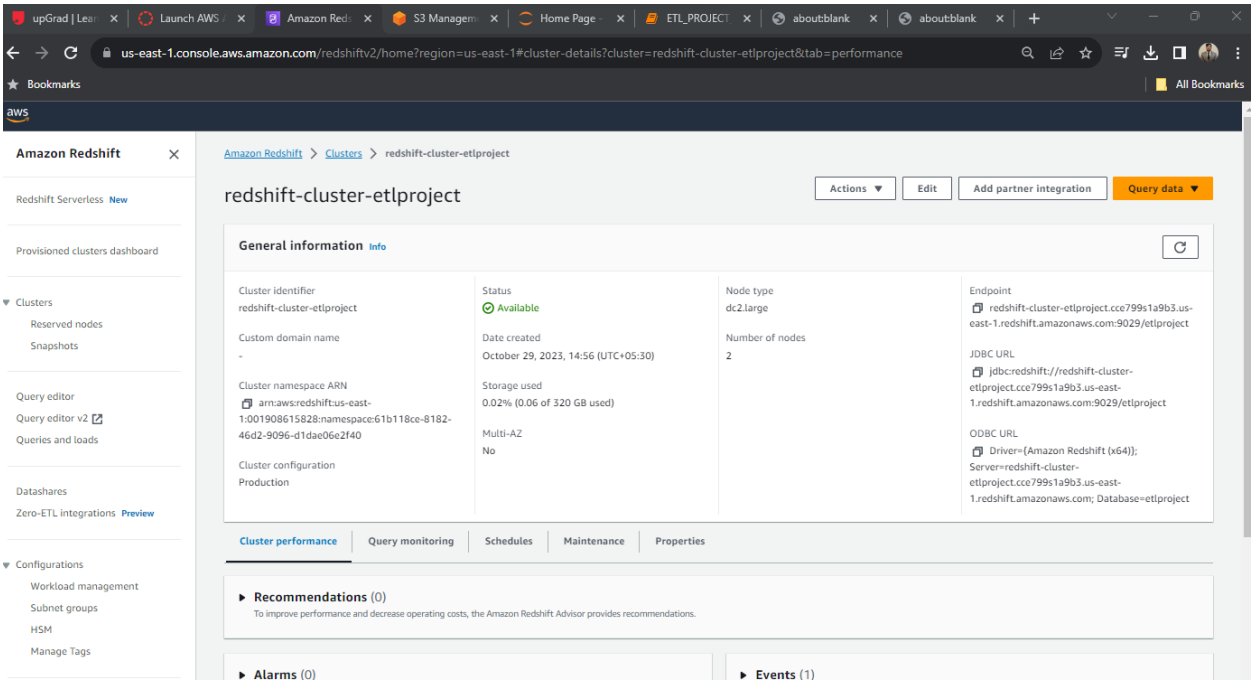
Creation of a Redshift Cluster

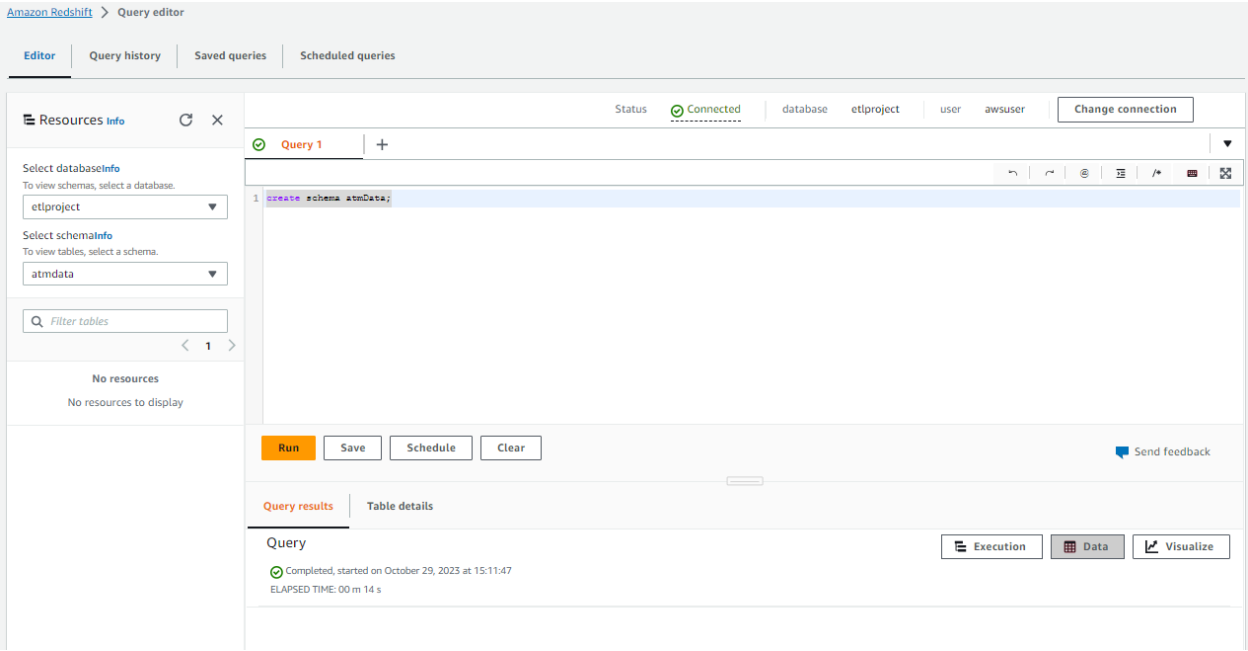
Screenshots of the configuration of the Redshift cluster that you have created:
Screenshot of the type of machine used along with number of nodes



