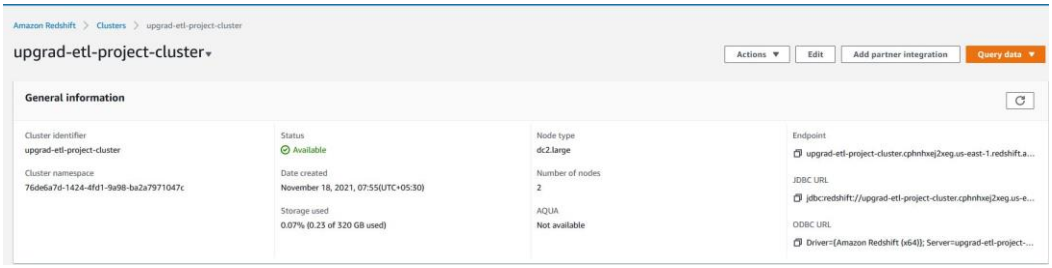
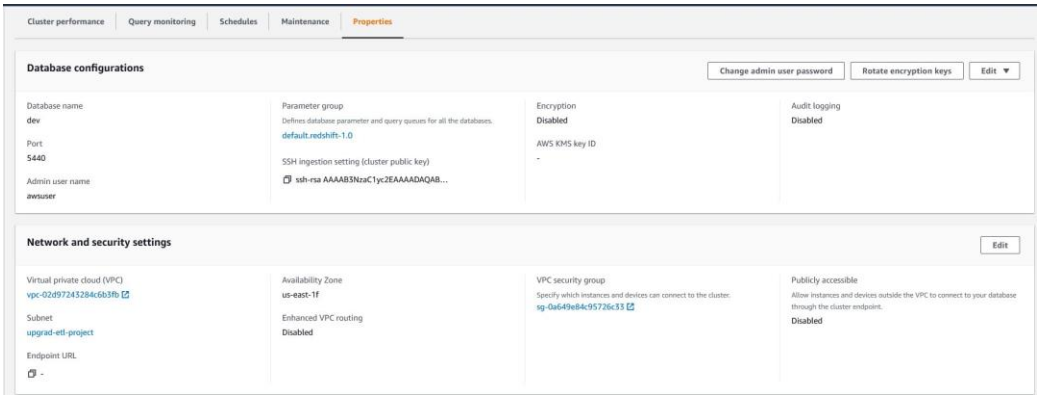
Creation of a Redshift Cluster

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

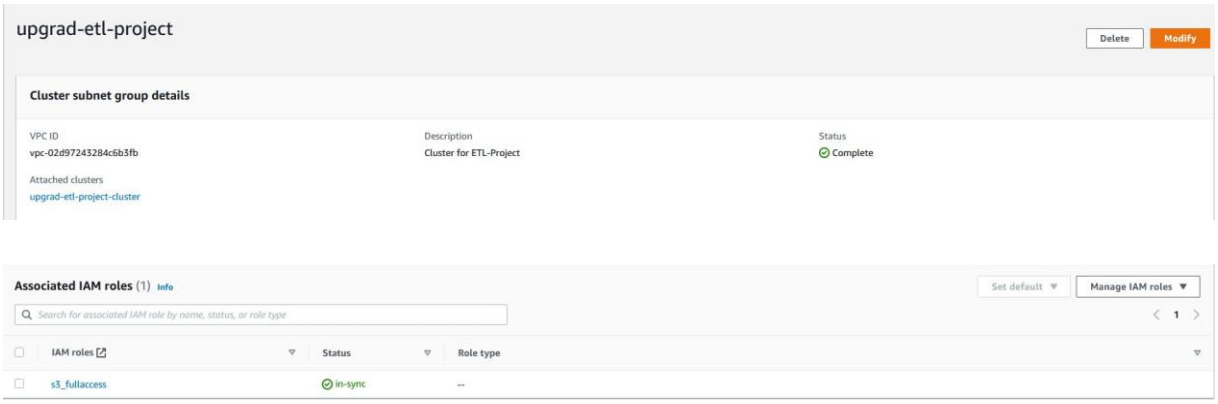
1. Screenshot for type of machine and number of nodes used.



