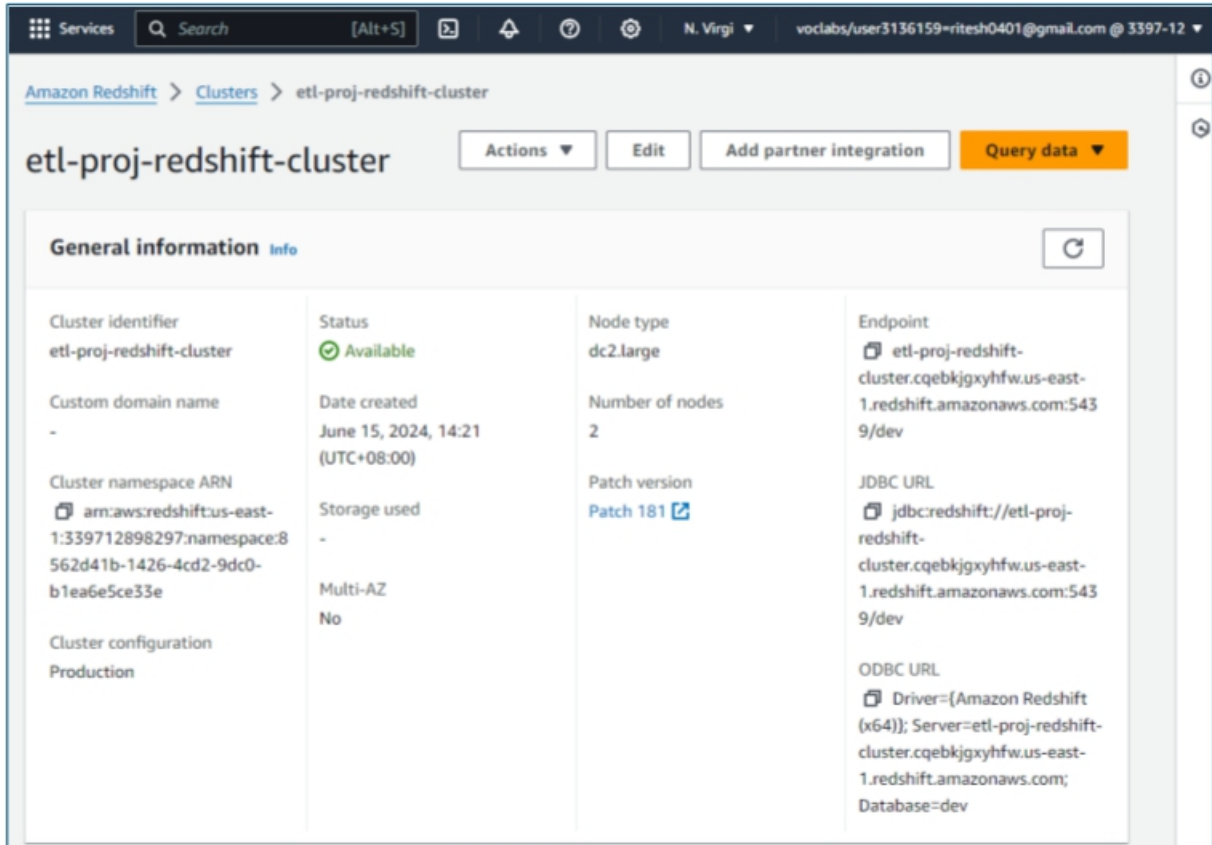


ETL Project : ATM Transactions

Redshift Set-up

1) Redshift cluster created:

