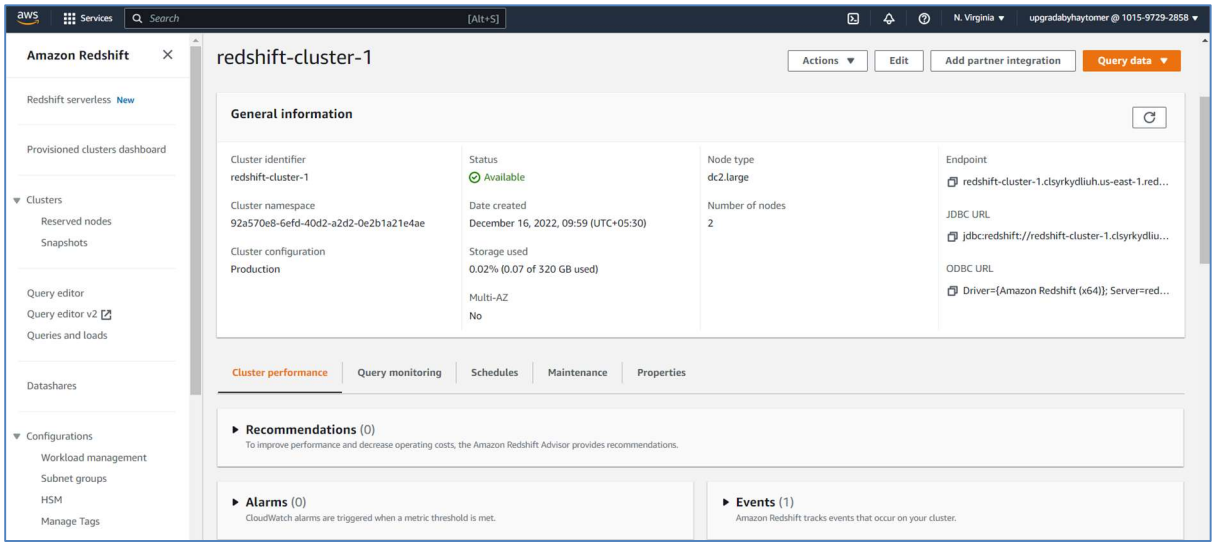
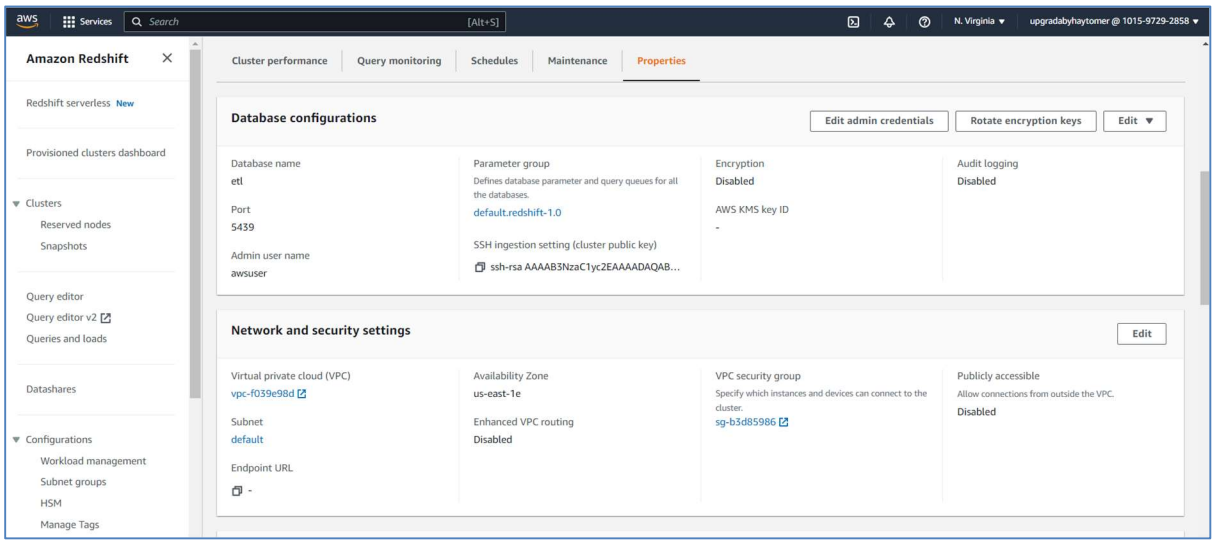
Creation of a Redshift Cluster

