number,  
  "NextToken": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

MaxResults

The maximum size of the response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

NextToken

A continuation token, if this is a continuation call.

Type: String

Required: No

Response Syntax

```
{  
  "ColumnStatisticsTaskRunIds": [ "string" ],  
  "NextToken": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

ColumnStatisticsTaskRunIds

A list of column statistics task run IDs.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 100 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

NextToken

A continuation token, if not all task run IDs have yet been returned.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)

- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListCrawlers

Retrieves the names of all crawler resources in this AWS account, or the resources with the specified tag. This operation allows you to see which resources are available in your account, and their names.

This operation takes the optional `Tags` field, which you can use as a filter on the response so that tagged resources can be retrieved as a group. If you choose to use tags filtering, only resources with the tag are retrieved.

Request Syntax

```
{
  "MaxResults": number,
  "NextToken": "string",
  "Tags": {
    "string" : "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

MaxResults

The maximum size of a list to return.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

NextToken

A continuation token, if this is a continuation request.

Type: String

Required: No

Tags

Specifies to return only these tagged resources.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Required: No

Response Syntax

```
{
  "CrawlerNames": [ "string" ],
  "NextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

CrawlerNames

The names of all crawlers in the account, or the crawlers with the specified tags.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 100 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

NextToken

A continuation token, if the returned list does not contain the last metric available.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListCrawls

Returns all the crawls of a specified crawler. Returns only the crawls that have occurred since the launch date of the crawler history feature, and only retains up to 12 months of crawls. Older crawls will not be returned.

You may use this API to:

- Retrieve all the crawls of a specified crawler.
- Retrieve all the crawls of a specified crawler within a limited count.
- Retrieve all the crawls of a specified crawler in a specific time range.
- Retrieve all the crawls of a specified crawler with a particular state, crawl ID, or DPU hour value.

Request Syntax

```
{
  "CrawlerName": "string",
  "Filters": [
    {
      "FieldName": "string",
      "FieldValue": "string",
      "FilterOperator": "string"
    }
  ],
  "MaxResults": number,
  "NextToken": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CrawlerName

The name of the crawler whose runs you want to retrieve.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Filters

Filters the crawls by the criteria you specify in a list of `CrawlsFilter` objects.

Type: Array of [CrawlsFilter](#) objects

Required: No

MaxResults

The maximum number of results to return. The default is 20, and maximum is 100.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

NextToken

A continuation token, if this is a continuation call.

Type: String

Required: No

Response Syntax

```
{
  "Crawls": [
    {
      "CrawlId": "string",
      "DPUHour": number,
      "EndTime": number,
      "ErrorMessage": "string",
      "LogGroup": "string",
      "LogStream": "string",
      "MessagePrefix": "string",
```

```
    "StartTime": number,
    "State": "string",
    "Summary": "string"
  }
],
"NextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Crawls

A list of `CrawlerHistory` objects representing the crawl runs that meet your criteria.

Type: Array of [CrawlerHistory](#) objects

NextToken

A continuation token for paginating the returned list of tokens, returned if the current segment of the list is not the last.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListCustomEntityTypes

Lists all the custom patterns that have been created.

Request Syntax

```
{
  "MaxResults": number,
  "NextToken": "string",
  "Tags": {
    "string" : "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

MaxResults

The maximum number of results to return.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

NextToken

A paginated token to offset the results.

Type: String

Required: No

Tags

A list of key-value pair tags.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Required: No

Response Syntax

```
{
  "CustomEntityTypes": [
    {
      "ContextWords": [ "string" ],
      "Name": "string",
      "RegexString": "string"
    }
  ],
  "NextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

CustomEntityTypes

A list of CustomEntityType objects representing custom patterns.

Type: Array of [CustomEntityType](#) objects

NextToken

A pagination token, if more results are available.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListDataQualityResults

Returns all data quality execution results for your account.

Request Syntax

```
{
  "Filter": {
    "DataSource": {
      "GlueTable": {
        "AdditionalOptions": {
          "string" : "string"
        },
        "CatalogId": "string",
        "ConnectionName": "string",
        "DatabaseName": "string",
        "TableName": "string"
      }
    },
    "JobName": "string",
    "JobRunId": "string",
    "StartedAfter": number,
    "StartedBefore": number
  },
  "MaxResults": number,
  "NextToken": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Filter

The filter criteria.

Type: [DataQualityResultFilterCriteria](#) object

Required: No

MaxResults

The maximum number of results to return.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

NextToken

A paginated token to offset the results.

Type: String

Required: No

Response Syntax

```
{
  "NextToken": "string",
  "Results": [
    {
      "DataSource": {
        "GlueTable": {
          "AdditionalOptions": {
            "string" : "string"
          },
          "CatalogId": "string",
          "ConnectionName": "string",
          "DatabaseName": "string",
          "TableName": "string"
        }
      },
      "JobName": "string",
      "JobRunId": "string",
      "ResultId": "string",
      "StartedOn": number
    }
  ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken

A pagination token, if more results are available.

Type: String

Results

A list of `DataQualityResultDescription` objects.

Type: Array of [DataQualityResultDescription](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListDataQualityRuleRecommendationRuns

Lists the recommendation runs meeting the filter criteria.

Request Syntax

```
{
  "Filter": {
    "DataSource": {
      "GlueTable": {
        "AdditionalOptions": {
          "string" : "string"
        },
        "CatalogId": "string",
        "ConnectionName": "string",
        "DatabaseName": "string",
        "TableName": "string"
      }
    },
    "StartedAfter": number,
    "StartedBefore": number
  },
  "MaxResults": number,
  "NextToken": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Filter

The filter criteria.

Type: [DataQualityRuleRecommendationRunFilter](#) object

Required: No

MaxResults

The maximum number of results to return.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

NextToken

A paginated token to offset the results.

Type: String

Required: No

Response Syntax

```
{
  "NextToken": "string",
  "Runs": [
    {
      "DataSource": {
        "GlueTable": {
          "AdditionalOptions": {
            "string": "string"
          },
          "CatalogId": "string",
          "ConnectionName": "string",
          "DatabaseName": "string",
          "TableName": "string"
        }
      },
      "RunId": "string",
      "StartedOn": number,
      "Status": "string"
    }
  ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken

A pagination token, if more results are available.

Type: String

Runs

A list of `DataQualityRuleRecommendationRunDescription` objects.

Type: Array of [DataQualityRuleRecommendationRunDescription](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)

- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListDataQualityRulesetEvaluationRuns

Lists all the runs meeting the filter criteria, where a ruleset is evaluated against a data source.

Request Syntax

```
{
  "Filter": {
    "DataSource": {
      "GlueTable": {
        "AdditionalOptions": {
          "string" : "string"
        },
        "CatalogId": "string",
        "ConnectionName": "string",
        "DatabaseName": "string",
        "TableName": "string"
      }
    },
    "StartedAfter": number,
    "StartedBefore": number
  },
  "MaxResults": number,
  "NextToken": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Filter

The filter criteria.

Type: [DataQualityRulesetEvaluationRunFilter](#) object

Required: No

MaxResults

The maximum number of results to return.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

NextToken

A paginated token to offset the results.

Type: String

Required: No

Response Syntax

```
{
  "NextToken": "string",
  "Runs": [
    {
      "DataSource": {
        "GlueTable": {
          "AdditionalOptions": {
            "string": "string"
          },
          "CatalogId": "string",
          "ConnectionName": "string",
          "DatabaseName": "string",
          "TableName": "string"
        }
      },
      "RunId": "string",
      "StartedOn": number,
      "Status": "string"
    }
  ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken

A pagination token, if more results are available.

Type: String

Runs

A list of `DataQualityRulesetEvaluationRunDescription` objects representing data quality ruleset runs.

Type: Array of [DataQualityRulesetEvaluationRunDescription](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)

- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListDataQualityRulesets

Returns a paginated list of rulesets for the specified list of AWS Glue tables.

Request Syntax

```
{
  "Filter": {
    "CreatedAfter": number,
    "CreatedBefore": number,
    "Description": "string",
    "LastModifiedAfter": number,
    "LastModifiedBefore": number,
    "Name": "string",
    "TargetTable": {
      "CatalogId": "string",
      "DatabaseName": "string",
      "TableName": "string"
    }
  },
  "MaxResults": number,
  "NextToken": "string",
  "Tags": {
    "string": "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Filter

The filter criteria.

Type: [DataQualityRulesetFilterCriteria](#) object

Required: No

MaxResults

The maximum number of results to return.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

NextToken

A paginated token to offset the results.

Type: String

Required: No

Tags

A list of key-value pair tags.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Required: No

Response Syntax

```
{
  "NextToken": "string",
  "Rulesets": [
    {
      "CreatedOn": number,
      "Description": "string",
      "LastModifiedOn": number,
      "Name": "string",
      "RecommendationRunId": "string",
      "RuleCount": number,
      "TargetTable": {
        "CatalogId": "string",
        "DatabaseName": "string",
        "TableName": "string"
      }
    }
  ]
}
```

```
}  
  }  
] }  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken

A pagination token, if more results are available.

Type: String

Rulesets

A paginated list of rulesets for the specified list of AWS Glue tables.

Type: Array of [DataQualityRulesetListDetails](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListDevEndpoints

Retrieves the names of all `DevEndpoint` resources in this AWS account, or the resources with the specified tag. This operation allows you to see which resources are available in your account, and their names.

This operation takes the optional `Tags` field, which you can use as a filter on the response so that tagged resources can be retrieved as a group. If you choose to use tags filtering, only resources with the tag are retrieved.

Request Syntax

```
{
  "MaxResults": number,
  "NextToken": "string",
  "Tags": {
    "string" : "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

MaxResults

The maximum size of a list to return.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

NextToken

A continuation token, if this is a continuation request.

Type: String

Required: No

Tags

Specifies to return only these tagged resources.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Required: No

Response Syntax

```
{
  "DevEndpointNames": [ "string" ],
  "NextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

DevEndpointNames

The names of all the DevEndpoints in the account, or the DevEndpoints with the specified tags.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

NextToken

A continuation token, if the returned list does not contain the last metric available.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)

- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListJobs

Retrieves the names of all job resources in this AWS account, or the resources with the specified tag. This operation allows you to see which resources are available in your account, and their names.

This operation takes the optional `Tags` field, which you can use as a filter on the response so that tagged resources can be retrieved as a group. If you choose to use tags filtering, only resources with the tag are retrieved.

Request Syntax

```
{
  "MaxResults": number,
  "NextToken": "string",
  "Tags": {
    "string" : "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

[MaxResults](#)

The maximum size of a list to return.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

[NextToken](#)

A continuation token, if this is a continuation request.

Type: String

Required: No

Tags

Specifies to return only these tagged resources.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Required: No

Response Syntax

```
{
  "JobNames": [ "string" ],
  "NextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

JobNames

The names of all jobs in the account, or the jobs with the specified tags.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

NextToken

A continuation token, if the returned list does not contain the last metric available.

Type: String

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)

- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListMLTransforms

Retrieves a sortable, filterable list of existing AWS Glue machine learning transforms in this AWS account, or the resources with the specified tag. This operation takes the optional `Tags` field, which you can use as a filter of the responses so that tagged resources can be retrieved as a group. If you choose to use tag filtering, only resources with the tags are retrieved.

Request Syntax

```
{
  "Filter": {
    "CreatedAfter": number,
    "CreatedBefore": number,
    "GlueVersion": "string",
    "LastModifiedAfter": number,
    "LastModifiedBefore": number,
    "Name": "string",
    "Schema": [
      {
        "DataType": "string",
        "Name": "string"
      }
    ],
    "Status": "string",
    "TransformType": "string"
  },
  "MaxResults": number,
  "NextToken": "string",
  "Sort": {
    "Column": "string",
    "SortDirection": "string"
  },
  "Tags": {
    "string" : "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Filter

A `TransformFilterCriteria` used to filter the machine learning transforms.

Type: [TransformFilterCriteria](#) object

Required: No

MaxResults

The maximum size of a list to return.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

NextToken

A continuation token, if this is a continuation request.

Type: String

Required: No

Sort

A `TransformSortCriteria` used to sort the machine learning transforms.

Type: [TransformSortCriteria](#) object

Required: No

Tags

Specifies to return only these tagged resources.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Required: No

Response Syntax

```
{
  "NextToken": "string",
  "TransformIds": [ "string" ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken

A continuation token, if the returned list does not contain the last metric available.

Type: String

TransformIds

The identifiers of all the machine learning transforms in the account, or the machine learning transforms with the specified tags.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListRegistries

Returns a list of registries that you have created, with minimal registry information. Registries in the `Deleting` status will not be included in the results. Empty results will be returned if there are no registries available.

Request Syntax

```
{
  "MaxResults": number,
  "NextToken": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

MaxResults

Maximum number of results required per page. If the value is not supplied, this will be defaulted to 25 per page.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

NextToken

A continuation token, if this is a continuation call.

Type: String

Required: No

Response Syntax

```
{
  "NextToken": "string",
}
```

```
"Registries": [  
  {  
    "CreatedTime": "string",  
    "Description": "string",  
    "RegistryArn": "string",  
    "RegistryName": "string",  
    "Status": "string",  
    "UpdateTime": "string"  
  }  
]
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

[NextToken](#)

A continuation token for paginating the returned list of tokens, returned if the current segment of the list is not the last.

Type: String

[Registries](#)

An array of RegistryDetailedListItem objects containing minimal details of each registry.

Type: Array of [RegistryListItem](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListSchemas

Returns a list of schemas with minimal details. Schemas in Deleting status will not be included in the results. Empty results will be returned if there are no schemas available.

When the RegistryId is not provided, all the schemas across registries will be part of the API response.

Request Syntax

```
{
  "MaxResults": number,
  "NextToken": "string",
  "RegistryId": {
    "RegistryArn": "string",
    "RegistryName": "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

MaxResults

Maximum number of results required per page. If the value is not supplied, this will be defaulted to 25 per page.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

NextToken

A continuation token, if this is a continuation call.

Type: String

Required: No

RegistryId

A wrapper structure that may contain the registry name and Amazon Resource Name (ARN).

Type: [RegistryId](#) object

Required: No

Response Syntax

```
{
  "NextToken": "string",
  "Schemas": [
    {
      "CreatedTime": "string",
      "Description": "string",
      "RegistryName": "string",
      "SchemaArn": "string",
      "SchemaName": "string",
      "SchemaStatus": "string",
      "UpdatedTime": "string"
    }
  ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken

A continuation token for paginating the returned list of tokens, returned if the current segment of the list is not the last.

Type: String

Schemas

An array of `SchemaListItem` objects containing details of each schema.

Type: Array of [SchemaListItem](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)

- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListSchemaVersions

Returns a list of schema versions that you have created, with minimal information. Schema versions in Deleted status will not be included in the results. Empty results will be returned if there are no schema versions available.

Request Syntax

```
{
  "MaxResults": number,
  "NextToken": "string",
  "SchemaId": {
    "RegistryName": "string",
    "SchemaArn": "string",
    "SchemaName": "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

MaxResults

Maximum number of results required per page. If the value is not supplied, this will be defaulted to 25 per page.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

NextToken

A continuation token, if this is a continuation call.

Type: String

Required: No

Schemald

This is a wrapper structure to contain schema identity fields. The structure contains:

- `Schemald$SchemaArn`: The Amazon Resource Name (ARN) of the schema. Either `SchemaArn` or `SchemaName` and `RegistryName` has to be provided.
- `Schemald$SchemaName`: The name of the schema. Either `SchemaArn` or `SchemaName` and `RegistryName` has to be provided.

Type: [Schemald](#) object

Required: Yes

Response Syntax

```
{
  "NextToken": "string",
  "Schemas": [
    {
      "CreatedTime": "string",
      "SchemaArn": "string",
      "SchemaVersionId": "string",
      "Status": "string",
      "VersionNumber": number
    }
  ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken

A continuation token for paginating the returned list of tokens, returned if the current segment of the list is not the last.

Type: String

[Schemas](#)

An array of `SchemaVersionList` objects containing details of each schema version.

Type: Array of [SchemaVersionListItem](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)

- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListSessions

Retrieve a list of sessions.

Request Syntax

```
{
  "MaxResults": number,
  "NextToken": "string",
  "RequestOrigin": "string",
  "Tags": {
    "string" : "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

[MaxResults](#)

The maximum number of results.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

[NextToken](#)

The token for the next set of results, or null if there are no more result.

Type: String

Length Constraints: Maximum length of 400000.

Required: No

[RequestOrigin](#)

The origin of the request.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[\.\-_\A-Za-z0-9]+`

Required: No

Tags

Tags belonging to the session.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Required: No

Response Syntax

```
{
  "Ids": [ "string" ],
  "NextToken": "string",
  "Sessions": [
    {
      "Command": {
        "Name": "string",
        "PythonVersion": "string"
      },
      "CompletedOn": number,
      "Connections": {
        "Connections": [ "string" ]
      },
      "CreatedOn": number,
      "DefaultArguments": {
        "string" : "string"
      },
      "Description": "string",
      "DPUSecods": number,
```


Type: Array of [Session](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)

- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListStatements

Lists statements for the session.

Request Syntax

```
{  
  "NextToken": "string",  
  "RequestOrigin": "string",  
  "SessionId": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

[NextToken](#)

A continuation token, if this is a continuation call.

Type: String

Length Constraints: Maximum length of 400000.

Required: No

[RequestOrigin](#)

The origin of the request to list statements.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[\.\-_\A-Za-z0-9]+`

Required: No

[SessionId](#)

The Session ID of the statements.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "NextToken": "string",
  "Statements": [
    {
      "Code": "string",
      "CompletedOn": number,
      "Id": number,
      "Output": {
        "Data": {
          "TextPlain": "string"
        },
        "ErrorName": "string",
        "ErrorValue": "string",
        "ExecutionCount": number,
        "Status": "string",
        "Traceback": [ "string" ]
      },
      "Progress": number,
      "StartedOn": number,
      "State": "string"
    }
  ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken

A continuation token, if not all statements have yet been returned.

Type: String

Length Constraints: Maximum length of 400000.

Statements

Returns the list of statements.

Type: Array of [Statement](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

IllegalSessionStateException

The session is in an invalid state to perform a requested operation.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListTableOptimizerRuns

Lists the history of previous optimizer runs for a specific table.

Request Syntax

```
{
  "CatalogId": "string",
  "DatabaseName": "string",
  "MaxResults": number,
  "NextToken": "string",
  "TableName": "string",
  "Type": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The Catalog ID of the table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

DatabaseName

The name of the database in the catalog in which the table resides.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

NextToken

A continuation token for paginating the returned list of optimizer runs, returned if the current segment of the list is not the last.

Type: String

TableName

The name of the table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

TableOptimizerRuns

A list of the optimizer runs associated with a table.

Type: Array of [TableOptimizerRun](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListTriggers

Retrieves the names of all trigger resources in this AWS account, or the resources with the specified tag. This operation allows you to see which resources are available in your account, and their names.

This operation takes the optional `Tags` field, which you can use as a filter on the response so that tagged resources can be retrieved as a group. If you choose to use tags filtering, only resources with the tag are retrieved.

Request Syntax

```
{
  "DependentJobName": "string",
  "MaxResults": number,
  "NextToken": "string",
  "Tags": {
    "string" : "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

DependentJobName

The name of the job for which to retrieve triggers. The trigger that can start this job is returned. If there is no such trigger, all triggers are returned.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

MaxResults

The maximum size of a list to return.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 200.

Required: No

NextToken

A continuation token, if this is a continuation request.

Type: String

Required: No

Tags

Specifies to return only these tagged resources.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Required: No

Response Syntax

```
{  
  "NextToken": "string",  
  "TriggerNames": [ "string" ]  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken

A continuation token, if the returned list does not contain the last metric available.

Type: String

TriggerNames

The names of all triggers in the account, or the triggers with the specified tags.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListUsageProfiles

List all the AWS Glue usage profiles.

Request Syntax

```
{  
  "MaxResults": number,  
  "NextToken": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

MaxResults

The maximum number of usage profiles to return in a single response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 200.

Required: No

NextToken

A continuation token, included if this is a continuation call.

Type: String

Length Constraints: Maximum length of 400000.

Required: No

Response Syntax

```
{  
  "NextToken": "string",  
}
```

```
"Profiles": [  
  {  
    "CreatedOn": number,  
    "Description": "string",  
    "LastModifiedOn": number,  
    "Name": "string"  
  }  
]
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken

A continuation token, present if the current list segment is not the last.

Type: String

Length Constraints: Maximum length of 400000.

Profiles

A list of usage profile (`UsageProfileDefinition`) objects.

Type: Array of [UsageProfileDefinition](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationNotSupportedException

The operation is not available in the region.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ListWorkflows

Lists names of workflows created in the account.

Request Syntax

```
{
  "MaxResults": number,
  "NextToken": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

MaxResults

The maximum size of a list to return.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 25.

Required: No

NextToken

A continuation token, if this is a continuation request.

Type: String

Required: No

Response Syntax

```
{
  "NextToken": "string",
  "Workflows": [ "string" ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken

A continuation token, if not all workflow names have been returned.

Type: String

Workflows

List of names of workflows in the account.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 25 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

[DataCatalogEncryptionSettings](#)

The security configuration to set.

Type: [DataCatalogEncryptionSettings](#) object

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)

- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

PutResourcePolicy

Sets the Data Catalog resource policy for access control.

Request Syntax

```
{
  "EnableHybrid": "string",
  "PolicyExistsCondition": "string",
  "PolicyHashCondition": "string",
  "PolicyInJson": "string",
  "ResourceArn": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

EnableHybrid

If 'TRUE', indicates that you are using both methods to grant cross-account access to Data Catalog resources:

- By directly updating the resource policy with `PutResourcePolicy`
- By using the **Grant permissions** command on the AWS Management Console.

Must be set to 'TRUE' if you have already used the Management Console to grant cross-account access, otherwise the call fails. Default is 'FALSE'.

Type: String

Valid Values: TRUE | FALSE

Required: No

PolicyExistsCondition

A value of `MUST_EXIST` is used to update a policy. A value of `NOT_EXIST` is used to create a new policy. If a value of `NONE` or a null value is used, the call does not depend on the existence of a policy.

Type: String

Valid Values: MUST_EXIST | NOT_EXIST | NONE

Required: No

PolicyHashCondition

The hash value returned when the previous policy was set using `PutResourcePolicy`. Its purpose is to prevent concurrent modifications of a policy. Do not use this parameter if no previous policy has been set.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

PolicyInJson

Contains the policy document to set, in JSON format.

Type: String

Length Constraints: Minimum length of 2.

Required: Yes

ResourceArn

Do not use. For internal use only.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 10240.

Pattern: `arn:(aws|aws-us-gov|aws-cn):glue:.*`

Required: No

Response Syntax

```
{
```

```
"PolicyHash": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

PolicyHash

A hash of the policy that has just been set. This must be included in a subsequent call that overwrites or updates this policy.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ConditionCheckFailureException

A specified condition was not satisfied.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

PutSchemaVersionMetadata

Puts the metadata key value pair for a specified schema version ID. A maximum of 10 key value pairs will be allowed per schema version. They can be added over one or more calls.

Request Syntax

```
{
  "MetadataKeyValue": {
    "MetadataKey": "string",
    "MetadataValue": "string"
  },
  "SchemaId": {
    "RegistryName": "string",
    "SchemaArn": "string",
    "SchemaName": "string"
  },
  "SchemaVersionId": "string",
  "SchemaVersionNumber": {
    "LatestVersion": boolean,
    "VersionNumber": number
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

MetadataKeyValue

The metadata key's corresponding value.

Type: [MetadataKeyValuePair](#) object

Required: Yes

SchemaId

The unique ID for the schema.

Type: [SchemaId](#) object

Required: No

SchemaVersionId

The unique version ID of the schema version.

Type: String

Length Constraints: Fixed length of 36.

Pattern: `[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

Required: No

SchemaVersionNumber

The version number of the schema.

Type: [SchemaVersionNumber](#) object

Required: No

Response Syntax

```
{
  "LatestVersion": boolean,
  "MetadataKey": "string",
  "MetadataValue": "string",
  "RegistryName": "string",
  "SchemaArn": "string",
  "SchemaName": "string",
  "SchemaVersionId": "string",
  "VersionNumber": number
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

LatestVersion

The latest version of the schema.

Type: Boolean

MetadataKey

The metadata key.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: [a-zA-Z0-9+-. _./@]+

MetadataValue

The value of the metadata key.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: [a-zA-Z0-9+-. _./@]+

RegistryName

The name for the registry.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: [a-zA-Z0-9-_\$#.]+

SchemaArn

The Amazon Resource Name (ARN) for the schema.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 10240.

Pattern: arn:(aws|aws-us-gov|aws-cn):glue:.*

SchemaName

The name for the schema.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: [a-zA-Z0-9-_\$#.]+

SchemaVersionId

The unique version ID of the schema version.

Type: String

Length Constraints: Fixed length of 36.

Pattern: [a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}

VersionNumber

The version number of the schema.

Type: Long

Valid Range: Minimum value of 1. Maximum value of 100000.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

PutWorkflowRunProperties

Puts the specified workflow run properties for the given workflow run. If a property already exists for the specified run, then it overrides the value otherwise adds the property to existing properties.

Request Syntax

```
{
  "Name": "string",
  "RunId": "string",
  "RunProperties": {
    "string" : "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Name

Name of the workflow which was run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

RunId

The ID of the workflow run for which the run properties should be updated.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

RunProperties

The properties to put for the specified run.

Type: String to string map

Key Length Constraints: Minimum length of 1. Maximum length of 255.

Key Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

QuerySchemaVersionMetadata

Queries for the schema version metadata information.

Request Syntax

```
{
  "MaxResults": number,
  "MetadataList": [
    {
      "MetadataKey": "string",
      "MetadataValue": "string"
    }
  ],
  "NextToken": "string",
  "SchemaId": {
    "RegistryName": "string",
    "SchemaArn": "string",
    "SchemaName": "string"
  },
  "SchemaVersionId": "string",
  "SchemaVersionNumber": {
    "LatestVersion": boolean,
    "VersionNumber": number
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

MaxResults

Maximum number of results required per page. If the value is not supplied, this will be defaulted to 25 per page.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 50.

Required: No

MetadataList

Search key-value pairs for metadata, if they are not provided all the metadata information will be fetched.

Type: Array of [MetadataKeyValuePair](#) objects

Required: No

NextToken

A continuation token, if this is a continuation call.

Type: String

Required: No

Schemald

A wrapper structure that may contain the schema name and Amazon Resource Name (ARN).

Type: [Schemald](#) object

Required: No

SchemaVersionId

The unique version ID of the schema version.

Type: String

Length Constraints: Fixed length of 36.

Pattern: `[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

Required: No

SchemaVersionNumber

The version number of the schema.

Type: [SchemaVersionNumber](#) object

Required: No

Response Syntax

```
{
  "MetadataInfoMap": {
    "string" : {
      "CreatedTime": "string",
      "MetadataValue": "string",
      "OtherMetadataValueList": [
        {
          "CreatedTime": "string",
          "MetadataValue": "string"
        }
      ]
    }
  },
  "NextToken": "string",
  "SchemaVersionId": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

MetadataInfoMap

A map of a metadata key and associated values.

Type: String to [MetadataInfo](#) object map

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Key Pattern: [a-zA-Z0-9+-. _./@]+

NextToken

A continuation token for paginating the returned list of tokens, returned if the current segment of the list is not the last.

Type: String

SchemaVersionId

The unique version ID of the schema version.

Type: String

Length Constraints: Fixed length of 36.

Pattern: `[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)

- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

RegisterSchemaVersion

Adds a new version to the existing schema. Returns an error if new version of schema does not meet the compatibility requirements of the schema set. This API will not create a new schema set and will return a 404 error if the schema set is not already present in the Schema Registry.

If this is the first schema definition to be registered in the Schema Registry, this API will store the schema version and return immediately. Otherwise, this call has the potential to run longer than other operations due to compatibility modes. You can call the `GetSchemaVersion` API with the `SchemaVersionId` to check compatibility modes.

If the same schema definition is already stored in Schema Registry as a version, the schema ID of the existing schema is returned to the caller.

Request Syntax

```
{
  "SchemaDefinition": "string",
  "SchemaId": {
    "RegistryName": "string",
    "SchemaArn": "string",
    "SchemaName": "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

SchemaDefinition

The schema definition using the `DataFormat` setting for the `SchemaName`.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 170000.

Pattern: `.*\S.*`

Required: Yes

Schemald

This is a wrapper structure to contain schema identity fields. The structure contains:

- `Schemald$SchemaArn`: The Amazon Resource Name (ARN) of the schema. Either `SchemaArn` or `SchemaName` and `RegistryName` has to be provided.
- `Schemald$SchemaName`: The name of the schema. Either `SchemaArn` or `SchemaName` and `RegistryName` has to be provided.

Type: [Schemald](#) object

Required: Yes

Response Syntax

```
{  
  "SchemaVersionId": "string",  
  "Status": "string",  
  "VersionNumber": number  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

SchemaVersionId

The unique ID that represents the version of this schema.

Type: String

Length Constraints: Fixed length of 36.

Pattern: `[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

Status

The status of the schema version.

Type: String

Valid Values: AVAILABLE | PENDING | FAILURE | DELETING

VersionNumber

The version of this schema (for sync flow only, in case this is the first version).

Type: Long

Valid Range: Minimum value of 1. Maximum value of 100000.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

RemoveSchemaVersionMetadata

Removes a key value pair from the schema version metadata for the specified schema version ID.

Request Syntax

```
{
  "MetadataKeyValue": {
    "MetadataKey": "string",
    "MetadataValue": "string"
  },
  "SchemaId": {
    "RegistryName": "string",
    "SchemaArn": "string",
    "SchemaName": "string"
  },
  "SchemaVersionId": "string",
  "SchemaVersionNumber": {
    "LatestVersion": boolean,
    "VersionNumber": number
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

MetadataKeyValue

The value of the metadata key.

Type: [MetadataKeyValuePair](#) object

Required: Yes

SchemaId

A wrapper structure that may contain the schema name and Amazon Resource Name (ARN).

Type: [SchemaId](#) object

Required: No

SchemaVersionId

The unique version ID of the schema version.

Type: String

Length Constraints: Fixed length of 36.

Pattern: `[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

Required: No

SchemaVersionNumber

The version number of the schema.

Type: [SchemaVersionNumber](#) object

Required: No

Response Syntax

```
{
  "LatestVersion": boolean,
  "MetadataKey": "string",
  "MetadataValue": "string",
  "RegistryName": "string",
  "SchemaArn": "string",
  "SchemaName": "string",
  "SchemaVersionId": "string",
  "VersionNumber": number
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

LatestVersion

The latest version of the schema.

Type: Boolean

MetadataKey

The metadata key.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: [a-zA-Z0-9+-. _./@]+

MetadataValue

The value of the metadata key.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: [a-zA-Z0-9+-. _./@]+

RegistryName

The name of the registry.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: [a-zA-Z0-9-_\$#.]+

SchemaArn

The Amazon Resource Name (ARN) of the schema.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 10240.

Pattern: arn:(aws|aws-us-gov|aws-cn):glue:.*

SchemaName

The name of the schema.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: [a-zA-Z0-9-_\$#.]+

SchemaVersionId

The version ID for the schema version.

Type: String

Length Constraints: Fixed length of 36.

Pattern: [a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}

VersionNumber

The version number of the schema.

Type: Long

Valid Range: Minimum value of 1. Maximum value of 100000.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ResetJobBookmark

Resets a bookmark entry.

For more information about enabling and using job bookmarks, see:

- [Tracking processed data using job bookmarks](#)
- [Job parameters used by AWS Glue](#)
- [Job structure](#)

Request Syntax

```
{  
  "JobName": "string",  
  "RunId": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

[JobName](#)

The name of the job in question.

Type: String

Required: Yes

[RunId](#)

The unique run identifier associated with this job run.

Type: String

Required: No

Response Syntax

```
{
  "JobBookmarkEntry": {
    "Attempt": number,
    "JobBookmark": "string",
    "JobName": "string",
    "PreviousRunId": "string",
    "Run": number,
    "RunId": "string",
    "Version": number
  }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

JobBookmarkEntry

The reset bookmark entry.

Type: [JobBookmarkEntry](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ResumeWorkflowRun

Restarts selected nodes of a previous partially completed workflow run and resumes the workflow run. The selected nodes and all nodes that are downstream from the selected nodes are run.

Request Syntax

```
{
  "Name": "string",
  "NodeIds": [ "string" ],
  "RunId": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Name

The name of the workflow to resume.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

NodeIds

A list of the node IDs for the nodes you want to restart. The nodes that are to be restarted must have a run attempt in the original run.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

RunId

The ID of the workflow run to resume.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "NodeIds": [ "string" ],
  "RunId": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NodeIds

A list of the node IDs for the nodes that were actually restarted.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

RunId

The new ID assigned to the resumed workflow run. Each resume of a workflow run will have a new run ID.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ConcurrentRunsExceededException

Too many jobs are being run concurrently.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

IllegalWorkflowStateException

The workflow is in an invalid state to perform a requested operation.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

RunStatement

Executes the statement.

Request Syntax

```
{  
  "Code": "string",  
  "RequestOrigin": "string",  
  "SessionId": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Code

The statement code to be run.

Type: String

Length Constraints: Maximum length of 68000.

Required: Yes

RequestOrigin

The origin of the request.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[\.\-_\A-Za-z0-9]+`

Required: No

SessionId

The Session Id of the statement to be run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{  
  "Id": number  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Id

Returns the Id of the statement that was run.

Type: Integer

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

IllegalSessionStateException

The session is in an invalid state to perform a requested operation.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

ValidationException

A value could not be validated.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)

- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

SearchTables

Searches a set of tables based on properties in the table metadata as well as on the parent database. You can search against text or filter conditions.

You can only get tables that you have access to based on the security policies defined in Lake Formation. You need at least a read-only access to the table for it to be returned. If you do not have access to all the columns in the table, these columns will not be searched against when returning the list of tables back to you. If you have access to the columns but not the data in the columns, those columns and the associated metadata for those columns will be included in the search.

Request Syntax

```
{
  "CatalogId": "string",
  "Filters": [
    {
      "Comparator": "string",
      "Key": "string",
      "Value": "string"
    }
  ],
  "MaxResults": number,
  "NextToken": "string",
  "ResourceShareType": "string",
  "SearchText": "string",
  "SortCriteria": [
    {
      "FieldName": "string",
      "Sort": "string"
    }
  ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

A unique identifier, consisting of `account_id` .

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Filters

A list of key-value pairs, and a comparator used to filter the search results. Returns all entities matching the predicate.

The `Comparator` member of the `PropertyPredicate` struct is used only for time fields, and can be omitted for other field types. Also, when comparing string values, such as when `Key=Name`, a fuzzy match algorithm is used. The `Key` field (for example, the value of the `Name` field) is split on certain punctuation characters, for example, `-`, `:`, `#`, etc. into tokens. Then each token is exact-match compared with the `Value` member of `PropertyPredicate`. For example, if `Key=Name` and `Value=link`, tables named `customer-link` and `xx-link-yy` are returned, but `xxlinkyy` is not returned.

Type: Array of [PropertyPredicate](#) objects

Required: No

MaxResults

The maximum number of tables to return in a single response.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 1000.

Required: No

NextToken

A continuation token, included if this is a continuation call.

Type: String

Required: No

ResourceShareType

Allows you to specify that you want to search the tables shared with your account. The allowable values are FOREIGN or ALL.

- If set to FOREIGN, will search the tables shared with your account.
- If set to ALL, will search the tables shared with your account, as well as the tables in your local account.

Type: String

Valid Values: FOREIGN | ALL | FEDERATED

Required: No

SearchText

A string used for a text search.

Specifying a value in quotes filters based on an exact match to the value.

Type: String

Length Constraints: Maximum length of 1024.

Required: No

SortCriteria

A list of criteria for sorting the results by a field name, in an ascending or descending order.

Type: Array of [SortCriterion](#) objects

Array Members: Minimum number of 0 items. Maximum number of 1 item.

Required: No

Response Syntax

```
{
  "NextToken": "string",
  "TableList": [
    {
      "CatalogId": "string",
      "CreatedBy": "string",
```

```
"CreateTime": number,
"DatabaseName": "string",
"Description": "string",
"FederatedTable": {
  "ConnectionName": "string",
  "DatabaseIdentifier": "string",
  "Identifier": "string"
},
"IsMultiDialectView": boolean,
"IsRegisteredWithLakeFormation": boolean,
"LastAccessTime": number,
"LastAnalyzedTime": number,
"Name": "string",
"Owner": "string",
"Parameters": {
  "string" : "string"
},
"PartitionKeys": [
  {
    "Comment": "string",
    "Name": "string",
    "Parameters": {
      "string" : "string"
    },
    "Type": "string"
  }
],
"Retention": number,
"StorageDescriptor": {
  "AdditionalLocations": [ "string" ],
  "BucketColumns": [ "string" ],
  "Columns": [
    {
      "Comment": "string",
      "Name": "string",
      "Parameters": {
        "string" : "string"
      },
      "Type": "string"
    }
  ],
  "Compressed": boolean,
  "InputFormat": "string",
  "Location": "string",
```

```
"NumberOfBuckets": number,
"OutputFormat": "string",
"Parameters": {
  "string" : "string"
},
"SchemaReference": {
  "SchemaId": {
    "RegistryName": "string",
    "SchemaArn": "string",
    "SchemaName": "string"
  },
  "SchemaVersionId": "string",
  "SchemaVersionNumber": number
},
"SerdeInfo": {
  "Name": "string",
  "Parameters": {
    "string" : "string"
  },
  "SerializationLibrary": "string"
},
"SkewedInfo": {
  "SkewedColumnNames": [ "string" ],
  "SkewedColumnValueLocationMaps": {
    "string" : "string"
  },
  "SkewedColumnValues": [ "string" ]
},
"SortColumns": [
  {
    "Column": "string",
    "SortOrder": number
  }
],
"StoredAsSubDirectories": boolean
},
"TableType": "string",
"TargetTable": {
  "CatalogId": "string",
  "DatabaseName": "string",
  "Name": "string",
  "Region": "string"
},
"UpdateTime": number,
```

```
"VersionId": "string",
"ViewDefinition": {
  "Definer": "string",
  "IsProtected": boolean,
  "Representations": [
    {
      "Dialect": "string",
      "DialectVersion": "string",
      "IsStale": boolean,
      "ValidationConnection": "string",
      "ViewExpandedText": "string",
      "ViewOriginalText": "string"
    }
  ],
  "SubObjects": [ "string" ]
},
"ViewExpandedText": "string",
"ViewOriginalText": "string"
}
]
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

NextToken

A continuation token, present if the current list segment is not the last.

Type: String

TableList

A list of the requested Table objects. The SearchTables response returns only the tables that you have access to.

Type: Array of [Table](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

StartBlueprintRun

Starts a new run of the specified blueprint.

Request Syntax

```
{  
  "BlueprintName": "string",  
  "Parameters": "string",  
  "RoleArn": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

BlueprintName

The name of the blueprint.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[\.\-_\A-Za-z0-9]+`

Required: Yes

Parameters

Specifies the parameters as a `BlueprintParameters` object.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 131072.

Required: No

RoleArn

Specifies the IAM role used to create the workflow.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `arn:aws[^:]*:iam::[0-9]*:role/.+`

Required: Yes

Response Syntax

```
{  
  "RunId": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

RunId

The run ID for this blueprint run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

IllegalBlueprintStateException

The blueprint is in an invalid state to perform a requested operation.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)

- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

DatabaseName

The name of the database where the table resides.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Role

The IAM role that the service assumes to generate statistics.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

SampleSize

The percentage of rows used to generate statistics. If none is supplied, the entire table will be used to generate stats.

Type: Double

Valid Range: Minimum value of 0. Maximum value of 100.

Required: No

SecurityConfiguration

Name of the security configuration that is used to encrypt CloudWatch logs for the column stats task run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

TableName

The name of the table to generate statistics.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "ColumnStatisticsTaskRunId": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

ColumnStatisticsTaskRunId

The identifier for the column statistics task run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

ColumnStatisticsTaskRunningException

An exception thrown when you try to start another job while running a column stats generation job.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)

- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

StartCrawler

Starts a crawl using the specified crawler, regardless of what is scheduled. If the crawler is already running, returns a [CrawlerRunningException](#).

Request Syntax

```
{
  "Name": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Name

Name of the crawler to start.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

CrawlerRunningException

The operation cannot be performed because the crawler is already running.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

StartCrawlerSchedule

Changes the schedule state of the specified crawler to SCHEDULED, unless the crawler is already running or the schedule state is already SCHEDULED.

Request Syntax

```
{  
  "CrawlerName": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CrawlerName

Name of the crawler to schedule.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

NoScheduleException

There is no applicable schedule.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

SchedulerRunningException

The specified scheduler is already running.

HTTP Status Code: 400

SchedulerTransitioningException

The specified scheduler is transitioning.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

StartDataQualityRuleRecommendationRun

Starts a recommendation run that is used to generate rules when you don't know what rules to write. AWS Glue Data Quality analyzes the data and comes up with recommendations for a potential ruleset. You can then triage the ruleset and modify the generated ruleset to your liking.

Recommendation runs are automatically deleted after 90 days.

Request Syntax

```
{
  "ClientToken": "string",
  "CreatedRulesetName": "string",
  "DataSource": {
    "GlueTable": {
      "AdditionalOptions": {
        "string" : "string"
      },
      "CatalogId": "string",
      "ConnectionName": "string",
      "DatabaseName": "string",
      "TableName": "string"
    }
  },
  "NumberOfWorkers": number,
  "Role": "string",
  "Timeout": number
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

ClientToken

Used for idempotency and is recommended to be set to a random ID (such as a UUID) to avoid creating or starting multiple instances of the same resource.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

CreatedRulesetName

A name for the ruleset.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

DataSource

The data source (AWS Glue table) associated with this run.

Type: [DataSource](#) object

Required: Yes

NumberOfWorkers

The number of G.1X workers to be used in the run. The default is 5.

Type: Integer

Required: No

Role

An IAM role supplied to encrypt the results of the run.

Type: String

Required: Yes

Timeout

The timeout for a run in minutes. This is the maximum time that a run can consume resources before it is terminated and enters TIMEOUT status. The default is 2,880 minutes (48 hours).

Type: Integer

Valid Range: Minimum value of 1.

Required: No

Response Syntax

```
{  
  "RunId": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

RunId

The unique run identifier associated with this run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ConflictException

The `CreatePartitions` API was called on a table that has indexes enabled.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

StartDataQualityRulesetEvaluationRun

Once you have a ruleset definition (either recommended or your own), you call this operation to evaluate the ruleset against a data source (AWS Glue table). The evaluation computes results which you can retrieve with the `GetDataQualityResult` API.

Request Syntax

```
{
  "AdditionalDataSources": {
    "string": {
      "GlueTable": {
        "AdditionalOptions": {
          "string": "string"
        },
        "CatalogId": "string",
        "ConnectionName": "string",
        "DatabaseName": "string",
        "TableName": "string"
      }
    }
  },
  "AdditionalRunOptions": {
    "CloudWatchMetricsEnabled": boolean,
    "CompositeRuleEvaluationMethod": "string",
    "ResultsS3Prefix": "string"
  },
  "ClientToken": "string",
  "DataSource": {
    "GlueTable": {
      "AdditionalOptions": {
        "string": "string"
      },
      "CatalogId": "string",
      "ConnectionName": "string",
      "DatabaseName": "string",
      "TableName": "string"
    }
  },
  "NumberOfWorkers": number,
  "Role": "string",
  "RulesetNames": [ "string" ],
}
```

```
"Timeout": number
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

AdditionalDataSources

A map of reference strings to additional data sources you can specify for an evaluation run.

Type: String to [DataSource](#) object map

Key Length Constraints: Minimum length of 1. Maximum length of 255.

Key Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

AdditionalRunOptions

Additional run options you can specify for an evaluation run.

Type: [DataQualityEvaluationRunAdditionalRunOptions](#) object

Required: No

ClientToken

Used for idempotency and is recommended to be set to a random ID (such as a UUID) to avoid creating or starting multiple instances of the same resource.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

DataSource

The data source (AWS Glue table) associated with this run.

Type: [DataSource](#) object

Required: Yes

[NumberOfWorkers](#)

The number of G.1X workers to be used in the run. The default is 5.

Type: Integer

Required: No

[Role](#)

An IAM role supplied to encrypt the results of the run.

Type: String

Required: Yes

[RulesetNames](#)

A list of ruleset names.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 10 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

[Timeout](#)

The timeout for a run in minutes. This is the maximum time that a run can consume resources before it is terminated and enters TIMEOUT status. The default is 2,880 minutes (48 hours).

Type: Integer

Valid Range: Minimum value of 1.

Required: No

Response Syntax

```
{  
  "RunId": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

RunId

The unique run identifier associated with this run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ConflictException

The `CreatePartitions` API was called on a table that has indexes enabled.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

StartExportLabelsTaskRun

Begins an asynchronous task to export all labeled data for a particular transform. This task is the only label-related API call that is not part of the typical active learning workflow. You typically use `StartExportLabelsTaskRun` when you want to work with all of your existing labels at the same time, such as when you want to remove or change labels that were previously submitted as truth. This API operation accepts the `TransformId` whose labels you want to export and an Amazon Simple Storage Service (Amazon S3) path to export the labels to. The operation returns a `TaskRunId`. You can check on the status of your task run by calling the `GetMLTaskRun` API.

Request Syntax

```
{
  "OutputS3Path": "string",
  "TransformId": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

[OutputS3Path](#)

The Amazon S3 path where you export the labels.

Type: String

Required: Yes

[TransformId](#)

The unique identifier of the machine learning transform.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{  
  "TaskRunId": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

TaskRunId

The unique identifier for the task run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

StartImportLabelsTaskRun

Enables you to provide additional labels (examples of truth) to be used to teach the machine learning transform and improve its quality. This API operation is generally used as part of the active learning workflow that starts with the `StartMLLabelingSetGenerationTaskRun` call and that ultimately results in improving the quality of your machine learning transform.

After the `StartMLLabelingSetGenerationTaskRun` finishes, AWS Glue machine learning will have generated a series of questions for humans to answer. (Answering these questions is often called 'labeling' in the machine learning workflows). In the case of the `FindMatches` transform, these questions are of the form, "What is the correct way to group these rows together into groups composed entirely of matching records?" After the labeling process is finished, users upload their answers/labels with a call to `StartImportLabelsTaskRun`. After `StartImportLabelsTaskRun` finishes, all future runs of the machine learning transform use the new and improved labels and perform a higher-quality transformation.

By default, `StartMLLabelingSetGenerationTaskRun` continually learns from and combines all labels that you upload unless you set `Replace` to true. If you set `Replace` to true, `StartImportLabelsTaskRun` deletes and forgets all previously uploaded labels and learns only from the exact set that you upload. Replacing labels can be helpful if you realize that you previously uploaded incorrect labels, and you believe that they are having a negative effect on your transform quality.

You can check on the status of your task run by calling the `GetMLTaskRun` operation.

Request Syntax

```
{
  "InputS3Path": "string",
  "ReplaceAllLabels": boolean,
  "TransformId": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

InputS3Path

The Amazon Simple Storage Service (Amazon S3) path from where you import the labels.

Type: String

Required: Yes

ReplaceAllLabels

Indicates whether to overwrite your existing labels.

Type: Boolean

Required: No

TransformId

The unique identifier of the machine learning transform.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "TaskRunId": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

TaskRunId

The unique identifier for the task run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

StartJobRun

Starts a job run using a job definition.

Request Syntax

```
{
  "AllocatedCapacity": number,
  "Arguments": {
    "string" : "string"
  },
  "ExecutionClass": "string",
  "JobName": "string",
  "JobRunId": "string",
  "MaxCapacity": number,
  "NotificationProperty": {
    "NotifyDelayAfter": number
  },
  "NumberOfWorkers": number,
  "SecurityConfiguration": "string",
  "Timeout": number,
  "WorkerType": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

AllocatedCapacity

This field is deprecated. Use `MaxCapacity` instead.

The number of AWS Glue data processing units (DPUs) to allocate to this JobRun. You can allocate a minimum of 2 DPUs; the default is 10. A DPU is a relative measure of processing power that consists of 4 vCPUs of compute capacity and 16 GB of memory. For more information, see the [AWS Glue pricing page](#).

Type: Integer

Required: No

Arguments

The job arguments associated with this run. For this job run, they replace the default arguments set in the job definition itself.

You can specify arguments here that your own job-execution script consumes, as well as arguments that AWS Glue itself consumes.

Job arguments may be logged. Do not pass plaintext secrets as arguments. Retrieve secrets from a AWS Glue Connection, AWS Secrets Manager or other secret management mechanism if you intend to keep them within the Job.

For information about how to specify and consume your own Job arguments, see the [Calling AWS Glue APIs in Python](#) topic in the developer guide.

For information about the arguments you can provide to this field when configuring Spark jobs, see the [Special Parameters Used by AWS Glue](#) topic in the developer guide.

For information about the arguments you can provide to this field when configuring Ray jobs, see [Using job parameters in Ray jobs](#) in the developer guide.

Type: String to string map

Required: No

ExecutionClass

Indicates whether the job is run with a standard or flexible execution class. The standard execution-class is ideal for time-sensitive workloads that require fast job startup and dedicated resources.

The flexible execution class is appropriate for time-insensitive jobs whose start and completion times may vary.

Only jobs with AWS Glue version 3.0 and above and command type `glueetl` will be allowed to set `ExecutionClass` to FLEX. The flexible execution class is available for Spark jobs.

Type: String

Length Constraints: Maximum length of 16.

Valid Values: FLEX | STANDARD

Required: No

JobName

The name of the job definition to use.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

JobRunId

The ID of a previous JobRun to retry.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

MaxCapacity

For Glue version 1.0 or earlier jobs, using the standard worker type, the number of AWS Glue data processing units (DPUs) that can be allocated when this job runs. A DPU is a relative measure of processing power that consists of 4 vCPUs of compute capacity and 16 GB of memory. For more information, see the [AWS Glue pricing page](#).

For Glue version 2.0+ jobs, you cannot specify a Maximum capacity. Instead, you should specify a Worker type and the Number of workers.

Do not set MaxCapacity if using WorkerType and NumberOfWorkers.

