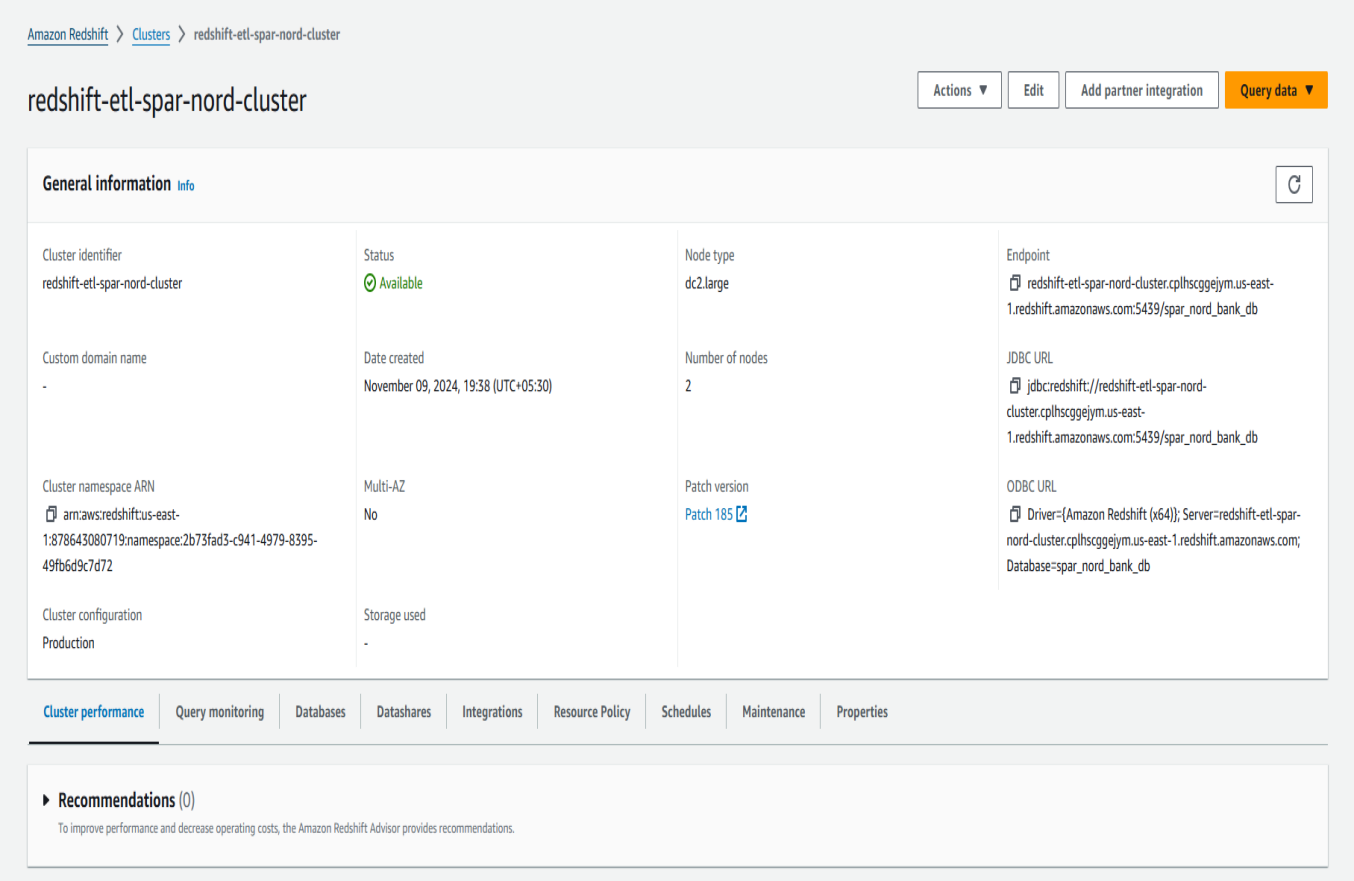
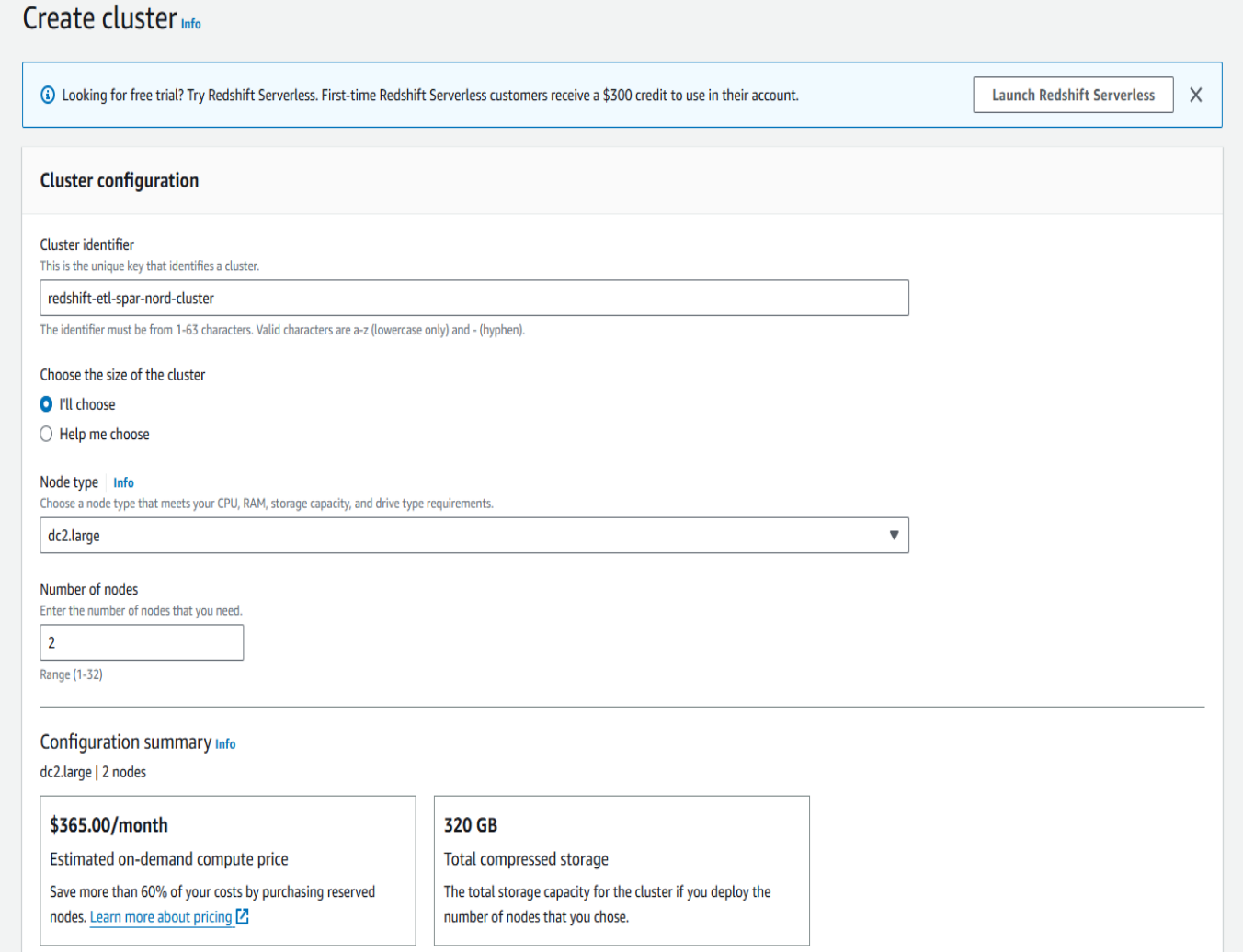
# Creation of a Redshift Cluster

