

Configuration of Redshift Cluster:

aws

Services

Search

[Alt+S]

N. Virginia

voclabs/user3142453=srutayanee@gmail.com @ 8517-2536-9539

Amazon Redshift > Clusters > sparnod-redshift-cluster-1

sparnod-redshift-cluster-1

Actions Edit Add partner integration Query data

General information

Cluster identifier sparnod-redshift-cluster-1	Status Available	Node type dc2.large	Endpoint sparnod-redshift-cluster-1.cy8prp7imilf.us-east-1.redshift.amazonaws.com:5439/dev
Custom domain name -	Date created June 04, 2024, 11:38 (UTC+05:30)	Number of nodes 2	JDBC URL jdbc:redshift://sparnod-redshift-cluster-1.cy8prp7imilf.us-east-1.redshift.amazonaws.com:5439/dev
Cluster namespace ARN arn:aws:redshift:us-east-1:851725369539:namespace:f927ea6a-7e24-4659-9fbe-3b3eb18f4b7b	Storage used 0.12% (0.39 of 320 GB used)	Patch version Patch 180	ODBC URL Driver={Amazon Redshift (x64)}; Server=sparnod-redshift-cluster-1.cy8prp7imilf.us-east-1.redshift.amazonaws.com; Database=dev
Cluster configuration Production	Multi-AZ No		

aws

Services

Search

[Alt+S]

N. Virginia

voclabs/user3142453=srutayanee@gmail.com @ 8517-2536-9539

Database configurations

Edit admin credentials Rotate encryption keys Edit

Database name dev	Parameter group default.redshift-1.0	Encryption Disabled	Audit logging Disabled
Port 5439	SSH ingestion setting (cluster public key) ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAQCAQRL BxDimwIYaCZBQJCXEoJ6ziX0YuN1B+a0ThgRft PalzjA4xn8mMFrCtScyEISNld6n6PZQuJyHY/L MqsegoWbug4rhVrUOTuYE1EjAZi6mzgdQIR TFwooznfwr05YMnt6U7MaUJGzh9KyIwdqKJ3 qolS+N9GzpV2zrsO9PvPHqbf5Uiivw0OcojGN QkSRL6+DHAo+d3GVNf/yamWvOSmWQ3H55 1EMR3YhkgDsoDialxgBjNgsO0a8fasTllwG3JIT wKHIXCOOLXPKqazIT0/5Aj3PtbGAeqLeCJ27 D3/9Sbuo5wAZIFh8XXFNyBXAUj48+rUhx5g4S T/hfN Amazon-Redshift	AWS KMS key ID -	
Admin user name awsuser			

aws

Services

Search

[Alt+S]

