

# EXP NO : 6

## Querying Data in S3 with Amazon Athena

NAME: SHUBHRA KUMARI

REG. NO.: RA2011028010093

**Aim:** Querying Data in S3 with Amazon Athena

Step 1: Go to buckets and create 2 buckets.

Amazon S3

**Buckets**

Access Points

Object Lambda Access Points

Multi-Region Access Points

Batch Operations

Access analyzer for S3

Block Public Access settings for this account

Storage Lens

Dashboards

AWS Organizations settings

Feature spotlight 3

AWS Marketplace for S3

Feedback Looking for language selection? Find it in the new Unified Settings

© 2022, Amazon Internet Services Private Ltd. or its affiliates. Privacy Terms Cookie preferences

Successfully created bucket "bucketawsexp62"

To upload files and folders, or to configure additional bucket settings choose [View details](#).

Amazon S3 > Buckets

**Account snapshot**

Storage lens provides visibility into storage usage and activity trends. [Learn more](#)

[View Storage Lens dashboard](#)

**Buckets (2)** [Info](#)

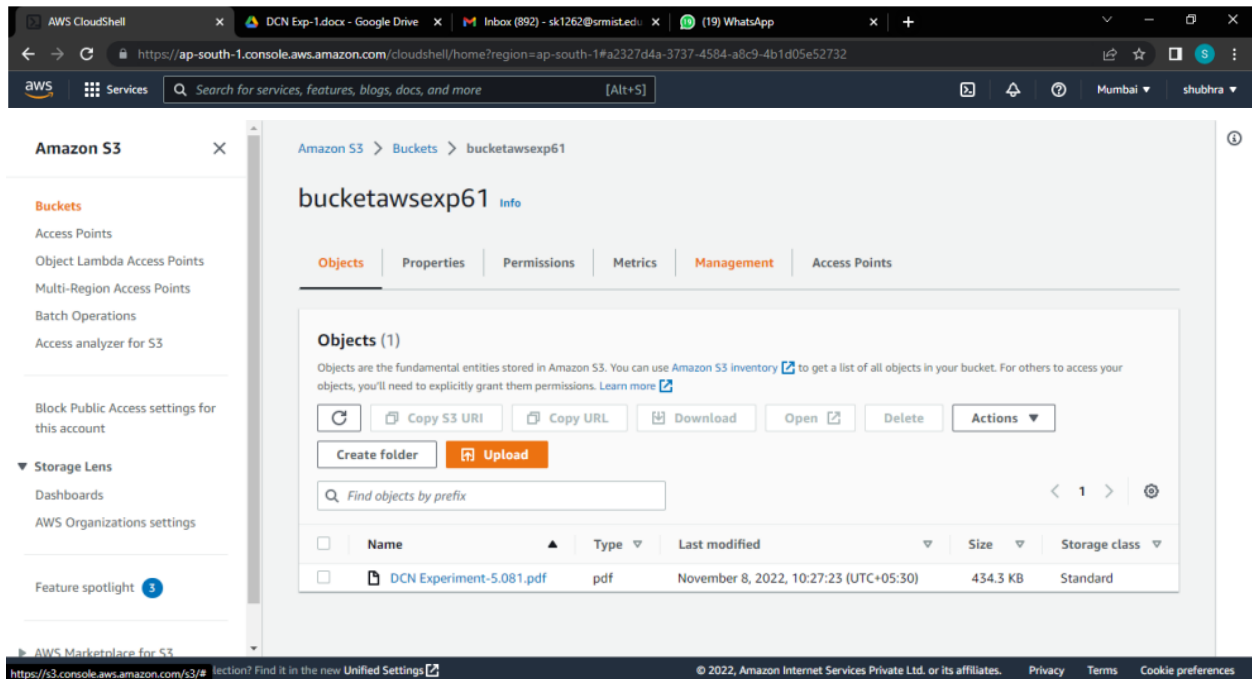
Buckets are containers for data stored in S3. [Learn more](#)

[Find buckets by name](#)