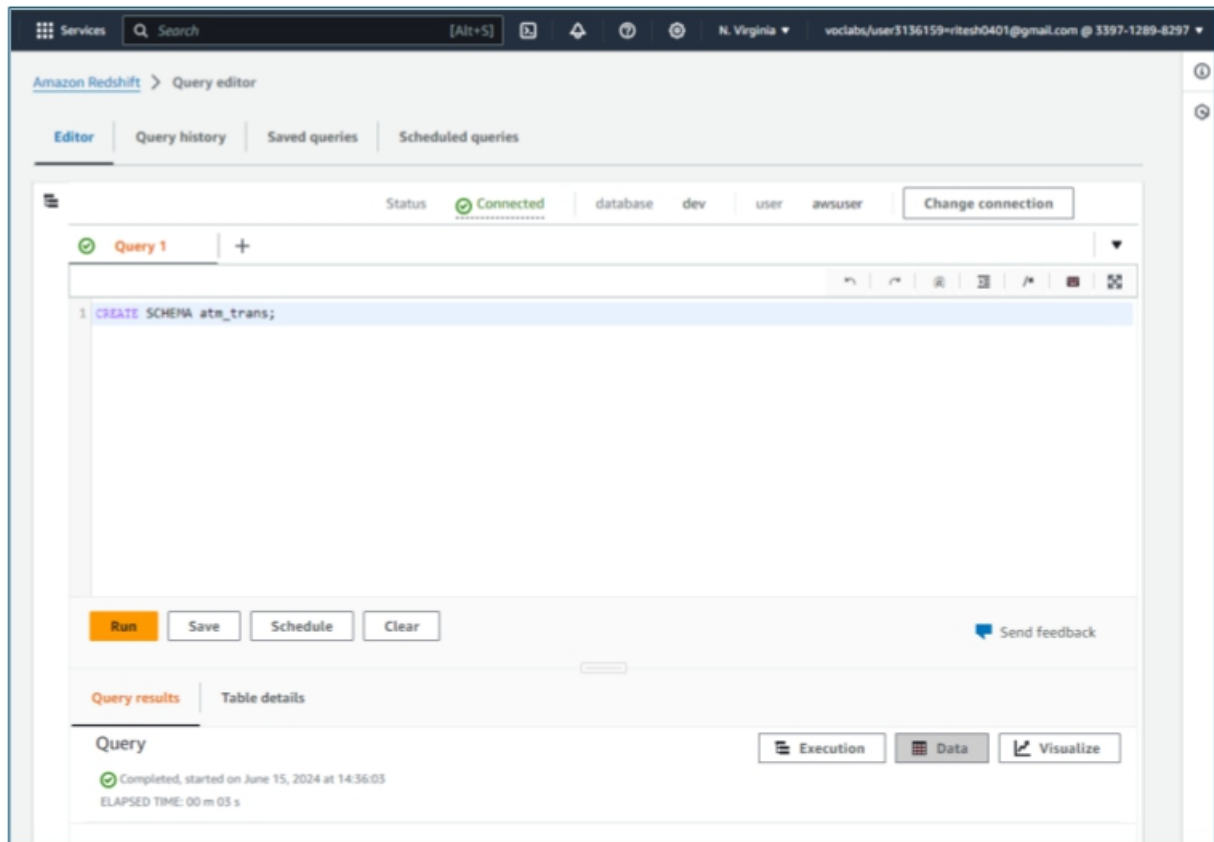


The screenshot displays the Amazon Redshift console interface. At the top, the navigation bar shows 'Services', a search bar, and the user profile 'N. Virgi'. The breadcrumb trail indicates the path: 'Amazon Redshift > Clusters > etl-proj-redshift-cluster'. The cluster name 'etl-proj-redshift-cluster' is prominently displayed, along with buttons for 'Actions', 'Edit', 'Add partner integration', and 'Query data'. Below this, the 'General information' tab is selected, showing a table of cluster details.

General information			
Cluster identifier	Status	Node type	Endpoint
etl-proj-redshift-cluster	Available	dc2.large	etl-proj-redshift-cluster.cqebkjgxyhfw.us-east-1.redshift.amazonaws.com:5439/dev
Custom domain name	Date created	Number of nodes	JDBC URL
-	June 15, 2024, 14:21 (UTC+08:00)	2	jdbc:redshift://etl-proj-redshift-cluster.cqebkjgxyhfw.us-east-1.redshift.amazonaws.com:5439/dev
Cluster namespace ARN	Storage used	Patch version	ODBC URL
arn:aws:redshift:us-east-1:339712898297:namespace:8562d41b-1426-4cd2-9dc0-b1ea6e5ce33e	-	Patch 181	Driver={Amazon Redshift (x64)}; Server=etl-proj-redshift-cluster.cqebkjgxyhfw.us-east-1.redshift.amazonaws.com; Database=dev
Cluster configuration	Multi-AZ		
Production	No		

2) Now create schema for the table:

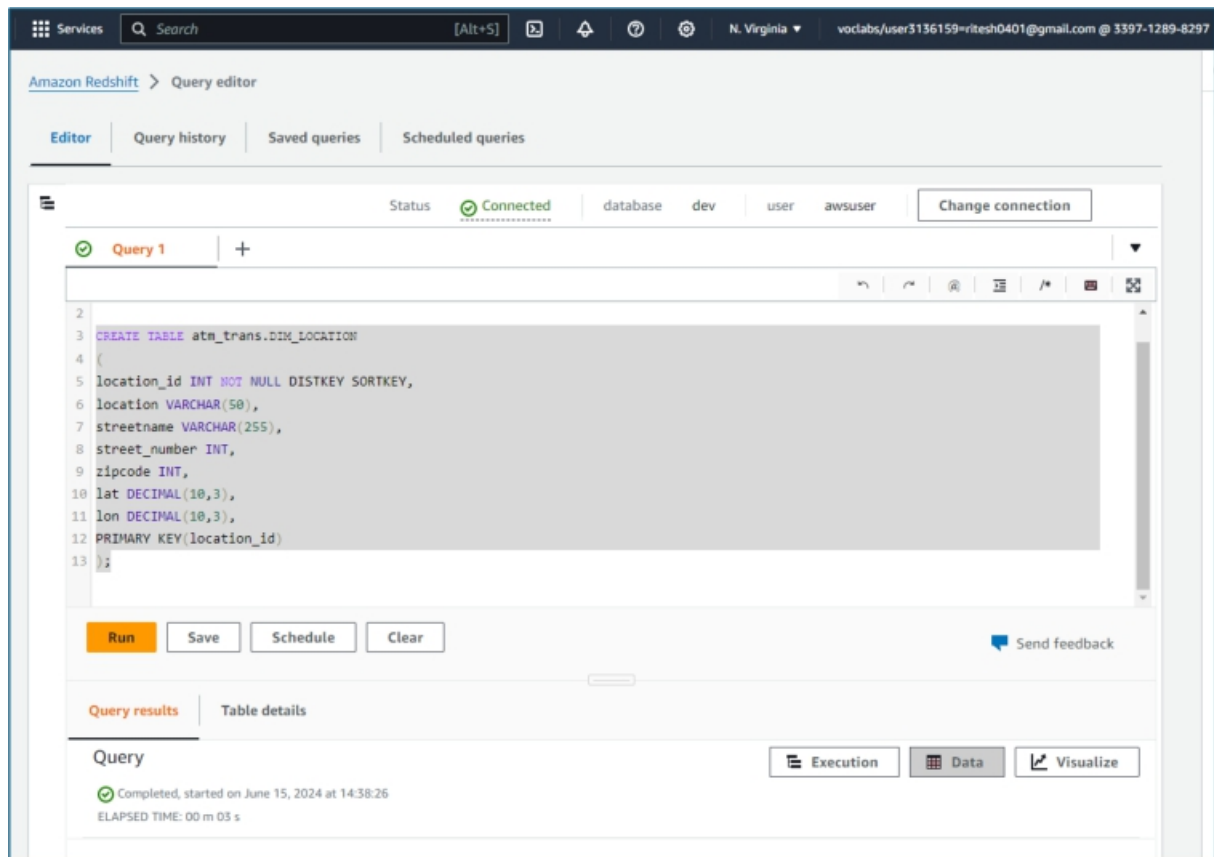
```
CREATE SCHEMA atm_trans;
```



3) Creating Dimension Tables.

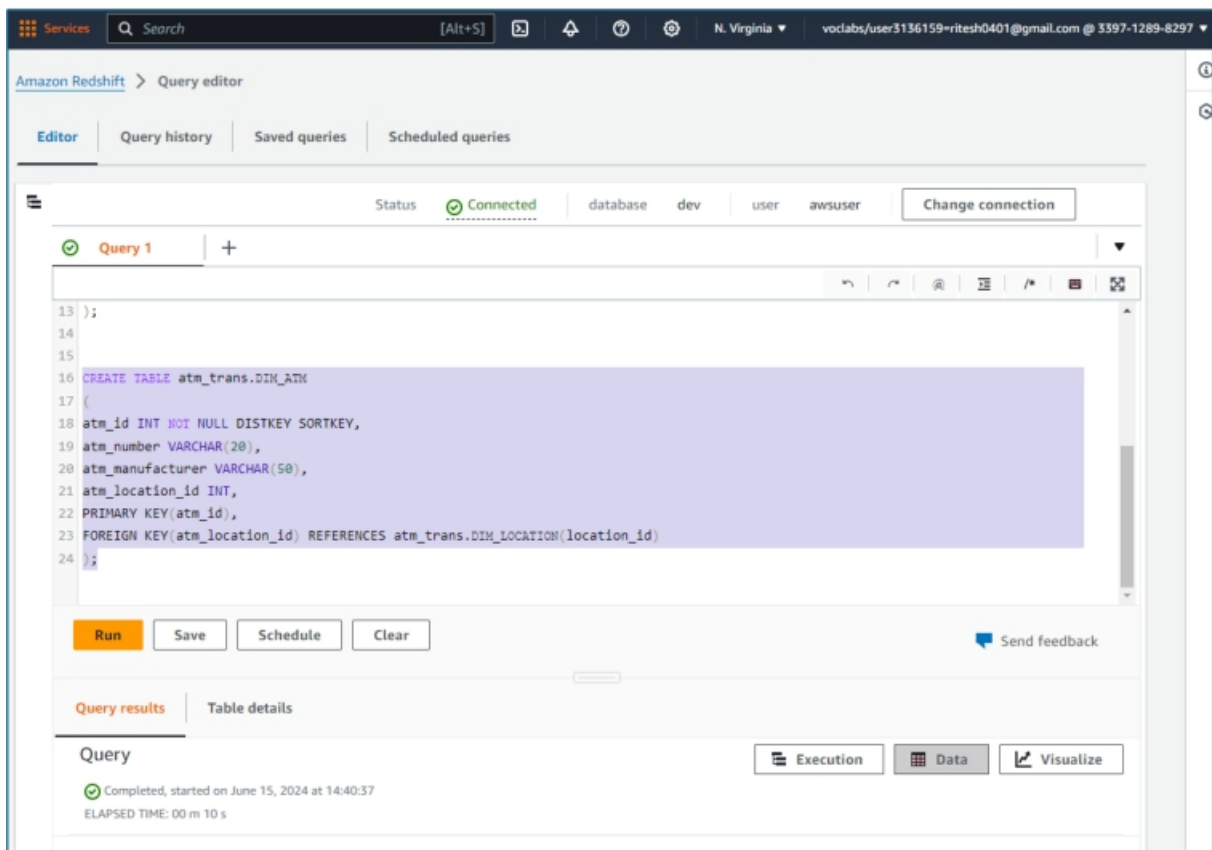
a) DIM_LOCATION :

```
CREATE TABLE atm_trans.DIM_LOCATION
(
  location_id INT NOT NULL DISTKEY SORTKEY,
  location VARCHAR(50),
  streetname VARCHAR(255),
  street_number INT,
  zipcode INT,
  lat DECIMAL(10,3),
  lon DECIMAL(10,3),
  PRIMARY KEY(location_id)
);
```