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

Creating Schema:
Query:

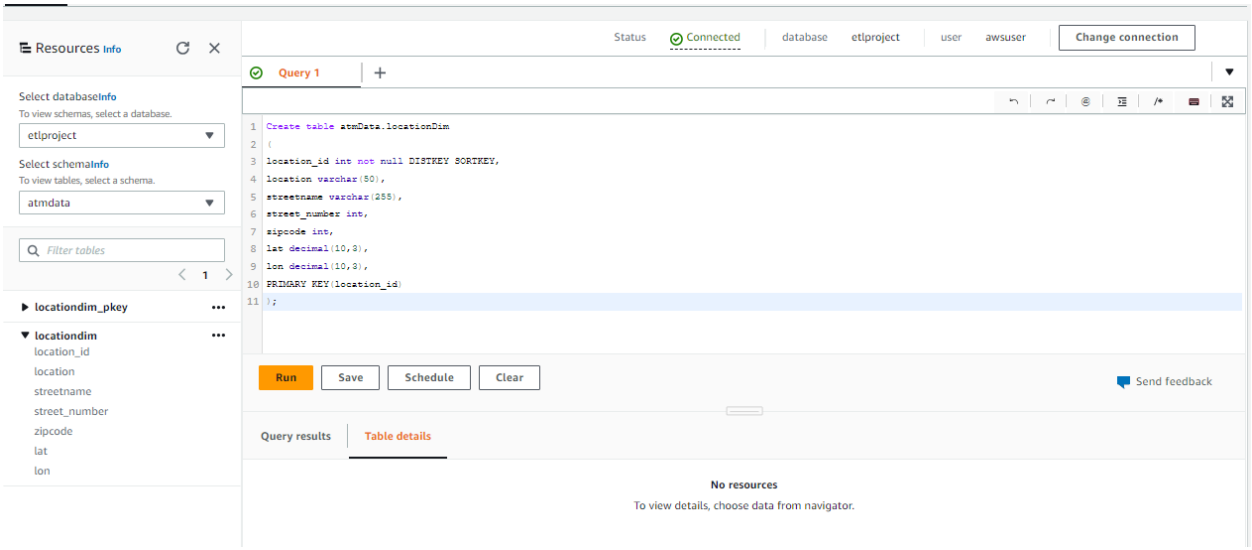
create schema atmData;