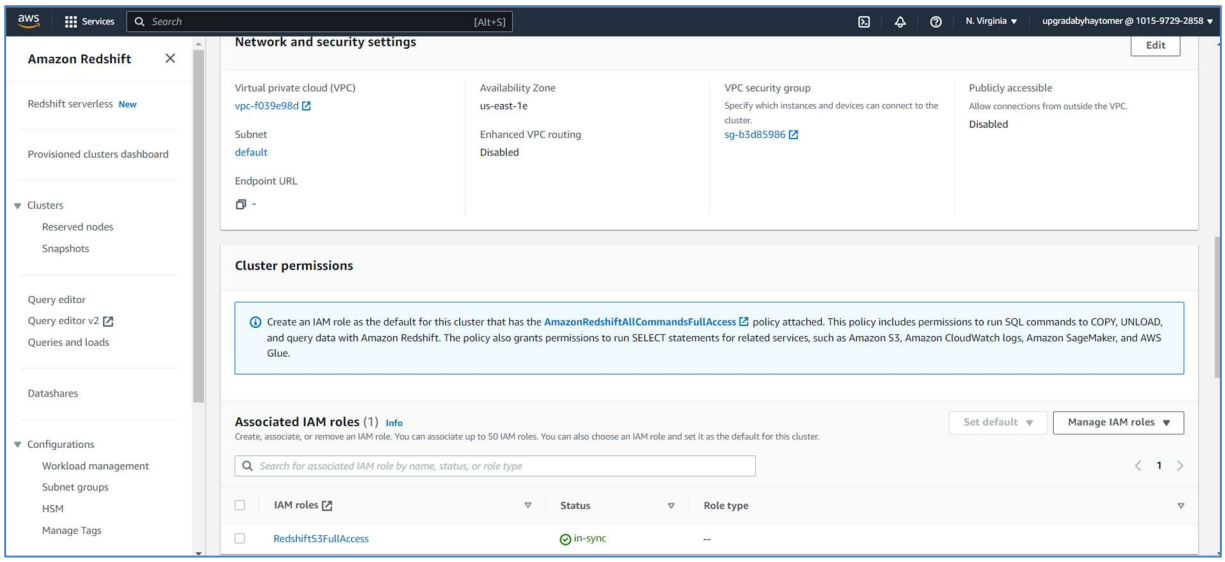
Screenshots of the configuration of the Redshift cluster that we have created:



Screenshot for Cluster Configuration and Network Security Settings:



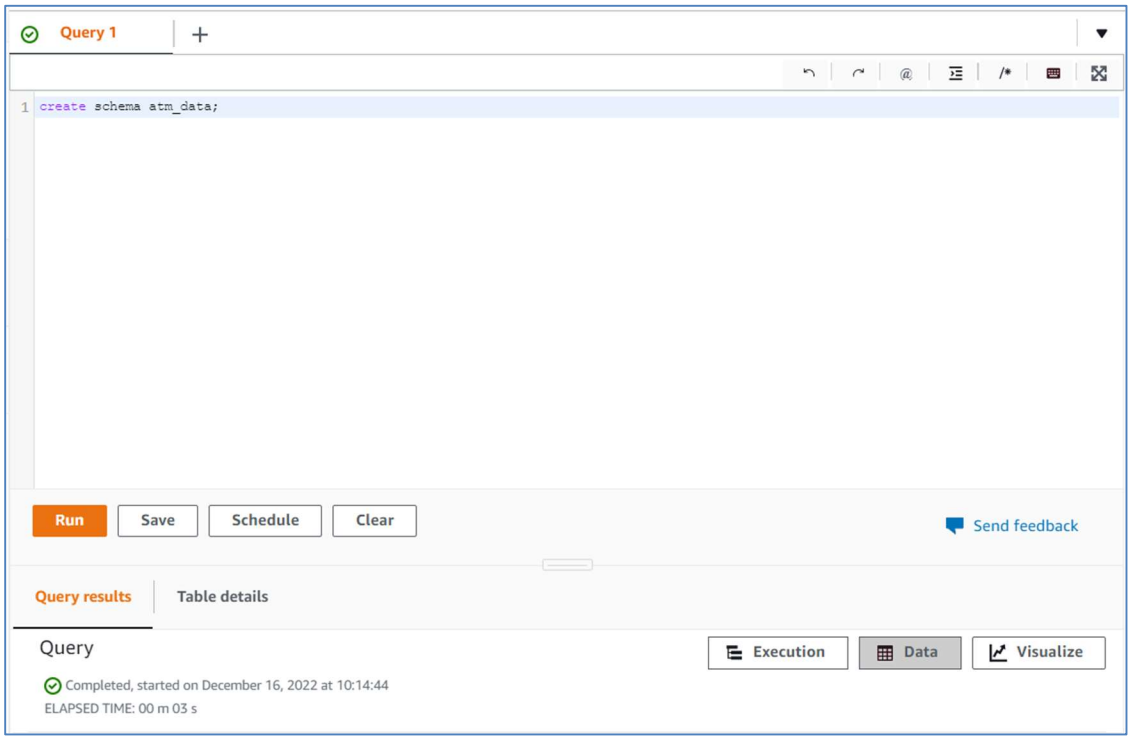
Screenshot of IAM Roles associated with Redshift Cluster for S3 Access



Setting up a database in the Redshift cluster and running queries to create the dimension and fact tables. Below are commands and related screenshots.

