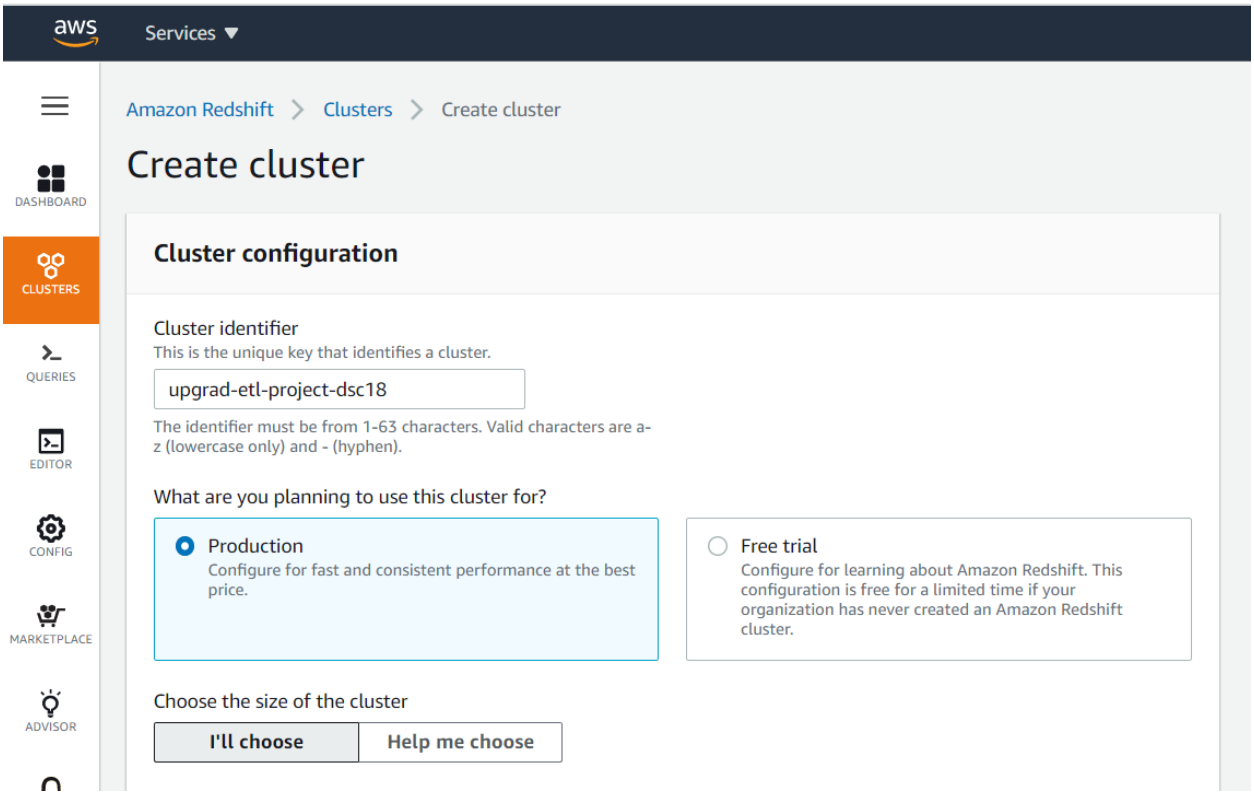
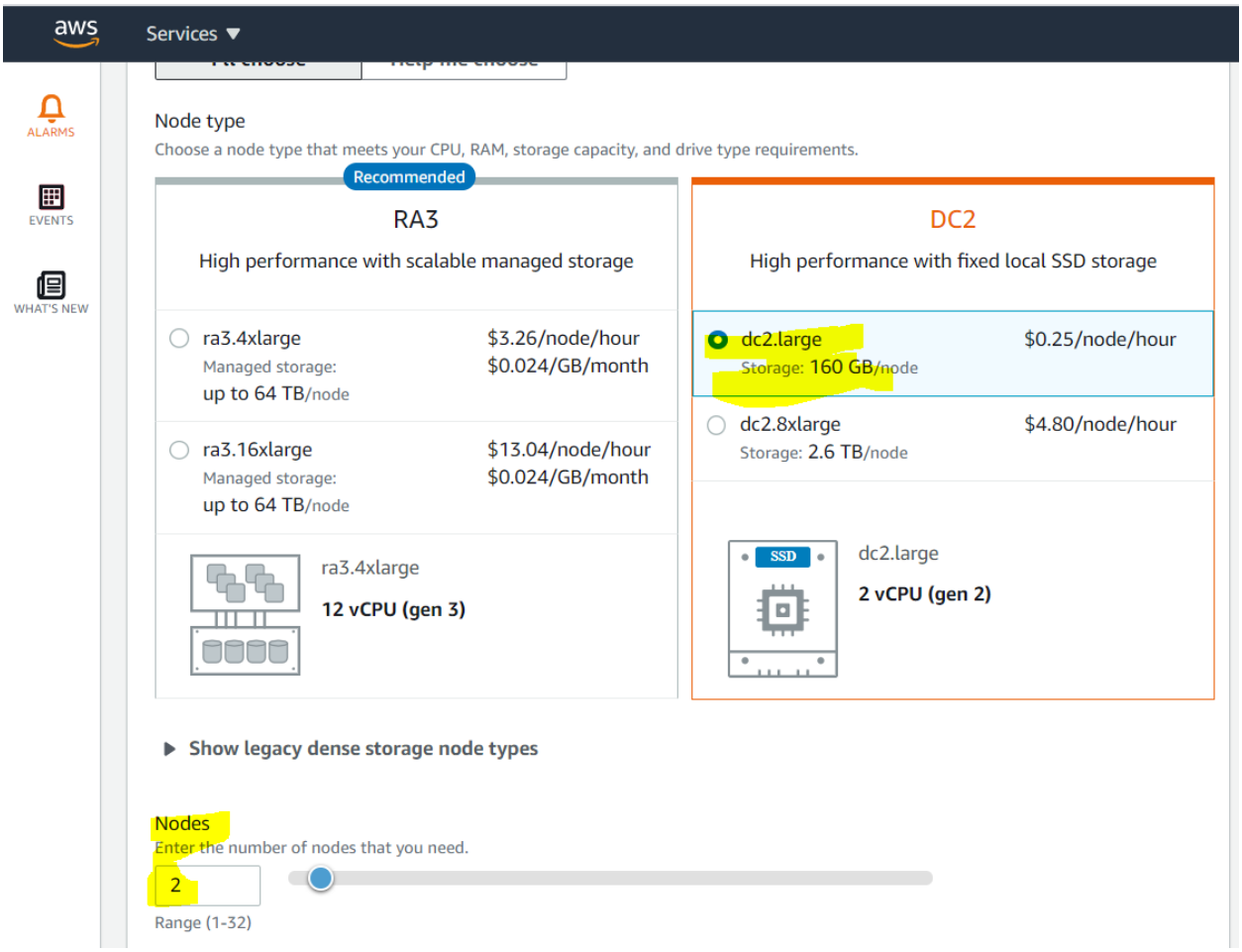
Creation of a RedShift Cluster

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





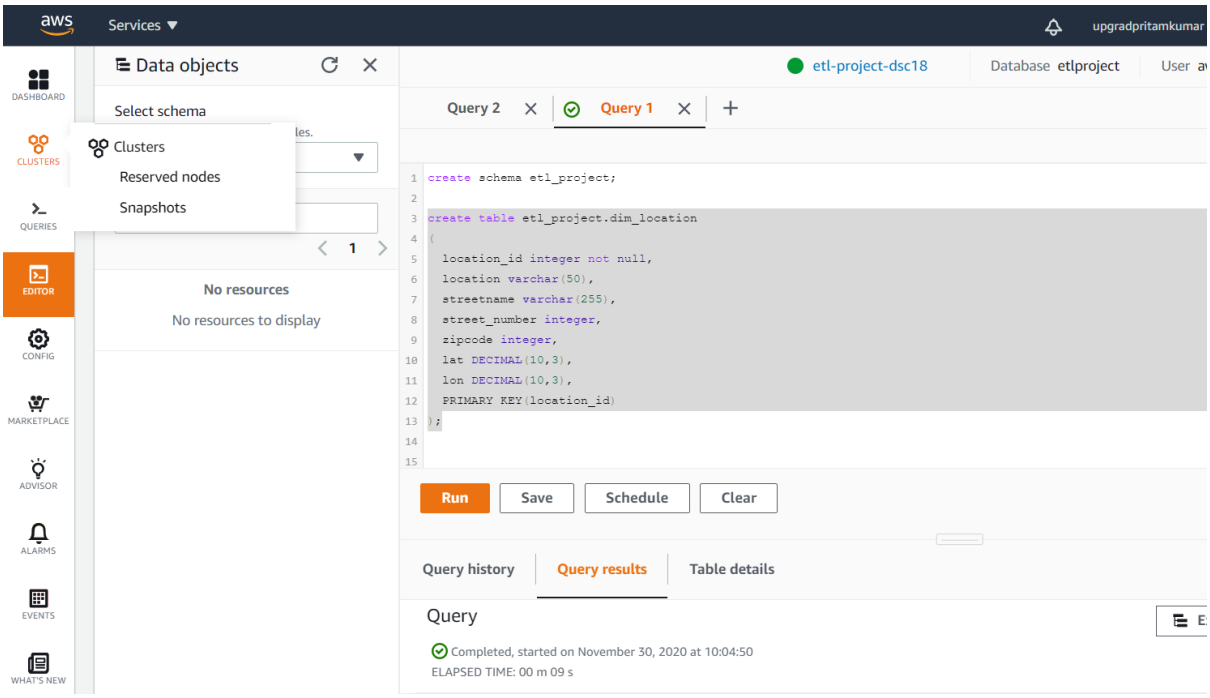
Setting up a database in the RedShift cluster and running queries to create the dimension and fact tables

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

1) **Schema creation →** create schema etl_project;

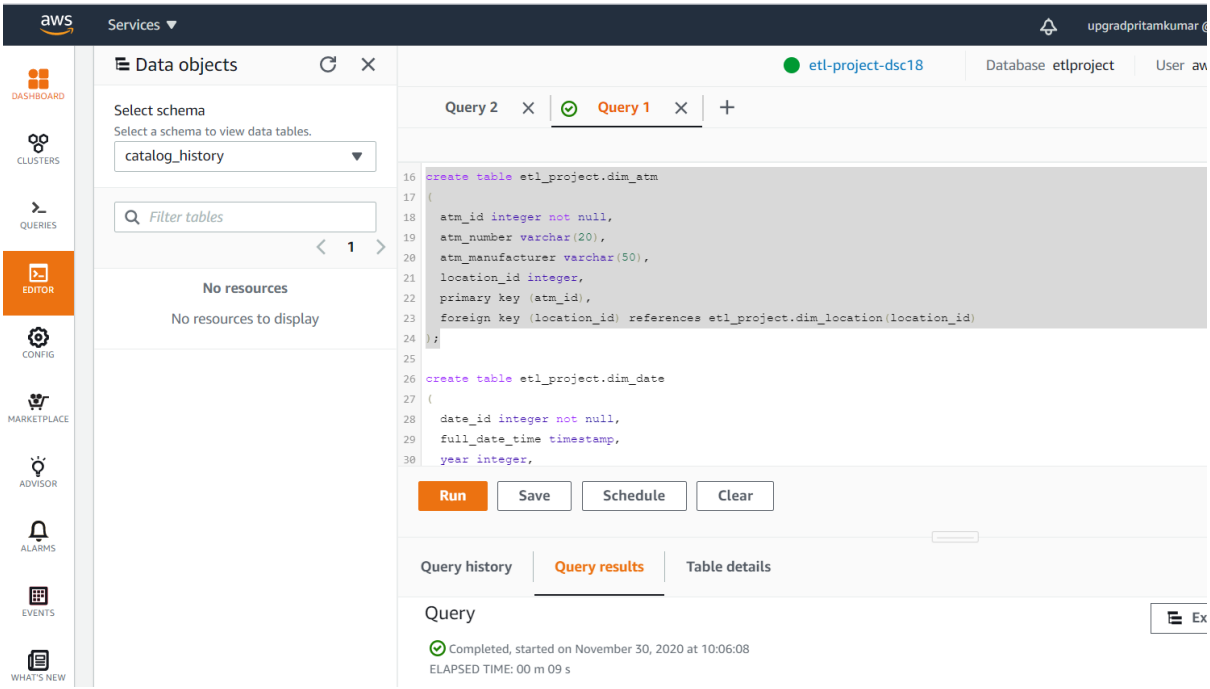
2) **Dim_Location Table creation →**

```
create table etl_project.dim_location
