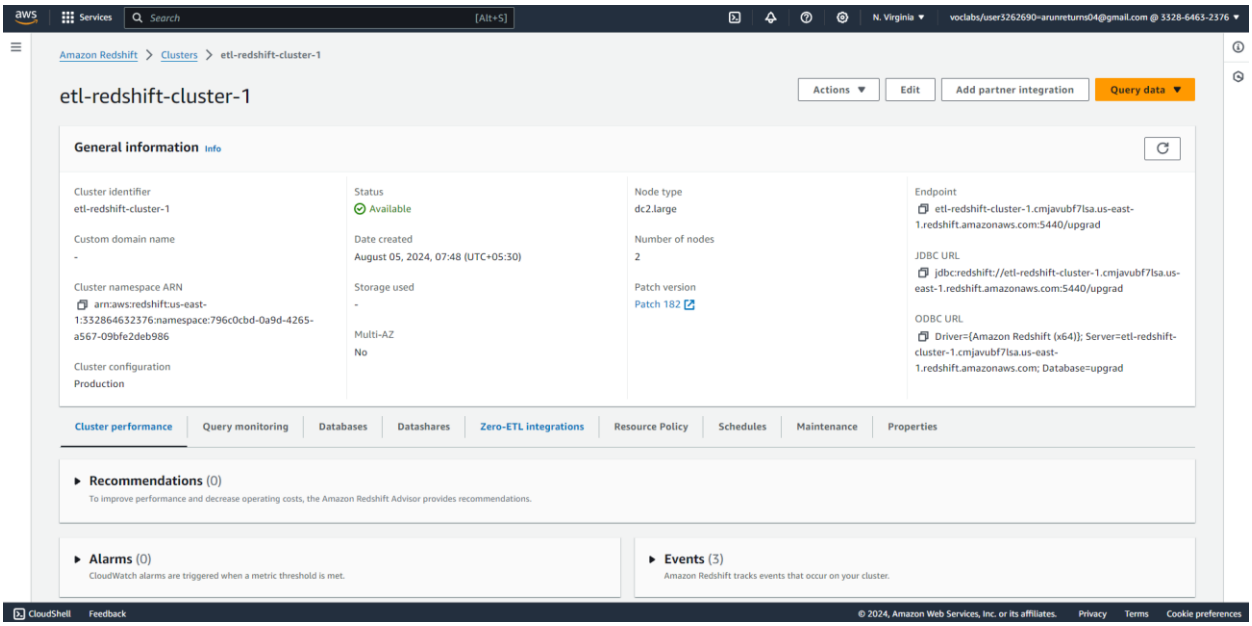
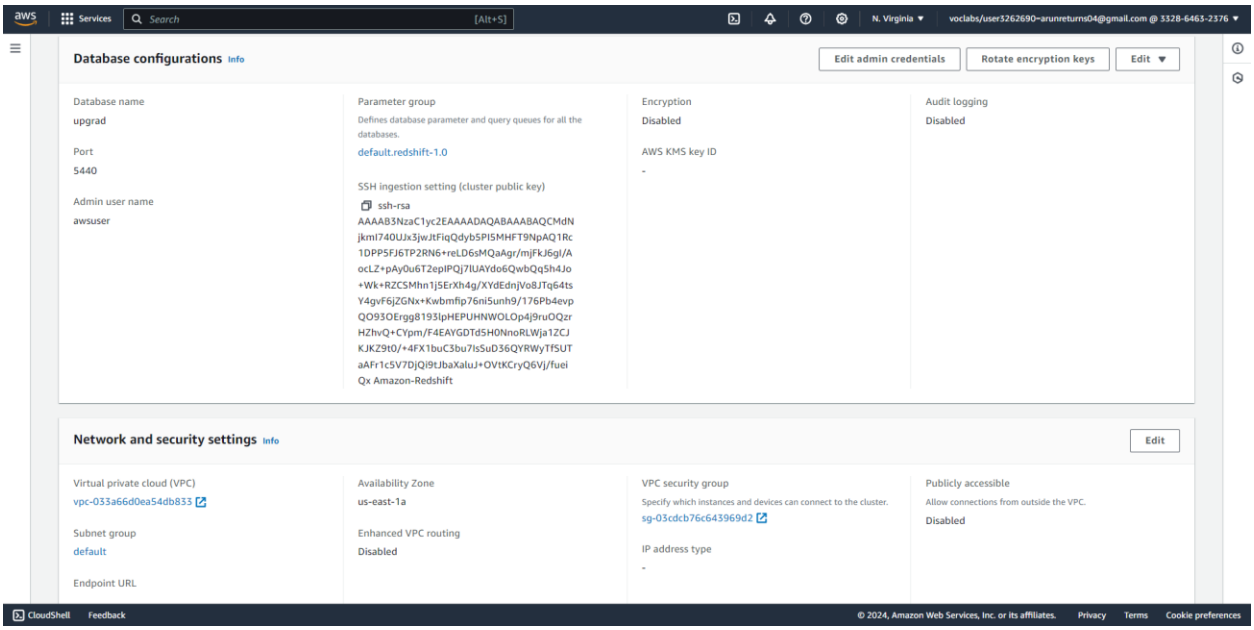
Creation of a Redshift Cluster