- **Creating schema atm_data:**

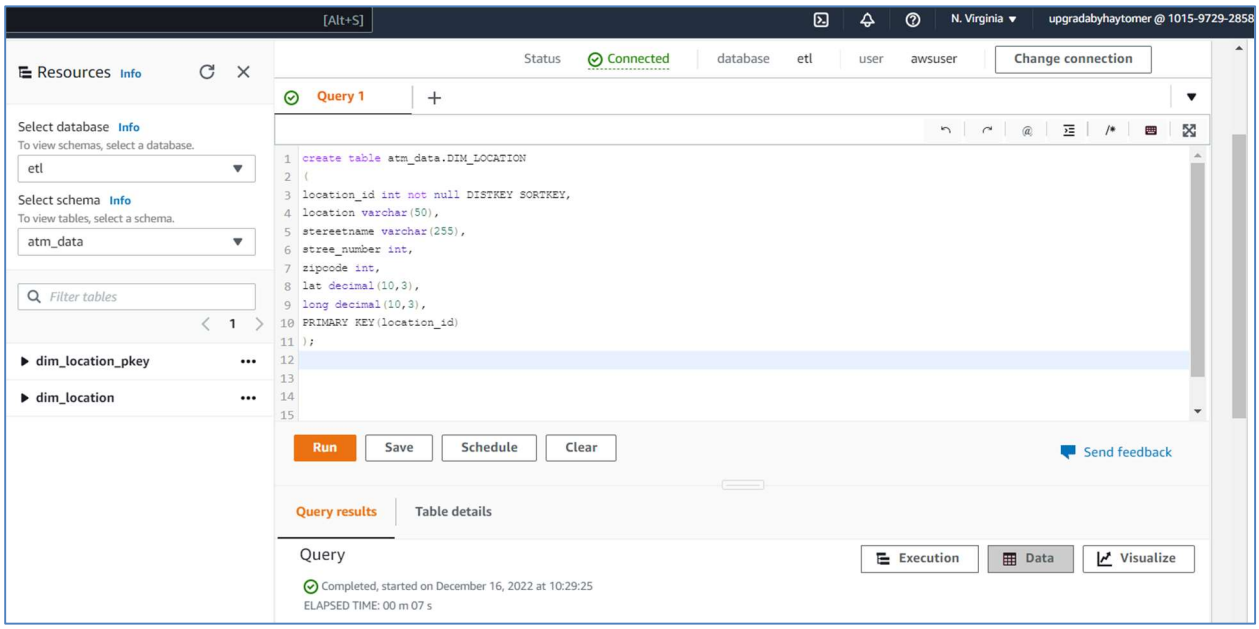
Create schema atm_data;



Query to create the various dimension and fact tables with appropriate primary and foreign keys:

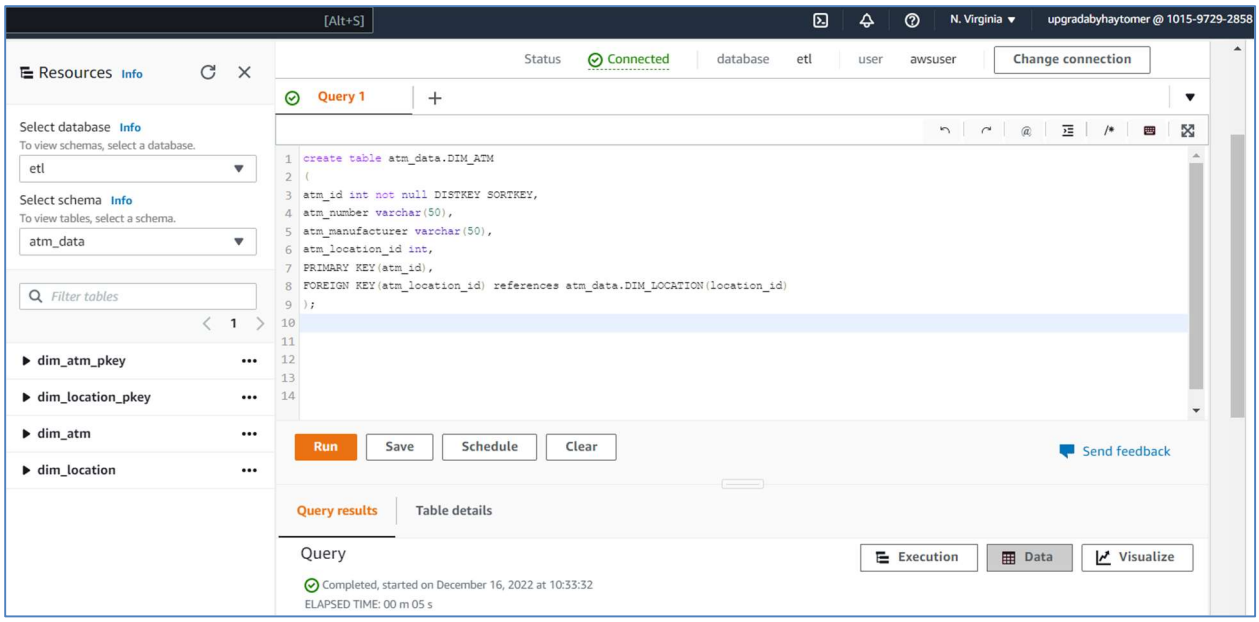
- **Creating Location dimension Table:**

```
create table atm_data.DIM_LOCATION
(
location_id int not null DISTKEY SORTKEY,
location varchar(50),
stereetname varchar(255),
stree_number int,
zipcode int,
lat decimal(10,3),
long decimal(10,3),
PRIMARY KEY(location_id)
);
```



