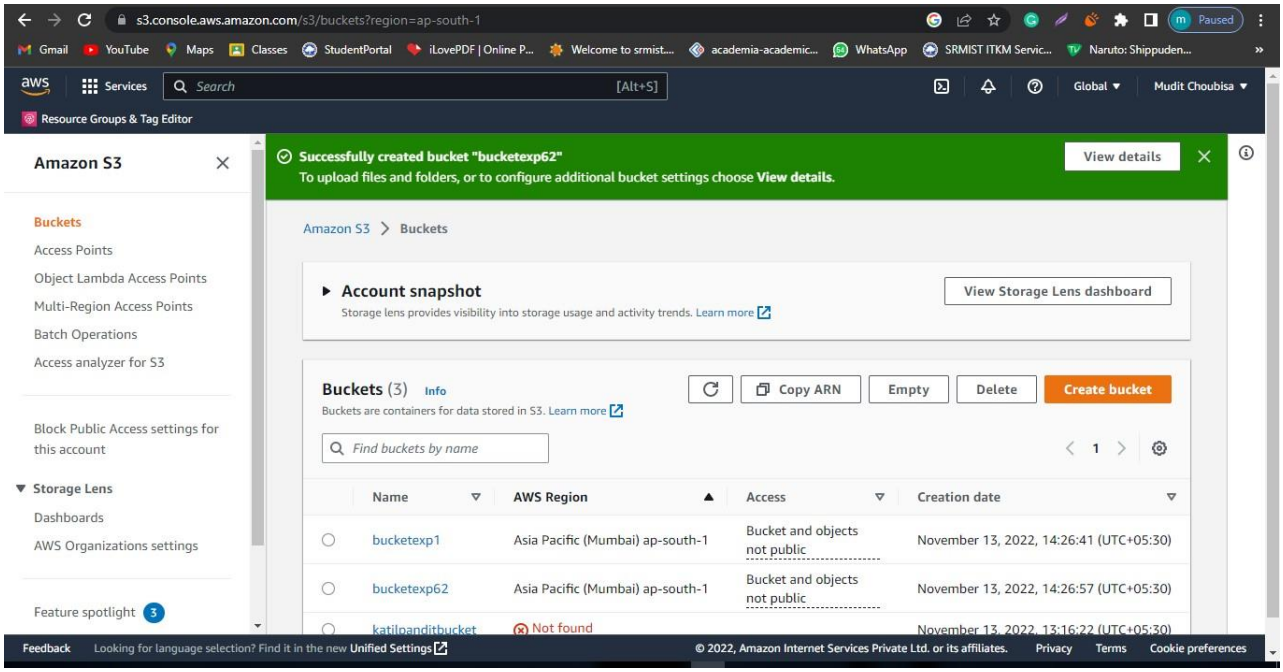
EXP NO : 6 Querying Data in S3 with Amazon Athena

