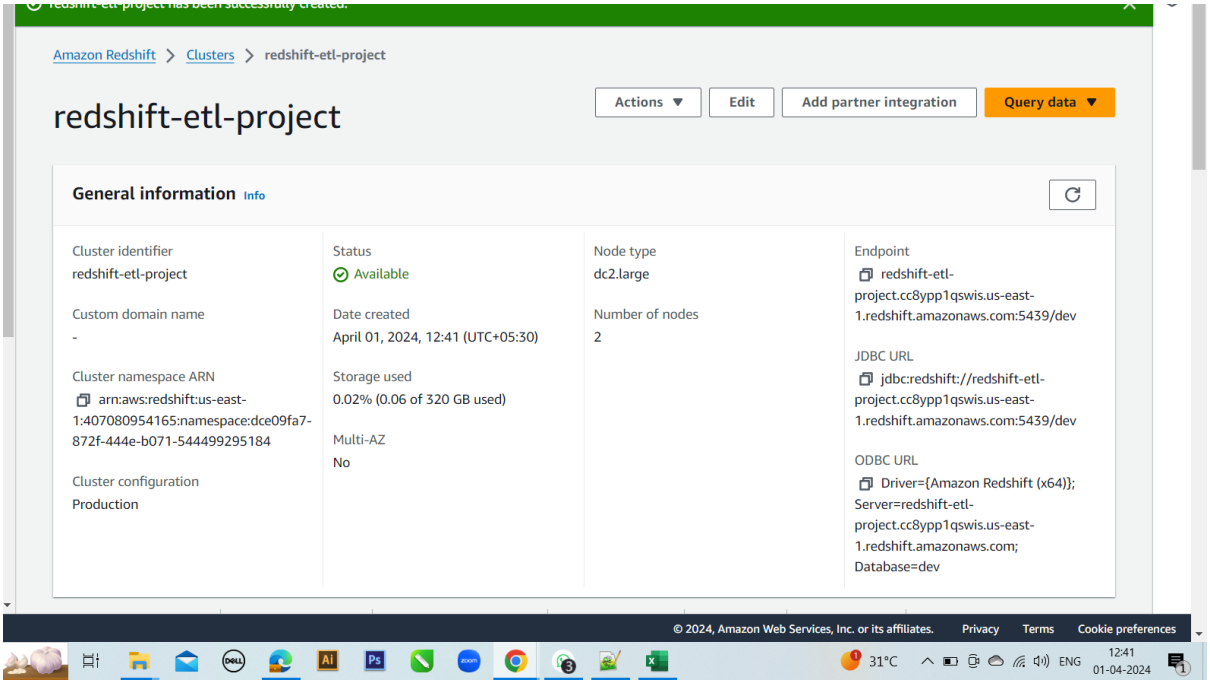
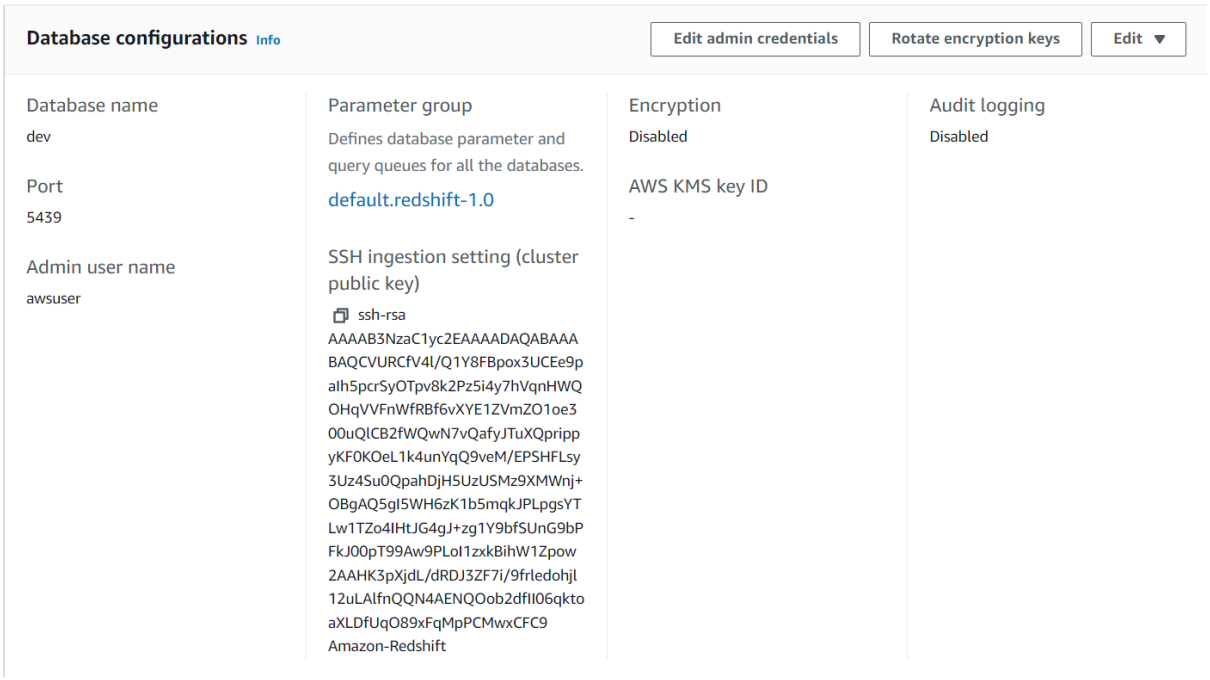
Creation of a Redshift Cluster