2. Screenshot for database configurations and network properties.



3. Screenshots subnet group and associated IAM rules.

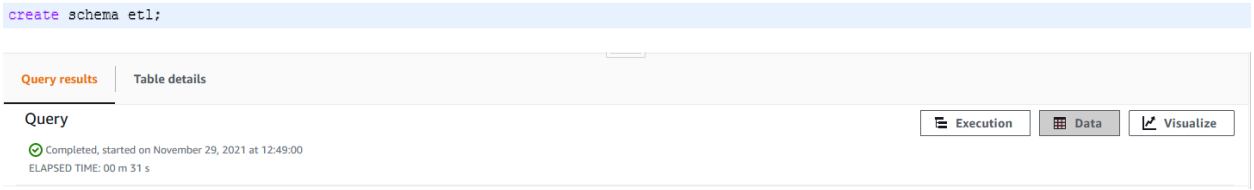


Setting up a database in the Redshift cluster and running queries to create the dimension and fact tables

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

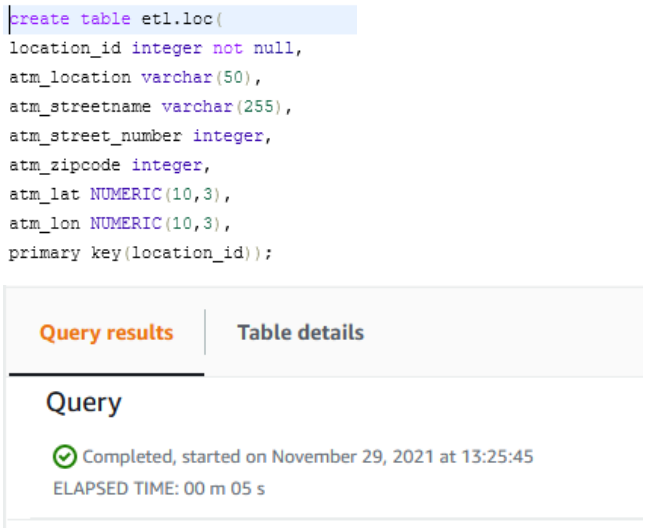
1. Query to create schema.

   create schema etl;



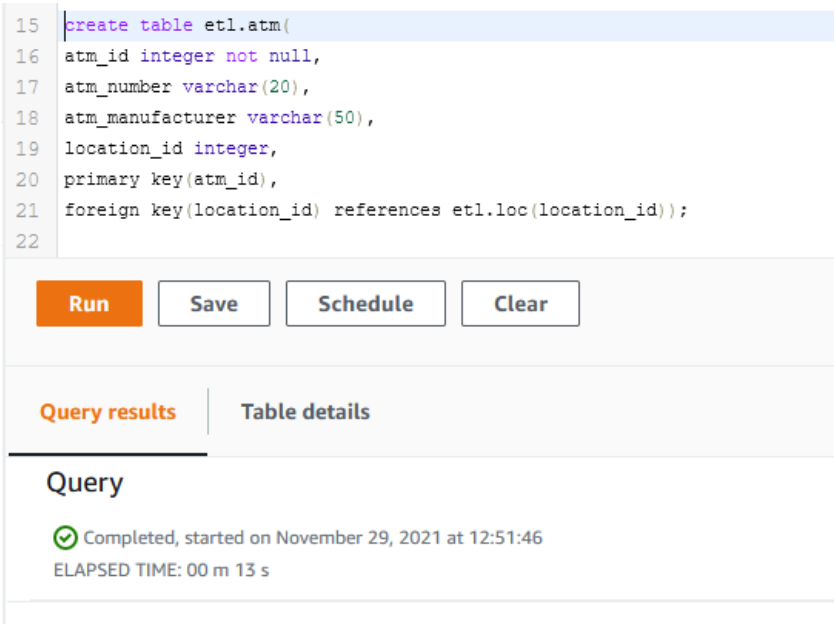
2. Query to create LOCATION dimension table.

   create table etl.loc(
   location_id integer not null,
   atm_location varchar(50),
   atm_streetname varchar(255),
   atm_street_number integer,
   atm_zipcode integer,
   atm_lat NUMERIC(10,3),
   atm_lon NUMERIC(10,3),
   primary key(location_id));



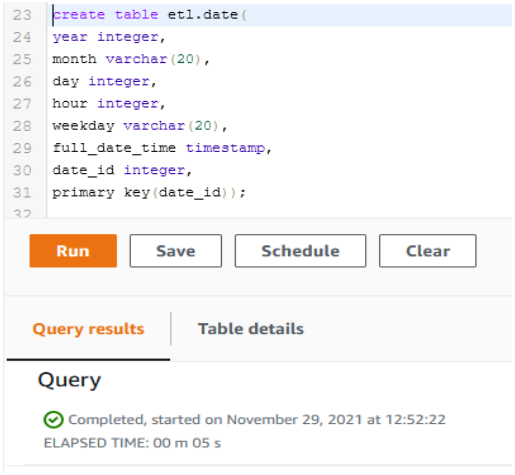
3. Query to create ATM dimension table.

```
create table etl.atm(
atm_id integer not null,
atm_number varchar(20),
atm_manufacturer varchar(50),
location_id integer,
primary key(atm_id),
foreign key(location_id) references etl.loc(location_id));
```