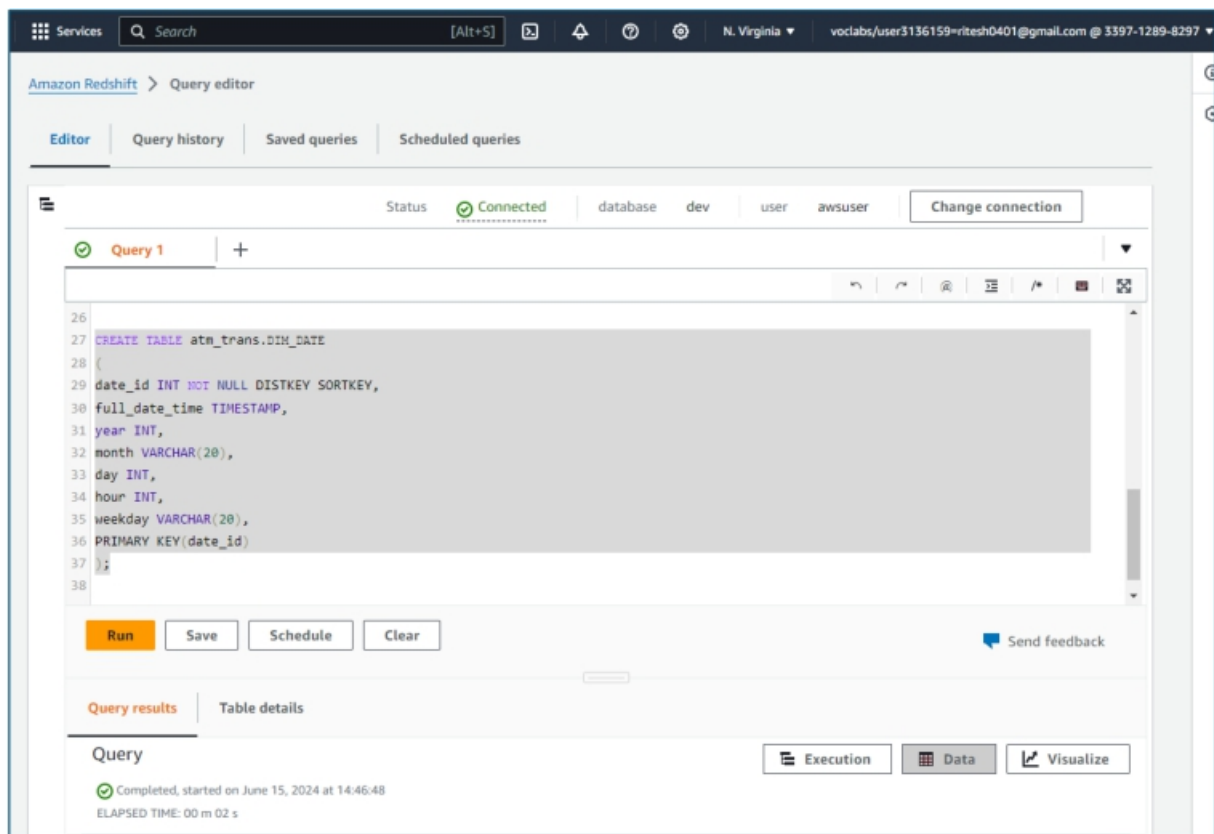
b) DIM_ATM

```
CREATE TABLE atm_trans.DIM_ATM
(
  atm_id INT NOT NULL DISTKEY SORTKEY,
  atm_number VARCHAR(20),
  atm_manufacturer VARCHAR(50),
  atm_location_id INT,
  PRIMARY KEY(atm_id),
  FOREIGN KEY(atm_location_id) REFERENCES
  atm_trans.DIM_LOCATION(location_id)
);
```



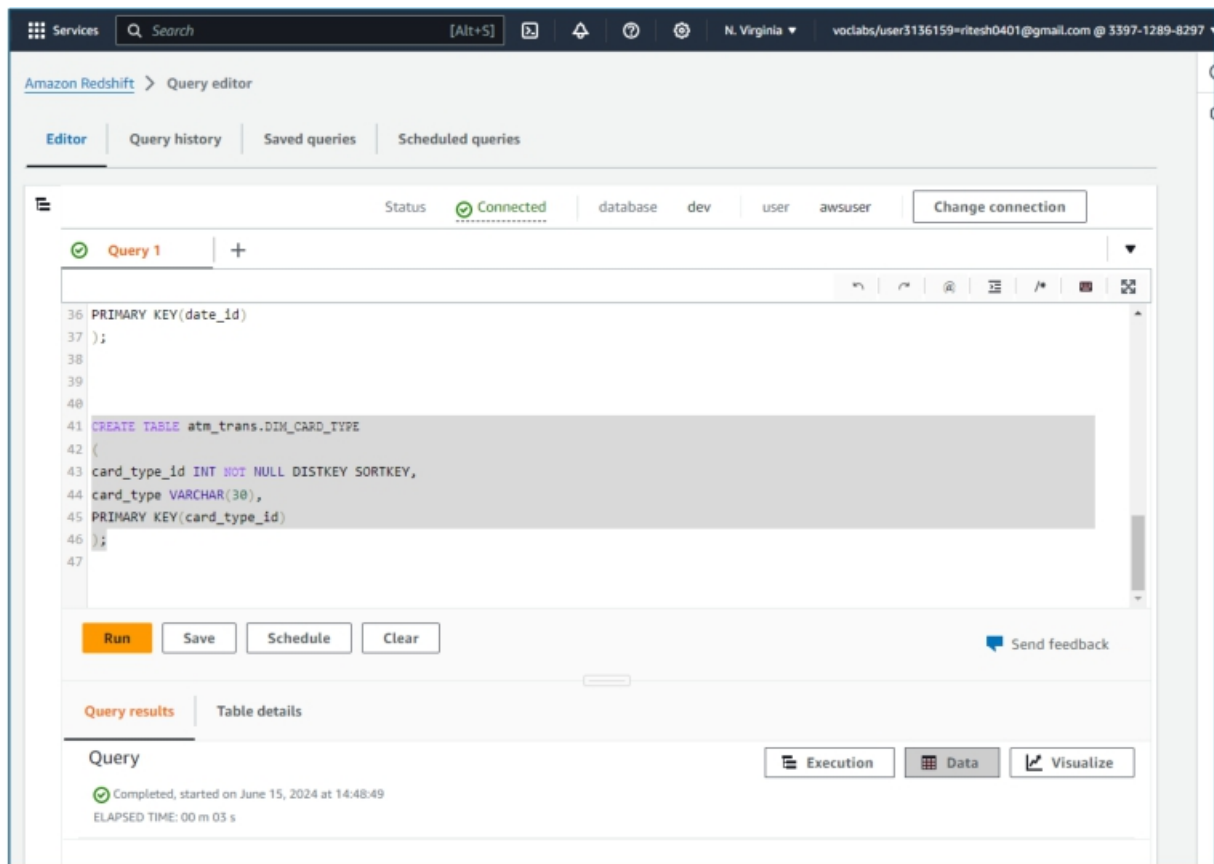
c) DIM_DATE

```
CREATE TABLE atm_trans.DIM_DATE (  
  date_id INT NOT NULL DISTKEY SORTKEY,  
  full_date_time TIMESTAMP,  
  year INT,  
  month VARCHAR(20),  
  day INT,  
  hour INT,  
  weekday VARCHAR(20),  
  PRIMARY KEY(date_id)  
);
```