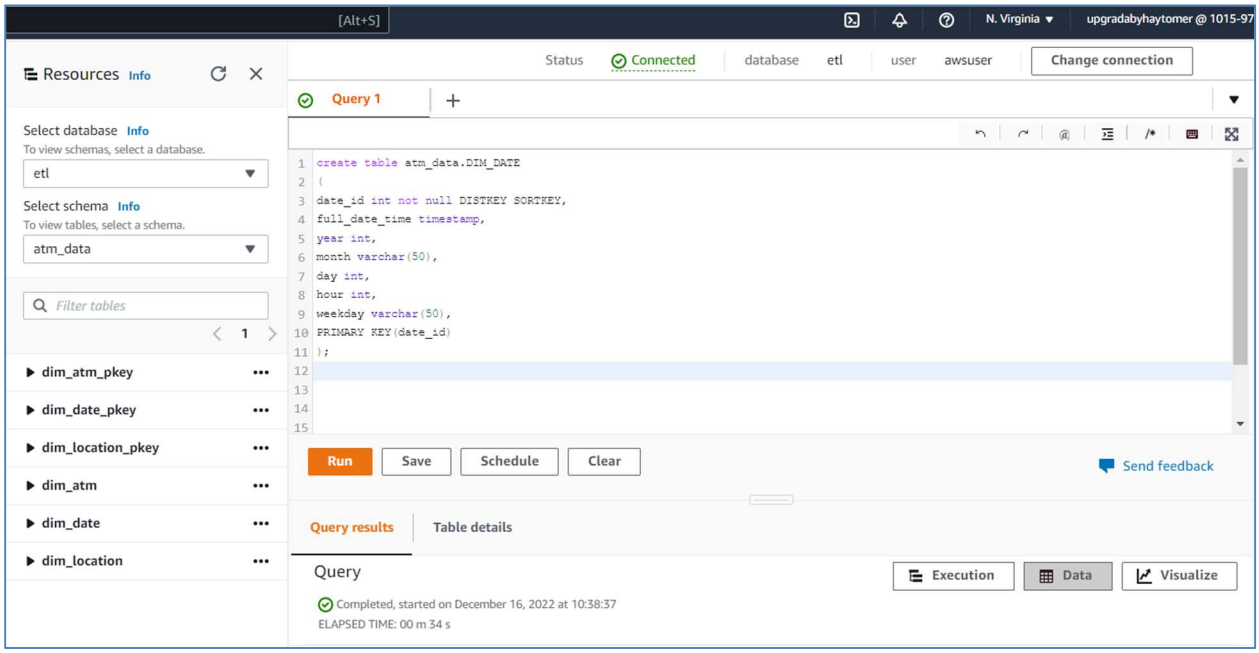
- **Creating ATM Dimension Table:**

```
create table atm_data.DIM_ATM
(
atm_id int not null DISTKEY SORTKEY,
atm_number varchar(50),
atm_manufacturer varchar(50),
atm_location_id int,
PRIMARY KEY(atm_id),
FOREIGN KEY(atm_location_id) references atm_data.DIM_LOCATION(location_id)
);
```