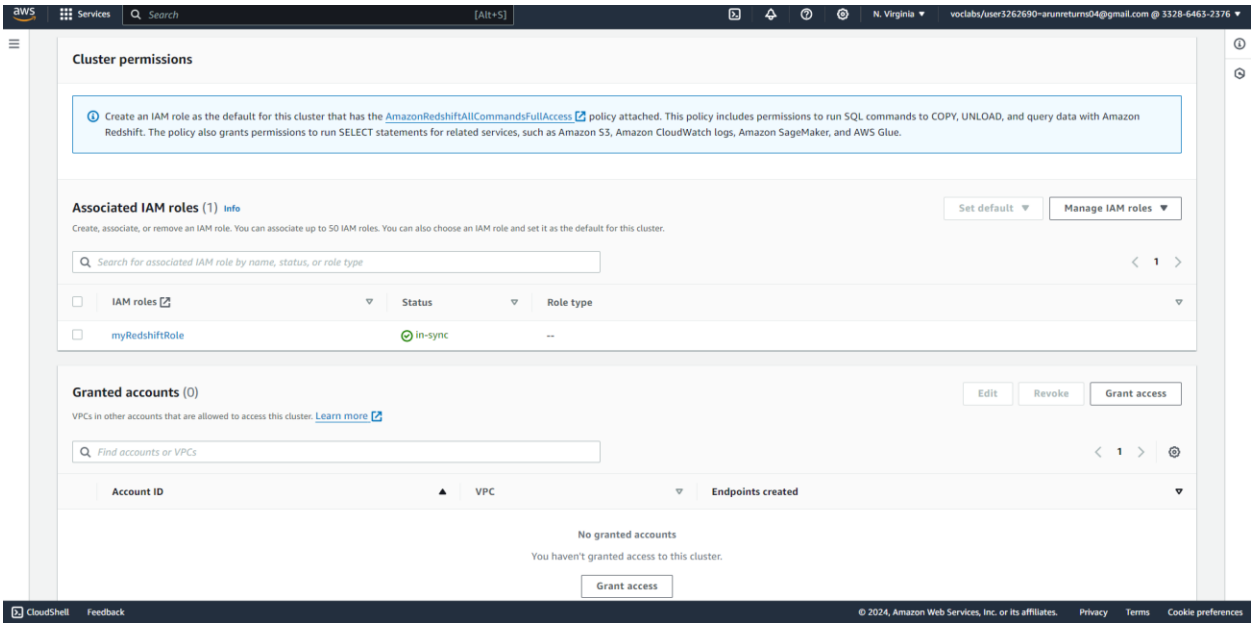
Screenshots of the configuration of the Redshift cluster that you have created:



Database configurations



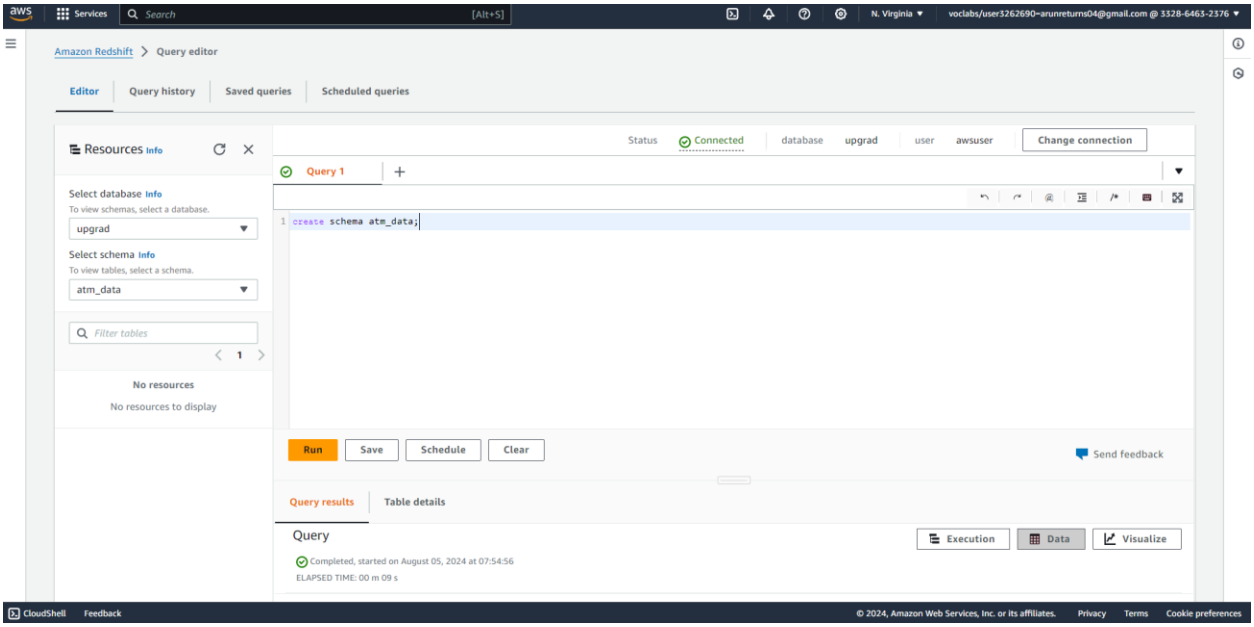
Cluster permissions



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

Query for creating a schema:

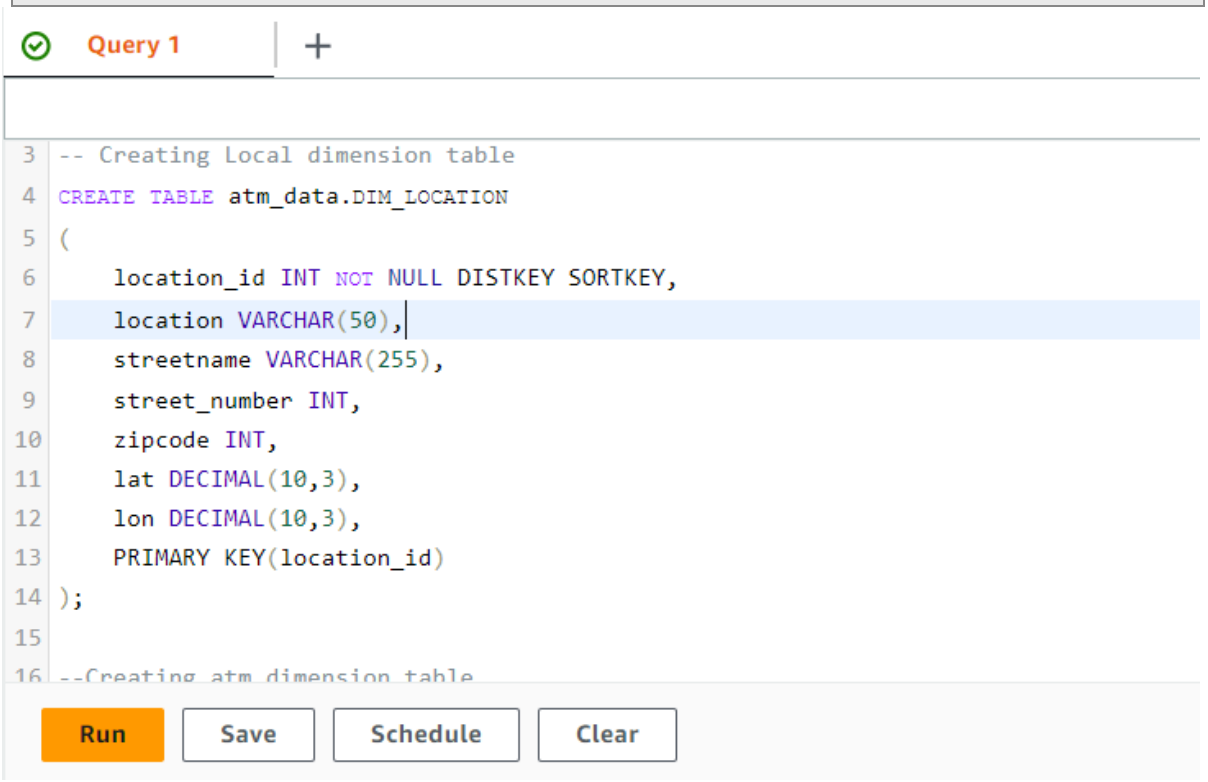
```
create schema atm_data;
```



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

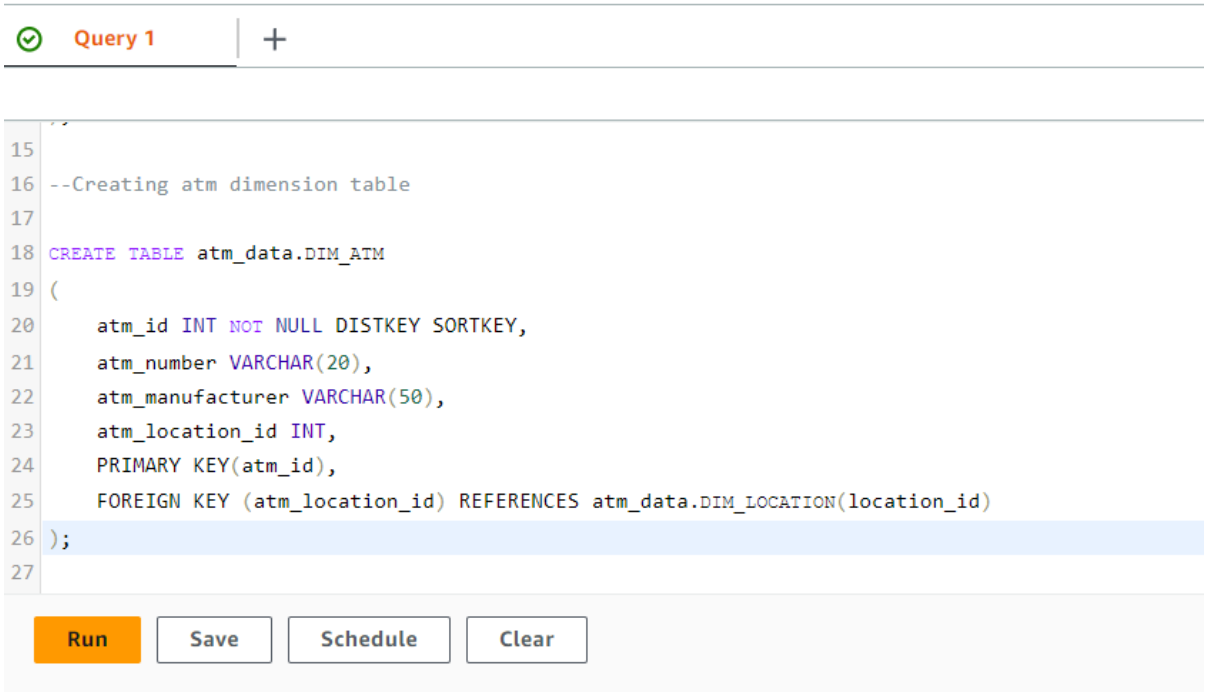
- **Creating Local dimension table**

```
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)
);
```