d) DIM_CARD_TYPE :

```
CREATE TABLE atm_trans.DIM_CARD_TYPE  
(  
  card_type_id INT NOT NULL DISTKEY SORTKEY,  
  card_type VARCHAR(30),  
  PRIMARYKEY(card_type_id)  
);
```



ETL Project : ATM Transactions - Redshift Setup

4) Creating fact table

a) FACT_ATM_TRANS

```
CREATE TABLE atm_trans.FACT_ATM_TRANS
(
  trans_id BIGINT NOT NULL DISTKEY SORTKEY,
  atm_id INT,
  weather_loc_id INT,
  date_id INT,
  card_type_id INT,
  atm_status VARCHAR(20),
  currency VARCHAR(10),
  service VARCHAR(20),
  transaction_amount INT,
  message_code VARCHAR(225),
  message_text VARCHAR(225),
  rain_3h DECIMAL(10,3),
  clouds_all INT,
  weather_id INT,
  weather_main VARCHAR(50),
  weather_description VARCHAR(255),
  PRIMARY KEY(trans_id),
  FOREIGN KEY(weather_loc_id) REFERENCES
atm_trans.DIM_LOCATION(location_id),
  FOREIGN KEY(atm_id) REFERENCES atm_trans.DIM_ATM(atm_id),
  FOREIGN KEY(date_id) REFERENCES atm_trans.DIM_DATE(date_id),
  FOREIGN KEY(card_type_id) REFERENCES
atm_trans.DIM_CARD_TYPE(card_type_id)
);
```

Services

Search

[Alt+S]

N. Virginia

voclabs/user3136159-ritesh0401@gmail.com @ 3397-1289-8297

47

48

49 CREATE TABLE atm_trans.FACT_ATM_TRANS

50 (

51 trans_id BIGINT NOT NULL DISTKEY SORTKEY,

52 atm_id INT,

53 weather_loc_id INT,

54 date_id INT,

55 card_type_id INT,

56 atm_status VARCHAR(20),

57 currency VARCHAR(10),

58 service VARCHAR(20),

59 transaction_amount INT,

60 message_code VARCHAR(225),

61 message_text VARCHAR(225),

62 rain_3h DECIMAL(10,3),

63 clouds_all INT,

64 weather_id INT,

65 weather_main VARCHAR(50),

66 weather_description VARCHAR(255),

67 PRIMARY KEY(trans_id),

68 FOREIGN KEY(weather_loc_id) REFERENCES atm_trans.DIM_LOCATION(location_id),

69 FOREIGN KEY(atm_id) REFERENCES atm_trans.DIM_ATM(atm_id),

70 FOREIGN KEY(date_id) REFERENCES atm_trans.DIM_DATE(date_id),

71 FOREIGN KEY(card_type_id) REFERENCES atm_trans.DIM_CARD_TYPE(card_type_id)

72);