N. Virginia

voclabs/user3142453=srutayanee@gmail.com @ 8517-2536-9539

Network and security settings

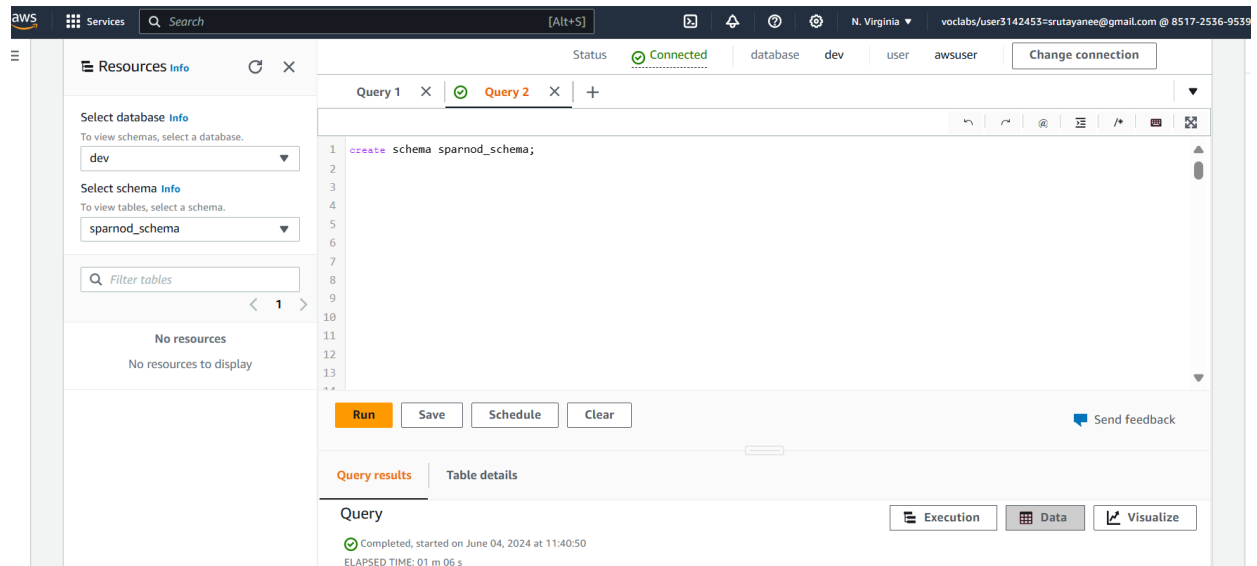
Edit

Virtual private cloud (VPC) vpc-05089ac4f2cc9fbc	Availability Zone us-east-1a	VPC security group Specify which instances and devices can connect to the cluster. sg-0649a9bd2a1fbbfd8	Publicly accessible Allow connections from outside the VPC. Disabled
Subnet group cluster-subnet-group-1	Enhanced VPC routing Disabled	IP address type -	
Endpoint URL -			

Queries used for creating the Schema, Dimension and Fact tables on the Redshift cluster:

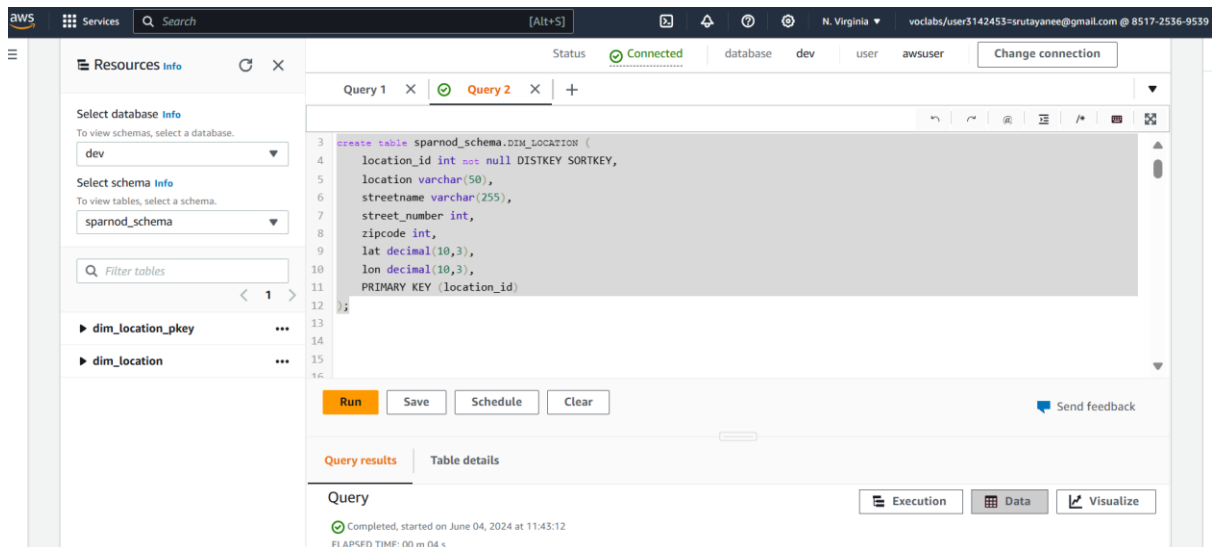
1. First, we will create the schema and for that we used the below command:

```
create schema sparnod_schema;
```



2. Next, we will create the dimension table for location and for that we used the below command:

```
create table sparnod_schema.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)  
);
```