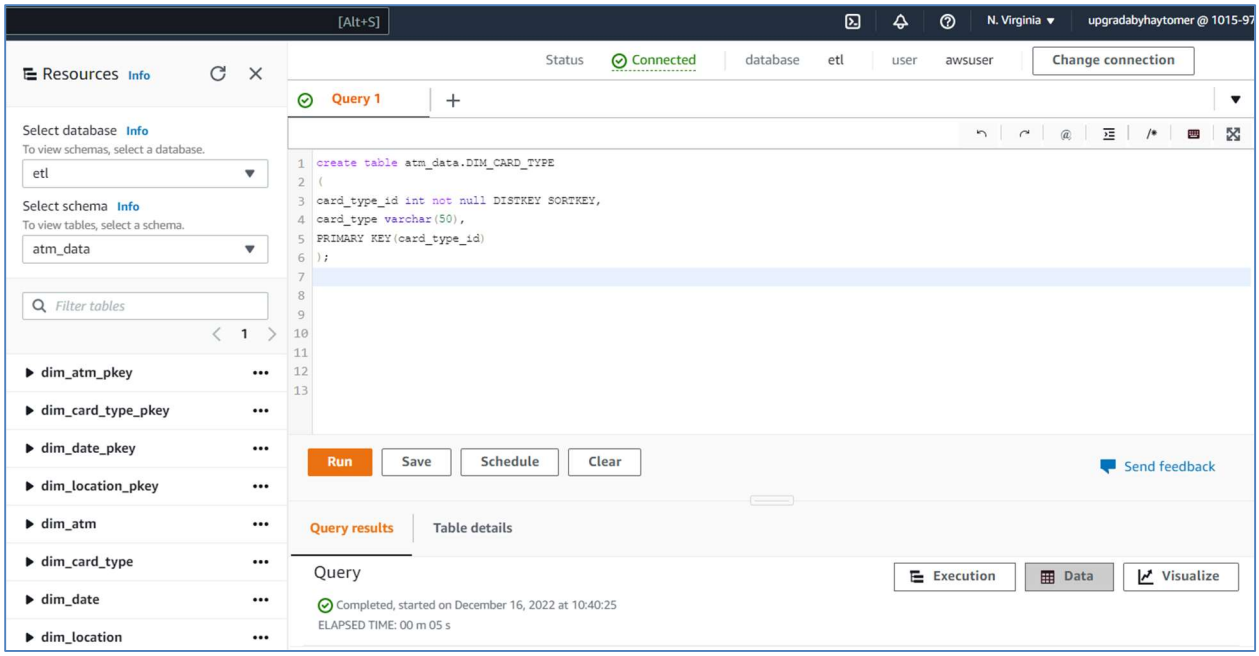
- **Creating Date Dimension Table:**

```
create table atm_data.DIM_DATE
(
date_id int not null DISTKEY SORTKEY,
full_date_time timestamp,
year int,
month varchar(50),
day int,
hour int,
weekday varchar(50),
PRIMARY KEY(date_id)
);
```



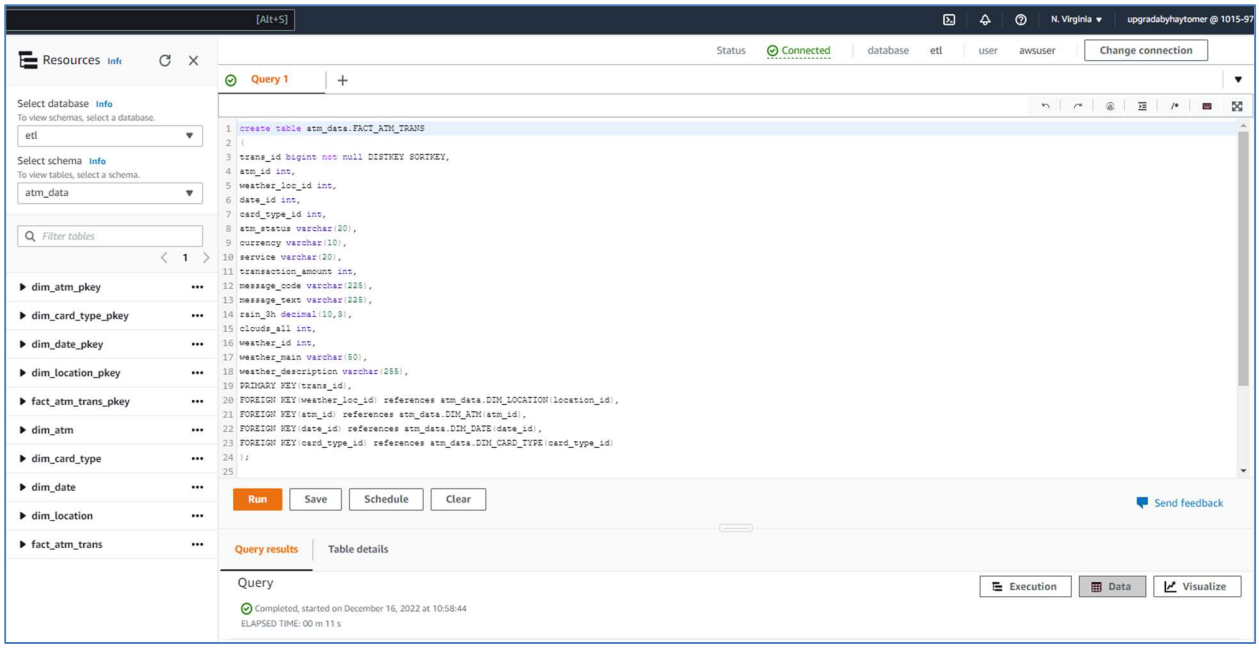
- **Creating Card Type Dimension Table:**

```
create table atm_data.DIM_CARD_TYPE
(
card_type_id int not null DISTKEY SORTKEY,
card_type varchar(50),
PRIMARY KEY(card_type_id)
);
```



