

MIS 6389.002

Lab 19 – Glue and Athena

Keshav Rajbux

Folder named queried-data

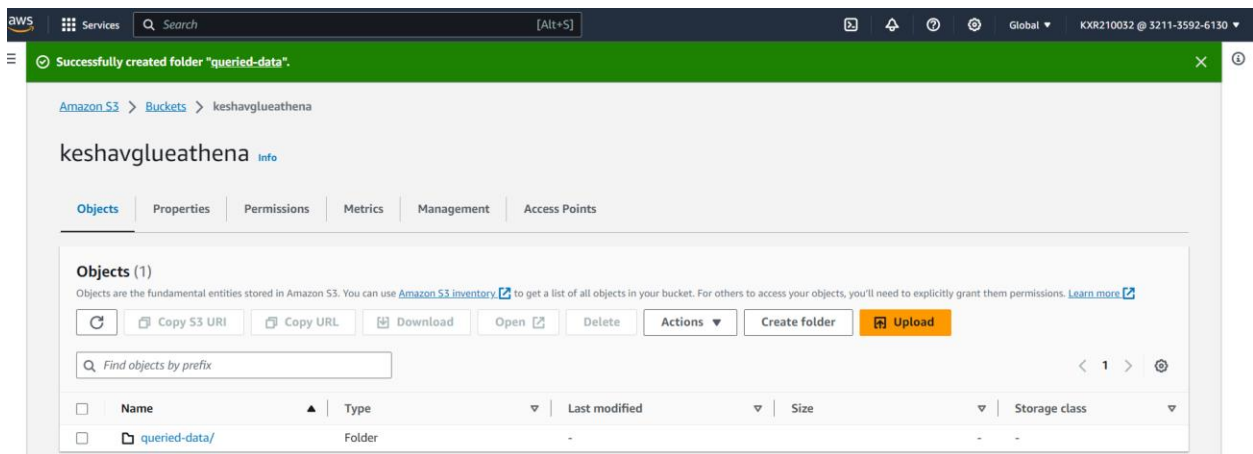


Figure 1 This is a deliverable

Table added after crawler run

