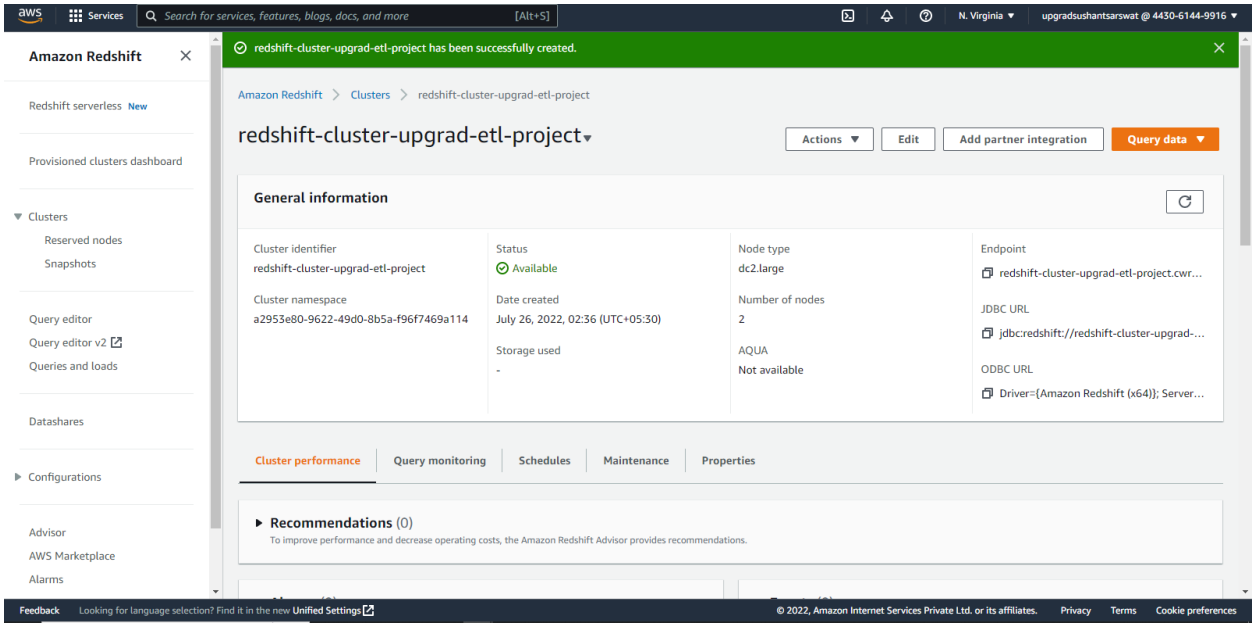
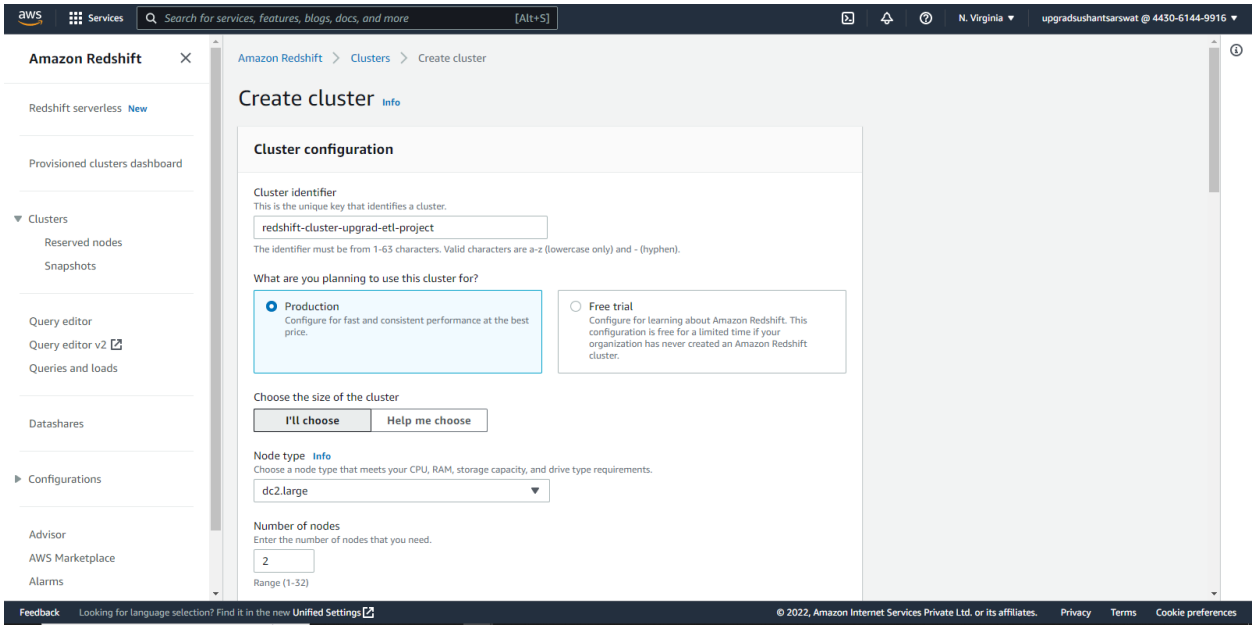