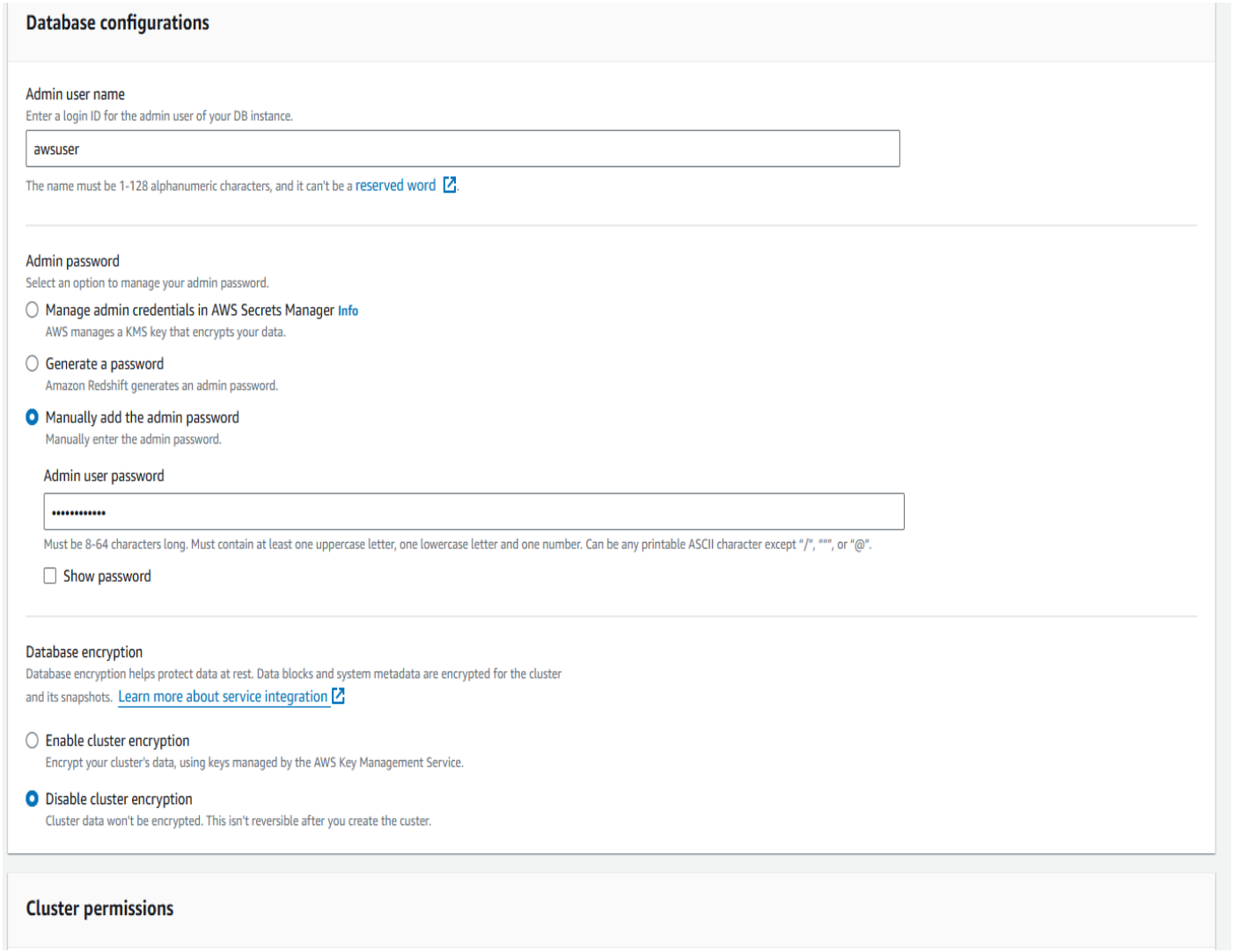
**Screenshots of the configuration of the Redshift cluster that I have created: redshift-etl-spar-nord-cluster**



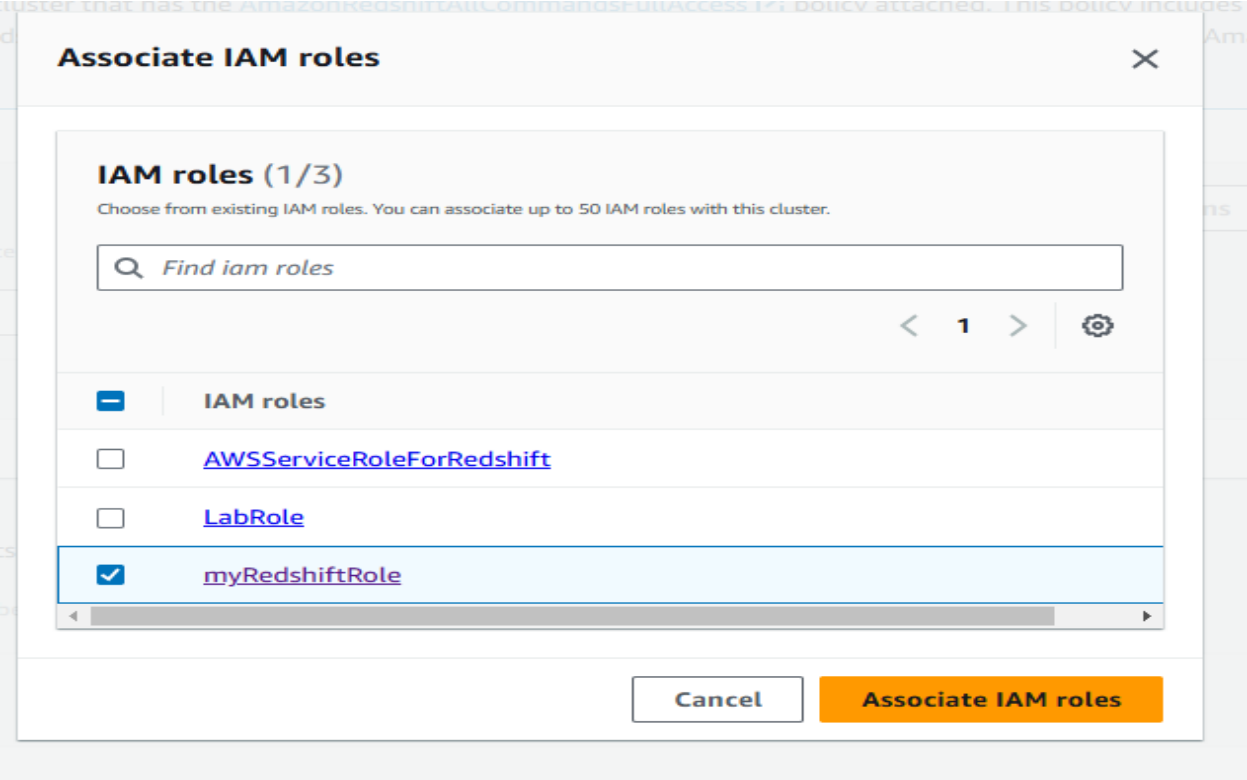
## Screenshots of configurations while creating cluster