The value that can be allocated for MaxCapacity depends on whether you are running a Python shell job, an Apache Spark ETL job, or an Apache Spark streaming ETL job:

- When you specify a Python shell job (`JobCommand.Name="pythonshell"`), you can allocate either 0.0625 or 1 DPU. The default is 0.0625 DPU.
- When you specify an Apache Spark ETL job (`JobCommand.Name="glueetl"`) or Apache Spark streaming ETL job (`JobCommand.Name="gluestreaming"`), you can allocate from 2 to 100 DPUs. The default is 10 DPUs. This job type cannot have a fractional DPU allocation.

Valid Range: Minimum value of 1.

Required: No

WorkerType

The type of predefined worker that is allocated when a job runs. Accepts a value of G.1X, G.2X, G.4X, G.8X or G.025X for Spark jobs. Accepts the value Z.2X for Ray jobs.

- For the G.1X worker type, each worker maps to 1 DPU (4 vCPUs, 16 GB of memory) with 84GB disk (approximately 34GB free), and provides 1 executor per worker. We recommend this worker type for workloads such as data transforms, joins, and queries, to offers a scalable and cost effective way to run most jobs.
- For the G.2X worker type, each worker maps to 2 DPU (8 vCPUs, 32 GB of memory) with 128GB disk (approximately 77GB free), and provides 1 executor per worker. We recommend this worker type for workloads such as data transforms, joins, and queries, to offers a scalable and cost effective way to run most jobs.
- For the G.4X worker type, each worker maps to 4 DPU (16 vCPUs, 64 GB of memory) with 256GB disk (approximately 235GB free), and provides 1 executor per worker. We recommend this worker type for jobs whose workloads contain your most demanding transforms, aggregations, joins, and queries. This worker type is available only for AWS Glue version 3.0 or later Spark ETL jobs in the following AWS Regions: US East (Ohio), US East (N. Virginia), US West (Oregon), Asia Pacific (Singapore), Asia Pacific (Sydney), Asia Pacific (Tokyo), Canada (Central), Europe (Frankfurt), Europe (Ireland), and Europe (Stockholm).
- For the G.8X worker type, each worker maps to 8 DPU (32 vCPUs, 128 GB of memory) with 512GB disk (approximately 487GB free), and provides 1 executor per worker. We recommend this worker type for jobs whose workloads contain your most demanding transforms, aggregations, joins, and queries. This worker type is available only for AWS Glue version 3.0 or later Spark ETL jobs, in the same AWS Regions as supported for the G.4X worker type.
- For the G.025X worker type, each worker maps to 0.25 DPU (2 vCPUs, 4 GB of memory) with 84GB disk (approximately 34GB free), and provides 1 executor per worker. We recommend this worker type for low volume streaming jobs. This worker type is only available for AWS Glue version 3.0 streaming jobs.
- For the Z.2X worker type, each worker maps to 2 M-DPU (8vCPUs, 64 GB of memory) with 128 GB disk (approximately 120GB free), and provides up to 8 Ray workers based on the autoscaler.

Type: String

Valid Values: Standard | G.1X | G.2X | G.025X | G.4X | G.8X | Z.2X

Required: No

Response Syntax

```
{  
  "JobRunId": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

JobRunId

The ID assigned to this job run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ConcurrentRunsExceededException

Too many jobs are being run concurrently.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

StartMLEvaluationTaskRun

Starts a task to estimate the quality of the transform.

When you provide label sets as examples of truth, AWS Glue machine learning uses some of those examples to learn from them. The rest of the labels are used as a test to estimate quality.

Returns a unique identifier for the run. You can call `GetMLTaskRun` to get more information about the stats of the `EvaluationTaskRun`.

Request Syntax

```
{  
  "TransformId": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

TransformId

The unique identifier of the machine learning transform.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{  
  "TaskRunId": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

TaskRunId

The unique identifier associated with this run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ConcurrentRunsExceededException

Too many jobs are being run concurrently.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

MLTransformNotReadyException

The machine learning transform is not ready to run.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

StartMLLabelingSetGenerationTaskRun

Starts the active learning workflow for your machine learning transform to improve the transform's quality by generating label sets and adding labels.

When the `StartMLLabelingSetGenerationTaskRun` finishes, AWS Glue will have generated a "labeling set" or a set of questions for humans to answer.

In the case of the `FindMatches` transform, these questions are of the form, "What is the correct way to group these rows together into groups composed entirely of matching records?"

After the labeling process is finished, you can upload your labels with a call to `StartImportLabelsTaskRun`. After `StartImportLabelsTaskRun` finishes, all future runs of the machine learning transform will use the new and improved labels and perform a higher-quality transformation.

Request Syntax

```
{
  "OutputS3Path": "string",
  "TransformId": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

OutputS3Path

The Amazon Simple Storage Service (Amazon S3) path where you generate the labeling set.

Type: String

Required: Yes

TransformId

The unique identifier of the machine learning transform.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{  
  "TaskRunId": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

[TaskRunId](#)

The unique run identifier that is associated with this task run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ConcurrentRunsExceededException

Too many jobs are being run concurrently.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

StartTrigger

Starts an existing trigger. See [Triggering Jobs](#) for information about how different types of trigger are started.

Request Syntax

```
{  
  "Name": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Name

The name of the trigger to start.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{  
  "Name": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Name

The name of the trigger that was started.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ConcurrentRunsExceededException

Too many jobs are being run concurrently.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

StartWorkflowRun

Starts a new run of the specified workflow.

Request Syntax

```
{
  "Name": "string",
  "RunProperties": {
    "string" : "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Name

The name of the workflow to start.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

RunProperties

The workflow run properties for the new workflow run.

Type: String to string map

Key Length Constraints: Minimum length of 1. Maximum length of 255.

Key Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Response Syntax

```
{  
  "RunId": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

RunId

An Id for the new run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ConcurrentRunsExceededException

Too many jobs are being run concurrently.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

StopColumnStatisticsTaskRun

Stops a task run for the specified table.

Request Syntax

```
{  
  "DatabaseName": "string",  
  "TableName": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

DatabaseName

The name of the database where the table resides.

Type: String

Required: Yes

TableName

The name of the table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ColumnStatisticsTaskNotRunningException

An exception thrown when you try to stop a task run when there is no task running.

HTTP Status Code: 400

ColumnStatisticsTaskStoppingException

An exception thrown when you try to stop a task run.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)

- [AWS SDK for Ruby V3](#)

StopCrawler

If the specified crawler is running, stops the crawl.

Request Syntax

```
{  
  "Name": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Name

Name of the crawler to stop.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

CrawlerNotRunningException

The specified crawler is not running.

HTTP Status Code: 400

CrawlerStoppingException

The specified crawler is stopping.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

SchedulerNotRunningException

The specified scheduler is not running.

HTTP Status Code: 400

SchedulerTransitioningException

The specified scheduler is transitioning.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

Response Syntax

```
{  
  "Id": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Id

Returns the Id of the stopped session.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

IllegalSessionStateException

The session is in an invalid state to perform a requested operation.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

StopTrigger

Stops a specified trigger.

Request Syntax

```
{  
  "Name": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Name

The name of the trigger to stop.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{  
  "Name": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Name

The name of the trigger that was stopped.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

StopWorkflowRun

Stops the execution of the specified workflow run.

Request Syntax

```
{  
  "Name": "string",  
  "RunId": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Name

The name of the workflow to stop.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

RunId

The ID of the workflow run to stop.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

IllegalWorkflowStateException

The workflow is in an invalid state to perform a requested operation.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)

- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

TagResource

Adds tags to a resource. A tag is a label you can assign to an AWS resource. In AWS Glue, you can tag only certain resources. For information about what resources you can tag, see [AWS Tags in AWS Glue](#).

In addition to the tagging permissions to call tag related APIs, you also need the `glue:GetConnection` permission to call tagging APIs on connections, and the `glue:GetDatabase` permission to call tagging APIs on databases.

Request Syntax

```
{
  "ResourceArn": "string",
  "TagsToAdd": {
    "string" : "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

[ResourceArn](#)

The ARN of the AWS Glue resource to which to add the tags. For more information about AWS Glue resource ARNs, see the [AWS Glue ARN string pattern](#).

Type: String

Length Constraints: Minimum length of 1. Maximum length of 10240.

Pattern: `arn:(aws|aws-us-gov|aws-cn):glue:.*`

Required: Yes

[TagsToAdd](#)

Tags to add to this resource.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 50 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

UntagResource

Removes tags from a resource.

Request Syntax

```
{  
  "ResourceArn": "string",  
  "TagsToRemove": [ "string" ]  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

ResourceArn

The Amazon Resource Name (ARN) of the resource from which to remove the tags.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 10240.

Pattern: `arn:(aws|aws-us-gov|aws-cn):glue:.*`

Required: Yes

TagsToRemove

Tags to remove from this resource.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Length Constraints: Minimum length of 1. Maximum length of 128.

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

UpdateBlueprint

Updates a registered blueprint.

Request Syntax

```
{  
  "BlueprintLocation": "string",  
  "Description": "string",  
  "Name": "string"  
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

BlueprintLocation

Specifies a path in Amazon S3 where the blueprint is published.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 8192.

Pattern: `^s3://([^\s/]+)/([^\s/]+)*([^\s/]+)$`

Required: Yes

Description

A description of the blueprint.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 512.

Required: No

Name

The name of the blueprint.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[\._\-A-Za-z0-9]+`

Required: Yes

Response Syntax

```
{  
  "Name": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Name

Returns the name of the blueprint that was updated.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

IllegalBlueprintStateException

The blueprint is in an invalid state to perform a requested operation.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

UpdateClassifier

Modifies an existing classifier (a GrokClassifier, an XMLClassifier, a JsonClassifier, or a CsvClassifier, depending on which field is present).

Request Syntax

```
{
  "CsvClassifier": {
    "AllowSingleColumn": boolean,
    "ContainsHeader": "string",
    "CustomDatatypeConfigured": boolean,
    "CustomDatatypes": [ "string" ],
    "Delimiter": "string",
    "DisableValueTrimming": boolean,
    "Header": [ "string" ],
    "Name": "string",
    "QuoteSymbol": "string",
    "Serde": "string"
  },
  "GrokClassifier": {
    "Classification": "string",
    "CustomPatterns": "string",
    "GrokPattern": "string",
    "Name": "string"
  },
  "JsonClassifier": {
    "JsonPath": "string",
    "Name": "string"
  },
  "XMLClassifier": {
    "Classification": "string",
    "Name": "string",
    "RowTag": "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CsvClassifier

A `CsvClassifier` object with updated fields.

Type: [UpdateCsvClassifierRequest](#) object

Required: No

GrokClassifier

A `GrokClassifier` object with updated fields.

Type: [UpdateGrokClassifierRequest](#) object

Required: No

JsonClassifier

A `JsonClassifier` object with updated fields.

Type: [UpdateJsonClassifierRequest](#) object

Required: No

XMLClassifier

An `XMLClassifier` object with updated fields.

Type: [UpdateXMLClassifierRequest](#) object

Required: No

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

VersionMismatchException

There was a version conflict.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

UpdateColumnStatisticsForPartition

Creates or updates partition statistics of columns.

The Identity and Access Management (IAM) permission required for this operation is `UpdatePartition`.

Request Syntax

```
{
  "CatalogId": "string",
  "ColumnStatisticsList": [
    {
      "AnalyzedTime": number,
      "ColumnName": "string",
      "ColumnType": "string",
      "StatisticsData": {
        "BinaryColumnStatisticsData": {
          "AverageLength": number,
          "MaximumLength": number,
          "NumberOfNulls": number
        },
        "BooleanColumnStatisticsData": {
          "NumberOfFalses": number,
          "NumberOfNulls": number,
          "NumberOfTrues": number
        },
        "DateColumnStatisticsData": {
          "MaximumValue": number,
          "MinimumValue": number,
          "NumberOfDistinctValues": number,
          "NumberOfNulls": number
        },
        "DecimalColumnStatisticsData": {
          "MaximumValue": {
            "Scale": number,
            "UnscaledValue": blob
          },
          "MinimumValue": {
            "Scale": number,
            "UnscaledValue": blob
          },
          "NumberOfDistinctValues": number,

```

```

        "NumberOfNulls": number
    },
    "DoubleColumnStatisticsData": {
        "MaximumValue": number,
        "MinimumValue": number,
        "NumberOfDistinctValues": number,
        "NumberOfNulls": number
    },
    "LongColumnStatisticsData": {
        "MaximumValue": number,
        "MinimumValue": number,
        "NumberOfDistinctValues": number,
        "NumberOfNulls": number
    },
    "StringColumnStatisticsData": {
        "AverageLength": number,
        "MaximumLength": number,
        "NumberOfDistinctValues": number,
        "NumberOfNulls": number
    },
    "Type": "string"
}
}
],
"DatabaseName": "string",
"PartitionValues": [ "string" ],
"TableName": "string"
}

```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog where the partitions in question reside. If none is supplied, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

ColumnStatisticsList

A list of the column statistics.

Type: Array of [ColumnStatistics](#) objects

Array Members: Minimum number of 0 items. Maximum number of 25 items.

Required: Yes

DatabaseName

The name of the catalog database where the partitions reside.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

PartitionValues

A list of partition values identifying the partition.

Type: Array of strings

Length Constraints: Maximum length of 1024.

Required: Yes

TableName

The name of the partitions' table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "Errors": [
    {
      "ColumnStatistics": {
        "AnalyzedTime": number,
        "ColumnName": "string",
        "ColumnType": "string",
        "StatisticsData": {
          "BinaryColumnStatisticsData": {
            "AverageLength": number,
            "MaximumLength": number,
            "NumberOfNulls": number
          },
          "BooleanColumnStatisticsData": {
            "NumberOfFalses": number,
            "NumberOfNulls": number,
            "NumberOfTrues": number
          },
          "DateColumnStatisticsData": {
            "MaximumValue": number,
            "MinimumValue": number,
            "NumberOfDistinctValues": number,
            "NumberOfNulls": number
          },
          "DecimalColumnStatisticsData": {
            "MaximumValue": {
              "Scale": number,
              "UnscaledValue": blob
            },
            "MinimumValue": {
              "Scale": number,
              "UnscaledValue": blob
            },
            "NumberOfDistinctValues": number,
            "NumberOfNulls": number
          },
          "DoubleColumnStatisticsData": {
            "MaximumValue": number,
            "MinimumValue": number,
            "NumberOfDistinctValues": number,
            "NumberOfNulls": number
          }
        }
      }
    }
  ]
}
```

```

    },
    "LongColumnStatisticsData": {
      "MaximumValue": number,
      "MinimumValue": number,
      "NumberOfDistinctValues": number,
      "NumberOfNulls": number
    },
    "StringColumnStatisticsData": {
      "AverageLength": number,
      "MaximumLength": number,
      "NumberOfDistinctValues": number,
      "NumberOfNulls": number
    },
    "Type": "string"
  }
},
"Error": {
  "ErrorCode": "string",
  "ErrorMessage": "string"
}
}
]
}

```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Errors

Error occurred during updating column statistics data.

Type: Array of [ColumnStatisticsError](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

UpdateColumnStatisticsForTable

Creates or updates table statistics of columns.

The Identity and Access Management (IAM) permission required for this operation is `UpdateTable`.

Request Syntax

```
{
  "CatalogId": "string",
  "ColumnStatisticsList": [
    {
      "AnalyzedTime": number,
      "ColumnName": "string",
      "ColumnType": "string",
      "StatisticsData": {
        "BinaryColumnStatisticsData": {
          "AverageLength": number,
          "MaximumLength": number,
          "NumberOfNulls": number
        },
        "BooleanColumnStatisticsData": {
          "NumberOfFalses": number,
          "NumberOfNulls": number,
          "NumberOfTrues": number
        },
        "DateColumnStatisticsData": {
          "MaximumValue": number,
          "MinimumValue": number,
          "NumberOfDistinctValues": number,
          "NumberOfNulls": number
        },
        "DecimalColumnStatisticsData": {
          "MaximumValue": {
            "Scale": number,
            "UnscaledValue": blob
          },
          "MinimumValue": {
            "Scale": number,
            "UnscaledValue": blob
          },
          "NumberOfDistinctValues": number,

```



```

        "NumberOfNulls": number
    },
    "DoubleColumnStatisticsData": {
        "MaximumValue": number,
        "MinimumValue": number,
        "NumberOfDistinctValues": number,
        "NumberOfNulls": number
    },
    "LongColumnStatisticsData": {
        "MaximumValue": number,
        "MinimumValue": number,
        "NumberOfDistinctValues": number,
        "NumberOfNulls": number
    },
    "StringColumnStatisticsData": {
        "AverageLength": number,
        "MaximumLength": number,
        "NumberOfDistinctValues": number,
        "NumberOfNulls": number
    },
    "Type": "string"
}
}
],
"DatabaseName": "string",
"TableName": "string"
}

```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog where the partitions in question reside. If none is supplied, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

ColumnStatisticsList

A list of the column statistics.

Type: Array of [ColumnStatistics](#) objects

Array Members: Minimum number of 0 items. Maximum number of 25 items.

Required: Yes

DatabaseName

The name of the catalog database where the partitions reside.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

TableName

The name of the partitions' table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "Errors": [
    {
      "ColumnStatistics": {
        "AnalyzedTime": number,
        "ColumnName": "string",
        "ColumnType": "string",
        "StatisticsData": {
```

```
"BinaryColumnStatisticsData": {
  "AverageLength": number,
  "MaximumLength": number,
  "NumberOfNulls": number
},
"BooleanColumnStatisticsData": {
  "NumberOfFalses": number,
  "NumberOfNulls": number,
  "NumberOfTrues": number
},
"DateColumnStatisticsData": {
  "MaximumValue": number,
  "MinimumValue": number,
  "NumberOfDistinctValues": number,
  "NumberOfNulls": number
},
"DecimalColumnStatisticsData": {
  "MaximumValue": {
    "Scale": number,
    "UnscaledValue": blob
  },
  "MinimumValue": {
    "Scale": number,
    "UnscaledValue": blob
  },
  "NumberOfDistinctValues": number,
  "NumberOfNulls": number
},
"DoubleColumnStatisticsData": {
  "MaximumValue": number,
  "MinimumValue": number,
  "NumberOfDistinctValues": number,
  "NumberOfNulls": number
},
"LongColumnStatisticsData": {
  "MaximumValue": number,
  "MinimumValue": number,
  "NumberOfDistinctValues": number,
  "NumberOfNulls": number
},
"StringColumnStatisticsData": {
  "AverageLength": number,
  "MaximumLength": number,
  "NumberOfDistinctValues": number,
```

```
        "NumberOfNulls": number
      },
      "Type": "string"
    }
  },
  "Error": {
    "ErrorCode": "string",
    "ErrorMessage": "string"
  }
}
]
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Errors

List of ColumnStatisticsErrors.

Type: Array of [ColumnStatisticsError](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

UpdateConnection

Updates a connection definition in the Data Catalog.

Request Syntax

```
{
  "CatalogId": "string",
  "ConnectionInput": {
    "AuthenticationConfiguration": {
      "AuthenticationType": "string",
      "OAuth2Properties": {
        "AuthorizationCodeProperties": {
          "AuthorizationCode": "string",
          "RedirectUri": "string"
        },
        "OAuth2ClientApplication": {
          "AWSManagedClientApplicationReference": "string",
          "UserManagedClientApplicationClientId": "string"
        },
        "OAuth2GrantType": "string",
        "TokenUrl": "string",
        "TokenUrlParametersMap": {
          "string": "string"
        }
      },
      "SecretArn": "string"
    },
    "ConnectionProperties": {
      "string": "string"
    },
    "ConnectionType": "string",
    "Description": "string",
    "MatchCriteria": [ "string" ],
    "Name": "string",
    "PhysicalConnectionRequirements": {
      "AvailabilityZone": "string",
      "SecurityGroupIdList": [ "string" ],
      "SubnetId": "string"
    },
    "ValidateCredentials": boolean
  },
  "Name": "string"
}
```

```
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog in which the connection resides. If none is provided, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

ConnectionInput

A `ConnectionInput` object that redefines the connection in question.

Type: [ConnectionInput](#) object

Required: Yes

Name

The name of the connection definition to update.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)

- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

UpdateCrawler

Updates a crawler. If a crawler is running, you must stop it using `StopCrawler` before updating it.

Request Syntax

```
{
  "Classifiers": [ "string" ],
  "Configuration": "string",
  "CrawlerSecurityConfiguration": "string",
  "DatabaseName": "string",
  "Description": "string",
  "LakeFormationConfiguration": {
    "AccountId": "string",
    "UseLakeFormationCredentials": boolean
  },
  "LineageConfiguration": {
    "CrawlerLineageSettings": "string"
  },
  "Name": "string",
  "RecrawlPolicy": {
    "RecrawlBehavior": "string"
  },
  "Role": "string",
  "Schedule": "string",
  "SchemaChangePolicy": {
    "DeleteBehavior": "string",
    "UpdateBehavior": "string"
  },
  "TablePrefix": "string",
  "Targets": {
    "CatalogTargets": [
      {
        "ConnectionName": "string",
        "DatabaseName": "string",
        "DlqEventQueueArn": "string",
        "EventQueueArn": "string",
        "Tables": [ "string" ]
      }
    ],
    "DeltaTargets": [
      {
        "ConnectionName": "string",
```

```
    "CreateNativeDeltaTable": boolean,
    "DeltaTables": [ "string" ],
    "WriteManifest": boolean
  }
],
"DynamoDBTargets": [
  {
    "Path": "string",
    "scanAll": boolean,
    "scanRate": number
  }
],
"HudiTargets": [
  {
    "ConnectionName": "string",
    "Exclusions": [ "string" ],
    "MaximumTraversalDepth": number,
    "Paths": [ "string" ]
  }
],
"IcebergTargets": [
  {
    "ConnectionName": "string",
    "Exclusions": [ "string" ],
    "MaximumTraversalDepth": number,
    "Paths": [ "string" ]
  }
],
"JdbcTargets": [
  {
    "ConnectionName": "string",
    "EnableAdditionalMetadada": [ "string" ],
    "Exclusions": [ "string" ],
    "Path": "string"
  }
],
"MongoDBTargets": [
  {
    "ConnectionName": "string",
    "Path": "string",
    "ScanAll": boolean
  }
],
"S3Targets": [
```


Type: String

Length Constraints: Minimum length of 0. Maximum length of 128.

Required: No

DatabaseName

The AWS Glue database where results are stored, such as: `arn:aws:daylight:us-east-1::database/sometable/*`.

Type: String

Required: No

Description

A description of the new crawler.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

LakeFormationConfiguration

Specifies AWS Lake Formation configuration settings for the crawler.

Type: [LakeFormationConfiguration](#) object

Required: No

LineageConfiguration

Specifies data lineage configuration settings for the crawler.

Type: [LineageConfiguration](#) object

Required: No

Name

Name of the new crawler.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

RecrawlPolicy

A policy that specifies whether to crawl the entire dataset again, or to crawl only folders that were added since the last crawler run.

Type: [RecrawlPolicy](#) object

Required: No

Role

The IAM role or Amazon Resource Name (ARN) of an IAM role that is used by the new crawler to access customer resources.

Type: String

Required: No

Schedule

A cron expression used to specify the schedule (see [Time-Based Schedules for Jobs and Crawlers](#)). For example, to run something every day at 12:15 UTC, you would specify: `cron(15 12 * * ? *)`.

Type: String

Required: No

SchemaChangePolicy

The policy for the crawler's update and deletion behavior.

Type: [SchemaChangePolicy](#) object

Required: No

TablePrefix

The table prefix used for catalog tables that are created.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 128.

Required: No

Targets

A list of targets to crawl.

Type: [CrawlerTargets](#) object

Required: No

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

CrawlerRunningException

The operation cannot be performed because the crawler is already running.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

VersionMismatchException

There was a version conflict.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

SchedulerTransitioningException

The specified scheduler is transitioning.

HTTP Status Code: 400

VersionMismatchException

There was a version conflict.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)

- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

UpdateDatabase

Updates an existing database definition in a Data Catalog.

Request Syntax

```
{
  "CatalogId": "string",
  "DatabaseInput": {
    "CreateTableDefaultPermissions": [
      {
        "Permissions": [ "string" ],
        "Principal": {
          "DataLakePrincipalIdentifier": "string"
        }
      }
    ],
    "Description": "string",
    "FederatedDatabase": {
      "ConnectionName": "string",
      "Identifier": "string"
    },
    "LocationUri": "string",
    "Name": "string",
    "Parameters": {
      "string" : "string"
    },
    "TargetDatabase": {
      "CatalogId": "string",
      "DatabaseName": "string",
      "Region": "string"
    }
  },
  "Name": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog in which the metadata database resides. If none is provided, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

DatabaseInput

A DatabaseInput object specifying the new definition of the metadata database in the catalog.

Type: [DatabaseInput](#) object

Required: Yes

Name

The name of the database to update in the catalog. For Hive compatibility, this is folded to lowercase.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)

- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

UpdateDataQualityRuleset

Updates the specified data quality ruleset.

Request Syntax

```
{
  "Description": "string",
  "Name": "string",
  "Ruleset": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Description

A description of the ruleset.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Name

The name of the data quality ruleset.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Ruleset

A Data Quality Definition Language (DQDL) ruleset. For more information, see the AWS Glue developer guide.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 65536.

Required: No

Response Syntax

```
{
  "Description": "string",
  "Name": "string",
  "Ruleset": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Description

A description of the ruleset.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Name

The name of the data quality ruleset.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF\\t]*`

Ruleset

A Data Quality Definition Language (DQDL) ruleset. For more information, see the AWS Glue developer guide.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 65536.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

IdempotentParameterMismatchException

The same unique identifier was associated with two different records.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

UpdateDevEndpoint

Updates a specified development endpoint.

Request Syntax

```
{
  "AddArguments": {
    "string": "string"
  },
  "AddPublicKeys": [ "string" ],
  "CustomLibraries": {
    "ExtraJarsS3Path": "string",
    "ExtraPythonLibsS3Path": "string"
  },
  "DeleteArguments": [ "string" ],
  "DeletePublicKeys": [ "string" ],
  "EndpointName": "string",
  "PublicKey": "string",
  "UpdateEtlLibraries": boolean
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

AddArguments

The map of arguments to add the map of arguments used to configure the DevEndpoint.

Valid arguments are:

- "--enable-glue-datacatalog": ""

You can specify a version of Python support for development endpoints by using the Arguments parameter in the CreateDevEndpoint or UpdateDevEndpoint APIs. If no arguments are provided, the version defaults to Python 2.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 100 items.

Required: No

AddPublicKeys

The list of public keys for the DevEndpoint to use.

Type: Array of strings

Array Members: Maximum number of 5 items.

Required: No

CustomLibraries

Custom Python or Java libraries to be loaded in the DevEndpoint.

Type: [DevEndpointCustomLibraries](#) object

Required: No

DeleteArguments

The list of argument keys to be deleted from the map of arguments used to configure the DevEndpoint.

Type: Array of strings

Required: No

DeletePublicKeys

The list of public keys to be deleted from the DevEndpoint.

Type: Array of strings

Array Members: Maximum number of 5 items.

Required: No

EndpointName

The name of the DevEndpoint to be updated.

Type: String

Required: Yes

PublicKey

The public key for the DevEndpoint to use.

Type: String

Required: No

UpdateEtlLibraries

True if the list of custom libraries to be loaded in the development endpoint needs to be updated, or False if otherwise.

Type: Boolean

Required: No

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ValidationException

A value could not be validated.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

UpdateJob

Updates an existing job definition. The previous job definition is completely overwritten by this information.

Request Syntax

```
{
  "JobName": "string",
  "JobUpdate": {
    "AllocatedCapacity": number,
    "CodeGenConfigurationNodes": {
      "string" : {
        "Aggregate": {
          "Aggs": [
            {
              "AggFunc": "string",
              "Column": [ "string" ]
            }
          ],
          "Groups": [
            [ "string" ]
          ],
          "Inputs": [ "string" ],
          "Name": "string"
        },
        "AmazonRedshiftSource": {
          "Data": {
            "AccessType": "string",
            "Action": "string",
            "AdvancedOptions": [
              {
                "Key": "string",
                "Value": "string"
              }
            ],
            "CatalogDatabase": {
              "Description": "string",
              "Label": "string",
              "Value": "string"
            },
            "CatalogRedshiftSchema": "string",
            "CatalogRedshiftTable": "string",
```



```
"CatalogTable": {
  "Description": "string",
  "Label": "string",
  "Value": "string"
},
"Connection": {
  "Description": "string",
  "Label": "string",
  "Value": "string"
},
"CrawlerConnection": "string",
"IamRole": {
  "Description": "string",
  "Label": "string",
  "Value": "string"
},
"MergeAction": "string",
"MergeClause": "string",
"MergeWhenMatched": "string",
"MergeWhenNotMatched": "string",
"PostAction": "string",
"PreAction": "string",
"SampleQuery": "string",
"Schema": {
  "Description": "string",
  "Label": "string",
  "Value": "string"
},
"SelectedColumns": [
  {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  }
],
"SourceType": "string",
"StagingTable": "string",
"Table": {
  "Description": "string",
  "Label": "string",
  "Value": "string"
},
"TablePrefix": "string",
"TableSchema": [
```

```
    {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    }
  ],
  "TempDir": "string",
  "Upsert": boolean
},
"Name": "string"
},
"AmazonRedshiftTarget": {
  "Data": {
    "AccessType": "string",
    "Action": "string",
    "AdvancedOptions": [
      {
        "Key": "string",
        "Value": "string"
      }
    ],
    "CatalogDatabase": {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    },
    "CatalogRedshiftSchema": "string",
    "CatalogRedshiftTable": "string",
    "CatalogTable": {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    },
    "Connection": {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    },
    "CrawlerConnection": "string",
    "IamRole": {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    }
  },

```

```
"MergeAction": "string",
"MergeClause": "string",
"MergeWhenMatched": "string",
"MergeWhenNotMatched": "string",
"PostAction": "string",
"PreAction": "string",
"SampleQuery": "string",
"Schema": {
  "Description": "string",
  "Label": "string",
  "Value": "string"
},
"SelectedColumns": [
  {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  }
],
"SourceType": "string",
"StagingTable": "string",
"Table": {
  "Description": "string",
  "Label": "string",
  "Value": "string"
},
"TablePrefix": "string",
"TableSchema": [
  {
    "Description": "string",
    "Label": "string",
    "Value": "string"
  }
],
"TempDir": "string",
"Upsert": boolean
},
"Inputs": [ "string" ],
"Name": "string"
},
"ApplyMapping": {
  "Inputs": [ "string" ],
  "Mapping": [
    {
```

```

        "Children": [
            "Mapping"
        ],
        "Dropped": boolean,
        "FromPath": [ "string" ],
        "FromType": "string",
        "ToKey": "string",
        "ToType": "string"
    }
],
    "Name": "string"
},
"AthenaConnectorSource": {
    "ConnectionName": "string",
    "ConnectionTable": "string",
    "ConnectionType": "string",
    "ConnectorName": "string",
    "Name": "string",
    "OutputSchemas": [
        {
            "Columns": [
                {
                    "Name": "string",
                    "Type": "string"
                }
            ]
        }
    ]
},
"SchemaName": "string"
},
"CatalogDeltaSource": {
    "AdditionalDeltaOptions": {
        "string" : "string"
    },
    "Database": "string",
    "Name": "string",
    "OutputSchemas": [
        {
            "Columns": [
                {
                    "Name": "string",
                    "Type": "string"
                }
            ]
        }
    ]
}

```

```

    }
  ],
  "Table": "string"
},
"CatalogHudiSource": {
  "AdditionalHudiOptions": {
    "string": "string"
  },
  "Database": "string",
  "Name": "string",
  "OutputSchemas": [
    {
      "Columns": [
        {
          "Name": "string",
          "Type": "string"
        }
      ]
    }
  ],
  "Table": "string"
},
"CatalogKafkaSource": {
  "Database": "string",
  "DataPreviewOptions": {
    "PollingTime": number,
    "RecordPollingLimit": number
  },
  "DetectSchema": boolean,
  "Name": "string",
  "StreamingOptions": {
    "AddRecordTimestamp": "string",
    "Assign": "string",
    "BootstrapServers": "string",
    "Classification": "string",
    "ConnectionName": "string",
    "Delimiter": "string",
    "EmitConsumerLagMetrics": "string",
    "EndingOffsets": "string",
    "IncludeHeaders": boolean,
    "MaxOffsetsPerTrigger": number,
    "MinPartitions": number,
    "NumRetries": number,
    "PollTimeoutMs": number,

```

```
    "RetryIntervalMs": number,
    "SecurityProtocol": "string",
    "StartingOffsets": "string",
    "StartingTimestamp": "string",
    "SubscribePattern": "string",
    "TopicName": "string"
  },
  "Table": "string",
  "WindowSize": number
},
"CatalogKinesisSource": {
  "Database": "string",
  "DataPreviewOptions": {
    "PollingTime": number,
    "RecordPollingLimit": number
  },
  "DetectSchema": boolean,
  "Name": "string",
  "StreamingOptions": {
    "AddIdleTimeBetweenReads": boolean,
    "AddRecordTimestamp": "string",
    "AvoidEmptyBatches": boolean,
    "Classification": "string",
    "Delimiter": "string",
    "DescribeShardInterval": number,
    "EmitConsumerLagMetrics": "string",
    "EndpointUrl": "string",
    "IdleTimeBetweenReadsInMs": number,
    "MaxFetchRecordsPerShard": number,
    "MaxFetchTimeInMs": number,
    "MaxRecordPerRead": number,
    "MaxRetryIntervalMs": number,
    "NumRetries": number,
    "RetryIntervalMs": number,
    "RoleArn": "string",
    "RoleSessionName": "string",
    "StartingPosition": "string",
    "StartingTimestamp": "string",
    "StreamArn": "string",
    "StreamName": "string"
  },
  "Table": "string",
  "WindowSize": number
},
```

```
"CatalogSource": {
  "Database": "string",
  "Name": "string",
  "Table": "string"
},
"CatalogTarget": {
  "Database": "string",
  "Inputs": [ "string" ],
  "Name": "string",
  "Table": "string"
},
"ConnectorDataSource": {
  "ConnectionType": "string",
  "Data": {
    "string" : "string"
  },
  "Name": "string",
  "OutputSchemas": [
    {
      "Columns": [
        {
          "Name": "string",
          "Type": "string"
        }
      ]
    }
  ]
},
"ConnectorDataTarget": {
  "ConnectionType": "string",
  "Data": {
    "string" : "string"
  },
  "Inputs": [ "string" ],
  "Name": "string"
},
"CustomCode": {
  "ClassName": "string",
  "Code": "string",
  "Inputs": [ "string" ],
  "Name": "string",
  "OutputSchemas": [
    {
      "Columns": [
```

```
        {
            "Name": "string",
            "Type": "string"
        }
    ]
}
],
"DirectJDBCSource": {
    "ConnectionName": "string",
    "ConnectionType": "string",
    "Database": "string",
    "Name": "string",
    "RedshiftTmpDir": "string",
    "Table": "string"
},
"DirectKafkaSource": {
    "DataPreviewOptions": {
        "PollingTime": number,
        "RecordPollingLimit": number
    },
    "DetectSchema": boolean,
    "Name": "string",
    "StreamingOptions": {
        "AddRecordTimestamp": "string",
        "Assign": "string",
        "BootstrapServers": "string",
        "Classification": "string",
        "ConnectionName": "string",
        "Delimiter": "string",
        "EmitConsumerLagMetrics": "string",
        "EndingOffsets": "string",
        "IncludeHeaders": boolean,
        "MaxOffsetsPerTrigger": number,
        "MinPartitions": number,
        "NumRetries": number,
        "PollTimeoutMs": number,
        "RetryIntervalMs": number,
        "SecurityProtocol": "string",
        "StartingOffsets": "string",
        "StartingTimestamp": "string",
        "SubscribePattern": "string",
        "TopicName": "string"
    }
},
```



```

    "WindowSize": number
  },
  "DirectKinesisSource": {
    "DataPreviewOptions": {
      "PollingTime": number,
      "RecordPollingLimit": number
    },
    "DetectSchema": boolean,
    "Name": "string",
    "StreamingOptions": {
      "AddIdleTimeBetweenReads": boolean,
      "AddRecordTimestamp": "string",
      "AvoidEmptyBatches": boolean,
      "Classification": "string",
      "Delimiter": "string",
      "DescribeShardInterval": number,
      "EmitConsumerLagMetrics": "string",
      "EndpointUrl": "string",
      "IdleTimeBetweenReadsInMs": number,
      "MaxFetchRecordsPerShard": number,
      "MaxFetchTimeInMs": number,
      "MaxRecordPerRead": number,
      "MaxRetryIntervalMs": number,
      "NumRetries": number,
      "RetryIntervalMs": number,
      "RoleArn": "string",
      "RoleSessionName": "string",
      "StartingPosition": "string",
      "StartingTimestamp": "string",
      "StreamArn": "string",
      "StreamName": "string"
    },
    "WindowSize": number
  },
  "DropDuplicates": {
    "Columns": [
      [ "string" ]
    ],
    "Inputs": [ "string" ],
    "Name": "string"
  },
  "DropFields": {
    "Inputs": [ "string" ],
    "Name": "string",

```

```
    "Paths": [
      [ "string" ]
    ]
  },
  "DropNullFields": {
    "Inputs": [ "string" ],
    "Name": "string",
    "NullCheckBoxList": {
      "IsEmpty": boolean,
      "IsNegOne": boolean,
      "IsNullString": boolean
    },
    "NullTextList": [
      {
        "Datatype": {
          "Id": "string",
          "Label": "string"
        },
        "Value": "string"
      }
    ]
  },
  "DynamicTransform": {
    "FunctionName": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ]
  },
  "Parameters": [
    {
      "IsOptional": boolean,
      "ListType": "string",
      "Name": "string",
      "Type": "string",
      "ValidationMessage": "string",
      "ValidationRule": "string",
    }
  ]
}
```

```

        "Value": [ "string" ]
    }
],
"Path": "string",
"TransformName": "string",
"Version": "string"
},
"DynamoDBCatalogSource": {
    "Database": "string",
    "Name": "string",
    "Table": "string"
},
"EvaluateDataQuality": {
    "Inputs": [ "string" ],
    "Name": "string",
    "Output": "string",
    "PublishingOptions": {
        "CloudWatchMetricsEnabled": boolean,
        "EvaluationContext": "string",
        "ResultsPublishingEnabled": boolean,
        "ResultsS3Prefix": "string"
    },
    "Ruleset": "string",
    "StopJobOnFailureOptions": {
        "StopJobOnFailureTiming": "string"
    }
},
"EvaluateDataQualityMultiFrame": {
    "AdditionalDataSources": {
        "string" : "string"
    },
    "AdditionalOptions": {
        "string" : "string"
    },
    "Inputs": [ "string" ],
    "Name": "string",
    "PublishingOptions": {
        "CloudWatchMetricsEnabled": boolean,
        "EvaluationContext": "string",
        "ResultsPublishingEnabled": boolean,
        "ResultsS3Prefix": "string"
    },
    "Ruleset": "string",
    "StopJobOnFailureOptions": {

```

```
        "StopJobOnFailureTiming": "string"
    }
},
"FillMissingValues": {
    "FilledPath": "string",
    "ImputedPath": "string",
    "Inputs": [ "string" ],
    "Name": "string"
},
"Filter": {
    "Filters": [
        {
            "Negated": boolean,
            "Operation": "string",
            "Values": [
                {
                    "Type": "string",
                    "Value": [ "string" ]
                }
            ]
        }
    ]
},
"Inputs": [ "string" ],
"LogicalOperator": "string",
"Name": "string"
},
"GovernedCatalogSource": {
    "AdditionalOptions": {
        "BoundedFiles": number,
        "BoundedSize": number
    },
    "Database": "string",
    "Name": "string",
    "PartitionPredicate": "string",
    "Table": "string"
},
"GovernedCatalogTarget": {
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
        [ "string" ]
    ],
    "SchemaChangePolicy": {
```

```

        "EnableUpdateCatalog": boolean,
        "UpdateBehavior": "string"
    },
    "Table": "string"
},
"JDBCConnectorSource": {
    "AdditionalOptions": {
        "DataTypeMapping": {
            "string" : "string"
        },
        "FilterPredicate": "string",
        "JobBookmarkKeys": [ "string" ],
        "JobBookmarkKeysSortOrder": "string",
        "LowerBound": number,
        "NumPartitions": number,
        "PartitionColumn": "string",
        "UpperBound": number
    },
    "ConnectionName": "string",
    "ConnectionTable": "string",
    "ConnectionType": "string",
    "ConnectorName": "string",
    "Name": "string",
    "OutputSchemas": [
        {
            "Columns": [
                {
                    "Name": "string",
                    "Type": "string"
                }
            ]
        }
    ]
},
    "Query": "string"
},
"JDBCConnectorTarget": {
    "AdditionalOptions": {
        "string" : "string"
    },
    "ConnectionName": "string",
    "ConnectionTable": "string",
    "ConnectionType": "string",
    "ConnectorName": "string",
    "Inputs": [ "string" ],

```

```
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
  },
  "Join": {
    "Columns": [
      {
        "From": "string",
        "Keys": [
          "string"
        ]
      }
    ]
  },
  "Inputs": [ "string" ],
  "JoinType": "string",
  "Name": "string"
},
"Merge": {
  "Inputs": [ "string" ],
  "Name": "string",
  "PrimaryKeys": [
    "string"
  ],
  "Source": "string"
},
"MicrosoftSQLServerCatalogSource": {
  "Database": "string",
  "Name": "string",
  "Table": "string"
},
"MicrosoftSQLServerCatalogTarget": {
  "Database": "string",
  "Inputs": [ "string" ],
  "Name": "string",
  "Table": "string"
},
}
```

```
"MySQLCatalogSource": {
  "Database": "string",
  "Name": "string",
  "Table": "string"
},
"MySQLCatalogTarget": {
  "Database": "string",
  "Inputs": [ "string" ],
  "Name": "string",
  "Table": "string"
},
"OracleSQLCatalogSource": {
  "Database": "string",
  "Name": "string",
  "Table": "string"
},
"OracleSQLCatalogTarget": {
  "Database": "string",
  "Inputs": [ "string" ],
  "Name": "string",
  "Table": "string"
},
"PIIDetection": {
  "EntityTypesToDetect": [ "string" ],
  "Inputs": [ "string" ],
  "MaskValue": "string",
  "Name": "string",
  "OutputColumnName": "string",
  "PiiType": "string",
  "SampleFraction": number,
  "ThresholdFraction": number
},
"PostgreSQLCatalogSource": {
  "Database": "string",
  "Name": "string",
  "Table": "string"
},
"PostgreSQLCatalogTarget": {
  "Database": "string",
  "Inputs": [ "string" ],
  "Name": "string",
  "Table": "string"
},
"Recipe": {
```

```
    "Inputs": [ "string" ],
    "Name": "string",
    "RecipeReference": {
      "RecipeArn": "string",
      "RecipeVersion": "string"
    }
  },
  "RedshiftSource": {
    "Database": "string",
    "Name": "string",
    "RedshiftTmpDir": "string",
    "Table": "string",
    "TmpDirIAMRole": "string"
  },
  "RedshiftTarget": {
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "RedshiftTmpDir": "string",
    "Table": "string",
    "TmpDirIAMRole": "string",
    "UpsertRedshiftOptions": {
      "ConnectionName": "string",
      "TableLocation": "string",
      "UpsertKeys": [ "string" ]
    }
  },
  "RelationalCatalogSource": {
    "Database": "string",
    "Name": "string",
    "Table": "string"
  },
  "RenameField": {
    "Inputs": [ "string" ],
    "Name": "string",
    "SourcePath": [ "string" ],
    "TargetPath": [ "string" ]
  },
  "S3CatalogDeltaSource": {
    "AdditionalDeltaOptions": {
      "string" : "string"
    },
    "Database": "string",
    "Name": "string",
```



```
"OutputSchemas": [
  {
    "Columns": [
      {
        "Name": "string",
        "Type": "string"
      }
    ]
  }
],
"Table": "string"
},
"S3CatalogHudiSource": {
  "AdditionalHudiOptions": {
    "string": "string"
  },
  "Database": "string",
  "Name": "string",
  "OutputSchemas": [
    {
      "Columns": [
        {
          "Name": "string",
          "Type": "string"
        }
      ]
    }
  ],
  "Table": "string"
},
"S3CatalogSource": {
  "AdditionalOptions": {
    "BoundedFiles": number,
    "BoundedSize": number
  },
  "Database": "string",
  "Name": "string",
  "PartitionPredicate": "string",
  "Table": "string"
},
"S3CatalogTarget": {
  "Database": "string",
  "Inputs": [ "string" ],
  "Name": "string",
```

```
    "PartitionKeys": [
      [ "string" ]
    ],
    "SchemaChangePolicy": {
      "EnableUpdateCatalog": boolean,
      "UpdateBehavior": "string"
    },
    "Table": "string"
  },
  "S3CsvSource": {
    "AdditionalOptions": {
      "BoundedFiles": number,
      "BoundedSize": number,
      "EnableSamplePath": boolean,
      "SamplePath": "string"
    },
    "CompressionType": "string",
    "Escaper": "string",
    "Exclusions": [ "string" ],
    "GroupFiles": "string",
    "GroupSize": "string",
    "MaxBand": number,
    "MaxFilesInBand": number,
    "Multiline": boolean,
    "Name": "string",
    "OptimizePerformance": boolean,
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
    "Paths": [ "string" ],
    "QuoteChar": "string",
    "Recurse": boolean,
    "Separator": "string",
    "SkipFirst": boolean,
    "WithHeader": boolean,
    "WriteHeader": boolean
  },
},
```

```
"S3DeltaCatalogTarget": {
  "AdditionalOptions": {
    "string" : "string"
  },
  "Database": "string",
  "Inputs": [ "string" ],
  "Name": "string",
  "PartitionKeys": [
    [ "string" ]
  ],
  "SchemaChangePolicy": {
    "EnableUpdateCatalog": boolean,
    "UpdateBehavior": "string"
  },
  "Table": "string"
},
"S3DeltaDirectTarget": {
  "AdditionalOptions": {
    "string" : "string"
  },
  "Compression": "string",
  "Format": "string",
  "Inputs": [ "string" ],
  "Name": "string",
  "PartitionKeys": [
    [ "string" ]
  ],
  "Path": "string",
  "SchemaChangePolicy": {
    "Database": "string",
    "EnableUpdateCatalog": boolean,
    "Table": "string",
    "UpdateBehavior": "string"
  }
},
"S3DeltaSource": {
  "AdditionalDeltaOptions": {
    "string" : "string"
  },
  "AdditionalOptions": {
    "BoundedFiles": number,
    "BoundedSize": number,
    "EnableSamplePath": boolean,
    "SamplePath": "string"
  }
}
```

```
    },
    "Name": "string",
    "OutputSchemas": [
      {
        "Columns": [
          {
            "Name": "string",
            "Type": "string"
          }
        ]
      }
    ],
    "Paths": [ "string" ]
  },
  "S3DirectTarget": {
    "Compression": "string",
    "Format": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
      [ "string" ]
    ],
  },
  "Path": "string",
  "SchemaChangePolicy": {
    "Database": "string",
    "EnableUpdateCatalog": boolean,
    "Table": "string",
    "UpdateBehavior": "string"
  }
},
"S3GlueParquetTarget": {
  "Compression": "string",
  "Inputs": [ "string" ],
  "Name": "string",
  "PartitionKeys": [
    [ "string" ]
  ],
},
"Path": "string",
"SchemaChangePolicy": {
  "Database": "string",
  "EnableUpdateCatalog": boolean,
  "Table": "string",
  "UpdateBehavior": "string"
}
```

```
},
  "S3HudiCatalogTarget": {
    "AdditionalOptions": {
      "string" : "string"
    },
    "Database": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
      [ "string" ]
    ],
    "SchemaChangePolicy": {
      "EnableUpdateCatalog": boolean,
      "UpdateBehavior": "string"
    },
    "Table": "string"
  },
  "S3HudiDirectTarget": {
    "AdditionalOptions": {
      "string" : "string"
    },
    "Compression": "string",
    "Format": "string",
    "Inputs": [ "string" ],
    "Name": "string",
    "PartitionKeys": [
      [ "string" ]
    ],
    "Path": "string",
    "SchemaChangePolicy": {
      "Database": "string",
      "EnableUpdateCatalog": boolean,
      "Table": "string",
      "UpdateBehavior": "string"
    }
  },
  "S3HudiSource": {
    "AdditionalHudiOptions": {
      "string" : "string"
    },
    "AdditionalOptions": {
      "BoundedFiles": number,
      "BoundedSize": number,
      "EnableSamplePath": boolean,
```

```
    "SamplePath": "string"
  },
  "Name": "string",
  "OutputSchemas": [
    {
      "Columns": [
        {
          "Name": "string",
          "Type": "string"
        }
      ]
    }
  ],
  "Paths": [ "string" ]
},
"S3JsonSource": {
  "AdditionalOptions": {
    "BoundedFiles": number,
    "BoundedSize": number,
    "EnableSamplePath": boolean,
    "SamplePath": "string"
  },
  "CompressionType": "string",
  "Exclusions": [ "string" ],
  "GroupFiles": "string",
  "GroupSize": "string",
  "JsonPath": "string",
  "MaxBand": number,
  "MaxFilesInBand": number,
  "Multiline": boolean,
  "Name": "string",
  "OutputSchemas": [
    {
      "Columns": [
        {
          "Name": "string",
          "Type": "string"
        }
      ]
    }
  ],
  "Paths": [ "string" ],
  "Recurse": boolean
},
```

```

    "S3ParquetSource": {
      "AdditionalOptions": {
        "BoundedFiles": number,
        "BoundedSize": number,
        "EnableSamplePath": boolean,
        "SamplePath": "string"
      },
      "CompressionType": "string",
      "Exclusions": [ "string" ],
      "GroupFiles": "string",
      "GroupSize": "string",
      "MaxBand": number,
      "MaxFilesInBand": number,
      "Name": "string",
      "OutputSchemas": [
        {
          "Columns": [
            {
              "Name": "string",
              "Type": "string"
            }
          ]
        }
      ],
      "Paths": [ "string" ],
      "Recurse": boolean
    },
    "SelectFields": {
      "Inputs": [ "string" ],
      "Name": "string",
      "Paths": [
        [ "string" ]
      ]
    }
  ],
  "SelectFromCollection": {
    "Index": number,
    "Inputs": [ "string" ],
    "Name": "string"
  },
  "SnowflakeSource": {
    "Data": {
      "Action": "string",
      "AdditionalOptions": {
        "string": "string"
      }
    }
  }
}

```

```
    },
    "AutoPushdown": boolean,
    "Connection": {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    },
    "Database": "string",
    "IamRole": {
      "Description": "string",
      "Label": "string",
      "Value": "string"
    },
    "MergeAction": "string",
    "MergeClause": "string",
    "MergeWhenMatched": "string",
    "MergeWhenNotMatched": "string",
    "PostAction": "string",
    "PreAction": "string",
    "SampleQuery": "string",
    "Schema": "string",
    "SelectedColumns": [
      {
        "Description": "string",
        "Label": "string",
        "Value": "string"
      }
    ],
    "SourceType": "string",
    "StagingTable": "string",
    "Table": "string",
    "TableSchema": [
      {
        "Description": "string",
        "Label": "string",
        "Value": "string"
      }
    ],
    "TempDir": "string",
    "Upsert": boolean
  },
  "Name": "string",
  "OutputSchemas": [
    {
```



```
        "Columns": [
            {
                "Name": "string",
                "Type": "string"
            }
        ]
    },
    ],
    "SnowflakeTarget": {
        "Data": {
            "Action": "string",
            "AdditionalOptions": {
                "string": "string"
            },
            "AutoPushdown": boolean,
            "Connection": {
                "Description": "string",
                "Label": "string",
                "Value": "string"
            },
            "Database": "string",
            "IamRole": {
                "Description": "string",
                "Label": "string",
                "Value": "string"
            },
            "MergeAction": "string",
            "MergeClause": "string",
            "MergeWhenMatched": "string",
            "MergeWhenNotMatched": "string",
            "PostAction": "string",
            "PreAction": "string",
            "SampleQuery": "string",
            "Schema": "string",
            "SelectedColumns": [
                {
                    "Description": "string",
                    "Label": "string",
                    "Value": "string"
                }
            ],
            "SourceType": "string",
            "StagingTable": "string",
```

```
    "Table": "string",
    "TableSchema": [
      {
        "Description": "string",
        "Label": "string",
        "Value": "string"
      }
    ],
    "TempDir": "string",
    "Upsert": boolean
  },
  "Inputs": [ "string" ],
  "Name": "string"
},
"SparkConnectorSource": {
  "AdditionalOptions": {
    "string" : "string"
  },
  "ConnectionName": "string",
  "ConnectionType": "string",
  "ConnectorName": "string",
  "Name": "string",
  "OutputSchemas": [
    {
      "Columns": [
        {
          "Name": "string",
          "Type": "string"
        }
      ]
    }
  ]
},
"SparkConnectorTarget": {
  "AdditionalOptions": {
    "string" : "string"
  },
  "ConnectionName": "string",
  "ConnectionType": "string",
  "ConnectorName": "string",
  "Inputs": [ "string" ],
  "Name": "string",
  "OutputSchemas": [
    {
```

```
        "Columns": [
            {
                "Name": "string",
                "Type": "string"
            }
        ]
    },
    ],
    "SparkSQL": {
        "Inputs": [ "string" ],
        "Name": "string",
        "OutputSchemas": [
            {
                "Columns": [
                    {
                        "Name": "string",
                        "Type": "string"
                    }
                ]
            }
        ],
        "SqlAliases": [
            {
                "Alias": "string",
                "From": "string"
            }
        ],
        "SqlQuery": "string"
    },
    "Spigot": {
        "Inputs": [ "string" ],
        "Name": "string",
        "Path": "string",
        "Prob": number,
        "Topk": number
    },
    "SplitFields": {
        "Inputs": [ "string" ],
        "Name": "string",
        "Paths": [
            [ "string" ]
        ]
    }
},
```

```
    "Union": {
      "Inputs": [ "string" ],
      "Name": "string",
      "UnionType": "string"
    }
  },
  "Command": {
    "Name": "string",
    "PythonVersion": "string",
    "Runtime": "string",
    "ScriptLocation": "string"
  },
  "Connections": {
    "Connections": [ "string" ]
  },
  "DefaultArguments": {
    "string" : "string"
  },
  "Description": "string",
  "ExecutionClass": "string",
  "ExecutionProperty": {
    "MaxConcurrentRuns": number
  },
  "GlueVersion": "string",
  "JobMode": "string",
  "LogUri": "string",
  "MaintenanceWindow": "string",
  "MaxCapacity": number,
  "MaxRetries": number,
  "NonOverridableArguments": {
    "string" : "string"
  },
  "NotificationProperty": {
    "NotifyDelayAfter": number
  },
  "NumberOfWorkers": number,
  "Role": "string",
  "SecurityConfiguration": "string",
  "SourceControlDetails": {
    "AuthStrategy": "string",
    "AuthToken": "string",
    "Branch": "string",
    "Folder": "string",
```

```
    "LastCommitId": "string",
    "Owner": "string",
    "Provider": "string",
    "Repository": "string"
  },
  "Timeout": number,
  "WorkerType": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

JobName

The name of the job definition to update.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

JobUpdate

Specifies the values with which to update the job definition. Unspecified configuration is removed or reset to default values.

Type: [JobUpdate](#) object

Required: Yes

Response Syntax

```
{
  "JobName": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

JobName

Returns the name of the updated job definition.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

UpdateJobFromSourceControl

Synchronizes a job from the source control repository. This operation takes the job artifacts that are located in the remote repository and updates the AWS Glue internal stores with these artifacts.