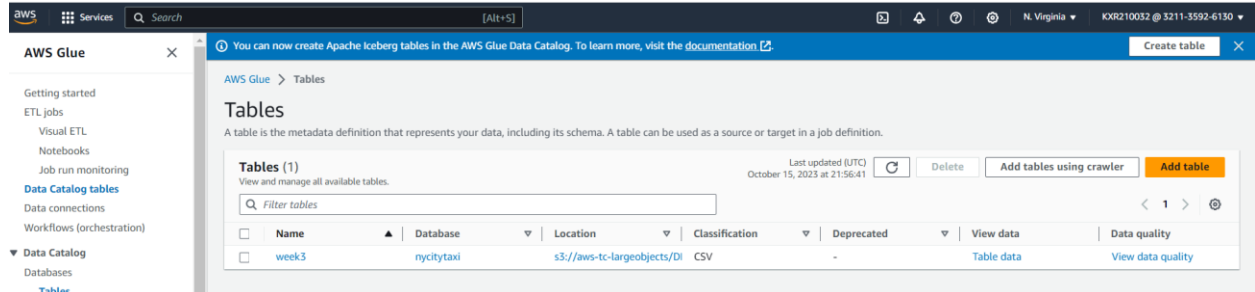


Figure 2 This is a deliverable

Editor Tab: Command1

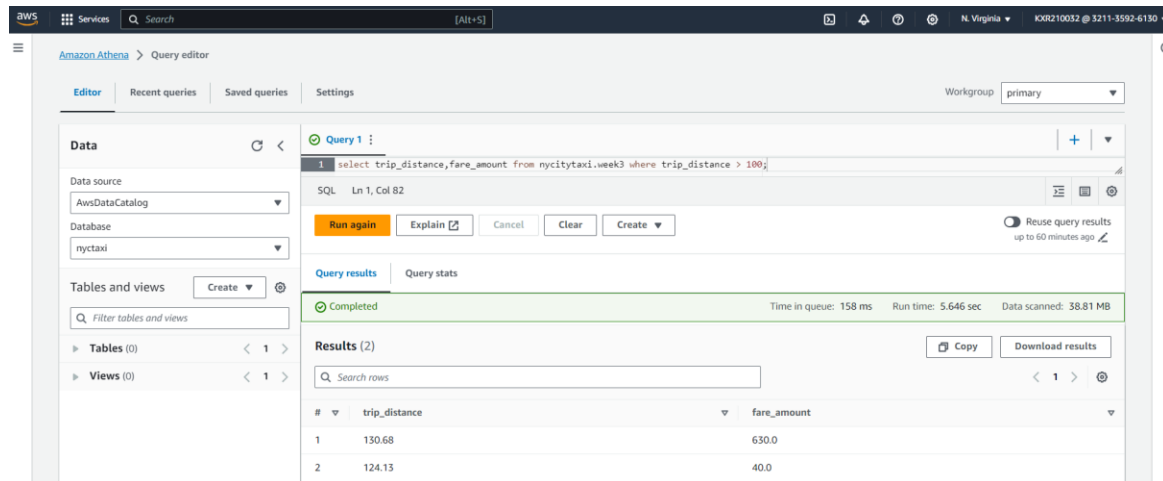
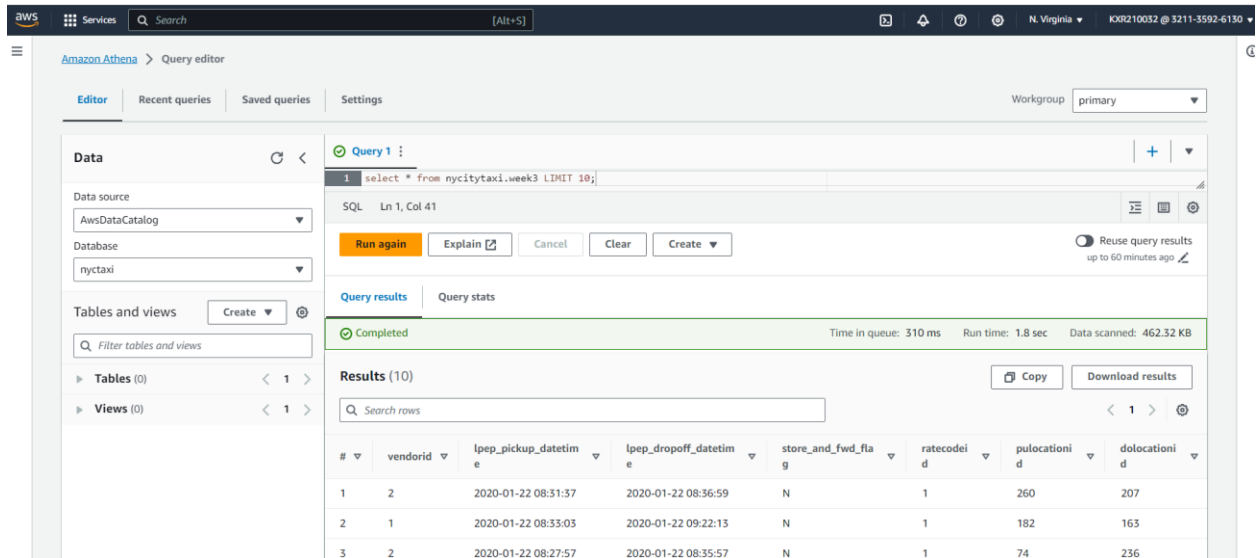