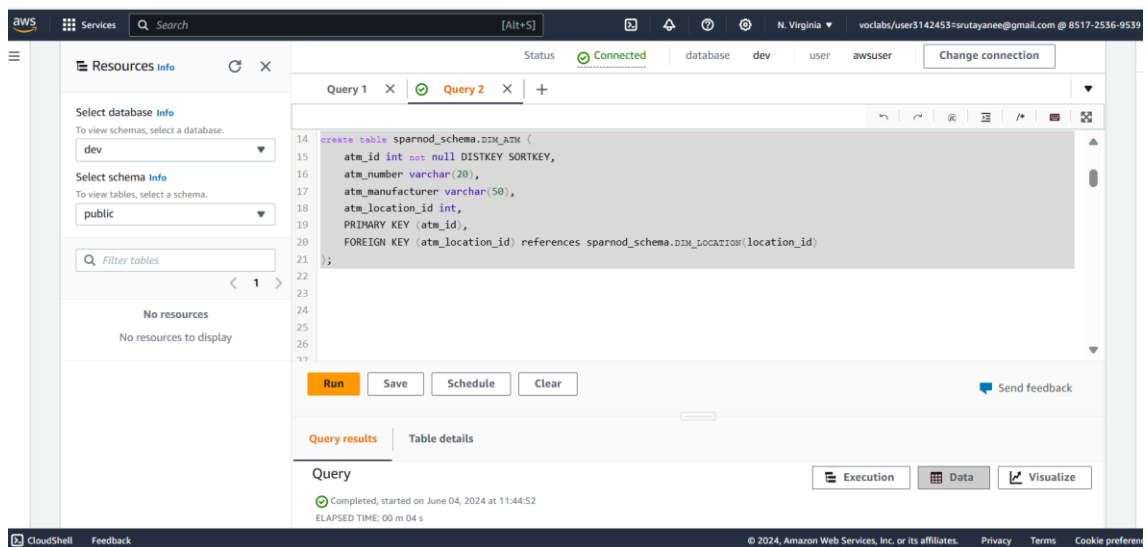


- Next, we will create the dimension table for ATM and for that we used the below command:

```

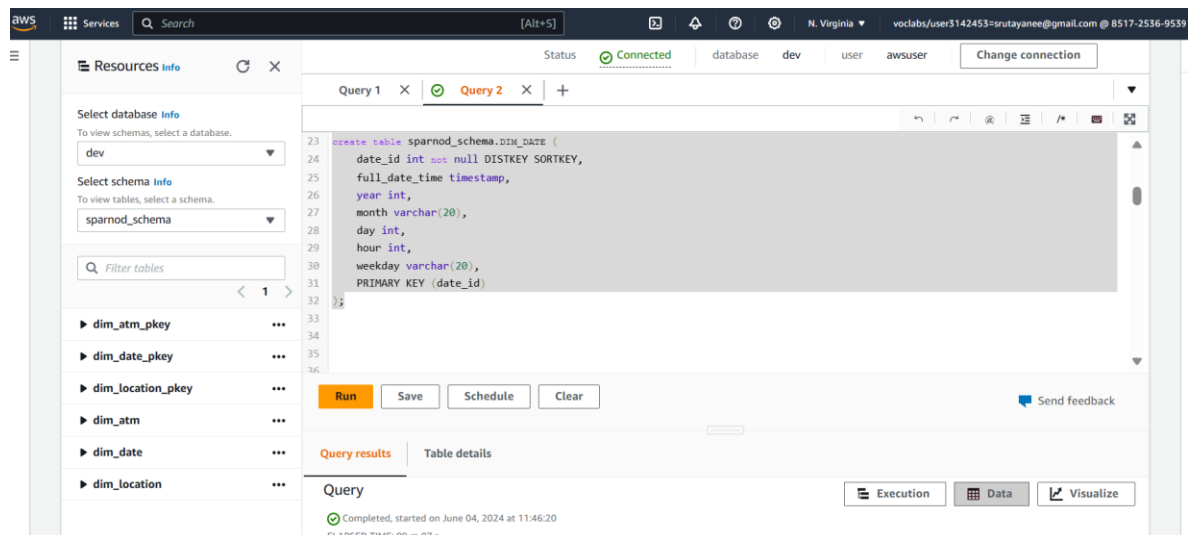
create table sparnod_schema.DIM_ATM (
  atm_id int not null DISTKEY SORTKEY,
  atm_number varchar (20),
  atm_manufacturer varchar (50),
  atm_location_id int,
  PRIMARY KEY (atm_id),
  FOREIGN KEY (atm_location_id) references sparnod_schema.DIM_LOCATION
(location_id)
);