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

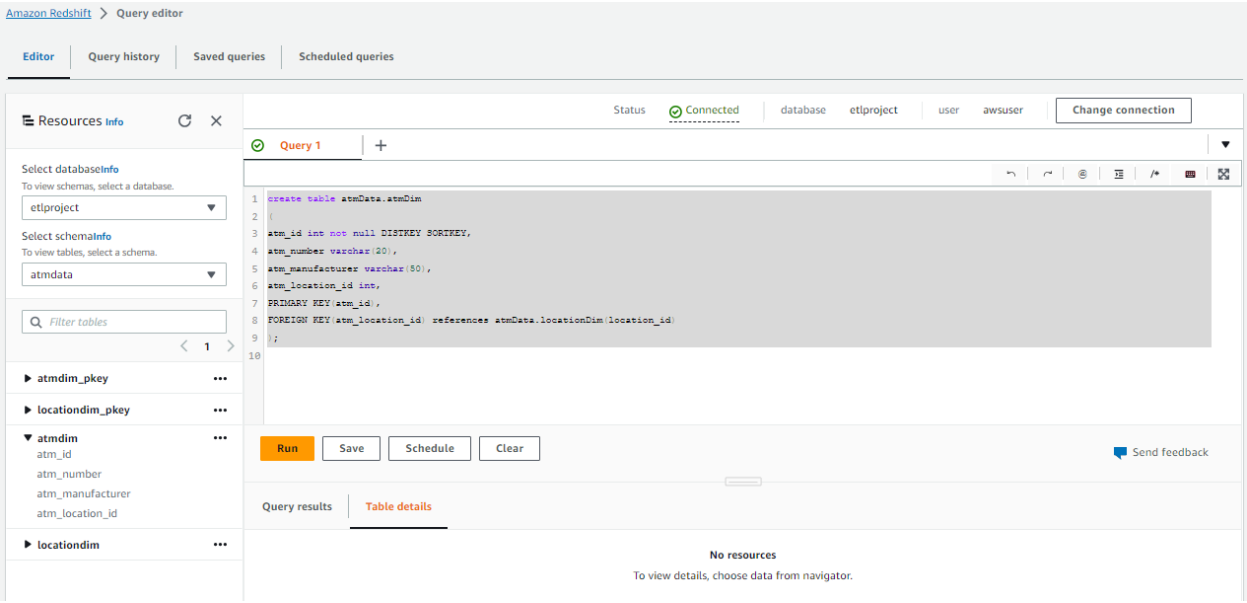
- Creating location dimension table:

```
Create table atmData.locationDimw
(
location varchar(255),
streetname varchar(255),
street_number int,
zipcode int,
lat decimal(10,3),
lon decimal(10,3),
location_id int not null DISTKEY SORTKEY,
PRIMARY KEY(location_id)
);
```