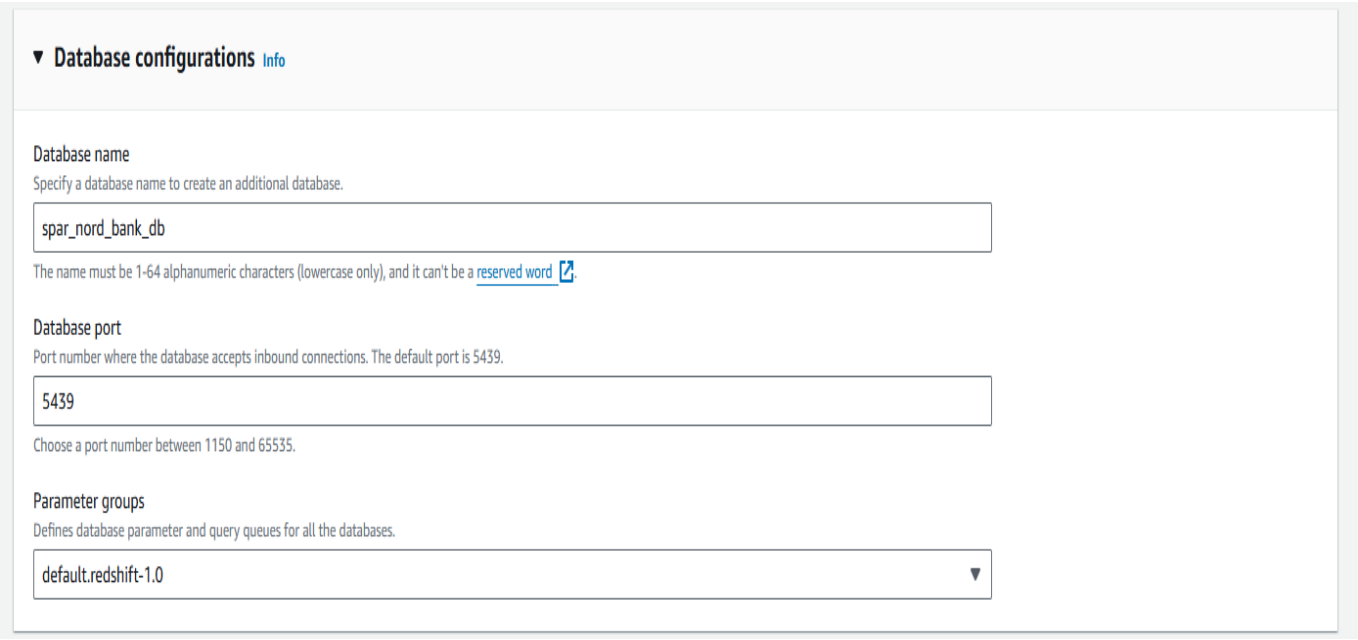


**Screenshot of database configurations:**



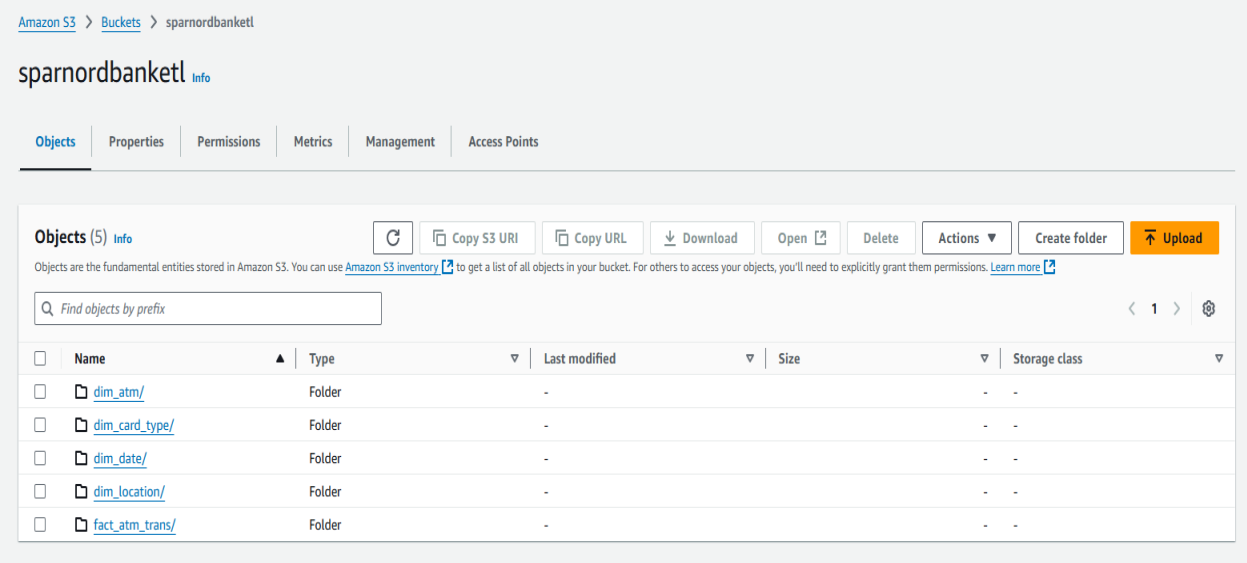
**Associating IAM role (myRedshiftRole) to my cluster and creating spar_nord_bank_db:**