Figure 3 This is a deliverable

Command2

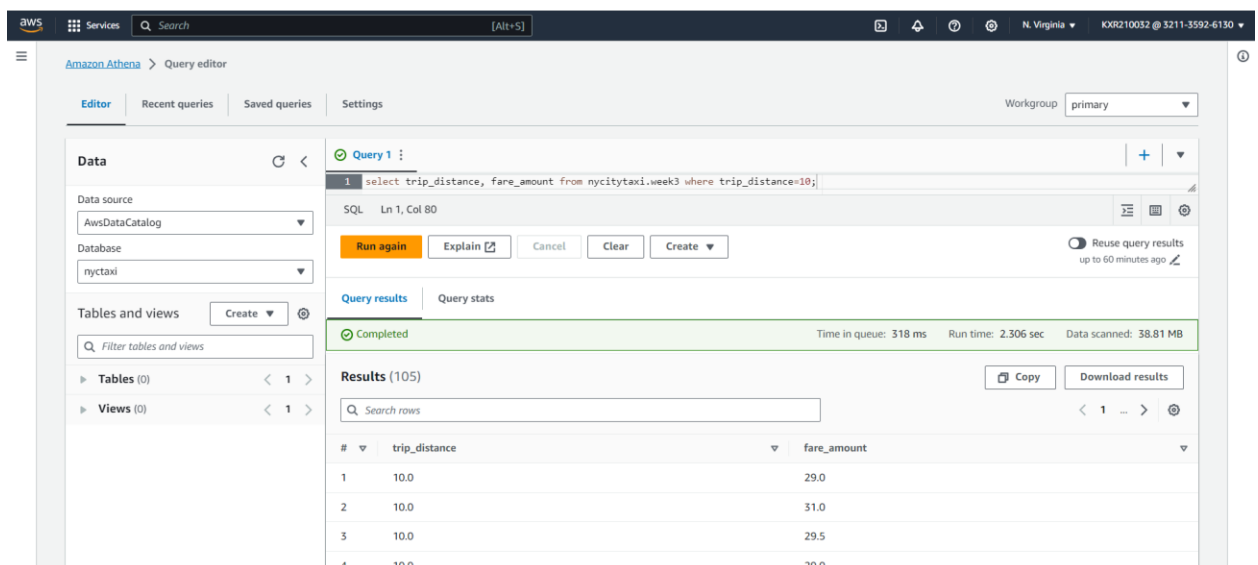


The screenshot displays the Amazon Athena Query Editor interface. On the left, the 'Data' panel shows the 'AwsDataCatalog' as the data source and 'nycitytaxi' as the database. The 'Query editor' tab is active, showing the SQL query: `select * from nycitytaxi.week3 LIMIT 10;`. The query has been executed successfully, as indicated by the 'Completed' status and the 'Query results' tab. The results show 10 rows of data, with columns including vendorid, lpep_pickup_datetime, lpep_dropoff_datetime, store_and_fwd_flag, ratecodeid, pulocationid, and dolocationid. The first three rows are visible in the table below.

#	vendorid	lpep_pickup_datetime	lpep_dropoff_datetime	store_and_fwd_flag	ratecodeid	pulocationid	dolocationid
1	2	2020-01-22 08:31:37	2020-01-22 08:36:59	N	1	260	207
2	1	2020-01-22 08:33:03	2020-01-22 09:22:13	N	1	182	163
3	2	2020-01-22 08:27:57	2020-01-22 08:35:57	N	1	74	236

Figure 4 This is a deliverable

Command 3 :



The screenshot displays the Amazon Athena Query Editor interface. On the left, the 'Data' panel shows the 'AwsDataCatalog' as the data source and 'nycitytaxi' as the database. The 'Query editor' tab is active, showing the SQL query: `select trip_distance, fare_amount from nycitytaxi.week3 where trip_distance=10;`. The query has been executed successfully, as indicated by the 'Completed' status and the 'Query results' tab. The results show 105 rows of data, with columns including trip_distance and fare_amount. The first four rows are visible in the table below.

#	trip_distance	fare_amount
1	10.0	29.0
2	10.0	31.0
3	10.0	29.5
4	10.0	29.0

Figure 5 This is a deliverable