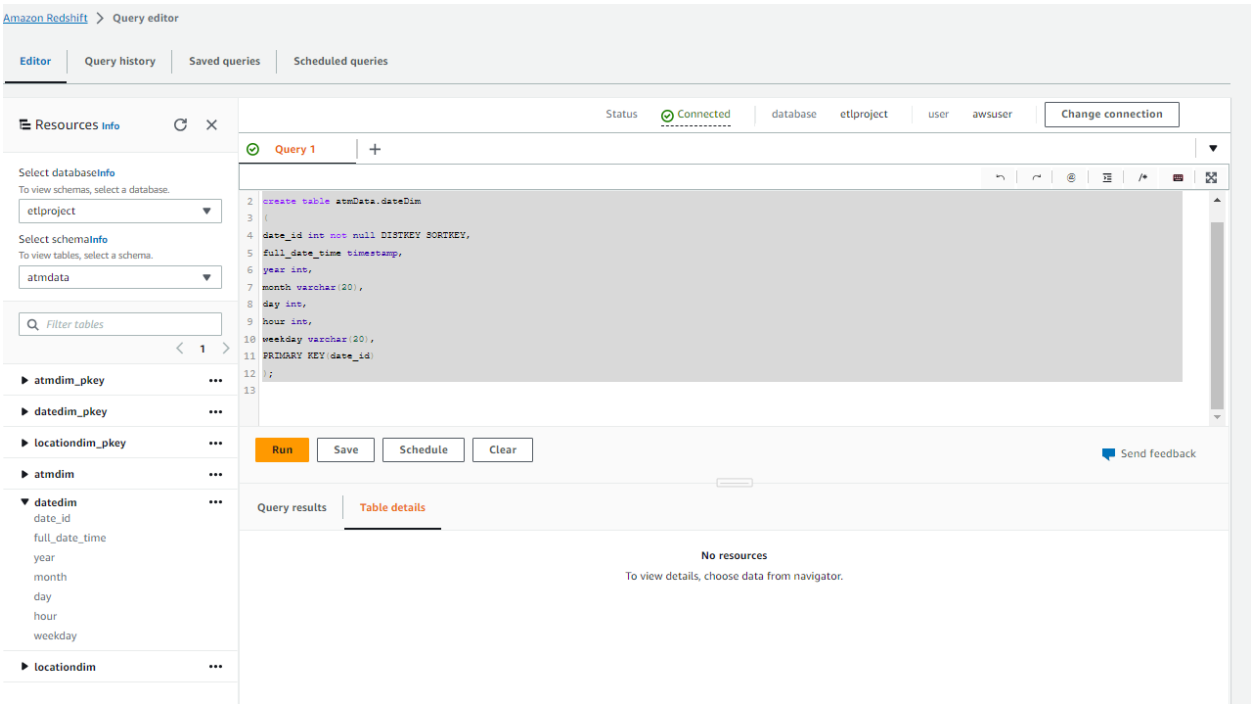


- Creating atm dimension table

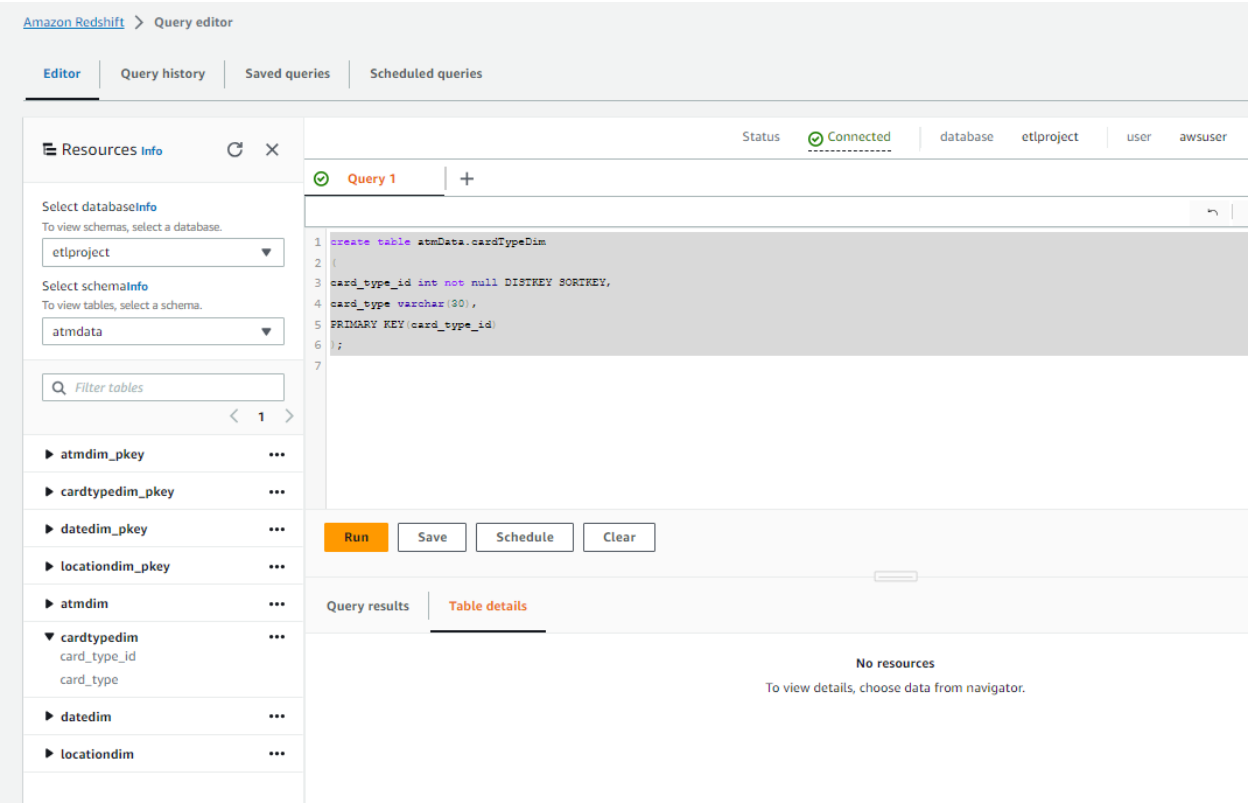
  create table atmData.atmDim
  (
  atm_id int not null DISTKEY SORTKEY,
  atm_number varchar(20),
  atm_manufacturer varchar(50),
  atm_location_id int,
  PRIMARY KEY(atm_id),
  FOREIGN KEY(atm_location_id) references atmData.locationDim(location_id)
  );



- Creating date dimension table
  create table atmData.dateDim
  (
  year int,
  month varchar(20),
  day int,
  hour int,
  weekday varchar(20),
  full_date_time timestamp,
  date_id int not null DISTKEY SORTKEY,
  PRIMARY KEY(date_id)
  );