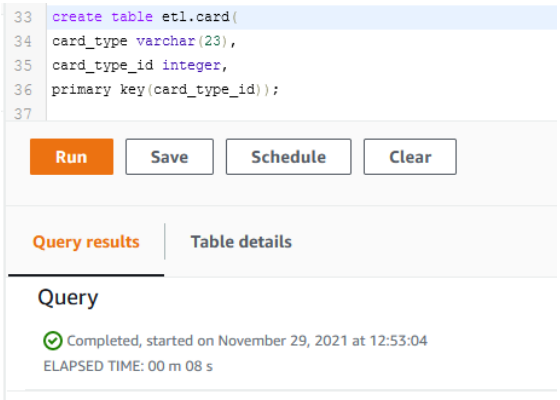
4. Query to create DATE dimension table.

```
create table etl.date(
year integer,
month varchar(20),
day integer,
hour integer,
weekday varchar(20),
full_date_time timestamp,
date_id integer,
primary key(date_id));
```



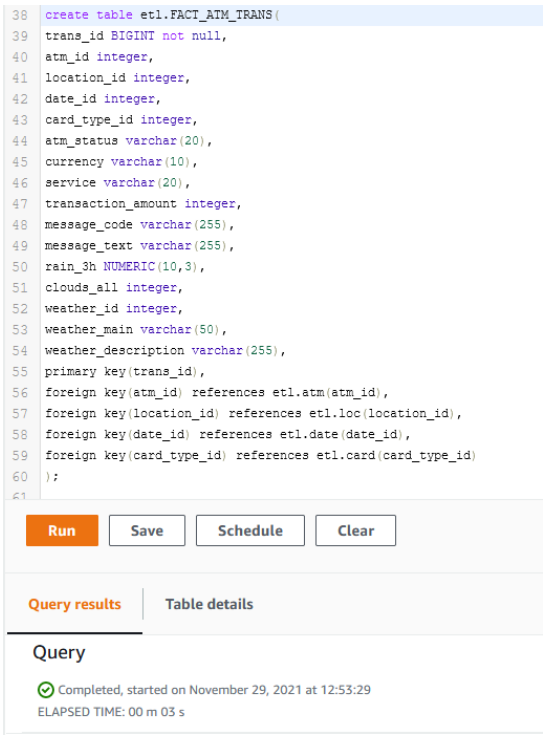
5. Query to create CARD dimension table.

```
create table etl.card(
card_type varchar(23),
card_type_id integer,
primary key(card_type_id));
```



6. Query to create FACT_ATM_TRANS table.

```
create table etl.FACT_ATM_TRANS(
trans_id BIGINT not null,
atm_id integer,
location_id integer,
date_id integer,
card_type_id integer,
atm_status varchar(20),
currency varchar(10),
service varchar(20),
transaction_amount integer,
message_code varchar(255),
message_text varchar(255),
rain_3h NUMERIC(10,3),
clouds_all integer,
weather_id integer,
weather_main varchar(50),
weather_description varchar(255),
primary key(trans_id),
foreign key(atm_id) references etl.atm(atm_id),
foreign key(location_id) references etl.loc(location_id),
foreign key(date_id) references etl.date(date_id),
foreign key(card_type_id) references etl.card(card_type_id));
```

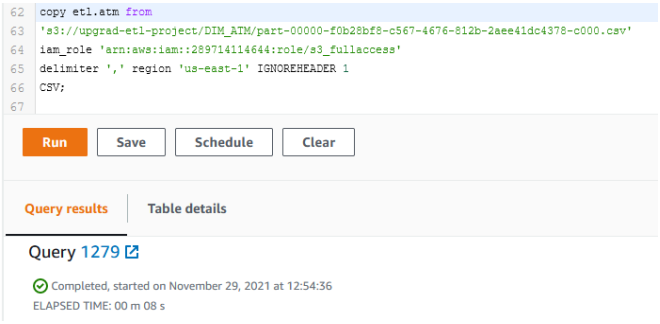


Loading data into a Redshift cluster from Amazon S3 bucket

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

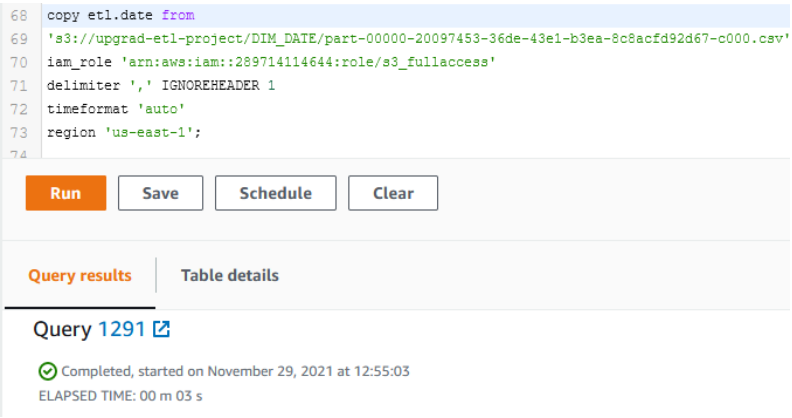
1. Query to load data into ATM dimension table.

```
copy etl.atm from
's3://upgrad-etl-project/DIM_ATM/part-00000-f0b28bf8-c567-4676-812b-2aee41dc4378-c000.csv'
iam_role 'arn:aws:iam::289714114644:role/s3_fullaccess'
delimiter ',' region 'us-east-1' IGNOREHEADER 1
CSV;
```