Run

Save

Schedule

Clear

Send feedback

Query results

Table details

Query

Execution

Data

Visualize

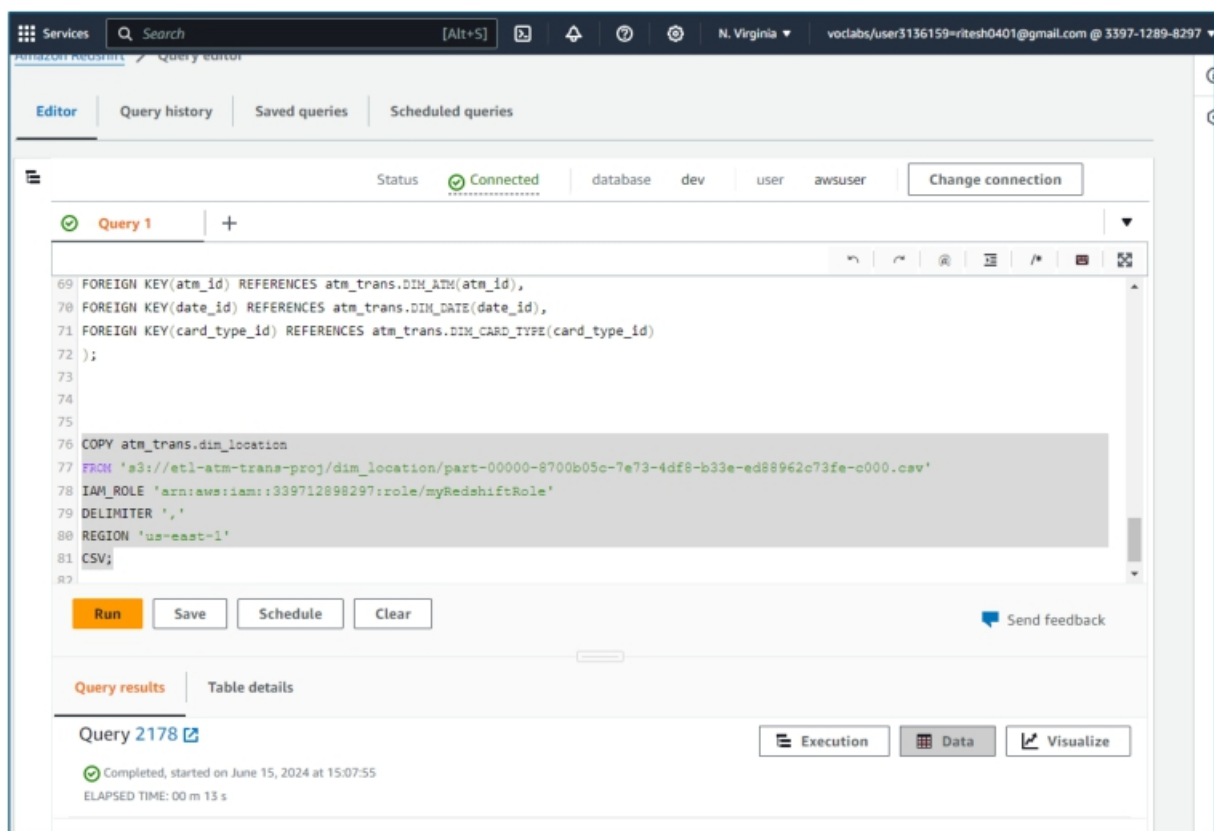
Completed, started on June 15, 2024 at 14:52:56

ELAPSED TIME: 00 m 20 s

5) Loading data from S3 Buckets to Redshift cluster

a) Copy the data to dimension table - dim_location

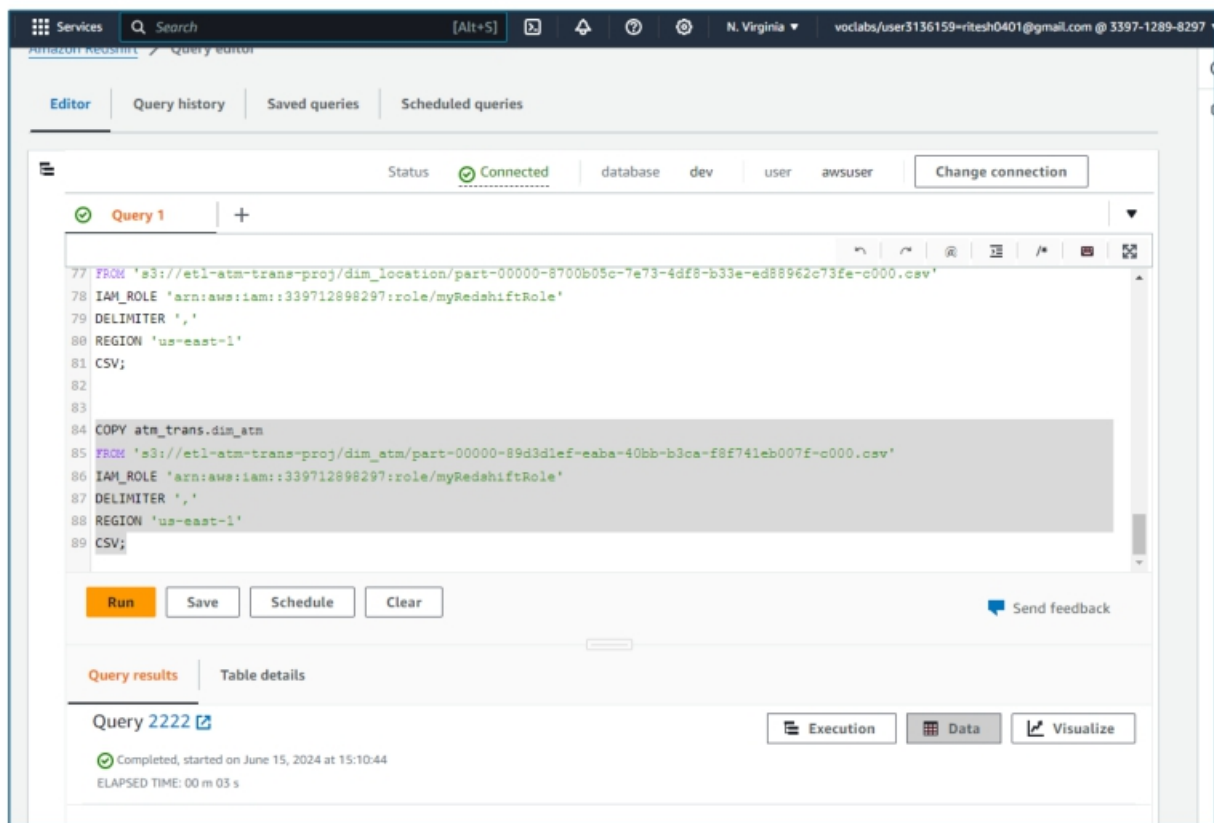
```
COPY atm_trans.dim_location
FROM 's3://etl-atm-trans-proj/dim_location/part-00000-8700b05c-7e73-4df8-
b33e-ed88962c73fe-c000.csv'
IAM_ROLE 'arn:aws:iam::339712898297:role/myRedshiftRole'
DELIMITER ','
REGION 'us-east-1'
CSV;
```