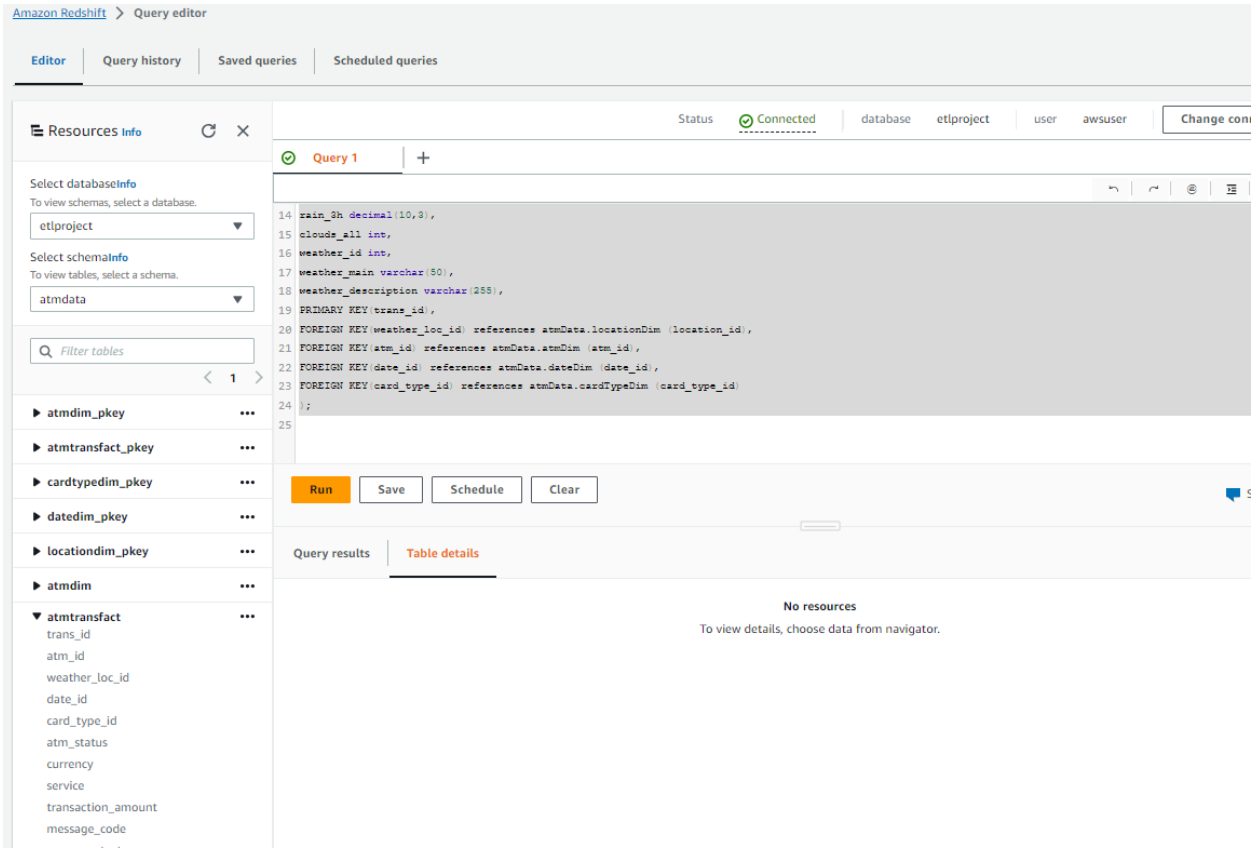


- Creating card type dimension table
  create table atmData.cardTypeDim
  (
  card_type_id int not null DISTKEY SORTKEY,
  card_type varchar(30),
  PRIMARY KEY(card_type_id)
  );



- Creating atm transactions fact table
  create table atmData.atmTransFact
  (
  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 atmData.locationDim (location_id),
  FOREIGN KEY(atm_id) references atmData.atmDim (atm_id),

FOREIGN KEY(date_id) references atmData.dateDim (date_id),
FOREIGN KEY(card_type_id) references atmData.cardTypeDim (card_type_id)
);