This API supports optional parameters which take in the repository information.

Request Syntax

```
{
  "AuthStrategy": "string",
  "AuthToken": "string",
  "BranchName": "string",
  "CommitId": "string",
  "Folder": "string",
  "JobName": "string",
  "Provider": "string",
  "RepositoryName": "string",
  "RepositoryOwner": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

[AuthStrategy](#)

The type of authentication, which can be an authentication token stored in AWS Secrets Manager, or a personal access token.

Type: String

Valid Values: PERSONAL_ACCESS_TOKEN | AWS_SECRETS_MANAGER

Required: No

[AuthToken](#)

The value of the authorization token.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

BranchName

An optional branch in the remote repository.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

CommitId

A commit ID for a commit in the remote repository.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 40.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Folder

An optional folder in the remote repository.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

JobName

The name of the AWS Glue job to be synchronized to or from the remote repository.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Provider

The provider for the remote repository. Possible values: GITHUB, AWS_CODE_COMMIT, GITLAB, BITBUCKET.

Type: String

Valid Values: GITHUB | GITLAB | BITBUCKET | AWS_CODE_COMMIT

Required: No

RepositoryName

The name of the remote repository that contains the job artifacts. For BitBucket providers, RepositoryName should include WorkspaceName. Use the format `<WorkspaceName>/<RepositoryName>`.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

RepositoryOwner

The owner of the remote repository that contains the job artifacts.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Response Syntax

```
{  
  "JobName": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

JobName

The name of the AWS Glue job.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ValidationException

A value could not be validated.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

UpdateMLTransform

Updates an existing machine learning transform. Call this operation to tune the algorithm parameters to achieve better results.

After calling this operation, you can call the `StartMLEvaluationTaskRun` operation to assess how well your new parameters achieved your goals (such as improving the quality of your machine learning transform, or making it more cost-effective).

Request Syntax

```
{
  "Description": "string",
  "GlueVersion": "string",
  "MaxCapacity": number,
  "MaxRetries": number,
  "Name": "string",
  "NumberOfWorkers": number,
  "Parameters": {
    "FindMatchesParameters": {
      "AccuracyCostTradeoff": number,
      "EnforceProvidedLabels": boolean,
      "PrecisionRecallTradeoff": number,
      "PrimaryKeyColumnName": "string"
    },
    "TransformType": "string"
  },
  "Role": "string",
  "Timeout": number,
  "TransformId": "string",
  "WorkerType": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Description

A description of the transform. The default is an empty string.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

GlueVersion

This value determines which version of AWS Glue this machine learning transform is compatible with. Glue 1.0 is recommended for most customers. If the value is not set, the Glue compatibility defaults to Glue 0.9. For more information, see [AWS Glue Versions](#) in the developer guide.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `^\w+\.\w+$`

Required: No

MaxCapacity

The number of AWS Glue data processing units (DPUs) that are allocated to task runs for this transform. You can allocate from 2 to 100 DPUs; the default is 10. A DPU is a relative measure of processing power that consists of 4 vCPUs of compute capacity and 16 GB of memory. For more information, see the [AWS Glue pricing page](#).

When the `WorkerType` field is set to a value other than `Standard`, the `MaxCapacity` field is set automatically and becomes read-only.

Type: Double

Required: No

MaxRetries

The maximum number of times to retry a task for this transform after a task run fails.

Type: Integer

Required: No

Name

The unique name that you gave the transform when you created it.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

NumberOfWorkers

The number of workers of a defined `workerType` that are allocated when this task runs.

Type: Integer

Required: No

Parameters

The configuration parameters that are specific to the transform type (algorithm) used. Conditionally dependent on the transform type.

Type: [TransformParameters](#) object

Required: No

Role

The name or Amazon Resource Name (ARN) of the IAM role with the required permissions.

Type: String

Required: No

Timeout

The timeout for a task run for this transform in minutes. This is the maximum time that a task run for this transform can consume resources before it is terminated and enters TIMEOUT status. The default is 2,880 minutes (48 hours).

Type: Integer

Valid Range: Minimum value of 1.

Required: No

TransformId

A unique identifier that was generated when the transform was created.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

WorkerType

The type of predefined worker that is allocated when this task runs. Accepts a value of Standard, G.1X, or G.2X.

- For the Standard worker type, each worker provides 4 vCPU, 16 GB of memory and a 50GB disk, and 2 executors per worker.
- For the G.1X worker type, each worker provides 4 vCPU, 16 GB of memory and a 64GB disk, and 1 executor per worker.
- For the G.2X worker type, each worker provides 8 vCPU, 32 GB of memory and a 128GB disk, and 1 executor per worker.

Type: String

Valid Values: Standard | G.1X | G.2X | G.025X | G.4X | G.8X | Z.2X

Required: No

Response Syntax

```
{  
  "TransformId": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

TransformId

The unique identifier for the transform that was updated.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

UpdatePartition

Updates a partition.

Request Syntax

```
{
  "CatalogId": "string",
  "DatabaseName": "string",
  "PartitionInput": {
    "LastAccessTime": number,
    "LastAnalyzedTime": number,
    "Parameters": {
      "string" : "string"
    },
    "StorageDescriptor": {
      "AdditionalLocations": [ "string" ],
      "BucketColumns": [ "string" ],
      "Columns": [
        {
          "Comment": "string",
          "Name": "string",
          "Parameters": {
            "string" : "string"
          },
          "Type": "string"
        }
      ],
      "Compressed": boolean,
      "InputFormat": "string",
      "Location": "string",
      "NumberOfBuckets": number,
      "OutputFormat": "string",
      "Parameters": {
        "string" : "string"
      },
      "SchemaReference": {
        "SchemaId": {
          "RegistryName": "string",
          "SchemaArn": "string",
          "SchemaName": "string"
        },
        "SchemaVersionId": "string",
```

```
    "SchemaVersionNumber": number
  },
  "SerdeInfo": {
    "Name": "string",
    "Parameters": {
      "string" : "string"
    },
    "SerializationLibrary": "string"
  },
  "SkewedInfo": {
    "SkewedColumnNames": [ "string" ],
    "SkewedColumnValueLocationMaps": {
      "string" : "string"
    },
    "SkewedColumnValues": [ "string" ]
  },
  "SortColumns": [
    {
      "Column": "string",
      "SortOrder": number
    }
  ],
  "StoredAsSubDirectories": boolean
},
"Values": [ "string" ]
},
"PartitionValueList": [ "string" ],
"TableName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog where the partition to be updated resides. If none is provided, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

DatabaseName

The name of the catalog database in which the table in question resides.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

PartitionInput

The new partition object to update the partition to.

The `Values` property can't be changed. If you want to change the partition key values for a partition, delete and recreate the partition.

Type: [PartitionInput](#) object

Required: Yes

PartitionValueList

List of partition key values that define the partition to update.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 100 items.

Length Constraints: Maximum length of 1024.

Required: Yes

TableName

The name of the table in which the partition to be updated is located.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

UpdateRegistry

Updates an existing registry which is used to hold a collection of schemas. The updated properties relate to the registry, and do not modify any of the schemas within the registry.

Request Syntax

```
{
  "Description": "string",
  "RegistryId": {
    "RegistryArn": "string",
    "RegistryName": "string"
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Description

A description of the registry. If description is not provided, this field will not be updated.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

RegistryId

This is a wrapper structure that may contain the registry name and Amazon Resource Name (ARN).

Type: [RegistryId](#) object

Required: Yes

Response Syntax

```
{  
  "RegistryArn": "string",  
  "RegistryName": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

RegistryArn

The Amazon Resource name (ARN) of the updated registry.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 10240.

Pattern: arn:(aws|aws-us-gov|aws-cn):glue:.*

RegistryName

The name of the updated registry.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: [a-zA-Z0-9-_\$#.]+

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

UpdateSchema

Updates the description, compatibility setting, or version checkpoint for a schema set.

For updating the compatibility setting, the call will not validate compatibility for the entire set of schema versions with the new compatibility setting. If the value for `Compatibility` is provided, the `VersionNumber` (a checkpoint) is also required. The API will validate the checkpoint version number for consistency.

If the value for the `VersionNumber` (checkpoint) is provided, `Compatibility` is optional and this can be used to set/reset a checkpoint for the schema.

This update will happen only if the schema is in the `AVAILABLE` state.

Request Syntax

```
{
  "Compatibility": "string",
  "Description": "string",
  "SchemaId": {
    "RegistryName": "string",
    "SchemaArn": "string",
    "SchemaName": "string"
  },
  "SchemaVersionNumber": {
    "LatestVersion": boolean,
    "VersionNumber": number
  }
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Compatibility

The new compatibility setting for the schema.

Type: String


```
"SchemaArn": "string",  
"SchemaName": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

[RegistryName](#)

The name of the registry that contains the schema.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: [a-zA-Z0-9-_\$#.]+

[SchemaArn](#)

The Amazon Resource Name (ARN) of the schema.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 10240.

Pattern: arn:(aws|aws-us-gov|aws-cn):glue:.*

[SchemaName](#)

The name of the schema.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: [a-zA-Z0-9-_\$#.]+

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)

- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

UpdateSourceControlFromJob

Synchronizes a job to the source control repository. This operation takes the job artifacts from the AWS Glue internal stores and makes a commit to the remote repository that is configured on the job.

This API supports optional parameters which take in the repository information.

Request Syntax

```
{
  "AuthStrategy": "string",
  "AuthToken": "string",
  "BranchName": "string",
  "CommitId": "string",
  "Folder": "string",
  "JobName": "string",
  "Provider": "string",
  "RepositoryName": "string",
  "RepositoryOwner": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

AuthStrategy

The type of authentication, which can be an authentication token stored in AWS Secrets Manager, or a personal access token.

Type: String

Valid Values: PERSONAL_ACCESS_TOKEN | AWS_SECRETS_MANAGER

Required: No

AuthToken

The value of the authorization token.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

BranchName

An optional branch in the remote repository.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

CommitId

A commit ID for a commit in the remote repository.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 40.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Folder

An optional folder in the remote repository.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

JobName

The name of the AWS Glue job to be synchronized to or from the remote repository.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Provider

The provider for the remote repository. Possible values: GITHUB, AWS_CODE_COMMIT, GITLAB, BITBUCKET.

Type: String

Valid Values: GITHUB | GITLAB | BITBUCKET | AWS_CODE_COMMIT

Required: No

RepositoryName

The name of the remote repository that contains the job artifacts. For BitBucket providers, `RepositoryName` should include `WorkspaceName`. Use the format `<WorkspaceName>/<RepositoryName>`.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

RepositoryOwner

The owner of the remote repository that contains the job artifacts.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Response Syntax

```
{  
  "JobName": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

JobName

The name of the AWS Glue job.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

AlreadyExistsException

A resource to be created or added already exists.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ValidationException

A value could not be validated.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

UpdateTable

Updates a metadata table in the Data Catalog.

Request Syntax

```
{
  "CatalogId": "string",
  "DatabaseName": "string",
  "Force": boolean,
  "SkipArchive": boolean,
  "TableInput": {
    "Description": "string",
    "LastAccessTime": number,
    "LastAnalyzedTime": number,
    "Name": "string",
    "Owner": "string",
    "Parameters": {
      "string" : "string"
    },
    "PartitionKeys": [
      {
        "Comment": "string",
        "Name": "string",
        "Parameters": {
          "string" : "string"
        },
        "Type": "string"
      }
    ],
    "Retention": number,
    "StorageDescriptor": {
      "AdditionalLocations": [ "string" ],
      "BucketColumns": [ "string" ],
      "Columns": [
        {
          "Comment": "string",
          "Name": "string",
          "Parameters": {
            "string" : "string"
          },
          "Type": "string"
        }
      ]
    }
  }
}
```

```
],
  "Compressed": boolean,
  "InputFormat": "string",
  "Location": "string",
  "NumberOfBuckets": number,
  "OutputFormat": "string",
  "Parameters": {
    "string" : "string"
  },
  "SchemaReference": {
    "SchemaId": {
      "RegistryName": "string",
      "SchemaArn": "string",
      "SchemaName": "string"
    },
    "SchemaVersionId": "string",
    "SchemaVersionNumber": number
  },
  "SerdeInfo": {
    "Name": "string",
    "Parameters": {
      "string" : "string"
    },
    "SerializationLibrary": "string"
  },
  "SkewedInfo": {
    "SkewedColumnNames": [ "string" ],
    "SkewedColumnValueLocationMaps": {
      "string" : "string"
    },
    "SkewedColumnValues": [ "string" ]
  },
  "SortColumns": [
    {
      "Column": "string",
      "SortOrder": number
    }
  ],
  "StoredAsSubDirectories": boolean
},
"TableType": "string",
"TargetTable": {
  "CatalogId": "string",
  "DatabaseName": "string",
```

```

    "Name": "string",
    "Region": "string"
  },
  "ViewDefinition": {
    "Definer": "string",
    "IsProtected": boolean,
    "Representations": [
      {
        "Dialect": "string",
        "DialectVersion": "string",
        "ValidationConnection": "string",
        "ViewExpandedText": "string",
        "ViewOriginalText": "string"
      }
    ],
    "SubObjects": [ "string" ]
  },
  "ViewExpandedText": "string",
  "ViewOriginalText": "string"
},
"TransactionId": "string",
"VersionId": "string",
"ViewUpdateAction": "string"
}

```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog where the table resides. If none is provided, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

DatabaseName

The name of the catalog database in which the table resides. For Hive compatibility, this name is entirely lowercase.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Force

A flag that can be set to true to ignore matching storage descriptor and subobject matching requirements.

Type: Boolean

Required: No

SkipArchive

By default, UpdateTable always creates an archived version of the table before updating it. However, if skipArchive is set to true, UpdateTable does not create the archived version.

Type: Boolean

Required: No

TableInput

An updated TableInput object to define the metadata table in the catalog.

Type: [TableInput](#) object

Required: Yes

TransactionId

The transaction ID at which to update the table contents.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\p{L}\p{N}\p{P}]*`

Required: No

VersionId

The version ID at which to update the table contents.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

ViewUpdateAction

The operation to be performed when updating the view.

Type: String

Valid Values: ADD | REPLACE | ADD_OR_REPLACE | DROP

Required: No

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

ResourceNotReadyException

A resource was not ready for a transaction.

HTTP Status Code: 400

ResourceNumberLimitExceededException

A resource numerical limit was exceeded.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)

- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

Required: Yes

TableName

The name of the table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

TableOptimizerConfiguration

A `TableOptimizerConfiguration` object representing the configuration of a table optimizer.

Type: [TableOptimizerConfiguration](#) object

Required: Yes

Type

The type of table optimizer. Currently, the only valid value is `compaction`.

Type: String

Valid Values: `compaction`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AccessDeniedException

Access to a resource was denied.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServerErrorException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

UpdateTrigger

Updates a trigger definition.

Request Syntax

```
{
  "Name": "string",
  "TriggerUpdate": {
    "Actions": [
      {
        "Arguments": {
          "string": "string"
        },
        "CrawlerName": "string",
        "JobName": "string",
        "NotificationProperty": {
          "NotifyDelayAfter": number
        },
        "SecurityConfiguration": "string",
        "Timeout": number
      }
    ],
    "Description": "string",
    "EventBatchingCondition": {
      "BatchSize": number,
      "BatchWindow": number
    },
    "Name": "string",
    "Predicate": {
      "Conditions": [
        {
          "CrawlerName": "string",
          "CrawlState": "string",
          "JobName": "string",
          "LogicalOperator": "string",
          "State": "string"
        }
      ],
      "Logical": "string"
    },
    "Schedule": "string"
  }
}
```

```
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Name

The name of the trigger to update.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

TriggerUpdate

The new values with which to update the trigger.

Type: [TriggerUpdate](#) object

Required: Yes

Response Syntax

```
{
  "Trigger": {
    "Actions": [
      {
        "Arguments": {
          "string" : "string"
        },
        "CrawlerName": "string",
        "JobName": "string",
        "NotificationProperty": {
          "NotifyDelayAfter": number
        },
        "SecurityConfiguration": "string",
        "Timeout": number
      }
    ]
  }
}
```



```
    }
  ],
  "Description": "string",
  "EventBatchingCondition": {
    "BatchSize": number,
    "BatchWindow": number
  },
  "Id": "string",
  "Name": "string",
  "Predicate": {
    "Conditions": [
      {
        "CrawlerName": "string",
        "CrawlState": "string",
        "JobName": "string",
        "LogicalOperator": "string",
        "State": "string"
      }
    ],
    "Logical": "string"
  },
  "Schedule": "string",
  "State": "string",
  "Type": "string",
  "WorkflowName": "string"
}
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Trigger

The resulting trigger definition.

Type: [Trigger](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)

- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

UpdateUsageProfile

Update an AWS Glue usage profile.

Request Syntax

```
{
  "Configuration": {
    "JobConfiguration": {
      "string": {
        "AllowedValues": [ "string" ],
        "DefaultValue": "string",
        "MaxValue": "string",
        "MinValue": "string"
      }
    },
    "SessionConfiguration": {
      "string": {
        "AllowedValues": [ "string" ],
        "DefaultValue": "string",
        "MaxValue": "string",
        "MinValue": "string"
      }
    }
  },
  "Description": "string",
  "Name": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

Configuration

A ProfileConfiguration object specifying the job and session values for the profile.

Type: [ProfileConfiguration](#) object

Required: Yes

Description

A description of the usage profile.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Name

The name of the usage profile.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{
  "Name": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Name

The name of the usage profile that was updated.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationNotSupportedException

The operation is not available in the region.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

UpdateUserDefinedFunction

Updates an existing function definition in the Data Catalog.

Request Syntax

```
{
  "CatalogId": "string",
  "DatabaseName": "string",
  "FunctionInput": {
    "ClassName": "string",
    "FunctionName": "string",
    "OwnerName": "string",
    "OwnerType": "string",
    "ResourceUris": [
      {
        "ResourceType": "string",
        "Uri": "string"
      }
    ]
  },
  "FunctionName": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

The request accepts the following data in JSON format.

CatalogId

The ID of the Data Catalog where the function to be updated is located. If none is provided, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF\\t]*`

Required: No

DatabaseName

The name of the catalog database where the function to be updated is located.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

FunctionInput

A `FunctionInput` object that redefines the function in the Data Catalog.

Type: [UserDefinedFunctionInput](#) object

Required: Yes

FunctionName

The name of the function.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

GlueEncryptionException

An encryption operation failed.

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

component jobs. If you leave this parameter blank, there is no limit to the number of concurrent workflow runs.

Type: Integer

Required: No

Name

Name of the workflow to be updated.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Response Syntax

```
{  
  "Name": "string"  
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Name

The name of the workflow which was specified in input.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ConcurrentModificationException

Two processes are trying to modify a resource simultaneously.

HTTP Status Code: 400

EntityNotFoundException

A specified entity does not exist

HTTP Status Code: 400

InternalServiceException

An internal service error occurred.

HTTP Status Code: 500

InvalidInputException

The input provided was not valid.

HTTP Status Code: 400

OperationTimeoutException

The operation timed out.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)

- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

Data Types

The AWS Glue API contains several data types that various actions use. This section describes each data type in detail.

Note

The order of each element in a data type structure is not guaranteed. Applications should not assume a particular order.

The following data types are supported:

- [Action](#)
- [Aggregate](#)
- [AggregateOperation](#)
- [AmazonRedshiftAdvancedOption](#)
- [AmazonRedshiftNodeData](#)
- [AmazonRedshiftSource](#)
- [AmazonRedshiftTarget](#)
- [ApplyMapping](#)
- [AthenaConnectorSource](#)
- [AuditContext](#)
- [AuthenticationConfiguration](#)
- [AuthenticationConfigurationInput](#)
- [AuthorizationCodeProperties](#)
- [BackfillError](#)
- [BasicCatalogTarget](#)
- [BatchGetTableOptimizerEntry](#)
- [BatchGetTableOptimizerError](#)
- [BatchStopJobRunError](#)
- [BatchStopJobRunSuccessfulSubmission](#)
- [BatchTableOptimizer](#)

- [BatchUpdatePartitionFailureEntry](#)
- [BatchUpdatePartitionRequestEntry](#)
- [BinaryColumnStatisticsData](#)
- [Blueprint](#)
- [BlueprintDetails](#)
- [BlueprintRun](#)
- [BooleanColumnStatisticsData](#)
- [CatalogDeltaSource](#)
- [CatalogEntry](#)
- [CatalogHudiSource](#)
- [CatalogImportStatus](#)
- [CatalogKafkaSource](#)
- [CatalogKinesisSource](#)
- [CatalogSchemaChangePolicy](#)
- [CatalogSource](#)
- [CatalogTarget](#)
- [Classifier](#)
- [CloudWatchEncryption](#)
- [CodeGenConfigurationNode](#)
- [CodeGenEdge](#)
- [CodeGenNode](#)
- [CodeGenNodeArg](#)
- [Column](#)
- [ColumnError](#)
- [ColumnImportance](#)
- [ColumnRowFilter](#)
- [ColumnStatistics](#)
- [ColumnStatisticsData](#)
- [ColumnStatisticsError](#)
- [ColumnStatisticsTaskRun](#)

- [Condition](#)
- [ConfigurationObject](#)
- [ConfusionMatrix](#)
- [Connection](#)
- [ConnectionInput](#)
- [ConnectionPasswordEncryption](#)
- [ConnectionsList](#)
- [ConnectorDataSource](#)
- [ConnectorDataTarget](#)
- [Crawl](#)
- [Crawler](#)
- [CrawlerHistory](#)
- [CrawlerMetrics](#)
- [CrawlerNodeDetails](#)
- [CrawlerTargets](#)
- [CrawlsFilter](#)
- [CreateCsvClassifierRequest](#)
- [CreateGrokClassifierRequest](#)
- [CreateJsonClassifierRequest](#)
- [CreateXMLClassifierRequest](#)
- [CsvClassifier](#)
- [CustomCode](#)
- [CustomEntityType](#)
- [Database](#)
- [DatabaseIdentifier](#)
- [DatabaseInput](#)
- [DataCatalogEncryptionSettings](#)
- [DataLakePrincipal](#)
- [DataQualityAnalyzerResult](#)
- [DataQualityEvaluationRunAdditionalRunOptions](#)

- [DataQualityMetricValues](#)
- [DataQualityObservation](#)
- [DataQualityResult](#)
- [DataQualityResultDescription](#)
- [DataQualityResultFilterCriteria](#)
- [DataQualityRuleRecommendationRunDescription](#)
- [DataQualityRuleRecommendationRunFilter](#)
- [DataQualityRuleResult](#)
- [DataQualityRulesetEvaluationRunDescription](#)
- [DataQualityRulesetEvaluationRunFilter](#)
- [DataQualityRulesetFilterCriteria](#)
- [DataQualityRulesetListDetails](#)
- [DataQualityTargetTable](#)
- [DataSource](#)
- [Datatype](#)
- [DateColumnStatisticsData](#)
- [DecimalColumnStatisticsData](#)
- [DecimalNumber](#)
- [DeltaTarget](#)
- [DevEndpoint](#)
- [DevEndpointCustomLibraries](#)
- [DirectJDBCSource](#)
- [DirectKafkaSource](#)
- [DirectKinesisSource](#)
- [DirectSchemaChangePolicy](#)
- [DoubleColumnStatisticsData](#)
- [DQResultsPublishingOptions](#)
- [DQStopJobOnFailureOptions](#)
- [DropDuplicates](#)
- [DropFields](#)

- [DropNullFields](#)
- [DynamicTransform](#)
- [DynamoDBCatalogSource](#)
- [DynamoDBTarget](#)
- [Edge](#)
- [EncryptionAtRest](#)
- [EncryptionConfiguration](#)
- [ErrorDetail](#)
- [ErrorDetails](#)
- [EvaluateDataQuality](#)
- [EvaluateDataQualityMultiFrame](#)
- [EvaluationMetrics](#)
- [EventBatchingCondition](#)
- [ExecutionProperty](#)
- [ExportLabelsTaskRunProperties](#)
- [FederatedDatabase](#)
- [FederatedTable](#)
- [FillMissingValues](#)
- [Filter](#)
- [FilterExpression](#)
- [FilterValue](#)
- [FindMatchesMetrics](#)
- [FindMatchesParameters](#)
- [FindMatchesTaskRunProperties](#)
- [GetConnectionsFilter](#)
- [GluePolicy](#)
- [GlueSchema](#)
- [GlueStudioSchemaColumn](#)
- [GlueTable](#)
- [GovernedCatalogSource](#)

- [GovernedCatalogTarget](#)
- [GrokClassifier](#)
- [HudiTarget](#)
- [IcebergInput](#)
- [IcebergTarget](#)
- [ImportLabelsTaskRunProperties](#)
- [JDBCConectorOptions](#)
- [JDBCConectorSource](#)
- [JDBCConectorTarget](#)
- [JdbcTarget](#)
- [Job](#)
- [JobBookmarkEntry](#)
- [JobBookmarksEncryption](#)
- [JobCommand](#)
- [JobNodeDetails](#)
- [JobRun](#)
- [JobUpdate](#)
- [Join](#)
- [JoinColumn](#)
- [JsonClassifier](#)
- [KafkaStreamingSourceOptions](#)
- [KeySchemaElement](#)
- [KinesisStreamingSourceOptions](#)
- [LabelingSetGenerationTaskRunProperties](#)
- [LakeFormationConfiguration](#)
- [LastActiveDefinition](#)
- [LastCrawlInfo](#)
- [LineageConfiguration](#)
- [Location](#)
- [LongColumnStatisticsData](#)

- [Mapping](#)
- [MappingEntry](#)
- [Merge](#)
- [MetadataInfo](#)
- [MetadataKeyValuePair](#)
- [MetricBasedObservation](#)
- [MicrosoftSQLServerCatalogSource](#)
- [MicrosoftSQLServerCatalogTarget](#)
- [MLTransform](#)
- [MLUserDataEncryption](#)
- [MongoDBTarget](#)
- [MySQLCatalogSource](#)
- [MySQLCatalogTarget](#)
- [Node](#)
- [NotificationProperty](#)
- [NullCheckBoxList](#)
- [NullValueField](#)
- [OAuth2ClientApplication](#)
- [OAuth2Properties](#)
- [OAuth2PropertiesInput](#)
- [OpenTableFormatInput](#)
- [Option](#)
- [OracleSQLCatalogSource](#)
- [OracleSQLCatalogTarget](#)
- [Order](#)
- [OtherMetadataValueListItem](#)
- [Partition](#)
- [PartitionError](#)
- [PartitionIndex](#)
- [PartitionIndexDescriptor](#)

- [PartitionInput](#)
- [PartitionValueList](#)
- [PhysicalConnectionRequirements](#)
- [PIIDetection](#)
- [PostgreSQLCatalogSource](#)
- [PostgreSQLCatalogTarget](#)
- [Predecessor](#)
- [Predicate](#)
- [PrincipalPermissions](#)
- [ProfileConfiguration](#)
- [PropertyPredicate](#)
- [QuerySessionContext](#)
- [Recipe](#)
- [RecipeReference](#)
- [RecrawlPolicy](#)
- [RedshiftSource](#)
- [RedshiftTarget](#)
- [RegistryId](#)
- [RegistryListItem](#)
- [RelationalCatalogSource](#)
- [RenameField](#)
- [ResourceUri](#)
- [RunMetrics](#)
- [S3CatalogDeltaSource](#)
- [S3CatalogHudiSource](#)
- [S3CatalogSource](#)
- [S3CatalogTarget](#)
- [S3CsvSource](#)
- [S3DeltaCatalogTarget](#)
- [S3DeltaDirectTarget](#)

- [S3DeltaSource](#)
- [S3DirectSourceAdditionalOptions](#)
- [S3DirectTarget](#)
- [S3Encryption](#)
- [S3GlueParquetTarget](#)
- [S3HudiCatalogTarget](#)
- [S3HudiDirectTarget](#)
- [S3HudiSource](#)
- [S3JsonSource](#)
- [S3ParquetSource](#)
- [S3SourceAdditionalOptions](#)
- [S3Target](#)
- [Schedule](#)
- [SchemaChangePolicy](#)
- [SchemaColumn](#)
- [SchemaId](#)
- [SchemaListItem](#)
- [SchemaReference](#)
- [SchemaVersionErrorItem](#)
- [SchemaVersionListItem](#)
- [SchemaVersionNumber](#)
- [SecurityConfiguration](#)
- [Segment](#)
- [SelectFields](#)
- [SelectFromCollection](#)
- [SerDeInfo](#)
- [Session](#)
- [SessionCommand](#)
- [SkewedInfo](#)
- [SnowflakeNodeData](#)

- [SnowflakeSource](#)
- [SnowflakeTarget](#)
- [SortCriterion](#)
- [SourceControlDetails](#)
- [SparkConnectorSource](#)
- [SparkConnectorTarget](#)
- [SparkSQL](#)
- [Spigot](#)
- [SplitFields](#)
- [SqlAlias](#)
- [StartingEventBatchCondition](#)
- [Statement](#)
- [StatementOutput](#)
- [StatementOutputData](#)
- [StorageDescriptor](#)
- [StreamingDataPreviewOptions](#)
- [StringColumnStatisticsData](#)
- [SupportedDialect](#)
- [Table](#)
- [TableError](#)
- [TableIdentifier](#)
- [TableInput](#)
- [TableOptimizer](#)
- [TableOptimizerConfiguration](#)
- [TableOptimizerRun](#)
- [TableVersion](#)
- [TableVersionError](#)
- [TaskRun](#)
- [TaskRunFilterCriteria](#)
- [TaskRunProperties](#)

- [TaskRunSortCriteria](#)
- [TransformConfigParameter](#)
- [TransformEncryption](#)
- [TransformFilterCriteria](#)
- [TransformParameters](#)
- [TransformSortCriteria](#)
- [Trigger](#)
- [TriggerNodeDetails](#)
- [TriggerUpdate](#)
- [UnfilteredPartition](#)
- [Union](#)
- [UpdateCsvClassifierRequest](#)
- [UpdateGrokClassifierRequest](#)
- [UpdateJsonClassifierRequest](#)
- [UpdateXMLClassifierRequest](#)
- [UpsertRedshiftTargetOptions](#)
- [UsageProfileDefinition](#)
- [UserDefinedFunction](#)
- [UserDefinedFunctionInput](#)
- [ViewDefinition](#)
- [ViewDefinitionInput](#)
- [ViewRepresentation](#)
- [ViewRepresentationInput](#)
- [Workflow](#)
- [WorkflowGraph](#)
- [WorkflowRun](#)
- [WorkflowRunStatistics](#)
- [XMLClassifier](#)

Action

Defines an action to be initiated by a trigger.

Contents

Arguments

The job arguments used when this trigger fires. For this job run, they replace the default arguments set in the job definition itself.

You can specify arguments here that your own job-execution script consumes, as well as arguments that AWS Glue itself consumes.

For information about how to specify and consume your own Job arguments, see the [Calling AWS Glue APIs in Python](#) topic in the developer guide.

For information about the key-value pairs that AWS Glue consumes to set up your job, see the [Special Parameters Used by AWS Glue](#) topic in the developer guide.

Type: String to string map

Required: No

CrawlerName

The name of the crawler to be used with this action.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

JobName

The name of a job to be run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF\\t]*`

Required: No

NotificationProperty

Specifies configuration properties of a job run notification.

Type: [NotificationProperty](#) object

Required: No

SecurityConfiguration

The name of the SecurityConfiguration structure to be used with this action.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF\\t]*`

Required: No

Timeout

The JobRun timeout in minutes. This is the maximum time that a job run can consume resources before it is terminated and enters TIMEOUT status. The default is 2,880 minutes (48 hours). This overrides the timeout value set in the parent job.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for Ruby V3](#)

Aggregate

Specifies a transform that groups rows by chosen fields and computes the aggregated value by specified function.

Contents

Aggs

Specifies the aggregate functions to be performed on specified fields.

Type: Array of [AggregateOperation](#) objects

Array Members: Minimum number of 1 item. Maximum number of 30 items.

Required: Yes

Groups

Specifies the fields to group by.

Type: Array of arrays of strings

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF] | [^\\S\\x\\n"']`)*

Required: Yes

Inputs

Specifies the fields and rows to use as inputs for the aggregate transform.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]`*

Required: Yes

Name

The name of the transform node.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]| [^\r\n]`)*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

AggregateOperation

Specifies the set of parameters needed to perform aggregation in the aggregate transform.

Contents

AggFunc

Specifies the aggregation function to apply.

Possible aggregation functions include: avg countDistinct, count, first, last, kurtosis, max, min, skewness, stddev_samp, stddev_pop, sum, sumDistinct, var_samp, var_pop

Type: String

Valid Values: avg | countDistinct | count | first | last | kurtosis | max | min | skewness | stddev_samp | stddev_pop | sum | sumDistinct | var_samp | var_pop

Required: Yes

Column

Specifies the column on the data set on which the aggregation function will be applied.

Type: Array of strings

Pattern: (`[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF]` | `[\^S\r\n"']`)*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

AmazonRedshiftAdvancedOption

Specifies an optional value when connecting to the Redshift cluster.

Contents

Key

The key for the additional connection option.

Type: String

Required: No

Value

The value for the additional connection option.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

AmazonRedshiftNodeData

Specifies an Amazon Redshift node.

Contents

AccessType

The access type for the Redshift connection. Can be a direct connection or catalog connections.

Type: String

Pattern: [A-Za-z0-9_-]*

Required: No

Action

Specifies how writing to a Redshift cluster will occur.

Type: String

Required: No

AdvancedOptions

Optional values when connecting to the Redshift cluster.

Type: Array of [AmazonRedshiftAdvancedOption](#) objects

Required: No

CatalogDatabase

The name of the AWS Glue Data Catalog database when working with a data catalog.

Type: [Option](#) object

Required: No

CatalogRedshiftSchema

The Redshift schema name when working with a data catalog.

Type: String

Required: No

CatalogRedshiftTable

The database table to read from.

Type: String

Required: No

CatalogTable

The AWS Glue Data Catalog table name when working with a data catalog.

Type: [Option](#) object

Required: No

Connection

The AWS Glue connection to the Redshift cluster.

Type: [Option](#) object

Required: No

CrawlerConnection

Specifies the name of the connection that is associated with the catalog table used.

Type: String

Required: No

IamRole

Optional. The role name use when connection to S3. The IAM role ill default to the role on the job when left blank.

Type: [Option](#) object

Required: No

MergeAction

The action used when to detemine how a MERGE in a Redshift sink will be handled.

Type: String

Pattern: [A-Za-z0-9_-]*

Required: No

MergeClause

The SQL used in a custom merge to deal with matching records.

Type: String

Required: No

MergeWhenMatched

The action used when to determine how a MERGE in a Redshift sink will be handled when an existing record matches a new record.

Type: String

Pattern: [A-Za-z0-9_-]*

Required: No

MergeWhenNotMatched

The action used when to determine how a MERGE in a Redshift sink will be handled when an existing record doesn't match a new record.

Type: String

Pattern: [A-Za-z0-9_-]*

Required: No

PostAction

The SQL used before a MERGE or APPEND with upsert is run.

Type: String

Required: No

PreAction

The SQL used before a MERGE or APPEND with upsert is run.

Type: String

Required: No

SampleQuery

The SQL used to fetch the data from a Redshift sources when the SourceType is 'query'.

Type: String

Required: No

Schema

The Redshift schema name when working with a direct connection.

Type: [Option](#) object

Required: No

SelectedColumns

The list of column names used to determine a matching record when doing a MERGE or APPEND with upsert.

Type: Array of [Option](#) objects

Required: No

SourceType

The source type to specify whether a specific table is the source or a custom query.

Type: String

Pattern: [A-Za-z0-9_-]*

Required: No

StagingTable

The name of the temporary staging table that is used when doing a MERGE or APPEND with upsert.

Type: String

Required: No

Table

The Redshift table name when working with a direct connection.

Type: [Option](#) object

Required: No

TablePrefix

Specifies the prefix to a table.

Type: String

Pattern: `[A-Za-z0-9_ -]*`

Required: No

TableSchema

The array of schema output for a given node.

Type: Array of [Option](#) objects

Required: No

TempDir

The Amazon S3 path where temporary data can be staged when copying out of the database.

Type: String

Pattern: `([\u0020-\u007F\u00E0-\u00FF\u0080-\u00FF | [^\S\r\n"'])*`

Required: No

Upsert

The action used on Redshift sinks when doing an APPEND.

Type: Boolean

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

AmazonRedshiftSource

Specifies an Amazon Redshift source.

Contents

Data

Specifies the data of the Amazon Redshift source node.

Type: [AmazonRedshiftNodeData](#) object

Required: No

Name

The name of the Amazon Redshift source.

Type: String

Pattern: (`[\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF]` | `[\^\r\n]`)*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

AmazonRedshiftTarget

Specifies an Amazon Redshift target.

Contents

Data

Specifies the data of the Amazon Redshift target node.

Type: [AmazonRedshiftNodeData](#) object

Required: No

Inputs

The nodes that are inputs to the data target.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: [A-Za-z0-9_-]*

Required: No

Name

The name of the Amazon Redshift target.

Type: String

Pattern: ([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\r\n])*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for Ruby V3](#)

ApplyMapping

Specifies a transform that maps data property keys in the data source to data property keys in the data target. You can rename keys, modify the data types for keys, and choose which keys to drop from the dataset.

Contents

Inputs

The data inputs identified by their node names.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]*`

Required: Yes

Mapping

Specifies the mapping of data property keys in the data source to data property keys in the data target.

Type: Array of [Mapping](#) objects

Required: Yes

Name

The name of the transform node.

Type: String

Pattern: `([\u0020-\u007F\u00E0-\u00FF\u0080-\u00FF\u0080-\u00FF\u0080-\u00FF\u0080-\u00FF] | [^\r\n])*`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

AthenaConnectorSource

Specifies a connector to an Amazon Athena data source.

Contents

ConnectionName

The name of the connection that is associated with the connector.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

ConnectionType

The type of connection, such as `marketplace.athena` or `custom.athena`, designating a connection to an Amazon Athena data store.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

ConnectorName

The name of a connector that assists with accessing the data store in AWS Glue Studio.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

Name

The name of the data source.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

SchemaName

The name of the Cloudwatch log group to read from. For example, `/aws-glue/jobs/output`.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

ConnectionTable

The name of the table in the data source.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n]`)*

Required: No

OutputSchemas

Specifies the data schema for the custom Athena source.

Type: Array of [GlueSchema](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

AuditContext

A structure containing the Lake Formation audit context.

Contents

AdditionalAuditContext

A string containing the additional audit context information.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Required: No

AllColumnsRequested

All columns request for audit.

Type: Boolean

Required: No

RequestedColumns

The requested columns for audit.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for Ruby V3](#)

AuthenticationConfiguration

A structure containing the authentication configuration.

Contents

AuthenticationType

A structure containing the authentication configuration.

Type: String

Valid Values: BASIC | OAUTH2 | CUSTOM

Required: No

OAuth2Properties

The properties for OAuth2 authentication.

Type: [OAuth2Properties](#) object

Required: No

SecretArn

The secret manager ARN to store credentials.

Type: String

Pattern: `^arn:aws(-(cn|us-gov|iso(-[bef]))?)?:secretsmanager:.*$`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

AuthenticationConfigurationInput

A structure containing the authentication configuration in the CreateConnection request.

Contents

AuthenticationType

A structure containing the authentication configuration in the CreateConnection request.

Type: String

Valid Values: BASIC | OAUTH2 | CUSTOM

Required: No

OAuth2Properties

The properties for OAuth2 authentication in the CreateConnection request.

Type: [OAuth2PropertiesInput](#) object

Required: No

SecretArn

The secret manager ARN to store credentials in the CreateConnection request.

Type: String

Pattern: `^arn:aws(-[cn|us-gov|iso(-[bef])?]):secretsmanager:.*$`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

AuthorizationCodeProperties

The set of properties required for the the OAuth2 AUTHORIZATION_CODE grant type workflow.

Contents

AuthorizationCode

An authorization code to be used in the third leg of the AUTHORIZATION_CODE grant workflow. This is a single-use code which becomes invalid once exchanged for an access token, thus it is acceptable to have this value as a request parameter.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 4096.

Pattern: \S+

Required: No

RedirectUri

The redirect URI where the user gets redirected to by authorization server when issuing an authorization code. The URI is subsequently used when the authorization code is exchanged for an access token.

Type: String

Length Constraints: Maximum length of 512.

Pattern: ^(https?):\:\/\/[^\s/\$.?\#] . [^\s]*\$

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for Ruby V3](#)

BackfillError

A list of errors that can occur when registering partition indexes for an existing table.

These errors give the details about why an index registration failed and provide a limited number of partitions in the response, so that you can fix the partitions at fault and try registering the index again. The most common set of errors that can occur are categorized as follows:

- **EncryptedPartitionError**: The partitions are encrypted.
- **InvalidPartitionTypeDataError**: The partition value doesn't match the data type for that partition column.
- **MissingPartitionValueError**: The partitions are encrypted.
- **UnsupportedPartitionCharacterError**: Characters inside the partition value are not supported. For example: U+0000 , U+0001, U+0002.
- **InternalError**: Any error which does not belong to other error codes.

Contents

Code

The error code for an error that occurred when registering partition indexes for an existing table.

Type: String

Valid Values: ENCRYPTED_PARTITION_ERROR | INTERNAL_ERROR | INVALID_PARTITION_TYPE_DATA_ERROR | MISSING_PARTITION_VALUE_ERROR | UNSUPPORTED_PARTITION_CHARACTER_ERROR

Required: No

Partitions

A list of a limited number of partitions in the response.

Type: Array of [PartitionValueList](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

BasicCatalogTarget

Specifies a target that uses a AWS Glue Data Catalog table.

Contents

Database

The database that contains the table you want to use as the target. This database must already exist in the Data Catalog.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

Inputs

The nodes that are inputs to the data target.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]`*

Required: Yes

Name

The name of your data target.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

Table

The table that defines the schema of your output data. This table must already exist in the Data Catalog.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF]|[\^\S\r\n"'])`)*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

BatchGetTableOptimizerEntry

Represents a table optimizer to retrieve in the `BatchGetTableOptimizer` operation.

Contents

catalogId

The Catalog ID of the table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

databaseName

The name of the database in the catalog in which the table resides.

Type: String

Length Constraints: Minimum length of 1.

Required: No

tableName

The name of the table.

Type: String

Length Constraints: Minimum length of 1.

Required: No

type

The type of table optimizer.

Type: String

Valid Values: `compaction`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

BatchGetTableOptimizerError

Contains details on one of the errors in the error list returned by the `BatchGetTableOptimizer` operation.

Contents

catalogId

The Catalog ID of the table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

databaseName

The name of the database in the catalog in which the table resides.

Type: String

Length Constraints: Minimum length of 1.

Required: No

error

An `ErrorDetail` object containing code and message details about the error.

Type: [ErrorDetail](#) object

Required: No

tableName

The name of the table.

Type: String

Length Constraints: Minimum length of 1.

Required: No

type

The type of table optimizer.

Type: String

Valid Values: `compaction`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

BatchStopJobRunError

Records an error that occurred when attempting to stop a specified job run.

Contents

ErrorDetail

Specifies details about the error that was encountered.

Type: [ErrorDetail](#) object

Required: No

JobName

The name of the job definition that is used in the job run in question.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

JobRunId

The JobRunId of the job run in question.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

BatchStopJobRunSuccessfulSubmission

Records a successful request to stop a specified JobRun.

Contents

JobName

The name of the job definition used in the job run that was stopped.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

JobRunId

The JobRunId of the job run that was stopped.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

BatchTableOptimizer

Contains details for one of the table optimizers returned by the `BatchGetTableOptimizer` operation.

Contents

catalogId

The Catalog ID of the table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

databaseName

The name of the database in the catalog in which the table resides.

Type: String

Length Constraints: Minimum length of 1.

Required: No

tableName

The name of the table.

Type: String

Length Constraints: Minimum length of 1.

Required: No

tableOptimizer

A `TableOptimizer` object that contains details on the configuration and last run of a table optimizer.

Type: [TableOptimizer](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

BatchUpdatePartitionFailureEntry

Contains information about a batch update partition error.

Contents

ErrorDetail

The details about the batch update partition error.

Type: [ErrorDetail](#) object

Required: No

PartitionValueList

A list of values defining the partitions.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 100 items.

Length Constraints: Maximum length of 1024.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

BatchUpdatePartitionRequestEntry

A structure that contains the values and structure used to update a partition.

Contents

PartitionInput

The structure used to update a partition.

Type: [PartitionInput](#) object

Required: Yes

PartitionValueList

A list of values defining the partitions.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 100 items.

Length Constraints: Maximum length of 1024.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

BinaryColumnStatisticsData

Defines column statistics supported for bit sequence data values.

Contents

AverageLength

The average bit sequence length in the column.

Type: Double

Valid Range: Minimum value of 0.0.

Required: Yes

MaximumLength

The size of the longest bit sequence in the column.

Type: Long

Valid Range: Minimum value of 0.

Required: Yes

NumberOfNulls

The number of null values in the column.

Type: Long

Valid Range: Minimum value of 0.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for Ruby V3](#)

Blueprint

The details of a blueprint.

Contents

BlueprintLocation

Specifies the path in Amazon S3 where the blueprint is published.