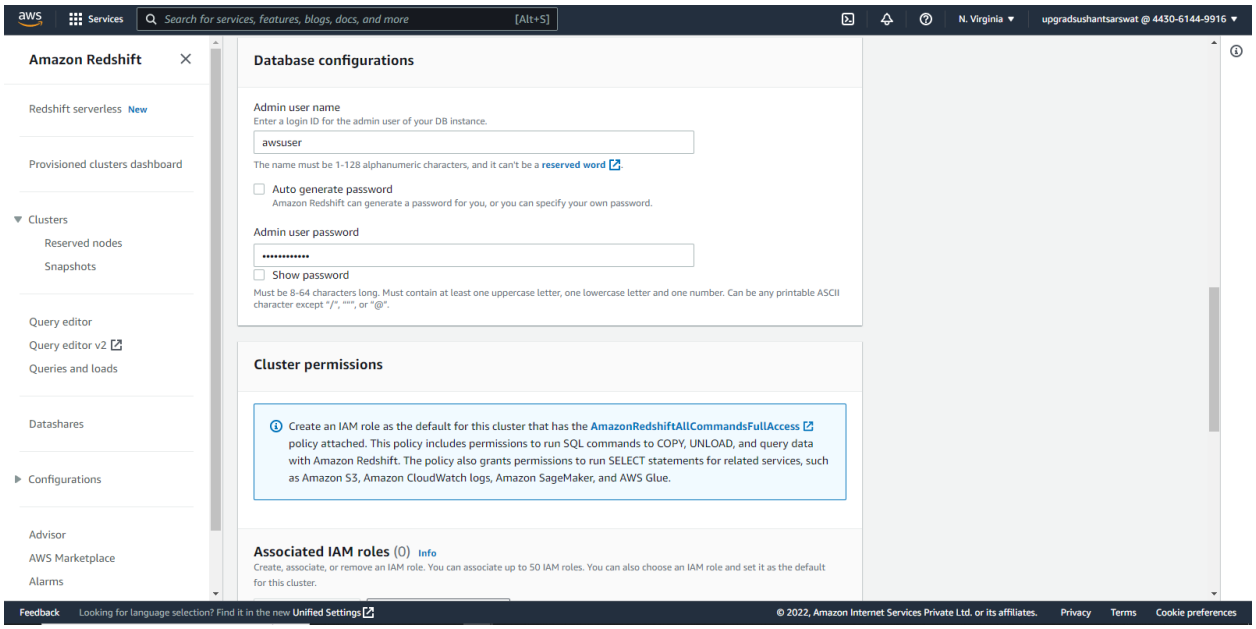
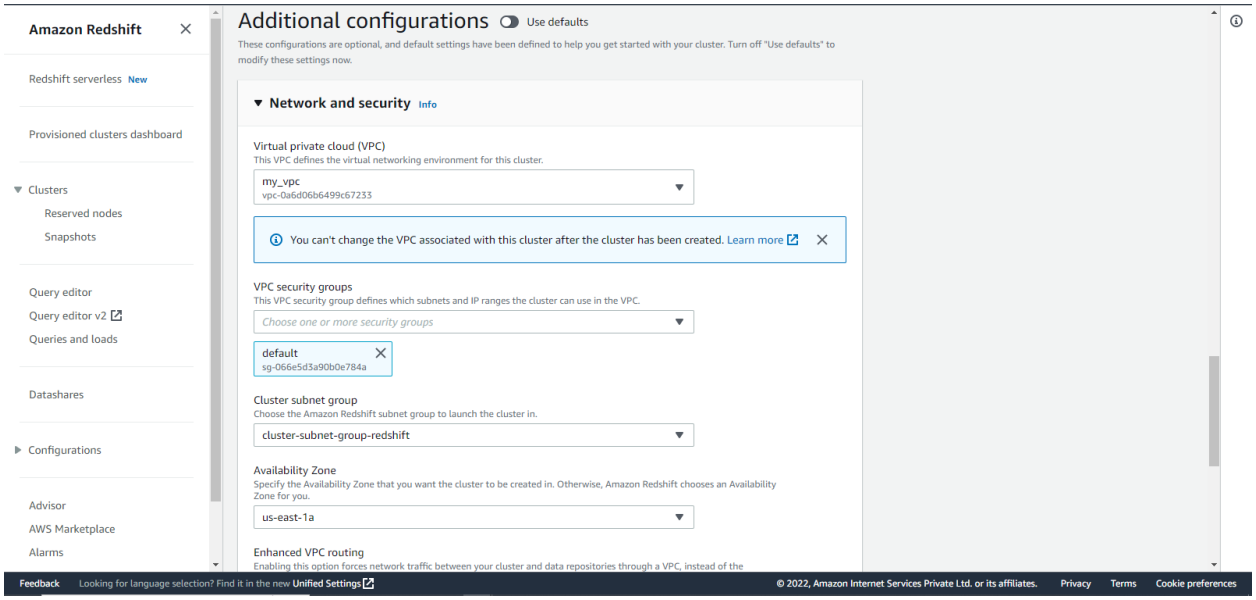
Creation of a Redshift Cluster