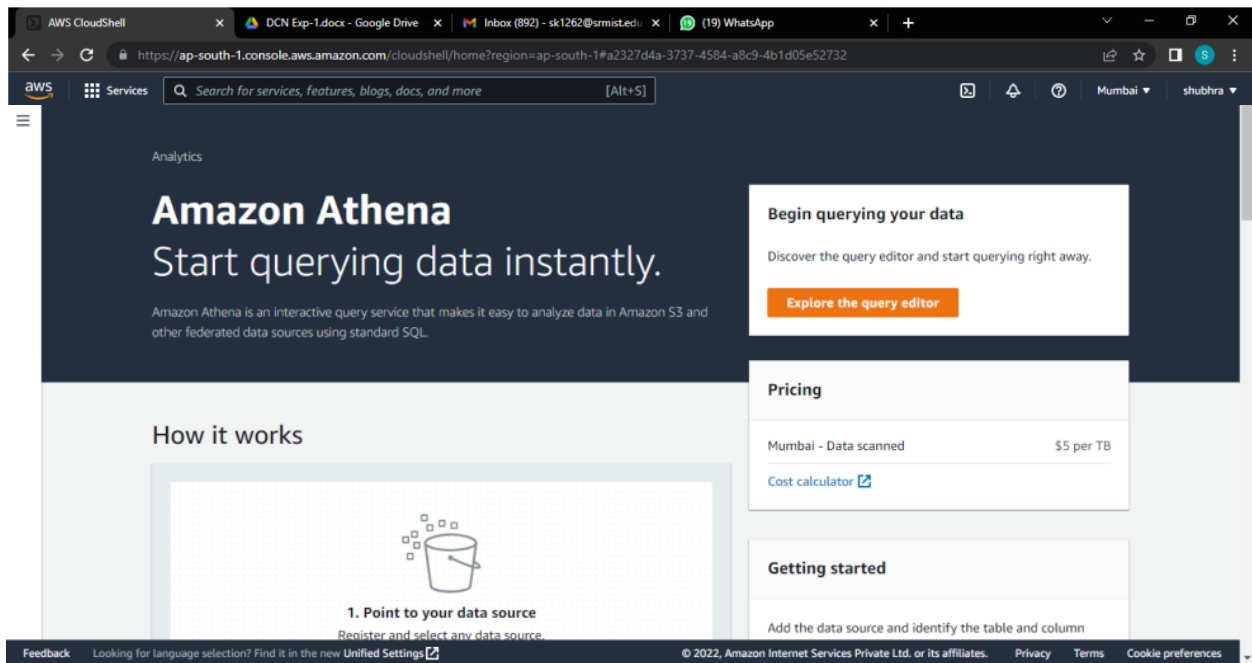
[Copy ARN](#) [Empty](#) [Delete](#) [Create bucket](#)

	Name	AWS Region	Access	Creation date
<input type="radio"/>	bucketawsexp61	Asia Pacific (Mumbai) ap-south-1	Bucket and objects not public	November 8, 2022, 10:25:51 (UTC+05:30)
<input type="radio"/>	bucketawsexp62	Asia Pacific (Mumbai) ap-south-1	Bucket and objects not public	November 8, 2022, 10:26:13 (UTC+05:30)