ETL Project : ATM Transactions - Redshift Setup

b) Copy the data to dimension table - dim_atm

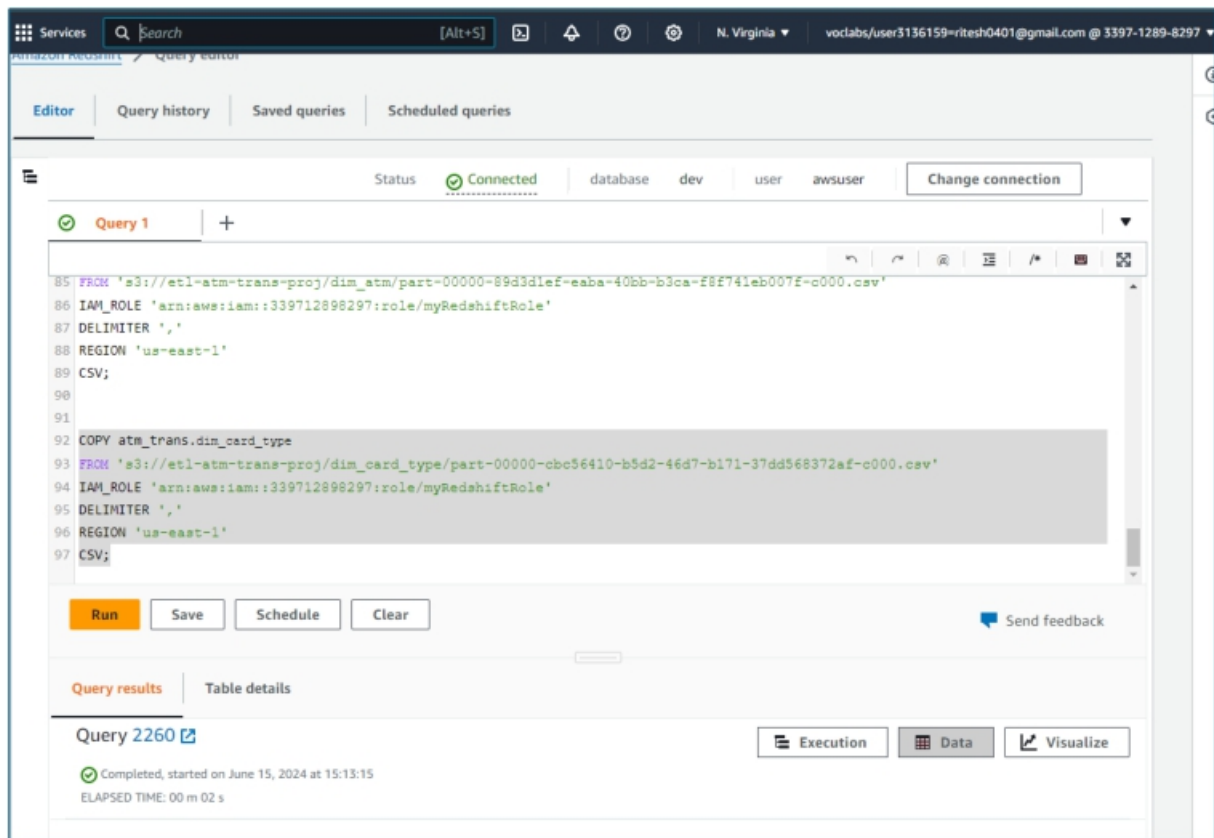
```
COPY atm_trans.dim_atm
FROM 's3://etl-atm-trans-proj/dim_atm/part-00000-89d3d1ef-eaba-40bb-b3ca-
f8f741eb007f-c000.csv'
IAM_ROLE 'arn:aws:iam::339712898297:role/myRedshiftRole'
DELIMITER ','
REGION 'us-east-1'
CSV;
```