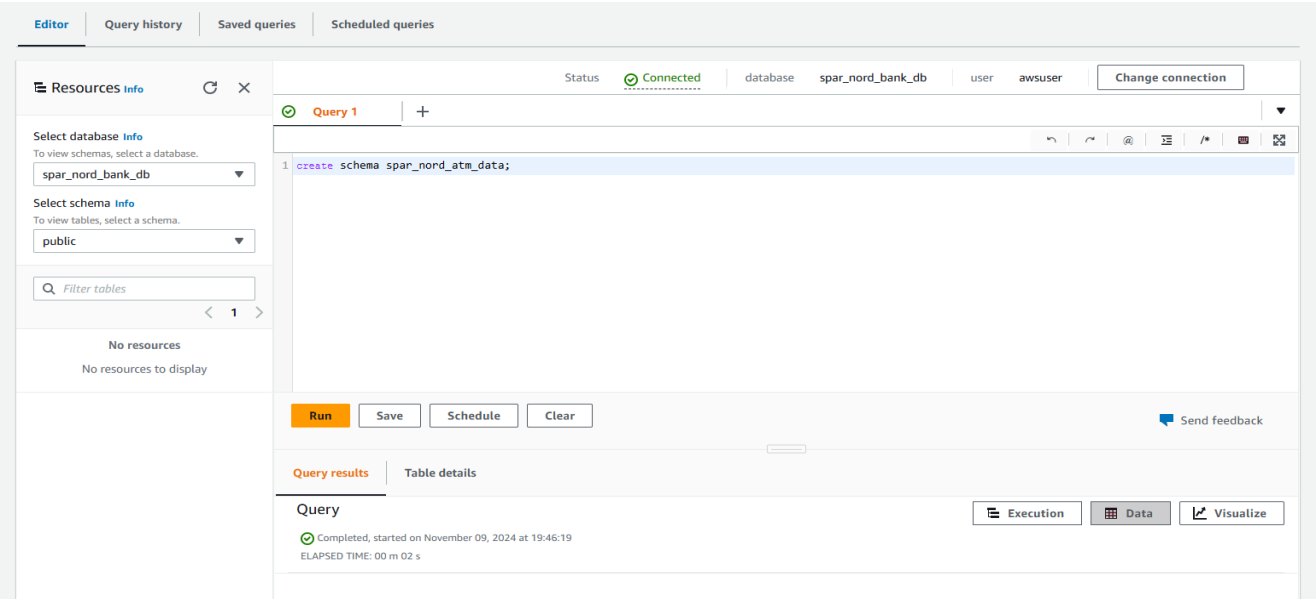
# Setting up a database in the Redshift cluster and running queries to create the dimension and fact tables