Type: String

Required: No

BlueprintServiceLocation

Specifies a path in Amazon S3 where the blueprint is copied when you call `CreateBlueprint/UpdateBlueprint` to register the blueprint in AWS Glue.

Type: String

Required: No

CreatedOn

The date and time the blueprint was registered.

Type: Timestamp

Required: No

Description

The description of the blueprint.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 512.

Required: No

ErrorMessage

An error message.

Type: String

Required: No

LastActiveDefinition

When there are multiple versions of a blueprint and the latest version has some errors, this attribute indicates the last successful blueprint definition that is available with the service.

Type: [LastActiveDefinition](#) object

Required: No

LastModifiedOn

The date and time the blueprint was last modified.

Type: Timestamp

Required: No

Name

The name of the blueprint.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[\.\-_\A-Za-z0-9]+`

Required: No

ParameterSpec

A JSON string that indicates the list of parameter specifications for the blueprint.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 131072.

Required: No

Status

The status of the blueprint registration.

- **Creating** — The blueprint registration is in progress.
- **Active** — The blueprint has been successfully registered.

- **Updating** — An update to the blueprint registration is in progress.
- **Failed** — The blueprint registration failed.

Type: String

Valid Values: CREATING | ACTIVE | UPDATING | FAILED

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

BlueprintDetails

The details of a blueprint.

Contents

BlueprintName

The name of the blueprint.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[\.\-_\A-Za-z0-9]+`

Required: No

RunId

The run ID for this blueprint.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

BlueprintRun

The details of a blueprint run.

Contents

BlueprintName

The name of the blueprint.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[\.\-_\A-Za-z0-9]+`

Required: No

CompletedOn

The date and time that the blueprint run completed.

Type: Timestamp

Required: No

ErrorMessage

Indicates any errors that are seen while running the blueprint.

Type: String

Required: No

Parameters

The blueprint parameters as a string. You will have to provide a value for each key that is required from the parameter spec that is defined in the `Blueprint$ParameterSpec`.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 131072.

Required: No

RoleArn

The role ARN. This role will be assumed by the AWS Glue service and will be used to create the workflow and other entities of a workflow.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `arn:aws[^\:]*:iam:[0-9]*:role/.+`

Required: No

RollbackErrorMessage

If there are any errors while creating the entities of a workflow, we try to roll back the created entities until that point and delete them. This attribute indicates the errors seen while trying to delete the entities that are created.

Type: String

Required: No

RunId

The run ID for this blueprint run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

StartedOn

The date and time that the blueprint run started.

Type: Timestamp

Required: No

State

The state of the blueprint run. Possible values are:

- **Running** — The blueprint run is in progress.
- **Succeeded** — The blueprint run completed successfully.
- **Failed** — The blueprint run failed and rollback is complete.
- **Rolling Back** — The blueprint run failed and rollback is in progress.

Type: String

Valid Values: RUNNING | SUCCEEDED | FAILED | ROLLING_BACK

Required: No

WorkflowName

The name of a workflow that is created as a result of a successful blueprint run. If a blueprint run has an error, there will not be a workflow created.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

BooleanColumnStatisticsData

Defines column statistics supported for Boolean data columns.

Contents

NumberOfFalses

The number of false values in the column.

Type: Long

Valid Range: Minimum value of 0.

Required: Yes

NumberOfNulls

The number of null values in the column.

Type: Long

Valid Range: Minimum value of 0.

Required: Yes

NumberOfTrues

The number of true values in the column.

Type: Long

Valid Range: Minimum value of 0.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for Ruby V3](#)

CatalogDeltaSource

Specifies a Delta Lake data source that is registered in the AWS Glue Data Catalog.

Contents

Database

The name of the database to read from.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

Name

The name of the Delta Lake data source.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

Table

The name of the table in the database to read from.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

AdditionalDeltaOptions

Specifies additional connection options.

Type: String to string map

Key Pattern: (`([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])`)*

Value Pattern: (`([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])`)*

Required: No

OutputSchemas

Specifies the data schema for the Delta Lake source.

Type: Array of [GlueSchema](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CatalogEntry

Specifies a table definition in the AWS Glue Data Catalog.

Contents

DatabaseName

The database in which the table metadata resides.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

TableName

The name of the table in question.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CatalogHudiSource

Specifies a Hudi data source that is registered in the AWS Glue Data Catalog.

Contents

Database

The name of the database to read from.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

Name

The name of the Hudi data source.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

Table

The name of the table in the database to read from.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

AdditionalHudiOptions

Specifies additional connection options.

Type: String to string map

Key Pattern: ([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF]|[\^\S\r\n"'])*

Value Pattern: ([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF]|[\^\S\r\n"'])*

Required: No

OutputSchemas

Specifies the data schema for the Hudi source.

Type: Array of [GlueSchema](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CatalogImportStatus

A structure containing migration status information.

Contents

ImportCompleted

True if the migration has completed, or False otherwise.

Type: Boolean

Required: No

ImportedBy

The name of the person who initiated the migration.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

ImportTime

The time that the migration was started.

Type: Timestamp

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CatalogKafkaSource

Specifies an Apache Kafka data store in the Data Catalog.

Contents

Database

The name of the database to read from.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])*`)

Required: Yes

Name

The name of the data store.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\r\n])*`)

Required: Yes

Table

The name of the table in the database to read from.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])*`)

Required: Yes

DataPreviewOptions

Specifies options related to data preview for viewing a sample of your data.

Type: [StreamingDataPreviewOptions](#) object

Required: No

DetectSchema

Whether to automatically determine the schema from the incoming data.

Type: Boolean

Required: No

StreamingOptions

Specifies the streaming options.

Type: [KafkaStreamingSourceOptions](#) object

Required: No

WindowSize

The amount of time to spend processing each micro batch.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CatalogKinesisSource

Specifies a Kinesis data source in the AWS Glue Data Catalog.

Contents

Database

The name of the database to read from.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])*`)

Required: Yes

Name

The name of the data source.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\r\n])*`)

Required: Yes

Table

The name of the table in the database to read from.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])*`)

Required: Yes

DataPreviewOptions

Additional options for data preview.

Type: [StreamingDataPreviewOptions](#) object

Required: No

DetectSchema

Whether to automatically determine the schema from the incoming data.

Type: Boolean

Required: No

StreamingOptions

Additional options for the Kinesis streaming data source.

Type: [KinesisStreamingSourceOptions](#) object

Required: No

WindowSize

The amount of time to spend processing each micro batch.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CatalogSchemaChangePolicy

A policy that specifies update behavior for the crawler.

Contents

EnableUpdateCatalog

Whether to use the specified update behavior when the crawler finds a changed schema.

Type: Boolean

Required: No

UpdateBehavior

The update behavior when the crawler finds a changed schema.

Type: String

Valid Values: UPDATE_IN_DATABASE | LOG

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CatalogSource

Specifies a data store in the AWS Glue Data Catalog.

Contents

Database

The name of the database to read from.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | (`^\\S\\r\\n"'`))*

Required: Yes

Name

The name of the data store.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | (`^\\r\\n`))*

Required: Yes

Table

The name of the table in the database to read from.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | (`^\\S\\r\\n"'`))*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CatalogTarget

Specifies an AWS Glue Data Catalog target.

Contents

DatabaseName

The name of the database to be synchronized.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Tables

A list of the tables to be synchronized.

Type: Array of strings

Array Members: Minimum number of 1 item.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

ConnectionName

The name of the connection for an Amazon S3-backed Data Catalog table to be a target of the crawl when using a Catalog connection type paired with a NETWORK Connection type.

Type: String

Required: No

DLqEventQueueArn

A valid Amazon dead-letter SQS ARN. For example,
`arn:aws:sqs:region:account:deadLetterQueue.`

Type: String

Required: No

EventQueueArn

A valid Amazon SQS ARN. For example, `arn:aws:sqs:region:account:sqs`.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Classifier

Classifiers are triggered during a crawl task. A classifier checks whether a given file is in a format it can handle. If it is, the classifier creates a schema in the form of a `StructType` object that matches that data format.

You can use the standard classifiers that AWS Glue provides, or you can write your own classifiers to best categorize your data sources and specify the appropriate schemas to use for them. A classifier can be a grok classifier, an XML classifier, a JSON classifier, or a custom CSV classifier, as specified in one of the fields in the `Classifier` object.

Contents

CsvClassifier

A classifier for comma-separated values (CSV).

Type: [CsvClassifier](#) object

Required: No

GrokClassifier

A classifier that uses `grok`.

Type: [GrokClassifier](#) object

Required: No

JsonClassifier

A classifier for JSON content.

Type: [JsonClassifier](#) object

Required: No

XMLClassifier

A classifier for XML content.

Type: [XMLClassifier](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CloudWatchEncryption

Specifies how Amazon CloudWatch data should be encrypted.

Contents

CloudWatchEncryptionMode

The encryption mode to use for CloudWatch data.

Type: String

Valid Values: DISABLED | SSE-KMS

Required: No

KmsKeyArn

The Amazon Resource Name (ARN) of the KMS key to be used to encrypt the data.

Type: String

Pattern: arn:aws:kms:.*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CodeGenConfigurationNode

CodeGenConfigurationNode enumerates all valid Node types. One and only one of its member variables can be populated.

Contents

Aggregate

Specifies a transform that groups rows by chosen fields and computes the aggregated value by specified function.

Type: [Aggregate](#) object

Required: No

AmazonRedshiftSource

Specifies a target that writes to a data source in Amazon Redshift.

Type: [AmazonRedshiftSource](#) object

Required: No

AmazonRedshiftTarget

Specifies a target that writes to a data target in Amazon Redshift.

Type: [AmazonRedshiftTarget](#) object

Required: No

ApplyMapping

Specifies a transform that maps data property keys in the data source to data property keys in the data target. You can rename keys, modify the data types for keys, and choose which keys to drop from the dataset.

Type: [ApplyMapping](#) object

Required: No

AthenaConnectorSource

Specifies a connector to an Amazon Athena data source.

Type: [AthenaConnectorSource](#) object

Required: No

CatalogDeltaSource

Specifies a Delta Lake data source that is registered in the AWS Glue Data Catalog.

Type: [CatalogDeltaSource](#) object

Required: No

CatalogHudiSource

Specifies a Hudi data source that is registered in the AWS Glue Data Catalog.

Type: [CatalogHudiSource](#) object

Required: No

CatalogKafkaSource

Specifies an Apache Kafka data store in the Data Catalog.

Type: [CatalogKafkaSource](#) object

Required: No

CatalogKinesisSource

Specifies a Kinesis data source in the AWS Glue Data Catalog.

Type: [CatalogKinesisSource](#) object

Required: No

CatalogSource

Specifies a data store in the AWS Glue Data Catalog.

Type: [CatalogSource](#) object

Required: No

CatalogTarget

Specifies a target that uses a AWS Glue Data Catalog table.

Type: [BasicCatalogTarget](#) object

Required: No

ConnectorDataSource

Specifies a source generated with standard connection options.

Type: [ConnectorDataSource](#) object

Required: No

ConnectorDataTarget

Specifies a target generated with standard connection options.

Type: [ConnectorDataTarget](#) object

Required: No

CustomCode

Specifies a transform that uses custom code you provide to perform the data transformation. The output is a collection of DynamicFrames.

Type: [CustomCode](#) object

Required: No

DirectJDBCSource

Specifies the direct JDBC source connection.

Type: [DirectJDBCSource](#) object

Required: No

DirectKafkaSource

Specifies an Apache Kafka data store.

Type: [DirectKafkaSource](#) object

Required: No

DirectKinesisSource

Specifies a direct Amazon Kinesis data source.

Type: [DirectKinesisSource](#) object

Required: No

DropDuplicates

Specifies a transform that removes rows of repeating data from a data set.

Type: [DropDuplicates](#) object

Required: No

DropFields

Specifies a transform that chooses the data property keys that you want to drop.

Type: [DropFields](#) object

Required: No

DropNullFields

Specifies a transform that removes columns from the dataset if all values in the column are 'null'. By default, AWS Glue Studio will recognize null objects, but some values such as empty strings, strings that are "null", -1 integers or other placeholders such as zeros, are not automatically recognized as nulls.

Type: [DropNullFields](#) object

Required: No

DynamicTransform

Specifies a custom visual transform created by a user.

Type: [DynamicTransform](#) object

Required: No

DynamoDBCatalogSource

Specifies a DynamoDBC Catalog data store in the AWS Glue Data Catalog.

Type: [DynamoDBCatalogSource](#) object

Required: No

EvaluateDataQuality

Specifies your data quality evaluation criteria.

Type: [EvaluateDataQuality](#) object

Required: No

EvaluateDataQualityMultiFrame

Specifies your data quality evaluation criteria. Allows multiple input data and returns a collection of Dynamic Frames.

Type: [EvaluateDataQualityMultiFrame](#) object

Required: No

FillMissingValues

Specifies a transform that locates records in the dataset that have missing values and adds a new field with a value determined by imputation. The input data set is used to train the machine learning model that determines what the missing value should be.

Type: [FillMissingValues](#) object

Required: No

Filter

Specifies a transform that splits a dataset into two, based on a filter condition.

Type: [Filter](#) object

Required: No

GovernedCatalogSource

Specifies a data source in a governed Data Catalog.

Type: [GovernedCatalogSource](#) object

Required: No

GovernedCatalogTarget

Specifies a data target that writes to a governed catalog.

Type: [GovernedCatalogTarget](#) object

Required: No

JDBCConnectorSource

Specifies a connector to a JDBC data source.

Type: [JDBCConnectorSource](#) object

Required: No

JDBCConnectorTarget

Specifies a data target that writes to Amazon S3 in Apache Parquet columnar storage.

Type: [JDBCConnectorTarget](#) object

Required: No

Join

Specifies a transform that joins two datasets into one dataset using a comparison phrase on the specified data property keys. You can use inner, outer, left, right, left semi, and left anti joins.

Type: [Join](#) object

Required: No

Merge

Specifies a transform that merges a `DynamicFrame` with a staging `DynamicFrame` based on the specified primary keys to identify records. Duplicate records (records with the same primary keys) are not de-duplicated.

Type: [Merge](#) object

Required: No

MicrosoftSQLServerCatalogSource

Specifies a Microsoft SQL server data source in the AWS Glue Data Catalog.

Type: [MicrosoftSQLServerCatalogSource](#) object

Required: No

MicrosoftSQLServerCatalogTarget

Specifies a target that uses Microsoft SQL.

Type: [MicrosoftSQLServerCatalogTarget](#) object

Required: No

MySQLCatalogSource

Specifies a MySQL data source in the AWS Glue Data Catalog.

Type: [MySQLCatalogSource](#) object

Required: No

MySQLCatalogTarget

Specifies a target that uses MySQL.

Type: [MySQLCatalogTarget](#) object

Required: No

OracleSQLCatalogSource

Specifies an Oracle data source in the AWS Glue Data Catalog.

Type: [OracleSQLCatalogSource](#) object

Required: No

OracleSQLCatalogTarget

Specifies a target that uses Oracle SQL.

Type: [OracleSQLCatalogTarget](#) object

Required: No

PIIDetection

Specifies a transform that identifies, removes or masks PII data.

Type: [PIIDetection](#) object

Required: No

PostgreSQLCatalogSource

Specifies a PostgreSQL data source in the AWS Glue Data Catalog.

Type: [PostgreSQLCatalogSource](#) object

Required: No

PostgreSQLCatalogTarget

Specifies a target that uses Postgres SQL.

Type: [PostgreSQLCatalogTarget](#) object

Required: No

Recipe

Specifies a AWS Glue DataBrew recipe node.

Type: [Recipe](#) object

Required: No

RedshiftSource

Specifies an Amazon Redshift data store.

Type: [RedshiftSource](#) object

Required: No

RedshiftTarget

Specifies a target that uses Amazon Redshift.

Type: [RedshiftTarget](#) object

Required: No

RelationalCatalogSource

Specifies a relational catalog data store in the AWS Glue Data Catalog.

Type: [RelationalCatalogSource](#) object

Required: No

RenameField

Specifies a transform that renames a single data property key.

Type: [RenameField](#) object

Required: No

S3CatalogDeltaSource

Specifies a Delta Lake data source that is registered in the AWS Glue Data Catalog. The data source must be stored in Amazon S3.

Type: [S3CatalogDeltaSource](#) object

Required: No

S3CatalogHudiSource

Specifies a Hudi data source that is registered in the AWS Glue Data Catalog. The data source must be stored in Amazon S3.

Type: [S3CatalogHudiSource](#) object

Required: No

S3CatalogSource

Specifies an Amazon S3 data store in the AWS Glue Data Catalog.

Type: [S3CatalogSource](#) object

Required: No

S3CatalogTarget

Specifies a data target that writes to Amazon S3 using the AWS Glue Data Catalog.

Type: [S3CatalogTarget](#) object

Required: No

S3CsvSource

Specifies a command-separated value (CSV) data store stored in Amazon S3.

Type: [S3CsvSource](#) object

Required: No

S3DeltaCatalogTarget

Specifies a target that writes to a Delta Lake data source in the AWS Glue Data Catalog.

Type: [S3DeltaCatalogTarget](#) object

Required: No

S3DeltaDirectTarget

Specifies a target that writes to a Delta Lake data source in Amazon S3.

Type: [S3DeltaDirectTarget](#) object

Required: No

S3DeltaSource

Specifies a Delta Lake data source stored in Amazon S3.

Type: [S3DeltaSource](#) object

Required: No

S3DirectTarget

Specifies a data target that writes to Amazon S3.

Type: [S3DirectTarget](#) object

Required: No

S3GlueParquetTarget

Specifies a data target that writes to Amazon S3 in Apache Parquet columnar storage.

Type: [S3GlueParquetTarget](#) object

Required: No

S3HudiCatalogTarget

Specifies a target that writes to a Hudi data source in the AWS Glue Data Catalog.

Type: [S3HudiCatalogTarget](#) object

Required: No

S3HudiDirectTarget

Specifies a target that writes to a Hudi data source in Amazon S3.

Type: [S3HudiDirectTarget](#) object

Required: No

S3HudiSource

Specifies a Hudi data source stored in Amazon S3.

Type: [S3HudiSource](#) object

Required: No

S3JsonSource

Specifies a JSON data store stored in Amazon S3.

Type: [S3JsonSource](#) object

Required: No

S3ParquetSource

Specifies an Apache Parquet data store stored in Amazon S3.

Type: [S3ParquetSource](#) object

Required: No

SelectFields

Specifies a transform that chooses the data property keys that you want to keep.

Type: [SelectFields](#) object

Required: No

SelectFromCollection

Specifies a transform that chooses one `DynamicFrame` from a collection of `DynamicFrames`. The output is the selected `DynamicFrame`

Type: [SelectFromCollection](#) object

Required: No

SnowflakeSource

Specifies a Snowflake data source.

Type: [SnowflakeSource](#) object

Required: No

SnowflakeTarget

Specifies a target that writes to a Snowflake data source.

Type: [SnowflakeTarget](#) object

Required: No

SparkConnectorSource

Specifies a connector to an Apache Spark data source.

Type: [SparkConnectorSource](#) object

Required: No

SparkConnectorTarget

Specifies a target that uses an Apache Spark connector.

Type: [SparkConnectorTarget](#) object

Required: No

SparkSQL

Specifies a transform where you enter a SQL query using Spark SQL syntax to transform the data. The output is a single `DynamicFrame`.

Type: [SparkSQL](#) object

Required: No

Spigot

Specifies a transform that writes samples of the data to an Amazon S3 bucket.

Type: [Spigot](#) object

Required: No

SplitFields

Specifies a transform that splits data property keys into two `DynamicFrames`. The output is a collection of `DynamicFrames`: one with selected data property keys, and one with the remaining data property keys.

Type: [SplitFields](#) object

Required: No

Union

Specifies a transform that combines the rows from two or more datasets into a single result.

Type: [Union](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CodeGenEdge

Represents a directional edge in a directed acyclic graph (DAG).

Contents

Source

The ID of the node at which the edge starts.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[A-Za-z_][A-Za-z0-9_]*`

Required: Yes

Target

The ID of the node at which the edge ends.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[A-Za-z_][A-Za-z0-9_]*`

Required: Yes

TargetParameter

The target of the edge.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CodeGenNode

Represents a node in a directed acyclic graph (DAG)

Contents

Args

Properties of the node, in the form of name-value pairs.

Type: Array of [CodeGenNodeArg](#) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: Yes

Id

A node identifier that is unique within the node's graph.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[A-Za-z_][A-Za-z0-9_]*`

Required: Yes

NodeType

The type of node that this is.

Type: String

Required: Yes

LineNumber

The line number of the node.

Type: Integer

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CodeGenNodeArg

An argument or property of a node.

Contents

Name

The name of the argument or property.

Type: String

Required: Yes

Value

The value of the argument or property.

Type: String

Required: Yes

Param

True if the value is used as a parameter.

Type: Boolean

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Column

A column in a Table.

Contents

Name

The name of the Column.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Comment

A free-form text comment.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Parameters

These key-value pairs define properties associated with the column.

Type: String to string map

Key Length Constraints: Minimum length of 1. Maximum length of 255.

Key Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Value Length Constraints: Maximum length of 512000.

Required: No

Type

The data type of the Column.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 131072.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ColumnError

Encapsulates a column name that failed and the reason for failure.

Contents

ColumnName

The name of the column that failed.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Error

An error message with the reason for the failure of an operation.

Type: [ErrorDetail](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ColumnImportance

A structure containing the column name and column importance score for a column.

Column importance helps you understand how columns contribute to your model, by identifying which columns in your records are more important than others.

Contents

ColumnName

The name of a column.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Importance

The column importance score for the column, as a decimal.

Type: Double

Valid Range: Minimum value of 0.0. Maximum value of 1.0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ColumnRowFilter

A filter that uses both column-level and row-level filtering.

Contents

ColumnName

A string containing the name of the column.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

RowFilterExpression

A string containing the row-level filter expression.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ColumnStatistics

Represents the generated column-level statistics for a table or partition.

Contents

AnalyzedTime

The timestamp of when column statistics were generated.

Type: Timestamp

Required: Yes

ColumnName

Name of column which statistics belong to.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

ColumnType

The data type of the column.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 20000.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

StatisticsData

A `ColumnStatisticData` object that contains the statistics data values.

Type: [ColumnStatisticsData](#) object

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ColumnStatisticsData

Contains the individual types of column statistics data. Only one data object should be set and indicated by the Type attribute.

Contents

Type

The type of column statistics data.

Type: String

Valid Values: BOOLEAN | DATE | DECIMAL | DOUBLE | LONG | STRING | BINARY

Required: Yes

BinaryColumnStatisticsData

Binary column statistics data.

Type: [BinaryColumnStatisticsData](#) object

Required: No

BooleanColumnStatisticsData

Boolean column statistics data.

Type: [BooleanColumnStatisticsData](#) object

Required: No

DateColumnStatisticsData

Date column statistics data.

Type: [DateColumnStatisticsData](#) object

Required: No

DecimalColumnStatisticsData

Decimal column statistics data. UnscaledValues within are Base64-encoded binary objects storing big-endian, two's complement representations of the decimal's unscaled value.

Type: [DecimalColumnStatisticsData](#) object

Required: No

DoubleColumnStatisticsData

Double column statistics data.

Type: [DoubleColumnStatisticsData](#) object

Required: No

LongColumnStatisticsData

Long column statistics data.

Type: [LongColumnStatisticsData](#) object

Required: No

StringColumnStatisticsData

String column statistics data.

Type: [StringColumnStatisticsData](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ColumnStatisticsError

Encapsulates a `ColumnStatistics` object that failed and the reason for failure.

Contents

ColumnStatistics

The `ColumnStatistics` of the column.

Type: [ColumnStatistics](#) object

Required: No

Error

An error message with the reason for the failure of an operation.

Type: [ErrorDetail](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ColumnStatisticsTaskRun

The object that shows the details of the column stats run.

Contents

CatalogID

The ID of the Data Catalog where the table resides. If none is supplied, the AWS account ID is used by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

ColumnNameList

A list of the column names. If none is supplied, all column names for the table will be used by default.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

ColumnStatisticsTaskRunId

The identifier for the particular column statistics task run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

CreationTime

The time that this task was created.

Type: Timestamp

Required: No

CustomerId

The AWS account ID.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 12.

Required: No

DatabaseName

The database where the table resides.

Type: String

Required: No

DPUSeconds

The calculated DPU usage in seconds for all autoscaled workers.

Type: Double

Valid Range: Minimum value of 0.0.

Required: No

EndTime

The end time of the task.

Type: Timestamp

Required: No

ErrorMessage

The error message for the job.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

LastUpdated

The last point in time when this task was modified.

Type: Timestamp

Required: No

NumberOfWorkers

The number of workers used to generate column statistics. The job is preconfigured to autoscale up to 25 instances.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

Role

The IAM role that the service assumes to generate statistics.

Type: String

Required: No

SampleSize

The percentage of rows used to generate statistics. If none is supplied, the entire table will be used to generate stats.

Type: Double

Valid Range: Minimum value of 0. Maximum value of 100.

Required: No

SecurityConfiguration

Name of the security configuration that is used to encrypt CloudWatch logs for the column stats task run.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 128.

Required: No

StartTime

The start time of the task.

Type: Timestamp

Required: No

Status

The status of the task run.

Type: String

Valid Values: STARTING | RUNNING | SUCCEEDED | FAILED | STOPPED

Required: No

TableName

The name of the table for which column statistics is generated.

Type: String

Required: No

WorkerType

The type of workers being used for generating stats. The default is g.1x.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Condition

Defines a condition under which a trigger fires.

Contents

CrawlerName

The name of the crawler to which this condition applies.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

CrawlState

The state of the crawler to which this condition applies.

Type: String

Valid Values: RUNNING | CANCELLING | CANCELLED | SUCCEEDED | FAILED | ERROR

Required: No

JobName

The name of the job whose JobRuns this condition applies to, and on which this trigger waits.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

LogicalOperator

A logical operator.

Type: String

Valid Values: EQUALS

Required: No

State

The condition state. Currently, the only job states that a trigger can listen for are SUCCEEDED, STOPPED, FAILED, and TIMEOUT. The only crawler states that a trigger can listen for are SUCCEEDED, FAILED, and CANCELLED.

Type: String

Valid Values: STARTING | RUNNING | STOPPING | STOPPED | SUCCEEDED | FAILED | TIMEOUT | ERROR | WAITING | EXPIRED

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ConfigurationObject

Specifies the values that an admin sets for each job or session parameter configured in a AWS Glue usage profile.

Contents

AllowedValues

A list of allowed values for the parameter.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: [a-zA-Z0-9_.-]+

Required: No

DefaultValue

A default value for the parameter.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: [a-zA-Z0-9_.-]+

Required: No

MaxValue

A maximum allowed value for the parameter.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: [a-zA-Z0-9_.-]+

Required: No

MinValue

A minimum allowed value for the parameter.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: [a-zA-Z0-9_.-]+

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ConfusionMatrix

The confusion matrix shows you what your transform is predicting accurately and what types of errors it is making.

For more information, see [Confusion matrix](#) in Wikipedia.

Contents

NumFalseNegatives

The number of matches in the data that the transform didn't find, in the confusion matrix for your transform.

Type: Long

Required: No

NumFalsePositives

The number of nonmatches in the data that the transform incorrectly classified as a match, in the confusion matrix for your transform.

Type: Long

Required: No

NumTrueNegatives

The number of nonmatches in the data that the transform correctly rejected, in the confusion matrix for your transform.

Type: Long

Required: No

NumTruePositives

The number of matches in the data that the transform correctly found, in the confusion matrix for your transform.

Type: Long

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Connection

Defines a connection to a data source.

Contents

AuthenticationConfiguration

The authentication properties of the connection.

Type: [AuthenticationConfiguration](#) object

Required: No

ConnectionProperties

These key-value pairs define parameters for the connection:

- **HOST** - The host URI: either the fully qualified domain name (FQDN) or the IPv4 address of the database host.
- **PORT** - The port number, between 1024 and 65535, of the port on which the database host is listening for database connections.
- **USER_NAME** - The name under which to log in to the database. The value string for **USER_NAME** is "USERNAME".
- **PASSWORD** - A password, if one is used, for the user name.
- **ENCRYPTED_PASSWORD** - When you enable connection password protection by setting `ConnectionPasswordEncryption` in the Data Catalog encryption settings, this field stores the encrypted password.
- **JDBC_DRIVER_JAR_URI** - The Amazon Simple Storage Service (Amazon S3) path of the JAR file that contains the JDBC driver to use.
- **JDBC_DRIVER_CLASS_NAME** - The class name of the JDBC driver to use.
- **JDBC_ENGINE** - The name of the JDBC engine to use.
- **JDBC_ENGINE_VERSION** - The version of the JDBC engine to use.
- **CONFIG_FILES** - (Reserved for future use.)
- **INSTANCE_ID** - The instance ID to use.
- **JDBC_CONNECTION_URL** - The URL for connecting to a JDBC data source.

- `JDBC_ENFORCE_SSL` - A Boolean string (true, false) specifying whether Secure Sockets Layer (SSL) with hostname matching is enforced for the JDBC connection on the client. The default is false.
- `CUSTOM_JDBC_CERT` - An Amazon S3 location specifying the customer's root certificate. AWS Glue uses this root certificate to validate the customer's certificate when connecting to the customer database. AWS Glue only handles X.509 certificates. The certificate provided must be DER-encoded and supplied in Base64 encoding PEM format.
- `SKIP_CUSTOM_JDBC_CERT_VALIDATION` - By default, this is false. AWS Glue validates the Signature algorithm and Subject Public Key Algorithm for the customer certificate. The only permitted algorithms for the Signature algorithm are SHA256withRSA, SHA384withRSA or SHA512withRSA. For the Subject Public Key Algorithm, the key length must be at least 2048. You can set the value of this property to true to skip AWS Glue's validation of the customer certificate.
- `CUSTOM_JDBC_CERT_STRING` - A custom JDBC certificate string which is used for domain match or distinguished name match to prevent a man-in-the-middle attack. In Oracle database, this is used as the `SSL_SERVER_CERT_DN`; in Microsoft SQL Server, this is used as the `hostNameInCertificate`.
- `CONNECTION_URL` - The URL for connecting to a general (non-JDBC) data source.
- `SECRET_ID` - The secret ID used for the secret manager of credentials.
- `CONNECTOR_URL` - The connector URL for a MARKETPLACE or CUSTOM connection.
- `CONNECTOR_TYPE` - The connector type for a MARKETPLACE or CUSTOM connection.
- `CONNECTOR_CLASS_NAME` - The connector class name for a MARKETPLACE or CUSTOM connection.
- `KAFKA_BOOTSTRAP_SERVERS` - A comma-separated list of host and port pairs that are the addresses of the Apache Kafka brokers in a Kafka cluster to which a Kafka client will connect to and bootstrap itself.
- `KAFKA_SSL_ENABLED` - Whether to enable or disable SSL on an Apache Kafka connection. Default value is "true".
- `KAFKA_CUSTOM_CERT` - The Amazon S3 URL for the private CA cert file (.pem format). The default is an empty string.
- `KAFKA_SKIP_CUSTOM_CERT_VALIDATION` - Whether to skip the validation of the CA cert file or not. AWS Glue validates for three algorithms: SHA256withRSA, SHA384withRSA and SHA512withRSA. Default value is "false".

- `KAFKA_CLIENT_KEYSTORE` - The Amazon S3 location of the client keystore file for Kafka client side authentication (Optional).
- `KAFKA_CLIENT_KEYSTORE_PASSWORD` - The password to access the provided keystore (Optional).
- `KAFKA_CLIENT_KEY_PASSWORD` - A keystore can consist of multiple keys, so this is the password to access the client key to be used with the Kafka server side key (Optional).
- `ENCRYPTED_KAFKA_CLIENT_KEYSTORE_PASSWORD` - The encrypted version of the Kafka client keystore password (if the user has the AWS Glue encrypt passwords setting selected).
- `ENCRYPTED_KAFKA_CLIENT_KEY_PASSWORD` - The encrypted version of the Kafka client key password (if the user has the AWS Glue encrypt passwords setting selected).
- `KAFKA_SASL_MECHANISM` - "SCRAM-SHA-512", "GSSAPI", "AWS_MSK_IAM", or "PLAIN". These are the supported [SASL Mechanisms](#).
- `KAFKA_SASL_PLAIN_USERNAME` - A plaintext username used to authenticate with the "PLAIN" mechanism.
- `KAFKA_SASL_PLAIN_PASSWORD` - A plaintext password used to authenticate with the "PLAIN" mechanism.
- `ENCRYPTED_KAFKA_SASL_PLAIN_PASSWORD` - The encrypted version of the Kafka SASL PLAIN password (if the user has the AWS Glue encrypt passwords setting selected).
- `KAFKA_SASL_SCRAM_USERNAME` - A plaintext username used to authenticate with the "SCRAM-SHA-512" mechanism.
- `KAFKA_SASL_SCRAM_PASSWORD` - A plaintext password used to authenticate with the "SCRAM-SHA-512" mechanism.
- `ENCRYPTED_KAFKA_SASL_SCRAM_PASSWORD` - The encrypted version of the Kafka SASL SCRAM password (if the user has the AWS Glue encrypt passwords setting selected).
- `KAFKA_SASL_SCRAM_SECRETS_ARN` - The Amazon Resource Name of a secret in AWS Secrets Manager.
- `KAFKA_SASL_GSSAPI_KEYTAB` - The S3 location of a Kerberos keytab file. A keytab stores long-term keys for one or more principals. For more information, see [MIT Kerberos Documentation: Keytab](#).
- `KAFKA_SASL_GSSAPI_KRB5_CONF` - The S3 location of a Kerberos `krb5.conf` file. A `krb5.conf` stores Kerberos configuration information, such as the location of the KDC server. For more information, see [MIT Kerberos Documentation: krb5.conf](#).

- `KAFKA_SASL_GSSAPI_SERVICE` - The Kerberos service name, as set with `sasl.kerberos.service.name` in your [Kafka Configuration](#).
- `KAFKA_SASL_GSSAPI_PRINCIPAL` - The name of the Kerberos principal used by AWS Glue. For more information, see [Kafka Documentation: Configuring Kafka Brokers](#).

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 100 items.

Valid Keys: `HOST` | `PORT` | `USERNAME` | `PASSWORD` | `ENCRYPTED_PASSWORD` | `JDBC_DRIVER_JAR_URI` | `JDBC_DRIVER_CLASS_NAME` | `JDBC_ENGINE` | `JDBC_ENGINE_VERSION` | `CONFIG_FILES` | `INSTANCE_ID` | `JDBC_CONNECTION_URL` | `JDBC_ENFORCE_SSL` | `CUSTOM_JDBC_CERT` | `SKIP_CUSTOM_JDBC_CERT_VALIDATION` | `CUSTOM_JDBC_CERT_STRING` | `CONNECTION_URL` | `KAFKA_BOOTSTRAP_SERVERS` | `KAFKA_SSL_ENABLED` | `KAFKA_CUSTOM_CERT` | `KAFKA_SKIP_CUSTOM_CERT_VALIDATION` | `KAFKA_CLIENT_KEYSTORE` | `KAFKA_CLIENT_KEYSTORE_PASSWORD` | `KAFKA_CLIENT_KEY_PASSWORD` | `ENCRYPTED_KAFKA_CLIENT_KEYSTORE_PASSWORD` | `ENCRYPTED_KAFKA_CLIENT_KEY_PASSWORD` | `SECRET_ID` | `CONNECTOR_URL` | `CONNECTOR_TYPE` | `CONNECTOR_CLASS_NAME` | `KAFKA_SASL_MECHANISM` | `KAFKA_SASL_PLAIN_USERNAME` | `KAFKA_SASL_PLAIN_PASSWORD` | `ENCRYPTED_KAFKA_SASL_PLAIN_PASSWORD` | `KAFKA_SASL_SCRAM_USERNAME` | `KAFKA_SASL_SCRAM_PASSWORD` | `KAFKA_SASL_SCRAM_SECRETS_ARN` | `ENCRYPTED_KAFKA_SASL_SCRAM_PASSWORD` | `KAFKA_SASL_GSSAPI_KEYTAB` | `KAFKA_SASL_GSSAPI_KRB5_CONF` | `KAFKA_SASL_GSSAPI_SERVICE` | `KAFKA_SASL_GSSAPI_PRINCIPAL` | `ROLE_ARN`

Value Length Constraints: Maximum length of 1024.

Required: No

ConnectionType

The type of the connection. Currently, SFTP is not supported.

Type: String

Valid Values: `JDBC` | `SFTP` | `MONGODB` | `KAFKA` | `NETWORK` | `MARKETPLACE` | `CUSTOM` | `SALESFORCE`

Required: No

CreationTime

The timestamp of the time that this connection definition was created.

Type: Timestamp

Required: No

Description

The description of the connection.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

LastConnectionValidationTime

A timestamp of the time this connection was last validated.

Type: Timestamp

Required: No

LastUpdatedBy

The user, group, or role that last updated this connection definition.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

LastUpdatedTime

The timestamp of the last time the connection definition was updated.

Type: Timestamp

Required: No

MatchCriteria

A list of criteria that can be used in selecting this connection.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 10 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Name

The name of the connection definition.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

PhysicalConnectionRequirements

The physical connection requirements, such as virtual private cloud (VPC) and SecurityGroup, that are needed to make this connection successfully.

Type: [PhysicalConnectionRequirements](#) object

Required: No

Status

The status of the connection. Can be one of: READY, IN_PROGRESS, or FAILED.

Type: String

Valid Values: READY | IN_PROGRESS | FAILED

Required: No

StatusReason

The reason for the connection status.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 16384.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ConnectionInput

A structure that is used to specify a connection to create or update.

Contents

ConnectionProperties

These key-value pairs define parameters for the connection.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 100 items.

Valid Keys: HOST | PORT | USERNAME | PASSWORD | ENCRYPTED_PASSWORD
| JDBC_DRIVER_JAR_URI | JDBC_DRIVER_CLASS_NAME | JDBC_ENGINE
| JDBC_ENGINE_VERSION | CONFIG_FILES | INSTANCE_ID |
JDBC_CONNECTION_URL | JDBC_ENFORCE_SSL | CUSTOM_JDBC_CERT |
SKIP_CUSTOM_JDBC_CERT_VALIDATION | CUSTOM_JDBC_CERT_STRING |
CONNECTION_URL | KAFKA_BOOTSTRAP_SERVERS | KAFKA_SSL_ENABLED
| KAFKA_CUSTOM_CERT | KAFKA_SKIP_CUSTOM_CERT_VALIDATION |
KAFKA_CLIENT_KEYSTORE | KAFKA_CLIENT_KEYSTORE_PASSWORD |
KAFKA_CLIENT_KEY_PASSWORD | ENCRYPTED_KAFKA_CLIENT_KEYSTORE_PASSWORD
| ENCRYPTED_KAFKA_CLIENT_KEY_PASSWORD | SECRET_ID | CONNECTOR_URL
| CONNECTOR_TYPE | CONNECTOR_CLASS_NAME | KAFKA_SASL_MECHANISM
| KAFKA_SASL_PLAIN_USERNAME | KAFKA_SASL_PLAIN_PASSWORD |
ENCRYPTED_KAFKA_SASL_PLAIN_PASSWORD | KAFKA_SASL_SCRAM_USERNAME
| KAFKA_SASL_SCRAM_PASSWORD | KAFKA_SASL_SCRAM_SECRETS_ARN |
ENCRYPTED_KAFKA_SASL_SCRAM_PASSWORD | KAFKA_SASL_GSSAPI_KEYTAB
| KAFKA_SASL_GSSAPI_KRB5_CONF | KAFKA_SASL_GSSAPI_SERVICE |
KAFKA_SASL_GSSAPI_PRINCIPAL | ROLE_ARN

Value Length Constraints: Maximum length of 1024.

Required: Yes

ConnectionType

The type of the connection. Currently, these types are supported:

- **JDBC** - Designates a connection to a database through Java Database Connectivity (JDBC).

JDBC Connections use the following ConnectionParameters.

- Required: All of (HOST, PORT, JDBC_ENGINE) or JDBC_CONNECTION_URL.
 - Required: All of (USERNAME, PASSWORD) or SECRET_ID.
 - Optional: JDBC_ENFORCE_SSL, CUSTOM_JDBC_CERT, CUSTOM_JDBC_CERT_STRING, SKIP_CUSTOM_JDBC_CERT_VALIDATION. These parameters are used to configure SSL with JDBC.
- **KAFKA** - Designates a connection to an Apache Kafka streaming platform.

KAFKA Connections use the following ConnectionParameters.

- Required: KAFKA_BOOTSTRAP_SERVERS.
 - Optional: KAFKA_SSL_ENABLED, KAFKA_CUSTOM_CERT, KAFKA_SKIP_CUSTOM_CERT_VALIDATION. These parameters are used to configure SSL with KAFKA.
 - Optional: KAFKA_CLIENT_KEYSTORE, KAFKA_CLIENT_KEYSTORE_PASSWORD, KAFKA_CLIENT_KEY_PASSWORD, ENCRYPTED_KAFKA_CLIENT_KEYSTORE_PASSWORD, ENCRYPTED_KAFKA_CLIENT_KEY_PASSWORD. These parameters are used to configure TLS client configuration with SSL in KAFKA.
 - Optional: KAFKA_SASL_MECHANISM. Can be specified as SCRAM-SHA-512, GSSAPI, or AWS_MSK_IAM.
 - Optional: KAFKA_SASL_SCRAM_USERNAME, KAFKA_SASL_SCRAM_PASSWORD, ENCRYPTED_KAFKA_SASL_SCRAM_PASSWORD. These parameters are used to configure SASL/SCRAM-SHA-512 authentication with KAFKA.
 - Optional: KAFKA_SASL_GSSAPI_KEYTAB, KAFKA_SASL_GSSAPI_KRB5_CONF, KAFKA_SASL_GSSAPI_SERVICE, KAFKA_SASL_GSSAPI_PRINCIPAL. These parameters are used to configure SASL/GSSAPI authentication with KAFKA.
- **MONGODB** - Designates a connection to a MongoDB document database.

MONGODB Connections use the following ConnectionParameters.

- Required: CONNECTION_URL.
 - Required: All of (USERNAME, PASSWORD) or SECRET_ID.
- **SALESFORCE** - Designates a connection to Salesforce using OAuth authentication.

- NETWORK - Designates a network connection to a data source within an Amazon Virtual Private Cloud environment (Amazon VPC).

NETWORK Connections do not require ConnectionParameters. Instead, provide a PhysicalConnectionRequirements.

- MARKETPLACE - Uses configuration settings contained in a connector purchased from AWS Marketplace to read from and write to data stores that are not natively supported by AWS Glue.

MARKETPLACE Connections use the following ConnectionParameters.

- Required: CONNECTOR_TYPE, CONNECTOR_URL, CONNECTOR_CLASS_NAME, CONNECTION_URL.
- Required for JDBC CONNECTOR_TYPE connections: All of (USERNAME, PASSWORD) or SECRET_ID.
- CUSTOM - Uses configuration settings contained in a custom connector to read from and write to data stores that are not natively supported by AWS Glue.

SFTP is not supported.

For more information about how optional ConnectionProperties are used to configure features in AWS Glue, consult [AWS Glue connection properties](#).

For more information about how optional ConnectionProperties are used to configure features in AWS Glue Studio, consult [Using connectors and connections](#).

Type: String

Valid Values: JDBC | SFTP | MONGODB | KAFKA | NETWORK | MARKETPLACE | CUSTOM | SALESFORCE

Required: Yes

Name

The name of the connection.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

AuthenticationConfiguration

The authentication properties of the connection. Used for a Salesforce connection.

Type: [AuthenticationConfigurationInput](#) object

Required: No

Description

The description of the connection.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

MatchCriteria

A list of criteria that can be used in selecting this connection.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 10 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

PhysicalConnectionRequirements

The physical connection requirements, such as virtual private cloud (VPC) and SecurityGroup, that are needed to successfully make this connection.

Type: [PhysicalConnectionRequirements](#) object

Required: No

ValidateCredentials

A flag to validate the credentials during create connection. Used for a Salesforce connection. Default is true.

Type: Boolean

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ConnectionPasswordEncryption

The data structure used by the Data Catalog to encrypt the password as part of `CreateConnection` or `UpdateConnection` and store it in the `ENCRYPTED_PASSWORD` field in the connection properties. You can enable catalog encryption or only password encryption.

When a `CreationConnection` request arrives containing a password, the Data Catalog first encrypts the password using your AWS KMS key. It then encrypts the whole connection object again if catalog encryption is also enabled.

This encryption requires that you set AWS KMS key permissions to enable or restrict access on the password key according to your security requirements. For example, you might want only administrators to have decrypt permission on the password key.

Contents

ReturnConnectionPasswordEncrypted

When the `ReturnConnectionPasswordEncrypted` flag is set to "true", passwords remain encrypted in the responses of `GetConnection` and `GetConnections`. This encryption takes effect independently from catalog encryption.

Type: Boolean

Required: Yes

AwsKmsKeyId

An AWS KMS key that is used to encrypt the connection password.

If connection password protection is enabled, the caller of `CreateConnection` and `UpdateConnection` needs at least `kms:Encrypt` permission on the specified AWS KMS key, to encrypt passwords before storing them in the Data Catalog.

You can set the decrypt permission to enable or restrict access on the password key according to your security requirements.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ConnectionsList

Specifies the connections used by a job.

Contents

Connections

A list of connections used by the job.

Type: Array of strings

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ConnectorDataSource

Specifies a source generated with standard connection options.

Contents

ConnectionType

The `connectionType`, as provided to the underlying AWS Glue library. This node type supports the following connection types:

- `opensearch`
- `azuresql`
- `azurecosmos`
- `bigquery`
- `saphana`
- `teradata`
- `vertica`

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF] | [^\\S\\r\\n"']`)*

Required: Yes

Data

A map specifying connection options for the node. You can find standard connection options for the corresponding connection type in the [Connection parameters](#) section of the AWS Glue documentation.

Type: String to string map

Required: Yes

Name

The name of this source node.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]| [^\\r\\n])*`

Required: Yes

OutputSchemas

Specifies the data schema for this source.

Type: Array of [GlueSchema](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ConnectorDataTarget

Specifies a target generated with standard connection options.

Contents

ConnectionType

The `connectionType`, as provided to the underlying AWS Glue library. This node type supports the following connection types:

- `opensearch`
- `azuresql`
- `azurecosmos`
- `bigquery`
- `saphana`
- `teradata`
- `vertica`

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF] | [^\\S\\r\\n"']`)*

Required: Yes

Data

A map specifying connection options for the node. You can find standard connection options for the corresponding connection type in the [Connection parameters](#) section of the AWS Glue documentation.

Type: String to string map

Required: Yes

Name

The name of this target node.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]| [^\\r\\n])*`

Required: Yes

Inputs

The nodes that are inputs to the data target.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Crawl

The details of a crawl in the workflow.

Contents

CompletedOn

The date and time on which the crawl completed.

Type: Timestamp

Required: No

ErrorMessage

The error message associated with the crawl.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\u007F\u00E0-\u00FF\u0080-\u00FF\u00DC-\u00FF\u00r\u00n\u00t]*`

Required: No

LogGroup

The log group associated with the crawl.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 512.

Pattern: `[\.\-_\/#A-Za-z0-9]+`

Required: No

LogStream

The log stream associated with the crawl.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 512.

Pattern: [^:*]*

Required: No

StartedOn

The date and time on which the crawl started.

Type: Timestamp

Required: No

State

The state of the crawler.

Type: String

Valid Values: RUNNING | CANCELLING | CANCELLED | SUCCEEDED | FAILED | ERROR

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Crawler

Specifies a crawler program that examines a data source and uses classifiers to try to determine its schema. If successful, the crawler records metadata concerning the data source in the AWS Glue Data Catalog.

Contents

Classifiers

A list of UTF-8 strings that specify the custom classifiers that are associated with the crawler.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Configuration

Crawler configuration information. This versioned JSON string allows users to specify aspects of a crawler's behavior. For more information, see [Setting crawler configuration options](#).

Type: String

Required: No

CrawlElapsedTime

If the crawler is running, contains the total time elapsed since the last crawl began.

Type: Long

Required: No

CrawlerSecurityConfiguration

The name of the SecurityConfiguration structure to be used by this crawler.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 128.

Required: No

CreationTime

The time that the crawler was created.

Type: Timestamp

Required: No

DatabaseName

The name of the database in which the crawler's output is stored.

Type: String

Required: No

Description

A description of the crawler.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

LakeFormationConfiguration

Specifies whether the crawler should use AWS Lake Formation credentials for the crawler instead of the IAM role credentials.

Type: [LakeFormationConfiguration](#) object

Required: No

LastCrawl

The status of the last crawl, and potentially error information if an error occurred.

Type: [LastCrawlInfo](#) object

Required: No

LastUpdated

The time that the crawler was last updated.

Type: Timestamp

Required: No

LineageConfiguration

A configuration that specifies whether data lineage is enabled for the crawler.

Type: [LineageConfiguration](#) object

Required: No

Name

The name of the crawler.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

RecrawlPolicy

A policy that specifies whether to crawl the entire dataset again, or to crawl only folders that were added since the last crawler run.

Type: [RecrawlPolicy](#) object

Required: No

Role

The Amazon Resource Name (ARN) of an IAM role that's used to access customer resources, such as Amazon Simple Storage Service (Amazon S3) data.

Type: String

Required: No

Schedule

For scheduled crawlers, the schedule when the crawler runs.

Type: [Schedule](#) object

Required: No

SchemaChangePolicy

The policy that specifies update and delete behaviors for the crawler.

Type: [SchemaChangePolicy](#) object

Required: No

State

Indicates whether the crawler is running, or whether a run is pending.

Type: String

Valid Values: READY | RUNNING | STOPPING

Required: No

TablePrefix

The prefix added to the names of tables that are created.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 128.

Required: No

Targets

A collection of targets to crawl.

Type: [CrawlerTargets](#) object

Required: No

Version

The version of the crawler.

Type: Long

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CrawlerHistory

Contains the information for a run of a crawler.

Contents

CrawlId

A UUID identifier for each crawl.

Type: String

Required: No

DPUHour

The number of data processing units (DPU) used in hours for the crawl.

Type: Double

Valid Range: Minimum value of 0.0.

Required: No

EndTime

The date and time on which the crawl ended.

Type: Timestamp

Required: No

ErrorMessage

If an error occurred, the error message associated with the crawl.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

LogGroup

The log group associated with the crawl.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 512.

