

## Creation of a Redshift Cluster

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

| redshift-cluster-etl-proj ▾                               |   |                        |  | Actions ▾ | Edit | Add partner integration | Query data ▾ |
|---|---|------------------------|--|-----------|------|-------------------------|--------------|
| General information                                       |   |                        |  |           |      |                         |              |
| Cluster identifier<br>redshift-cluster-etl-proj           | Status<br>Available                             | Node type<br>dc2.large | Endpoint<br>redshift-cluster-etl-proj.cnoseukjer1.us-east-1.redshift.amaz...   |           |      |                         |              |
| Cluster namespace<br>0ba7465e-f139-41f9-a1c6-6246b291e569 | Date created<br>May 04, 2022, 23:13 (UTC+05:30) | Number of nodes<br>2   | JDBC URL<br>jdbc:redshift://redshift-cluster-etl-proj.cnoseukjer1.us-east-1... |           |      |                         |              |
|   | Storage used<br>0.09% (0.29 of 320 GB used)     | AQUA<br>Not available  | ODBC URL<br>Driver={Amazon Redshift (x64)}; Server=redshift-cluster-etl-p...   |           |      |                         |              |

| Database configurations    |   |                        |                           | Change admin user password | Rotate encryption keys | Edit ▾ |
|----------------------------|---|------------------------|---------------------------|----------------------------|------------------------|--------|
| Database name<br>dev       | Parameter group<br>Defines database parameter and query queues for all the databases.<br>default.redshift-1.0 | Encryption<br>Disabled | Audit logging<br>Disabled |                            |                        |        |
| Port<br>5436               | SSH ingestion setting (cluster public key)  | AWS KMS key ID<br>-    |                           |                            |                        |        |
| Admin user name<br>awsuser | ssh-rsa AAAAB3NzaC1yc2EAAAADAQAB...   |                        |                           |                            |                        |        |

| Network and security settings                       |                                  |   |  | Edit |
|---|----------------------------------|---|--|------|
| Virtual private cloud (VPC)<br>vpc-0430be62ae940a9f | Availability Zone<br>us-east-1a  | VPC security group<br>Specify which instances and devices can connect to the cluster.<br>sg-0d4a2cb80a4995179<br>sg-03f99c4b768471234 | Publicly accessible<br>Allow instances and devices outside the VPC to connect to your database through the cluster endpoint.<br>Disabled |      |
| Subnet<br>cluster-subnet-group-1-redshift-upgrad    | Enhanced VPC routing<br>Disabled |   |  |      |
| Endpoint URL<br>-                                   |                                  |   |  |      |

| Associated IAM roles (1) info                                |                                |         |           | Set default ▾ | Manage IAM roles ▾ |
|--|--------------------------------|---------|-----------|---------------|--------------------|
| Search for associated IAM role by name, status, or role type |                                |         |           |               |                    |
| <input type="checkbox"/>                                     | IAM roles                      | Status  | Role type |               |                    |
| <input type="checkbox"/>                                     | Upgrad_Redshift_S3_Full_Access | in-sync | --        |               |                    |

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 etl\_project;

```
create table etl_project.location_dim(  
location_id integer not null distkey sortkey,  
location varchar(30),  
streetname varchar(30),  
street_number integer,  
zipcode integer,  
lat decimal(10,3),  
lon decimal(10,3));
```

```
create table etl_project.atm_dim(  
atm_id integer not null distkey sortkey,  
atm_number integer,  
atm_manufacturer varchar(30),  
atm_location_id integer,  
street_number integer);
```

```
create table etl_project.date_dim(  
date_id date not null distkey sortkey,  
full_date_time timestamp,  
year integer,  
month varchar(20),  
day integer,  
hour integer,  
weekday varchar(20));
```

```
create table etl_project.card_type_dim(  
card_type_id integer not null distkey sortkey,  
card_type varchar(30));
```

-- TRANSACTION TABLE:

```
create table etl_project.atm_trans_fact(  
trans_id bigint not null distkey sortkey,  
atm_id integer,  
weather_loc_id integer,  
date_id date,  
card_type_id integer,  
atm_status varchar(20),  
currency varchar(10),  
service varchar(20),  
message_code varchar(255),  
message_text varchar(255),  
rain_3h decimal(10,3),  
clouds_all integer,  
weather_id integer,  
weather_main varchar(50),  
weather_description varchar(255),  
transaction_amount integer);
```

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 etl_project.location_dim from  
's3://etl-proj/location-dimension/location_dim.csv/'  
iam_role 'arn:aws:iam::521912254229:role/Upgrad_Redshift_S3_Full_Access'  
delimiter ',' region 'us-east-1'  
IGNOREHEADER 1;
```

```
copy etl_project.atm_dim from  
's3://etl-proj/location-dimension/atm_dim.csv/'  
iam_role 'arn:aws:iam::521912254229:role/Upgrad_Redshift_S3_Full_Access'  
delimiter ',' region 'us-east-1'  
IGNOREHEADER 1;
```

```
copy etl_project.date_dim from  
's3://etl-proj/date-dimension/date_dim.csv/'  
iam_role 'arn:aws:iam::521912254229:role/Upgrad_Redshift_S3_Full_Access'
```

```
delimiter ',' region 'us-east-1'  
IGNOREHEADER 1;
```

```
copy etl_project.card_type_dim from  
's3://etl-proj/card-type-dimension/card_type_dim.csv/'  
iam_role 'arn:aws:iam::521912254229:role/Upgrad_Redshift_S3_Full_Access'  
delimiter ',' region 'us-east-1'  
IGNOREHEADER 1;
```

```
copy etl_project.atm_trans_fact from  
's3://etl-proj/transaction-fact/atm_trans_fact.csv/'  
iam_role 'arn:aws:iam::521912254229:role/Upgrad_Redshift_S3_Full_Access'  
delimiter ',' region 'us-east-1'  
IGNOREHEADER 1;
```