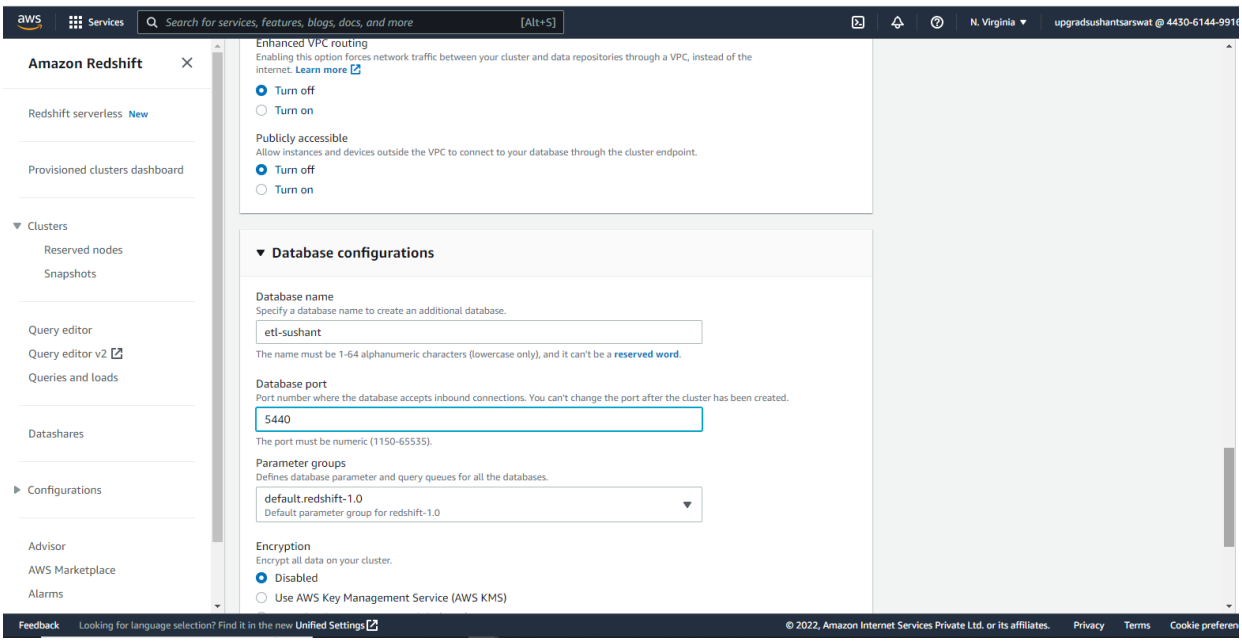
Screenshots of the configuration of the Redshift cluster that I have created:





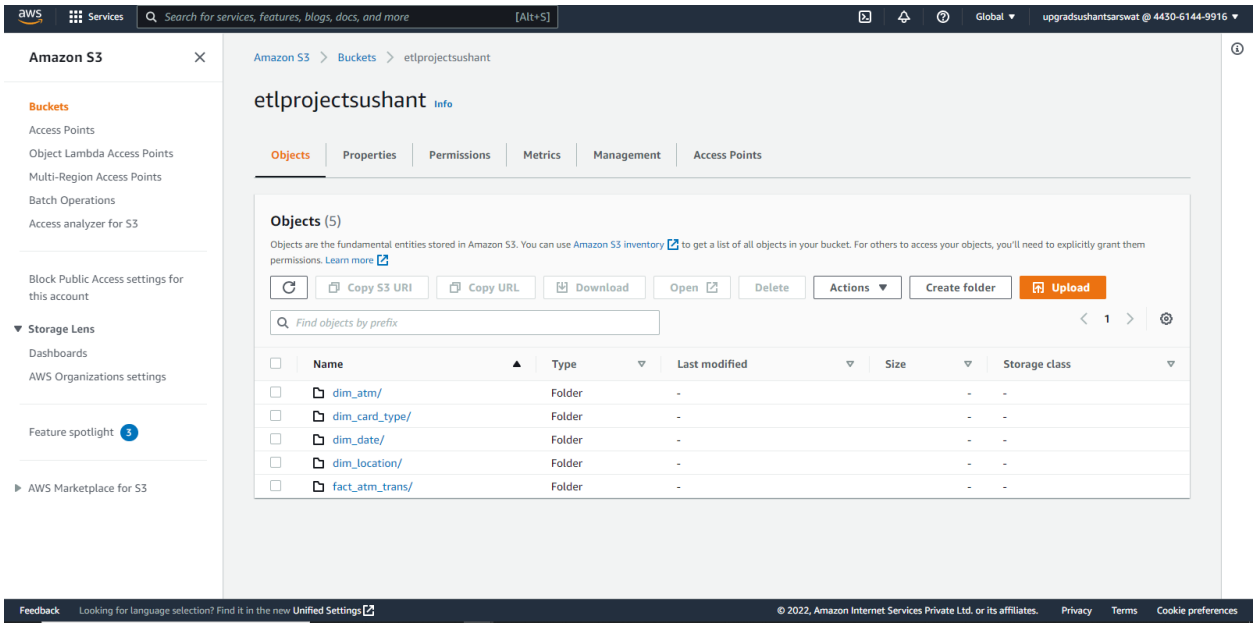






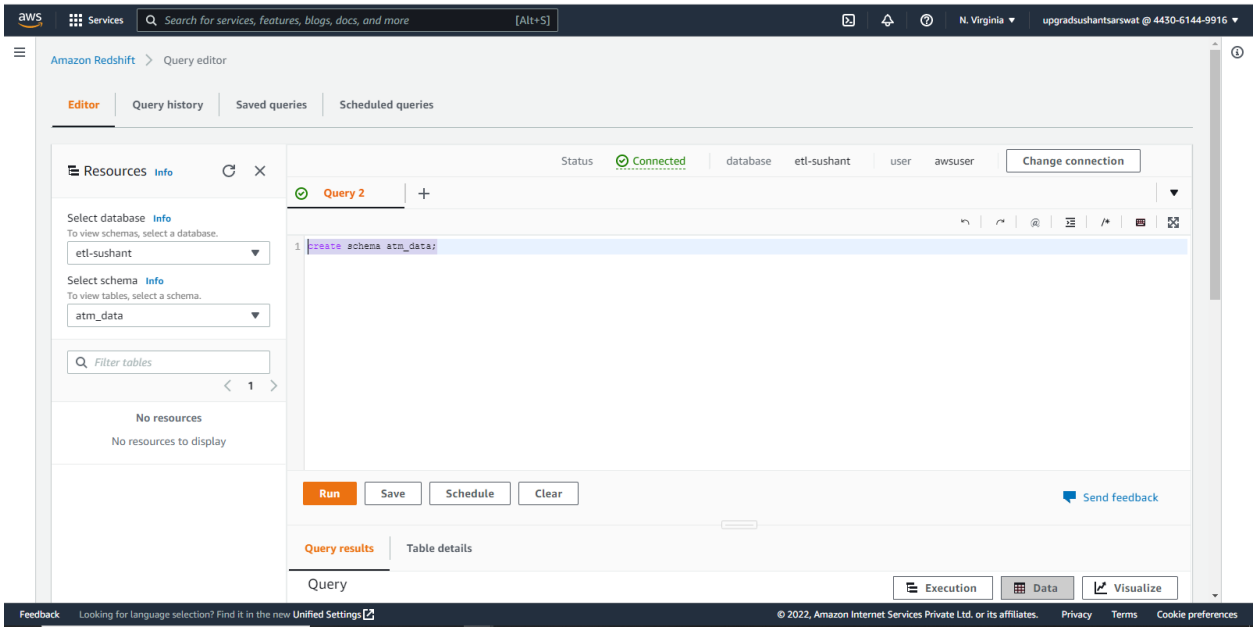
Setting up a database in the Redshift cluster and running queries to create the dimension and fact tables

Viewing all data in Amazon S3 Bucket



Query to create a schema for dimensions and fact tables:

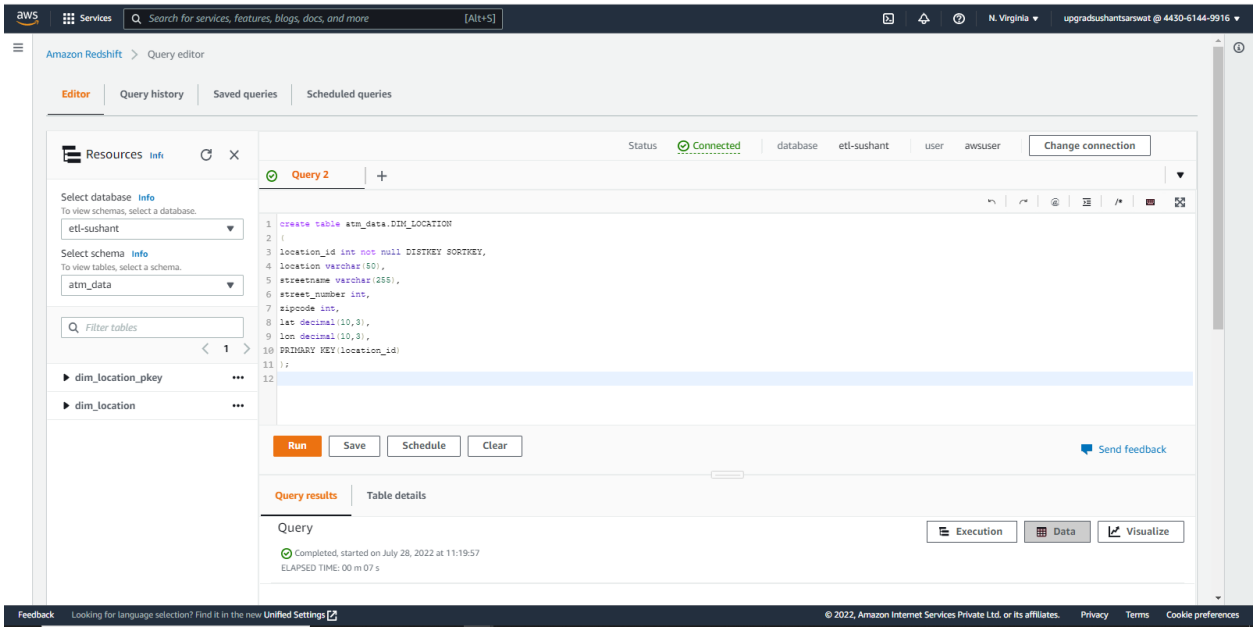
```
create schema atm_data;
```



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

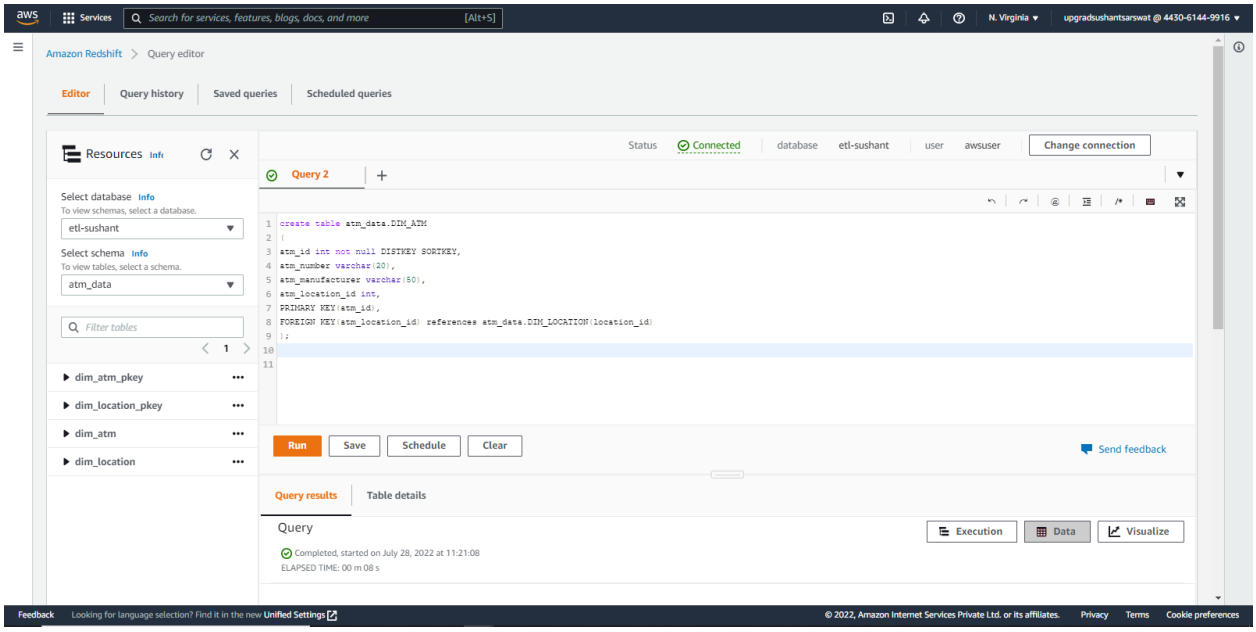
- **Creating location dimension table**

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



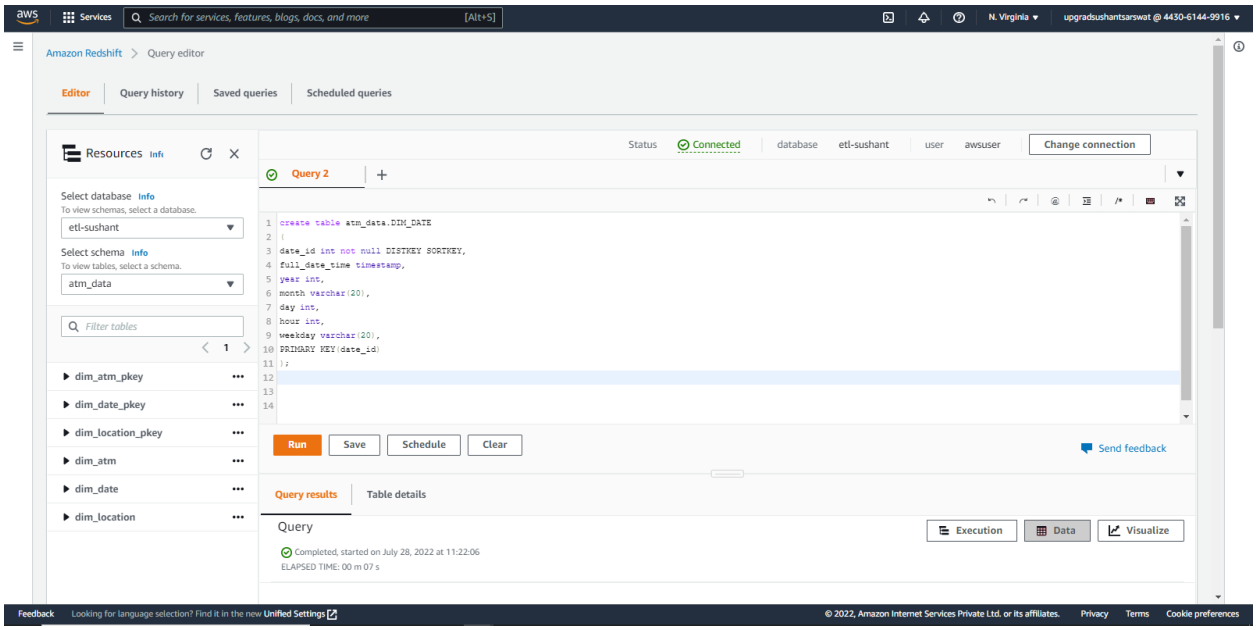
- **Creating atm dimension table**

```
create table atm_data.DIM_ATM
(
atm_id int not null DISTKEY SORTKEY,
atm_number varchar(20),
atm_manufacturer varchar(50),
atm_location_id int,
PRIMARY KEY(atm_id),
FOREIGN KEY(atm_location_id) references atm_data.DIM_LOCATION(location_id)
);
```