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

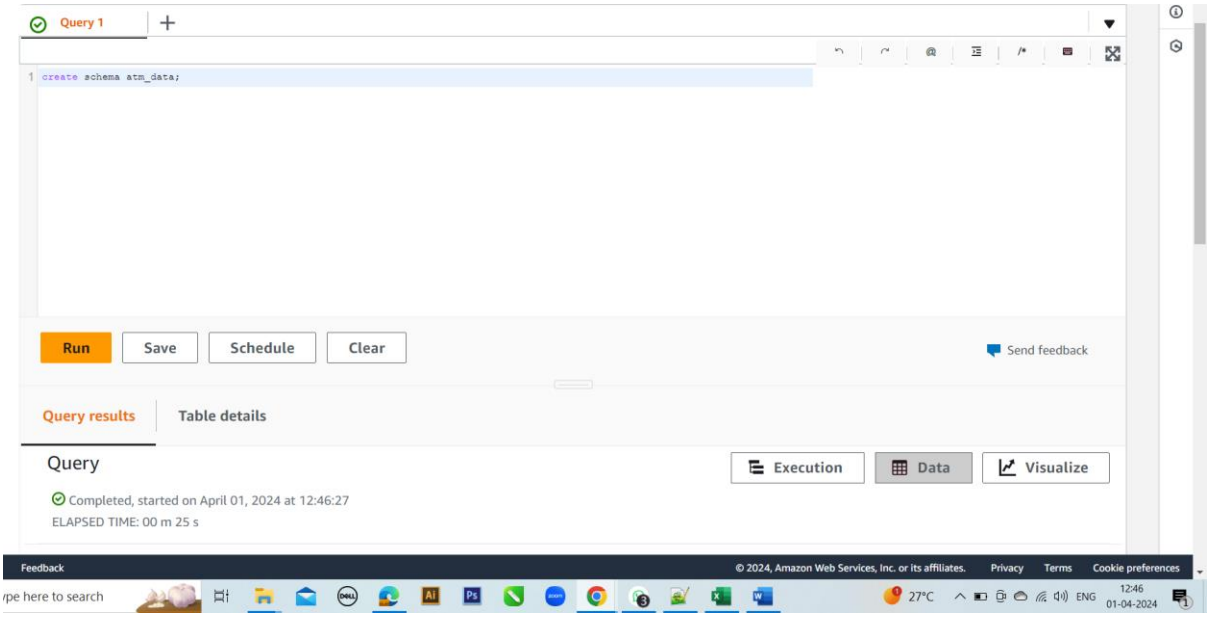




Setting up a database in the Redshift cluster and running queries to create the dimension and fact tables