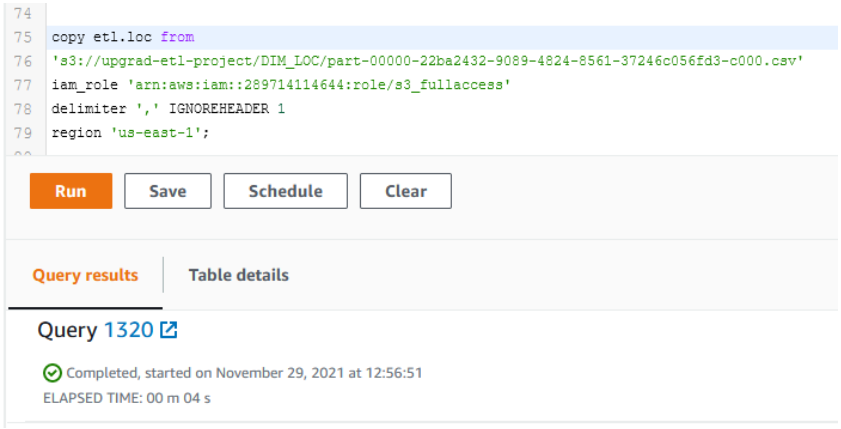
2. Query to load data into DATE dimension table.

```
copy etl.date from
's3://upgrad-etl-project/DIM_DATE/part-00000-20097453-36de-43e1-b3ea-8c8acfd92d67-c000.csv'
iam_role 'arn:aws:iam::289714114644:role/s3_fullaccess'
delimiter ',' IGNOREHEADER 1
timeformat 'auto'
region 'us-east-1';
```



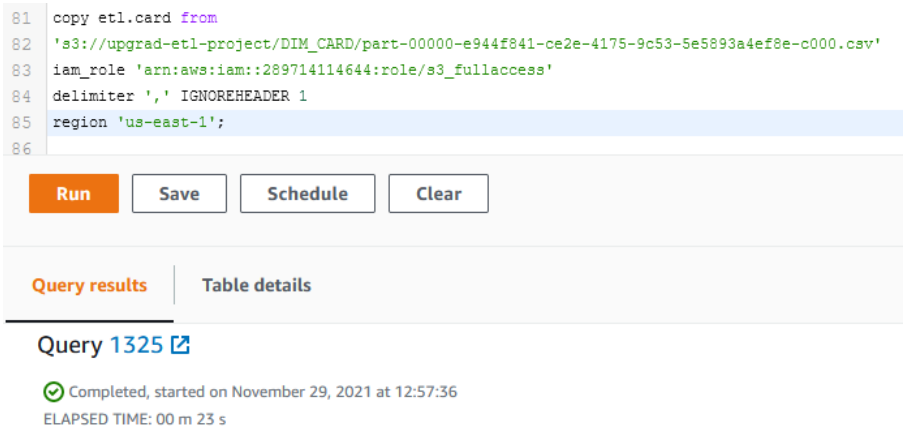
3. Query to load data into LOCATION dimension table.

```
copy etl.loc from
's3://upgrad-etl-project/DIM_LOC/part-00000-22ba2432-9089-4824-8561-37246c056fd3-c000.csv'
iam_role 'arn:aws:iam::289714114644:role/s3_fullaccess'
delimiter ',' IGNOREHEADER 1
region 'us-east-1';
```



4. Query to load data into CARD dimension table.

```
copy etl.card from
's3://upgrad-etl-project/DIM_CARD/part-00000-e944f841-ce2e-4175-9c53-5e5893a4ef8e-c000.csv'
iam_role 'arn:aws:iam::289714114644:role/s3_fullaccess'
delimiter ',' IGNOREHEADER 1
region 'us-east-1';
```



5. Query to load data into FACT_ATM_TRANS table.

```
copy etl.FACT_ATM_TRANS from
's3://upgrad-etl-project/FACT_ATM_TRANS/part-00000-2b0eccf5-78f5-4bed-9d6e-b2206b942fb2-c000.csv'
iam_role 'arn:aws:iam::289714114644:role/s3_fullaccess'
delimiter ',' IGNOREHEADER 1
region 'us-east-1'
TRUNCATECOLUMNS
CSV;
```