(
  location_id integer not null,
  location varchar(50),
  streetname varchar(255),
  street_number integer,
  zipcode integer,
  lat DECIMAL(10,3),
  lon DECIMAL(10,3),
  PRIMARY KEY(location_id)
);
```



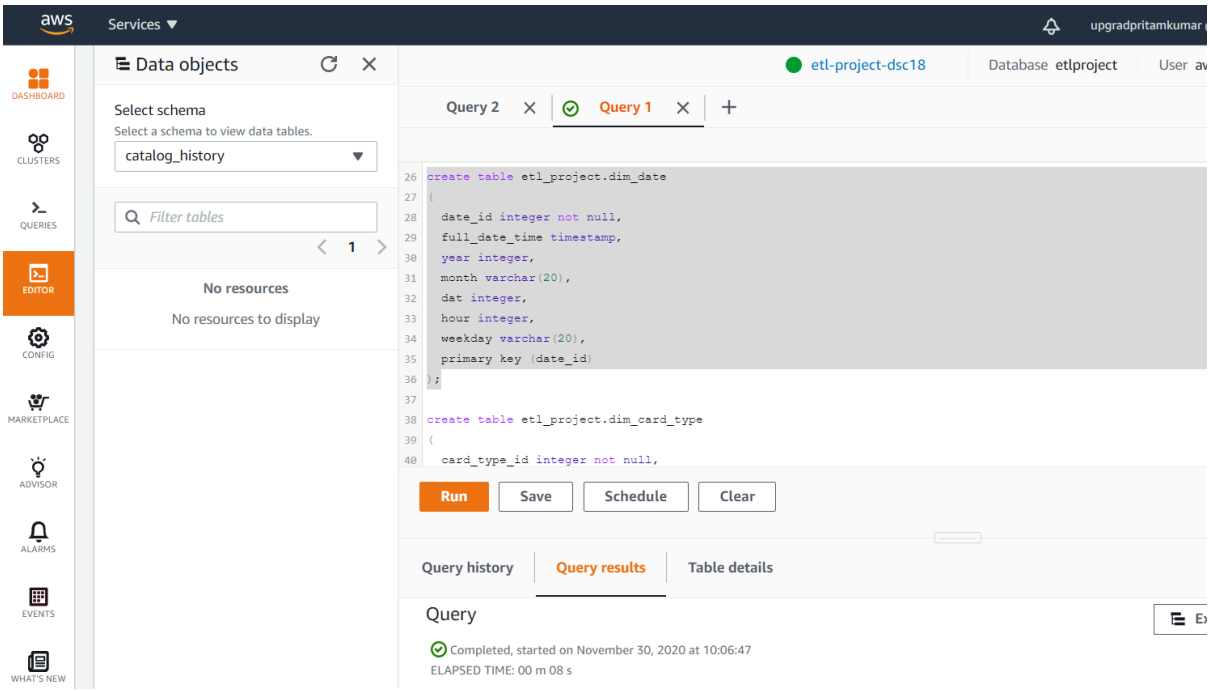
3) DIM_ATM table creation →

```
create table etl_project.dim_atm
(
  atm_id integer not null,
  atm_number varchar(20),
  atm_manufacturer varchar(50),
  location_id integer,
  primary key (atm_id),
  foreign key (location_id) references etl_project.dim_location(location_id)
