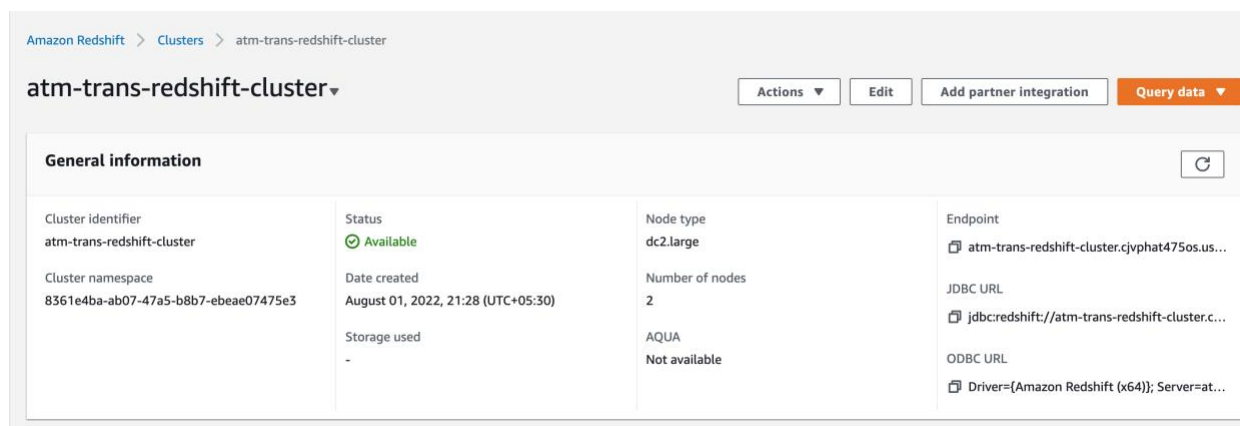Creation of a Redshift Cluster

1. **Screenshots of the configuration of the Redshift cluster that you have created:**

<Screenshot of the type of machine used along with number of nodes>



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

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

   **Schema creation:**
   create schema atmTransData;

   **Location Dimension table creation:**
   create table atmTransData.dim_location(
   location_id INT,
   location VARCHAR(50),
   streetname VARCHAR(255),
   street_number INT,
   zipcode INT,
   lat DECIMAL(10,3),
   lon DECIMAL(10,3));

**Atm Dimension  table creation:**
```
create table atmTransData.dim_atm(
atm_id INT,
atm_number VARCHAR(20),
atm_manufacturer VARCHAR(50),
atm_location_id INT);
```

**Date Dimension table creation:**
```
create table atmTransData.dim_date(
date_id INT,
full_date_time TIMESTAMP,
year INT,
month VARCHAR(20),
day INT,
hour INT,
weekday VARCHAR(20));
```
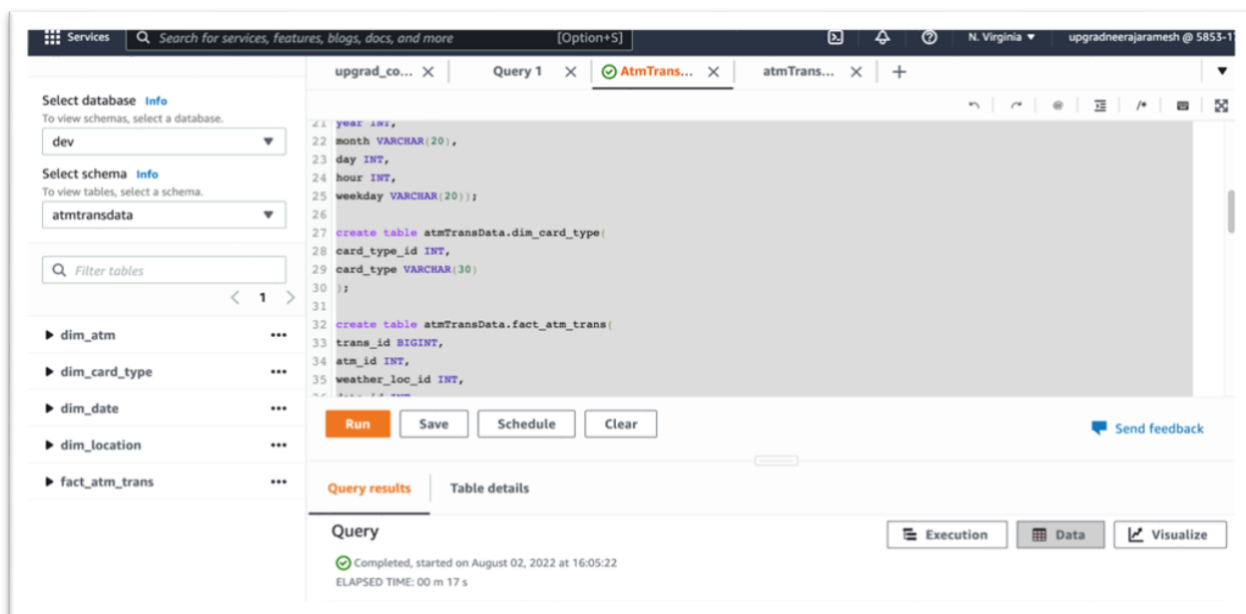
**Card Dimension table creation:**
```
create table atmTransData.dim_card_type(
card_type_id INT,
card_type VARCHAR(30)
);
```

