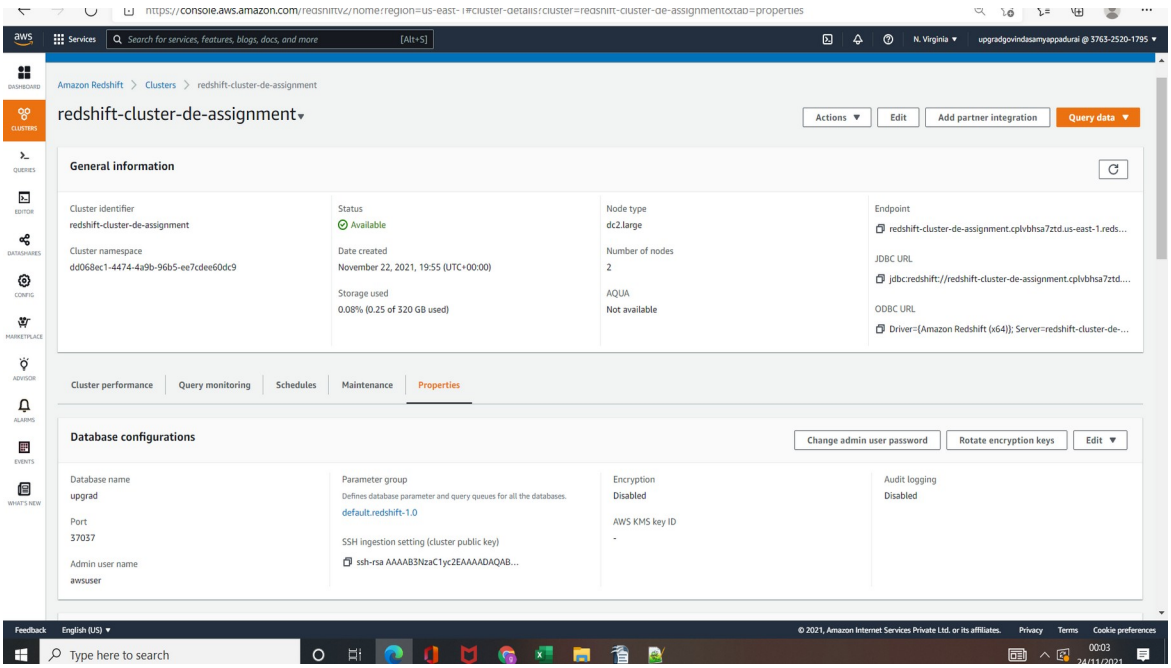
Creation of a Redshift Cluster

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



Steps I have carried out:

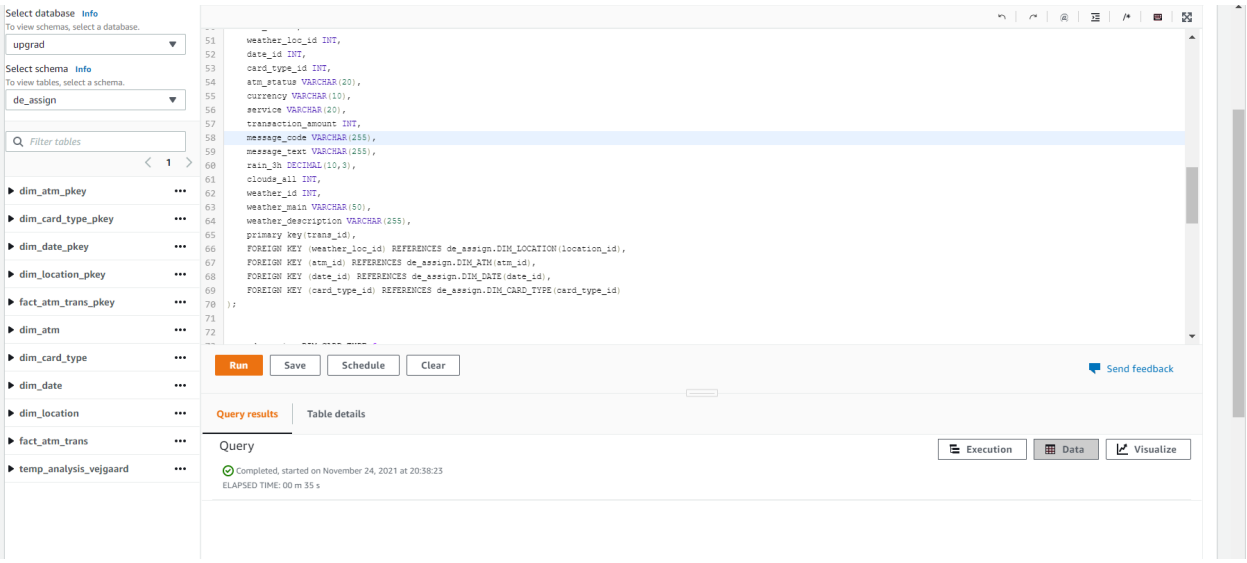
1. Created IAM role for accessing S3 from Redshift cluster
2. Assigned existing VPC
3. Configured Subnet groups
4. Associated the IAM to the cluster
5. Created cluster as above and used for analysis purpose