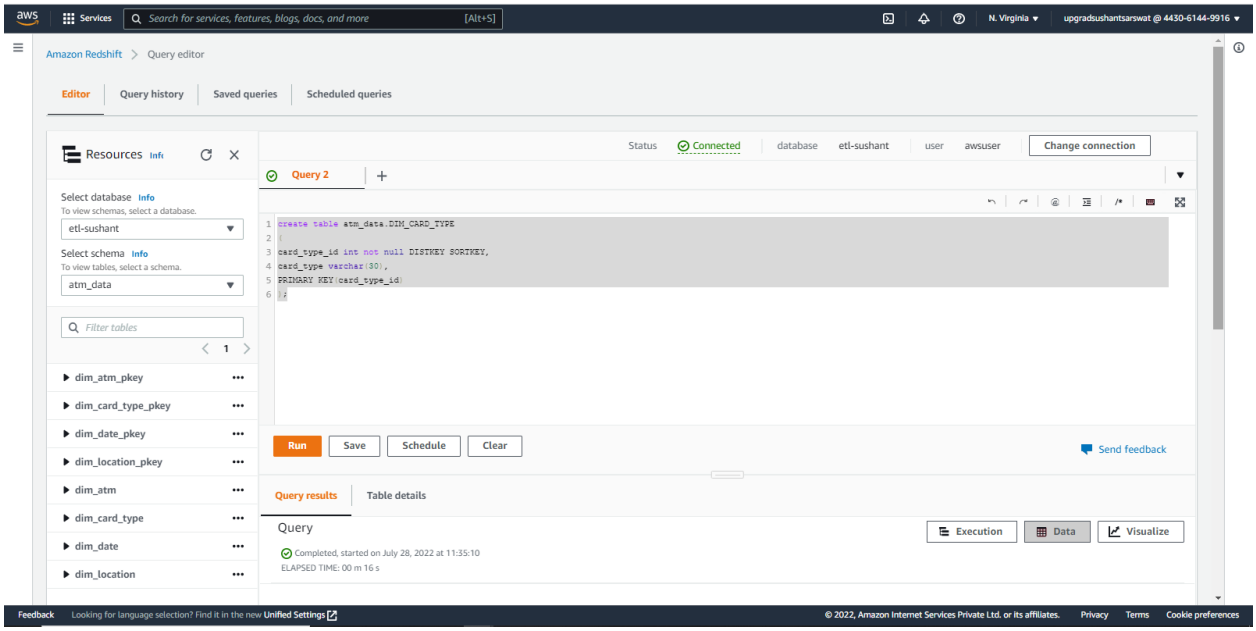
- **Creating date dimension table**

```
create table atm_data.DIM_DATE
(
date_id int not null DISTKEY SORTKEY,
full_date_time timestamp,
year int,
month varchar(20),
day int,
hour int,
weekday varchar(20),
PRIMARY KEY(date_id)
);
```



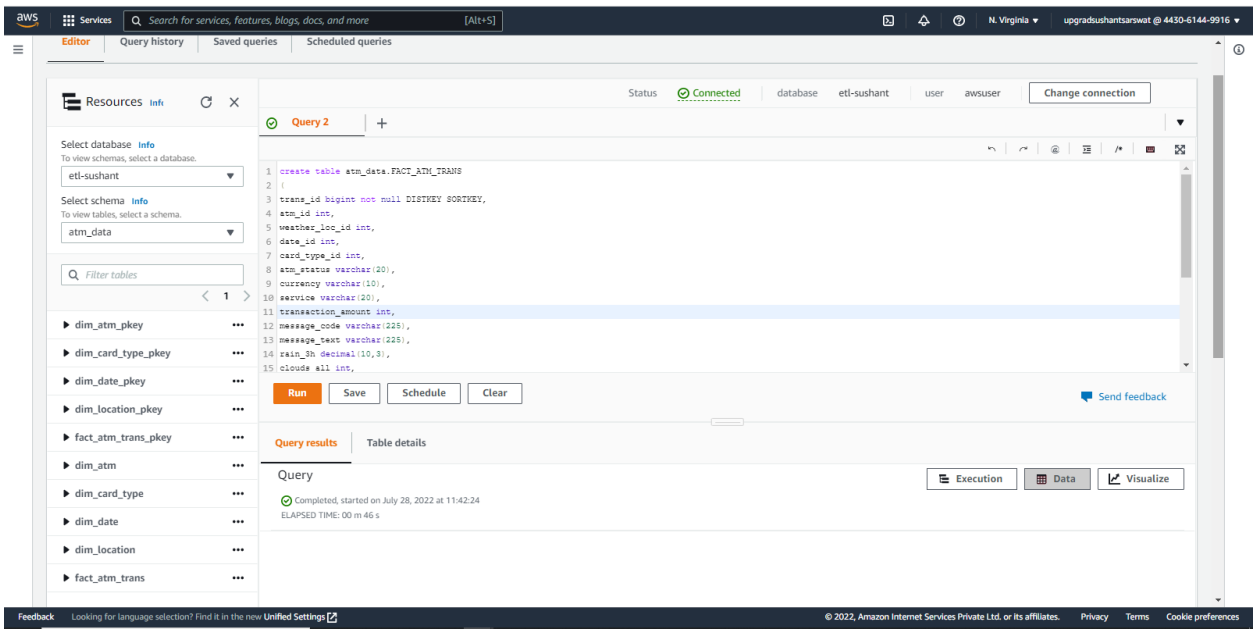
- **Creating card type dimension table**

```
create table atm_data.DIM_CARD_TYPE
(
card_type_id int not null DISTKEY SORTKEY,
card_type varchar(30)
PRIMARY KEY(card_type_id)
);
```