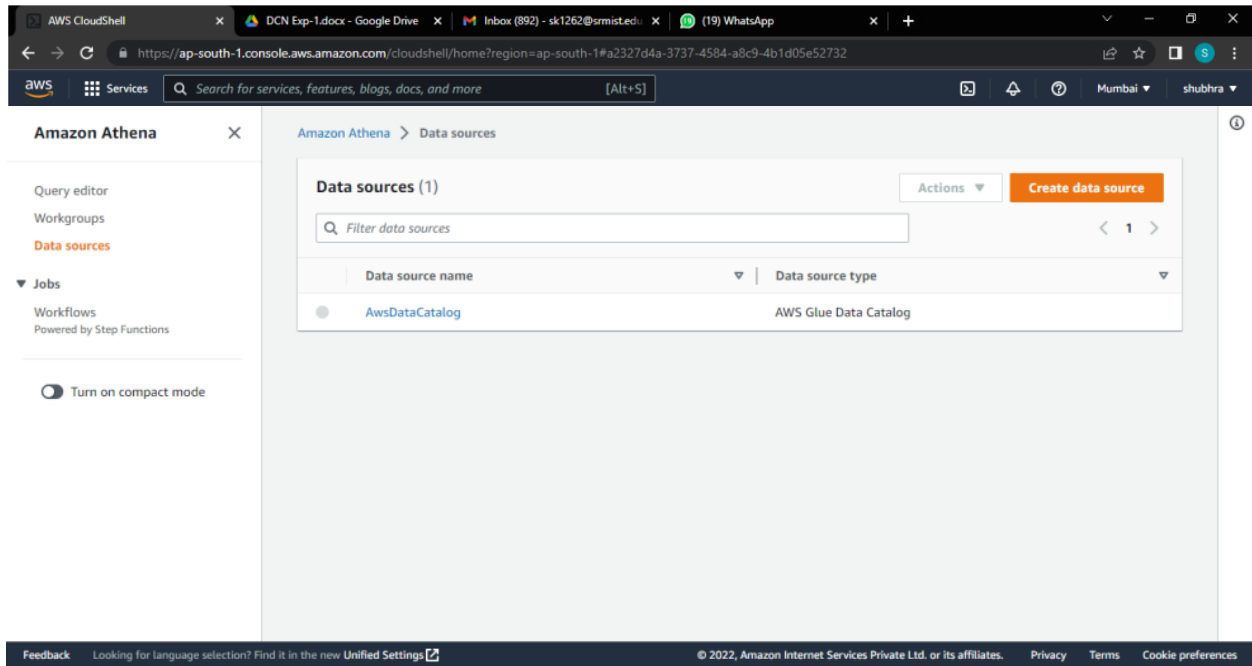
Step 2 : After clicking onto the bucket add files to it.



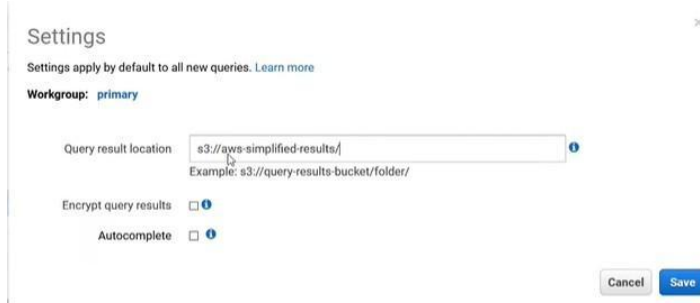
Step 3 : Now go to Amazon athena.



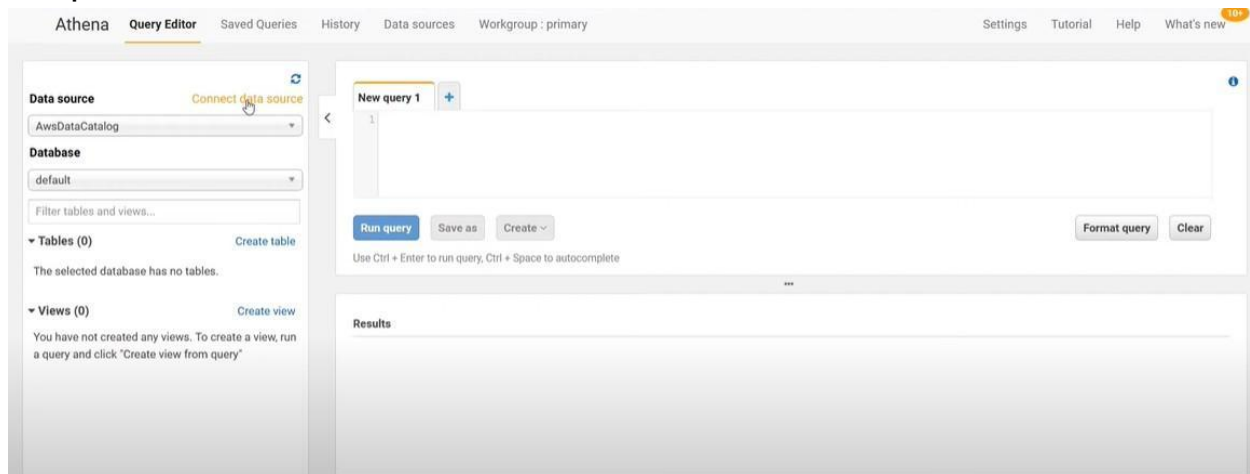
Step 4: Select AwsDataCatalog in the left side which is present in the data source tab.