Creation of schema atm_data on the Redshift cluster

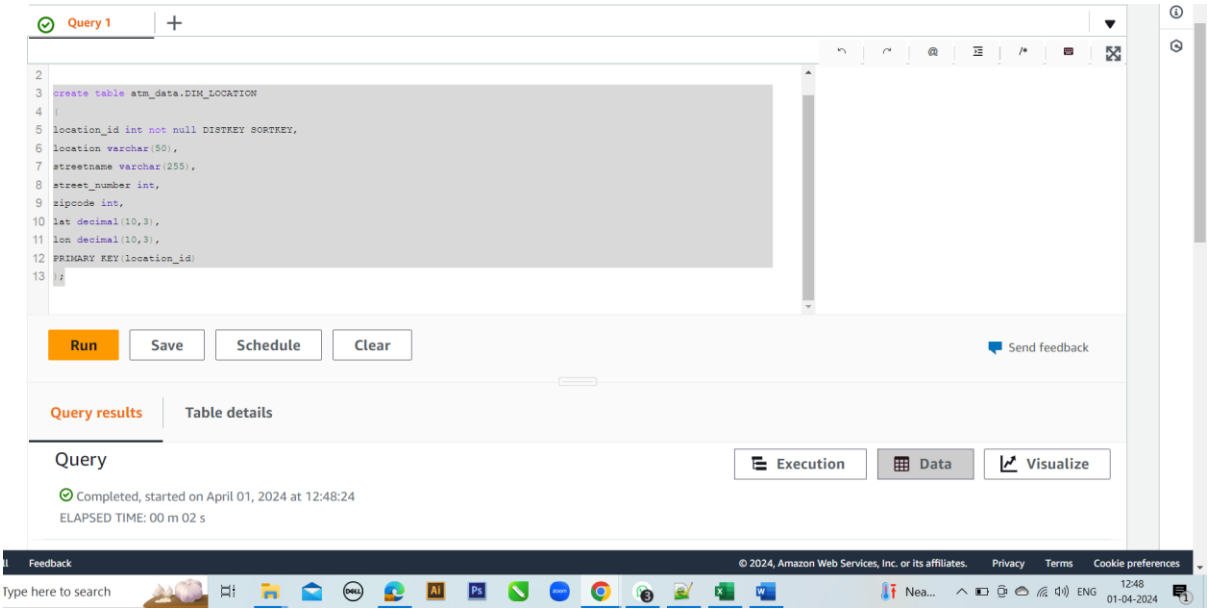
- **create schema atm_data;**



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

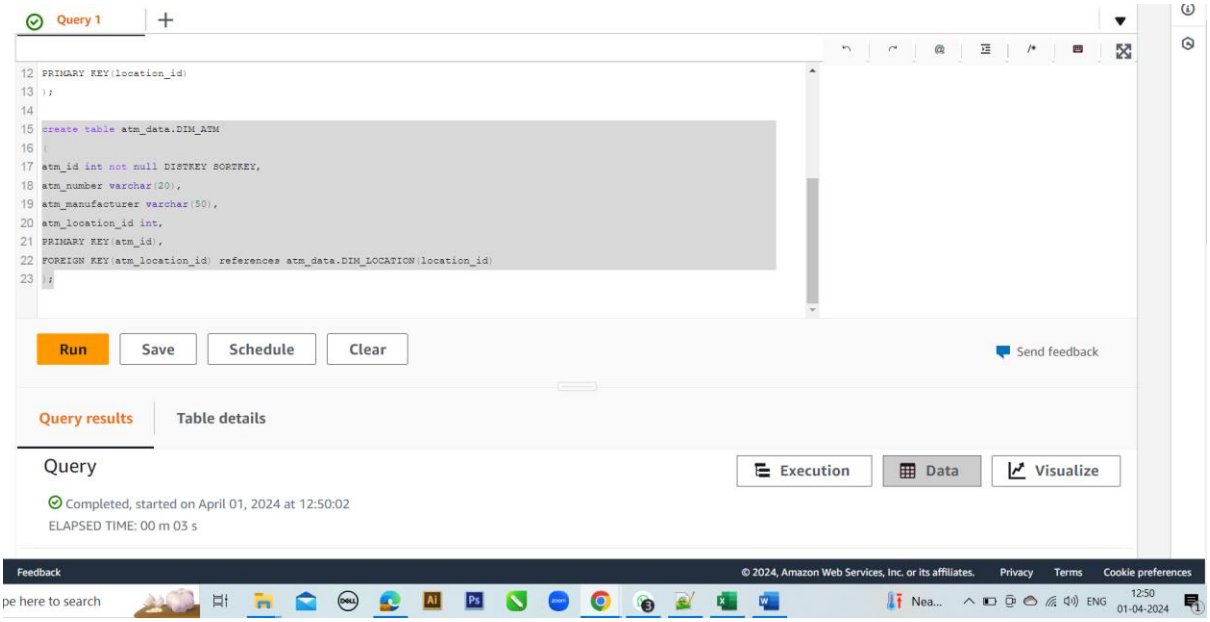
Creation of location dimension table DIM_LOCATION

- **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)
  );