- **Creating atm transactions fact 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)
);
```

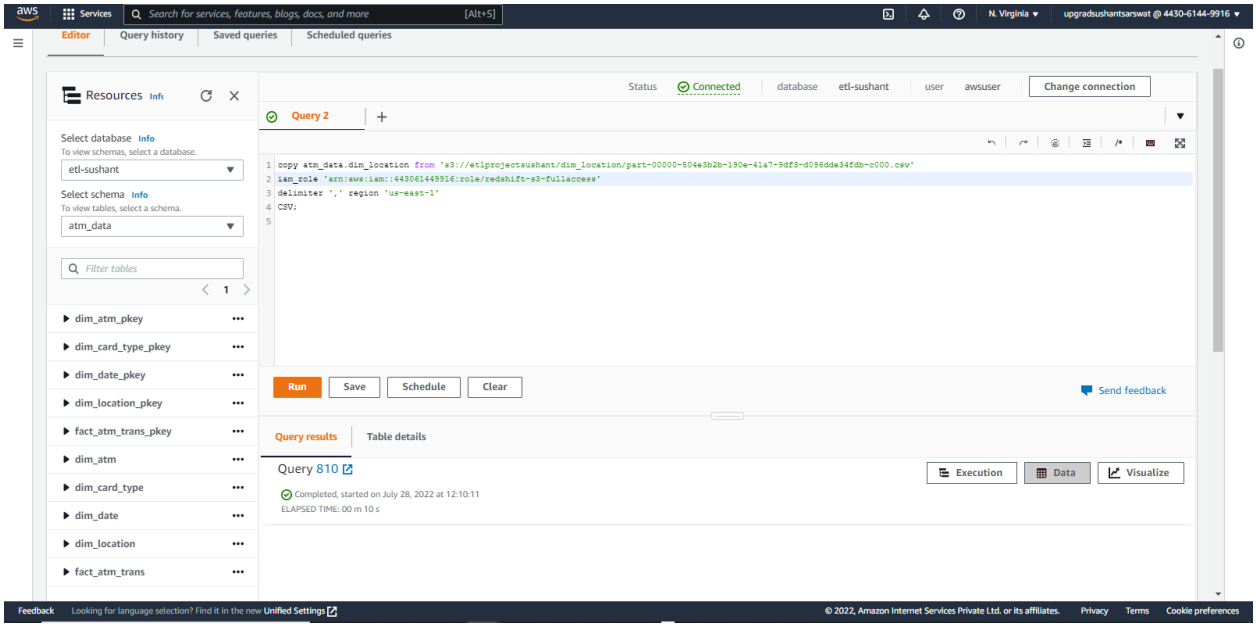


Loading data into a Redshift cluster from Amazon S3 bucket

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

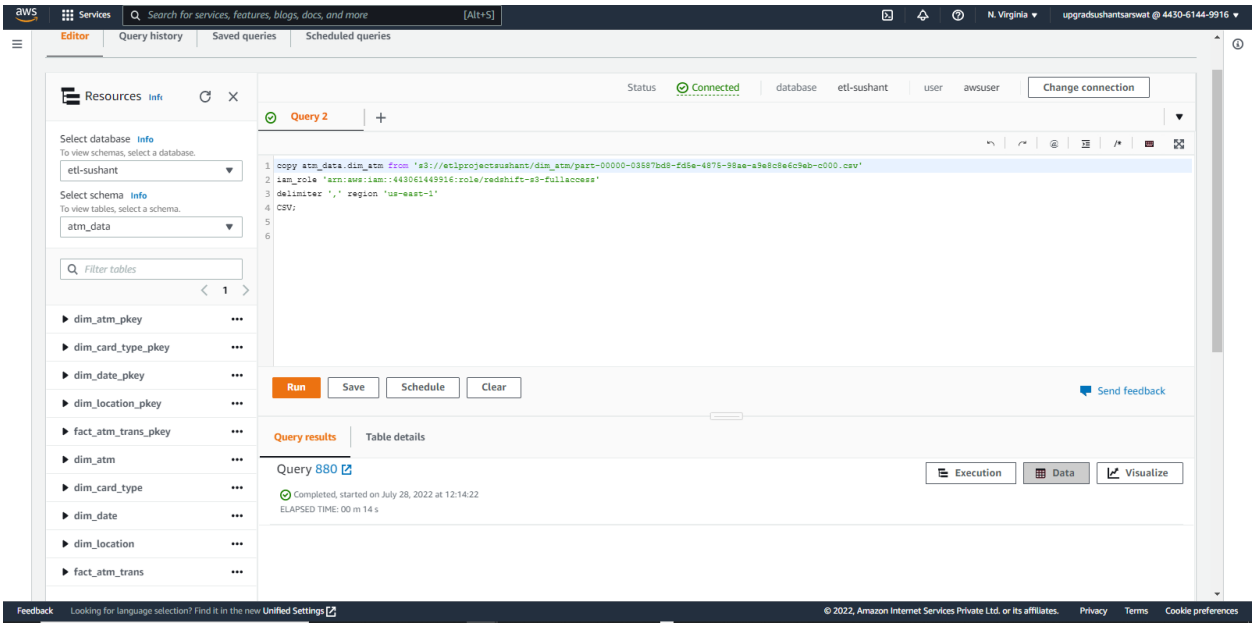
- **Copying the data to dim_location table**

```
copy atm_data.dim_location from 's3://etlprojectsushant/dim_location/part-00000-504e3b2b-190e-41a7-9df3-d096dde34fdb-c000.csv'
iam_role 'arn:aws:iam::443061449916:role/redshift-s3-fullaccess'
delimiter ',' region 'us-east-1'
CSV;
```



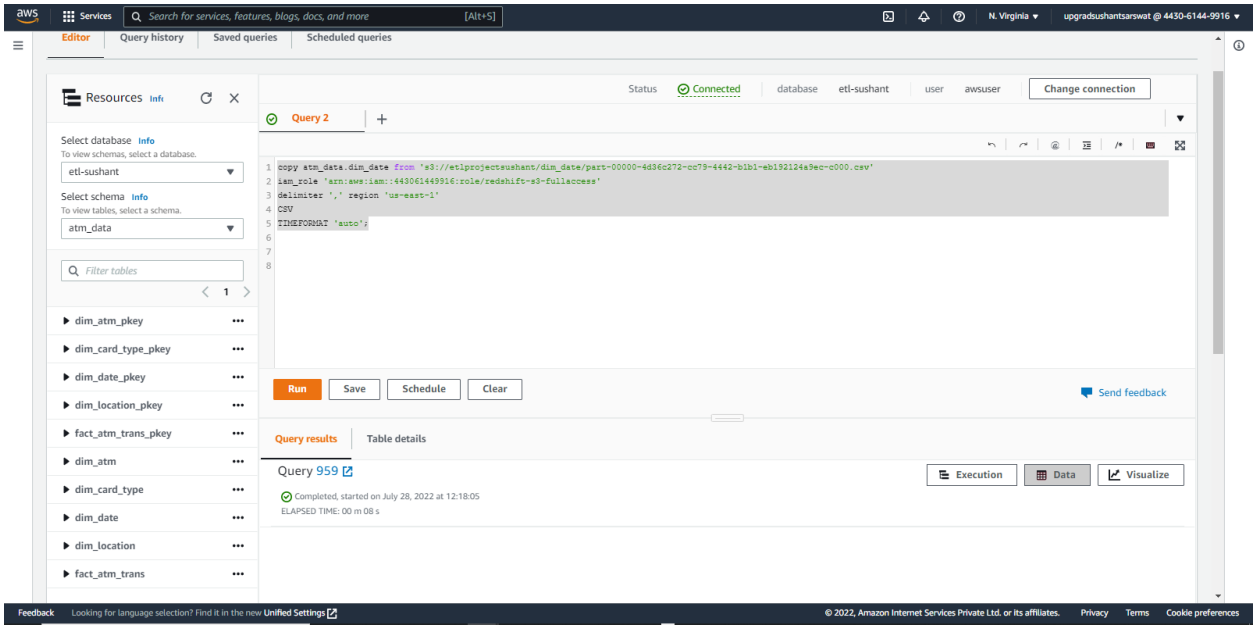
- **Copying the data to dim_atm table**