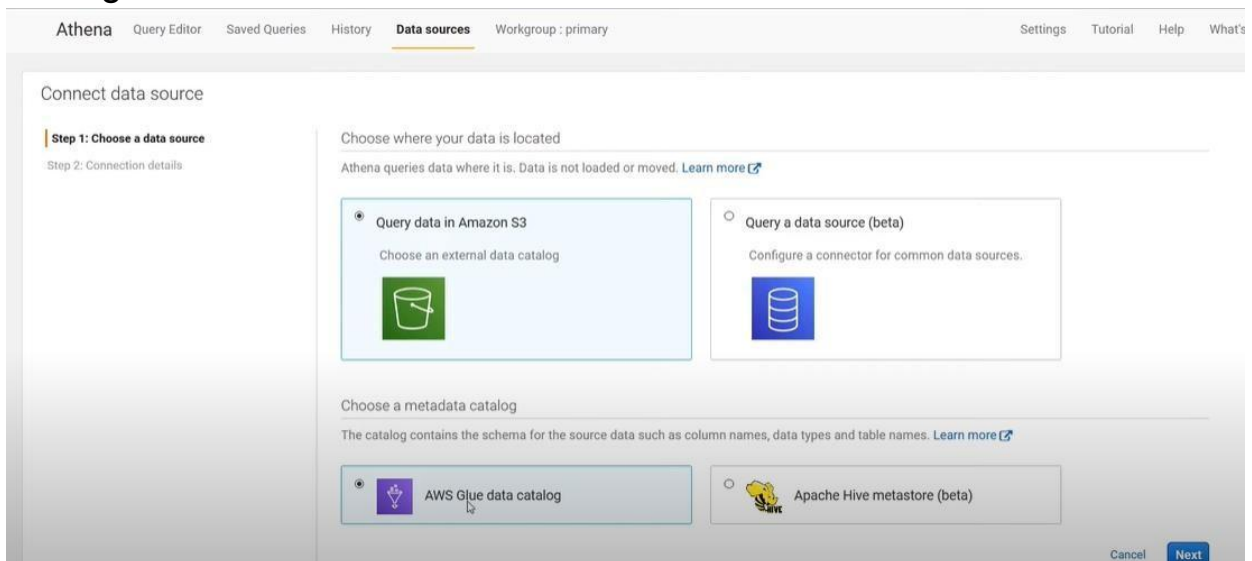
Step 5: After that go to settings and specify an output path.