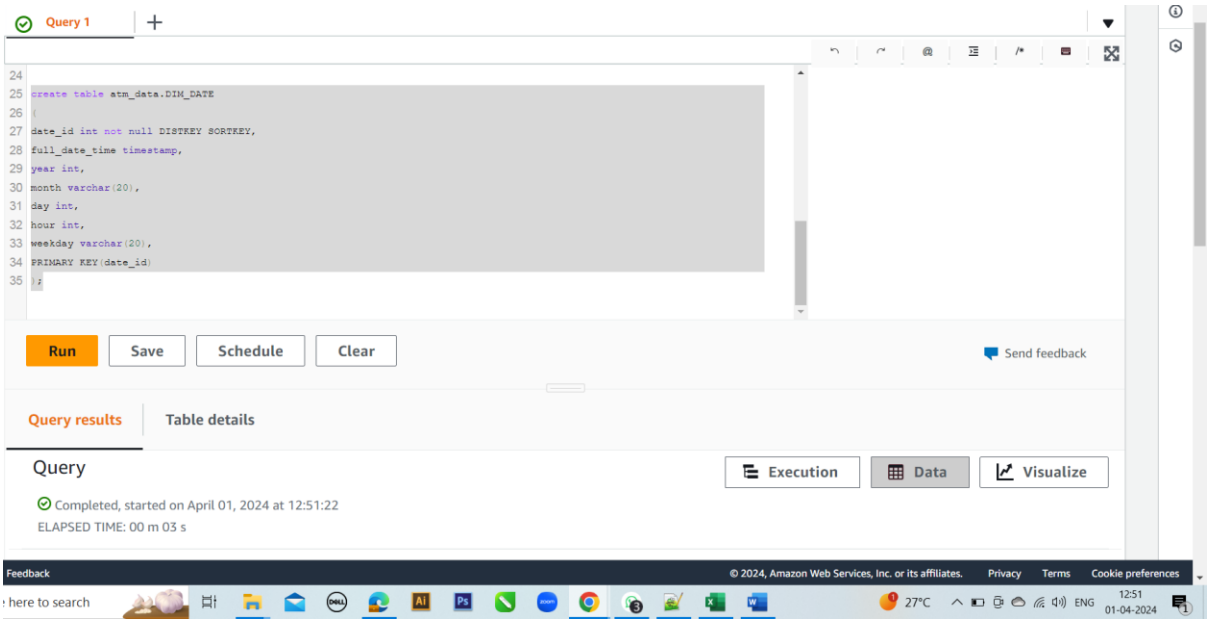
Creation of atm dimension table DIM_ATM

- **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)
  );



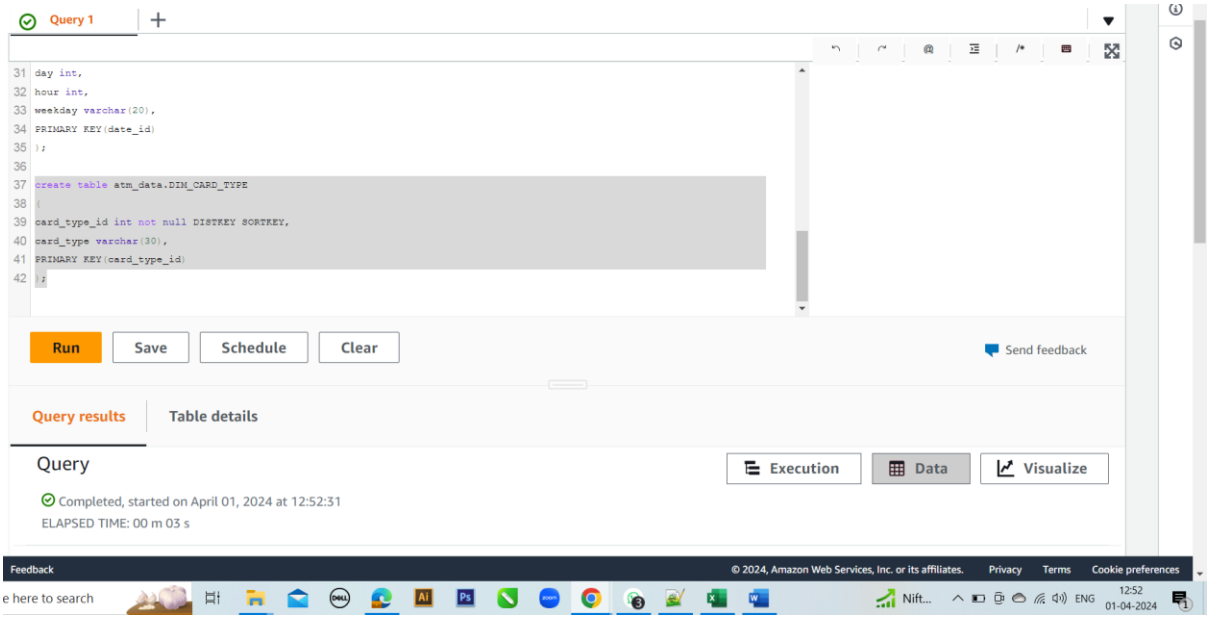
Creation of date dimension table DIM_DATE

- **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)
  );