## Making Sure the files are present in Amazon S3 bucket: sparnordbanketl



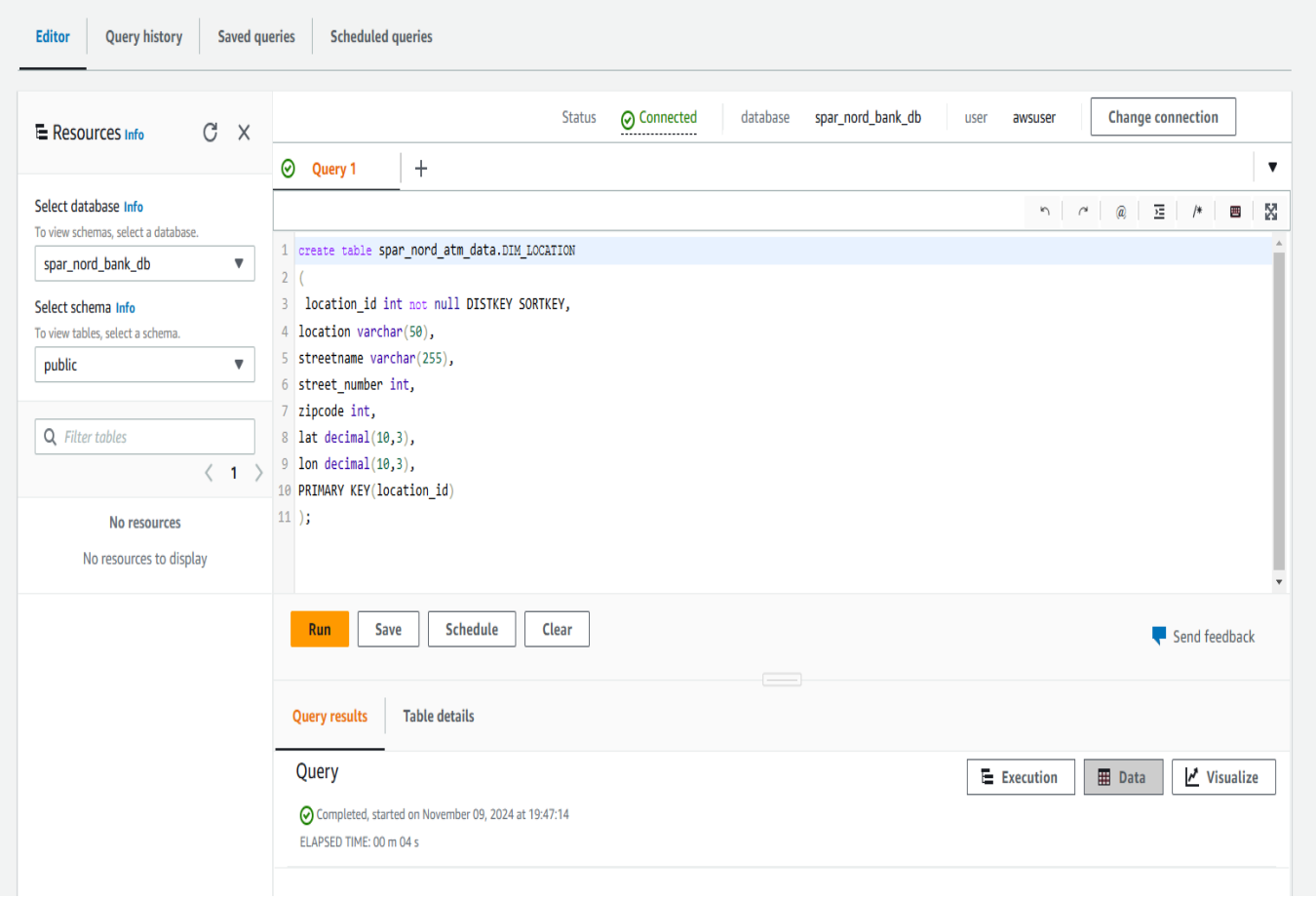

## Query to create Schema for dimension and fact tables:

## Create schema spar_nord_atm_data;



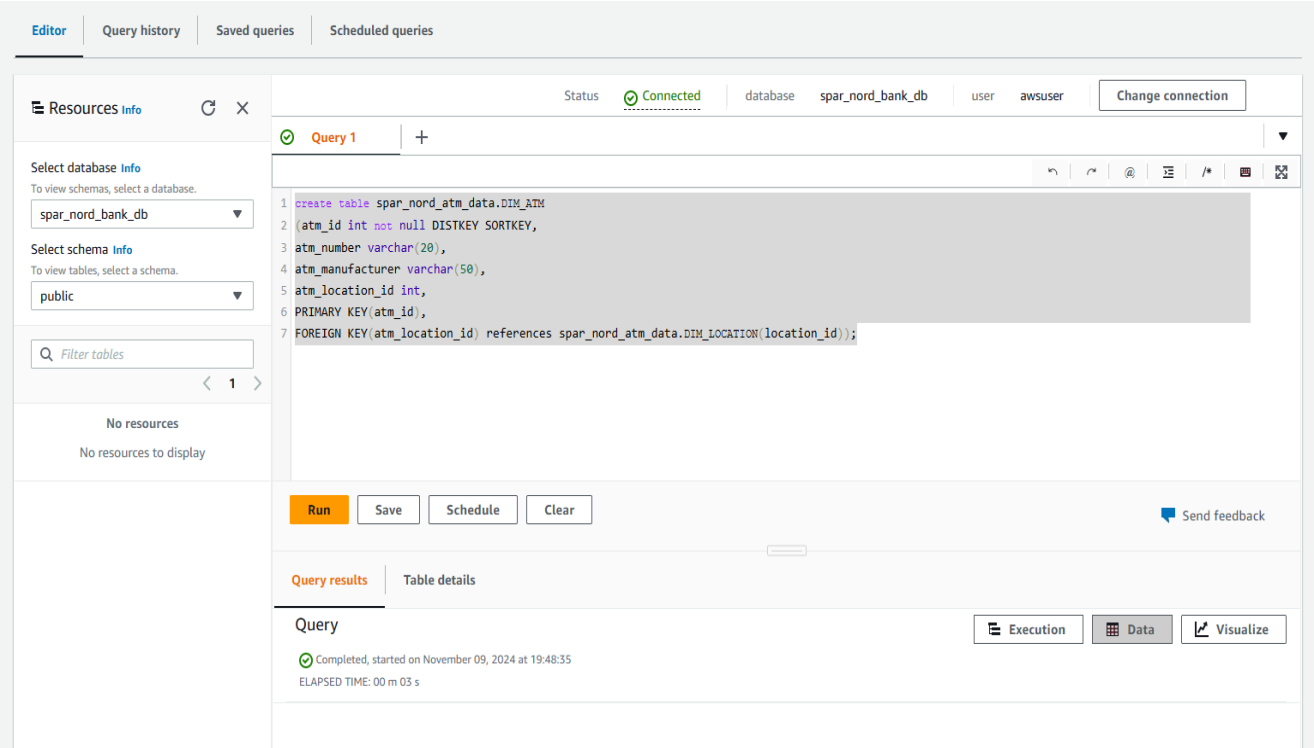
**Query to create DIM_LOACTION dimension table:**

```
create table spar_nord_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));
```



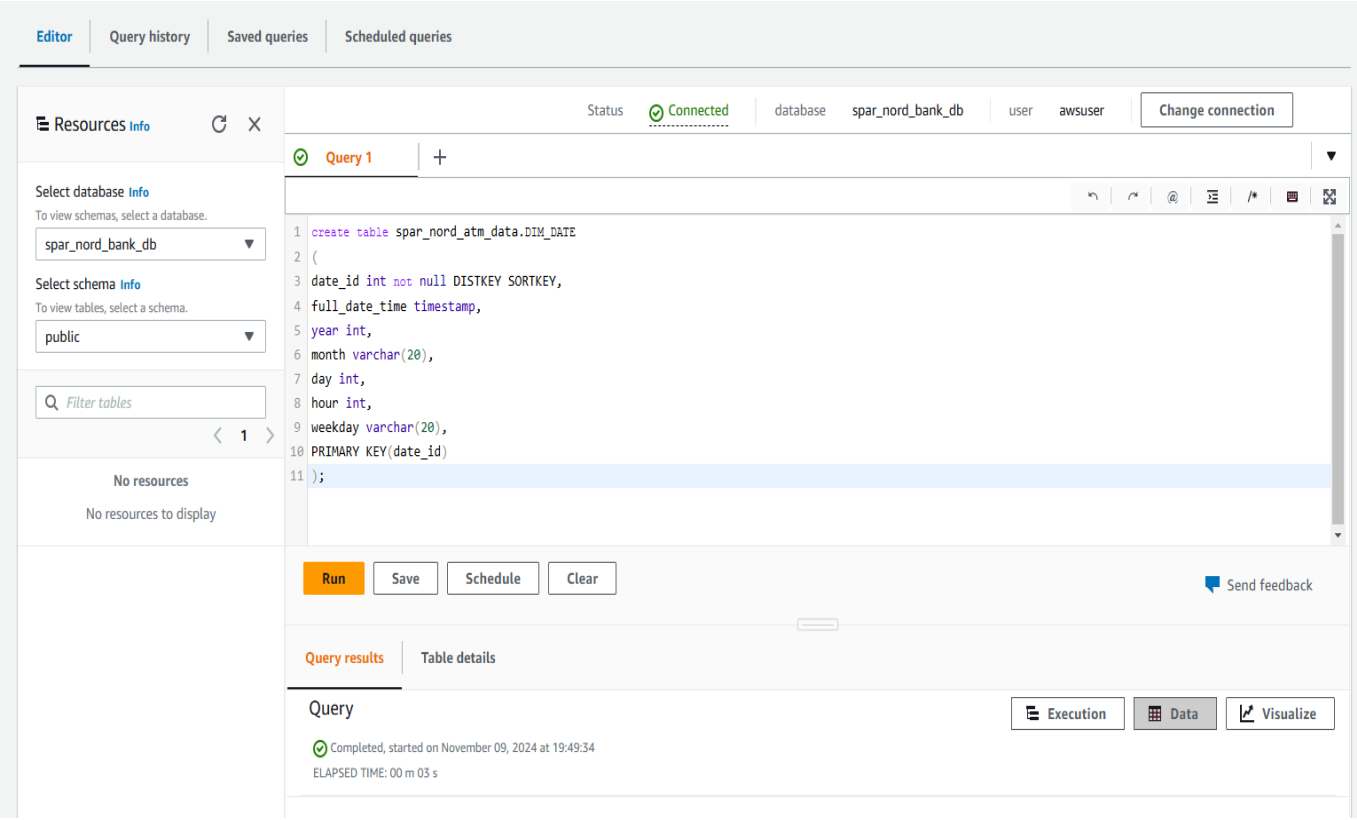
**Query to create DIM_ATM dimension table:**

```
create table spar_nord_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
spar_nord_atm_data.DIM_LOCATION(location_id));
```