```



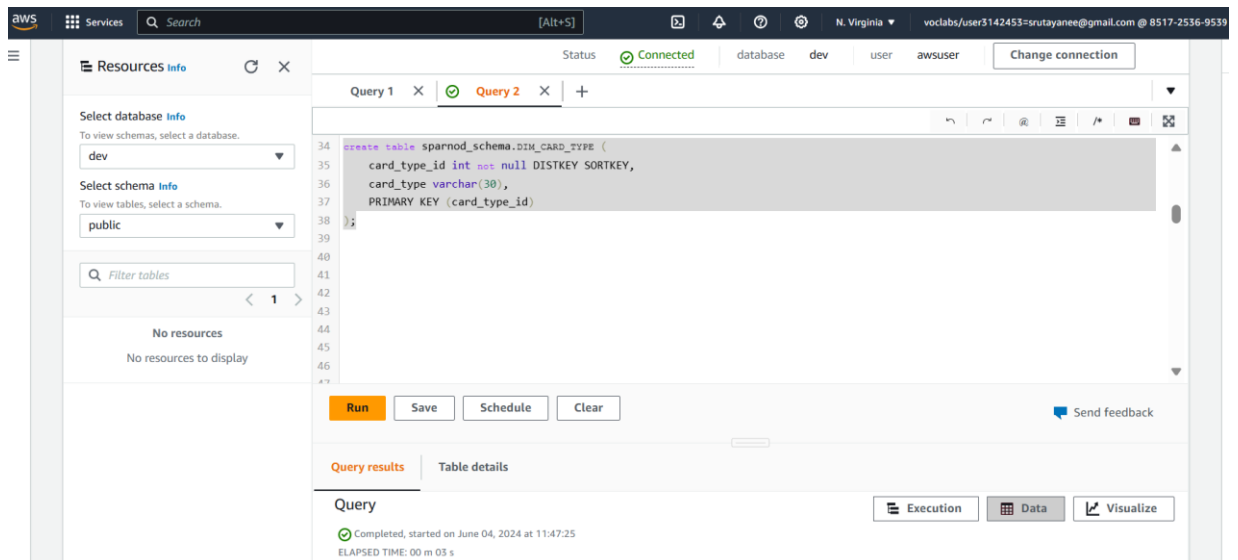
4. Next, we will create the dimension table for Date and for that we used the below command:

```
create table sparnod_schema.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)  
);
```



5. Next, we will create the dimension table for card type and for that we used the below command:

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



- Finally, we will create the fact table called **FACT_ATM_TRANS** and for that we used the below command:

```

create table sparnod_schema.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 (255),
    message_text varchar (255),
    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 sparnod_schema.DIM_LOCATION
(location_id),

FOREIGN KEY (atm_id) references sparnod_schema.DIM_ATM (atm_id),

FOREIGN KEY (date_id) references sparnod_schema.DIM_DATE (date_id)

FOREIGN KEY (card_type_id) references sparnod_schema.DIM_CARD_TYPE
(card_type_id)

);

The screenshot shows the AWS Redshift console interface. On the left, the 'Resources' panel is open, showing the 'dev' database and 'sparnod_schema' schema. A list of tables is displayed, including 'dim_atm_pkey', 'dim_card_type_pkey', 'dim_date_pkey', 'dim_location_pkey', 'fact_atm_trans_pkey', 'dim_atm', 'dim_card_type', and 'dim_date'. The main panel shows a SQL query for creating a fact table 'sparnod_schema.FACT_ATM_TRANS' with various columns and foreign key constraints. The query is as follows:

```
40 create table sparnod_schema.FACT_ATM_TRANS (  
41   trans_id bigint not null DISTKEY SORTKEY,  
42   atm_id int,  
43   weather_loc_id int,  
44   date_id int,  
45   card_type_id int,  
46   atm_status varchar(20),  
47   currency varchar(10),  
48   service varchar(20),  
49   transaction_amount int,  
50   message_code varchar(255),  
51   message_text varchar(255),  
52   rain_3h decimal(10,3),  
53   clouds_all int,  
54   weather_id int,  
55   weather_main varchar(50),  
56   weather_description varchar(255),  
57   PRIMARY KEY (trans_id),  
58   FOREIGN KEY (weather_loc_id) references sparnod_schema.DIM_LOCATION(location_id),  
59   FOREIGN KEY (atm_id) references sparnod_schema.DIM_ATM(atm_id),  
60   FOREIGN KEY (date_id) references sparnod_schema.DIM_DATE(date_id),  
61   FOREIGN KEY (card_type_id) references sparnod_schema.DIM_CARD_TYPE(card_type_id)  
62 );
```

Below the query editor, the 'Query results' tab is selected, showing the status 'Completed, started on June 04, 2024 at 11:52:01' and 'ELAPSED TIME: 00 m 14 s'.

Queries used for loading the data into the Dimension and Fact tables in the Redshift cluster from the S3 bucket:

1. Let's, copy the data from S3 bucket specific to location into DIM_LOCATION table in Redshift cluster and below is the command we used:

copy sparnod_schema.DIM_LOCATION

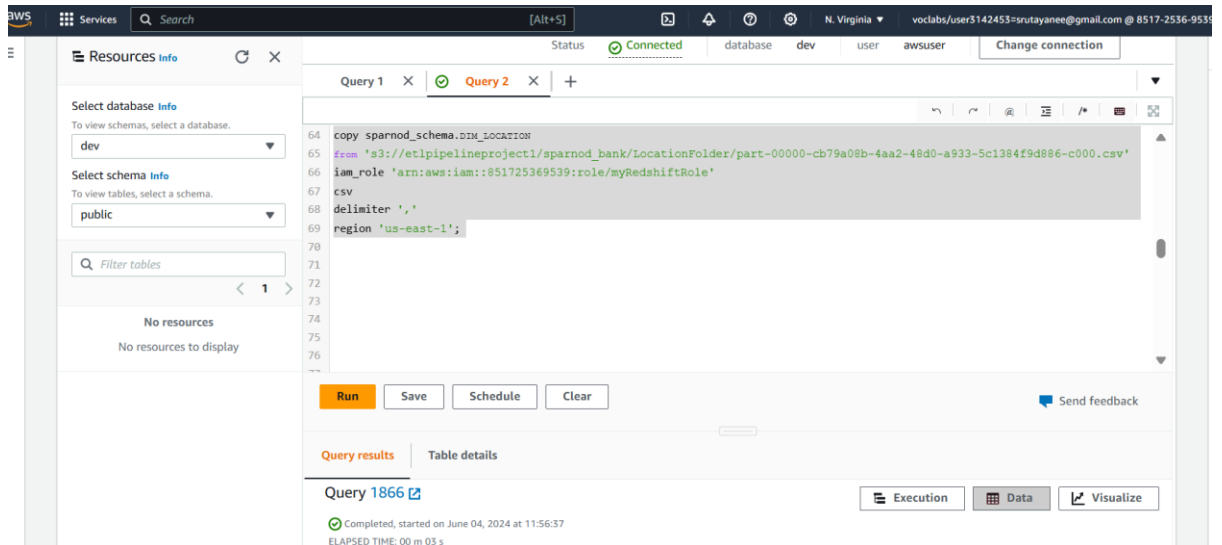
from 's3://etlpipelineproject1/sparnod_bank/LocationFolder/part-00000-cb79a08b-4aa2-48d0-a933-5c1384f9d886-c000.csv'

iam_role 'arn:aws:iam::851725369539:role/myRedshiftRole'

csv

delimiter ','

region 'us-east-1';



- Let's, copy the data from S3 bucket specific to location into DIM_ATM table in Redshift cluster and below is the command we used:

copy sparnod_schema.DIM_ATM

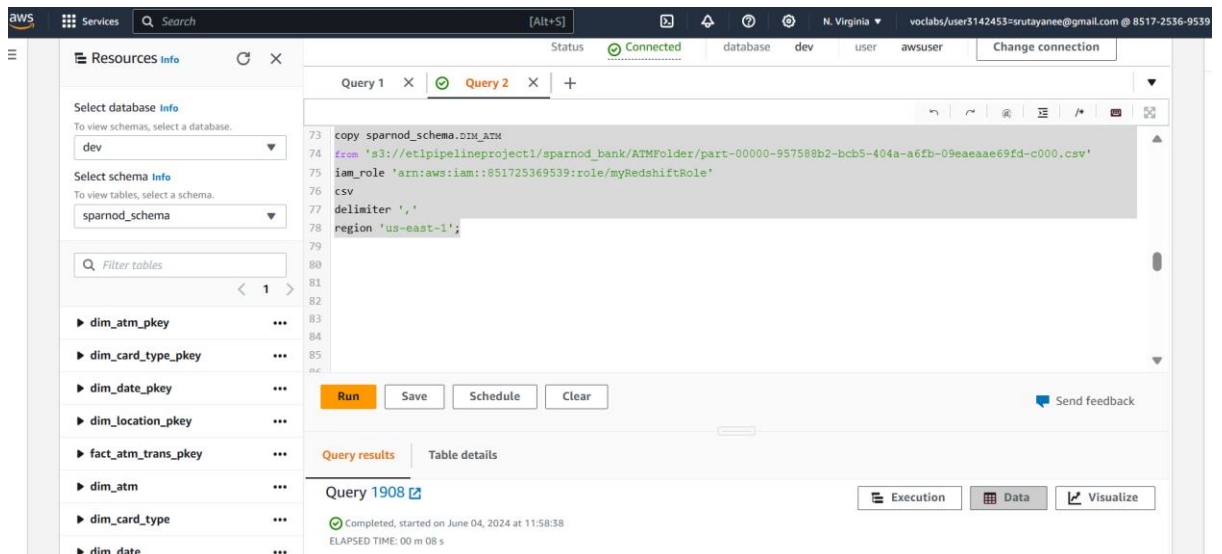
from 's3://etlpipelineproject1/sparnod_bank/ATMFolder/part-00000-957588b2-bcb5-404a-a6fb-09eaeaae69fd-c000.csv'

iam_role 'arn:aws:iam::851725369539:role/myRedshiftRole'

csv

delimiter ','

region 'us-east-1';



3. Let's, copy the data from S3 bucket specific to location into DIM_DATE table in Redshift cluster and below is the command we used:

```
copy sparnod_schema.DIM_DATE
```

```
from 's3://etlpipelineproject1/sparnod_bank/DateFolder/part-00000-82f65a72-62fb-4561-8208-5a28203aa310-c000.csv'
```

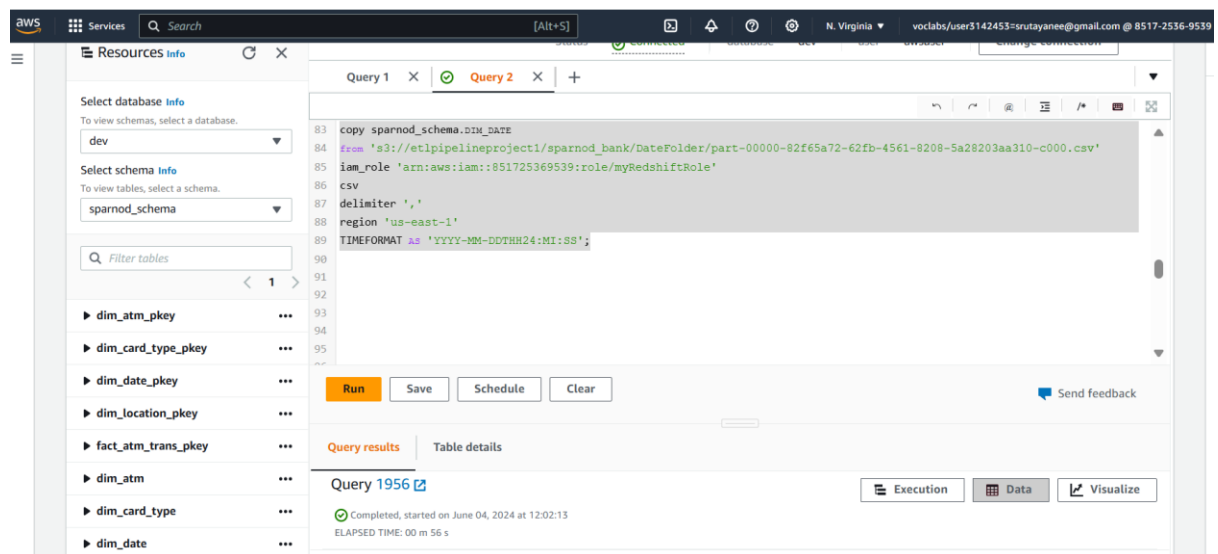
```
iam_role 'arn:aws:iam::851725369539:role/myRedshiftRole'
```