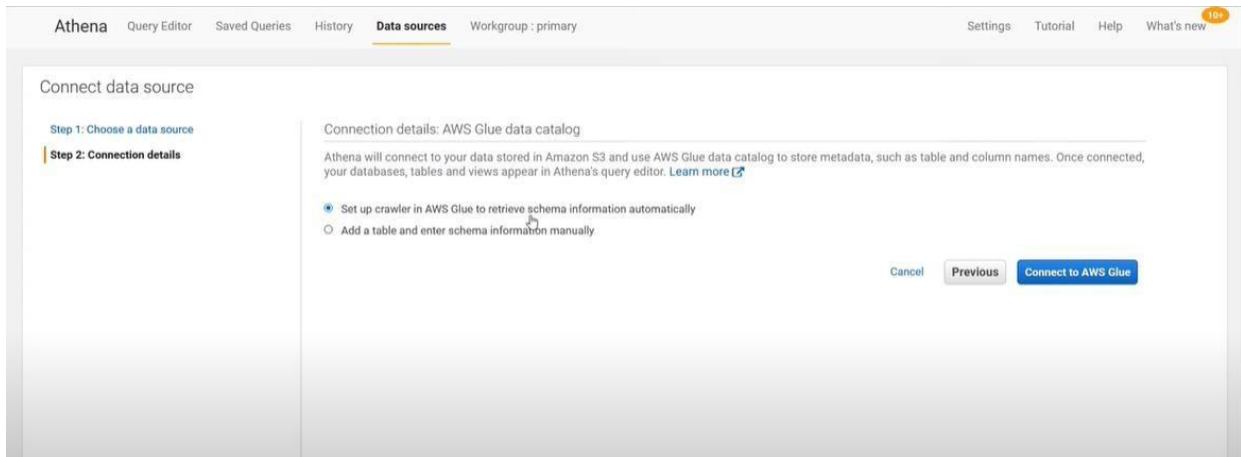
Step 6: Click on connect data source.



Step 7: After clicking choose a query in amazon s3 and Aws glue data catalog.



Step 8: Click on next and select setup a crawler in AWS glue to retrieve schema information automatically.



Athena Query Editor Saved Queries History **Data sources** Workgroup : primary Settings Tutorial Help What's new 164

Connect data source

Step 1: Choose a data source

**Step 2: Connection details**

Connection details: AWS Glue data catalog

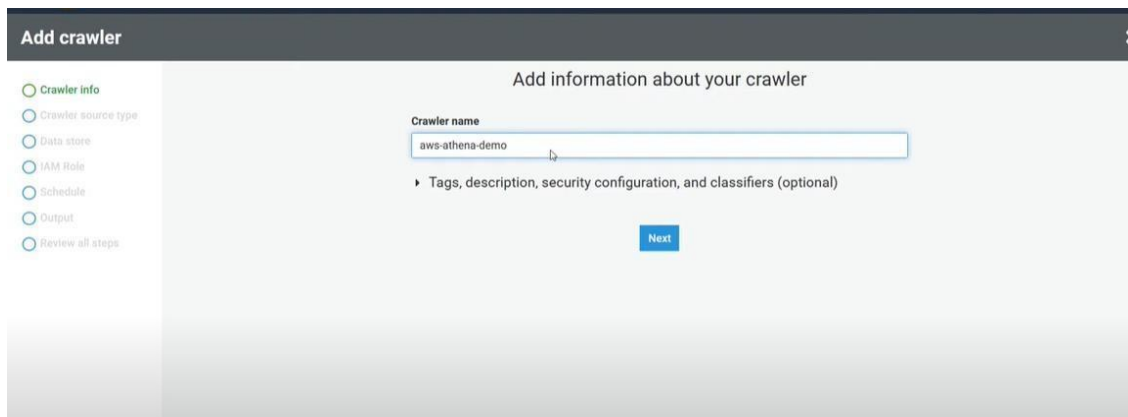
Athena will connect to your data stored in Amazon S3 and use AWS Glue data catalog to store metadata, such as table and column names. Once connected, your databases, tables and views appear in Athena's query editor. [Learn more](#)

☒ Set up crawler in AWS Glue to retrieve schema information automatically

☐ Add a table and enter schema information manually

Cancel Previous **Connect to AWS Glue**

Step 9: After selecting that it will redirect to a new page and add crawler and follow below steps to add a new crawler after setting up click on finish.