);
```



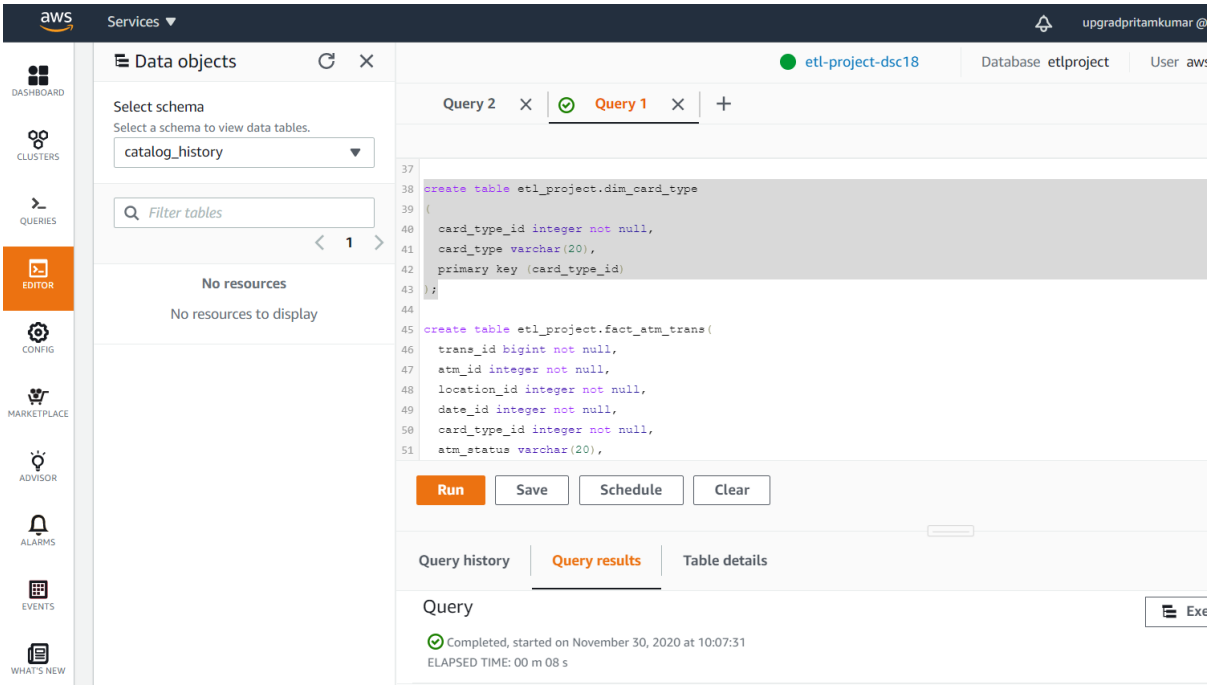
4) DIM_DATE table creation →

```
create table etl_project.dim_date
(
  date_id integer not null,
  full_date_time timestamp,
  year integer,
  month varchar(20),
  dat integer,
  hour integer,
  weekday varchar(20),
  primary key (date_id)
);
```



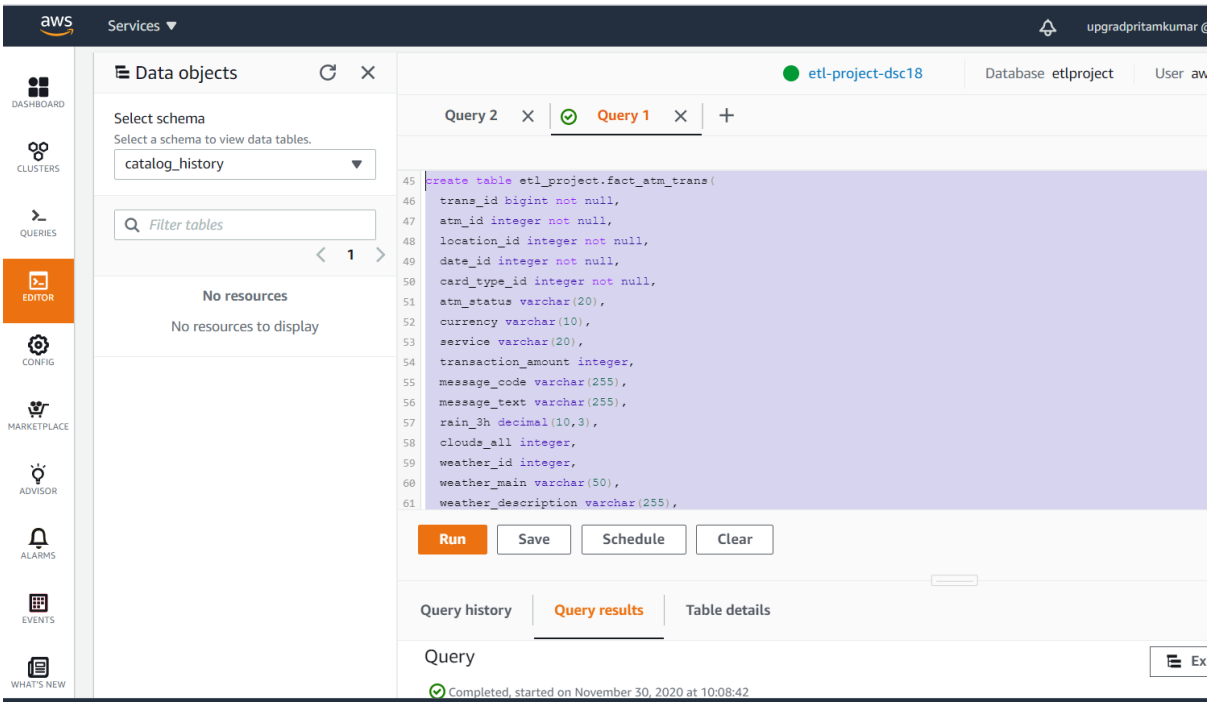
5) DIM_CARD_TYPE table creation →

```
create table etl_project.dim_card_type
(
  card_type_id integer not null,
  card_type varchar(20),
  primary key (card_type_id)
);
```



