

Amazon Athena using AWS CLI

Firstly, need to configure the AWS CLI with appropriate user with AmazonAthenaFullAccess permission.

To list the data base using AWS CLI command:

aws athena list-databases --catalog-name AwsDataCatalog

```
pchra@DESKTOP-06V8C9M MINGW64 ~/Desktop (master)
$ aws configure
AWS Access Key ID [*****ANX6]: AKIARHAWEQ3PCD5SANX6
AWS Secret Access Key [*****d/na]: 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
Default region name [us-east-2]:
Default output format [None]:

pchra@DESKTOP-06V8C9M MINGW64 ~/Desktop (master)
$ aws athena list-databases --catalog-name AwsDataCatalog
{
  "DatabaseList": [
    {
      "Name": "default"
    },
    {
      "Name": "myreatil"
    },
    {
      "Name": "retail_db"
    }
  ]
}
```

Athena Workgroups using AWS CLI

```
pchra@DESKTOP-06V8C9M MINGW64 ~/Desktop (master)
$ aws athena list-work-groups
{
  "WorkGroups": [
    {
      "Name": "primary",
      "State": "ENABLED",
      "Description": "",
      "CreationTime": "2022-02-26T15:44:49.218000-06:00",
      "EngineVersion": {
        "SelectedEngineVersion": "AUTO",
        "EffectiveEngineVersion": "Athena engine version 2"
      }
    }
  ]
}
```

Configure query result for the workgroup (go to athena > workflows there we need to configure the s3 bucket for the query results)

Validate whether the query can be executed using command line interface

```
pchra@DESKTOP-06V8C9M MINGW64 ~/Desktop (master)
$ aws athena start-query-execution --query-string "select count(1) from myretail.orders"
{
  "QueryExecutionId": "26d9f254-20ec-4d36-b545-fcdcf1b88768"
}

pchra@DESKTOP-06V8C9M MINGW64 ~/Desktop (master)
$
```

Run Athena Queries using AWS CLI

```
aws athena start-query-execution --query-string "SELECT * FROM retail_db.orders LIMIT 5"
{
  "QueryExecutionId": "acaf0d7c-1f80-44b3-a65d-0a3da2b3dbc6"
}

pchra@DESKTOP-06V8C9M MINGW64 ~/Desktop (master)
$ aws athena get-query-execution --query-execution-id
usage: aws [options] <command> [<subcommand> ...] [parameters]
to see help text, you can run:

    aws help
    aws <command> help
    aws <command> <subcommand> help

aws.exe: error: argument --query-execution-id: expected one argument

pchra@DESKTOP-06V8C9M MINGW64 ~/Desktop (master)
$ aws athena get-query-execution --query-execution-id acaf0d7c-1f80-44b3-a65d-0a3da2b3dbc6
{
  "QueryExecution": {
    "QueryExecutionId": "acaf0d7c-1f80-44b3-a65d-0a3da2b3dbc6",
    "Query": "SELECT * FROM retail_db.orders LIMIT 5",
    "StatementType": "DML",
    "ResultConfiguration": {
      "OutputLocation": "s3://itvathena1/wgprimary1/acaf0d7c-1f80-44b3-a65d-0a3da2b3dbc6.csv"
    },
    "QueryExecutionContext": {},
    "Status": {
      "State": "SUCCEEDED",
      "SubmissionDateTime": "2022-03-10T23:11:20.948000-06:00",
      "CompletionDateTime": "2022-03-10T23:11:21.711000-06:00"
    },
    "Statistics": {
      "EngineExecutionTimeInMillis": 604,
      "DataScannedInBytes": 788848,
      "TotalExecutionTimeInMillis": 763,
      "QueryQueueTimeInMillis": 111,
      "QueryPlanningTimeInMillis": 96,
      "ServiceProcessingTimeInMillis": 48
    },
    "WorkGroup": "primary",
    "EngineVersion": {
      "SelectedEngineVersion": "AUTO",
      "EffectiveEngineVersion": "Athena engine version 2"
    }
  }
}
```

```
pchra@DESKTOP-06V8C9M MINGW64 ~/Desktop (master)
$ aws s3 cp s3://itvathena1/wgprimary1/acaf0d7c-1f80-44b3-a65d-0a3da2b3dbc6.csv .
download: s3://itvathena1/wgprimary1/acaf0d7c-1f80-44b3-a65d-0a3da2b3dbc6.csv to .\acaf0d7c-1f80-44b3-a65d-0a3da2b3dbc6.csv

pchra@DESKTOP-06V8C9M MINGW64 ~/Desktop (master)
$ view acaf0d7c-1f80-44b3-a65d-0a3da2b3dbc6.csv

pchra@DESKTOP-06V8C9M MINGW64 ~/Desktop (master)
```

```
MINGW64:/c/Users/pchra/Desktop
order_id","order_date","order_customer_id","order_status"
1","2013-07-25 00:00:00.0","11599","CLOSED"
2","2013-07-25 00:00:00.0","256","PENDING_PAYMENT"
3","2013-07-25 00:00:00.0","12111","COMPLETE"
4","2013-07-25 00:00:00.0","8827","CLOSED"
5","2013-07-25 00:00:00.0","11318","COMPLETE"
```