```
copy atm_data.dim_atm from 's3://etlprojectsushant/dim_atm/part-00000-03587bd8-fd5e-4875-98ae-a9e8c8e6c9eb-c000.csv'
iam_role 'arn:aws:iam::443061449916:role/redshift-s3-fullaccess'
delimiter ',' region 'us-east-1'
CSV;
```



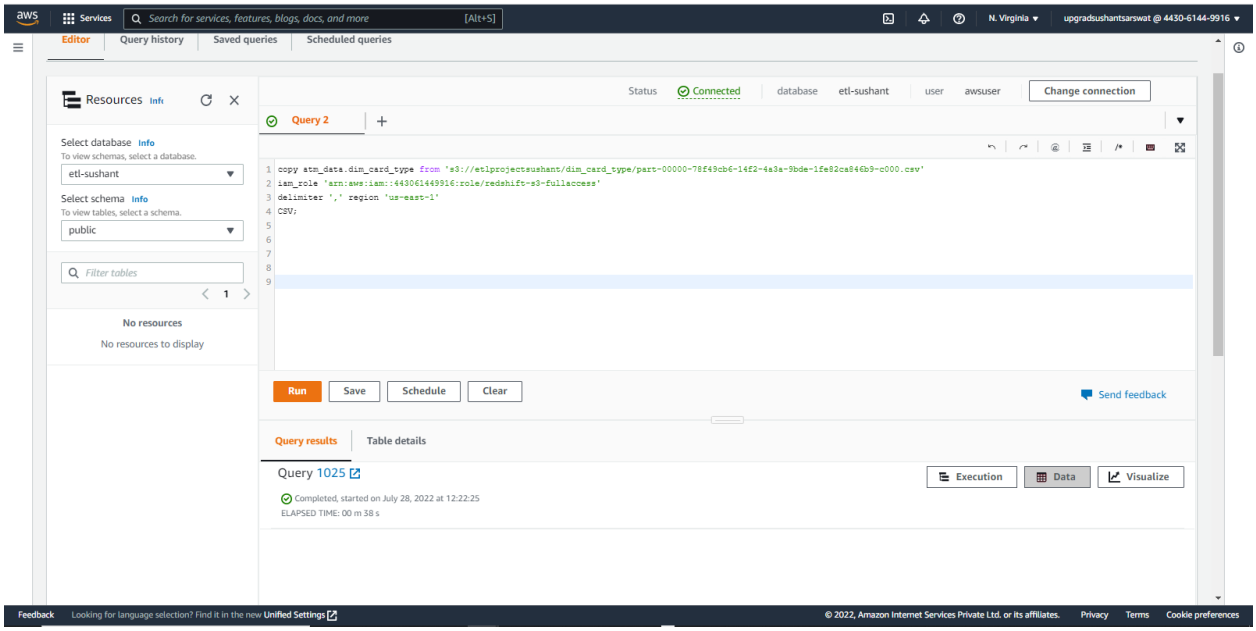
- **Copying the data to dim_date table**

```
copy atm_data.dim_date from 's3://etlprojectsushant/dim_date/part-00000-4d36c272-cc79-4442-b1b1-eb192124a9ec-c000.csv'
iam_role 'arn:aws:iam::443061449916:role/redshift-s3-fullaccess'
delimiter ',' region 'us-east-1'
CSV
TIMEFORMAT 'auto';
```



- **Copying the data to dim_card_type table**

```
copy atm_data.dim_card_type from 's3://etlprojectsushant/dim_card_type/part-00000-78f49cb6-14f2-4a3a-9bde-1fe82ca846b9-c000.csv'
iam_role 'arn:aws:iam::443061449916:role/redshift-s3-fullaccess'
delimiter ',' region 'us-east-1'
CSV;
```



- **Copying the data to fact_atm_trans table**

```
copy atm_data.fact_atm_trans from 's3://etlprojectsushant/fact_atm_trans/part-00000-a2c998ee-bb0f-4278-89e6-8cfdd11f576c-c000.csv'
iam_role 'arn:aws:iam::443061449916:role/redshift-s3-fullaccess'
delimiter ',' region 'us-east-1'
CSV;
```