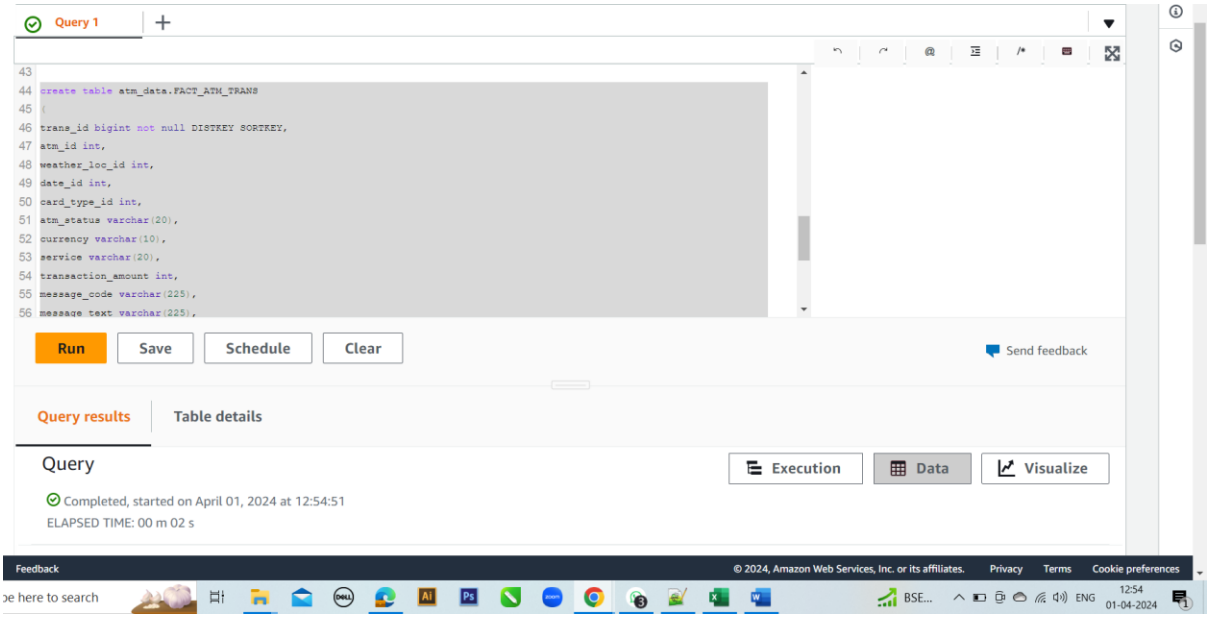
Creation of card type dimension table DIM_CARD_TYPE

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



Creation of atm transaction fact table FACT_ATM_TRANS

- **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

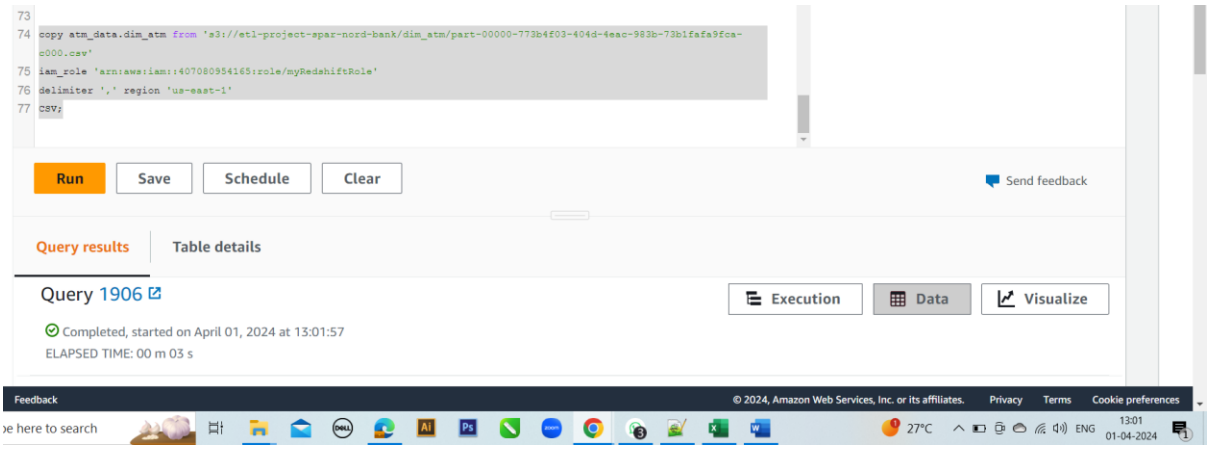
Queries to copy the data from S3 buckets to the Redshift cluster in the appropriate tables

Copy query for DIM_LOCATION table

- **copy atm_data.dim_location from 's3://etl-project-spar-nord-bank/dim_location/part-00000-298b1e66-7cc2-4b86-9225-ab52ce97a389-c000.csv'**
  iam_role 'arn:aws:iam::407080954165:role/myRedshiftRole'
  delimiter ',' region 'us-east-1'
  CSV;