Pattern: `[\.\-_/#A-Za-z0-9]+`

Required: No

LogStream

The log stream associated with the crawl.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 512.

Pattern: `[^:]*`

Required: No

MessagePrefix

The prefix for a CloudWatch message about this crawl.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

StartTime

The date and time on which the crawl started.

Type: Timestamp

Required: No

State

The state of the crawl.

Type: String

Valid Values: RUNNING | COMPLETED | FAILED | STOPPED

Required: No

Summary

A run summary for the specific crawl in JSON. Contains the catalog tables and partitions that were added, updated, or deleted.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CrawlerMetrics

Metrics for a specified crawler.

Contents

CrawlerName

The name of the crawler.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

LastRuntimeSeconds

The duration of the crawler's most recent run, in seconds.

Type: Double

Valid Range: Minimum value of 0.0.

Required: No

MedianRuntimeSeconds

The median duration of this crawler's runs, in seconds.

Type: Double

Valid Range: Minimum value of 0.0.

Required: No

StillEstimating

True if the crawler is still estimating how long it will take to complete this run.

Type: Boolean

Required: No

TablesCreated

The number of tables created by this crawler.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

TablesDeleted

The number of tables deleted by this crawler.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

TablesUpdated

The number of tables updated by this crawler.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

TimeLeftSeconds

The estimated time left to complete a running crawl.

Type: Double

Valid Range: Minimum value of 0.0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CrawlerNodeDetails

The details of a Crawler node present in the workflow.

Contents

Crawls

A list of crawls represented by the crawl node.

Type: Array of [Crawl](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CrawlerTargets

Specifies data stores to crawl.

Contents

CatalogTargets

Specifies AWS Glue Data Catalog targets.

Type: Array of [CatalogTarget](#) objects

Required: No

DeltaTargets

Specifies Delta data store targets.

Type: Array of [DeltaTarget](#) objects

Required: No

DynamoDBTargets

Specifies Amazon DynamoDB targets.

Type: Array of [DynamoDBTarget](#) objects

Required: No

HudiTargets

Specifies Apache Hudi data store targets.

Type: Array of [HudiTarget](#) objects

Required: No

IcebergTargets

Specifies Apache Iceberg data store targets.

Type: Array of [IcebergTarget](#) objects

Required: No

JdbcTargets

Specifies JDBC targets.

Type: Array of [JdbcTarget](#) objects

Required: No

MongoDBTargets

Specifies Amazon DocumentDB or MongoDB targets.

Type: Array of [MongoDBTarget](#) objects

Required: No

S3Targets

Specifies Amazon Simple Storage Service (Amazon S3) targets.

Type: Array of [S3Target](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CrawlsFilter

A list of fields, comparators and value that you can use to filter the crawler runs for a specified crawler.

Contents

FieldName

A key used to filter the crawler runs for a specified crawler. Valid values for each of the field names are:

- `CRAWL_ID`: A string representing the UUID identifier for a crawl.
- `STATE`: A string representing the state of the crawl.
- `START_TIME` and `END_TIME`: The epoch timestamp in milliseconds.
- `DPU_HOUR`: The number of data processing unit (DPU) hours used for the crawl.

Type: String

Valid Values: `CRAWL_ID` | `STATE` | `START_TIME` | `END_TIME` | `DPU_HOUR`

Required: No

FieldValue

The value provided for comparison on the crawl field.

Type: String

Required: No

FilterOperator

A defined comparator that operates on the value. The available operators are:

- `GT`: Greater than.
- `GE`: Greater than or equal to.
- `LT`: Less than.
- `LE`: Less than or equal to.
- `EQ`: Equal to.
- `NE`: Not equal to.

Type: String

Valid Values: GT | GE | LT | LE | EQ | NE

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CreateCsvClassifierRequest

Specifies a custom CSV classifier for `CreateClassifier` to create.

Contents

Name

The name of the classifier.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

AllowSingleColumn

Enables the processing of files that contain only one column.

Type: Boolean

Required: No

ContainsHeader

Indicates whether the CSV file contains a header.

Type: String

Valid Values: UNKNOWN | PRESENT | ABSENT

Required: No

CustomDatatypeConfigured

Enables the configuration of custom datatypes.

Type: Boolean

Required: No

CustomDatatypes

Creates a list of supported custom datatypes.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Delimiter

A custom symbol to denote what separates each column entry in the row.

Type: String

Length Constraints: Fixed length of 1.

Pattern: `[\^\r\n]`

Required: No

DisableValueTrimming

Specifies not to trim values before identifying the type of column values. The default value is true.

Type: Boolean

Required: No

Header

A list of strings representing column names.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

QuoteSymbol

A custom symbol to denote what combines content into a single column value. Must be different from the column delimiter.

Type: String

Length Constraints: Fixed length of 1.

Pattern: [^\r\n]

Required: No

Serde

Sets the SerDe for processing CSV in the classifier, which will be applied in the Data Catalog. Valid values are `OpenCSVSerde`, `LazySimpleSerDe`, and `None`. You can specify the `None` value when you want the crawler to do the detection.

Type: String

Valid Values: `OpenCSVSerde` | `LazySimpleSerDe` | `None`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CreateGrokClassifierRequest

Specifies a grok classifier for `CreateClassifier` to create.

Contents

Classification

An identifier of the data format that the classifier matches, such as Twitter, JSON, Omniture logs, Amazon CloudWatch Logs, and so on.

Type: String

Required: Yes

GrokPattern

The grok pattern used by this classifier.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\t]*`

Required: Yes

Name

The name of the new classifier.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

CustomPatterns

Optional custom grok patterns used by this classifier.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 16000.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CreateJsonClassifierRequest

Specifies a JSON classifier for `CreateClassifier` to create.

Contents

JsonPath

A `JsonPath` string defining the JSON data for the classifier to classify. AWS Glue supports a subset of `JsonPath`, as described in [Writing JsonPath Custom Classifiers](#).

Type: String

Required: Yes

Name

The name of the classifier.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CreateXMLClassifierRequest

Specifies an XML classifier for `CreateClassifier` to create.

Contents

Classification

An identifier of the data format that the classifier matches.

Type: String

Required: Yes

Name

The name of the classifier.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

RowTag

The XML tag designating the element that contains each record in an XML document being parsed. This can't identify a self-closing element (closed by `/>`). An empty row element that contains only attributes can be parsed as long as it ends with a closing tag (for example, `<row item_a="A" item_b="B"></row>` is okay, but `<row item_a="A" item_b="B" />` is not).

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)

- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CsvClassifier

A classifier for custom CSV content.

Contents

Name

The name of the classifier.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

AllowSingleColumn

Enables the processing of files that contain only one column.

Type: Boolean

Required: No

ContainsHeader

Indicates whether the CSV file contains a header.

Type: String

Valid Values: UNKNOWN | PRESENT | ABSENT

Required: No

CreationTime

The time that this classifier was registered.

Type: Timestamp

Required: No

CustomDatatypeConfigured

Enables the custom datatype to be configured.

Type: Boolean

Required: No

CustomDatatypes

A list of custom datatypes including "BINARY", "BOOLEAN", "DATE", "DECIMAL", "DOUBLE", "FLOAT", "INT", "LONG", "SHORT", "STRING", "TIMESTAMP".

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Delimiter

A custom symbol to denote what separates each column entry in the row.

Type: String

Length Constraints: Fixed length of 1.

Pattern: `[^\r\n]`

Required: No

DisableValueTrimming

Specifies not to trim values before identifying the type of column values. The default value is `true`.

Type: Boolean

Required: No

Header

A list of strings representing column names.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

LastUpdated

The time that this classifier was last updated.

Type: Timestamp

Required: No

QuoteSymbol

A custom symbol to denote what combines content into a single column value. It must be different from the column delimiter.

Type: String

Length Constraints: Fixed length of 1.

Pattern: `[^\r\n]`

Required: No

Serde

Sets the SerDe for processing CSV in the classifier, which will be applied in the Data Catalog. Valid values are `OpenCSVSerde`, `LazySimpleSerDe`, and `None`. You can specify the `None` value when you want the crawler to do the detection.

Type: String

Valid Values: `OpenCSVSerde` | `LazySimpleSerDe` | `None`

Required: No

Version

The version of this classifier.

Type: Long

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CustomCode

Specifies a transform that uses custom code you provide to perform the data transformation. The output is a collection of DynamicFrames.

Contents

ClassName

The name defined for the custom code node class.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

Code

The custom code that is used to perform the data transformation.

Type: String

Pattern: `[\\s\\S]*`

Required: Yes

Inputs

The data inputs identified by their node names.

Type: Array of strings

Array Members: Minimum number of 1 item.

Pattern: `[A-Za-z0-9_ -]*`

Required: Yes

Name

The name of the transform node.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]| [^\r\n])*`

Required: Yes

OutputSchemas

Specifies the data schema for the custom code transform.

Type: Array of [GlueSchema](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CustomEntityType

An object representing a custom pattern for detecting sensitive data across the columns and rows of your structured data.

Contents

Name

A name for the custom pattern that allows it to be retrieved or deleted later. This name must be unique per AWS account.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

RegexString

A regular expression string that is used for detecting sensitive data in a custom pattern.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

ContextWords

A list of context words. If none of these context words are found within the vicinity of the regular expression the data will not be detected as sensitive data.

If no context words are passed only a regular expression is checked.

Type: Array of strings

Array Members: Minimum number of 1 item. Maximum number of 20 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Database

The Database object represents a logical grouping of tables that might reside in a Hive metastore or an RDBMS.

Contents

Name

The name of the database. For Hive compatibility, this is folded to lowercase when it is stored.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

CatalogId

The ID of the Data Catalog in which the database resides.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

CreateTableDefaultPermissions

Creates a set of default permissions on the table for principals. Used by AWS Lake Formation. Not used in the normal course of AWS Glue operations.

Type: Array of [PrincipalPermissions](#) objects

Required: No

CreateTime

The time at which the metadata database was created in the catalog.

Type: Timestamp

Required: No

Description

A description of the database.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

FederatedDatabase

A FederatedDatabase structure that references an entity outside the AWS Glue Data Catalog.

Type: [FederatedDatabase](#) object

Required: No

LocationUri

The location of the database (for example, an HDFS path).

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Parameters

These key-value pairs define parameters and properties of the database.

Type: String to string map

Key Length Constraints: Minimum length of 1. Maximum length of 255.

Key Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Value Length Constraints: Maximum length of 512000.

Required: No

TargetDatabase

A `DatabaseIdentifier` structure that describes a target database for resource linking.

Type: [DatabaseIdentifier](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Databaseldentifier

A structure that describes a target database for resource linking.

Contents

CatalogId

The ID of the Data Catalog in which the database resides.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

DatabaseName

The name of the catalog database.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Region

Region of the target database.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DatabaseInput

The structure used to create or update a database.

Contents

Name

The name of the database. For Hive compatibility, this is folded to lowercase when it is stored.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

CreateTableDefaultPermissions

Creates a set of default permissions on the table for principals. Used by AWS Lake Formation. Not used in the normal course of AWS Glue operations.

Type: Array of [PrincipalPermissions](#) objects

Required: No

Description

A description of the database.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

FederatedDatabase

A FederatedDatabase structure that references an entity outside the AWS Glue Data Catalog.

Type: [FederatedDatabase](#) object

Required: No

LocationUri

The location of the database (for example, an HDFS path).

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Parameters

These key-value pairs define parameters and properties of the database.

These key-value pairs define parameters and properties of the database.

Type: String to string map

Key Length Constraints: Minimum length of 1. Maximum length of 255.

Key Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Value Length Constraints: Maximum length of 512000.

Required: No

TargetDatabase

A `DatabaseIdentifier` structure that describes a target database for resource linking.

Type: [DatabaseIdentifier](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)

- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DataCatalogEncryptionSettings

Contains configuration information for maintaining Data Catalog security.

Contents

ConnectionPasswordEncryption

When connection password protection is enabled, the Data Catalog uses a customer-provided key to encrypt the password as part of `CreateConnection` or `UpdateConnection` and store it in the `ENCRYPTED_PASSWORD` field in the connection properties. You can enable catalog encryption or only password encryption.

Type: [ConnectionPasswordEncryption](#) object

Required: No

EncryptionAtRest

Specifies the encryption-at-rest configuration for the Data Catalog.

Type: [EncryptionAtRest](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DataLakePrincipal

The AWS Lake Formation principal.

Contents

DataLakePrincipalIdentifier

An identifier for the AWS Lake Formation principal.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DataQualityAnalyzerResult

Describes the result of the evaluation of a data quality analyzer.

Contents

Description

A description of the data quality analyzer.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

EvaluatedMetrics

A map of metrics associated with the evaluation of the analyzer.

Type: String to double map

Key Length Constraints: Minimum length of 1. Maximum length of 255.

Key Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

EvaluationMessage

An evaluation message.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Name

The name of the data quality analyzer.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DataQualityEvaluationRunAdditionalRunOptions

Additional run options you can specify for an evaluation run.

Contents

CloudWatchMetricsEnabled

Whether or not to enable CloudWatch metrics.

Type: Boolean

Required: No

CompositeRuleEvaluationMethod

Set the evaluation method for composite rules in the ruleset to ROW/COLUMN

Type: String

Valid Values: COLUMN | ROW

Required: No

ResultsS3Prefix

Prefix for Amazon S3 to store results.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DataQualityMetricValues

Describes the data quality metric value according to the analysis of historical data.

Contents

ActualValue

The actual value of the data quality metric.

Type: Double

Required: No

ExpectedValue

The expected value of the data quality metric according to the analysis of historical data.

Type: Double

Required: No

LowerLimit

The lower limit of the data quality metric value according to the analysis of historical data.

Type: Double

Required: No

UpperLimit

The upper limit of the data quality metric value according to the analysis of historical data.

Type: Double

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)

- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DataQualityObservation

Describes the observation generated after evaluating the rules and analyzers.

Contents

Description

A description of the data quality observation.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

MetricBasedObservation

An object of type `MetricBasedObservation` representing the observation that is based on evaluated data quality metrics.

Type: [MetricBasedObservation](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DataQualityResult

Describes a data quality result.

Contents

AnalyzerResults

A list of `DataQualityAnalyzerResult` objects representing the results for each analyzer.

Type: Array of [DataQualityAnalyzerResult](#) objects

Array Members: Minimum number of 0 items. Maximum number of 2000 items.

Required: No

CompletedOn

The date and time when this data quality run completed.

Type: Timestamp

Required: No

DataSource

The table associated with the data quality result, if any.

Type: [DataSource](#) object

Required: No

EvaluationContext

In the context of a job in AWS Glue Studio, each node in the canvas is typically assigned some sort of name and data quality nodes will have names. In the case of multiple nodes, the `evaluationContext` can differentiate the nodes.

Type: String

Required: No

JobName

The job name associated with the data quality result, if any.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

JobRunId

The job run ID associated with the data quality result, if any.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Observations

A list of `DataQualityObservation` objects representing the observations generated after evaluating the rules and analyzers.

Type: Array of [DataQualityObservation](#) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

ResultId

A unique result ID for the data quality result.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

RuleResults

A list of `DataQualityRuleResult` objects representing the results for each rule.

Type: Array of [DataQualityRuleResult](#) objects

Array Members: Minimum number of 0 items. Maximum number of 2000 items.

Required: No

RulesetEvaluationRunId

The unique run ID for the ruleset evaluation for this data quality result.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

RulesetName

The name of the ruleset associated with the data quality result.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Score

An aggregate data quality score. Represents the ratio of rules that passed to the total number of rules.

Type: Double

Valid Range: Minimum value of 0.0. Maximum value of 1.0.

Required: No

StartedOn

The date and time when this data quality run started.

Type: Timestamp

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DataQualityResultDescription

Describes a data quality result.

Contents

DataSource

The table name associated with the data quality result.

Type: [DataSource](#) object

Required: No

JobName

The job name associated with the data quality result.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

JobRunId

The job run ID associated with the data quality result.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

ResultId

The unique result ID for this data quality result.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

StartedOn

The time that the run started for this data quality result.

Type: Timestamp

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DataQualityResultFilterCriteria

Criteria used to return data quality results.

Contents

DataSource

Filter results by the specified data source. For example, retrieving all results for an AWS Glue table.

Type: [DataSource](#) object

Required: No

JobName

Filter results by the specified job name.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

JobRunId

Filter results by the specified job run ID.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

StartedAfter

Filter results by runs that started after this time.

Type: Timestamp

Required: No

StartedBefore

Filter results by runs that started before this time.

Type: Timestamp

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DataQualityRuleRecommendationRunDescription

Describes the result of a data quality rule recommendation run.

Contents

DataSource

The data source (AWS Glue table) associated with the recommendation run.

Type: [DataSource](#) object

Required: No

RunId

The unique run identifier associated with this run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

StartedOn

The date and time when this run started.

Type: Timestamp

Required: No

Status

The status for this run.

Type: String

Valid Values: STARTING | RUNNING | STOPPING | STOPPED | SUCCEEDED | FAILED | TIMEOUT

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DataQualityRuleRecommendationRunFilter

A filter for listing data quality recommendation runs.

Contents

DataSource

Filter based on a specified data source (AWS Glue table).

Type: [DataSource](#) object

Required: Yes

StartedAfter

Filter based on time for results started after provided time.

Type: Timestamp

Required: No

StartedBefore

Filter based on time for results started before provided time.

Type: Timestamp

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DataQualityRuleResult

Describes the result of the evaluation of a data quality rule.

Contents

Description

A description of the data quality rule.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

EvaluatedMetrics

A map of metrics associated with the evaluation of the rule.

Type: String to double map

Key Length Constraints: Minimum length of 1. Maximum length of 255.

Key Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

EvaluationMessage

An evaluation message.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Name

The name of the data quality rule.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Result

A pass or fail status for the rule.

Type: String

Valid Values: PASS | FAIL | ERROR

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DataQualityRulesetEvaluationRunDescription

Describes the result of a data quality ruleset evaluation run.

Contents

DataSource

The data source (an AWS Glue table) associated with the run.

Type: [DataSource](#) object

Required: No

RunId

The unique run identifier associated with this run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

StartedOn

The date and time when the run started.

Type: Timestamp

Required: No

Status

The status for this run.

Type: String

Valid Values: STARTING | RUNNING | STOPPING | STOPPED | SUCCEEDED | FAILED | TIMEOUT

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DataQualityRulesetEvaluationRunFilter

The filter criteria.

Contents

DataSource

Filter based on a data source (an AWS Glue table) associated with the run.

Type: [DataSource](#) object

Required: Yes

StartedAfter

Filter results by runs that started after this time.

Type: Timestamp

Required: No

StartedBefore

Filter results by runs that started before this time.

Type: Timestamp

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DataQualityRulesetFilterCriteria

The criteria used to filter data quality rulesets.

Contents

CreatedAfter

Filter on rulesets created after this date.

Type: Timestamp

Required: No

CreatedBefore

Filter on rulesets created before this date.

Type: Timestamp

Required: No

Description

The description of the ruleset filter criteria.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

LastModifiedAfter

Filter on rulesets last modified after this date.

Type: Timestamp

Required: No

LastModifiedBefore

Filter on rulesets last modified before this date.

Type: Timestamp

Required: No

Name

The name of the ruleset filter criteria.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

TargetTable

The name and database name of the target table.

Type: [DataQualityTargetTable](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DataQualityRulesetListDetails

Describes a data quality ruleset returned by `GetDataQualityRuleset`.

Contents

CreatedOn

The date and time the data quality ruleset was created.

Type: Timestamp

Required: No

Description

A description of the data quality ruleset.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

LastModifiedOn

The date and time the data quality ruleset was last modified.

Type: Timestamp

Required: No

Name

The name of the data quality ruleset.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

RecommendationRunId

When a ruleset was created from a recommendation run, this run ID is generated to link the two together.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

RuleCount

The number of rules in the ruleset.

Type: Integer

Required: No

TargetTable

An object representing an AWS Glue table.

Type: [DataQualityTargetTable](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DataQualityTargetTable

An object representing an AWS Glue table.

Contents

DatabaseName

The name of the database where the AWS Glue table exists.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

TableName

The name of the AWS Glue table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

CatalogId

The catalog id where the AWS Glue table exists.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DataSource

A data source (an AWS Glue table) for which you want data quality results.

Contents

GlueTable

An AWS Glue table.

Type: [GlueTable](#) object

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Datatype

A structure representing the datatype of the value.

Contents

Id

The datatype of the value.

Type: String

Pattern: [A-Za-z0-9_-]*

Required: Yes

Label

A label assigned to the datatype.

Type: String

Pattern: [A-Za-z0-9_-]*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DateColumnStatisticsData

Defines column statistics supported for timestamp data columns.

Contents

NumberOfDistinctValues

The number of distinct values in a column.

Type: Long

Valid Range: Minimum value of 0.

Required: Yes

NumberOfNulls

The number of null values in the column.

Type: Long

Valid Range: Minimum value of 0.

Required: Yes

MaximumValue

The highest value in the column.

Type: Timestamp

Required: No

MinimumValue

The lowest value in the column.

Type: Timestamp

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DecimalColumnStatisticsData

Defines column statistics supported for fixed-point number data columns.

Contents

NumberOfDistinctValues

The number of distinct values in a column.

Type: Long

Valid Range: Minimum value of 0.

Required: Yes

NumberOfNulls

The number of null values in the column.

Type: Long

Valid Range: Minimum value of 0.

Required: Yes

MaximumValue

The highest value in the column.

Type: [DecimalNumber](#) object

Required: No

MinimumValue

The lowest value in the column.

Type: [DecimalNumber](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DecimalNumber

Contains a numeric value in decimal format.

Contents

Scale

The scale that determines where the decimal point falls in the unscaled value.

Type: Integer

Required: Yes

UnscaledValue

The unscaled numeric value.

Type: Base64-encoded binary data object

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DeltaTarget

Specifies a Delta data store to crawl one or more Delta tables.

Contents

ConnectionName

The name of the connection to use to connect to the Delta table target.

Type: String

Required: No

CreateNativeDeltaTable

Specifies whether the crawler will create native tables, to allow integration with query engines that support querying of the Delta transaction log directly.

Type: Boolean

Required: No

DeltaTables

A list of the Amazon S3 paths to the Delta tables.

Type: Array of strings

Required: No

WriteManifest

Specifies whether to write the manifest files to the Delta table path.

Type: Boolean

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DevEndpoint

A development endpoint where a developer can remotely debug extract, transform, and load (ETL) scripts.

Contents

Arguments

A map of arguments used to configure the DevEndpoint.

Valid arguments are:

- `--enable-glue-datacatalog`: ""

You can specify a version of Python support for development endpoints by using the `Arguments` parameter in the `CreateDevEndpoint` or `UpdateDevEndpoint` APIs. If no arguments are provided, the version defaults to Python 2.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 100 items.

Required: No

AvailabilityZone

The AWS Availability Zone where this DevEndpoint is located.

Type: String

Required: No

CreatedTimestamp

The point in time at which this DevEndpoint was created.

Type: Timestamp

Required: No

EndpointName

The name of the DevEndpoint.

Type: String

Required: No

ExtraJarsS3Path

The path to one or more Java `.jar` files in an S3 bucket that should be loaded in your DevEndpoint.

 **Note**


You can only use pure Java/Scala libraries with a DevEndpoint.

Type: String

Required: No

ExtraPythonLibsS3Path

The paths to one or more Python libraries in an Amazon S3 bucket that should be loaded in your DevEndpoint. Multiple values must be complete paths separated by a comma.

 **Note**

You can only use pure Python libraries with a DevEndpoint. Libraries that rely on C extensions, such as the [pandas](#) Python data analysis library, are not currently supported.

Type: String

Required: No

FailureReason

The reason for a current failure in this DevEndpoint.

Type: String

Required: No

GlueVersion

Glue version determines the versions of Apache Spark and Python that AWS Glue supports. The Python version indicates the version supported for running your ETL scripts on development endpoints.

For more information about the available AWS Glue versions and corresponding Spark and Python versions, see [Glue version](#) in the developer guide.

Development endpoints that are created without specifying a Glue version default to Glue 0.9.

You can specify a version of Python support for development endpoints by using the `Arguments` parameter in the `CreateDevEndpoint` or `UpdateDevEndpoint` APIs. If no arguments are provided, the version defaults to Python 2.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `^\w+\.\w+$`

Required: No

LastModifiedTimestamp

The point in time at which this `DevEndpoint` was last modified.

Type: Timestamp

Required: No

LastUpdateStatus

The status of the last update.

Type: String

Required: No

NumberOfNodes

The number of AWS Glue Data Processing Units (DPUs) allocated to this `DevEndpoint`.

Type: Integer

Required: No

NumberOfWorkers

The number of workers of a defined `workerType` that are allocated to the development endpoint.

The maximum number of workers you can define are 299 for G.1X, and 149 for G.2X.

Type: Integer

Required: No

PrivateAddress

A private IP address to access the DevEndpoint within a VPC if the DevEndpoint is created within one. The PrivateAddress field is present only when you create the DevEndpoint within your VPC.

Type: String

Required: No

PublicAddress

The public IP address used by this DevEndpoint. The PublicAddress field is present only when you create a non-virtual private cloud (VPC) DevEndpoint.

Type: String

Required: No

PublicKey

The public key to be used by this DevEndpoint for authentication. This attribute is provided for backward compatibility because the recommended attribute to use is public keys.

Type: String

Required: No

PublicKeys

A list of public keys to be used by the DevEndpoints for authentication. Using this attribute is preferred over a single public key because the public keys allow you to have a different private key per client.

Note

If you previously created an endpoint with a public key, you must remove that key to be able to set a list of public keys. Call the UpdateDevEndpoint API operation with the

public key content in the `deletePublicKeys` attribute, and the list of new keys in the `addPublicKeys` attribute.

Type: Array of strings

Array Members: Maximum number of 5 items.

Required: No

RoleArn

The Amazon Resource Name (ARN) of the IAM role used in this `DevEndpoint`.

Type: String

Pattern: `arn:aws:iam::\d{12}:role/.*`

Required: No

SecurityConfiguration

The name of the `SecurityConfiguration` structure to be used with this `DevEndpoint`.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

SecurityGroupIds

A list of security group identifiers used in this `DevEndpoint`.

Type: Array of strings

Required: No

Status

The current status of this `DevEndpoint`.

Type: String

Required: No

SubnetId

The subnet ID for this DevEndpoint.

Type: String

Required: No

VpcId

The ID of the virtual private cloud (VPC) used by this DevEndpoint.

Type: String

Required: No

WorkerType

The type of predefined worker that is allocated to the development endpoint. Accepts a value of Standard, G.1X, or G.2X.

- For the Standard worker type, each worker provides 4 vCPU, 16 GB of memory and a 50GB disk, and 2 executors per worker.
- For the G.1X worker type, each worker maps to 1 DPU (4 vCPU, 16 GB of memory, 64 GB disk), and provides 1 executor per worker. We recommend this worker type for memory-intensive jobs.
- For the G.2X worker type, each worker maps to 2 DPU (8 vCPU, 32 GB of memory, 128 GB disk), and provides 1 executor per worker. We recommend this worker type for memory-intensive jobs.

Known issue: when a development endpoint is created with the G.2X WorkerType configuration, the Spark drivers for the development endpoint will run on 4 vCPU, 16 GB of memory, and a 64 GB disk.

Type: String

Valid Values: Standard | G.1X | G.2X | G.025X | G.4X | G.8X | Z.2X

Required: No

YarnEndpointAddress

The YARN endpoint address used by this DevEndpoint.

Type: String

Required: No

ZeppelinRemoteSparkInterpreterPort

The Apache Zeppelin port for the remote Apache Spark interpreter.

Type: Integer

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)


DevEndpointCustomLibraries

Custom libraries to be loaded into a development endpoint.

Contents

ExtraJarsS3Path

The path to one or more Java `.jar` files in an S3 bucket that should be loaded in your DevEndpoint.

 **Note**


You can only use pure Java/Scala libraries with a DevEndpoint.

Type: String

Required: No

ExtraPythonLibsS3Path

The paths to one or more Python libraries in an Amazon Simple Storage Service (Amazon S3) bucket that should be loaded in your DevEndpoint. Multiple values must be complete paths separated by a comma.

 **Note**

You can only use pure Python libraries with a DevEndpoint. Libraries that rely on C extensions, such as the [pandas](#) Python data analysis library, are not currently supported.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DirectJDBCSource

Specifies the direct JDBC source connection.

Contents

ConnectionName

The connection name of the JDBC source.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

ConnectionType

The connection type of the JDBC source.

Type: String

Valid Values: `sqlserver` | `mysql` | `oracle` | `postgresql` | `redshift`

Required: Yes

Database

The database of the JDBC source connection.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

Name

The name of the JDBC source connection.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

Table

The table of the JDBC source connection.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

RedshiftTmpDir

The temp directory of the JDBC Redshift source.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DirectKafkaSource

Specifies an Apache Kafka data store.

Contents

Name

The name of the data store.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

DataPreviewOptions

Specifies options related to data preview for viewing a sample of your data.

Type: [StreamingDataPreviewOptions](#) object

Required: No

DetectSchema

Whether to automatically determine the schema from the incoming data.

Type: Boolean

Required: No

StreamingOptions

Specifies the streaming options.

Type: [KafkaStreamingSourceOptions](#) object

Required: No

WindowSize

The amount of time to spend processing each micro batch.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DirectKinesisSource

Specifies a direct Amazon Kinesis data source.

Contents

Name

The name of the data source.

Type: String

Pattern: (`[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF]` | `[\^\r\n]`)*

Required: Yes

DataPreviewOptions

Additional options for data preview.

Type: [StreamingDataPreviewOptions](#) object

Required: No

DetectSchema

Whether to automatically determine the schema from the incoming data.

Type: Boolean

Required: No

StreamingOptions

Additional options for the Kinesis streaming data source.

Type: [KinesisStreamingSourceOptions](#) object

Required: No

WindowSize

The amount of time to spend processing each micro batch.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DirectSchemaChangePolicy

A policy that specifies update behavior for the crawler.

Contents

Database

Specifies the database that the schema change policy applies to.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

EnableUpdateCatalog

Whether to use the specified update behavior when the crawler finds a changed schema.

Type: Boolean

Required: No

Table

Specifies the table in the database that the schema change policy applies to.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

UpdateBehavior

The update behavior when the crawler finds a changed schema.

Type: String

Valid Values: UPDATE_IN_DATABASE | LOG

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DoubleColumnStatisticsData

Defines column statistics supported for floating-point number data columns.

Contents

NumberOfDistinctValues

The number of distinct values in a column.

Type: Long

Valid Range: Minimum value of 0.

Required: Yes

NumberOfNulls

The number of null values in the column.

Type: Long

Valid Range: Minimum value of 0.

Required: Yes

MaximumValue

The highest value in the column.

Type: Double

Required: No

MinimumValue

The lowest value in the column.

Type: Double

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DQResultsPublishingOptions

Options to configure how your data quality evaluation results are published.

Contents

CloudWatchMetricsEnabled

Enable metrics for your data quality results.

Type: Boolean

Required: No

EvaluationContext

The context of the evaluation.

Type: String

Pattern: [A-Za-z0-9_-]*

Required: No

ResultsPublishingEnabled

Enable publishing for your data quality results.

Type: Boolean

Required: No

ResultsS3Prefix

The Amazon S3 prefix prepended to the results.

Type: String

Pattern: ([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF]|[\^\\S\\r\\n"'])*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DQStopJobOnFailureOptions

Options to configure how your job will stop if your data quality evaluation fails.

Contents

StopJobOnFailureTiming

When to stop job if your data quality evaluation fails. Options are Immediate or AfterDataLoad.

Type: String

Valid Values: Immediate | AfterDataLoad

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DropDuplicates

Specifies a transform that removes rows of repeating data from a data set.

Contents

Inputs

The data inputs identified by their node names.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: [A-Za-z0-9_-]*

Required: Yes

Name

The name of the transform node.

Type: String

Pattern: ([\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF]|[\^\r\n])*

Required: Yes

Columns

The name of the columns to be merged or removed if repeating.

Type: Array of arrays of strings

Pattern: [A-Za-z0-9_-]*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DropFields

Specifies a transform that chooses the data property keys that you want to drop.

Contents

Inputs

The data inputs identified by their node names.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]*`

Required: Yes

Name

The name of the transform node.

Type: String

Pattern: `([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\r\n])*`

Required: Yes

Paths

A JSON path to a variable in the data structure.

Type: Array of arrays of strings

Pattern: `([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])*`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DropNullFields

Specifies a transform that removes columns from the dataset if all values in the column are 'null'. By default, AWS Glue Studio will recognize null objects, but some values such as empty strings, strings that are "null", -1 integers or other placeholders such as zeros, are not automatically recognized as nulls.

Contents

Inputs

The data inputs identified by their node names.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: [A-Za-z0-9_-]*

Required: Yes

Name

The name of the transform node.

Type: String

Pattern: ([\u0020-\u007F\u00E0-\u00FF\u0080-\u00FF\u0080-\u00FF\u0080-\u00FF\u0080-\u00FF\u0080-\u00FF] | [^\r\n])*

Required: Yes

NullCheckBoxList

A structure that represents whether certain values are recognized as null values for removal.

Type: [NullCheckBoxList](#) object

Required: No

NullTextList

A structure that specifies a list of NullValueField structures that represent a custom null value such as zero or other value being used as a null placeholder unique to the dataset.

The `DropNullFields` transform removes custom null values only if both the value of the null placeholder and the datatype match the data.

Type: Array of [NullValueField](#) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DynamicTransform

Specifies the set of parameters needed to perform the dynamic transform.

Contents

FunctionName

Specifies the name of the function of the dynamic transform.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])*`)

Required: Yes

Inputs

Specifies the inputs for the dynamic transform that are required.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]*`

Required: Yes

Name

Specifies the name of the dynamic transform.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])*`)

Required: Yes

Path

Specifies the path of the dynamic transform source and config files.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

TransformName

Specifies the name of the dynamic transform as it appears in the AWS Glue Studio visual editor.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

OutputSchemas

Specifies the data schema for the dynamic transform.

Type: Array of [GlueSchema](#) objects

Required: No

Parameters

Specifies the parameters of the dynamic transform.

Type: Array of [TransformConfigParameter](#) objects

Required: No

Version

This field is not used and will be deprecated in future release.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DynamoDBCatalogSource

Specifies a DynamoDB data source in the AWS Glue Data Catalog.

Contents

Database

The name of the database to read from.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

Name

The name of the data source.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

Table

The name of the table in the database to read from.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DynamoDBTarget

Specifies an Amazon DynamoDB table to crawl.

Contents

Path

The name of the DynamoDB table to crawl.

Type: String

Required: No

scanAll

Indicates whether to scan all the records, or to sample rows from the table. Scanning all the records can take a long time when the table is not a high throughput table.

A value of `true` means to scan all records, while a value of `false` means to sample the records. If no value is specified, the value defaults to `true`.

Type: Boolean

Required: No

scanRate

The percentage of the configured read capacity units to use by the AWS Glue crawler. Read capacity units is a term defined by DynamoDB, and is a numeric value that acts as rate limiter for the number of reads that can be performed on that table per second.

The valid values are null or a value between 0.1 to 1.5. A null value is used when user does not provide a value, and defaults to 0.5 of the configured Read Capacity Unit (for provisioned tables), or 0.25 of the max configured Read Capacity Unit (for tables using on-demand mode).

Type: Double

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Edge

An edge represents a directed connection between two AWS Glue components that are part of the workflow the edge belongs to.

Contents

DestinationId

The unique of the node within the workflow where the edge ends.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

SourceId

The unique of the node within the workflow where the edge starts.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

EncryptionAtRest

Specifies the encryption-at-rest configuration for the Data Catalog.

Contents

CatalogEncryptionMode

The encryption-at-rest mode for encrypting Data Catalog data.

Type: String

Valid Values: DISABLED | SSE-KMS | SSE-KMS-WITH-SERVICE-ROLE

Required: Yes

CatalogEncryptionServiceRole

The role that AWS Glue assumes to encrypt and decrypt the Data Catalog objects on the caller's behalf.

Type: String

Pattern: `^arn:aws(-[cn|us-gov|iso(-[bef])?]):iam::[0-9]{12}:role/.+`

Required: No

SseAwsKmsKeyId

The ID of the AWS KMS key to use for encryption at rest.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

EncryptionConfiguration

Specifies an encryption configuration.

Contents

CloudWatchEncryption

The encryption configuration for Amazon CloudWatch.

Type: [CloudWatchEncryption](#) object

Required: No

JobBookmarksEncryption

The encryption configuration for job bookmarks.

Type: [JobBookmarksEncryption](#) object

Required: No

S3Encryption

The encryption configuration for Amazon Simple Storage Service (Amazon S3) data.

Type: Array of [S3Encryption](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ErrorDetail

Contains details about an error.

Contents

ErrorCode

The code associated with this error.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

ErrorMessage

A message describing the error.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ErrorDetails

An object containing error details.

Contents

ErrorCode

The error code for an error.

Type: String

Required: No

ErrorMessage

The error message for an error.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

EvaluateDataQuality

Specifies your data quality evaluation criteria.

Contents

Inputs

The inputs of your data quality evaluation.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]*`

Required: Yes

Name

The name of the data quality evaluation.

Type: String

Pattern: `([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF]|[\^\r\n])*`

Required: Yes

Ruleset

The ruleset for your data quality evaluation.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 65536.

Pattern: `([\u0020-\u007E\r\s\n])*`

Required: Yes

Output

The output of your data quality evaluation.

Type: String

Valid Values: PrimaryInput | EvaluationResults

Required: No

PublishingOptions

Options to configure how your results are published.

Type: [DQResultsPublishingOptions](#) object

Required: No

StopJobOnFailureOptions

Options to configure how your job will stop if your data quality evaluation fails.

Type: [DQStopJobOnFailureOptions](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

EvaluateDataQualityMultiFrame

Specifies your data quality evaluation criteria.

Contents

Inputs

The inputs of your data quality evaluation. The first input in this list is the primary data source.

Type: Array of strings

Array Members: Minimum number of 1 item.

Pattern: [A-Za-z0-9_-]*

Required: Yes

Name

The name of the data quality evaluation.

Type: String

Pattern: ([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF]|[\^\r\n])*

Required: Yes

Ruleset

The ruleset for your data quality evaluation.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 65536.

Pattern: ([\u0020-\u007E\r\s\n])*

Required: Yes

AdditionalDataSources

The aliases of all data sources except primary.

Type: String to string map

Key Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Value Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

AdditionalOptions

Options to configure runtime behavior of the transform.

Type: String to string map

Valid Keys: `performanceTuning.caching` | `observations.scope`

Required: No

PublishingOptions

Options to configure how your results are published.

Type: [DQResultsPublishingOptions](#) object

Required: No

StopJobOnFailureOptions

Options to configure how your job will stop if your data quality evaluation fails.

Type: [DQStopJobOnFailureOptions](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

EvaluationMetrics

Evaluation metrics provide an estimate of the quality of your machine learning transform.

Contents

TransformType

The type of machine learning transform.

Type: String

Valid Values: FIND_MATCHES

Required: Yes

FindMatchesMetrics

The evaluation metrics for the find matches algorithm.

Type: [FindMatchesMetrics](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

EventBatchingCondition

Batch condition that must be met (specified number of events received or batch time window expired) before EventBridge event trigger fires.

Contents

BatchSize

Number of events that must be received from Amazon EventBridge before EventBridge event trigger fires.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 100.

Required: Yes

BatchWindow

Window of time in seconds after which EventBridge event trigger fires. Window starts when first event is received.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 900.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ExecutionProperty

An execution property of a job.

Contents

MaxConcurrentRuns

The maximum number of concurrent runs allowed for the job. The default is 1. An error is returned when this threshold is reached. The maximum value you can specify is controlled by a service limit.

Type: Integer

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ExportLabelsTaskRunProperties

Specifies configuration properties for an exporting labels task run.

Contents

OutputS3Path

The Amazon Simple Storage Service (Amazon S3) path where you will export the labels.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

FederatedDatabase

A database that points to an entity outside the AWS Glue Data Catalog.

Contents

ConnectionName

The name of the connection to the external metastore.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Identifier

A unique identifier for the federated database.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 512.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

FederatedTable

A table that points to an entity outside the AWS Glue Data Catalog.

Contents

ConnectionName

The name of the connection to the external metastore.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

DatabaseIdentifier

A unique identifier for the federated database.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 512.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Identifier

A unique identifier for the federated table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 512.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

FillMissingValues

Specifies a transform that locates records in the dataset that have missing values and adds a new field with a value determined by imputation. The input data set is used to train the machine learning model that determines what the missing value should be.

Contents

ImputedPath

A JSON path to a variable in the data structure for the dataset that is imputed.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | (`^\\S\\r\\n''`))*

Required: Yes

Inputs

The data inputs identified by their node names.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]`*

Required: Yes

Name

The name of the transform node.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | (`^\\r\\n`))*

Required: Yes

FilledPath

A JSON path to a variable in the data structure for the dataset that is filled.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF]|[\^\S\r\n"'])*`)

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Filter

Specifies a transform that splits a dataset into two, based on a filter condition.

Contents

Filters

Specifies a filter expression.

Type: Array of [FilterExpression](#) objects

Required: Yes

Inputs

The data inputs identified by their node names.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]*`

Required: Yes

LogicalOperator

The operator used to filter rows by comparing the key value to a specified value.

Type: String

Valid Values: AND | OR

Required: Yes

Name

The name of the transform node.

Type: String

Pattern: `([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\r\n])*`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

FilterExpression

Specifies a filter expression.

Contents

Operation

The type of operation to perform in the expression.

Type: String

Valid Values: EQ | LT | GT | LTE | GTE | REGEX | ISNULL

Required: Yes

Values

A list of filter values.

Type: Array of [FilterValue](#) objects

Required: Yes

Negated

Whether the expression is to be negated.

Type: Boolean

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

FilterValue

Represents a single entry in the list of values for a `FilterExpression`.

Contents

Type

The type of filter value.

Type: String

Valid Values: COLUMNEXTRACTED | CONSTANT

Required: Yes

Value

The value to be associated.

Type: Array of strings

Pattern: (`[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\S\x\n"']`)*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

FindMatchesMetrics

The evaluation metrics for the find matches algorithm. The quality of your machine learning transform is measured by getting your transform to predict some matches and comparing the results to known matches from the same dataset. The quality metrics are based on a subset of your data, so they are not precise.

Contents

AreaUnderPRCurve

The area under the precision/recall curve (AUPRC) is a single number measuring the overall quality of the transform, that is independent of the choice made for precision vs. recall. Higher values indicate that you have a more attractive precision vs. recall tradeoff.

For more information, see [Precision and recall](#) in Wikipedia.

Type: Double

Valid Range: Minimum value of 0.0. Maximum value of 1.0.

Required: No

ColumnImportances

A list of `ColumnImportance` structures containing column importance metrics, sorted in order of descending importance.

Type: Array of [ColumnImportance](#) objects

Array Members: Minimum number of 0 items. Maximum number of 100 items.

Required: No

ConfusionMatrix

The confusion matrix shows you what your transform is predicting accurately and what types of errors it is making.

For more information, see [Confusion matrix](#) in Wikipedia.

Type: [ConfusionMatrix](#) object

Required: No

F1

The maximum F1 metric indicates the transform's accuracy between 0 and 1, where 1 is the best accuracy.

For more information, see [F1 score](#) in Wikipedia.

Type: Double

Valid Range: Minimum value of 0.0. Maximum value of 1.0.

Required: No

Precision

The precision metric indicates when often your transform is correct when it predicts a match. Specifically, it measures how well the transform finds true positives from the total true positives possible.

For more information, see [Precision and recall](#) in Wikipedia.

Type: Double

Valid Range: Minimum value of 0.0. Maximum value of 1.0.

Required: No

Recall

The recall metric indicates that for an actual match, how often your transform predicts the match. Specifically, it measures how well the transform finds true positives from the total records in the source data.

For more information, see [Precision and recall](#) in Wikipedia.

Type: Double

Valid Range: Minimum value of 0.0. Maximum value of 1.0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

FindMatchesParameters

The parameters to configure the find matches transform.

Contents

AccuracyCostTradeoff

The value that is selected when tuning your transform for a balance between accuracy and cost. A value of 0.5 means that the system balances accuracy and cost concerns. A value of 1.0 means a bias purely for accuracy, which typically results in a higher cost, sometimes substantially higher. A value of 0.0 means a bias purely for cost, which results in a less accurate FindMatches transform, sometimes with unacceptable accuracy.

Accuracy measures how well the transform finds true positives and true negatives. Increasing accuracy requires more machine resources and cost. But it also results in increased recall.

Cost measures how many compute resources, and thus money, are consumed to run the transform.

Type: Double

Valid Range: Minimum value of 0.0. Maximum value of 1.0.

Required: No

EnforceProvidedLabels

The value to switch on or off to force the output to match the provided labels from users. If the value is `True`, the find matches transform forces the output to match the provided labels. The results override the normal conflation results. If the value is `False`, the find matches transform does not ensure all the labels provided are respected, and the results rely on the trained model.

Note that setting this value to true may increase the conflation execution time.

Type: Boolean

Required: No

PrecisionRecallTradeoff

The value selected when tuning your transform for a balance between precision and recall. A value of 0.5 means no preference; a value of 1.0 means a bias purely for precision, and a value

of 0.0 means a bias for recall. Because this is a tradeoff, choosing values close to 1.0 means very low recall, and choosing values close to 0.0 results in very low precision.

The precision metric indicates how often your model is correct when it predicts a match.

The recall metric indicates that for an actual match, how often your model predicts the match.

Type: Double

Valid Range: Minimum value of 0.0. Maximum value of 1.0.

Required: No

PrimaryKeyColumnName

The name of a column that uniquely identifies rows in the source table. Used to help identify matching records.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

FindMatchesTaskRunProperties

Specifies configuration properties for a Find Matches task run.

Contents

JobId

The job ID for the Find Matches task run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

JobName

The name assigned to the job for the Find Matches task run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

JobRunId

The job run ID for the Find Matches task run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

GetConnectionsFilter

Filters the connection definitions that are returned by the `GetConnections` API operation.

Contents

ConnectionType

The type of connections to return. Currently, SFTP is not supported.

Type: String

Valid Values: JDBC | SFTP | MONGODB | KAFKA | NETWORK | MARKETPLACE | CUSTOM
| SALESFORCE

Required: No

MatchCriteria

A criteria string that must match the criteria recorded in the connection definition for that connection definition to be returned.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 10 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

GluePolicy

A structure for returning a resource policy.

Contents

CreateTime

The date and time at which the policy was created.

Type: Timestamp

Required: No

PolicyHash

Contains the hash value associated with this policy.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

PolicyInJson

Contains the requested policy document, in JSON format.

Type: String

Length Constraints: Minimum length of 2.

Required: No

UpdateTime

The date and time at which the policy was last updated.

Type: Timestamp

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

GlueSchema

Specifies a user-defined schema when a schema cannot be determined by AWS Glue.

Contents

Columns

Specifies the column definitions that make up a AWS Glue schema.

Type: Array of [GlueStudioSchemaColumn](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

GlueStudioSchemaColumn

Specifies a single column in a AWS Glue schema definition.

Contents

Name

The name of the column in the AWS Glue Studio schema.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 1024.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Type

The hive type for this column in the AWS Glue Studio schema.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 131072.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

GlueTable

The database and table in the AWS Glue Data Catalog that is used for input or output data.

Contents

DatabaseName

A database name in the AWS Glue Data Catalog.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

TableName

A table name in the AWS Glue Data Catalog.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

AdditionalOptions

Additional options for the table. Currently there are two keys supported:

- `pushDownPredicate`: to filter on partitions without having to list and read all the files in your dataset.
- `catalogPartitionPredicate`: to use server-side partition pruning using partition indexes in the AWS Glue Data Catalog.

Type: String to string map

Map Entries: Maximum number of 10 items.

Key Length Constraints: Minimum length of 1. Maximum length of 255.

Key Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Value Length Constraints: Minimum length of 0. Maximum length of 2048.

Value Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

CatalogId

A unique identifier for the AWS Glue Data Catalog.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

ConnectionName

The name of the connection to the AWS Glue Data Catalog.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

GovernedCatalogSource

Specifies the data store in the governed AWS Glue Data Catalog.

Contents

Database

The database to read from.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])*`)

Required: Yes

Name

The name of the data store.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\r\n])*`)

Required: Yes

Table

The database table to read from.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])*`)

Required: Yes

AdditionalOptions

Specifies additional connection options.

Type: [S3SourceAdditionalOptions](#) object

Required: No

PartitionPredicate

Partitions satisfying this predicate are deleted. Files within the retention period in these partitions are not deleted. Set to "" – empty by default.

Type: String

Pattern: (`[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF]` | `[\S\r\n"']`)*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

GovernedCatalogTarget

Specifies a data target that writes to Amazon S3 using the AWS Glue Data Catalog.

Contents

Database

The name of the database to write to.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]|[^\\S\\r\\n"']`)*

Required: Yes

Inputs

The nodes that are inputs to the data target.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]`*

Required: Yes

Name

The name of the data target.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]|[^\\r\\n]`)*

Required: Yes

Table

The name of the table in the database to write to.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]|[^\\S\\r\\n"']`)*

Required: Yes

PartitionKeys

Specifies native partitioning using a sequence of keys.

Type: Array of arrays of strings

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]|[^\\S\\r\\n"']`)*

Required: No

SchemaChangePolicy

A policy that specifies update behavior for the governed catalog.

Type: [CatalogSchemaChangePolicy](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

GrokClassifier

A classifier that uses grok patterns.

Contents

Classification

An identifier of the data format that the classifier matches, such as Twitter, JSON, Omniture logs, and so on.

Type: String

Required: Yes

GrokPattern

The grok pattern applied to a data store by this classifier. For more information, see built-in patterns in [Writing Custom Classifiers](#).

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\t]*`

Required: Yes

Name

The name of the classifier.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

CreationTime

The time that this classifier was registered.

Type: Timestamp

Required: No

CustomPatterns

Optional custom grok patterns defined by this classifier. For more information, see custom patterns in [Writing Custom Classifiers](#).

Type: String

Length Constraints: Minimum length of 0. Maximum length of 16000.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

LastUpdated

The time that this classifier was last updated.

Type: Timestamp

Required: No

Version

The version of this classifier.

Type: Long

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

HudiTarget

Specifies an Apache Hudi data source.

Contents

ConnectionName

The name of the connection to use to connect to the Hudi target. If your Hudi files are stored in buckets that require VPC authorization, you can set their connection properties here.