**Fact_Atm_Trans fact table creation:**
```
create table atmTransData.fact_atm_trans(
trans_id BIGINT,
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(1000),
message_text VARCHAR(1000),
rain_3h DECIMAL(10,3),
clouds_all INT,
weather_id INT,
weather_main VARCHAR(50),
weather_description VARCHAR(255));
```

Screenshot of success state of query

3. 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**

copy atmTransData.dim_location from
's3://atmtransdata/dimLoc/part-00000-16949c74-9c28-493b-af2e-f818c66388e4-c000.csv'
iam_role 'arn:aws:iam::585317706400:role/redshift_s3_full_access'
delimiter ',' region 'us-east-1' IGNOREHEADER 1;

 copy atmTransData.dim_atm from
's3://atmtransdata/dimAtm/part-00000-2d0bee25-6e48-4564-a96d-77c320ea95a3-c000.csv'
iam_role 'arn:aws:iam::585317706400:role/redshift_s3_full_access'
delimiter ',' region 'us-east-1' IGNOREHEADER 1;

copy atmTransData.dim_card_type from
's3://atmtransdata/dimCard/part-00000-f764053d-601d-436e-b42b-19942e3d1e3d-c000.csv'
iam_role 'arn:aws:iam::585317706400:role/redshift_s3_full_access'
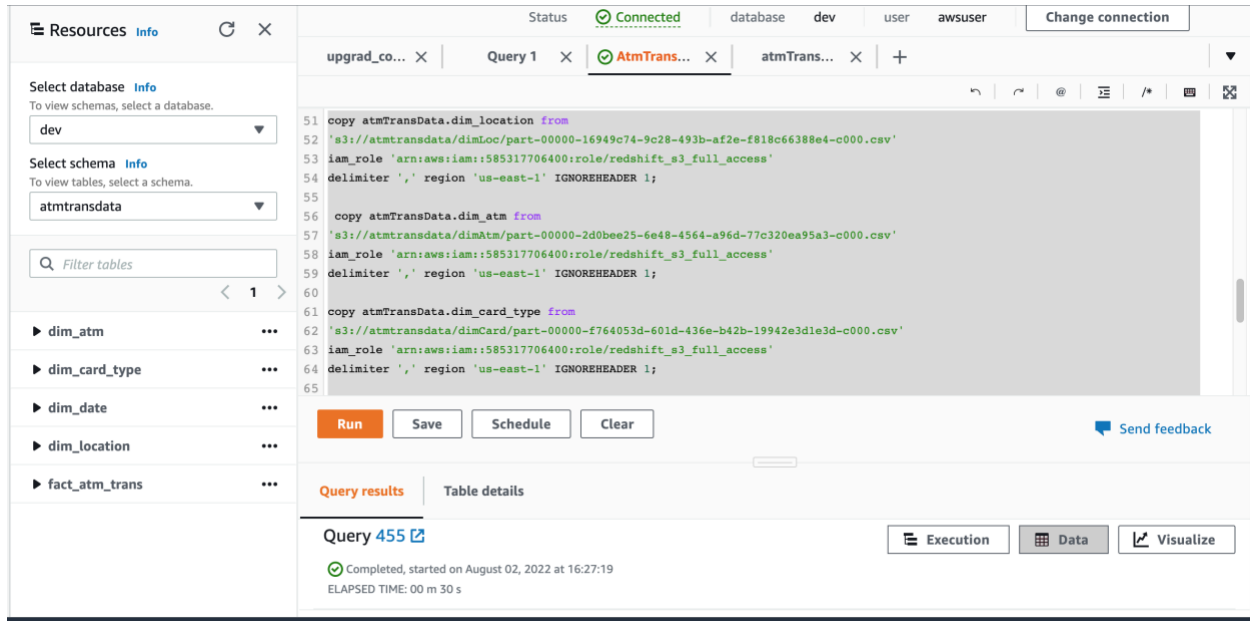delimiter ',' region 'us-east-1' IGNOREHEADER 1;

copy atmTransData.dim_date from
's3://atmtransdata/dimDate/part-00000-dadd0ed7-9b39-4db0-9ce1-dd50fae07258-c000.csv'
iam_role 'arn:aws:iam::585317706400:role/redshift_s3_full_access'
timeformat 'auto'
delimiter ',' region 'us-east-1' IGNOREHEADER 1;

copy atmTransData.fact_atm_trans from
's3://atmtransdata/factAtmTrans/part-00000-acb5589d-8ef4-475d-be0d-177eb086bdfd-c000.csv'
iam_role 'arn:aws:iam::585317706400:role/redshift_s3_full_access'
delimiter ',' region 'us-east-1' IGNOREHEADER 1;

**Queries to verify count in dimension and fact tables:**

select count(1) from atmTransData.dim_location;
select count(1) from atmTransData.dim_card_type;
select count(1) from atmTransData.dim_atm;
select count(1) from atmTransData.dim_date;
select count(1) from atmTransData.fact_atm_trans;

**Success screenshot:**