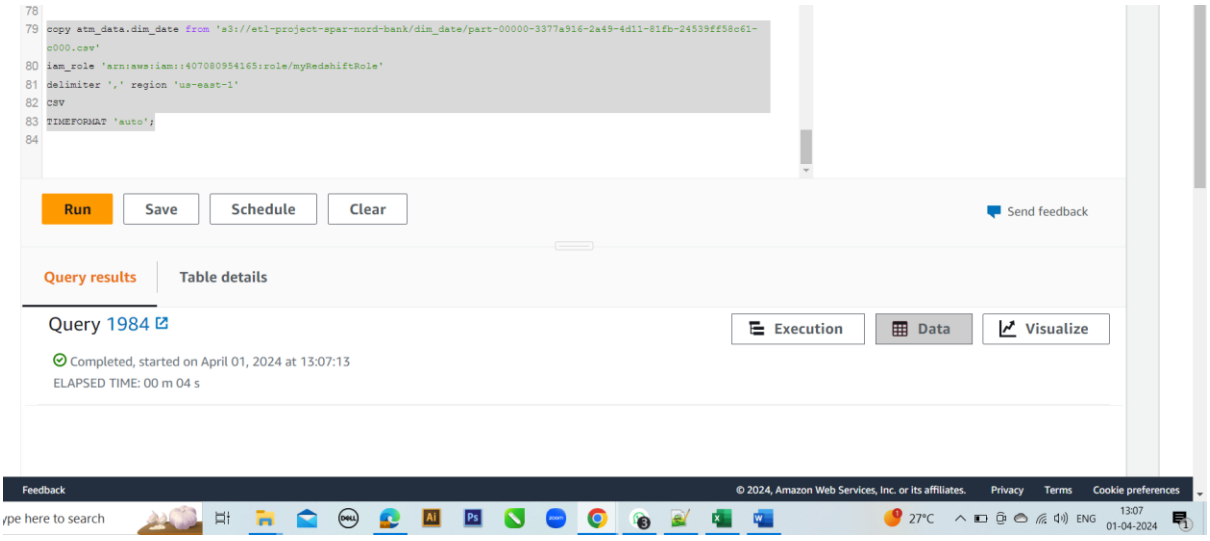
Copy query for DIM_ATM table

- **copy atm_data.dim_atm from 's3://etl-project-spar-nord-bank/dim_atm/part-00000-773b4f03-404d-4eac-983b-73b1fafa9fca-c000.csv'**
  iam_role 'arn:aws:iam::407080954165:role/myRedshiftRole'
  delimiter ',' region 'us-east-1'
  CSV;



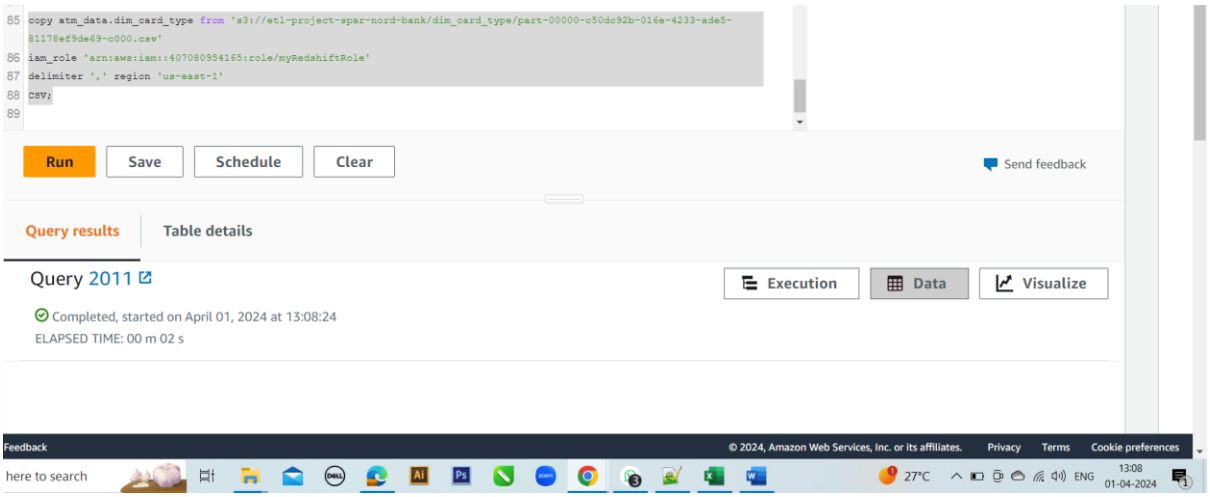
Copy query for DIM_DATE table

- **copy atm_data.dim_date from 's3://etl-project-spar-nord-bank/dim_date/part-00000-3377a916-2a49-4d11-81fb-24539ff58c61-c000.csv'**
  iam_role 'arn:aws:iam::407080954165:role/myRedshiftRole'
  delimiter ',' region 'us-east-1'
  CSV
  TIMEFORMAT 'auto';