Type: String

Required: No

Exclusions

A list of glob patterns used to exclude from the crawl. For more information, see [Catalog Tables with a Crawler](#).

Type: Array of strings

Required: No

MaximumTraversalDepth

The maximum depth of Amazon S3 paths that the crawler can traverse to discover the Hudi metadata folder in your Amazon S3 path. Used to limit the crawler run time.

Type: Integer

Required: No

Paths

An array of Amazon S3 location strings for Hudi, each indicating the root folder with which the metadata files for a Hudi table resides. The Hudi folder may be located in a child folder of the root folder.

The crawler will scan all folders underneath a path for a Hudi folder.

Type: Array of strings

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

IcebergInput

A structure that defines an Apache Iceberg metadata table to create in the catalog.

Contents

MetadataOperation

A required metadata operation. Can only be set to CREATE.

Type: String

Valid Values: CREATE

Required: Yes

Version

The table version for the Iceberg table. Defaults to 2.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

IcebergTarget

Specifies an Apache Iceberg data source where Iceberg tables are stored in Amazon S3.

Contents

ConnectionName

The name of the connection to use to connect to the Iceberg target.

Type: String

Required: No

Exclusions

A list of glob patterns used to exclude from the crawl. For more information, see [Catalog Tables with a Crawler](#).

Type: Array of strings

Required: No

MaximumTraversalDepth

The maximum depth of Amazon S3 paths that the crawler can traverse to discover the Iceberg metadata folder in your Amazon S3 path. Used to limit the crawler run time.

Type: Integer

Required: No

Paths

One or more Amazon S3 paths that contains Iceberg metadata folders as `s3://bucket/prefix`.

Type: Array of strings

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ImportLabelsTaskRunProperties

Specifies configuration properties for an importing labels task run.

Contents

InputS3Path

The Amazon Simple Storage Service (Amazon S3) path from where you will import the labels.

Type: String

Required: No

Replace

Indicates whether to overwrite your existing labels.

Type: Boolean

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

JDBCConectorOptions

Additional connection options for the connector.

Contents

DataTypeMapping

Custom data type mapping that builds a mapping from a JDBC data type to an AWS Glue data type. For example, the option "dataTypeMapping":{"FLOAT":"STRING"} maps data fields of JDBC type FLOAT into the Java String type by calling the `ResultSet.getString()` method of the driver, and uses it to build the AWS Glue record. The `ResultSet` object is implemented by each driver, so the behavior is specific to the driver you use. Refer to the documentation for your JDBC driver to understand how the driver performs the conversions.

Type: String to string map

Valid Keys: ARRAY | BIGINT | BINARY | BIT | BLOB | BOOLEAN | CHAR | CLOB | DATALINK | DATE | DECIMAL | DISTINCT | DOUBLE | FLOAT | INTEGER | JAVA_OBJECT | LONGNVARCHAR | LONGVARBINARY | LONGVARCHAR | NCHAR | NCLOB | NULL | NUMERIC | NVARCHAR | OTHER | REAL | REF | REF_CURSOR | ROWID | SMALLINT | SQLXML | STRUCT | TIME | TIME_WITH_TIMEZONE | TIMESTAMP | TIMESTAMP_WITH_TIMEZONE | TINYINT | VARBINARY | VARCHAR

Valid Values: DATE | STRING | TIMESTAMP | INT | FLOAT | LONG | BIGDECIMAL | BYTE | SHORT | DOUBLE

Required: No

FilterPredicate

Extra condition clause to filter data from source. For example:

```
BillingCity='Mountain View'
```

When using a query instead of a table name, you should validate that the query works with the specified `filterPredicate`.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

JobBookmarkKeys

The name of the job bookmark keys on which to sort.

Type: Array of strings

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | (`^[S\\r\\n"']`))*

Required: No

JobBookmarkKeysSortOrder

Specifies an ascending or descending sort order.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | (`^[S\\r\\n"']`))*

Required: No

LowerBound

The minimum value of `partitionColumn` that is used to decide partition stride.

Type: Long

Valid Range: Minimum value of 0.

Required: No

NumPartitions

The number of partitions. This value, along with `lowerBound` (inclusive) and `upperBound` (exclusive), form partition strides for generated `WHERE` clause expressions that are used to split the `partitionColumn`.

Type: Long

Valid Range: Minimum value of 0.

Required: No

PartitionColumn

The name of an integer column that is used for partitioning. This option works only when it's included with `lowerBound`, `upperBound`, and `numPartitions`. This option works the same way as in the Spark SQL JDBC reader.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

UpperBound

The maximum value of `partitionColumn` that is used to decide partition stride.

Type: Long

Valid Range: Minimum value of 0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

JDBCConectorSource

Specifies a connector to a JDBC data source.

Contents

ConnectionName

The name of the connection that is associated with the connector.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n'']`)*

Required: Yes

ConnectionType

The type of connection, such as `marketplace.jdbc` or `custom.jdbc`, designating a connection to a JDBC data store.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n'']`)*

Required: Yes

ConnectorName

The name of a connector that assists with accessing the data store in AWS Glue Studio.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n'']`)*

Required: Yes

Name

The name of the data source.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

AdditionalOptions

Additional connection options for the connector.

Type: [JDBCConnectorOptions](#) object

Required: No

ConnectionTable

The name of the table in the data source.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n]`)*

Required: No

OutputSchemas

Specifies the data schema for the custom JDBC source.

Type: Array of [GlueSchema](#) objects

Required: No

Query

The table or SQL query to get the data from. You can specify either `ConnectionTable` or `query`, but not both.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF\\s]`)*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

JDBCConectorTarget

Specifies a data target that writes to Amazon S3 in Apache Parquet columnar storage.

Contents

ConnectionName

The name of the connection that is associated with the connector.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | (`^[S\\r\\n"']`))*

Required: Yes

ConnectionTable

The name of the table in the data target.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | (`^[S\\r\\n]`))*

Required: Yes

ConnectionType

The type of connection, such as `marketplace.jdbc` or `custom.jdbc`, designating a connection to a JDBC data target.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | (`^[S\\r\\n"']`))*

Required: Yes

ConnectorName

The name of a connector that will be used.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

Inputs

The nodes that are inputs to the data target.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]`*

Required: Yes

Name

The name of the data target.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

AdditionalOptions

Additional connection options for the connector.

Type: String to string map

Key Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Value Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

OutputSchemas

Specifies the data schema for the JDBC target.

Type: Array of [GlueSchema](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

JdbcTarget

Specifies a JDBC data store to crawl.

Contents

ConnectionName

The name of the connection to use to connect to the JDBC target.

Type: String

Required: No

EnableAdditionalMetadata

Specify a value of RAWTYPES or COMMENTS to enable additional metadata in table responses. RAWTYPES provides the native-level datatype. COMMENTS provides comments associated with a column or table in the database.

If you do not need additional metadata, keep the field empty.

Type: Array of strings

Valid Values: COMMENTS | RAWTYPES

Required: No

Exclusions

A list of glob patterns used to exclude from the crawl. For more information, see [Catalog Tables with a Crawler](#).

Type: Array of strings

Required: No

Path

The path of the JDBC target.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Job

Specifies a job definition.

Contents

AllocatedCapacity

This field is deprecated. Use `MaxCapacity` instead.

The number of AWS Glue data processing units (DPUs) allocated to runs of this job. You can allocate a minimum of 2 DPUs; the default is 10. A DPU is a relative measure of processing power that consists of 4 vCPUs of compute capacity and 16 GB of memory. For more information, see the [AWS Glue pricing page](#).

Type: Integer

Required: No

CodeGenConfigurationNodes

The representation of a directed acyclic graph on which both the Glue Studio visual component and Glue Studio code generation is based.

Type: String to [CodeGenConfigurationNode](#) object map

Key Pattern: `[A-Za-z0-9_-]*`

Required: No

Command

The `JobCommand` that runs this job.

Type: [JobCommand](#) object

Required: No

Connections

The connections used for this job.

Type: [ConnectionsList](#) object

Required: No

CreatedOn

The time and date that this job definition was created.

Type: Timestamp

Required: No

DefaultArguments

The default arguments for every run of this job, specified as name-value pairs.

You can specify arguments here that your own job-execution script consumes, as well as arguments that AWS Glue itself consumes.

Job arguments may be logged. Do not pass plaintext secrets as arguments. Retrieve secrets from a AWS Glue Connection, AWS Secrets Manager or other secret management mechanism if you intend to keep them within the Job.

For information about how to specify and consume your own Job arguments, see the [Calling AWS Glue APIs in Python](#) topic in the developer guide.

For information about the arguments you can provide to this field when configuring Spark jobs, see the [Special Parameters Used by AWS Glue](#) topic in the developer guide.

For information about the arguments you can provide to this field when configuring Ray jobs, see [Using job parameters in Ray jobs](#) in the developer guide.

Type: String to string map

Required: No

Description

A description of the job.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

ExecutionClass

Indicates whether the job is run with a standard or flexible execution class. The standard execution class is ideal for time-sensitive workloads that require fast job startup and dedicated resources.

The flexible execution class is appropriate for time-insensitive jobs whose start and completion times may vary.

Only jobs with AWS Glue version 3.0 and above and command type `glueetl` will be allowed to set `ExecutionClass` to `FLEX`. The flexible execution class is available for Spark jobs.

Type: String

Length Constraints: Maximum length of 16.

Valid Values: `FLEX` | `STANDARD`

Required: No

ExecutionProperty

An `ExecutionProperty` specifying the maximum number of concurrent runs allowed for this job.

Type: [ExecutionProperty](#) object

Required: No

GlueVersion

In Spark jobs, `GlueVersion` determines the versions of Apache Spark and Python that AWS Glue available in a job. The Python version indicates the version supported for jobs of type Spark.

Ray jobs should set `GlueVersion` to `4.0` or greater. However, the versions of Ray, Python and additional libraries available in your Ray job are determined by the `Runtime` parameter of the Job command.

For more information about the available AWS Glue versions and corresponding Spark and Python versions, see [Glue version](#) in the developer guide.

Jobs that are created without specifying a Glue version default to Glue 0.9.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `^\w+\.\w+$`

Required: No

JobMode

A mode that describes how a job was created. Valid values are:

- `SCRIPT` - The job was created using the AWS Glue Studio script editor.
- `VISUAL` - The job was created using the AWS Glue Studio visual editor.
- `NOTEBOOK` - The job was created using an interactive sessions notebook.

When the JobMode field is missing or null, `SCRIPT` is assigned as the default value.

Type: String

Valid Values: `SCRIPT` | `VISUAL` | `NOTEBOOK`

Required: No

LastModifiedOn

The last point in time when this job definition was modified.

Type: Timestamp

Required: No

LogUri

This field is reserved for future use.

Type: String

Required: No

MaintenanceWindow

This field specifies a day of the week and hour for a maintenance window for streaming jobs. AWS Glue periodically performs maintenance activities. During these maintenance windows, AWS Glue will need to restart your streaming jobs.

AWS Glue will restart the job within 3 hours of the specified maintenance window. For instance, if you set up the maintenance window for Monday at 10:00AM GMT, your jobs will be restarted between 10:00AM GMT to 1:00PM GMT.

Type: String

Pattern: `^(Sun|Mon|Tue|Wed|Thu|Fri|Sat):([01]?[0-9]|2[0-3])$`

Required: No

MaxCapacity

For Glue version 1.0 or earlier jobs, using the standard worker type, the number of AWS Glue data processing units (DPUs) that can be allocated when this job runs. A DPU is a relative measure of processing power that consists of 4 vCPUs of compute capacity and 16 GB of memory. For more information, see the [AWS Glue pricing page](#).

For Glue version 2.0 or later jobs, you cannot specify a `MaximumCapacity`. Instead, you should specify a `WorkerType` and the `NumberOfWorkers`.

Do not set `MaxCapacity` if using `WorkerType` and `NumberOfWorkers`.

The value that can be allocated for `MaxCapacity` depends on whether you are running a Python shell job, an Apache Spark ETL job, or an Apache Spark streaming ETL job:

- When you specify a Python shell job (`JobCommand.Name="pythonshell"`), you can allocate either 0.0625 or 1 DPU. The default is 0.0625 DPU.
- When you specify an Apache Spark ETL job (`JobCommand.Name="glueetl"`) or Apache Spark streaming ETL job (`JobCommand.Name="gluestreaming"`), you can allocate from 2 to 100 DPUs. The default is 10 DPUs. This job type cannot have a fractional DPU allocation.

Type: Double

Required: No

MaxRetries

The maximum number of times to retry this job after a `JobRun` fails.

Type: Integer

Required: No

Name

The name you assign to this job definition.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

NonOverridableArguments

Arguments for this job that are not overridden when providing job arguments in a job run, specified as name-value pairs.

Type: String to string map

Required: No

NotificationProperty

Specifies configuration properties of a job notification.

Type: [NotificationProperty](#) object

Required: No

NumberOfWorkers

The number of workers of a defined `workerType` that are allocated when a job runs.

Type: Integer

Required: No

ProfileName

The name of an AWS Glue usage profile associated with the job.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Role

The name or Amazon Resource Name (ARN) of the IAM role associated with this job.

Type: String

Required: No

SecurityConfiguration

The name of the SecurityConfiguration structure to be used with this job.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

SourceControlDetails

The details for a source control configuration for a job, allowing synchronization of job artifacts to or from a remote repository.

Type: [SourceControlDetails](#) object

Required: No

Timeout

The job timeout in minutes. This is the maximum time that a job run can consume resources before it is terminated and enters TIMEOUT status. The default is 2,880 minutes (48 hours) for batch jobs.

Streaming jobs must have timeout values less than 7 days or 10080 minutes. When the value is left blank, the job will be restarted after 7 days based if you have not setup a maintenance window. If you have setup maintenance window, it will be restarted during the maintenance window after 7 days.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

WorkerType

The type of predefined worker that is allocated when a job runs. Accepts a value of G.1X, G.2X, G.4X, G.8X or G.025X for Spark jobs. Accepts the value Z.2X for Ray jobs.

- For the G.1X worker type, each worker maps to 1 DPU (4 vCPUs, 16 GB of memory) with 84GB disk (approximately 34GB free), and provides 1 executor per worker. We recommend this worker type for workloads such as data transforms, joins, and queries, to offers a scalable and cost effective way to run most jobs.
- For the G.2X worker type, each worker maps to 2 DPU (8 vCPUs, 32 GB of memory) with 128GB disk (approximately 77GB free), and provides 1 executor per worker. We recommend this worker type for workloads such as data transforms, joins, and queries, to offers a scalable and cost effective way to run most jobs.
- For the G.4X worker type, each worker maps to 4 DPU (16 vCPUs, 64 GB of memory) with 256GB disk (approximately 235GB free), and provides 1 executor per worker. We recommend this worker type for jobs whose workloads contain your most demanding transforms, aggregations, joins, and queries. This worker type is available only for AWS Glue version 3.0 or later Spark ETL jobs in the following AWS Regions: US East (Ohio), US East (N. Virginia), US West (Oregon), Asia Pacific (Singapore), Asia Pacific (Sydney), Asia Pacific (Tokyo), Canada (Central), Europe (Frankfurt), Europe (Ireland), and Europe (Stockholm).
- For the G.8X worker type, each worker maps to 8 DPU (32 vCPUs, 128 GB of memory) with 512GB disk (approximately 487GB free), and provides 1 executor per worker. We recommend this worker type for jobs whose workloads contain your most demanding transforms, aggregations, joins, and queries. This worker type is available only for AWS Glue version 3.0 or later Spark ETL jobs, in the same AWS Regions as supported for the G.4X worker type.
- For the G.025X worker type, each worker maps to 0.25 DPU (2 vCPUs, 4 GB of memory) with 84GB disk (approximately 34GB free), and provides 1 executor per worker. We recommend this worker type for low volume streaming jobs. This worker type is only available for AWS Glue version 3.0 streaming jobs.
- For the Z.2X worker type, each worker maps to 2 M-DPU (8vCPUs, 64 GB of memory) with 128 GB disk (approximately 120GB free), and provides up to 8 Ray workers based on the autoscaler.

Type: String

Valid Values: Standard | G.1X | G.2X | G.025X | G.4X | G.8X | Z.2X

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

JobBookmarkEntry

Defines a point that a job can resume processing.

Contents

Attempt

The attempt ID number.

Type: Integer

Required: No

JobBookmark

The bookmark itself.

Type: String

Required: No

JobName

The name of the job in question.

Type: String

Required: No

PreviousRunId

The unique run identifier associated with the previous job run.

Type: String

Required: No

Run

The run ID number.

Type: Integer

Required: No

RunId

The run ID number.

Type: String

Required: No

Version

The version of the job.

Type: Integer

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

JobBookmarksEncryption

Specifies how job bookmark data should be encrypted.

Contents

JobBookmarksEncryptionMode

The encryption mode to use for job bookmarks data.

Type: String

Valid Values: DISABLED | CSE-KMS

Required: No

KmsKeyArn

The Amazon Resource Name (ARN) of the KMS key to be used to encrypt the data.

Type: String

Pattern: arn:aws:kms:.*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

JobCommand

Specifies code that runs when a job is run.

Contents

Name

The name of the job command. For an Apache Spark ETL job, this must be `glueetl`. For a Python shell job, it must be `pythonshell`. For an Apache Spark streaming ETL job, this must be `gluestreaming`. For a Ray job, this must be `glueray`.

Type: String

Required: No

PythonVersion

The Python version being used to run a Python shell job. Allowed values are 2 or 3.

Type: String

Pattern: `^[2-3]|3[.]9)$`

Required: No

Runtime

In Ray jobs, Runtime is used to specify the versions of Ray, Python and additional libraries available in your environment. This field is not used in other job types. For supported runtime environment values, see [Supported Ray runtime environments](#) in the AWS Glue Developer Guide.

Type: String

Length Constraints: Maximum length of 64.

Pattern: `.*`

Required: No

ScriptLocation

Specifies the Amazon Simple Storage Service (Amazon S3) path to a script that runs a job.

Type: String

Length Constraints: Maximum length of 400000.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

JobNodeDetails

The details of a Job node present in the workflow.

Contents

JobRuns

The information for the job runs represented by the job node.

Type: Array of [JobRun](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

JobRun

Contains information about a job run.

Contents

AllocatedCapacity

This field is deprecated. Use `MaxCapacity` instead.

The number of AWS Glue data processing units (DPUs) allocated to this JobRun. From 2 to 100 DPUs can be allocated; the default is 10. A DPU is a relative measure of processing power that consists of 4 vCPUs of compute capacity and 16 GB of memory. For more information, see the [AWS Glue pricing page](#).

Type: Integer

Required: No

Arguments

The job arguments associated with this run. For this job run, they replace the default arguments set in the job definition itself.

You can specify arguments here that your own job-execution script consumes, as well as arguments that AWS Glue itself consumes.

Job arguments may be logged. Do not pass plaintext secrets as arguments. Retrieve secrets from a AWS Glue Connection, AWS Secrets Manager or other secret management mechanism if you intend to keep them within the Job.

For information about how to specify and consume your own Job arguments, see the [Calling AWS Glue APIs in Python](#) topic in the developer guide.

For information about the arguments you can provide to this field when configuring Spark jobs, see the [Special Parameters Used by AWS Glue](#) topic in the developer guide.

For information about the arguments you can provide to this field when configuring Ray jobs, see [Using job parameters in Ray jobs](#) in the developer guide.

Type: String to string map

Required: No

Attempt

The number of the attempt to run this job.

Type: Integer

Required: No

CompletedOn

The date and time that this job run completed.

Type: Timestamp

Required: No

DPUSecods

This field can be set for either job runs with execution class FLEX or when Auto Scaling is enabled, and represents the total time each executor ran during the lifecycle of a job run in seconds, multiplied by a DPU factor (1 for G.1X, 2 for G.2X, or 0.25 for G.025X workers). This value may be different than the `executionEngineRuntime * MaxCapacity` as in the case of Auto Scaling jobs, as the number of executors running at a given time may be less than the `MaxCapacity`. Therefore, it is possible that the value of `DPUSecods` is less than `executionEngineRuntime * MaxCapacity`.

Type: Double

Required: No

ErrorMessage

An error message associated with this job run.

Type: String

Required: No

ExecutionClass

Indicates whether the job is run with a standard or flexible execution class. The standard execution-class is ideal for time-sensitive workloads that require fast job startup and dedicated resources.

The flexible execution class is appropriate for time-insensitive jobs whose start and completion times may vary.

Only jobs with AWS Glue version 3.0 and above and command type `glueetl` will be allowed to set `ExecutionClass` to FLEX. The flexible execution class is available for Spark jobs.

Type: String

Length Constraints: Maximum length of 16.

Valid Values: FLEX | STANDARD

Required: No

ExecutionTime

The amount of time (in seconds) that the job run consumed resources.

Type: Integer

Required: No

GlueVersion

In Spark jobs, `GlueVersion` determines the versions of Apache Spark and Python that AWS Glue available in a job. The Python version indicates the version supported for jobs of type Spark.

Ray jobs should set `GlueVersion` to `4.0` or greater. However, the versions of Ray, Python and additional libraries available in your Ray job are determined by the `Runtime` parameter of the Job command.

For more information about the available AWS Glue versions and corresponding Spark and Python versions, see [Glue version](#) in the developer guide.

Jobs that are created without specifying a Glue version default to Glue 0.9.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `^\w+\.\w+$`

Required: No

Id

The ID of this job run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

JobMode

A mode that describes how a job was created. Valid values are:

- **SCRIPT** - The job was created using the AWS Glue Studio script editor.
- **VISUAL** - The job was created using the AWS Glue Studio visual editor.
- **NOTEBOOK** - The job was created using an interactive sessions notebook.

When the JobMode field is missing or null, **SCRIPT** is assigned as the default value.

Type: String

Valid Values: **SCRIPT** | **VISUAL** | **NOTEBOOK**

Required: No

JobName

The name of the job definition being used in this run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

JobRunState

The current state of the job run. For more information about the statuses of jobs that have terminated abnormally, see [AWS Glue Job Run Statuses](#).

Type: String

Valid Values: STARTING | RUNNING | STOPPING | STOPPED | SUCCEEDED | FAILED | TIMEOUT | ERROR | WAITING | EXPIRED

Required: No

LastModifiedOn

The last time that this job run was modified.

Type: Timestamp

Required: No

LogGroupName

The name of the log group for secure logging that can be server-side encrypted in Amazon CloudWatch using AWS KMS. This name can be `/aws-glue/jobs/`, in which case the default encryption is NONE. If you add a role name and SecurityConfiguration name (in other words, `/aws-glue/jobs-yourRoleName-yourSecurityConfigurationName/`), then that security configuration is used to encrypt the log group.

Type: String

Required: No

MaintenanceWindow

This field specifies a day of the week and hour for a maintenance window for streaming jobs. AWS Glue periodically performs maintenance activities. During these maintenance windows, AWS Glue will need to restart your streaming jobs.

AWS Glue will restart the job within 3 hours of the specified maintenance window. For instance, if you set up the maintenance window for Monday at 10:00AM GMT, your jobs will be restarted between 10:00AM GMT to 1:00PM GMT.

Type: String

Pattern: `^(Sun|Mon|Tue|Wed|Thu|Fri|Sat):([01]?[0-9]|2[0-3])$`

Required: No

MaxCapacity

For Glue version 1.0 or earlier jobs, using the standard worker type, the number of AWS Glue data processing units (DPUs) that can be allocated when this job runs. A DPU is a relative

measure of processing power that consists of 4 vCPUs of compute capacity and 16 GB of memory. For more information, see the [AWS Glue pricing page](#).

For Glue version 2.0+ jobs, you cannot specify a `MaximumCapacity`. Instead, you should specify a `WorkerType` and the `NumberOfWorkers`.

Do not set `MaxCapacity` if using `WorkerType` and `NumberOfWorkers`.

The value that can be allocated for `MaxCapacity` depends on whether you are running a Python shell job, an Apache Spark ETL job, or an Apache Spark streaming ETL job:

- When you specify a Python shell job (`JobCommand.Name="pythonshell"`), you can allocate either 0.0625 or 1 DPU. The default is 0.0625 DPU.
- When you specify an Apache Spark ETL job (`JobCommand.Name="glueetl"`) or Apache Spark streaming ETL job (`JobCommand.Name="gluestreaming"`), you can allocate from 2 to 100 DPUs. The default is 10 DPUs. This job type cannot have a fractional DPU allocation.

Type: Double

Required: No

NotificationProperty

Specifies configuration properties of a job run notification.

Type: [NotificationProperty](#) object

Required: No

NumberOfWorkers

The number of workers of a defined `workerType` that are allocated when a job runs.

Type: Integer

Required: No

PredecessorRuns

A list of predecessors to this job run.

Type: Array of [Predecessor](#) objects

Required: No

PreviousRunId

The ID of the previous run of this job. For example, the JobRunId specified in the StartJobRun action.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

ProfileName

The name of an AWS Glue usage profile associated with the job run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

SecurityConfiguration

The name of the SecurityConfiguration structure to be used with this job run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

StartedOn

The date and time at which this job run was started.

Type: Timestamp

Required: No

Timeout

The JobRun timeout in minutes. This is the maximum time that a job run can consume resources before it is terminated and enters TIMEOUT status. This value overrides the timeout value set in the parent job.

Streaming jobs must have timeout values less than 7 days or 10080 minutes. When the value is left blank, the job will be restarted after 7 days based if you have not setup a maintenance window. If you have setup maintenance window, it will be restarted during the maintenance window after 7 days.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

TriggerName

The name of the trigger that started this job run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

WorkerType

The type of predefined worker that is allocated when a job runs. Accepts a value of G.1X, G.2X, G.4X, G.8X or G.025X for Spark jobs. Accepts the value Z.2X for Ray jobs.

- For the G.1X worker type, each worker maps to 1 DPU (4 vCPUs, 16 GB of memory) with 84GB disk (approximately 34GB free), and provides 1 executor per worker. We recommend this worker type for workloads such as data transforms, joins, and queries, to offers a scalable and cost effective way to run most jobs.
- For the G.2X worker type, each worker maps to 2 DPU (8 vCPUs, 32 GB of memory) with 128GB disk (approximately 77GB free), and provides 1 executor per worker. We recommend this worker type for workloads such as data transforms, joins, and queries, to offers a scalable and cost effective way to run most jobs.

- For the G.4X worker type, each worker maps to 4 DPU (16 vCPUs, 64 GB of memory) with 256GB disk (approximately 235GB free), and provides 1 executor per worker. We recommend this worker type for jobs whose workloads contain your most demanding transforms, aggregations, joins, and queries. This worker type is available only for AWS Glue version 3.0 or later Spark ETL jobs in the following AWS Regions: US East (Ohio), US East (N. Virginia), US West (Oregon), Asia Pacific (Singapore), Asia Pacific (Sydney), Asia Pacific (Tokyo), Canada (Central), Europe (Frankfurt), Europe (Ireland), and Europe (Stockholm).
- For the G.8X worker type, each worker maps to 8 DPU (32 vCPUs, 128 GB of memory) with 512GB disk (approximately 487GB free), and provides 1 executor per worker. We recommend this worker type for jobs whose workloads contain your most demanding transforms, aggregations, joins, and queries. This worker type is available only for AWS Glue version 3.0 or later Spark ETL jobs, in the same AWS Regions as supported for the G.4X worker type.
- For the G.025X worker type, each worker maps to 0.25 DPU (2 vCPUs, 4 GB of memory) with 84GB disk (approximately 34GB free), and provides 1 executor per worker. We recommend this worker type for low volume streaming jobs. This worker type is only available for AWS Glue version 3.0 streaming jobs.
- For the Z.2X worker type, each worker maps to 2 M-DPU (8vCPUs, 64 GB of memory) with 128 GB disk (approximately 120GB free), and provides up to 8 Ray workers based on the autoscaler.

Type: String

Valid Values: Standard | G.1X | G.2X | G.025X | G.4X | G.8X | Z.2X

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

JobUpdate

Specifies information used to update an existing job definition. The previous job definition is completely overwritten by this information.

Contents

AllocatedCapacity

This field is deprecated. Use `MaxCapacity` instead.

The number of AWS Glue data processing units (DPUs) to allocate to this job. You can allocate a minimum of 2 DPUs; the default is 10. A DPU is a relative measure of processing power that consists of 4 vCPUs of compute capacity and 16 GB of memory. For more information, see the [AWS Glue pricing page](#).

Type: Integer

Required: No

CodeGenConfigurationNodes

The representation of a directed acyclic graph on which both the Glue Studio visual component and Glue Studio code generation is based.

Type: String to [CodeGenConfigurationNode](#) object map

Key Pattern: `[A-Za-z0-9_-]*`

Required: No

Command

The `JobCommand` that runs this job (required).

Type: [JobCommand](#) object

Required: No

Connections

The connections used for this job.

Type: [ConnectionsList](#) object

Required: No

DefaultArguments

The default arguments for every run of this job, specified as name-value pairs.

You can specify arguments here that your own job-execution script consumes, as well as arguments that AWS Glue itself consumes.

Job arguments may be logged. Do not pass plaintext secrets as arguments. Retrieve secrets from a AWS Glue Connection, AWS Secrets Manager or other secret management mechanism if you intend to keep them within the Job.

For information about how to specify and consume your own Job arguments, see the [Calling AWS Glue APIs in Python](#) topic in the developer guide.

For information about the arguments you can provide to this field when configuring Spark jobs, see the [Special Parameters Used by AWS Glue](#) topic in the developer guide.

For information about the arguments you can provide to this field when configuring Ray jobs, see [Using job parameters in Ray jobs](#) in the developer guide.

Type: String to string map

Required: No

Description

Description of the job being defined.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

ExecutionClass

Indicates whether the job is run with a standard or flexible execution class. The standard execution-class is ideal for time-sensitive workloads that require fast job startup and dedicated resources.

The flexible execution class is appropriate for time-insensitive jobs whose start and completion times may vary.

Only jobs with AWS Glue version 3.0 and above and command type `glueetl` will be allowed to set `ExecutionClass` to `FLEX`. The flexible execution class is available for Spark jobs.

Type: String

Length Constraints: Maximum length of 16.

Valid Values: `FLEX` | `STANDARD`

Required: No

ExecutionProperty

An `ExecutionProperty` specifying the maximum number of concurrent runs allowed for this job.

Type: [ExecutionProperty](#) object

Required: No

GlueVersion

In Spark jobs, `GlueVersion` determines the versions of Apache Spark and Python that AWS Glue available in a job. The Python version indicates the version supported for jobs of type Spark.

Ray jobs should set `GlueVersion` to `4.0` or greater. However, the versions of Ray, Python and additional libraries available in your Ray job are determined by the `Runtime` parameter of the Job command.

For more information about the available AWS Glue versions and corresponding Spark and Python versions, see [Glue version](#) in the developer guide.

Jobs that are created without specifying a Glue version default to Glue 0.9.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `^\w+\.\w+$`

Required: No

JobMode

A mode that describes how a job was created. Valid values are:

- SCRIPT - The job was created using the AWS Glue Studio script editor.
- VISUAL - The job was created using the AWS Glue Studio visual editor.
- NOTEBOOK - The job was created using an interactive sessions notebook.

When the JobMode field is missing or null, SCRIPT is assigned as the default value.

Type: String

Valid Values: SCRIPT | VISUAL | NOTEBOOK

Required: No

LogUri

This field is reserved for future use.

Type: String

Required: No

MaintenanceWindow

This field specifies a day of the week and hour for a maintenance window for streaming jobs. AWS Glue periodically performs maintenance activities. During these maintenance windows, AWS Glue will need to restart your streaming jobs.

AWS Glue will restart the job within 3 hours of the specified maintenance window. For instance, if you set up the maintenance window for Monday at 10:00AM GMT, your jobs will be restarted between 10:00AM GMT to 1:00PM GMT.

Type: String

Pattern: `^(Sun|Mon|Tue|Wed|Thu|Fri|Sat):([01]?[0-9]|2[0-3])$`

Required: No

MaxCapacity

For Glue version 1.0 or earlier jobs, using the standard worker type, the number of AWS Glue data processing units (DPUs) that can be allocated when this job runs. A DPU is a relative

measure of processing power that consists of 4 vCPUs of compute capacity and 16 GB of memory. For more information, see the [AWS Glue pricing page](#).

For Glue version 2.0+ jobs, you cannot specify a `MaximumCapacity`. Instead, you should specify a `WorkerType` and the `NumberOfWorkers`.

Do not set `MaxCapacity` if using `WorkerType` and `NumberOfWorkers`.

The value that can be allocated for `MaxCapacity` depends on whether you are running a Python shell job, an Apache Spark ETL job, or an Apache Spark streaming ETL job:

- When you specify a Python shell job (`JobCommand.Name="pythonshell"`), you can allocate either 0.0625 or 1 DPU. The default is 0.0625 DPU.
- When you specify an Apache Spark ETL job (`JobCommand.Name="glueetl"`) or Apache Spark streaming ETL job (`JobCommand.Name="gluestreaming"`), you can allocate from 2 to 100 DPUs. The default is 10 DPUs. This job type cannot have a fractional DPU allocation.

Type: Double

Required: No

MaxRetries

The maximum number of times to retry this job if it fails.

Type: Integer

Required: No

NonOverridableArguments

Arguments for this job that are not overridden when providing job arguments in a job run, specified as name-value pairs.

Type: String to string map

Required: No

NotificationProperty

Specifies the configuration properties of a job notification.

Type: [NotificationProperty](#) object

Required: No

NumberOfWorkers

The number of workers of a defined `workerType` that are allocated when a job runs.

Type: Integer

Required: No

Role

The name or Amazon Resource Name (ARN) of the IAM role associated with this job (required).

Type: String

Required: No

SecurityConfiguration

The name of the `SecurityConfiguration` structure to be used with this job.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

SourceControlDetails

The details for a source control configuration for a job, allowing synchronization of job artifacts to or from a remote repository.

Type: [SourceControlDetails](#) object

Required: No

Timeout

The job timeout in minutes. This is the maximum time that a job run can consume resources before it is terminated and enters `TIMEOUT` status. The default is 2,880 minutes (48 hours) for batch jobs.

Streaming jobs must have timeout values less than 7 days or 10080 minutes. When the value is left blank, the job will be restarted after 7 days based if you have not setup a maintenance window. If you have setup maintenance window, it will be restarted during the maintenance window after 7 days.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

WorkerType

The type of predefined worker that is allocated when a job runs. Accepts a value of G.1X, G.2X, G.4X, G.8X or G.025X for Spark jobs. Accepts the value Z.2X for Ray jobs.

- For the G.1X worker type, each worker maps to 1 DPU (4 vCPUs, 16 GB of memory) with 84GB disk (approximately 34GB free), and provides 1 executor per worker. We recommend this worker type for workloads such as data transforms, joins, and queries, to offers a scalable and cost effective way to run most jobs.
- For the G.2X worker type, each worker maps to 2 DPU (8 vCPUs, 32 GB of memory) with 128GB disk (approximately 77GB free), and provides 1 executor per worker. We recommend this worker type for workloads such as data transforms, joins, and queries, to offers a scalable and cost effective way to run most jobs.
- For the G.4X worker type, each worker maps to 4 DPU (16 vCPUs, 64 GB of memory) with 256GB disk (approximately 235GB free), and provides 1 executor per worker. We recommend this worker type for jobs whose workloads contain your most demanding transforms, aggregations, joins, and queries. This worker type is available only for AWS Glue version 3.0 or later Spark ETL jobs in the following AWS Regions: US East (Ohio), US East (N. Virginia), US West (Oregon), Asia Pacific (Singapore), Asia Pacific (Sydney), Asia Pacific (Tokyo), Canada (Central), Europe (Frankfurt), Europe (Ireland), and Europe (Stockholm).
- For the G.8X worker type, each worker maps to 8 DPU (32 vCPUs, 128 GB of memory) with 512GB disk (approximately 487GB free), and provides 1 executor per worker. We recommend this worker type for jobs whose workloads contain your most demanding transforms, aggregations, joins, and queries. This worker type is available only for AWS Glue version 3.0 or later Spark ETL jobs, in the same AWS Regions as supported for the G.4X worker type.
- For the G.025X worker type, each worker maps to 0.25 DPU (2 vCPUs, 4 GB of memory) with 84GB disk (approximately 34GB free), and provides 1 executor per worker. We recommend this worker type for low volume streaming jobs. This worker type is only available for AWS Glue version 3.0 streaming jobs.
- For the Z.2X worker type, each worker maps to 2 M-DPU (8vCPUs, 64 GB of memory) with 128 GB disk (approximately 120GB free), and provides up to 8 Ray workers based on the autoscaler.

Type: String

Valid Values: Standard | G.1X | G.2X | G.025X | G.4X | G.8X | Z.2X

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Join

Specifies a transform that joins two datasets into one dataset using a comparison phrase on the specified data property keys. You can use inner, outer, left, right, left semi, and left anti joins.

Contents

Columns

A list of the two columns to be joined.

Type: Array of [JoinColumn](#) objects

Array Members: Fixed number of 2 items.

Required: Yes

Inputs

The data inputs identified by their node names.

Type: Array of strings

Array Members: Fixed number of 2 items.

Pattern: `[A-Za-z0-9_-]*`

Required: Yes

JoinType

Specifies the type of join to be performed on the datasets.

Type: String

Valid Values: `equi``join` | `left` | `right` | `outer` | `leftsemi` | `leftanti`

Required: Yes

Name

The name of the transform node.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]| [^\\r\\n]`)*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

JoinColumn

Specifies a column to be joined.

Contents

From

The column to be joined.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF]|[\^\S\r\n"'])*`)

Required: Yes

Keys

The key of the column to be joined.

Type: Array of arrays of strings

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF]|[\^\S\r\n"'])*`)

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

JsonClassifier

A classifier for JSON content.

Contents

JsonPath

A JsonPath string defining the JSON data for the classifier to classify. AWS Glue supports a subset of JsonPath, as described in [Writing JsonPath Custom Classifiers](#).

Type: String

Required: Yes

Name

The name of the classifier.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

CreationTime

The time that this classifier was registered.

Type: Timestamp

Required: No

LastUpdated

The time that this classifier was last updated.

Type: Timestamp

Required: No

Version

The version of this classifier.

Type: Long

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

KafkaStreamingSourceOptions

Additional options for streaming.

Contents

AddRecordTimestamp

When this option is set to 'true', the data output will contain an additional column named "__src_timestamp" that indicates the time when the corresponding record received by the topic. The default value is 'false'. This option is supported in AWS Glue version 4.0 or later.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

Assign

The specific TopicPartitions to consume. You must specify at least one of "topicName", "assign" or "subscribePattern".

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

BootstrapServers

A list of bootstrap server URLs, for example, as `b-1.vpc-test-2.o4q88o.c6.kafka.us-east-1.amazonaws.com:9094`. This option must be specified in the API call or defined in the table metadata in the Data Catalog.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

Classification

An optional classification.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

ConnectionName

The name of the connection.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

Delimiter

Specifies the delimiter character.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

EmitConsumerLagMetrics

When this option is set to 'true', for each batch, it will emit the metrics for the duration between the oldest record received by the topic and the time it arrives in AWS Glue to CloudWatch. The metric's name is "glue.driver.streaming.maxConsumerLagInMs". The default value is 'false'. This option is supported in AWS Glue version 4.0 or later.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

EndingOffsets

The end point when a batch query is ended. Possible values are either "latest" or a JSON string that specifies an ending offset for each TopicPartition.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])*`)

Required: No

IncludeHeaders

Whether to include the Kafka headers. When the option is set to "true", the data output will contain an additional column named "glue_streaming_kafka_headers" with type `Array[Struct(key: String, value: String)]`. The default value is "false". This option is available in AWS Glue version 3.0 or later only.

Type: Boolean

Required: No

MaxOffsetsPerTrigger

The rate limit on the maximum number of offsets that are processed per trigger interval. The specified total number of offsets is proportionally split across topicPartitions of different volumes. The default value is null, which means that the consumer reads all offsets until the known latest offset.

Type: Long

Valid Range: Minimum value of 0.

Required: No

MinPartitions

The desired minimum number of partitions to read from Kafka. The default value is null, which means that the number of spark partitions is equal to the number of Kafka partitions.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

NumRetries

The number of times to retry before failing to fetch Kafka offsets. The default value is 3.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

PollTimeoutMs

The timeout in milliseconds to poll data from Kafka in Spark job executors. The default value is 512.

Type: Long

Valid Range: Minimum value of 0.

Required: No

RetryIntervalMs

The time in milliseconds to wait before retrying to fetch Kafka offsets. The default value is 10.

Type: Long

Valid Range: Minimum value of 0.

Required: No

SecurityProtocol

The protocol used to communicate with brokers. The possible values are "SSL" or "PLAINTEXT".

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

StartingOffsets

The starting position in the Kafka topic to read data from. The possible values are "earliest" or "latest". The default value is "latest".

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n'']`)*

Required: No

StartingTimestamp

The timestamp of the record in the Kafka topic to start reading data from. The possible values are a timestamp string in UTC format of the pattern `yyyy-mm-ddTHH:MM:SSZ` (where Z represents a UTC timezone offset with a +/-). For example: "2023-04-04T08:00:00+08:00".

Only one of StartingTimestamp or StartingOffsets must be set.

Type: Timestamp

Required: No

SubscribePattern

A Java regex string that identifies the topic list to subscribe to. You must specify at least one of "topicName", "assign" or "subscribePattern".

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n'']`)*

Required: No

TopicName

The topic name as specified in Apache Kafka. You must specify at least one of "topicName", "assign" or "subscribePattern".

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n'']`)*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

KeySchemaElement

A partition key pair consisting of a name and a type.

Contents

Name

The name of a partition key.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Type

The type of a partition key.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 131072.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

KinesisStreamingSourceOptions

Additional options for the Amazon Kinesis streaming data source.

Contents

AddIdleTimeBetweenReads

Adds a time delay between two consecutive `getRecords` operations. The default value is `"False"`. This option is only configurable for Glue version 2.0 and above.

Type: Boolean

Required: No

AddRecordTimestamp

When this option is set to `'true'`, the data output will contain an additional column named `"__src_timestamp"` that indicates the time when the corresponding record received by the stream. The default value is `'false'`. This option is supported in AWS Glue version 4.0 or later.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

AvoidEmptyBatches

Avoids creating an empty microbatch job by checking for unread data in the Kinesis data stream before the batch is started. The default value is `"False"`.

Type: Boolean

Required: No

Classification

An optional classification.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

Delimiter

Specifies the delimiter character.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

DescribeShardInterval

The minimum time interval between two ListShards API calls for your script to consider resharding. The default value is 1s.

Type: Long

Valid Range: Minimum value of 0.

Required: No

EmitConsumerLagMetrics

When this option is set to 'true', for each batch, it will emit the metrics for the duration between the oldest record received by the stream and the time it arrives in AWS Glue to CloudWatch. The metric's name is "glue.driver.streaming.maxConsumerLagInMs". The default value is 'false'. This option is supported in AWS Glue version 4.0 or later.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

EndpointUrl

The URL of the Kinesis endpoint.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

IdleTimeBetweenReadsInMs

The minimum time delay between two consecutive `getRecords` operations, specified in ms. The default value is `1000`. This option is only configurable for Glue version 2.0 and above.

Type: Long

Valid Range: Minimum value of 0.

Required: No

MaxFetchRecordsPerShard

The maximum number of records to fetch per shard in the Kinesis data stream per microbatch. Note: The client can exceed this limit if the streaming job has already read extra records from Kinesis (in the same `get-records` call). If `MaxFetchRecordsPerShard` needs to be strict then it needs to be a multiple of `MaxRecordPerRead`. The default value is `100000`.

Type: Long

Valid Range: Minimum value of 0.

Required: No

MaxFetchTimeInMs

The maximum time spent for the job executor to read records for the current batch from the Kinesis data stream, specified in milliseconds (ms). Multiple `GetRecords` API calls may be made within this time. The default value is `1000`.

Type: Long

Valid Range: Minimum value of 0.

Required: No

MaxRecordPerRead

The maximum number of records to fetch from the Kinesis data stream in each `getRecords` operation. The default value is `10000`.

Type: Long

Valid Range: Minimum value of 0.

Required: No

MaxRetryIntervalMs

The maximum cool-off time period (specified in ms) between two retries of a Kinesis Data Streams API call. The default value is 10000.

Type: Long

Valid Range: Minimum value of 0.

Required: No

NumRetries

The maximum number of retries for Kinesis Data Streams API requests. The default value is 3.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

RetryIntervalMs

The cool-off time period (specified in ms) before retrying the Kinesis Data Streams API call. The default value is 1000.

Type: Long

Valid Range: Minimum value of 0.

Required: No

RoleArn

The Amazon Resource Name (ARN) of the role to assume using AWS Security Token Service (AWS STS). This role must have permissions for describe or read record operations for the Kinesis data stream. You must use this parameter when accessing a data stream in a different account. Used in conjunction with "awsSTSSessionName".

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

RoleSessionName

An identifier for the session assuming the role using AWS STS. You must use this parameter when accessing a data stream in a different account. Used in conjunction with "awsSTSRoleARN".

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

StartingPosition

The starting position in the Kinesis data stream to read data from. The possible values are "latest", "trim_horizon", "earliest", or a timestamp string in UTC format in the pattern yyyy-mm-ddTHH:MM:SSZ (where Z represents a UTC timezone offset with a +/-). For example: "2023-04-04T08:00:00-04:00". The default value is "latest".

Note: Using a value that is a timestamp string in UTC format for "startingPosition" is supported only for AWS Glue version 4.0 or later.

Type: String

Valid Values: latest | trim_horizon | earliest | timestamp

Required: No

StartingTimestamp

The timestamp of the record in the Kinesis data stream to start reading data from. The possible values are a timestamp string in UTC format of the pattern yyyy-mm-ddTHH:MM:SSZ (where Z represents a UTC timezone offset with a +/-). For example: "2023-04-04T08:00:00+08:00".

Type: Timestamp

Required: No

StreamArn

The Amazon Resource Name (ARN) of the Kinesis data stream.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])`)*

Required: No

StreamName

The name of the Kinesis data stream.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])`)*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

LabelingSetGenerationTaskRunProperties

Specifies configuration properties for a labeling set generation task run.

Contents

OutputS3Path

The Amazon Simple Storage Service (Amazon S3) path where you will generate the labeling set.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

LakeFormationConfiguration

Specifies AWS Lake Formation configuration settings for the crawler.

Contents

AccountId

Required for cross account crawls. For same account crawls as the target data, this can be left as null.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 12.

Required: No

UseLakeFormationCredentials

Specifies whether to use AWS Lake Formation credentials for the crawler instead of the IAM role credentials.

Type: Boolean

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

LastActiveDefinition

When there are multiple versions of a blueprint and the latest version has some errors, this attribute indicates the last successful blueprint definition that is available with the service.

Contents

BlueprintLocation

Specifies a path in Amazon S3 where the blueprint is published by the AWS Glue developer.

Type: String

Required: No

BlueprintServiceLocation

Specifies a path in Amazon S3 where the blueprint is copied when you create or update the blueprint.

Type: String

Required: No

Description

The description of the blueprint.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 512.

Required: No

LastModifiedOn

The date and time the blueprint was last modified.

Type: Timestamp

Required: No

ParameterSpec

A JSON string specifying the parameters for the blueprint.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 131072.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

LastCrawlInfo

Status and error information about the most recent crawl.

Contents

ErrorMessage

If an error occurred, the error information about the last crawl.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\x\n\t]*`

Required: No

LogGroup

The log group for the last crawl.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 512.

Pattern: `[\.\-_\/#A-Za-z0-9]+`

Required: No

LogStream

The log stream for the last crawl.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 512.

Pattern: `[\^:]*`

Required: No

MessagePrefix

The prefix for a message about this crawl.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

StartTime

The time at which the crawl started.

Type: Timestamp

Required: No

Status

Status of the last crawl.

Type: String

Valid Values: SUCCEEDED | CANCELLED | FAILED

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

LineageConfiguration

Specifies data lineage configuration settings for the crawler.

Contents

CrawlerLineageSettings

Specifies whether data lineage is enabled for the crawler. Valid values are:

- **ENABLE**: enables data lineage for the crawler
- **DISABLE**: disables data lineage for the crawler

Type: String

Valid Values: ENABLE | DISABLE

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Location

The location of resources.

Contents

DynamoDB

An Amazon DynamoDB table location.

Type: Array of [CodeGenNodeArg](#) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

Jdbc

A JDBC location.

Type: Array of [CodeGenNodeArg](#) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

S3

An Amazon Simple Storage Service (Amazon S3) location.

Type: Array of [CodeGenNodeArg](#) objects

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for Ruby V3](#)

LongColumnStatisticsData

Defines column statistics supported for integer data columns.

Contents

NumberOfDistinctValues

The number of distinct values in a column.

Type: Long

Valid Range: Minimum value of 0.

Required: Yes

NumberOfNulls

The number of null values in the column.

Type: Long

Valid Range: Minimum value of 0.

Required: Yes

MaximumValue

The highest value in the column.

Type: Long

Required: No

MinimumValue

The lowest value in the column.

Type: Long

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Mapping

Specifies the mapping of data property keys.

Contents

Children

Only applicable to nested data structures. If you want to change the parent structure, but also one of its children, you can fill out this data structure. It is also Mapping, but its FromPath will be the parent's FromPath plus the FromPath from this structure.

For the children part, suppose you have the structure:

```
{ "FromPath": "OuterStructure", "ToKey": "OuterStructure", "ToType":  
"Struct", "Dropped": false, "Children": [{ "FromPath": "inner", "ToKey":  
"inner", "ToType": "Double", "Dropped": false, }] }
```

You can specify a Mapping that looks like:

```
{ "FromPath": "OuterStructure", "ToKey": "OuterStructure", "ToType":  
"Struct", "Dropped": false, "Children": [{ "FromPath": "inner", "ToKey":  
"inner", "ToType": "Double", "Dropped": false, }] }
```

Type: Array of [Mapping](#) objects

Required: No

Dropped

If true, then the column is removed.

Type: Boolean

Required: No

FromPath

The table or column to be modified.

Type: Array of strings

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]|[^\\S\\r\\n"']`)*

Required: No

FromType

The type of the data to be modified.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

ToKey

After the apply mapping, what the name of the column should be. Can be the same as `FromPath`.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

ToType

The data type that the data is to be modified to.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for Ruby V3](#)

MappingEntry

Defines a mapping.

Contents

SourcePath

The source path.

Type: String

Required: No

SourceTable

The name of the source table.

Type: String

Required: No

SourceType

The source type.

Type: String

Required: No

TargetPath

The target path.