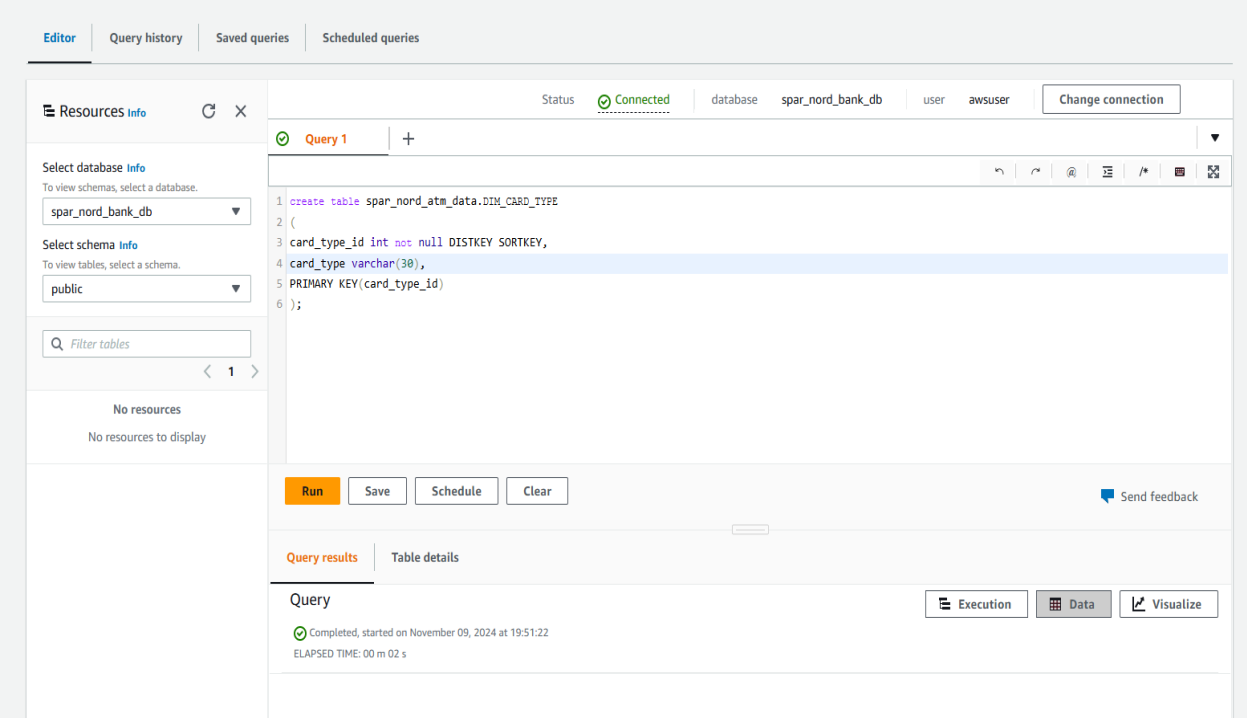
**Query to create DIM_DATE dimension table:**

```
create table spar_nord_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));
```



**Query to create DIM_CARD_TYPE dimension table:**

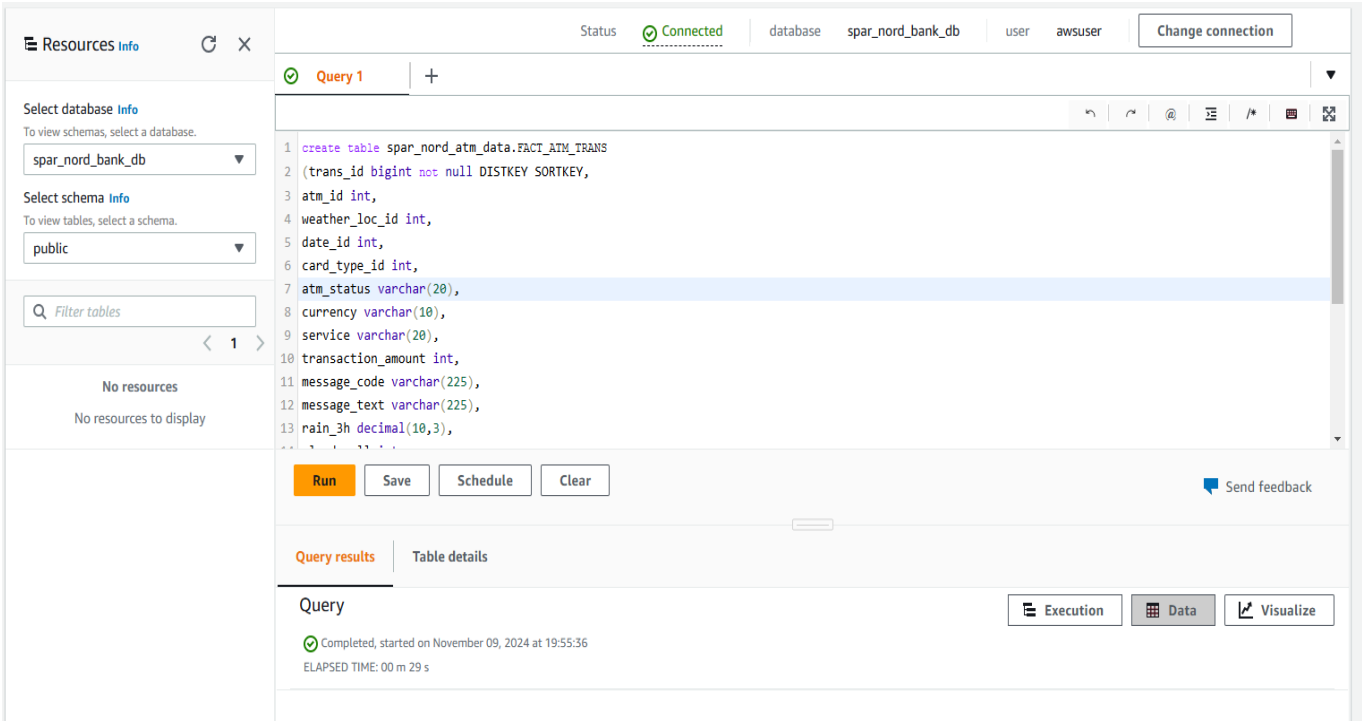
```
create table spar_nord_atm_data.DIM_CARD_TYPE
(card_type_id int not null DISTKEY SORTKEY,
card_type varchar(30),
PRIMARY KEY(card_type_id));
```




**Query to create FACT_ATM_TRANS Fact table:**

```
create table spar_nord_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
spar_nord_atm_data.DIM_LOCATION(location_id),
FOREIGN KEY(atm_id) references spar_nord_atm_data.DIM_ATM(atm_id),
FOREIGN KEY(date_id) references spar_nord_atm_data.DIM_DATE(date_id),
FOREIGN KEY(card_type_id) references
spar_nord_atm_data.DIM_CARD_TYPE(card_type_id));
```