- **Creating Fact ATM Trans Table:**

```
create table atm_data.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_data.DIM_LOCATION(location_id),
FOREIGN KEY(atm_id) references atm_data.DIM_ATM(atm_id),
FOREIGN KEY(date_id) references atm_data.DIM_DATE(date_id),
FOREIGN KEY(card_type_id) references atm_data.DIM_CARD_TYPE(card_type_id)
);
```

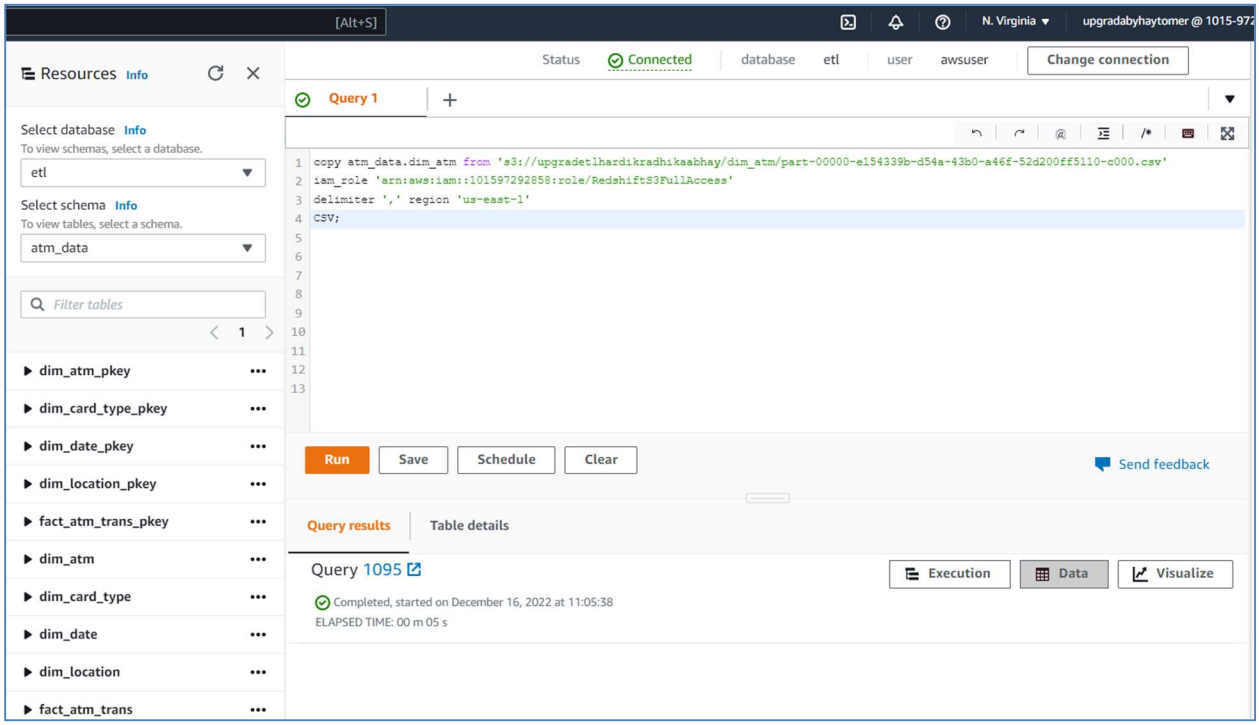


Loading data into a Redshift cluster from Amazon S3 bucket

Queries to copy the data from S3 buckets to the Redshift cluster in the appropriate tables