NAME: PIYUSH KUMAR
REG. NO.:RA2011028010084

Aim: Querying Data in S3 with Amazon Athena

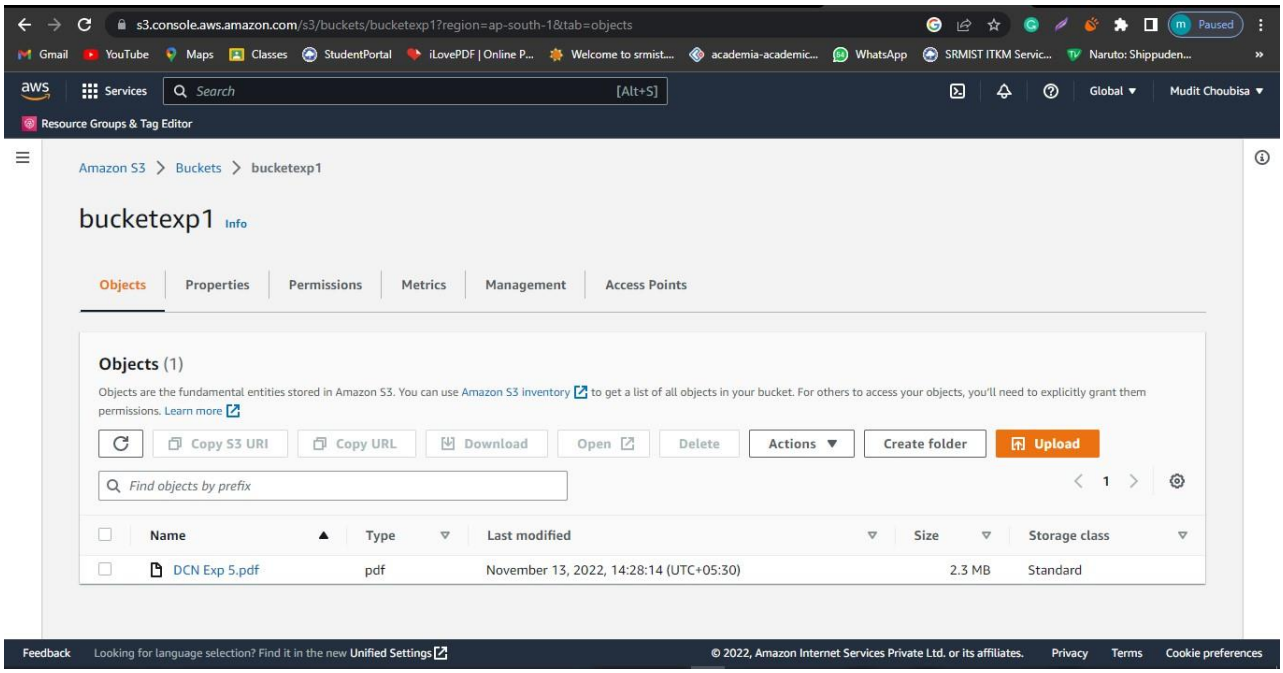
Step 1:
Go to buckets and create two buckets.



2

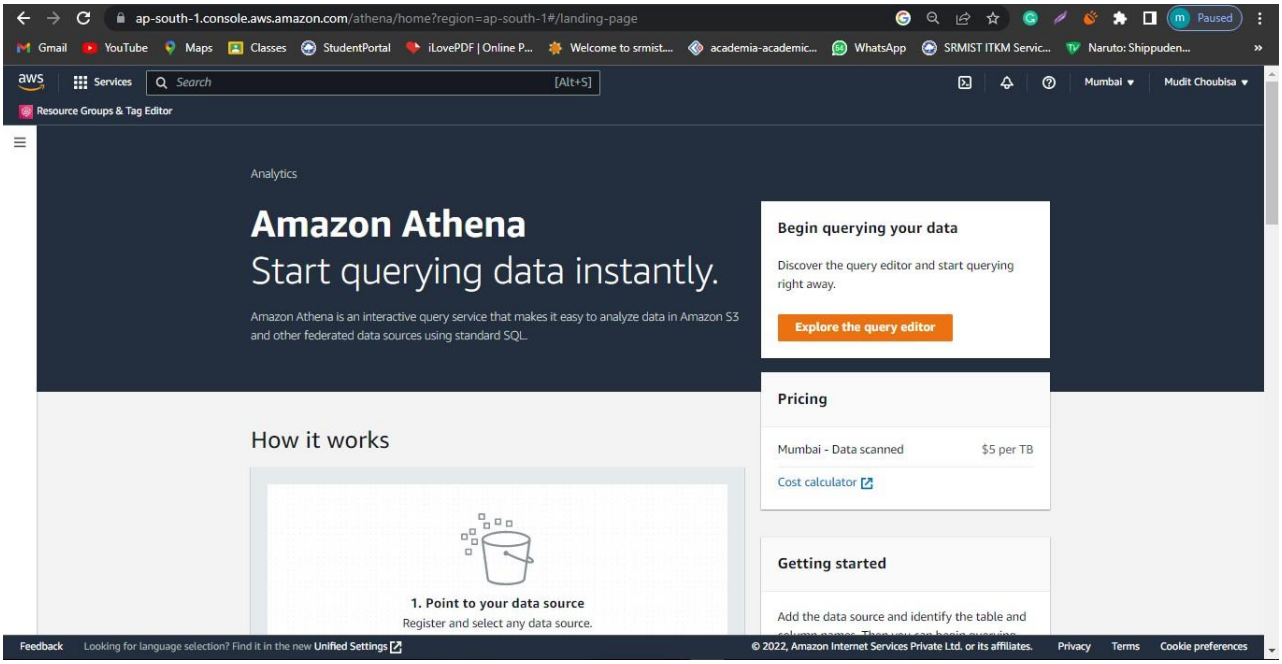
After clicking onto the bucket add files to it.