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:

```
create schema DE_ASSIGN;

CREATE TABLE DE_ASSIGN.DIM_LOCATION
(
        location_id int not null,
        location varchar(50),
        streetname varchar(255),
        street_number int,
        zipcode int,
        lat decimal(10,3),
        lon decimal(10,3),
        primary key(location_id)
);


CREATE TABLE DE_ASSIGN.DIM_ATM
(
        atm_id int not null,
        atm_number VARCHAR(20),
        atm_manufacturer VARCHAR(50),
        atm_location_id INT,
        PRIMARY KEY(atm_id),
        foreign key(atm_location_id) references DE_ASSIGN.dim_location(location_id)
)
;


CREATE TABLE DE_ASSIGN.DIM_DATE
(
        date_id int not null,
        full_date_time TIMESTAMP,
        year int,
        month VARCHAR(20),
        day int,
        hour int,
        weekday varchar(20),
```

```
        primary key(date_id)
);

CREATE TABLE DE_ASSIGN.DIM_CARD_TYPE
(
        card_type_id int,
        card_type varchar(30),
        primary key(card_type_id)
);
CREATE TABLE DE_ASSIGN.FACT_ATM_TRANS
(
        trans_id BIGINT,
        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
de_assign.DIM_LOCATION(location_id),
        FOREIGN KEY (atm_id) REFERENCES de_assign.DIM_ATM(atm_id),
        FOREIGN KEY (date_id) REFERENCES de_assign.DIM_DATE(date_id),
   FOREIGN KEY (card_type_id) REFERENCES
de_assign.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 de_assign.DIM_CARD_TYPE from
's3://de-assignment-gov/dim_card_type/part-00000-4dc336c8-1efc-4eb4-adc3-22f468512297-c000.csv'
iam_role 'arn:aws:iam::376325201795:role/de_assign_role' IGNOREHEADER 1
delimiter ',' region 'us-east-1';
```

```
copy de_assign.DIM_LOCATION from
's3://de-assignment-gov/dim_location/part-00000-367ea558-2914-48a3-8212-00197012ff71-c000.csv'
iam_role 'arn:aws:iam::376325201795:role/de_assign_role' IGNOREHEADER 1
delimiter ',' region 'us-east-1';
```

```
copy de_assign.DIM_DATE from
's3://de-assignment-gov/dim_date/part-00000-fabc4e1b-2fa1-47ca-b41d-4b9eaa4946fc-c000.csv'
iam_role 'arn:aws:iam::376325201795:role/de_assign_role' IGNOREHEADER 1
delimiter ',' region 'us-east-1';
```

copy de_assign.DIM_ATM from
's3://de-assignment-gov/dim_atm/part-00000-317af816-9829-4e93-bc47-a0b6d69c5a94-c000.csv'
iam_role 'arn:aws:iam::376325201795:role/de_assign_role' IGNOREHEADER 1
delimiter ',' region 'us-east-1';

copy de_assign.fact_atm_trans from
's3://de-assignment-gov/fact_atm_trans/part-00000-6d7e6e0a-c195-40f5-90c8-3a0ee63b0e19-c000.csv'
iam_role 'arn:aws:iam::376325201795:role/de_assign_role' IGNOREHEADER 1
delimiter ',' region 'us-east-1';