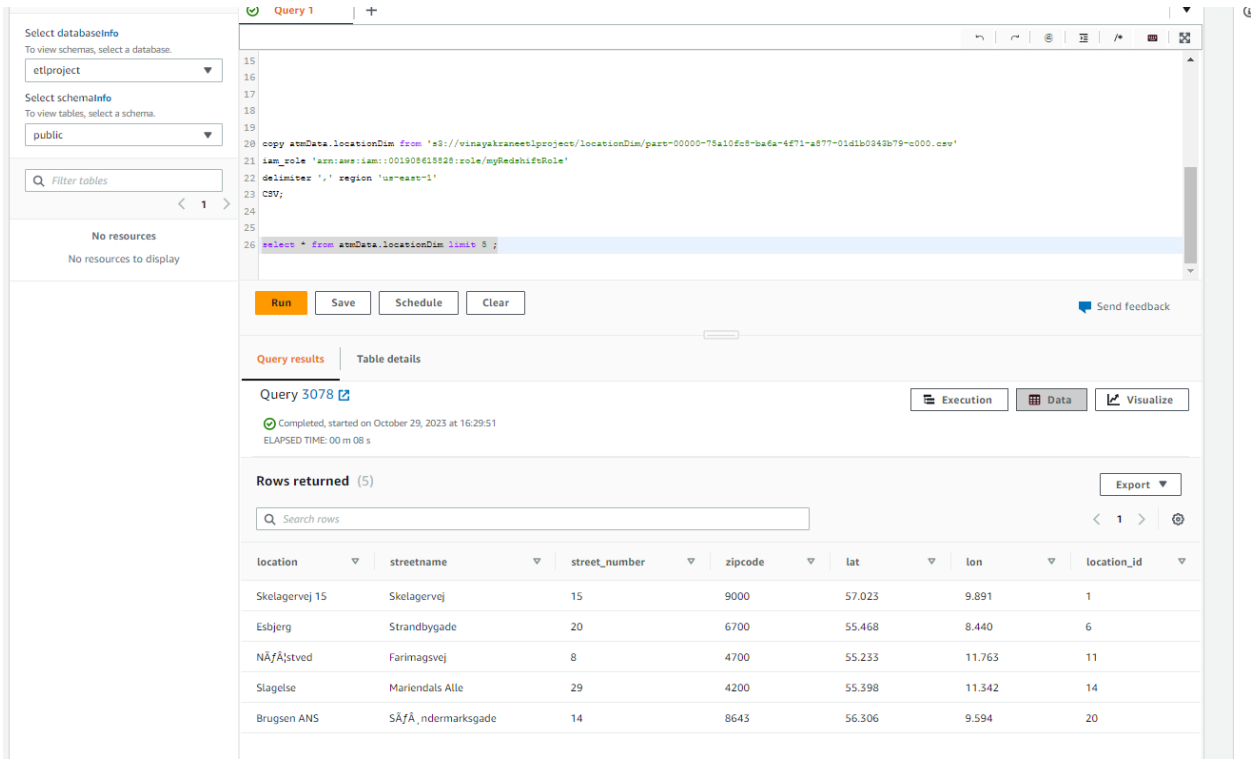
Loading data into a Redshift cluster from Amazon S3 bucket

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

Queries:

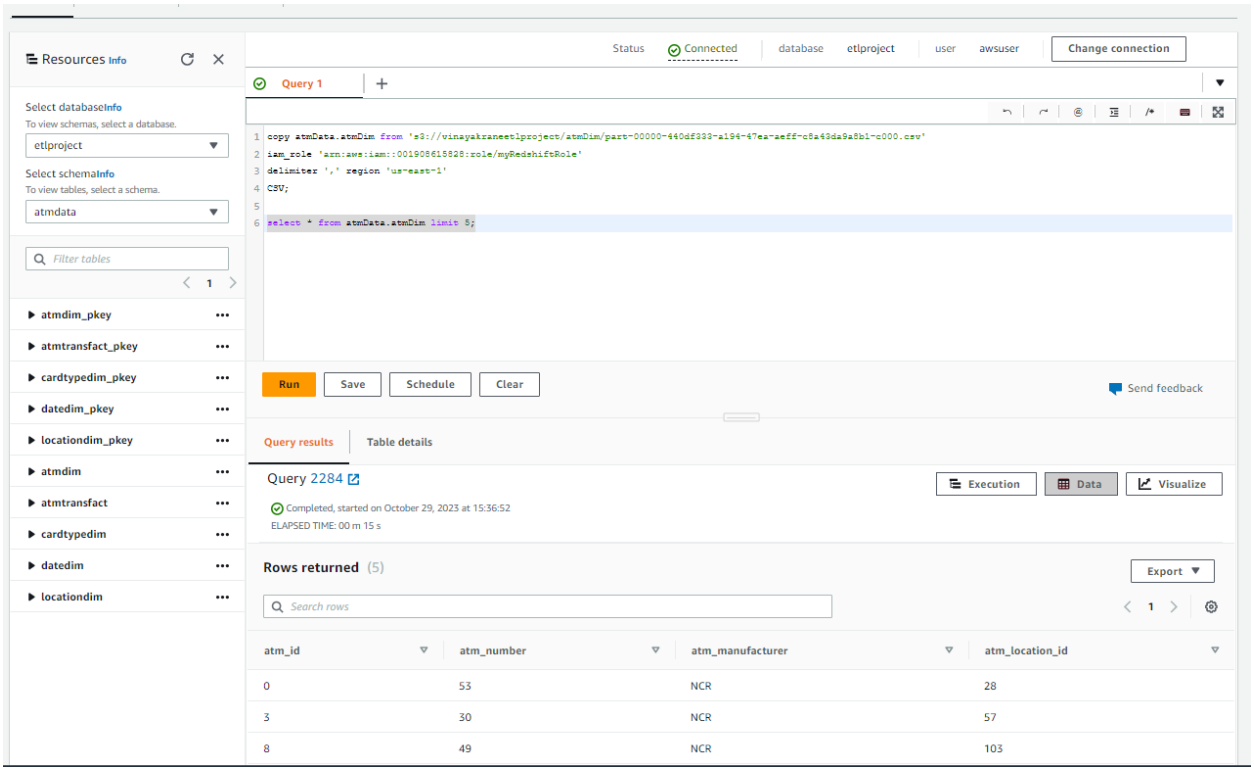
- **Copying the data to locationDim table**

```
copy atmData.locationDim from 's3://vinayakraneetlproject/locationDim/part-00000-
b220c916-33e4-4550-bc76-777be5d02a7d-c000.csv'
iam_role 'arn:aws:iam::001908615828:role/myRedshiftRole'
delimiter ',' region 'us-east-1'
CSV;
```