Step :


s3.console.aws.amazon.com/s3/buckets/bucketexp1?region=ap-south-1&tab=objects
Paused
Gmail
YouTube
Maps
Classes
StudentPortal
iLovePDF | Online P...
Welcome to srmist....
academia-academic...
WhatsApp
SRMIST ITKM Servic...
Naruto: Shippuden...
aws
Services
Search
[Alt+S]
Global
Mudit Choubisa
Resource Groups & Tag Editor
Amazon S3 > Buckets > bucketexp1
bucketexp1 Info
Objects
Properties
Permissions
Metrics
Management
Access Points
Objects (1)
Objects are the fundamental entities stored in Amazon S3. You can use Amazon S3 inventory to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. Learn more
Copy S3 URI
Copy URL
Download
Open
Delete
Actions
Create folder
Upload
Find objects by prefix
1
Name
Type
Last modified
Size
Storage class
DCN Exp 5.pdf
pdf
November 13, 2022, 14:28:14 (UTC+05:30)
2.3 MB
Standard
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

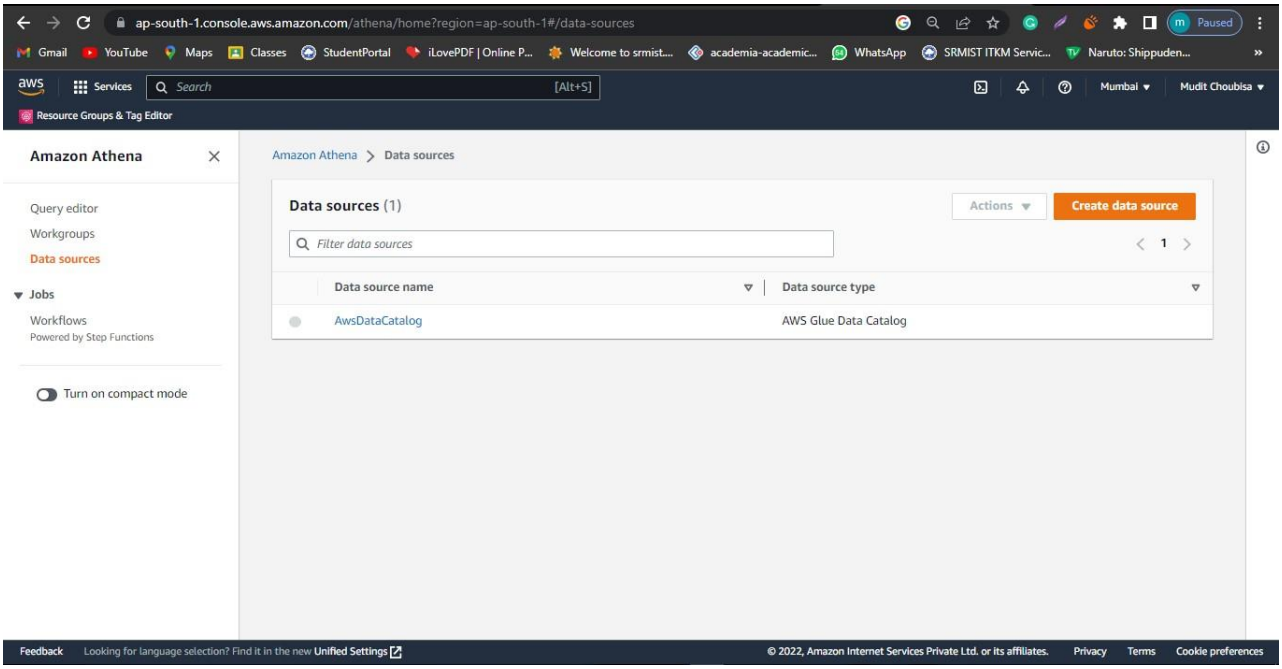
Step :

3 Now go to Amazon athena.



Step 4 :

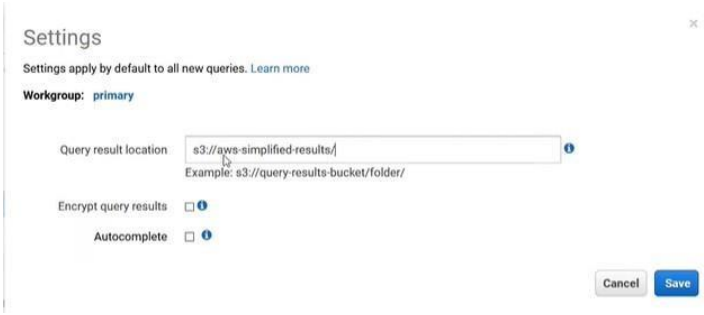
Select AwsDataCatalog in the left side which is present in the data source tab.