**Add crawler**

Add information about your crawler

**Crawler info**

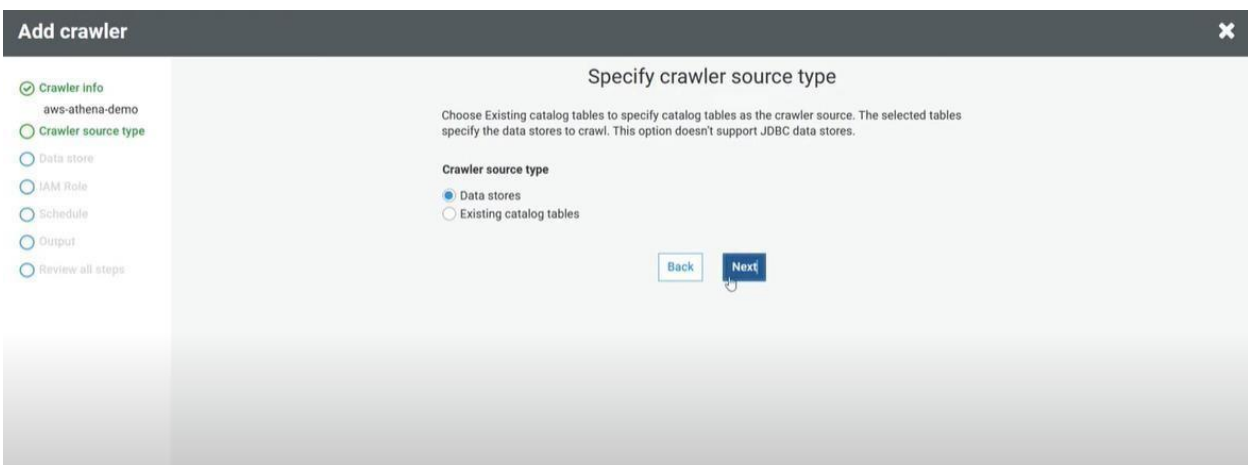
- ☒ Crawler info
- ☐ Crawler source type
- ☐ Data store
- ☐ IAM Role
- ☐ Schedule
- ☐ Output
- ☐ Review all steps

Crawler name

aws-athena-demo

Tags, description, security configuration, and classifiers (optional)

**Next**



**Add crawler**

Specify crawler source type

Choose Existing catalog tables to specify catalog tables as the crawler source. The selected tables specify the data stores to crawl. This option doesn't support JDBC data stores.

**Crawler source type**

☒ Data stores

☐ Existing catalog tables

**Back** **Next**

Add crawler

Crawler info

aws-athena-demo

Crawler source type

Data stores

Data store

IAM Role

Schedule

Output

Review all steps

Add a data store

Choose a data store

S3

Crawl data in

Specified path in my account

Specified path in another account

Include path

s3://bucket/prefix/object

Exclude patterns (optional)

Back

Next

Add crawler

Crawler info

aws-athena-demo

Crawler source type

Data stores

Data store

S3: s3://aws-simplifi...

IAM Role

Schedule

Output

Review all steps

Choose an IAM role

The IAM role allows the crawler to run and access your Amazon S3 data stores. [Learn more](#)

Update a policy in an IAM role

Choose an existing IAM role

Create an IAM role

IAM role ⓘ

AWSGlueServiceRole- demo

To create an IAM role, you must have **CreateRole**, **CreatePolicy**, and **AttachRolePolicy** permissions. Create an IAM role named **"AWSGlueServiceRole-rolename"** and attach the AWS managed policy, **AWSGlueServiceRole**, plus an inline policy that allows read access to:

- s3://aws-simplified-athena-demo/

You can also create an IAM role on the [IAM console](#).

Back

Next

Crawler info

aws-athena-demo

Crawler source type

Data stores

Data store

S3: s3://aws-simplifi...

IAM Role

am:aws:iam::398447858632:role/service-role/AWSGlueServiceRole-demo

Schedule

Output

Review all steps

Create a schedule for this crawler

Frequency

Run on demand

Back

Next

Add crawler

Crawler info

aws-athena-demo

Crawler source type

Data stores

Data store

S3: s3://aws-simplifi...