Get Metadata for all the tables from the selected database using AWSCLI

```
chra@DESKTOP-06V8C9M MINGW64 ~/Desktop (master)
aws athena list-table-metadata --catalog-name AwsDatacatalog --database-name myretail

{"TableMetadataList": [
  {
    "Name": "orders",
    "CreateTime": "2022-03-10T22:38:30-06:00",
    "LastAccessTime": "1969-12-31T18:00:00-06:00",
    "TableType": "EXTERNAL_TABLE",
    "Columns": [
      {
        "Name": "order_id",
        "Type": "bigint"
      },
      {
        "Name": "order_date",
        "Type": "string"
      },
      {
        "Name": "order_customer_id",
        "Type": "bigint"
      },
      {
        "Name": "order_status",
        "Type": "string"
      }
    ]
  },
  "PartitionKeys": [],
  "Parameters": {
    "CrawlerSchemaDeserializerVersion": "1.0",
    "CrawlerSchemaSerializerVersion": "1.0",
    "EXTERNAL": "TRUE",
    "UPDATED_BY_CRAWLER": "retail_crawler",
    "areColumnsQuoted": "false",
    "averageRecordSize": "39",
    "classification": "csv",
    "columnsOrdered": "true",
    "compressionType": "none",
    "delimiter": ",",
    "inputformat": "org.apache.hadoop.mapred.TextInputFormat",
    "location": "s3://iretail-db/orders",
    "objectCount": "1",
    "outputformat": "org.apache.hadoop.hive ql.io.HiveIgnoreKeyTextOutputFormat",
```

Get Metadata for particular Table from database.

```
chra@DESKTOP-06V8C9M MINGW64 ~/Desktop (master)
aws athena get-table-metadata --catalog-name AwsDataCatalog --database-name myretail --table-name orders

{"TableMetadata": {
  "Name": "orders",
  "CreateTime": "2022-03-10T22:38:30-06:00",
  "LastAccessTime": "1969-12-31T18:00:00-06:00",
  "TableType": "EXTERNAL_TABLE",
  "Columns": [
    {
      "Name": "order_id",
      "Type": "bigint"
    },
    {
      "Name": "order_date",
      "Type": "string"
    },
    {
      "Name": "order_customer_id",
      "Type": "bigint"
    },
    {
      "Name": "order_status",
      "Type": "string"
    }
  ],
  "PartitionKeys": [],
  "Parameters": {
    "CrawlerSchemaSerializerVersion": "1.0",
    "CrawlerSchemaSerializerVersion": "1.0",
    "EXTERNAL": "TRUE",
    "UPDATED_BY_CRAWLER": "retail_crawler",
    "arcColumnsQuoted": "false",
    "averageRecordSize": "39",
    "classification": "csv",
    "columnsOrdered": "true",
    "compressionType": "none",
    "delimiter": ",",
    "inputFormat": "org.apache.hadoop.mapred.TextInputFormat",
    "location": "s3://iretail-db/orders",
    "objectCount": "1",
    "outputFormat": "org.apache.hadoop.hive.q1.io.HiveIgnoreKeyTextOutputFormat",
    "recordCount": "78687",
    "serde.param.field.delim": "",
    "serde.param.serialization.format": "",
    "serde.serialization.lib": "org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe",
  }
}
```