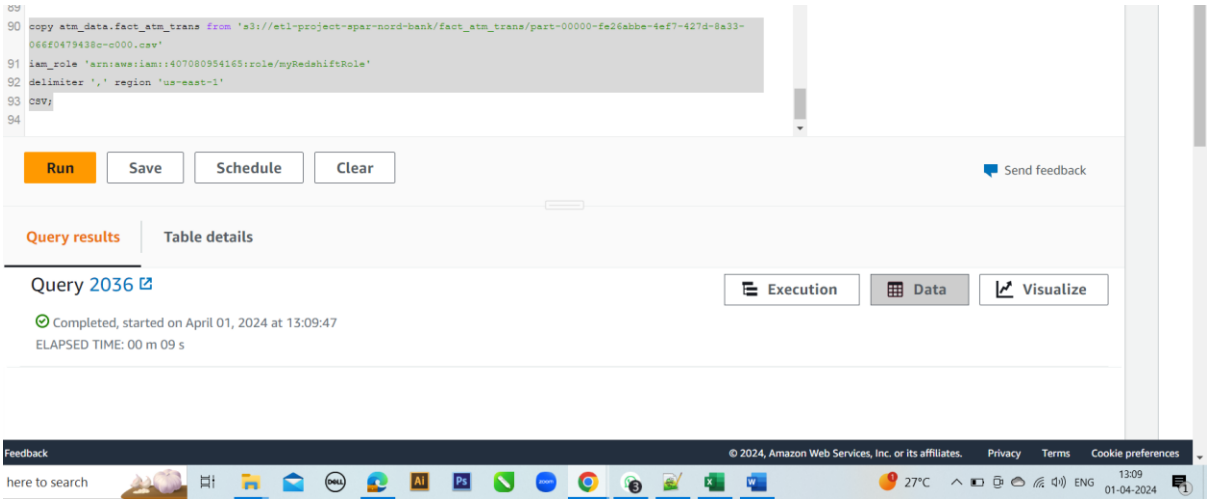
Copy query for DIM_CARD_TYPE table

- **copy atm_data.dim_card_type from 's3://etl-project-spar-nord-bank/dim_card_type/part-00000-c50dc92b-016e-4233-ade5-81178ef9de69-c000.csv'**
  iam_role 'arn:aws:iam::407080954165:role/myRedshiftRole'
  delimiter ',' region 'us-east-1'
  CSV;

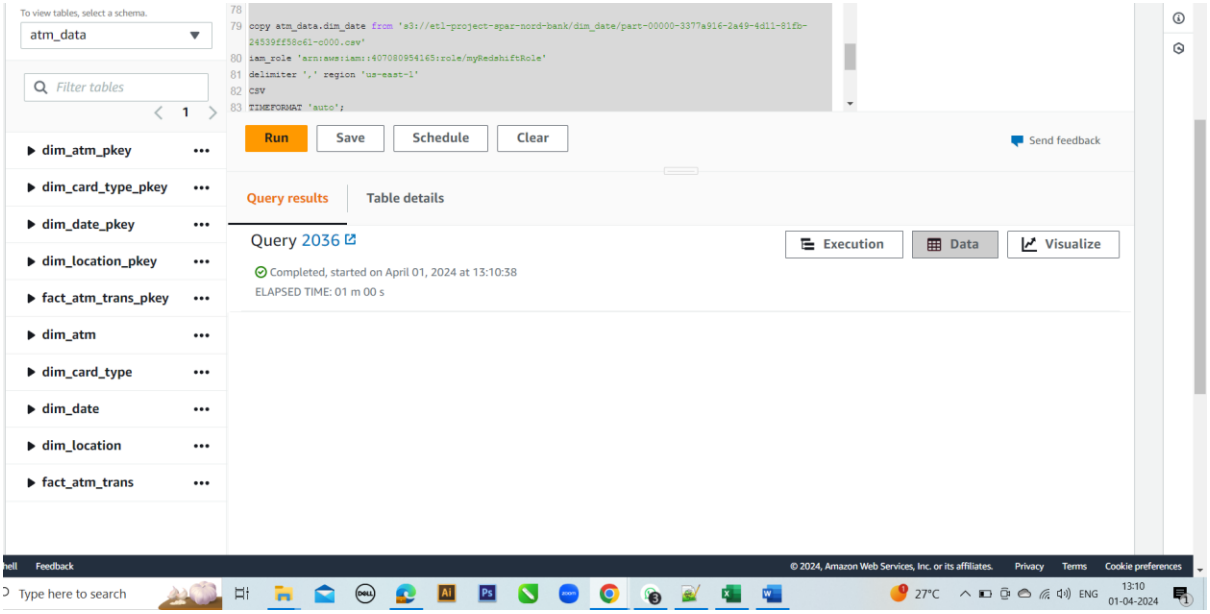


Copy query for FACT_ATM_TRANS table

- **copy atm_data.fact_atm_trans from 's3://etl-project-spar-nord-bank/fact_atm_trans/part-00000-fe26abbe-4ef7-427d-8a33-066f0479438c-c000.csv'**
  iam_role 'arn:aws:iam::407080954165:role/myRedshiftRole'
  delimiter ',' region 'us-east-1'
  CSV;



Final screenshot of tables in the schema



Count verification in the tables

- **Count of FACT_ATM_TRANS**

```
16 select count(*) from atm_data.FACT_ATM_TRANS;
```