ETL Project : ATM Transactions - Redshift Setup

c) Copy the data to dimension table - dim_card_type

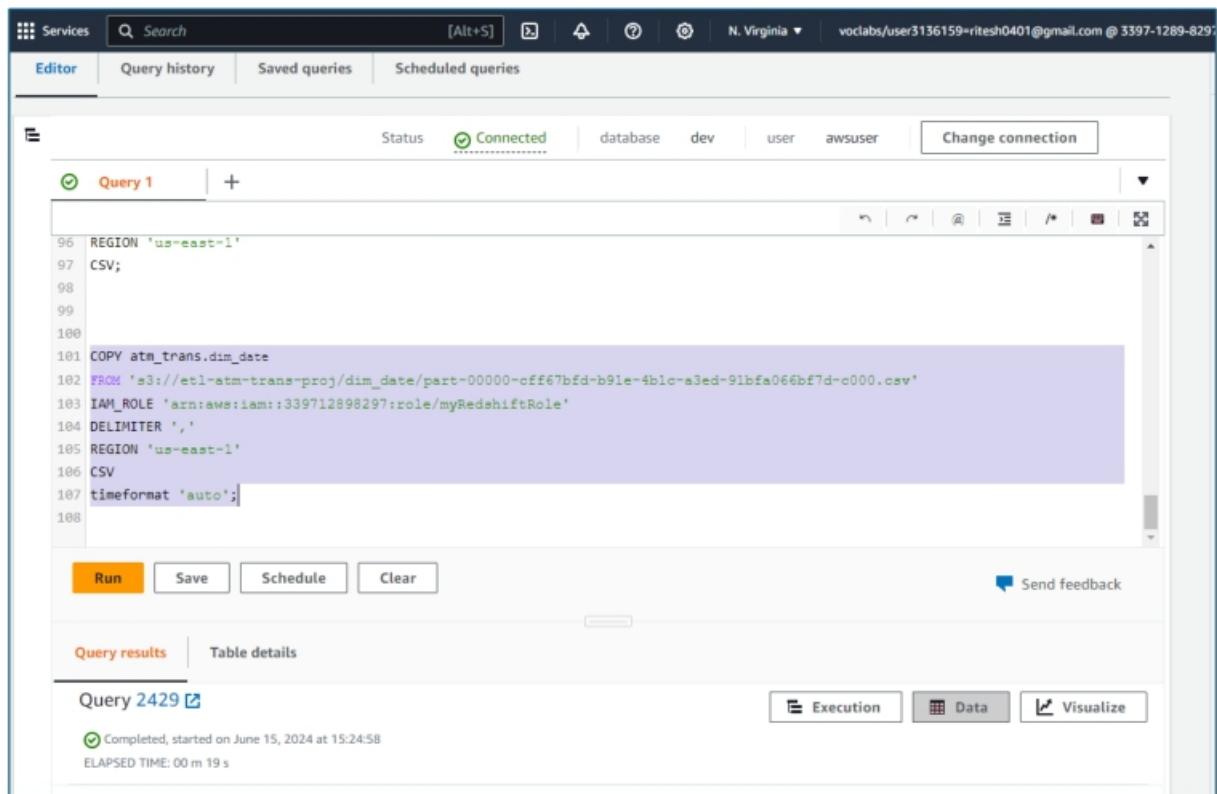
```
COPY atm_trans.dim_card_type
FROM 's3://etl-atm-trans-proj/dim_card_type/part-00000-cbc56410-b5d2-46d7-
b171-37dd568372af-c000.csv'
IAM_ROLE 'arn:aws:iam::339712898297:role/myRedshiftRole'
DELIMITER ','
REGION 'us-east-1'
CSV;
```



ETL Project : ATM Transactions - Redshift Setup

d) Copy the data to dimension table - dim_date

```
COPY atm_trans.dim_date
FROM 's3://etl-atm-trans-proj/dim_date/part-00000-cff67bfd-b91e-4b1c-a3ed-
91bfa066bf7d-c000.csv'
IAM_ROLE 'arn:aws:iam::339712898297:role/myRedshiftRole'
DELIMITER ','
REGION 'us-east-1'
CSV
timeformat 'auto';
```



ETL Project : ATM Transactions - Redshift Setup

e) Copy data to fact table – fact_atm_trans

```
COPY atm_trans.fact_atm_trans
FROM 's3://etl-atm-trans-proj/fact_atm_trans/part-00000-aadd058d-77ef-48a7-
b4ed-e0869b129602-c000.csv'
IAM_ROLE 'arn:aws:iam::339712898297:role/myRedshiftRole'
DELIMITER ','
REGION 'us-east-1'
CSV;
```

The screenshot displays the Amazon Redshift Query Editor interface. At the top, there's a navigation bar with 'Services', a search bar, and user information. Below this is a tabbed interface with 'Editor', 'Query history', 'Saved queries', and 'Scheduled queries'. The 'Editor' tab is active, showing a SQL query for 'Query 1'. The query is as follows:

```
104 DELIMITER ','
105 REGION 'us-east-1'
106 CSV
107 timeformat 'auto';
108
109
110
111 COPY atm_trans.fact_atm_trans
112 FROM 's3://etl-atm-trans-proj/fact_atm_trans/part-00000-aadd058d-77ef-48a7-b4ed-e0869b129602-c000.csv'
113 IAM_ROLE 'arn:aws:iam::339712898297:role/myRedshiftRole'
114 DELIMITER ','
115 REGION 'us-east-1'
116 CSV;
```

Below the query editor, there are buttons for 'Run', 'Save', 'Schedule', and 'Clear'. To the right is a 'Send feedback' link. The bottom section shows 'Query results' and 'Table details' tabs. The 'Query results' tab is active, displaying 'Query 2497' with a status of 'Completed, started on June 15, 2024 at 15:28:51' and an 'ELAPSED TIME: 00 m 08 s'. There are also buttons for 'Execution', 'Data', and 'Visualize'.