```
csv
```

```
delimiter ','
```

```
region 'us-east-1'
```

```
TIMEFORMAT AS 'YYYY-MM-DDTHH24:MI:SS';
```



4. Let's, copy the data from S3 bucket specific to location into DIM_CARD_TYPE table in Redshift cluster and below is the command we used:

```
copy sparnod_schema.DIM_CARD_TYPE
```

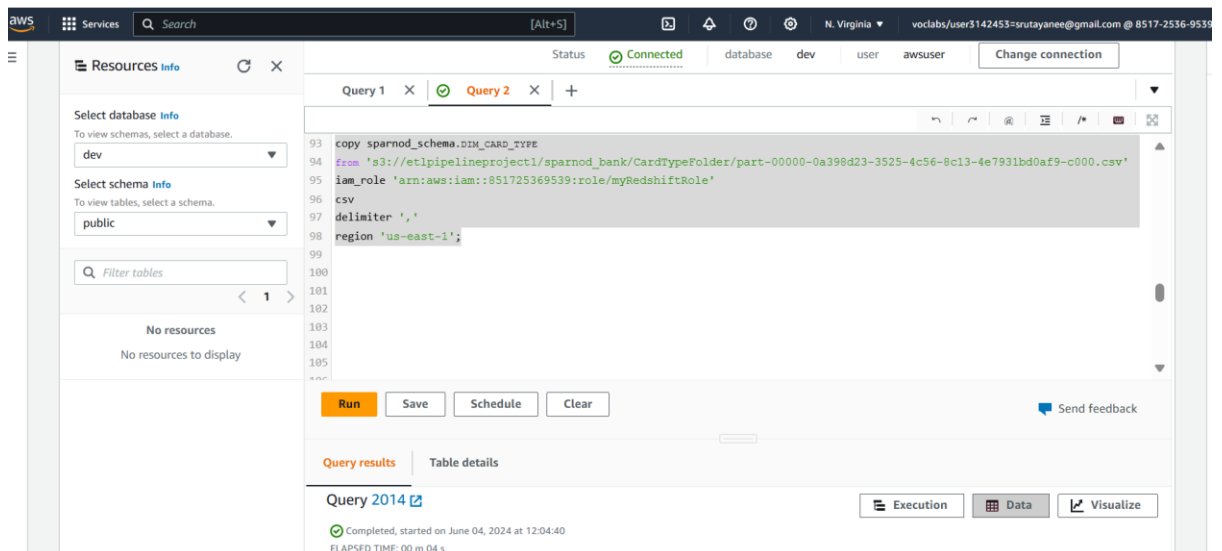
```
from 's3://etlpipelineproject1/sparnod_bank/CardTypeFolder/part-00000-0a398d23-3525-4c56-8c13-4e7931bd0af9-c000.csv'
```

```
iam_role 'arn:aws:iam::851725369539:role/myRedshiftRole'
```

```
csv
```

```
delimiter ','
```

```
region 'us-east-1';
```

- Let's, copy the data from S3 bucket specific to location into FACT_ATM_TRANS table in Redshift cluster and below is the command we used:

```
copy sparnod_schema.FACT_ATM_TRANS
```

```
from 's3://etlpipelineproject1/sparnod_bank/FactATMFolder/part-00000-b6c38d06-037c-4604-9a9d-c700ae8475c3-c000.csv'
```

```
iam_role 'arn:aws:iam::851725369539:role/myRedshiftRole'
```

```
csv
```

```
delimiter ','
```

```
region 'us-east-1';
```