Run | Save | Schedule | Clear

Query results | Table details

Query 2128

Completed, started on April 01, 2024 at 13:13:56

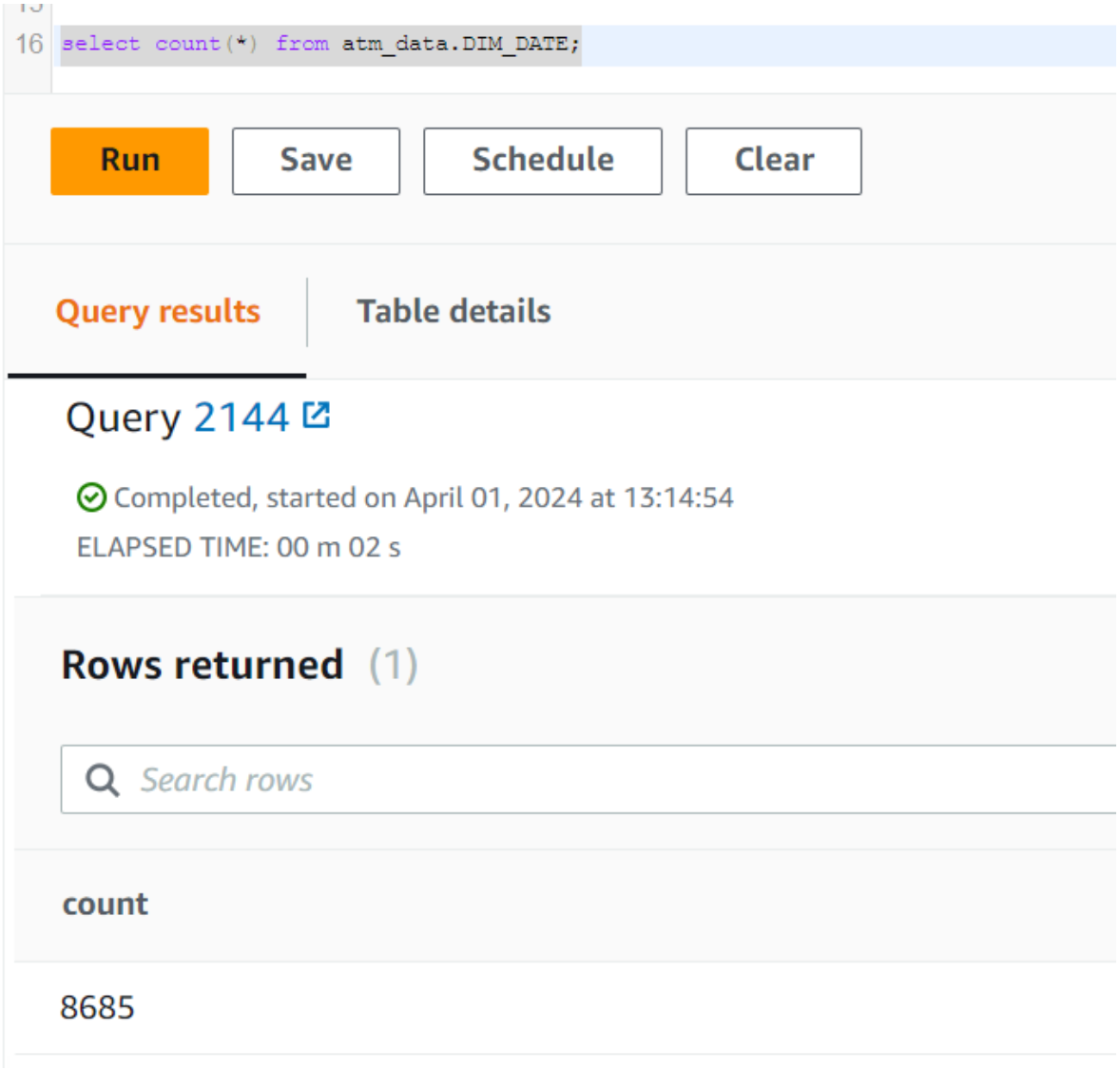
ELAPSED TIME: 00 m 02 s

Rows returned (1)

Search rows

count
2468572

- **Count of DIM_DATE**



- **Count of DIM_CARD_TYPE**

```
15
16 select count(*) from atm_data.DIM_CARD_TYPE;
```

Run | Save | Schedule | Clear

Query results | Table details

Query 2158

Completed, started on April 01, 2024 at 13:15:32

ELAPSED TIME: 00 m 02 s

Rows returned (1)

Search rows

count
12

- **Count of DIM_ATM**

```
15
16 select count(*) from atm_data.DIM_ATM;
```

Run | Save | Schedule | Clear

Query results | Table details

Query 2172

Completed, started on April 01, 2024 at 13:16:20

ELAPSED TIME: 00 m 03 s

Rows returned (1)

Search rows

count
113

- **Count of DIM_LOCATION**

```
16 select count(*) from atm_data.DIM_LOCATION;
```

Run | Save | Schedule | Clear

Query results | Table details

Query 2182

Completed, started on April 01, 2024 at 13:16:48

ELAPSED TIME: 00 m 03 s

Rows returned (1)

Search rows

count
109

******** Thank you *********