Type: String

Required: No

TargetTable

The target table.

Type: String

Required: No

TargetType

The target type.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Merge

Specifies a transform that merges a `DynamicFrame` with a staging `DynamicFrame` based on the specified primary keys to identify records. Duplicate records (records with the same primary keys) are not de-duplicated.

Contents

Inputs

The data inputs identified by their node names.

Type: Array of strings

Array Members: Fixed number of 2 items.

Pattern: `[A-Za-z0-9_-]*`

Required: Yes

Name

The name of the transform node.

Type: String

Pattern: `([\u0020-\u007F\u00E0-\u00FF\u0080-\u00FF\u0080-\u00FF\u0080-\u00FF\u0080-\u00FF] | [^\x\n])*`

Required: Yes

PrimaryKeys

The list of primary key fields to match records from the source and staging dynamic frames.

Type: Array of arrays of strings

Pattern: `([\u0020-\u007F\u00E0-\u00FF\u0080-\u00FF\u0080-\u00FF\u0080-\u00FF\u0080-\u00FF] | [^\S\x\n"'])*`

Required: Yes

Source

The source `DynamicFrame` that will be merged with a staging `DynamicFrame`.

Type: String

Pattern: [A-Za-z0-9_-]*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MetadataInfo

A structure containing metadata information for a schema version.

Contents

CreatedTime

The time at which the entry was created.

Type: String

Required: No

MetadataValue

The metadata key's corresponding value.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: [a-zA-Z0-9+-. _./@]+

Required: No

OtherMetadataValueList

Other metadata belonging to the same metadata key.

Type: Array of [OtherMetadataValueListItem](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MetadataKeyValuePair

A structure containing a key value pair for metadata.

Contents

MetadataKey

A metadata key.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: [a-zA-Z0-9+-. _./@]+

Required: No

MetadataValue

A metadata key's corresponding value.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: [a-zA-Z0-9+-. _./@]+

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MetricBasedObservation

Describes the metric based observation generated based on evaluated data quality metrics.

Contents

MetricName

The name of the data quality metric used for generating the observation.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

MetricValues

An object of type `DataQualityMetricValues` representing the analysis of the data quality metric value.

Type: [DataQualityMetricValues](#) object

Required: No

NewRules

A list of new data quality rules generated as part of the observation based on the data quality metric value.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MicrosoftSQLServerCatalogSource

Specifies a Microsoft SQL server data source in the AWS Glue Data Catalog.

Contents

Database

The name of the database to read from.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

Name

The name of the data source.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

Table

The name of the table in the database to read from.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MicrosoftSQLServerCatalogTarget

Specifies a target that uses Microsoft SQL.

Contents

Database

The name of the database to write to.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

Inputs

The nodes that are inputs to the data target.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]`*

Required: Yes

Name

The name of the data target.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

Table

The name of the table in the database to write to.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]|[^\\S\\r\\n"']`)*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MLTransform

A structure for a machine learning transform.

Contents

CreatedOn

A timestamp. The time and date that this machine learning transform was created.

Type: Timestamp

Required: No

Description

A user-defined, long-form description text for the machine learning transform. Descriptions are not guaranteed to be unique and can be changed at any time.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

EvaluationMetrics

An `EvaluationMetrics` object. Evaluation metrics provide an estimate of the quality of your machine learning transform.

Type: [EvaluationMetrics](#) object

Required: No

GlueVersion

This value determines which version of AWS Glue this machine learning transform is compatible with. Glue 1.0 is recommended for most customers. If the value is not set, the Glue compatibility defaults to Glue 0.9. For more information, see [AWS Glue Versions](#) in the developer guide.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `^\w+\.\w+$`

Required: No

InputRecordTables

A list of AWS Glue table definitions used by the transform.

Type: Array of [GlueTable](#) objects

Array Members: Minimum number of 0 items. Maximum number of 10 items.

Required: No

LabelCount

A count identifier for the labeling files generated by AWS Glue for this transform. As you create a better transform, you can iteratively download, label, and upload the labeling file.

Type: Integer

Required: No

LastModifiedOn

A timestamp. The last point in time when this machine learning transform was modified.

Type: Timestamp

Required: No

MaxCapacity

The number of AWS Glue data processing units (DPUs) that are allocated to task runs for this transform. You can allocate from 2 to 100 DPUs; the default is 10. A DPU is a relative measure of processing power that consists of 4 vCPUs of compute capacity and 16 GB of memory. For more information, see the [AWS Glue pricing page](#).

MaxCapacity is a mutually exclusive option with NumberOfWorkers and WorkerType.

- If either NumberOfWorkers or WorkerType is set, then MaxCapacity cannot be set.
- If MaxCapacity is set then neither NumberOfWorkers or WorkerType can be set.
- If WorkerType is set, then NumberOfWorkers is required (and vice versa).
- MaxCapacity and NumberOfWorkers must both be at least 1.

When the `WorkerType` field is set to a value other than `Standard`, the `MaxCapacity` field is set automatically and becomes read-only.

Type: Double

Required: No

MaxRetries

The maximum number of times to retry after an `MLTaskRun` of the machine learning transform fails.

Type: Integer

Required: No

Name

A user-defined name for the machine learning transform. Names are not guaranteed unique and can be changed at any time.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

NumberOfWorkers

The number of workers of a defined `workerType` that are allocated when a task of the transform runs.

If `WorkerType` is set, then `NumberOfWorkers` is required (and vice versa).

Type: Integer

Required: No

Parameters

A `TransformParameters` object. You can use parameters to tune (customize) the behavior of the machine learning transform by specifying what data it learns from and your preference on various tradeoffs (such as precision vs. recall, or accuracy vs. cost).

Type: [TransformParameters](#) object

Required: No

Role

The name or Amazon Resource Name (ARN) of the IAM role with the required permissions. The required permissions include both AWS Glue service role permissions to AWS Glue resources, and Amazon S3 permissions required by the transform.

- This role needs AWS Glue service role permissions to allow access to resources in AWS Glue. See [Attach a Policy to IAM Users That Access AWS Glue](#).
- This role needs permission to your Amazon Simple Storage Service (Amazon S3) sources, targets, temporary directory, scripts, and any libraries used by the task run for this transform.

Type: String

Required: No

Schema

A map of key-value pairs representing the columns and data types that this transform can run against. Has an upper bound of 100 columns.

Type: Array of [SchemaColumn](#) objects

Array Members: Maximum number of 100 items.

Required: No

Status

The current status of the machine learning transform.

Type: String

Valid Values: NOT_READY | READY | DELETING

Required: No

Timeout

The timeout in minutes of the machine learning transform.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

TransformEncryption

The encryption-at-rest settings of the transform that apply to accessing user data. Machine learning transforms can access user data encrypted in Amazon S3 using KMS.

Type: [TransformEncryption](#) object

Required: No

TransformId

The unique transform ID that is generated for the machine learning transform. The ID is guaranteed to be unique and does not change.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

WorkerType

The type of predefined worker that is allocated when a task of this transform runs. Accepts a value of Standard, G.1X, or G.2X.

- For the Standard worker type, each worker provides 4 vCPU, 16 GB of memory and a 50GB disk, and 2 executors per worker.
- For the G.1X worker type, each worker provides 4 vCPU, 16 GB of memory and a 64GB disk, and 1 executor per worker.
- For the G.2X worker type, each worker provides 8 vCPU, 32 GB of memory and a 128GB disk, and 1 executor per worker.

MaxCapacity is a mutually exclusive option with NumberOfWorkers and WorkerType.

- If either NumberOfWorkers or WorkerType is set, then MaxCapacity cannot be set.
- If MaxCapacity is set then neither NumberOfWorkers or WorkerType can be set.
- If WorkerType is set, then NumberOfWorkers is required (and vice versa).

- `MaxCapacity` and `NumberOfWorkers` must both be at least 1.

Type: String

Valid Values: Standard | G.1X | G.2X | G.025X | G.4X | G.8X | Z.2X

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MLUserDataEncryption

The encryption-at-rest settings of the transform that apply to accessing user data.

Contents

MLUserDataEncryptionMode

The encryption mode applied to user data. Valid values are:

- **DISABLED**: encryption is disabled
- **SSEKMS**: use of server-side encryption with AWS Key Management Service (SSE-KMS) for user data stored in Amazon S3.

Type: String

Valid Values: DISABLED | SSE-KMS

Required: Yes

KmsKeyId

The ID for the customer-provided KMS key.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MongoDBTarget

Specifies an Amazon DocumentDB or MongoDB data store to crawl.

Contents

ConnectionName

The name of the connection to use to connect to the Amazon DocumentDB or MongoDB target.

Type: String

Required: No

Path

The path of the Amazon DocumentDB or MongoDB target (database/collection).

Type: String

Required: No

ScanAll

Indicates whether to scan all the records, or to sample rows from the table. Scanning all the records can take a long time when the table is not a high throughput table.

A value of `true` means to scan all records, while a value of `false` means to sample the records. If no value is specified, the value defaults to `true`.

Type: Boolean

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MySQLCatalogSource

Specifies a MySQL data source in the AWS Glue Data Catalog.

Contents

Database

The name of the database to read from.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

Name

The name of the data source.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

Table

The name of the table in the database to read from.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MySQLCatalogTarget

Specifies a target that uses MySQL.

Contents

Database

The name of the database to write to.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

Inputs

The nodes that are inputs to the data target.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]`*

Required: Yes

Name

The name of the data target.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

Table

The name of the table in the database to write to.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]|[^\\S\\r\\n"']`)*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Node

A node represents an AWS Glue component (trigger, crawler, or job) on a workflow graph.

Contents

CrawlerDetails

Details of the crawler when the node represents a crawler.

Type: [CrawlerNodeDetails](#) object

Required: No

JobDetails

Details of the Job when the node represents a Job.

Type: [JobNodeDetails](#) object

Required: No

Name

The name of the AWS Glue component represented by the node.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

TriggerDetails

Details of the Trigger when the node represents a Trigger.

Type: [TriggerNodeDetails](#) object

Required: No

Type

The type of AWS Glue component represented by the node.

Type: String

Valid Values: CRAWLER | JOB | TRIGGER

Required: No

Uniqueld

The unique Id assigned to the node within the workflow.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

NotificationProperty

Specifies configuration properties of a notification.

Contents

NotifyDelayAfter

After a job run starts, the number of minutes to wait before sending a job run delay notification.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

NullCheckBoxList

Represents whether certain values are recognized as null values for removal.

Contents

IsEmpty

Specifies that an empty string is considered as a null value.

Type: Boolean

Required: No

IsNegOne

Specifies that an integer value of -1 is considered as a null value.

Type: Boolean

Required: No

IsNullString

Specifies that a value spelling out the word 'null' is considered as a null value.

Type: Boolean

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

NullValueField

Represents a custom null value such as a zeros or other value being used as a null placeholder unique to the dataset.

Contents

Datatype

The datatype of the value.

Type: [Datatype](#) object

Required: Yes

Value

The value of the null placeholder.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])`)*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

OAuth2ClientApplication

The OAuth2 client app used for the connection.

Contents

AWSManagedClientApplicationReference

The reference to the SaaS-side client app that is AWS managed.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: \S+

Required: No

UserManagedClientApplicationClientId

The client application clientID if the ClientAppType is USER_MANAGED.

Type: String

Length Constraints: Maximum length of 2048.

Pattern: \S+

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

OAuth2Properties

A structure containing properties for OAuth2 authentication.

Contents

OAuth2ClientApplication

The client application type. For example, `AWS_MANAGED` or `USER_MANAGED`.

Type: [OAuth2ClientApplication](#) object

Required: No

OAuth2GrantType

The OAuth2 grant type. For example, `AUTHORIZATION_CODE`, `JWT_BEARER`, or `CLIENT_CREDENTIALS`.

Type: String

Valid Values: `AUTHORIZATION_CODE` | `CLIENT_CREDENTIALS` | `JWT_BEARER`

Required: No

TokenUrl

The URL of the provider's authentication server, to exchange an authorization code for an access token.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `^(https?):\/\/[-a-zA-Z0-9+&@#/%?~_]|!:,.;]*[-a-zA-Z0-9+&@#/%=~_]|`

Required: No

TokenUrlParametersMap

A map of parameters that are added to the token GET request.

Type: String to string map

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 1. Maximum length of 512.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

OAuth2PropertiesInput

A structure containing properties for OAuth2 in the CreateConnection request.

Contents

AuthorizationCodeProperties

The set of properties required for the the OAuth2 AUTHORIZATION_CODE grant type.

Type: [AuthorizationCodeProperties](#) object

Required: No

OAuth2ClientApplication

The client application type in the CreateConnection request. For example, AWS_MANAGED or USER_MANAGED.

Type: [OAuth2ClientApplication](#) object

Required: No

OAuth2GrantType

The OAuth2 grant type in the CreateConnection request. For example, AUTHORIZATION_CODE, JWT_BEARER, or CLIENT_CREDENTIALS.

Type: String

Valid Values: AUTHORIZATION_CODE | CLIENT_CREDENTIALS | JWT_BEARER

Required: No

TokenUrl

The URL of the provider's authentication server, to exchange an authorization code for an access token.

Type: String

Length Constraints: Maximum length of 256.

Pattern: `^(https?):\/\/[-a-zA-Z0-9+&@#/%?~_|!:,.;]*[-a-zA-Z0-9+&@#/%~_|]`

Required: No

TokenUrlParametersMap

A map of parameters that are added to the token GET request.

Type: String to string map

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 1. Maximum length of 512.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

OpenTableFormatInput

A structure representing an open format table.

Contents

IcebergInput

Specifies an IcebergInput structure that defines an Apache Iceberg metadata table.

Type: [IcebergInput](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Option

Specifies an option value.

Contents

Description

Specifies the description of the option.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])*`)

Required: No

Label

Specifies the label of the option.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])*`)

Required: No

Value

Specifies the value of the option.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])*`)

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

OracleSQLCatalogSource

Specifies an Oracle data source in the AWS Glue Data Catalog.

Contents

Database

The name of the database to read from.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

Name

The name of the data source.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

Table

The name of the table in the database to read from.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

OracleSQLCatalogTarget

Specifies a target that uses Oracle SQL.

Contents

Database

The name of the database to write to.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

Inputs

The nodes that are inputs to the data target.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]`*

Required: Yes

Name

The name of the data target.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

Table

The name of the table in the database to write to.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]|[^\\S\\r\\n"']`)*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Order

Specifies the sort order of a sorted column.

Contents

Column

The name of the column.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

SortOrder

Indicates that the column is sorted in ascending order (`= 1`), or in descending order (`= 0`).

Type: Integer

Valid Range: Minimum value of 0. Maximum value of 1.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

OtherMetadataValueListItem

A structure containing other metadata for a schema version belonging to the same metadata key.

Contents

CreatedTime

The time at which the entry was created.

Type: String

Required: No

MetadataValue

The metadata key's corresponding value for the other metadata belonging to the same metadata key.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: [a-zA-Z0-9+-. _./@]+

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Partition

Represents a slice of table data.

Contents

CatalogId

The ID of the Data Catalog in which the partition resides.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

CreationTime

The time at which the partition was created.

Type: Timestamp

Required: No

DatabaseName

The name of the catalog database in which to create the partition.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

LastAccessTime

The last time at which the partition was accessed.

Type: Timestamp

Required: No

LastAnalyzedTime

The last time at which column statistics were computed for this partition.

Type: Timestamp

Required: No

Parameters

These key-value pairs define partition parameters.

Type: String to string map

Key Length Constraints: Minimum length of 1. Maximum length of 255.

Key Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Value Length Constraints: Maximum length of 512000.

Required: No

StorageDescriptor

Provides information about the physical location where the partition is stored.

Type: [StorageDescriptor](#) object

Required: No

TableName

The name of the database table in which to create the partition.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Values

The values of the partition.

Type: Array of strings

Length Constraints: Maximum length of 1024.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PartitionError

Contains information about a partition error.

Contents

ErrorDetail

The details about the partition error.

Type: [ErrorDetail](#) object

Required: No

PartitionValues

The values that define the partition.

Type: Array of strings

Length Constraints: Maximum length of 1024.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PartitionIndex

A structure for a partition index.

Contents

IndexName

The name of the partition index.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Keys

The keys for the partition index.

Type: Array of strings

Array Members: Minimum number of 1 item.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Required: Yes

BackfillErrors

A list of errors that can occur when registering partition indexes for an existing table.

Type: Array of [BackfillError](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PartitionInput

The structure used to create and update a partition.

Contents

LastAccessTime

The last time at which the partition was accessed.

Type: Timestamp

Required: No

LastAnalyzedTime

The last time at which column statistics were computed for this partition.

Type: Timestamp

Required: No

Parameters

These key-value pairs define partition parameters.

Type: String to string map

Key Length Constraints: Minimum length of 1. Maximum length of 255.

Key Pattern: `[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF\\t]*`

Value Length Constraints: Maximum length of 512000.

Required: No

StorageDescriptor

Provides information about the physical location where the partition is stored.

Type: [StorageDescriptor](#) object

Required: No

Values

The values of the partition. Although this parameter is not required by the SDK, you must specify this parameter for a valid input.

The values for the keys for the new partition must be passed as an array of String objects that must be ordered in the same order as the partition keys appearing in the Amazon S3 prefix. Otherwise AWS Glue will add the values to the wrong keys.

Type: Array of strings

Length Constraints: Maximum length of 1024.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PartitionValueList

Contains a list of values defining partitions.

Contents

Values

The list of values.

Type: Array of strings

Length Constraints: Maximum length of 1024.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PhysicalConnectionRequirements

The OAuth client app in GetConnection response.

Contents

AvailabilityZone

The connection's Availability Zone.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

SecurityGroupIdList

The security group ID list used by the connection.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 50 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

SubnetId

The subnet ID used by the connection.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PIIDetection

Specifies a transform that identifies, removes or masks PII data.

Contents

EntityTypesToDetect

Indicates the types of entities the PIIDetection transform will identify as PII data.

PII type entities include: PERSON_NAME, DATE, USA_SNN, EMAIL, USA_ITIN, USA_PASSPORT_NUMBER, PHONE_NUMBER, BANK_ACCOUNT, IP_ADDRESS, MAC_ADDRESS, USA_CPT_CODE, USA_HCPCS_CODE, USA_NATIONAL_DRUG_CODE, USA_MEDICARE_BENEFICIARY_IDENTIFIER, USA_HEALTH_INSURANCE_CLAIM_NUMBER, CREDIT_CARD, USA_NATIONAL_PROVIDER_IDENTIFIER, USA

Type: Array of strings

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

Inputs

The node ID inputs to the transform.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]`*

Required: Yes

Name

The name of the transform node.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

PiiType

Indicates the type of PII Detection transform.

Type: String

Valid Values: RowAudit | RowMasking | ColumnAudit | ColumnMasking

Required: Yes

MaskValue

Indicates the value that will replace the detected entity.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 256.

Pattern: [*A-Za-z0-9_-]*

Required: No

OutputColumnName

Indicates the output column name that will contain any entity type detected in that row.

Type: String

Pattern: ([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])*

Required: No

SampleFraction

Indicates the fraction of the data to sample when scanning for PII entities.

Type: Double

Valid Range: Minimum value of 0. Maximum value of 1.

Required: No

ThresholdFraction

Indicates the fraction of the data that must be met in order for a column to be identified as PII data.

Type: Double

Valid Range: Minimum value of 0. Maximum value of 1.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PostgreSQLCatalogSource

Specifies a PostgreSQL data source in the AWS Glue Data Catalog.

Contents

Database

The name of the database to read from.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

Name

The name of the data source.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

Table

The name of the table in the database to read from.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PostgreSQLCatalogTarget

Specifies a target that uses Postgres SQL.

Contents

Database

The name of the database to write to.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

Inputs

The nodes that are inputs to the data target.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]`*

Required: Yes

Name

The name of the data target.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

Table

The name of the table in the database to write to.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]|[^\\S\\r\\n"']`)*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Predecessor

A job run that was used in the predicate of a conditional trigger that triggered this job run.

Contents

JobName

The name of the job definition used by the predecessor job run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

RunId

The job-run ID of the predecessor job run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Predicate

Defines the predicate of the trigger, which determines when it fires.

Contents

Conditions

A list of the conditions that determine when the trigger will fire.

Type: Array of [Condition](#) objects

Required: No

Logical

An optional field if only one condition is listed. If multiple conditions are listed, then this field is required.

Type: String

Valid Values: AND | ANY

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PrincipalPermissions

Permissions granted to a principal.

Contents

Permissions

The permissions that are granted to the principal.

Type: Array of strings

Valid Values: ALL | SELECT | ALTER | DROP | DELETE | INSERT |
CREATE_DATABASE | CREATE_TABLE | DATA_LOCATION_ACCESS

Required: No

Principal

The principal who is granted permissions.

Type: [DataLakePrincipal](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ProfileConfiguration

Specifies the job and session values that an admin configures in an AWS Glue usage profile.

Contents

JobConfiguration

A key-value map of configuration parameters for AWS Glue jobs.

Type: String to [ConfigurationObject](#) object map

Key Length Constraints: Minimum length of 1. Maximum length of 255.

Key Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

SessionConfiguration

A key-value map of configuration parameters for AWS Glue sessions.

Type: String to [ConfigurationObject](#) object map

Key Length Constraints: Minimum length of 1. Maximum length of 255.

Key Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PropertyPredicate

Defines a property predicate.

Contents

Comparator

The comparator used to compare this property to others.

Type: String

Valid Values: EQUALS | GREATER_THAN | LESS_THAN | GREATER_THAN_EQUALS | LESS_THAN_EQUALS

Required: No

Key

The key of the property.

Type: String

Length Constraints: Maximum length of 1024.

Required: No

Value

The value of the property.

Type: String

Length Constraints: Maximum length of 1024.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)

- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

QuerySessionContext

A structure used as a protocol between query engines and Lake Formation or AWS Glue. Contains both a Lake Formation generated authorization identifier and information from the request's authorization context.

Contents

AdditionalContext

An opaque string-string map passed by the query engine.

Type: String to string map

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Required: No

ClusterId

An identifier string for the consumer cluster.

Type: String

Required: No

QueryAuthorizationId

A cryptographically generated query identifier generated by AWS Glue or Lake Formation.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

QueryId

A unique identifier generated by the query engine for the query.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

QueryStartTime

A timestamp provided by the query engine for when the query started.

Type: Timestamp

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Recipe

A AWS Glue Studio node that uses a AWS Glue DataBrew recipe in AWS Glue jobs.

Contents

Inputs

The nodes that are inputs to the recipe node, identified by id.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]*`

Required: Yes

Name

The name of the AWS Glue Studio node.

Type: String

Pattern: `([\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF]|[\^\r\n])*`

Required: Yes

RecipeReference

A reference to the DataBrew recipe used by the node.

Type: [RecipeReference](#) object

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for Ruby V3](#)

RecipeReference

A reference to a AWS Glue DataBrew recipe.

Contents

RecipeArn

The ARN of the DataBrew recipe.

Type: String

Pattern: (`[\u0020-\u007F\u00E0\u0000-\u00FF\u0080\u00DC\u0000-\u00DB\u00FF\u00DF\u00FF] | [^\S\r\n"']`)*

Required: Yes

RecipeVersion

The RecipeVersion of the DataBrew recipe.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 16.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

RecrawlPolicy

When crawling an Amazon S3 data source after the first crawl is complete, specifies whether to crawl the entire dataset again or to crawl only folders that were added since the last crawler run. For more information, see [Incremental Crawls in AWS Glue](#) in the developer guide.

Contents

RecrawlBehavior

Specifies whether to crawl the entire dataset again or to crawl only folders that were added since the last crawler run.

A value of `CRAWL_EVERYTHING` specifies crawling the entire dataset again.

A value of `CRAWL_NEW_FOLDERS_ONLY` specifies crawling only folders that were added since the last crawler run.

A value of `CRAWL_EVENT_MODE` specifies crawling only the changes identified by Amazon S3 events.

Type: String

Valid Values: `CRAWL_EVERYTHING` | `CRAWL_NEW_FOLDERS_ONLY` | `CRAWL_EVENT_MODE`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

RedshiftSource

Specifies an Amazon Redshift data store.

Contents

Database

The database to read from.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

Name

The name of the Amazon Redshift data store.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

Table

The database table to read from.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

RedshiftTmpDir

The Amazon S3 path where temporary data can be staged when copying out of the database.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF]|[\^\S\r\n"'])*`)^{*}

Required: No

TmpDirIAMRole

The IAM role with permissions.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF]|[\^\S\r\n"'])*`)^{*}

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

RedshiftTarget

Specifies a target that uses Amazon Redshift.

Contents

Database

The name of the database to write to.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]|[^\\S\\r\\n"']`)*

Required: Yes

Inputs

The nodes that are inputs to the data target.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]`*

Required: Yes

Name

The name of the data target.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]|[^\\r\\n]`)*

Required: Yes

Table

The name of the table in the database to write to.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF]|[\^S\r\n"'])`)*

Required: Yes

RedshiftTmpDir

The Amazon S3 path where temporary data can be staged when copying out of the database.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF]|[\^S\r\n"'])`)*

Required: No

TmpDirIAMRole

The IAM role with permissions.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF]|[\^S\r\n"'])`)*

Required: No

UpsertRedshiftOptions

The set of options to configure an upsert operation when writing to a Redshift target.

Type: [UpsertRedshiftTargetOptions](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

RegistryId

A wrapper structure that may contain the registry name and Amazon Resource Name (ARN).

Contents

RegistryArn

Arn of the registry to be updated. One of RegistryArn or RegistryName has to be provided.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 10240.

Pattern: `arn:(aws|aws-us-gov|aws-cn):glue:.*`

Required: No

RegistryName

Name of the registry. Used only for lookup. One of RegistryArn or RegistryName has to be provided.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[a-zA-Z0-9-_$#.]+`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

RegistryListItem

A structure containing the details for a registry.

Contents

CreatedTime

The data the registry was created.

Type: String

Required: No

Description

A description of the registry.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

RegistryArn

The Amazon Resource Name (ARN) of the registry.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 10240.

Pattern: `arn:(aws|aws-us-gov|aws-cn):glue:.*`

Required: No

RegistryName

The name of the registry.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: [a-zA-Z0-9-_\$#.]+

Required: No

Status

The status of the registry.

Type: String

Valid Values: AVAILABLE | DELETING

Required: No

UpdateTime

The date the registry was updated.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

RelationalCatalogSource

Specifies a Relational database data source in the AWS Glue Data Catalog.

Contents

Database

The name of the database to read from.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

Name

The name of the data source.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

Table

The name of the table in the database to read from.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

RenameField

Specifies a transform that renames a single data property key.

Contents

Inputs

The data inputs identified by their node names.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]*`

Required: Yes

Name

The name of the transform node.

Type: String

Pattern: `([\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\r\n])*`

Required: Yes

SourcePath

A JSON path to a variable in the data structure for the source data.

Type: Array of strings

Pattern: `([\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])*`

Required: Yes

TargetPath

A JSON path to a variable in the data structure for the target data.

Type: Array of strings

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]|[^\\S\\r\\n"']`)*

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ResourceUri

The URIs for function resources.

Contents

ResourceType

The type of the resource.

Type: String

Valid Values: JAR | FILE | ARCHIVE

Required: No

Uri

The URI for accessing the resource.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

RunMetrics

Metrics for the optimizer run.

Contents

JobDurationInHour

The duration of the job in hours.

Type: String

Required: No

NumberOfBytesCompacted

The number of bytes removed by the compaction job run.

Type: String

Required: No

NumberOfDpus

The number of DPU hours consumed by the job.

Type: String

Required: No

NumberOfFilesCompacted

The number of files removed by the compaction job run.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)

- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

S3CatalogDeltaSource

Specifies a Delta Lake data source that is registered in the AWS Glue Data Catalog. The data source must be stored in Amazon S3.

Contents

Database

The name of the database to read from.

Type: String

Pattern: (`([\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])*`)

Required: Yes

Name

The name of the Delta Lake data source.

Type: String

Pattern: (`([\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\r\n])*`)

Required: Yes

Table

The name of the table in the database to read from.

Type: String

Pattern: (`([\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])*`)

Required: Yes

AdditionalDeltaOptions

Specifies additional connection options.

Type: String to string map

Key Pattern: (`([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])*`)

Value Pattern: (`([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])*`)

Required: No

OutputSchemas

Specifies the data schema for the Delta Lake source.

Type: Array of [GlueSchema](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

S3CatalogHudiSource

Specifies a Hudi data source that is registered in the AWS Glue Data Catalog. The Hudi data source must be stored in Amazon S3.

Contents

Database

The name of the database to read from.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

Name

The name of the Hudi data source.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

Table

The name of the table in the database to read from.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

AdditionalHudiOptions

Specifies additional connection options.

Type: String to string map

Key Pattern: (`([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF]|[\^\S\r\n"'])*`)

Value Pattern: (`([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF]|[\^\S\r\n"'])*`)

Required: No

OutputSchemas

Specifies the data schema for the Hudi source.

Type: Array of [GlueSchema](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PartitionPredicate

Partitions satisfying this predicate are deleted. Files within the retention period in these partitions are not deleted. Set to "" – empty by default.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

S3CatalogTarget

Specifies a data target that writes to Amazon S3 using the AWS Glue Data Catalog.

Contents

Database

The name of the database to write to.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

Inputs

The nodes that are inputs to the data target.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]`*

Required: Yes

Name

The name of the data target.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

Table

The name of the table in the database to write to.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF]|[\^\S\r\n"'])`)*

Required: Yes

PartitionKeys

Specifies native partitioning using a sequence of keys.

Type: Array of arrays of strings

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF]|[\^\S\r\n"'])`)*

Required: No

SchemaChangePolicy

A policy that specifies update behavior for the crawler.

Type: [CatalogSchemaChangePolicy](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

S3CsvSource

Specifies a command-separated value (CSV) data store stored in Amazon S3.

Contents

Name

The name of the data store.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

Paths

A list of the Amazon S3 paths to read from.

Type: Array of strings

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

QuoteChar

Specifies the character to use for quoting. The default is a double quote: `'"`. Set this to `-1` to turn off quoting entirely.

Type: String

Valid Values: `quote` | `quilletmet` | `single_quote` | `disabled`

Required: Yes

Separator

Specifies the delimiter character. The default is a comma: `","`, but any other character can be specified.

Type: String

Valid Values: comma | ctrl-a | pipe | semicolon | tab

Required: Yes

AdditionalOptions

Specifies additional connection options.

Type: [S3DirectSourceAdditionalOptions](#) object

Required: No

CompressionType

Specifies how the data is compressed. This is generally not necessary if the data has a standard file extension. Possible values are "gzip" and "bzip").

Type: String

Valid Values: gzip | bzip2

Required: No

Escaper

Specifies a character to use for escaping. This option is used only when reading CSV files. The default value is none. If enabled, the character which immediately follows is used as-is, except for a small set of well-known escapes (`\n`, `\r`, `\t`, and `\0`).

Type: String

Pattern: (`[\u0020-\u007F\u00E0-\u00FF\u0080-\u00FF\u0080-\u00FF\u0080-\u00FF\u0080-\u00FF\u0080-\u00FF]` | `[\S\r\n]`)*

Required: No

Exclusions

A string containing a JSON list of Unix-style glob patterns to exclude. For example, `["**/*.pdf"]` excludes all PDF files.

Type: Array of strings

Pattern: (`[\u0020-\u007F\u00E0-\u00FF\u0080-\u00FF\u0080-\u00FF\u0080-\u00FF\u0080-\u00FF\u0080-\u00FF]` | `[\S\r\n"']`)*

Required: No

GroupFiles

Grouping files is turned on by default when the input contains more than 50,000 files. To turn on grouping with fewer than 50,000 files, set this parameter to "inPartition". To disable grouping when there are more than 50,000 files, set this parameter to "none".

Type: String

Pattern: ([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF]|[\^\S\r\n"'])*

Required: No

GroupSize

The target group size in bytes. The default is computed based on the input data size and the size of your cluster. When there are fewer than 50,000 input files, "groupFiles" must be set to "inPartition" for this to take effect.

Type: String

Pattern: ([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF]|[\^\S\r\n"'])*

Required: No

MaxBand

This option controls the duration in milliseconds after which the s3 listing is likely to be consistent. Files with modification timestamps falling within the last maxBand milliseconds are tracked specially when using JobBookmarks to account for Amazon S3 eventual consistency. Most users don't need to set this option. The default is 900000 milliseconds, or 15 minutes.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

MaxFilesInBand

This option specifies the maximum number of files to save from the last maxBand seconds. If this number is exceeded, extra files are skipped and only processed in the next job run.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

Multiline

A Boolean value that specifies whether a single record can span multiple lines. This can occur when a field contains a quoted new-line character. You must set this option to True if any record spans multiple lines. The default value is `False`, which allows for more aggressive file-splitting during parsing.

Type: Boolean

Required: No

OptimizePerformance

A Boolean value that specifies whether to use the advanced SIMD CSV reader along with Apache Arrow based columnar memory formats. Only available in AWS Glue version 3.0.

Type: Boolean

Required: No

OutputSchemas

Specifies the data schema for the S3 CSV source.

Type: Array of [GlueSchema](#) objects

Required: No

Recurse

If set to true, recursively reads files in all subdirectories under the specified paths.

Type: Boolean

Required: No

SkipFirst

A Boolean value that specifies whether to skip the first data line. The default value is `False`.

Type: Boolean

Required: No

WithHeader

A Boolean value that specifies whether to treat the first line as a header. The default value is `False`.

Type: Boolean

Required: No

WriteHeader

A Boolean value that specifies whether to write the header to output. The default value is `True`.

Type: Boolean

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

S3DeltaCatalogTarget

Specifies a target that writes to a Delta Lake data source in the AWS Glue Data Catalog.

Contents

Database

The name of the database to write to.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

Inputs

The nodes that are inputs to the data target.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]`*

Required: Yes

Name

The name of the data target.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

Table

The name of the table in the database to write to.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF]|[\^\\S\\r\\n"'])*`)^{*}

Required: Yes

AdditionalOptions

Specifies additional connection options for the connector.

Type: String to string map

Key Pattern: (`([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF]|[\^\\S\\r\\n"'])*`)^{*}

Value Pattern: (`([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF]|[\^\\S\\r\\n"'])*`)^{*}

Required: No

PartitionKeys

Specifies native partitioning using a sequence of keys.

Type: Array of arrays of strings

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF]|[\^\\S\\r\\n"'])*`)^{*}

Required: No

SchemaChangePolicy

A policy that specifies update behavior for the crawler.

Type: [CatalogSchemaChangePolicy](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)

- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

S3DeltaDirectTarget

Specifies a target that writes to a Delta Lake data source in Amazon S3.

Contents

Compression

Specifies how the data is compressed. This is generally not necessary if the data has a standard file extension. Possible values are "gzip" and "bzip").

Type: String

Valid Values: uncompressed | snappy

Required: Yes

Format

Specifies the data output format for the target.

Type: String

Valid Values: json | csv | avro | orc | parquet | hudi | delta

Required: Yes

Inputs

The nodes that are inputs to the data target.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: [A-Za-z0-9_-]*

Required: Yes

Name

The name of the data target.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

Path

The Amazon S3 path of your Delta Lake data source to write to.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

AdditionalOptions

Specifies additional connection options for the connector.

Type: String to string map

Key Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Value Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

PartitionKeys

Specifies native partitioning using a sequence of keys.

Type: Array of arrays of strings

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

SchemaChangePolicy

A policy that specifies update behavior for the crawler.

Type: [DirectSchemaChangePolicy](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

S3DeltaSource

Specifies a Delta Lake data source stored in Amazon S3.

Contents

Name

The name of the Delta Lake source.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

Paths

A list of the Amazon S3 paths to read from.

Type: Array of strings

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

AdditionalDeltaOptions

Specifies additional connection options.

Type: String to string map

Key Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Value Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

AdditionalOptions

Specifies additional options for the connector.

Type: [S3DirectSourceAdditionalOptions](#) object

Required: No

OutputSchemas

Specifies the data schema for the Delta Lake source.

Type: Array of [GlueSchema](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

S3DirectSourceAdditionalOptions

Specifies additional connection options for the Amazon S3 data store.

Contents

BoundedFiles

Sets the upper limit for the target number of files that will be processed.

Type: Long

Required: No

BoundedSize

Sets the upper limit for the target size of the dataset in bytes that will be processed.

Type: Long

Required: No

EnableSamplePath

Sets option to enable a sample path.

Type: Boolean

Required: No

SamplePath

If enabled, specifies the sample path.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

S3DirectTarget

Specifies a data target that writes to Amazon S3.

Contents

Format

Specifies the data output format for the target.

Type: String

Valid Values: json | csv | avro | orc | parquet | hudi | delta

Required: Yes

Inputs

The nodes that are inputs to the data target.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: [A-Za-z0-9_-]*

Required: Yes

Name

The name of the data target.

Type: String

Pattern: ([\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF]|[\^\r\n])*

Required: Yes

Path

A single Amazon S3 path to write to.

Type: String

Pattern: ([\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF]|[\^\S\r\n"'])*

Required: Yes

Compression

Specifies how the data is compressed. This is generally not necessary if the data has a standard file extension. Possible values are "gzip" and "bzip").

Type: String

Pattern: ([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"']) *

Required: No

PartitionKeys

Specifies native partitioning using a sequence of keys.

Type: Array of arrays of strings

Pattern: ([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"']) *

Required: No

SchemaChangePolicy

A policy that specifies update behavior for the crawler.

Type: [DirectSchemaChangePolicy](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

S3Encryption

Specifies how Amazon Simple Storage Service (Amazon S3) data should be encrypted.

Contents

KmsKeyArn

The Amazon Resource Name (ARN) of the KMS key to be used to encrypt the data.

Type: String

Pattern: `arn:aws:kms:.*`

Required: No

S3EncryptionMode

The encryption mode to use for Amazon S3 data.

Type: String

Valid Values: `DISABLED` | `SSE-KMS` | `SSE-S3`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

S3GlueParquetTarget

Specifies a data target that writes to Amazon S3 in Apache Parquet columnar storage.

Contents

Inputs

The nodes that are inputs to the data target.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]*`

Required: Yes

Name

The name of the data target.

Type: String

Pattern: `([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF]|[\^\r\n])*`

Required: Yes

Path

A single Amazon S3 path to write to.

Type: String

Pattern: `([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF]|[\^\S\r\n"'])*`

Required: Yes

Compression

Specifies how the data is compressed. This is generally not necessary if the data has a standard file extension. Possible values are "gzip" and "bzip").

Type: String

S3HudiCatalogTarget

Specifies a target that writes to a Hudi data source in the AWS Glue Data Catalog.

Contents

AdditionalOptions

Specifies additional connection options for the connector.

Type: String to string map

Key Pattern: (`([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n''])*`)

Value Pattern: (`([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n''])*`)

Required: Yes

Database

The name of the database to write to.

Type: String

Pattern: (`([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n''])*`)

Required: Yes

Inputs

The nodes that are inputs to the data target.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]*`

Required: Yes

Name

The name of the data target.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

Table

The name of the table in the database to write to.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n'']`)*

Required: Yes

PartitionKeys

Specifies native partitioning using a sequence of keys.

Type: Array of arrays of strings

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n'']`)*

Required: No

SchemaChangePolicy

A policy that specifies update behavior for the crawler.

Type: [CatalogSchemaChangePolicy](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

S3HudiDirectTarget

Specifies a target that writes to a Hudi data source in Amazon S3.

Contents

AdditionalOptions

Specifies additional connection options for the connector.

Type: String to string map

Key Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n'']`)*

Value Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n'']`)*

Required: Yes

Compression

Specifies how the data is compressed. This is generally not necessary if the data has a standard file extension. Possible values are "gzip" and "bzip").

Type: String

Valid Values: `gzip` | `lzo` | `uncompressed` | `snappy`

Required: Yes

Format

Specifies the data output format for the target.

Type: String

Valid Values: `json` | `csv` | `avro` | `orc` | `parquet` | `hudi` | `delta`

Required: Yes

Inputs

The nodes that are inputs to the data target.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: [A-Za-z0-9_-]*

Required: Yes

Name

The name of the data target.

Type: String

Pattern: ([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF] | [^\r\n])*

Required: Yes

Path

The Amazon S3 path of your Hudi data source to write to.

Type: String

Pattern: ([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])*

Required: Yes

PartitionKeys

Specifies native partitioning using a sequence of keys.

Type: Array of arrays of strings

Pattern: ([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])*

Required: No

SchemaChangePolicy

A policy that specifies update behavior for the crawler.

Type: [DirectSchemaChangePolicy](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

S3HudiSource

Specifies a Hudi data source stored in Amazon S3.

Contents

Name

The name of the Hudi source.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

Paths

A list of the Amazon S3 paths to read from.

Type: Array of strings

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

AdditionalHudiOptions

Specifies additional connection options.

Type: String to string map

Key Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Value Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

AdditionalOptions

Specifies additional options for the connector.

Type: [S3DirectSourceAdditionalOptions](#) object

Required: No

OutputSchemas

Specifies the data schema for the Hudi source.

Type: Array of [GlueSchema](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

S3JsonSource

Specifies a JSON data store stored in Amazon S3.

Contents

Name

The name of the data store.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

Paths

A list of the Amazon S3 paths to read from.

Type: Array of strings

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

AdditionalOptions

Specifies additional connection options.

Type: [S3DirectSourceAdditionalOptions](#) object

Required: No

CompressionType

Specifies how the data is compressed. This is generally not necessary if the data has a standard file extension. Possible values are "gzip" and "bzip").

Type: String

Valid Values: `gzip` | `bzip2`

Required: No

Exclusions

A string containing a JSON list of Unix-style glob patterns to exclude. For example, "[\"**.*pdf\"]" excludes all PDF files.

Type: Array of strings

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

GroupFiles

Grouping files is turned on by default when the input contains more than 50,000 files. To turn on grouping with fewer than 50,000 files, set this parameter to "inPartition". To disable grouping when there are more than 50,000 files, set this parameter to "none".

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

GroupSize

The target group size in bytes. The default is computed based on the input data size and the size of your cluster. When there are fewer than 50,000 input files, "groupFiles" must be set to "inPartition" for this to take effect.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

JsonPath

A JsonPath string defining the JSON data.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]|[^\\S\\r\\n"']`)*

Required: No

MaxBand

This option controls the duration in milliseconds after which the s3 listing is likely to be consistent. Files with modification timestamps falling within the last maxBand milliseconds are tracked specially when using JobBookmarks to account for Amazon S3 eventual consistency. Most users don't need to set this option. The default is 900000 milliseconds, or 15 minutes.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

MaxFilesInBand

This option specifies the maximum number of files to save from the last maxBand seconds. If this number is exceeded, extra files are skipped and only processed in the next job run.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

Multiline

A Boolean value that specifies whether a single record can span multiple lines. This can occur when a field contains a quoted new-line character. You must set this option to True if any record spans multiple lines. The default value is False, which allows for more aggressive file-splitting during parsing.

Type: Boolean

Required: No

OutputSchemas

Specifies the data schema for the S3 JSON source.

Type: Array of [GlueSchema](#) objects

Required: No

Recurse

If set to true, recursively reads files in all subdirectories under the specified paths.

Type: Boolean

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

S3ParquetSource

Specifies an Apache Parquet data store stored in Amazon S3.

Contents

Name

The name of the data store.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

Paths

A list of the Amazon S3 paths to read from.

Type: Array of strings

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

AdditionalOptions

Specifies additional connection options.

Type: [S3DirectSourceAdditionalOptions](#) object

Required: No

CompressionType

Specifies how the data is compressed. This is generally not necessary if the data has a standard file extension. Possible values are "gzip" and "bzip").

Type: String

Valid Values: snappy | lzo | gzip | uncompressed | none

Required: No

Exclusions

A string containing a JSON list of Unix-style glob patterns to exclude. For example, "[\"**\\.pdf\"]" excludes all PDF files.

Type: Array of strings

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

GroupFiles

Grouping files is turned on by default when the input contains more than 50,000 files. To turn on grouping with fewer than 50,000 files, set this parameter to "inPartition". To disable grouping when there are more than 50,000 files, set this parameter to "none".

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

GroupSize

The target group size in bytes. The default is computed based on the input data size and the size of your cluster. When there are fewer than 50,000 input files, "groupFiles" must be set to "inPartition" for this to take effect.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

MaxBand

This option controls the duration in milliseconds after which the s3 listing is likely to be consistent. Files with modification timestamps falling within the last maxBand milliseconds are tracked specially when using JobBookmarks to account for Amazon S3 eventual consistency. Most users don't need to set this option. The default is 900000 milliseconds, or 15 minutes.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

MaxFilesInBand

This option specifies the maximum number of files to save from the last maxBand seconds. If this number is exceeded, extra files are skipped and only processed in the next job run.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

OutputSchemas

Specifies the data schema for the S3 Parquet source.

Type: Array of [GlueSchema](#) objects

Required: No

Recurse

If set to true, recursively reads files in all subdirectories under the specified paths.

Type: Boolean

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

S3SourceAdditionalOptions

Specifies additional connection options for the Amazon S3 data store.

Contents

BoundedFiles

Sets the upper limit for the target number of files that will be processed.

Type: Long

Required: No

BoundedSize

Sets the upper limit for the target size of the dataset in bytes that will be processed.

Type: Long

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

S3Target

Specifies a data store in Amazon Simple Storage Service (Amazon S3).

Contents

ConnectionName

The name of a connection which allows a job or crawler to access data in Amazon S3 within an Amazon Virtual Private Cloud environment (Amazon VPC).

Type: String

Required: No

DlqEventQueueArn

A valid Amazon dead-letter SQS ARN. For example, `arn:aws:sqs:region:account:deadLetterQueue`.

Type: String

Required: No

EventQueueArn

A valid Amazon SQS ARN. For example, `arn:aws:sqs:region:account:sqs`.

Type: String

Required: No

Exclusions

A list of glob patterns used to exclude from the crawl. For more information, see [Catalog Tables with a Crawler](#).

Type: Array of strings

Required: No

Path

The path to the Amazon S3 target.

Type: String

Required: No

SampleSize

Sets the number of files in each leaf folder to be crawled when crawling sample files in a dataset. If not set, all the files are crawled. A valid value is an integer between 1 and 249.

Type: Integer

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Schedule

A scheduling object using a cron statement to schedule an event.

Contents

ScheduleExpression

A cron expression used to specify the schedule (see [Time-Based Schedules for Jobs and Crawlers](#)). For example, to run something every day at 12:15 UTC, you would specify: `cron(15 12 * * ? *)`.

Type: String

Required: No

State

The state of the schedule.

Type: String

Valid Values: SCHEDULED | NOT_SCHEDULED | TRANSITIONING

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SchemaChangePolicy

A policy that specifies update and deletion behaviors for the crawler.

Contents

DeleteBehavior

The deletion behavior when the crawler finds a deleted object.

Type: String

Valid Values: LOG | DELETE_FROM_DATABASE | DEPRECATE_IN_DATABASE

Required: No

UpdateBehavior

The update behavior when the crawler finds a changed schema.

Type: String

Valid Values: LOG | UPDATE_IN_DATABASE

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SchemaColumn

A key-value pair representing a column and data type that this transform can run against. The Schema parameter of the `MLTransform` may contain up to 100 of these structures.

Contents

Data Type

The type of data in the column.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 131072.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Name

The name of the column.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Schemald

The unique ID of the schema in the AWS Glue schema registry.

Contents

RegistryName

The name of the schema registry that contains the schema.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[a-zA-Z0-9-_$#.]+`

Required: No

SchemaArn

The Amazon Resource Name (ARN) of the schema. One of `SchemaArn` or `SchemaName` has to be provided.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 10240.

Pattern: `arn:(aws|aws-us-gov|aws-cn):glue:.*`

Required: No

SchemaName

The name of the schema. One of `SchemaArn` or `SchemaName` has to be provided.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[a-zA-Z0-9-_$#.]+`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SchemaListItem

An object that contains minimal details for a schema.

Contents

CreatedTime

The date and time that a schema was created.

Type: String

Required: No

Description

A description for the schema.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

RegistryName

the name of the registry where the schema resides.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[a-zA-Z0-9-_$#.]+`

Required: No

SchemaArn

The Amazon Resource Name (ARN) for the schema.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 10240.

Pattern: `arn:(aws|aws-us-gov|aws-cn):glue:.*`

Required: No

SchemaName

The name of the schema.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[a-zA-Z0-9-_$#.]+`

Required: No

SchemaStatus

The status of the schema.

Type: String

Valid Values: AVAILABLE | PENDING | DELETING

Required: No

UpdatedTime

The date and time that a schema was updated.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SchemaReference

An object that references a schema stored in the AWS Glue Schema Registry.

Contents

SchemaId

A structure that contains schema identity fields. Either this or the `SchemaVersionId` has to be provided.

Type: [SchemaId](#) object

Required: No

SchemaVersionId

The unique ID assigned to a version of the schema. Either this or the `SchemaId` has to be provided.

Type: String

Length Constraints: Fixed length of 36.

Pattern: `[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

Required: No

SchemaVersionNumber

The version number of the schema.

Type: Long

Valid Range: Minimum value of 1. Maximum value of 100000.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SchemaVersionErrorItem

An object that contains the error details for an operation on a schema version.

Contents

ErrorDetails

The details of the error for the schema version.

Type: [ErrorDetails](#) object

Required: No

VersionNumber

The version number of the schema.

Type: Long

Valid Range: Minimum value of 1. Maximum value of 100000.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SchemaVersionListItem

An object containing the details about a schema version.

Contents

CreatedTime

The date and time the schema version was created.

Type: String

Required: No

SchemaArn

The Amazon Resource Name (ARN) of the schema.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 10240.

Pattern: `arn:(aws|aws-us-gov|aws-cn):glue:.*`

Required: No

SchemaVersionId

The unique identifier of the schema version.

Type: String

Length Constraints: Fixed length of 36.

Pattern: `[a-f0-9]{8}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{4}-[a-f0-9]{12}`

Required: No

Status

The status of the schema version.

Type: String

Valid Values: AVAILABLE | PENDING | FAILURE | DELETING

Required: No

VersionNumber

The version number of the schema.

Type: Long

Valid Range: Minimum value of 1. Maximum value of 100000.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SchemaVersionNumber

A structure containing the schema version information.

Contents

LatestVersion

The latest version available for the schema.

Type: Boolean

Required: No

VersionNumber

The version number of the schema.

Type: Long

Valid Range: Minimum value of 1. Maximum value of 100000.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Segment

Defines a non-overlapping region of a table's partitions, allowing multiple requests to be run in parallel.