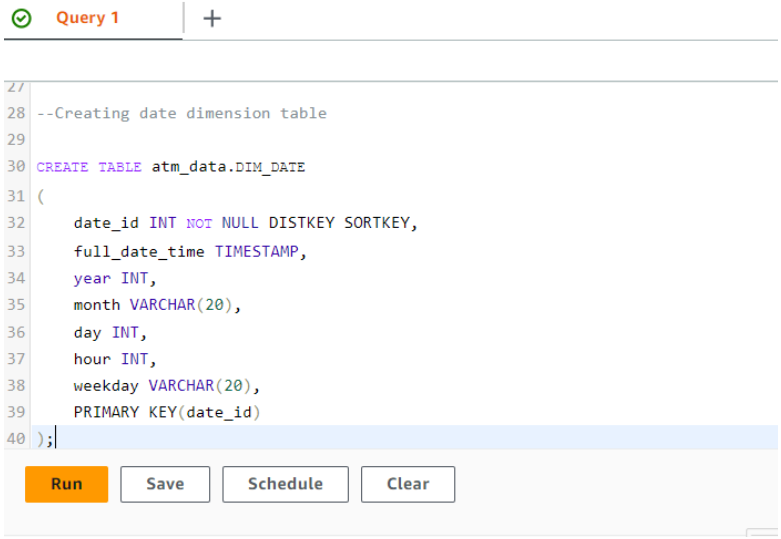
- **Creating atm dimension table**

```
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)
);
```



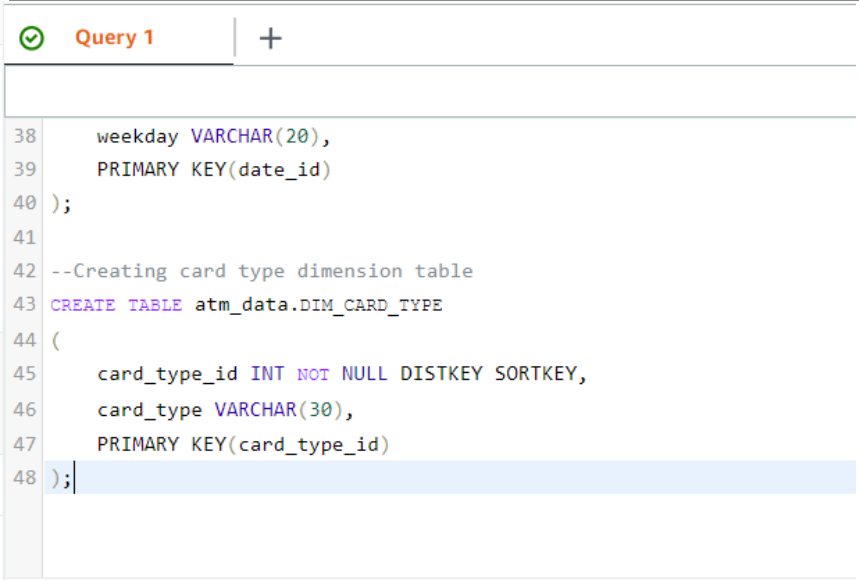
- **Creating date dimension table**

```
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)
);
```



- **Creating card type dimension table**

```
CREATE TABLE atm_data.DIM_CARD_TYPE
(
    card_type_id INT NOT NULL DISTKEY SORTKEY,
    card_type VARCHAR(30),
    PRIMARY KEY(card_type_id)
);
```



- **Creating atm transactions fact table**

```
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)
);
```

```
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)
);
```

Run Save Schedule Clear

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 the data to dim_location table**

```
COPY atm_data.dim_location
FROM 's3://upgrad-bucket-s3-1/dim_location/part-00000-6edaea43-bb5d-4253-a499-c618a3e566a8-c000.csv'
IAM_ROLE 'arn:aws:iam::332864632376:role/myRedshiftRole'
DELIMITER ','
REGION 'us-east-1'
CSV;
```

- **Copying the data to dim_atm table**

Query 1

```
--• Copying the data to dim_location table

FROM 's3://upgrad-bucket-s3-1/dim_location/part-00000-340cf3f0-da30-4d93-b125-bc3c543a4830-c000.csv'
IAM_ROLE 'arn:aws:iam::332864632376:role/myRedshiftRole'
DELIMITER ','
REGION 'us-east-1'
CSV;
```