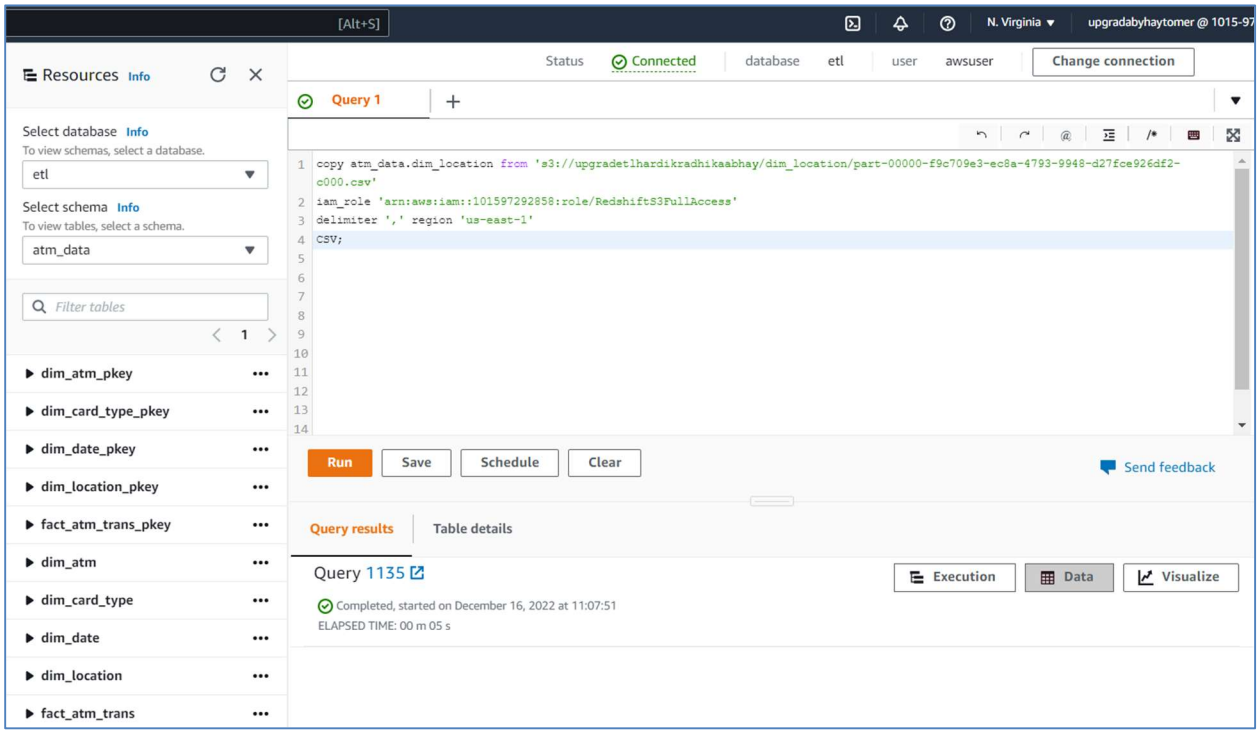
- **Copying the data to dim_atm from S3 location:**

```
copy atm_data.dim_atm from 's3://upgradetlhardikradhikaabhay/dim_atm/part-00000-e154339b-d54a-43b0-a46f-52d200ff5110-c000.csv'
iam_role 'arn:aws:iam::101597292858:role/RedshiftS3FullAccess'
delimiter ',' region 'us-east-1'
CSV;
```



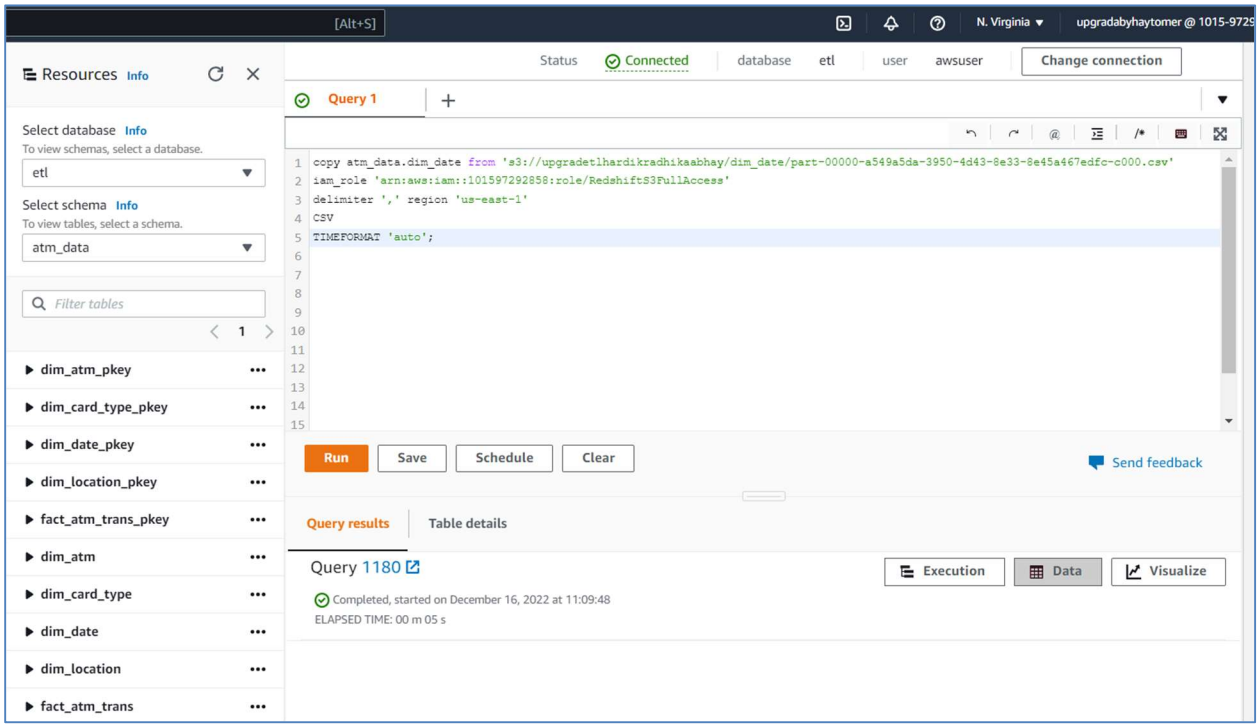
- **Copying the data to dim_location from S3 location:**

```
copy atm_data.dim_location from 's3://upgradetlhardikradhikaabhay/dim_location/part-00000-f9c709e3-ec8a-4793-9948-d27fce926df2-c000.csv'
iam_role 'arn:aws:iam::101597292858:role/RedshiftS3FullAccess'
delimiter ',' region 'us-east-1'
CSV;
```