Contents

SegmentNumber

The zero-based index number of the segment. For example, if the total number of segments is 4, `SegmentNumber` values range from 0 through 3.

Type: Integer

Valid Range: Minimum value of 0.

Required: Yes

TotalSegments

The total number of segments.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 10.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SelectFields

Specifies a transform that chooses the data property keys that you want to keep.

Contents

Inputs

The data inputs identified by their node names.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]*`

Required: Yes

Name

The name of the transform node.

Type: String

Pattern: `([\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF]|[\^\r\n])*`

Required: Yes

Paths

A JSON path to a variable in the data structure.

Type: Array of arrays of strings

Pattern: `([\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF]|[\^\S\r\n"'])*`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SelectFromCollection

Specifies a transform that chooses one `DynamicFrame` from a collection of `DynamicFrames`. The output is the selected `DynamicFrame`

Contents

Index

The index for the `DynamicFrame` to be selected.

Type: Integer

Valid Range: Minimum value of 0.

Required: Yes

Inputs

The data inputs identified by their node names.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]*`

Required: Yes

Name

The name of the transform node.

Type: String

Pattern: `([\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF]|[\^\r\n])*`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SerDeInfo

Information about a serialization/deserialization program (SerDe) that serves as an extractor and loader.

Contents

Name

Name of the SerDe.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Parameters

These key-value pairs define initialization parameters for the SerDe.

Type: String to string map

Key Length Constraints: Minimum length of 1. Maximum length of 255.

Key Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Value Length Constraints: Maximum length of 512000.

Required: No

SerializationLibrary

Usually the class that implements the SerDe. An example is `org.apache.hadoop.hive.serde2.columnar.ColumnarSerDe`.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Session

The period in which a remote Spark runtime environment is running.

Contents

Command

The command object. See `SessionCommand`.

Type: [SessionCommand](#) object

Required: No

CompletedOn

The date and time that this session is completed.

Type: Timestamp

Required: No

Connections

The number of connections used for the session.

Type: [ConnectionsList](#) object

Required: No

CreatedOn

The time and date when the session was created.

Type: Timestamp

Required: No

DefaultArguments

A map array of key-value pairs. Max is 75 pairs.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 75 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Key Pattern: `[\.\-_\A-Za-z0-9]+`

Value Length Constraints: Minimum length of 0. Maximum length of 4096.

Value Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Description

The description of the session.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

DPUSecods

The DPUs consumed by the session (formula: $ExecutionTime * MaxCapacity$).

Type: Double

Required: No

ErrorMessage

The error message displayed during the session.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

ExecutionTime

The total time the session ran for.

Type: Double

Required: No

GlueVersion

The AWS Glue version determines the versions of Apache Spark and Python that AWS Glue supports. The GlueVersion must be greater than 2.0.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `^\w+\.\w+$`

Required: No

Id

The ID of the session.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

IdleTimeout

The number of minutes when idle before the session times out.

Type: Integer

Required: No

MaxCapacity

The number of AWS Glue data processing units (DPUs) that can be allocated when the job runs. A DPU is a relative measure of processing power that consists of 4 vCPUs of compute capacity and 16 GB memory.

Type: Double

Required: No

NumberOfWorkers

The number of workers of a defined `WorkerType` to use for the session.

Type: Integer

Required: No

ProfileName

The name of an AWS Glue usage profile associated with the session.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Progress

The code execution progress of the session.

Type: Double

Required: No

Role

The name or Amazon Resource Name (ARN) of the IAM role associated with the Session.

Type: String

Length Constraints: Minimum length of 20. Maximum length of 2048.

Pattern: `arn:aws[^:]*:iam:[0-9]*:role/.+`

Required: No

SecurityConfiguration

The name of the SecurityConfiguration structure to be used with the session.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Status

The session status.

Type: String

Valid Values: PROVISIONING | READY | FAILED | TIMEOUT | STOPPING | STOPPED

Required: No

WorkerType

The type of predefined worker that is allocated when a session runs. Accepts a value of G.1X, G.2X, G.4X, or G.8X for Spark sessions. Accepts the value Z.2X for Ray sessions.

Type: String

Valid Values: Standard | G.1X | G.2X | G.025X | G.4X | G.8X | Z.2X

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SessionCommand

The SessionCommand that runs the job.

Contents

Name

Specifies the name of the SessionCommand. Can be 'glueetl' or 'gluestreaming'.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

PythonVersion

Specifies the Python version. The Python version indicates the version supported for jobs of type Spark.

Type: String

Pattern: `^[2-3]|3[.]9$`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SkewedInfo

Specifies skewed values in a table. Skewed values are those that occur with very high frequency.

Contents

SkewedColumnNames

A list of names of columns that contain skewed values.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

SkewedColumnValueLocationMaps

A mapping of skewed values to the columns that contain them.

Type: String to string map

Required: No

SkewedColumnValues

A list of values that appear so frequently as to be considered skewed.

Type: Array of strings

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SnowflakeNodeData

Specifies configuration for Snowflake nodes in AWS Glue Studio.

Contents

Action

Specifies what action to take when writing to a table with preexisting data. Valid values: `append`, `merge`, `truncate`, `drop`.

Type: String

Required: No

AdditionalOptions

Specifies additional options passed to the Snowflake connector. If options are specified elsewhere in this node, this will take precedence.

Type: String to string map

Key Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Value Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

AutoPushdown

Specifies whether automatic query pushdown is enabled. If pushdown is enabled, then when a query is run on Spark, if part of the query can be "pushed down" to the Snowflake server, it is pushed down. This improves performance of some queries.

Type: Boolean

Required: No

Connection

Specifies a AWS Glue Data Catalog Connection to a Snowflake endpoint.

Type: [Option](#) object

Required: No

Database

Specifies a Snowflake database for your node to use.

Type: String

Required: No

IamRole

Not currently used.

Type: [Option](#) object

Required: No

MergeAction

Specifies a merge action. Valid values: `simple`, `custom`. If `simple`, merge behavior is defined by `MergeWhenMatched` and `MergeWhenNotMatched`. If `custom`, defined by `MergeClause`.

Type: String

Pattern: `[A-Za-z0-9_-]*`

Required: No

MergeClause

A SQL statement that specifies a custom merge behavior.

Type: String

Required: No

MergeWhenMatched

Specifies how to resolve records that match preexisting data when merging. Valid values: `update`, `delete`.

Type: String

Pattern: [A-Za-z0-9_-]*

Required: No

MergeWhenNotMatched

Specifies how to process records that do not match preexisting data when merging. Valid values: `insert`, `none`.

Type: String

Pattern: [A-Za-z0-9_-]*

Required: No

PostAction

A SQL string run after the Snowflake connector performs its standard actions.

Type: String

Required: No

PreAction

A SQL string run before the Snowflake connector performs its standard actions.

Type: String

Required: No

SampleQuery

A SQL string used to retrieve data with the query sourcetype.

Type: String

Required: No

Schema

Specifies a Snowflake database schema for your node to use.

Type: String

Required: No

SelectedColumns

Specifies the columns combined to identify a record when detecting matches for merges and upserts. A list of structures with `value`, `label` and `description` keys. Each structure describes a column.

Type: Array of [Option](#) objects

Required: No

SourceType

Specifies how retrieved data is specified. Valid values: "table", "query".

Type: String

Pattern: [A-Za-z0-9_-]*

Required: No

StagingTable

The name of a staging table used when performing merge or upsert append actions. Data is written to this table, then moved to `table` by a generated postaction.

Type: String

Required: No

Table

Specifies a Snowflake table for your node to use.

Type: String

Required: No

TableSchema

Manually defines the target schema for the node. A list of structures with `value`, `label` and `description` keys. Each structure defines a column.

Type: Array of [Option](#) objects

Required: No

TempDir

Not currently used.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]|[^\\S\\r\\n"']`)*

Required: No

Upsert

Used when Action is append. Specifies the resolution behavior when a row already exists. If true, preexisting rows will be updated. If false, those rows will be inserted.

Type: Boolean

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SnowflakeSource

Specifies a Snowflake data source.

Contents

Data

Configuration for the Snowflake data source.

Type: [SnowflakeNodeData](#) object

Required: Yes

Name

The name of the Snowflake data source.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

OutputSchemas

Specifies user-defined schemas for your output data.

Type: Array of [GlueSchema](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SnowflakeTarget

Specifies a Snowflake target.

Contents

Data

Specifies the data of the Snowflake target node.

Type: [SnowflakeNodeData](#) object

Required: Yes

Name

The name of the Snowflake target.

Type: String

Pattern: (`[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF]` | `[\^\r\n]`)*

Required: Yes

Inputs

The nodes that are inputs to the data target.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for Ruby V3](#)

SortCriterion

Specifies a field to sort by and a sort order.

Contents

FieldName

The name of the field on which to sort.

Type: String

Length Constraints: Maximum length of 1024.

Required: No

Sort

An ascending or descending sort.

Type: String

Valid Values: ASC | DESC

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SourceControlDetails

The details for a source control configuration for a job, allowing synchronization of job artifacts to or from a remote repository.

Contents

AuthStrategy

The type of authentication, which can be an authentication token stored in AWS Secrets Manager, or a personal access token.

Type: String

Valid Values: PERSONAL_ACCESS_TOKEN | AWS_SECRETS_MANAGER

Required: No

AuthToken

The value of an authorization token.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 512.

Required: No

Branch

An optional branch in the remote repository.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 512.

Required: No

Folder

An optional folder in the remote repository.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 512.

Required: No

LastCommitId

The last commit ID for a commit in the remote repository.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 512.

Required: No

Owner

The owner of the remote repository that contains the job artifacts.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 512.

Required: No

Provider

The provider for the remote repository.

Type: String

Valid Values: GITHUB | GITLAB | BITBUCKET | AWS_CODE_COMMIT

Required: No

Repository

The name of the remote repository that contains the job artifacts.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 512.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SparkConnectorSource

Specifies a connector to an Apache Spark data source.

Contents

ConnectionName

The name of the connection that is associated with the connector.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n'']`)*

Required: Yes

ConnectionType

The type of connection, such as `marketplace.spark` or `custom.spark`, designating a connection to an Apache Spark data store.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n'']`)*

Required: Yes

ConnectorName

The name of a connector that assists with accessing the data store in AWS Glue Studio.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n'']`)*

Required: Yes

Name

The name of the data source.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\r\\n]`)*

Required: Yes

AdditionalOptions

Additional connection options for the connector.

Type: String to string map

Key Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n''']`)*

Value Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n''']`)*

Required: No

OutputSchemas

Specifies data schema for the custom spark source.

Type: Array of [GlueSchema](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SparkConnectorTarget

Specifies a target that uses an Apache Spark connector.

Contents

ConnectionName

The name of a connection for an Apache Spark connector.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

ConnectionType

The type of connection, such as `marketplace.spark` or `custom.spark`, designating a connection to an Apache Spark data store.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

ConnectorName

The name of an Apache Spark connector.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: Yes

Inputs

The nodes that are inputs to the data target.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: [A-Za-z0-9_-]*

Required: Yes

Name

The name of the data target.

Type: String

Pattern: ([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF]|[\^\r\n])*

Required: Yes

AdditionalOptions

Additional connection options for the connector.

Type: String to string map

Key Pattern: ([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF]|[\^\S\r\n"'])*

Value Pattern: ([\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF]|[\^\S\r\n"'])*

Required: No

OutputSchemas

Specifies the data schema for the custom spark target.

Type: Array of [GlueSchema](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SparkSQL

Specifies a transform where you enter a SQL query using Spark SQL syntax to transform the data. The output is a single `DynamicFrame`.

Contents

Inputs

The data inputs identified by their node names. You can associate a table name with each input node to use in the SQL query. The name you choose must meet the Spark SQL naming restrictions.

Type: Array of strings

Array Members: Minimum number of 1 item.

Pattern: `[A-Za-z0-9_-]*`

Required: Yes

Name

The name of the transform node.

Type: String

Pattern: `([\u0020-\u007F\u00E0-\u00FF\u0080-\u00FF\u0080-\u00FF\u0080-\u00FF\u0080-\u00FF] | [^\r\n])*`

Required: Yes

SqlAliases

A list of aliases. An alias allows you to specify what name to use in the SQL for a given input. For example, you have a datasource named "MyDataSource". If you specify `From` as `MyDataSource`, and `Alias` as `SqlName`, then in your SQL you can do:

```
select * from SqlName
```

and that gets data from `MyDataSource`.

Type: Array of [SqlAlias](#) objects

Required: Yes

SqlQuery

A SQL query that must use Spark SQL syntax and return a single data set.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF\\s]`)*

Required: Yes

OutputSchemas

Specifies the data schema for the SparkSQL transform.

Type: Array of [GlueSchema](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Spigot

Specifies a transform that writes samples of the data to an Amazon S3 bucket.

Contents

Inputs

The data inputs identified by their node names.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: [A-Za-z0-9_-]*

Required: Yes

Name

The name of the transform node.

Type: String

Pattern: ([\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\r\n])*

Required: Yes

Path

A path in Amazon S3 where the transform will write a subset of records from the dataset to a JSON file in an Amazon S3 bucket.

Type: String

Pattern: ([\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF] | [^\S\r\n"'])*

Required: Yes

Prob

The probability (a decimal value with a maximum value of 1) of picking any given record. A value of 1 indicates that each row read from the dataset should be included in the sample output.

Type: Double

Valid Range: Minimum value of 0. Maximum value of 1.

Required: No

Topk

Specifies a number of records to write starting from the beginning of the dataset.

Type: Integer

Valid Range: Minimum value of 0. Maximum value of 100.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SplitFields

Specifies a transform that splits data property keys into two `DynamicFrames`. The output is a collection of `DynamicFrames`: one with selected data property keys, and one with the remaining data property keys.

Contents

Inputs

The data inputs identified by their node names.

Type: Array of strings

Array Members: Fixed number of 1 item.

Pattern: `[A-Za-z0-9_-]*`

Required: Yes

Name

The name of the transform node.

Type: String

Pattern: `([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF]|[\^\r\n])*`

Required: Yes

Paths

A JSON path to a variable in the data structure.

Type: Array of arrays of strings

Pattern: `([\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF]|[\^\S\r\n"'])*`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SqlAlias

Represents a single entry in the list of values for `SqlAliases`.

Contents

Alias

A temporary name given to a table, or a column in a table.

Type: String

Pattern: (`[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF]` | `[\S\r\n]`)*

Required: Yes

From

A table, or a column in a table.

Type: String

Pattern: `[A-Za-z0-9_-]*`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

StartingEventBatchCondition

The batch condition that started the workflow run. Either the number of events in the batch size arrived, in which case the BatchSize member is non-zero, or the batch window expired, in which case the BatchWindow member is non-zero.

Contents

BatchSize

Number of events in the batch.

Type: Integer

Required: No

BatchWindow

Duration of the batch window in seconds.

Type: Integer

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Statement

The statement or request for a particular action to occur in a session.

Contents

Code

The execution code of the statement.

Type: String

Required: No

CompletedOn

The unix time and date that the job definition was completed.

Type: Long

Required: No

Id

The ID of the statement.

Type: Integer

Required: No

Output

The output in JSON.

Type: [StatementOutput](#) object

Required: No

Progress

The code execution progress.

Type: Double

Required: No

StartedOn

The unix time and date that the job definition was started.

Type: Long

Required: No

State

The state while request is actioned.

Type: String

Valid Values: WAITING | RUNNING | AVAILABLE | CANCELLING | CANCELLED | ERROR

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

StatementOutput

The code execution output in JSON format.

Contents

Data

The code execution output.

Type: [StatementOutputData](#) object

Required: No

ErrorName

The name of the error in the output.

Type: String

Required: No

ErrorValue

The error value of the output.

Type: String

Required: No

ExecutionCount

The execution count of the output.

Type: Integer

Required: No

Status

The status of the code execution output.

Type: String

Valid Values: WAITING | RUNNING | AVAILABLE | CANCELLING | CANCELLED | ERROR

Required: No

Traceback

The traceback of the output.

Type: Array of strings

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

StatementOutputData

The code execution output in JSON format.

Contents

TextPlain

The code execution output in text format.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

StorageDescriptor

Describes the physical storage of table data.

Contents

AdditionalLocations

A list of locations that point to the path where a Delta table is located.

Type: Array of strings

Length Constraints: Maximum length of 2056.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

BucketColumns

A list of reducer grouping columns, clustering columns, and bucketing columns in the table.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Columns

A list of the Columns in the table.

Type: Array of [Column](#) objects

Required: No

Compressed

True if the data in the table is compressed, or False if not.

Type: Boolean

Required: No

InputFormat

The input format: `SequenceFileInputFormat` (binary), or `TextInputFormat`, or a custom format.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Location

The physical location of the table. By default, this takes the form of the warehouse location, followed by the database location in the warehouse, followed by the table name.

Type: String

Length Constraints: Maximum length of 2056.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

NumberOfBuckets

Must be specified if the table contains any dimension columns.

Type: Integer

Required: No

OutputFormat

The output format: `SequenceFileOutputFormat` (binary), or `IgnoreKeyTextOutputFormat`, or a custom format.

Type: String

Length Constraints: Maximum length of 128.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Parameters

The user-supplied properties in key-value form.

Type: String to string map

Key Length Constraints: Minimum length of 1. Maximum length of 255.

Key Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Value Length Constraints: Maximum length of 512000.

Required: No

SchemaReference

An object that references a schema stored in the AWS Glue Schema Registry.

When creating a table, you can pass an empty list of columns for the schema, and instead use a schema reference.

Type: [SchemaReference](#) object

Required: No

SerdeInfo

The serialization/deserialization (SerDe) information.

Type: [SerDeInfo](#) object

Required: No

SkewedInfo

The information about values that appear frequently in a column (skewed values).

Type: [SkewedInfo](#) object

Required: No

SortColumns

A list specifying the sort order of each bucket in the table.

Type: Array of [Order](#) objects

Required: No

StoredAsSubDirectories

True if the table data is stored in subdirectories, or False if not.

Type: Boolean

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

StreamingDataPreviewOptions

Specifies options related to data preview for viewing a sample of your data.

Contents

PollingTime

The polling time in milliseconds.

Type: Long

Valid Range: Minimum value of 10.

Required: No

RecordPollingLimit

The limit to the number of records polled.

Type: Long

Valid Range: Minimum value of 1.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

StringColumnStatisticsData

Defines column statistics supported for character sequence data values.

Contents

AverageLength

The average string length in the column.

Type: Double

Valid Range: Minimum value of 0.0.

Required: Yes

MaximumLength

The size of the longest string in the column.

Type: Long

Valid Range: Minimum value of 0.

Required: Yes

NumberOfDistinctValues

The number of distinct values in a column.

Type: Long

Valid Range: Minimum value of 0.

Required: Yes

NumberOfNulls

The number of null values in the column.

Type: Long

Valid Range: Minimum value of 0.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SupportedDialect

A structure specifying the dialect and dialect version used by the query engine.

Contents

Dialect

The dialect of the query engine.

Type: String

Valid Values: REDSHIFT | ATHENA | SPARK

Required: No

DialectVersion

The version of the dialect of the query engine. For example, 3.0.0.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Table

Represents a collection of related data organized in columns and rows.

Contents

Name

The table name. For Hive compatibility, this must be entirely lowercase.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

CatalogId

The ID of the Data Catalog in which the table resides.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

CreatedBy

The person or entity who created the table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

CreateTime

The time when the table definition was created in the Data Catalog.

Type: Timestamp

Required: No

DatabaseName

The name of the database where the table metadata resides. For Hive compatibility, this must be all lowercase.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Description

A description of the table.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

FederatedTable

A `FederatedTable` structure that references an entity outside the AWS Glue Data Catalog.

Type: [FederatedTable](#) object

Required: No

IsMultiDialectView

Specifies whether the view supports the SQL dialects of one or more different query engines and can therefore be read by those engines.

Type: Boolean

Required: No

IsRegisteredWithLakeFormation

Indicates whether the table has been registered with AWS Lake Formation.

Type: Boolean

Required: No

LastAccessTime

The last time that the table was accessed. This is usually taken from HDFS, and might not be reliable.

Type: Timestamp

Required: No

LastAnalyzedTime

The last time that column statistics were computed for this table.

Type: Timestamp

Required: No

Owner

The owner of the table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Parameters

These key-value pairs define properties associated with the table.

Type: String to string map

Key Length Constraints: Minimum length of 1. Maximum length of 255.

Key Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Value Length Constraints: Maximum length of 512000.

Required: No

PartitionKeys

A list of columns by which the table is partitioned. Only primitive types are supported as partition keys.

When you create a table used by Amazon Athena, and you do not specify any `partitionKeys`, you must at least set the value of `partitionKeys` to an empty list. For example:

```
"PartitionKeys": []
```

Type: Array of [Column](#) objects

Required: No

Retention

The retention time for this table.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

StorageDescriptor

A storage descriptor containing information about the physical storage of this table.

Type: [StorageDescriptor](#) object

Required: No

TableType

The type of this table. AWS Glue will create tables with the `EXTERNAL_TABLE` type. Other services, such as Athena, may create tables with additional table types.

AWS Glue related table types:

`EXTERNAL_TABLE`

Hive compatible attribute - indicates a non-Hive managed table.

GOVERNED

Used by AWS Lake Formation. The AWS Glue Data Catalog understands GOVERNED.

Type: String

Length Constraints: Maximum length of 255.

Required: No

TargetTable

A `TableIdentifier` structure that describes a target table for resource linking.

Type: [TableIdentifier](#) object

Required: No

UpdateTime

The last time that the table was updated.

Type: Timestamp

Required: No

VersionId

The ID of the table version.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

ViewDefinition

A structure that contains all the information that defines the view, including the dialect or dialects for the view, and the query.

Type: [ViewDefinition](#) object

Required: No

ViewExpandedText

Included for Apache Hive compatibility. Not used in the normal course of AWS Glue operations.

Type: String

Length Constraints: Maximum length of 409600.

Required: No

ViewOriginalText

Included for Apache Hive compatibility. Not used in the normal course of AWS Glue operations. If the table is a VIRTUAL_VIEW, certain Athena configuration encoded in base64.

Type: String

Length Constraints: Maximum length of 409600.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

TableError

An error record for table operations.

Contents

ErrorDetail

The details about the error.

Type: [ErrorDetail](#) object

Required: No

TableName

The name of the table. For Hive compatibility, this must be entirely lowercase.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

TableIdentifier

A structure that describes a target table for resource linking.

Contents

CatalogId

The ID of the Data Catalog in which the table resides.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

DatabaseName

The name of the catalog database that contains the target table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Name

The name of the target table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Region

Region of the target table.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

TableInput

A structure used to define a table.

Contents

Name

The table name. For Hive compatibility, this is folded to lowercase when it is stored.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Description

A description of the table.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

LastAccessTime

The last time that the table was accessed.

Type: Timestamp

Required: No

LastAnalyzedTime

The last time that column statistics were computed for this table.

Type: Timestamp

Required: No

Owner

The table owner. Included for Apache Hive compatibility. Not used in the normal course of AWS Glue operations.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Parameters

These key-value pairs define properties associated with the table.

Type: String to string map

Key Length Constraints: Minimum length of 1. Maximum length of 255.

Key Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Value Length Constraints: Maximum length of 512000.

Required: No

PartitionKeys

A list of columns by which the table is partitioned. Only primitive types are supported as partition keys.

When you create a table used by Amazon Athena, and you do not specify any `partitionKeys`, you must at least set the value of `partitionKeys` to an empty list. For example:

```
"PartitionKeys": []
```

Type: Array of [Column](#) objects

Required: No

Retention

The retention time for this table.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

StorageDescriptor

A storage descriptor containing information about the physical storage of this table.

Type: [StorageDescriptor](#) object

Required: No

TableType

The type of this table. AWS Glue will create tables with the EXTERNAL_TABLE type. Other services, such as Athena, may create tables with additional table types.

AWS Glue related table types:

EXTERNAL_TABLE

Hive compatible attribute - indicates a non-Hive managed table.

GOVERNED

Used by AWS Lake Formation. The AWS Glue Data Catalog understands GOVERNED.

Type: String

Length Constraints: Maximum length of 255.

Required: No

TargetTable

A `TableIdentifier` structure that describes a target table for resource linking.

Type: [TableIdentifier](#) object

Required: No

ViewDefinition

A structure that contains all the information that defines the view, including the dialect or dialects for the view, and the query.

Type: [ViewDefinitionInput](#) object

Required: No

ViewExpandedText

Included for Apache Hive compatibility. Not used in the normal course of AWS Glue operations.

Type: String

Length Constraints: Maximum length of 409600.

Required: No

ViewOriginalText

Included for Apache Hive compatibility. Not used in the normal course of AWS Glue operations. If the table is a VIRTUAL_VIEW, certain Athena configuration encoded in base64.

Type: String

Length Constraints: Maximum length of 409600.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

TableOptimizer

Contains details about an optimizer associated with a table.

Contents

configuration

A `TableOptimizerConfiguration` object that was specified when creating or updating a table optimizer.

Type: [TableOptimizerConfiguration](#) object

Required: No

lastRun

A `TableOptimizerRun` object representing the last run of the table optimizer.

Type: [TableOptimizerRun](#) object

Required: No

type

The type of table optimizer. Currently, the only valid value is `compaction`.

Type: String

Valid Values: `compaction`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

TableOptimizerConfiguration

Contains details on the configuration of a table optimizer. You pass this configuration when creating or updating a table optimizer.

Contents

enabled

Whether table optimization is enabled.

Type: Boolean

Required: No

roleArn

A role passed by the caller which gives the service permission to update the resources associated with the optimizer on the caller's behalf.

Type: String

Length Constraints: Minimum length of 20. Maximum length of 2048.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

TableOptimizerRun

Contains details for a table optimizer run.

Contents

endTimeStamp

Represents the epoch timestamp at which the compaction job ended.

Type: Timestamp

Required: No

error

An error that occurred during the optimizer run.

Type: String

Required: No

eventType

An event type representing the status of the table optimizer run.

Type: String

Valid Values: `starting` | `completed` | `failed` | `in_progress`

Required: No

metrics

A `RunMetrics` object containing metrics for the optimizer run.

Type: [RunMetrics](#) object

Required: No

startTimeStamp

Represents the epoch timestamp at which the compaction job was started within Lake Formation.

Type: Timestamp

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

TableVersion

Specifies a version of a table.

Contents

Table

The table in question.

Type: [Table](#) object

Required: No

VersionId

The ID value that identifies this table version. A `VersionId` is a string representation of an integer. Each version is incremented by 1.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

TableVersionError

An error record for table-version operations.

Contents

ErrorDetail

The details about the error.

Type: [ErrorDetail](#) object

Required: No

TableName

The name of the table in question.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

VersionId

The ID value of the version in question. A `VersionID` is a string representation of an integer. Each version is incremented by 1.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

TaskRun

The sampling parameters that are associated with the machine learning transform.

Contents

CompletedOn

The last point in time that the requested task run was completed.

Type: Timestamp

Required: No

ErrorString

The list of error strings associated with this task run.

Type: String

Required: No

ExecutionTime

The amount of time (in seconds) that the task run consumed resources.

Type: Integer

Required: No

LastModifiedOn

The last point in time that the requested task run was updated.

Type: Timestamp

Required: No

LogGroupName

The names of the log group for secure logging, associated with this task run.

Type: String

Required: No

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

TaskRunFilterCriteria

The criteria that are used to filter the task runs for the machine learning transform.

Contents

StartedAfter

Filter on task runs started after this date.

Type: Timestamp

Required: No

StartedBefore

Filter on task runs started before this date.

Type: Timestamp

Required: No

Status

The current status of the task run.

Type: String

Valid Values: STARTING | RUNNING | STOPPING | STOPPED | SUCCEEDED | FAILED | TIMEOUT

Required: No

TaskRunType

The type of task run.

Type: String

Valid Values: EVALUATION | LABELING_SET_GENERATION | IMPORT_LABELS | EXPORT_LABELS | FIND_MATCHES

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

TaskRunProperties

The configuration properties for the task run.

Contents

ExportLabelsTaskRunProperties

The configuration properties for an exporting labels task run.

Type: [ExportLabelsTaskRunProperties](#) object

Required: No

FindMatchesTaskRunProperties

The configuration properties for a find matches task run.

Type: [FindMatchesTaskRunProperties](#) object

Required: No

ImportLabelsTaskRunProperties

The configuration properties for an importing labels task run.

Type: [ImportLabelsTaskRunProperties](#) object

Required: No

LabelingSetGenerationTaskRunProperties

The configuration properties for a labeling set generation task run.

Type: [LabelingSetGenerationTaskRunProperties](#) object

Required: No

TaskType

The type of task run.

Type: String

Valid Values: EVALUATION | LABELING_SET_GENERATION | IMPORT_LABELS | EXPORT_LABELS | FIND_MATCHES

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

TaskRunSortCriteria

The sorting criteria that are used to sort the list of task runs for the machine learning transform.

Contents

Column

The column to be used to sort the list of task runs for the machine learning transform.

Type: String

Valid Values: TASK_RUN_TYPE | STATUS | STARTED

Required: Yes

SortDirection

The sort direction to be used to sort the list of task runs for the machine learning transform.

Type: String

Valid Values: DESCENDING | ASCENDING

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

TransformConfigParameter

Specifies the parameters in the config file of the dynamic transform.

Contents

Name

Specifies the name of the parameter in the config file of the dynamic transform.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF]|[^\\S\\r\\n"']`)*

Required: Yes

Type

Specifies the parameter type in the config file of the dynamic transform.

Type: String

Valid Values: `str` | `int` | `float` | `complex` | `bool` | `list` | `null`

Required: Yes

IsOptional

Specifies whether the parameter is optional or not in the config file of the dynamic transform.

Type: Boolean

Required: No

ListType

Specifies the list type of the parameter in the config file of the dynamic transform.

Type: String

Valid Values: `str` | `int` | `float` | `complex` | `bool` | `list` | `null`

Required: No

ValidationMessage

Specifies the validation message in the config file of the dynamic transform.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

ValidationRule

Specifies the validation rule in the config file of the dynamic transform.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

Value

Specifies the value of the parameter in the config file of the dynamic transform.

Type: Array of strings

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

TransformFilterCriteria

The criteria used to filter the machine learning transforms.

Contents

CreatedAfter

The time and date after which the transforms were created.

Type: Timestamp

Required: No

CreatedBefore

The time and date before which the transforms were created.

Type: Timestamp

Required: No

GlueVersion

This value determines which version of AWS Glue this machine learning transform is compatible with. Glue 1.0 is recommended for most customers. If the value is not set, the Glue compatibility defaults to Glue 0.9. For more information, see [AWS Glue Versions](#) in the developer guide.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `^\w+\.\w+$`

Required: No

LastModifiedAfter

Filter on transforms last modified after this date.

Type: Timestamp

Required: No

LastModifiedBefore

Filter on transforms last modified before this date.

Type: Timestamp

Required: No

Name

A unique transform name that is used to filter the machine learning transforms.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Schema

Filters on datasets with a specific schema. The `Map<Column, Type>` object is an array of key-value pairs representing the schema this transform accepts, where `Column` is the name of a column, and `Type` is the type of the data such as an integer or string. Has an upper bound of 100 columns.

Type: Array of [SchemaColumn](#) objects

Array Members: Maximum number of 100 items.

Required: No

Status

Filters the list of machine learning transforms by the last known status of the transforms (to indicate whether a transform can be used or not). One of "NOT_READY", "READY", or "DELETING".

Type: String

Valid Values: NOT_READY | READY | DELETING

Required: No

TransformType

The type of machine learning transform that is used to filter the machine learning transforms.

Type: String

Valid Values: FIND_MATCHES

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

TransformParameters

The algorithm-specific parameters that are associated with the machine learning transform.

Contents

TransformType

The type of machine learning transform.

For information about the types of machine learning transforms, see [Creating Machine Learning Transforms](#).

Type: String

Valid Values: FIND_MATCHES

Required: Yes

FindMatchesParameters

The parameters for the find matches algorithm.

Type: [FindMatchesParameters](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

TransformSortCriteria

The sorting criteria that are associated with the machine learning transform.

Contents

Column

The column to be used in the sorting criteria that are associated with the machine learning transform.

Type: String

Valid Values: NAME | TRANSFORM_TYPE | STATUS | CREATED | LAST_MODIFIED

Required: Yes

SortDirection

The sort direction to be used in the sorting criteria that are associated with the machine learning transform.

Type: String

Valid Values: DESCENDING | ASCENDING

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Trigger

Information about a specific trigger.

Contents

Actions

The actions initiated by this trigger.

Type: Array of [Action](#) objects

Required: No

Description

A description of this trigger.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

EventBatchingCondition

Batch condition that must be met (specified number of events received or batch time window expired) before EventBridge event trigger fires.

Type: [EventBatchingCondition](#) object

Required: No

Id

Reserved for future use.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Name

The name of the trigger.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Predicate

The predicate of this trigger, which defines when it will fire.

Type: [Predicate](#) object

Required: No

Schedule

A cron expression used to specify the schedule (see [Time-Based Schedules for Jobs and Crawlers](#)). For example, to run something every day at 12:15 UTC, you would specify: `cron(15 12 * * ? *)`.

Type: String

Required: No

State

The current state of the trigger.

Type: String

Valid Values: CREATING | CREATED | ACTIVATING | ACTIVATED | DEACTIVATING | DEACTIVATED | DELETING | UPDATING

Required: No

Type

The type of trigger that this is.

Type: String

Valid Values: SCHEDULED | CONDITIONAL | ON_DEMAND | EVENT

Required: No

WorkflowName

The name of the workflow associated with the trigger.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

TriggerNodeDetails

The details of a Trigger node present in the workflow.

Contents

Trigger

The information of the trigger represented by the trigger node.

Type: [Trigger](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

TriggerUpdate

A structure used to provide information used to update a trigger. This object updates the previous trigger definition by overwriting it completely.

Contents

Actions

The actions initiated by this trigger.

Type: Array of [Action](#) objects

Required: No

Description

A description of this trigger.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\x\n\t]*`

Required: No

EventBatchingCondition

Batch condition that must be met (specified number of events received or batch time window expired) before EventBridge event trigger fires.

Type: [EventBatchingCondition](#) object

Required: No

Name

Reserved for future use.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Predicate

The predicate of this trigger, which defines when it will fire.

Type: [Predicate](#) object

Required: No

Schedule

A cron expression used to specify the schedule (see [Time-Based Schedules for Jobs and Crawlers](#)). For example, to run something every day at 12:15 UTC, you would specify: `cron(15 12 * * ? *)`.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

UnfilteredPartition

A partition that contains unfiltered metadata.

Contents

AuthorizedColumns

The list of columns the user has permissions to access.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

IsRegisteredWithLakeFormation

A Boolean value indicating that the partition location is registered with Lake Formation.

Type: Boolean

Required: No

Partition

The partition object.

Type: [Partition](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Union

Specifies a transform that combines the rows from two or more datasets into a single result.

Contents

Inputs

The node ID inputs to the transform.

Type: Array of strings

Array Members: Fixed number of 2 items.

Pattern: `[A-Za-z0-9_-]*`

Required: Yes

Name

The name of the transform node.

Type: String

Pattern: `([\u0020-\u007F\u00E0-\u00FF\u0080-\u00FF\u0080-\u00FF\u0080-\u00FF\u0080-\u00FF\u0080-\u00FF] | [^\r\n])*`

Required: Yes

UnionType

Indicates the type of Union transform.

Specify ALL to join all rows from data sources to the resulting DynamicFrame. The resulting union does not remove duplicate rows.

Specify DISTINCT to remove duplicate rows in the resulting DynamicFrame.

Type: String

Valid Values: ALL | DISTINCT

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

UpdateCsvClassifierRequest

Specifies a custom CSV classifier to be updated.

Contents

Name

The name of the classifier.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

AllowSingleColumn

Enables the processing of files that contain only one column.

Type: Boolean

Required: No

ContainsHeader

Indicates whether the CSV file contains a header.

Type: String

Valid Values: UNKNOWN | PRESENT | ABSENT

Required: No

CustomDatatypeConfigured

Specifies the configuration of custom datatypes.

Type: Boolean

Required: No

CustomDatatypes

Specifies a list of supported custom datatypes.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

Delimiter

A custom symbol to denote what separates each column entry in the row.

Type: String

Length Constraints: Fixed length of 1.

Pattern: `[\^\r\n]`

Required: No

DisableValueTrimming

Specifies not to trim values before identifying the type of column values. The default value is true.

Type: Boolean

Required: No

Header

A list of strings representing column names.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

QuoteSymbol

A custom symbol to denote what combines content into a single column value. It must be different from the column delimiter.

Type: String

Length Constraints: Fixed length of 1.

Pattern: [^\r\n]

Required: No

Serde

Sets the SerDe for processing CSV in the classifier, which will be applied in the Data Catalog. Valid values are `OpenCSVSerde`, `LazySimpleSerDe`, and `None`. You can specify the `None` value when you want the crawler to do the detection.

Type: String

Valid Values: `OpenCSVSerde` | `LazySimpleSerDe` | `None`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

UpdateGrokClassifierRequest

Specifies a grok classifier to update when passed to `UpdateClassifier`.

Contents

Name

The name of the `GrokClassifier`.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Classification

An identifier of the data format that the classifier matches, such as Twitter, JSON, Omniture logs, Amazon CloudWatch Logs, and so on.

Type: String

Required: No

CustomPatterns

Optional custom grok patterns used by this classifier.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 16000.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

GrokPattern

The grok pattern used by this classifier.

Type: String

UpdateJsonClassifierRequest

Specifies a JSON classifier to be updated.

Contents

Name

The name of the classifier.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

JsonPath

A JsonPath string defining the JSON data for the classifier to classify. AWS Glue supports a subset of JsonPath, as described in [Writing JsonPath Custom Classifiers](#).

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

UpdateXMLClassifierRequest

Specifies an XML classifier to be updated.

Contents

Name

The name of the classifier.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

Classification

An identifier of the data format that the classifier matches.

Type: String

Required: No

RowTag

The XML tag designating the element that contains each record in an XML document being parsed. This cannot identify a self-closing element (closed by `/>`). An empty row element that contains only attributes can be parsed as long as it ends with a closing tag (for example, `<row item_a="A" item_b="B"></row>` is okay, but `<row item_a="A" item_b="B" />` is not).

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)

- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

UpsertRedshiftTargetOptions

The options to configure an upsert operation when writing to a Redshift target .

Contents

ConnectionName

The name of the connection to use to write to Redshift.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

TableLocation

The physical location of the Redshift table.

Type: String

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

UpsertKeys

The keys used to determine whether to perform an update or insert.

Type: Array of strings

Pattern: (`[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF]` | `[^\\S\\r\\n"']`)*

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

UsageProfileDefinition

Describes an AWS Glue usage profile.

Contents

CreatedOn

The date and time when the usage profile was created.

Type: Timestamp

Required: No

Description

A description of the usage profile.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 2048.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

LastModifiedOn

The date and time when the usage profile was last modified.

Type: Timestamp

Required: No

Name

The name of the usage profile.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

UserDefinedFunction

Represents the equivalent of a Hive user-defined function (UDF) definition.

Contents

CatalogId

The ID of the Data Catalog in which the function resides.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

ClassName

The Java class that contains the function code.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

CreateTime

The time at which the function was created.

Type: Timestamp

Required: No

DatabaseName

The name of the catalog database that contains the function.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

FunctionName

The name of the function.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

OwnerName

The owner of the function.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

OwnerType

The owner type.

Type: String

Valid Values: USER | ROLE | GROUP

Required: No

ResourceUris

The resource URIs for the function.

Type: Array of [ResourceUri](#) objects

Array Members: Minimum number of 0 items. Maximum number of 1000 items.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

UserDefinedFunctionInput

A structure used to create or update a user-defined function.

Contents

ClassName

The Java class that contains the function code.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

FunctionName

The name of the function.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

OwnerName

The owner of the function.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

OwnerType

The owner type.

Type: String

Valid Values: USER | ROLE | GROUP

Required: No

ResourceUris

The resource URIs for the function.

Type: Array of [ResourceUri](#) objects

Array Members: Minimum number of 0 items. Maximum number of 1000 items.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ViewDefinition

A structure containing details for representations.

Contents

Definer

The definer of a view in SQL.

Type: String

Length Constraints: Minimum length of 20. Maximum length of 2048.

Required: No

IsProtected

You can set this flag as true to instruct the engine not to push user-provided operations into the logical plan of the view during query planning. However, setting this flag does not guarantee that the engine will comply. Refer to the engine's documentation to understand the guarantees provided, if any.

Type: Boolean

Required: No

Representations

A list of representations.

Type: Array of [ViewRepresentation](#) objects

Array Members: Minimum number of 1 item. Maximum number of 1000 items.

Required: No

SubObjects

A list of table Amazon Resource Names (ARNs).

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 10 items.

Length Constraints: Minimum length of 20. Maximum length of 2048.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ViewDefinitionInput

A structure containing details for creating or updating an AWS Glue view.

Contents

Definer

The definer of a view in SQL.

Type: String

Length Constraints: Minimum length of 20. Maximum length of 2048.

Required: No

IsProtected

You can set this flag as true to instruct the engine not to push user-provided operations into the logical plan of the view during query planning. However, setting this flag does not guarantee that the engine will comply. Refer to the engine's documentation to understand the guarantees provided, if any.

Type: Boolean

Required: No

Representations

A list of structures that contains the dialect of the view, and the query that defines the view.

Type: Array of [ViewRepresentationInput](#) objects

Array Members: Minimum number of 1 item. Maximum number of 10 items.

Required: No

SubObjects

A list of base table ARNs that make up the view.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 10 items.

Length Constraints: Minimum length of 20. Maximum length of 2048.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ViewRepresentation

A structure that contains the dialect of the view, and the query that defines the view.

Contents

Dialect

The dialect of the query engine.

Type: String

Valid Values: REDSHIFT | ATHENA | SPARK

Required: No

DialectVersion

The version of the dialect of the query engine. For example, 3.0.0.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Required: No

IsStale

Dialects marked as stale are no longer valid and must be updated before they can be queried in their respective query engines.

Type: Boolean

Required: No

ValidationConnection

The name of the connection to be used to validate the specific representation of the view.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

ViewExpandedText

The expanded SQL for the view. This SQL is used by engines while processing a query on a view. Engines may perform operations during view creation to transform `ViewOriginalText` to `ViewExpandedText`. For example:

- Fully qualified identifiers: `SELECT * from table1 -> SELECT * from db1.table1`

Type: String

Length Constraints: Maximum length of 409600.

Required: No

ViewOriginalText

The `SELECT` query provided by the customer during `CREATE VIEW DDL`. This SQL is not used during a query on a view (`ViewExpandedText` is used instead). `ViewOriginalText` is used for cases like `SHOW CREATE VIEW` where users want to see the original DDL command that created the view.

Type: String

Length Constraints: Maximum length of 409600.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ViewRepresentationInput

A structure containing details of a representation to update or create a Lake Formation view.

Contents

Dialect

A parameter that specifies the engine type of a specific representation.

Type: String

Valid Values: REDSHIFT | ATHENA | SPARK

Required: No

DialectVersion

A parameter that specifies the version of the engine of a specific representation.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Required: No

ValidationConnection

The name of the connection to be used to validate the specific representation of the view.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

ViewExpandedText

A string that represents the SQL query that describes the view with expanded resource ARNs

Type: String

Length Constraints: Maximum length of 409600.

Required: No

ViewOriginalText

A string that represents the original SQL query that describes the view.

Type: String

Length Constraints: Maximum length of 409600.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Graph

The graph representing all the AWS Glue components that belong to the workflow as nodes and directed connections between them as edges.

Type: [WorkflowGraph](#) object

Required: No

LastModifiedOn

The date and time when the workflow was last modified.

Type: Timestamp

Required: No

LastRun

The information about the last execution of the workflow.

Type: [WorkflowRun](#) object

Required: No

MaxConcurrentRuns

You can use this parameter to prevent unwanted multiple updates to data, to control costs, or in some cases, to prevent exceeding the maximum number of concurrent runs of any of the component jobs. If you leave this parameter blank, there is no limit to the number of concurrent workflow runs.

Type: Integer

Required: No

Name

The name of the workflow.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

WorkflowGraph

A workflow graph represents the complete workflow containing all the AWS Glue components present in the workflow and all the directed connections between them.

Contents

Edges

A list of all the directed connections between the nodes belonging to the workflow.

Type: Array of [Edge](#) objects

Required: No

Nodes

A list of the the AWS Glue components belong to the workflow represented as nodes.

Type: Array of [Node](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PreviousRunId

The ID of the previous workflow run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

StartedOn

The date and time when the workflow run was started.

Type: Timestamp

Required: No

StartingEventBatchCondition

The batch condition that started the workflow run.

Type: [StartingEventBatchCondition](#) object

Required: No

Statistics

The statistics of the run.

Type: [WorkflowRunStatistics](#) object

Required: No

Status

The status of the workflow run.

Type: String

Valid Values: RUNNING | COMPLETED | STOPPING | STOPPED | ERROR

Required: No

WorkflowRunId

The ID of this workflow run.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

WorkflowRunProperties

The workflow run properties which were set during the run.

Type: String to string map

Key Length Constraints: Minimum length of 1. Maximum length of 255.

Key Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

WorkflowRunStatistics

Workflow run statistics provides statistics about the workflow run.

Contents

ErroredActions

Indicates the count of job runs in the ERROR state in the workflow run.

Type: Integer

Required: No

FailedActions

Total number of Actions that have failed.

Type: Integer

Required: No

RunningActions

Total number Actions in running state.

Type: Integer

Required: No

StoppedActions

Total number of Actions that have stopped.

Type: Integer

Required: No

SucceededActions

Total number of Actions that have succeeded.

Type: Integer

Required: No

TimeoutActions

Total number of Actions that timed out.

Type: Integer

Required: No

TotalActions

Total number of Actions in the workflow run.

Type: Integer

Required: No

WaitingActions

Indicates the count of job runs in WAITING state in the workflow run.

Type: Integer

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

XMLClassifier

A classifier for XML content.

Contents

Classification

An identifier of the data format that the classifier matches.

Type: String

Required: Yes

Name

The name of the classifier.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`

Required: Yes

CreationTime

The time that this classifier was registered.

Type: Timestamp

Required: No

LastUpdated

The time that this classifier was last updated.

Type: Timestamp

Required: No

RowTag

The XML tag designating the element that contains each record in an XML document being parsed. This can't identify a self-closing element (closed by `/>`). An empty row element that

contains only attributes can be parsed as long as it ends with a closing tag (for example, `<row item_a="A" item_b="B"></row>` is okay, but `<row item_a="A" item_b="B" />` is not).

Type: String

Required: No

Version

The version of this classifier.

Type: Long

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Common Parameters

The following list contains the parameters that all actions use for signing Signature Version 4 requests with a query string. Any action-specific parameters are listed in the topic for that action. For more information about Signature Version 4, see [Signing AWS API requests](#) in the *IAM User Guide*.

Action

The action to be performed.

Type: string

Required: Yes

Version

The API version that the request is written for, expressed in the format YYYY-MM-DD.

Type: string

Required: Yes

X-Amz-Algorithm

The hash algorithm that you used to create the request signature.

Condition: Specify this parameter when you include authentication information in a query string instead of in the HTTP authorization header.

Type: string

Valid Values: AWS4-HMAC-SHA256

Required: Conditional

X-Amz-Credential

The credential scope value, which is a string that includes your access key, the date, the region you are targeting, the service you are requesting, and a termination string ("aws4_request"). The value is expressed in the following format: *access_key/YYYYMMDD/region/service/aws4_request*.

For more information, see [Create a signed AWS API request](#) in the *IAM User Guide*.

Condition: Specify this parameter when you include authentication information in a query string instead of in the HTTP authorization header.

Type: string

Required: Conditional

X-Amz-Date

The date that is used to create the signature. The format must be ISO 8601 basic format (YYYYMMDD'T'HHMMSS'Z'). For example, the following date time is a valid X-Amz-Date value: 20120325T120000Z.

Condition: X-Amz-Date is optional for all requests; it can be used to override the date used for signing requests. If the Date header is specified in the ISO 8601 basic format, X-Amz-Date is not required. When X-Amz-Date is used, it always overrides the value of the Date header. For more information, see [Elements of an AWS API request signature](#) in the *IAM User Guide*.

Type: string

Required: Conditional

X-Amz-Security-Token

The temporary security token that was obtained through a call to AWS Security Token Service (AWS STS). For a list of services that support temporary security credentials from AWS STS, see [AWS services that work with IAM](#) in the *IAM User Guide*.

Condition: If you're using temporary security credentials from AWS STS, you must include the security token.

Type: string

Required: Conditional

X-Amz-Signature

Specifies the hex-encoded signature that was calculated from the string to sign and the derived signing key.

Condition: Specify this parameter when you include authentication information in a query string instead of in the HTTP authorization header.

Type: string

Required: Conditional

X-Amz-SignedHeaders

Specifies all the HTTP headers that were included as part of the canonical request. For more information about specifying signed headers, see [Create a signed AWS API request](#) in the *IAM User Guide*.

Condition: Specify this parameter when you include authentication information in a query string instead of in the HTTP authorization header.

Type: string

Required: Conditional

Common Errors

This section lists the errors common to the API actions of all AWS services. For errors specific to an API action for this service, see the topic for that API action.

AccessDeniedException

You do not have sufficient access to perform this action.

HTTP Status Code: 400

IncompleteSignature

The request signature does not conform to AWS standards.

HTTP Status Code: 400

InternalFailure

The request processing has failed because of an unknown error, exception or failure.

HTTP Status Code: 500

InvalidAction

The action or operation requested is invalid. Verify that the action is typed correctly.

HTTP Status Code: 400

InvalidClientTokenId

The X.509 certificate or AWS access key ID provided does not exist in our records.

HTTP Status Code: 403

NotAuthorized

You do not have permission to perform this action.

HTTP Status Code: 400

OptInRequired

The AWS access key ID needs a subscription for the service.

HTTP Status Code: 403

RequestExpired

The request reached the service more than 15 minutes after the date stamp on the request or more than 15 minutes after the request expiration date (such as for pre-signed URLs), or the date stamp on the request is more than 15 minutes in the future.

HTTP Status Code: 400

ServiceUnavailable

The request has failed due to a temporary failure of the server.

HTTP Status Code: 503

ThrottlingException

The request was denied due to request throttling.

HTTP Status Code: 400

ValidationError

The input fails to satisfy the constraints specified by an AWS service.

HTTP Status Code: 400