Query results | Table details

dim_location

atm_data.dim_location

Show schema | Preview data

Search rows

location_id	location	streetname	street_number	zipcode	lat	lon
0	Kolding	Vejlevej	135	6000	55.505	9.457
3	Odense	FÃƒÂ¦lledvej	3	5000	55.394	10.370
8	Skive	Adelgade	8	7800	56.567	9.027
12	Hobro	Adelgade	31	9500	56.638	9.794
13	Nordkraft	Kjellerups Torv	1	9000	57.047	9.932
17	Hasseris	Hasserisvej	113	9000	57.044	9.898
21	Silkeborg	Borgergade	36	8600	56.179	9.552
26	NykÃƒÂ¸bing Mors	Kirketorvet	1	7900	56.795	8.860
30	Menu KÃƒÂ¸benhavn Klarup	Klarupvej	52	9270	57.013	10.046
31	Aarhus	SÃƒÂ¸nder Alle	11	8000	56.153	10.206

```
COPY atm_data.dim_atm
FROM 's3://upgrad-bucket-s3-1/dim_atm/part-00000-5bdb3bbc-35fc-44c2-ac36-05aea46e0bf8-c000.csv'
IAM_ROLE 'arn:aws:iam::332864632376:role/myRedshiftRole'
DELIMITER ','
REGION 'us-east-1'
CSV;
```

```
COPY atm_data.dim_atm
FROM 's3://upgrad-bucket-s3-1/dim_atm/part-00000-e95db0ed-a34e-4d59-a11d-2fdd5798ee9c-c000.csv'
IAM_ROLE 'arn:aws:iam::332864632376:role/myRedshiftRole'
DELIMITER ','
REGION 'us-east-1'
CSV;
```

- **Copying the data to dim_date table**

```
COPY atm_data.dim_date
FROM 's3://upgrad-bucket-s3-1/dim_date/part-00000-10328b66-88c7-402f-ac6f-cd69109ef5bd-c000.csv'
IAM_ROLE 'arn:aws:iam::332864632376:role/myRedshiftRole'
DELIMITER ','
REGION 'us-east-1'
CSV
TIMEFORMAT 'auto';
```

- **Copying the data to dim_card_type table**

```
COPY atm_data.dim_card_type
FROM 's3://upgrad-bucket-s3-1/dim_card_type/part-00000-3529d67c-af4c-4310-98a5-624c707e575c-c000.csv'
IAM_ROLE 'arn:aws:iam::332864632376:role/myRedshiftRole'
DELIMITER ','
REGION 'us-east-1'
CSV;
```

- **Copying the data to fact_atm_trans table**

```
COPY atm_data.fact_atm_trans
FROM 's3://upgrad-bucket-s3-1/fact_atm_trans/part-00000-b00b2638-1f9b-4e95-8ec2-52c4936926ef-c000.csv'
IAM_ROLE 'arn:aws:iam::332864632376:role/myRedshiftRole'
DELIMITER ','
REGION 'us-east-1'
CSV;
```

```
--Copying the data to fact_atm_trans table
COPY atm_data.fact_atm_trans
FROM 's3://upgrad-bucket-s3-1/fact_atm_trans/part-00000-38e20993-698e-4785-b511-38fef6b4559f-c000.csv'
IAM_ROLE 'arn:aws:iam::332864632376:role/myRedshiftRole'
DELIMITER ','
REGION 'us-east-1'
CSV;
```

Run | Save | Schedule | Clear

Checking whether the data is loaded or not

```
select * from atm_data.fact_atm_trans
limit 1;
```

```
--Checking whether the data is loaded or not
select * from atm_data.fact_atm_trans
limit 3;
```

Run | Save | Schedule | Clear | Send feedback

Query results | Table details

Query 2072

Execution | Data | Visualize

Completed, started on August 05, 2024 at 08:11:35
ELAPSED TIME: 00 m 02 s

Rows returned (3)

Export

Search rows

1

trans_id	atm_id	weather_loc_id	date_id	card_type_id	atm_status	currency	service	transaction_amount	message_c
5	133	56	0	1	Active	DKK	Withdrawal	8981	
15	142	90	0	5	Active	DKK	Withdrawal	5508	