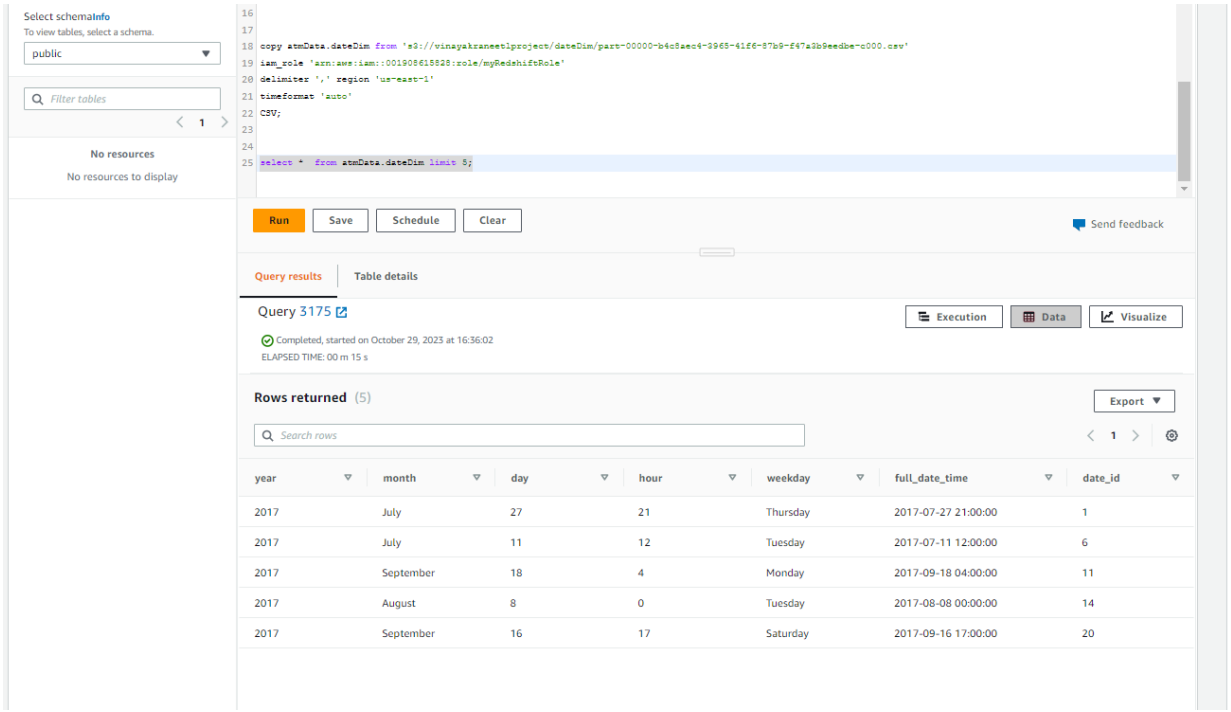
- **Copying the data to atmDim table**

  copy atmData.atmDim from 's3://vinayakraneetlproject/atmDim/part-00000-440df333-a194-47ea-aeff-c8a43da9a8b1-c000.csv'
  iam_role 'arn:aws:iam::001908615828:role/myRedshiftRole'
  delimiter ',' region 'us-east-1'
  CSV;