Drop table using AWS CLI

```
chra@DESKTOP-06V8C9M MINGW64 ~/Desktop (master)
aws athena start-query-execution --query-string "DROP TABLE myretail.orders"

{"QueryExecutionId": "4862b585-15a9-49ba-b919-194371a3fad0"}

chra@DESKTOP-06V8C9M MINGW64 ~/Desktop (master)
aws athena get-query-execution --query-execution-id 4862b585-15a9-49ba-b919-194371a3fad0

{"QueryExecution": {
  "QueryExecutionId": "4862b585-15a9-49ba-b919-194371a3fad0",
  "Query": "DROP TABLE myretail.orders",
  "StatementType": "DDL",
  "ResultConfiguration": {
    "OutputLocation": "s3://itvathena1/wgprimary1/4862b585-15a9-49ba-b919-194371a3fad0.txt"
  },
  "QueryExecutionContext": {},
  "Status": {
    "State": "SUCCEEDED",
    "SubmissionDateTime": "2022-03-10T23:43:15.383000-06:00",
    "CompletionDateTime": "2022-03-10T23:43:16.038000-06:00"
  },
  "Statistics": {
    "EngineExecutionTimeInMillis": 585,
    "DataScannedInBytes": 0,
    "TotalExecutionTimeInMillis": 655,
    "QueryQueueTimeInMillis": 45,
    "ServiceProcessingTimeInMillis": 25
  },
  "WorkGroup": "primary",
  "EngineVersion": {
    "SelectedEngineVersion": "AUTO",
    "EffectiveEngineVersion": "Athena engine version 2"
  }
}

chra@DESKTOP-06V8C9M MINGW64 ~/Desktop (master)
aws s3 ls s3://iretail-db

An error occurred (NoSuchBucket) when calling the ListObjectsV2 operation: The specified bucket does not exist

chra@DESKTOP-06V8C9M MINGW64 ~/Desktop (master)
aws s3 ls s3://iretail-db

chra@DESKTOP-06V8C9M MINGW64 ~/Desktop (master)
```

Run Create Table AS SELECT under Athena using AWS CLI:

```
chra@DESKTOP-06V8C9M MINGW64 ~/Desktop (master)
aws athena start-query-execution --query-string "CREATE TABLE myretail.order_items WITH ( format = 'TEXTFILE', external_location = 's3://iretail-db/orders_items/', field_delimiter = ',' ) AS SELECT * FROM retail_db.order_items"

"QueryExecutionId": "425476ce-b36d-447c-98c6-537611c6df2d"

chra@DESKTOP-06V8C9M MINGW64 ~/Desktop (master)
aws athena get-query-execution --query-execution-id 425476ce-b36d-447c-98c6-537611c6df2d

{
  "QueryExecution": {
    "QueryExecutionId": "425476ce-b36d-447c-98c6-537611c6df2d",
    "Query": "CREATE TABLE myretail.order_items WITH ( format = 'TEXTFILE', external_location = 's3://iretail-db/orders_items/', field_delimiter = ',' ) AS SELECT * FROM retail_db.order_items",
    "StatementType": "DDL",
    "ResultConfiguration": {
      "OutputLocation": "s3://itvathena1/wgprimary1/tables/425476ce-b36d-447c-98c6-537611c6df2d"
    },
    "QueryExecutionContext": {},
    "Status": {
      "State": "SUCCEEDED",
      "SubmissionDateTime": "2022-03-10T23:54:11.621000-06:00",
      "CompletionDateTime": "2022-03-10T23:54:13.407000-06:00"
    },
    "Statistics": {
      "EngineExecutionTimeInMillis": 1602,
      "DataScannedInBytes": 5581078,
      "DataManifestLocation": "s3://itvathena1/wgprimary1/tables/425476ce-b36d-447c-98c6-537611c6df2d-manifest.csv",
      "TotalExecutionTimeInMillis": 1786,
      "QueryQueueTimeInMillis": 116,
      "QueryPlanningTimeInMillis": 106,
      "ServiceProcessingTimeInMillis": 68
    },
    "WorkGroup": "primary",
    "EngineVersion": {
      "SelectedEngineVersion": "AUTO",
      "EffectiveEngineVersion": "Athena engine version 2"
    }
  }
}

chra@DESKTOP-06V8C9M MINGW64 ~/Desktop (master)
```