IAM Role

am:aws:iam::398447858632:role/service-role/AWSGlueServiceRole-demo

Schedule

Run on demand

Output

default

Crawler info

Name

aws-athena-demo

Tags

-

IAM role

IAM role

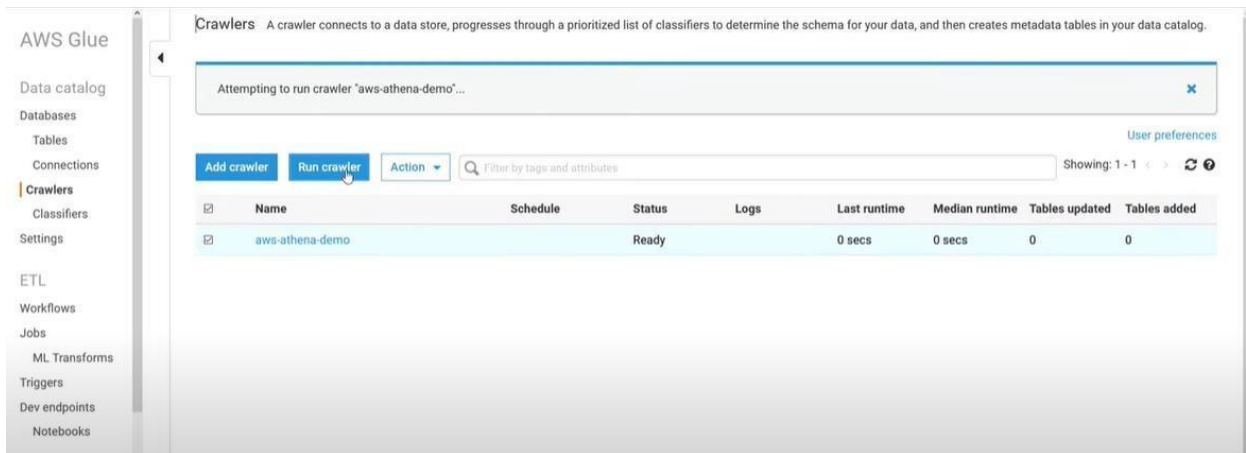
am:aws:iam::398447858632:role/service-role/AWSGlueServiceRole-demo

Schedule

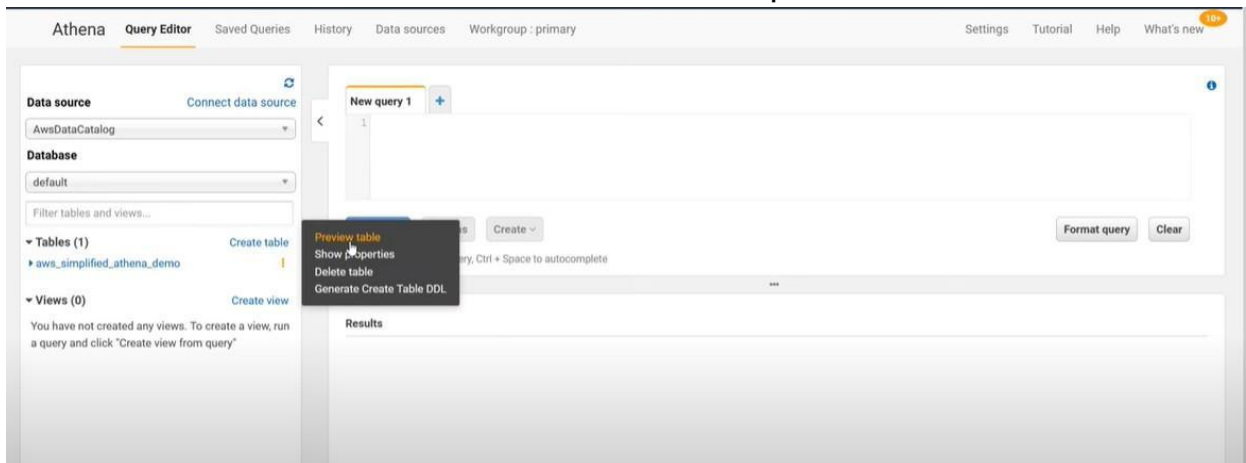
Schedule

Run on demand

Step 10: Crawler is successfully created and now click on the crawler and click run crawler.



Step 11: After running the crawler go back to athena you will see a table created on table column select that and click on preview table.



Step 12: Now the query can be executed.



**RESULT:** Querying Data in S3 with Amazon Athena is done and output is verified.