Step :

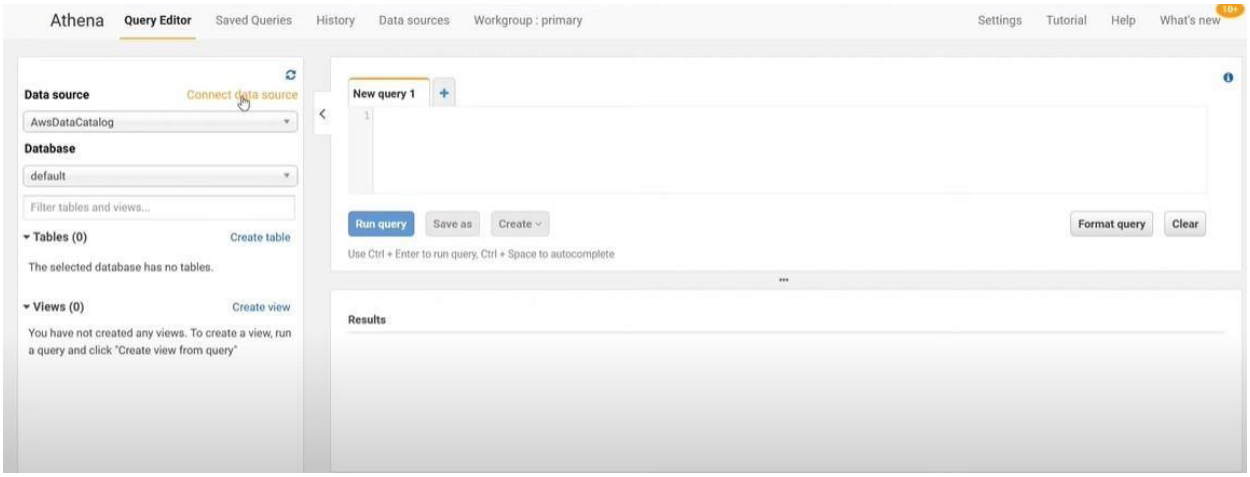
5

After that go to settings and specify an output path.



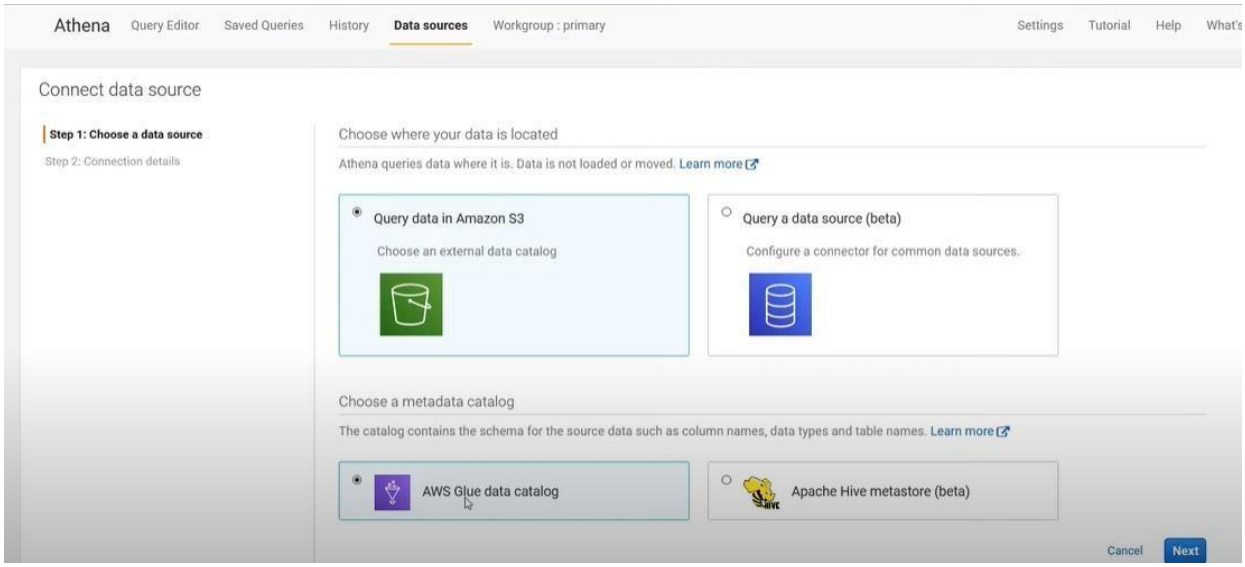
Step 6:

Click on connect data source.



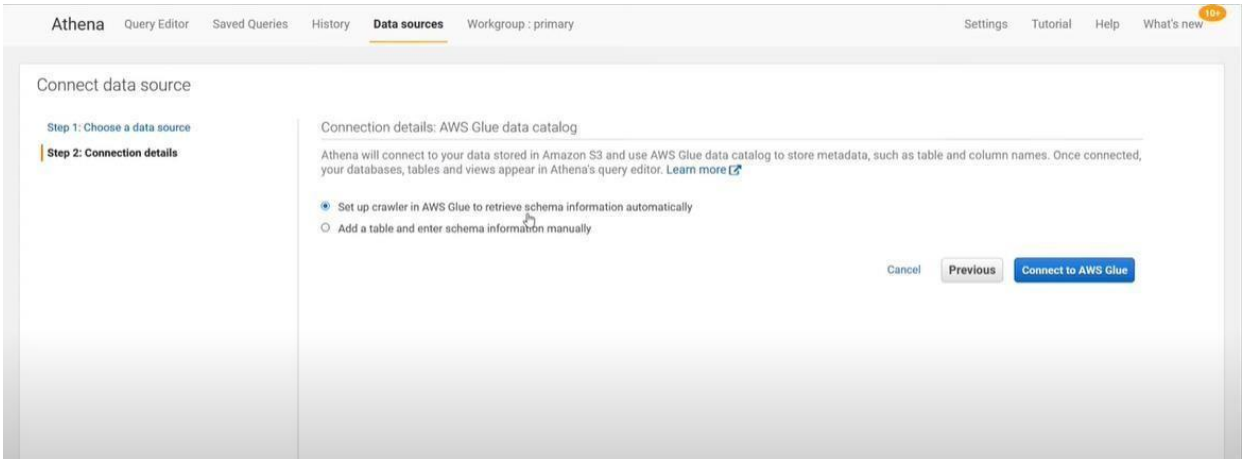
Step 7:

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



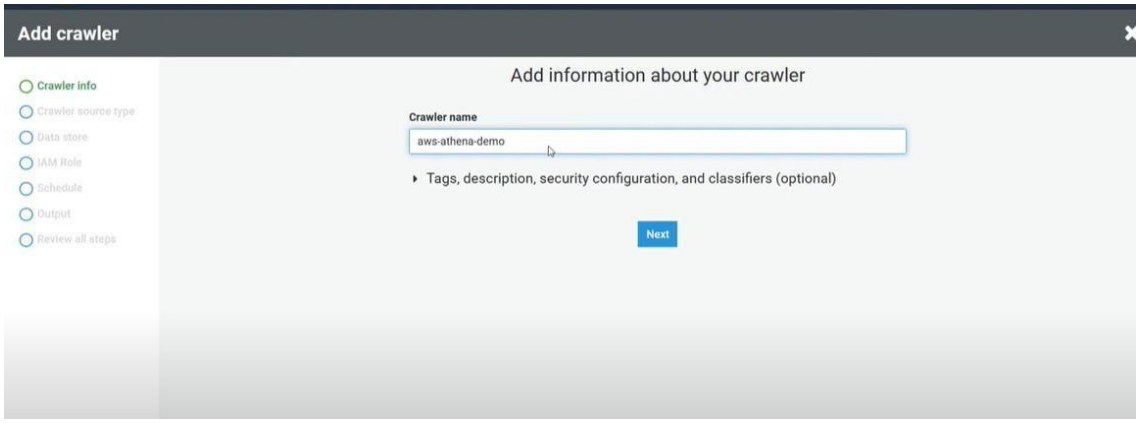
Step 8:

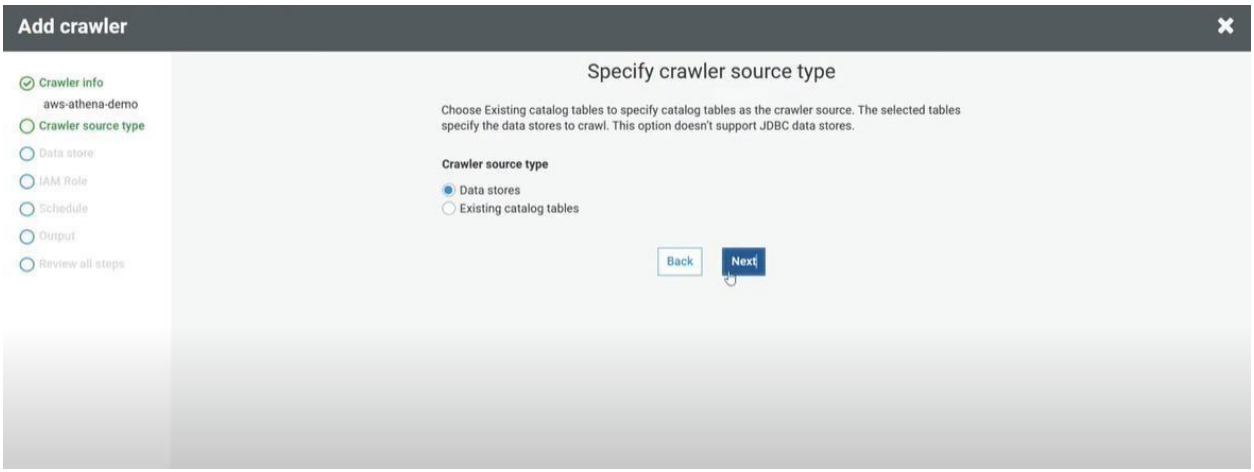
Click on next and select setup a crawler in AWS glue to retrieve schema information automatically.