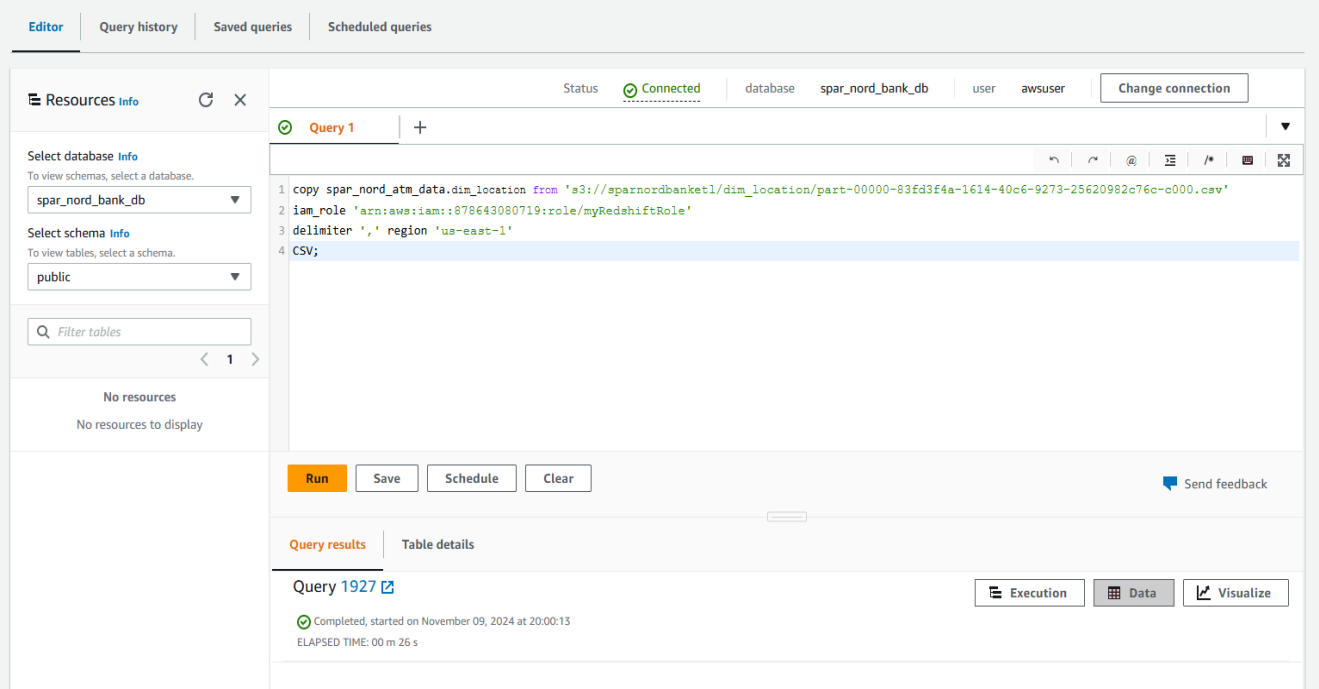


# Loading data into a Redshift cluster from Amazon S3 bucket

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

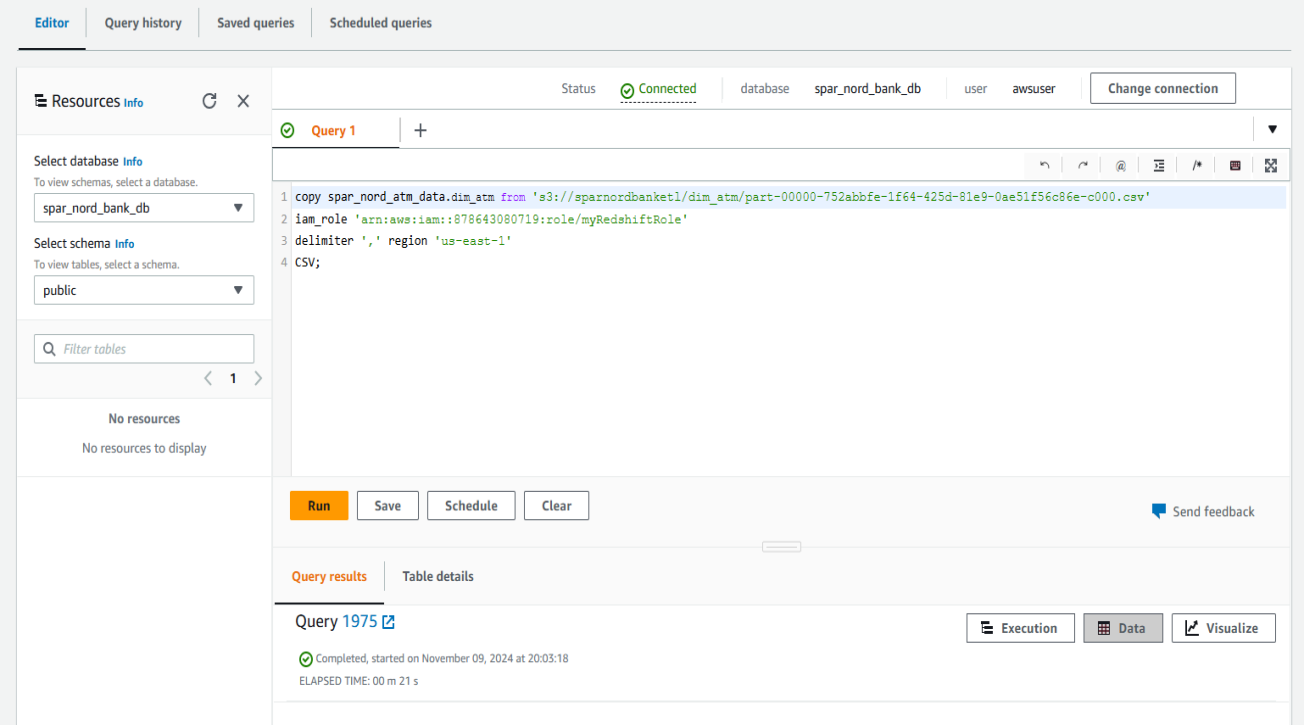
**Copying data to dim_location table from S3 bucket:**

```
copy spar_nord_atm_data.dim_location from 's3://sparnordbanketl/dim_location/part-
00000-83fd3f4a-1614-40c6-9273-25620982c76c-c000.csv'
iam_role 'arn:aws:iam::878643080719:role/myRedshiftRole'
delimiter ',' region 'us-east-1'
CSV;
```