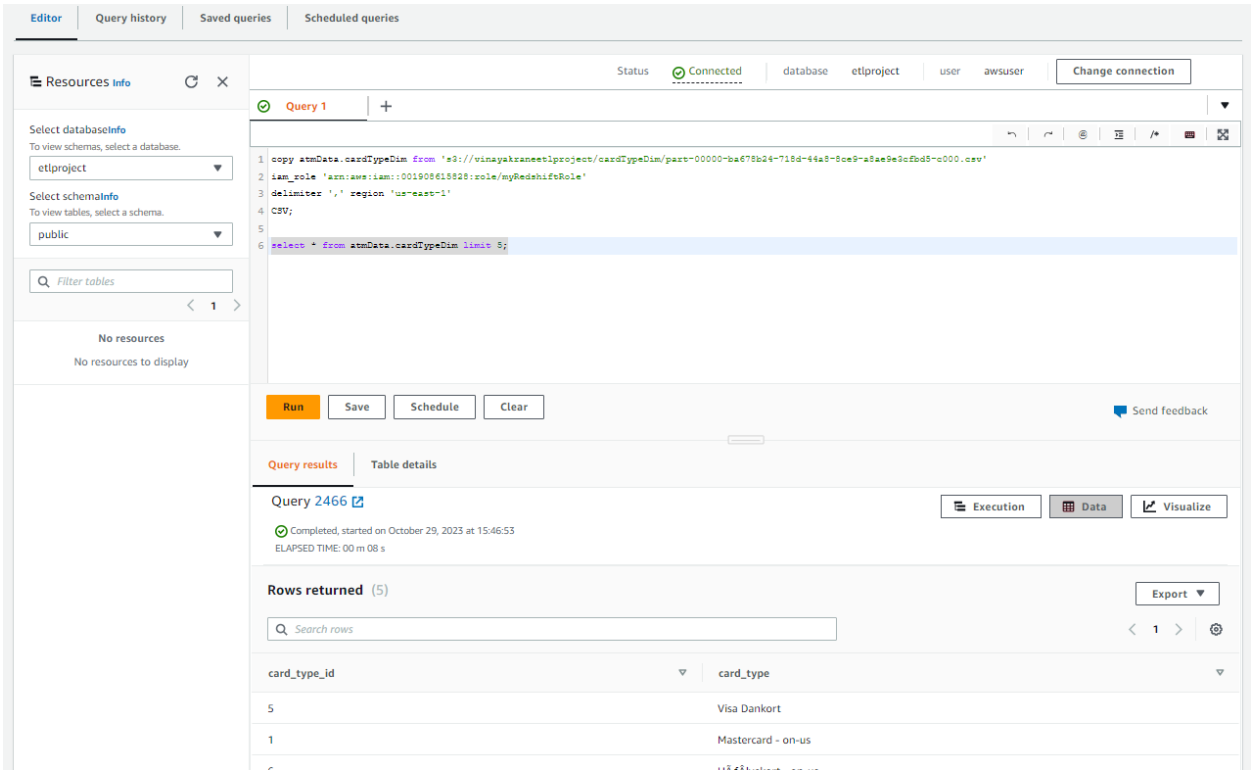
- **Copying the data to dateDim table**

  copy atmData.dateDim from 's3://vinayakraneetlproject/dateDim/part-00000-b4c8aec4-3965-41f6-87b9-f47a3b9eedbe-c000.csv'
  iam_role 'arn:aws:iam::001908615828:role/myRedshiftRole'
  delimiter ',' region 'us-east-1'
  timeformat 'auto'
  CSV;



- **Copying the data to cardTypeDim table**

  copy atmData.cardTypeDim from 's3://vinayakraneetlproject/cardTypeDim/part-00000-ba678b24-718d-44a8-8ce9-a8ae9e3cfbd5-c000.csv'
  iam_role 'arn:aws:iam::001908615828:role/myRedshiftRole'
  delimiter ',' region 'us-east-1'
  CSV;



- **Copying the data to atmTransFact table**

  copy atmData.atmTransFact from 's3://vinayakraneetlproject/atmTransFact/part-00000-9b72c7a0-0fa7-4260-9399-31c61edd9b6b-c000.csv'
  iam_role 'arn:aws:iam::001908615828:role/myRedshiftRole'
  delimiter ',' region 'us-east-1'
  CSV;