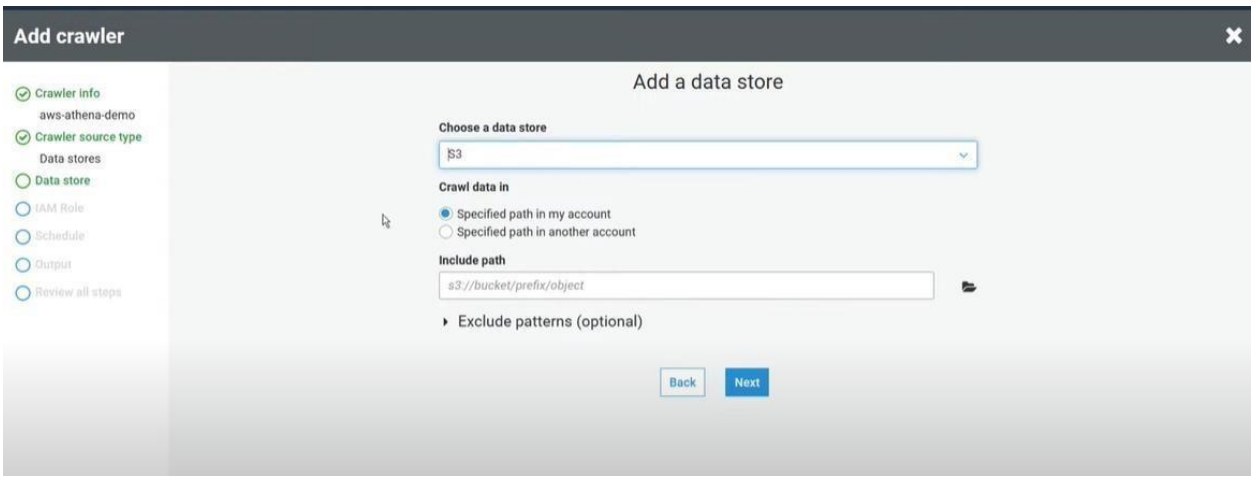


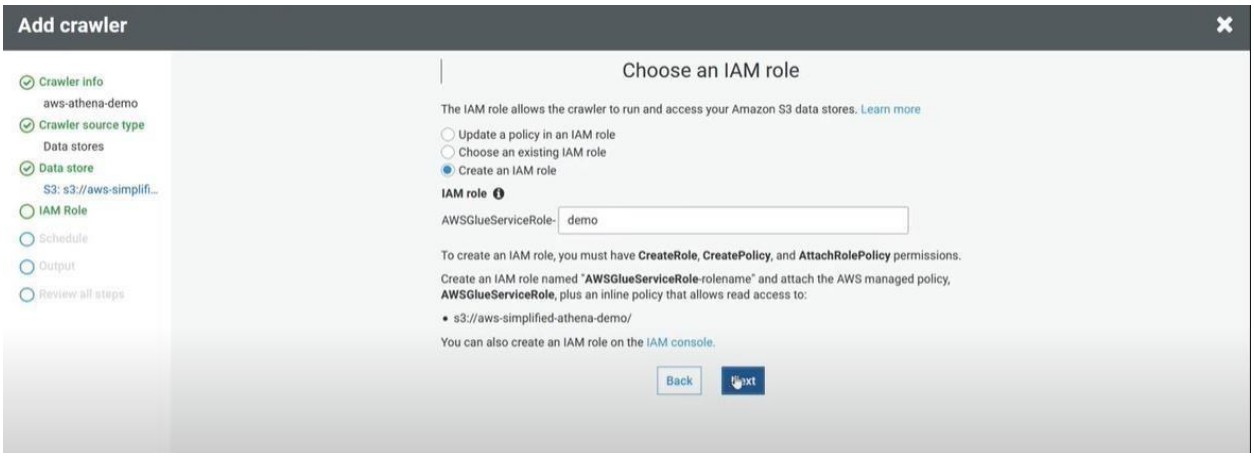
Step 9:

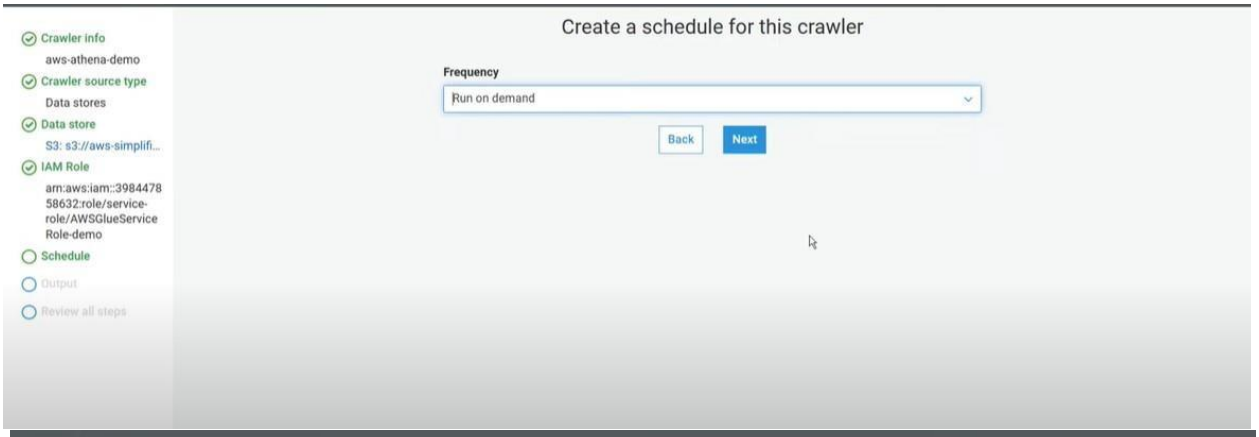
After selecting that it will redirect to a new page and add crawler and follow below steps to add a new crawler after setting up click on finish.


Add crawler
Crawler info
Crawler source type
Data store
IAM Role
Schedule
Output
Review all steps
Add information about your crawler
Crawler name
aws-athena-demo
Tags, description, security configuration, and classifiers (optional)
Next


Add crawler
Crawler info
aws-athena-demo
Crawler source type
Data store
IAM Role
Schedule
Output
Review all steps
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