View data

```
chra@DESKTOP-06V8C9M MINGW64 ~/Desktop (master)
aws athena start-query-execution --query-string "SELECT * FROM retail_db.order_items limit 10"

"QueryExecutionId": "93baea55-b5de-4c1a-993f-b79522228908"

chra@DESKTOP-06V8C9M MINGW64 ~/Desktop (master)
aws athena get-query-execution --query-execution-id 93baea55-b5de-4c1a-993f-b79522228908

{
  "QueryExecution": {
    "QueryExecutionId": "93baea55-b5de-4c1a-993f-b79522228908",
    "Query": "SELECT * FROM retail_db.order_items limit 10",
    "StatementType": "DML",
    "ResultConfiguration": {
      "OutputLocation": "s3://itvathena1/wgprimary1/93baea55-b5de-4c1a-993f-b79522228908.csv"
    },
    "QueryExecutionContext": {},
    "Status": {
      "State": "SUCCEEDED",
      "SubmissionDateTime": "2022-03-10T23:58:59.268000-06:00",
      "CompletionDateTime": "2022-03-10T23:59:00.062000-06:00"
    },
    "Statistics": {
      "EngineExecutionTimeInMillis": 579,
      "DataScannedInBytes": 591726,
      "TotalExecutionTimeInMillis": 794,
      "QueryQueueTimeInMillis": 130,
      "QueryPlanningTimeInMillis": 67,
      "ServiceProcessingTimeInMillis": 85
    },
    "WorkGroup": "primary",
    "EngineVersion": {
      "SelectedEngineVersion": "AUTO",
      "EffectiveEngineVersion": "Athena engine version 2"
    }
  }
}

chra@DESKTOP-06V8C9M MINGW64 ~/Desktop (master)
aws s3 cp s3://itvathena1/wgprimary1/93baea55-b5de-4c1a-993f-b79522228908.csv .
download: s3://itvathena1/wgprimary1/93baea55-b5de-4c1a-993f-b79522228908.csv to .\93baea55-b5de-4c1a-993f-b79522228908.csv

chra@DESKTOP-06V8C9M MINGW64 ~/Desktop (master)
cat 93baea55-b5de-4c1a-993f-b79522228908.csv
```

\\192.168.1.104\c\USCIS\plm\plm\Desktop

```
col0","col1","col2","col3","col4","col5"
1","1","957","1","299.98","299.98"
2","2","1073","1","199.99","199.99"
3","2","502","5","250.0","50.0"
4","2","403","1","129.99","129.99"
5","4","897","2","49.98","24.99"
6","4","365","5","299.95","59.99"
7","4","502","3","150.0","50.0"
8","4","1014","4","199.92","49.98"
9","5","957","1","299.98","299.98"
10","5","365","5","299.95","59.99"
```