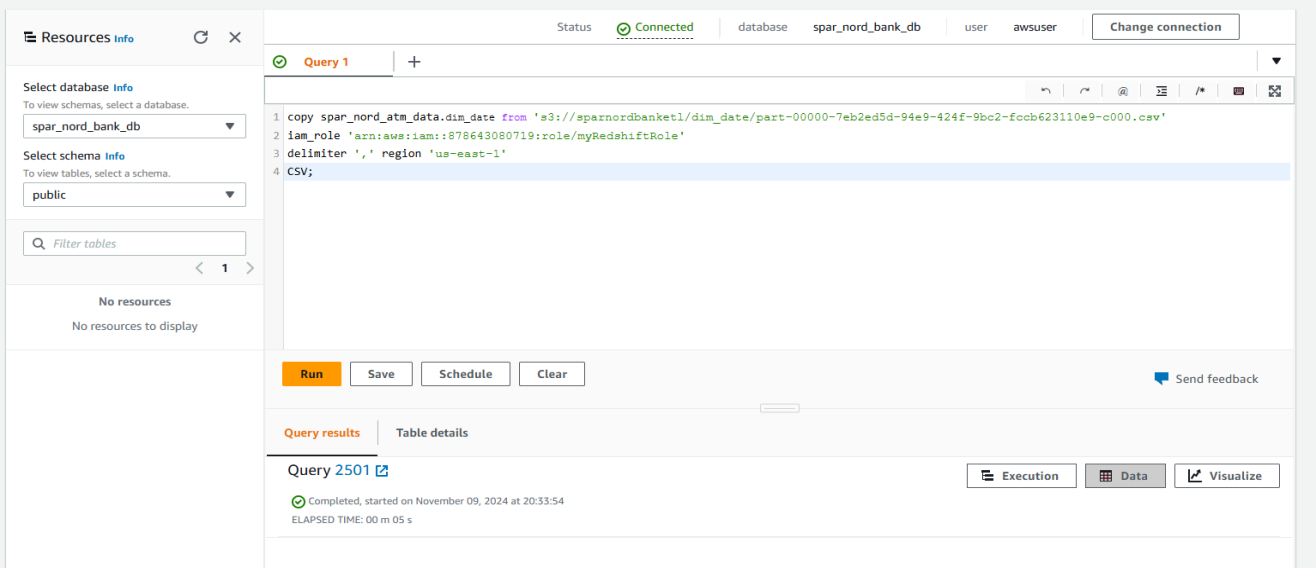

**Copying data to dim_atm table from S3 bucket:**

```
copy spar_nord_atm_data.dim_atm from 's3://sparnordbanketl/dim_atm/part-00000-
752abbfe-1f64-425d-81e9-0ae51f56c86e-c000.csv'
iam_role 'arn:aws:iam::878643080719:role/myRedshiftRole'
delimiter ',' region 'us-east-1'
CSV;
```



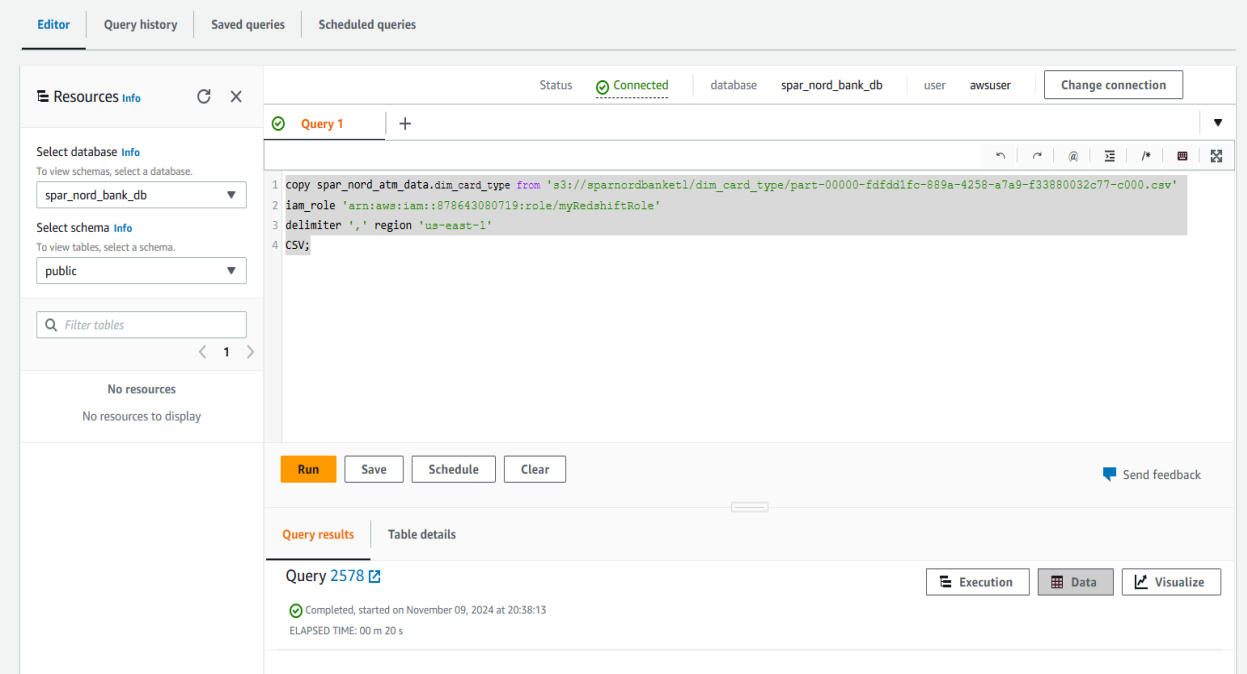

**Copying data to dim_date table from S3 bucket:**

**copy spar_nord_atm_data.dim_date from 's3://sparnordbanketl/dim_date/part-00000-7eb2ed5d-94e9-424f-9bc2-fccb623110e9-c000.csv'**
**iam_role 'arn:aws:iam::878643080719:role/myRedshiftRole'**
**delimiter ',' region 'us-east-1'**
**CSV;**



**Copying data to dim_card_type table from S3 bucket:**

```
copy spar_nord_atm_data.dim_card_type from 's3://sparnordbanketl/dim_card_type/part-00000-fdfdd1fc-889a-4258-a7a9-f33880032c77-c000.csv'
iam_role 'arn:aws:iam::878643080719:role/myRedshiftRole'
delimiter ',' region 'us-east-1'
CSV;
```




**Copying data to fact_atm_trans table from S3 bucket:**

```
copy spar_nord_atm_data.fact_atm_trans from
's3://sparnordbanketl/fact_atm_trans/part-00000-f3a78bbe-4298-409f-bb9f-b6ed344ffb4f-c000.csv'
iam_role 'arn:aws:iam::878643080719:role/myRedshiftRole'
delimiter ',' region 'us-east-1'
CSV;
```