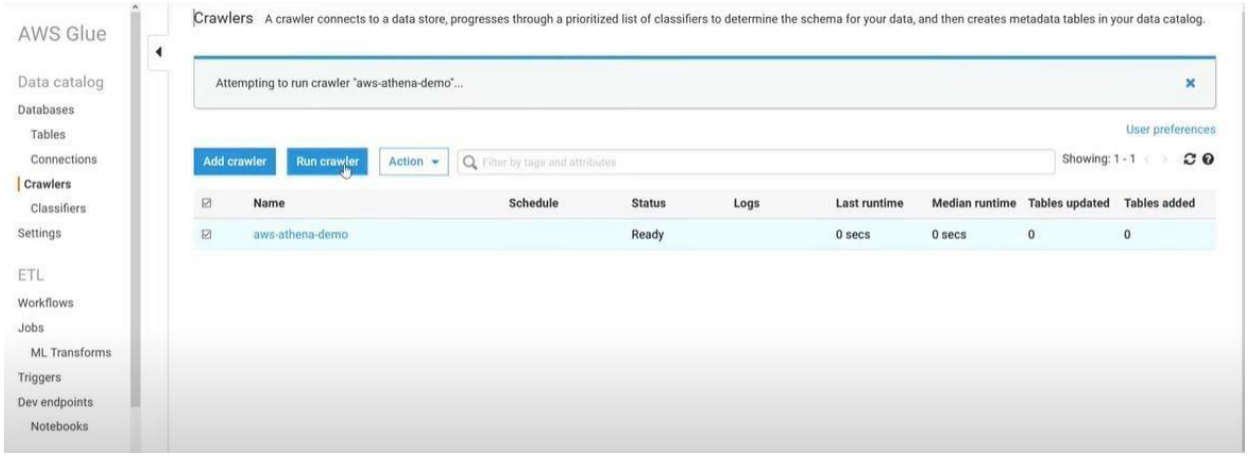

Add crawler
Crawler info
aws-athena-demo
Crawler source type
Data stores
Data store
S3: s3://aws-simplifi...
IAM Role
Schedule
Output
Review all steps
Choose an IAM role
The IAM role allows the crawler to run and access your Amazon S3 data stores. Learn more
Update a policy in an IAM role
Choose an existing 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
arn: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
arn:aws:iam::398447858632:role/service-role/AWSGlueServiceRole-demo
Schedule
Run on demand
Output
default
Review all steps
Crawler info
Name aws-athena-demo
Tags -
IAM role
IAM role arn:aws:iam::398447858632:role/service-role/AWSGlueServiceRole-demo
Schedule
Schedule Run on demand
Output
Database default
Prefix added to tables (optional)
Create a single schema for each S3 path false

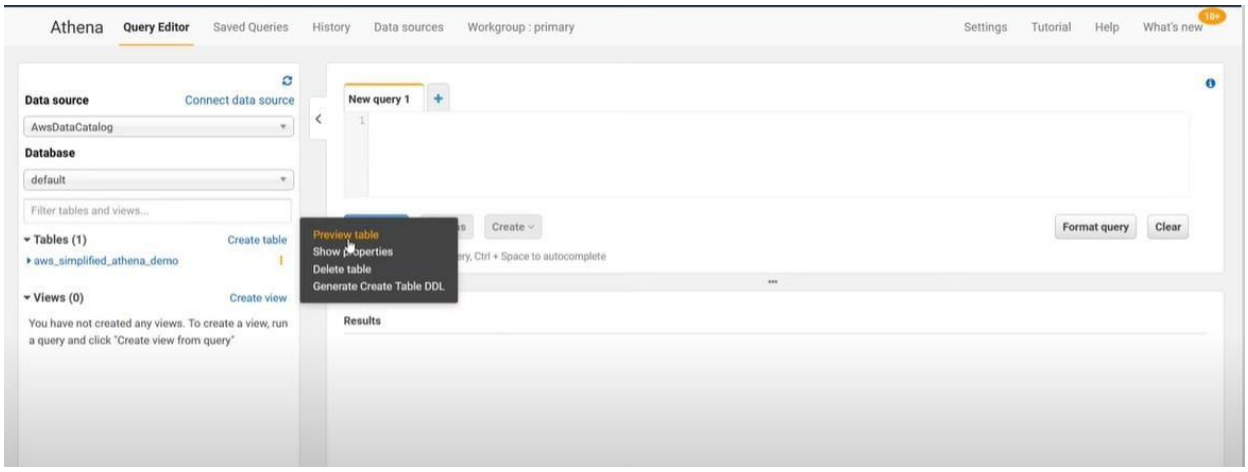
Step 10:

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.