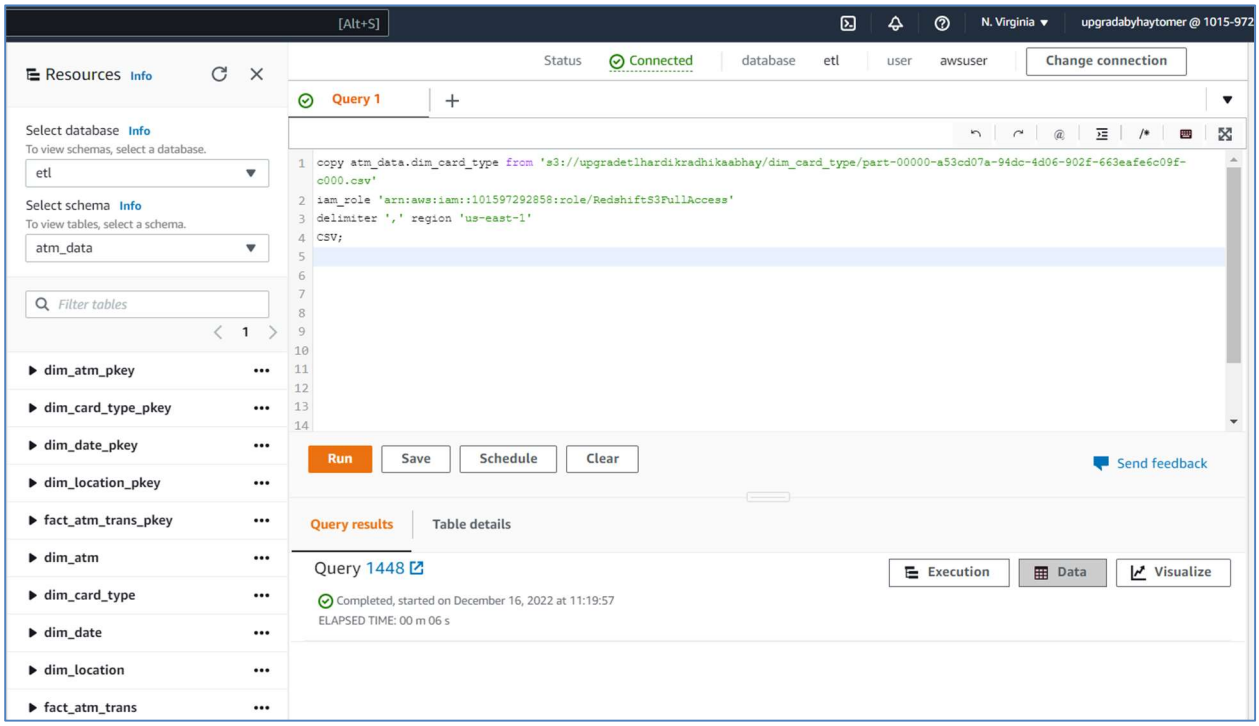
- **Copying the data to dim_date from S3 location:**

```
copy atm_data.dim_date from 's3://upgradetlhardikradhikaabhay/dim_date/part-00000-a549a5da-3950-4d43-8e33-8e45a467edfc-c000.csv'
iam_role 'arn:aws:iam::101597292858:role/RedshiftS3FullAccess'
delimiter ',' region 'us-east-1'
CSV
TIMEFORMAT 'auto';
```



- **Copying the data to dim_card_type from S3 location:**

copy atm_data.dim_card_type from 's3://upgradetlhardikradhikaabhay/dim_card_type/part-00000-a53cd07a-94dc-4d06-902f-663eafe6c09f-c000.csv'
iam_role 'arn:aws:iam::101597292858:role/RedshiftS3FullAccess'
delimiter ',' region 'us-east-1'
CSV;



- **Copying the data to fact_atm_trans table from S3 location:**

copy atm_data.fact_atm_trans from 's3://upgradetlhardikradhikaabhay/fact_atm_trans/part-00000-ea5f0521-d9cd-44af-b159-4d6fd35a29f3-c000.csv'
iam_role 'arn:aws:iam::101597292858:role/RedshiftS3FullAccess'
delimiter ',' region 'us-east-1'
CSV;