6) FACT_ATM_TRANS table creation →

```
create table etl_project.fact_atm_trans(
  trans_id bigint not null,
  atm_id integer not null,
  location_id integer not null,
  date_id integer not null,
  card_type_id integer not null,
  atm_status varchar(20),
  currency varchar(10),
  service varchar(20),
  transaction_amount integer,
  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),
  primary key (trans_id),
  foreign key (location_id) references etl_project.dim_location(location_id),
  foreign key (atm_id) references etl_project.dim_atm(atm_id),
  foreign key (date_id) references etl_project.dim_date(date_id),
  foreign key (card_type_id) references etl_project.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

1) copy etl_project.dim_location from 's3://etl-project-dsc18/location-dim/part-00000-615b2af3-571b-4f08-99a2-2593db79654f-c000.csv'
iam_role 'arn:aws:iam::113692286195:role/upgrad-etl-project-dsc18' csv IGNOREHEADER 1 delimiter ',' region 'us-east-1';
2) copy etl_project.dim_atm from 's3://etl-project-dsc18/atm-dim/part-00000-e9e6d7de-1169-466e-b67b-38a780751d41-c000.csv'
iam_role 'arn:aws:iam::113692286195:role/upgrad-etl-project-dsc18' csv IGNOREHEADER 1 delimiter ',' region 'us-east-1';
3) copy etl_project.dim_card_type from 's3://etl-project-dsc18/card-type-dim/part-00000-d688fcb7-c70f-42a7-a262-987479a3adcb-c000.csv'
iam_role 'arn:aws:iam::113692286195:role/upgrad-etl-project-dsc18' csv IGNOREHEADER 1 delimiter ',' region 'us-east-1';

4) copy etl_project.dim_date from 's3://etl-project-dsc18/date-dim/part-00000-f409154c-f0b1-447e-807b-9ff31ae4c6a8-c000.csv'
iam_role 'arn:aws:iam::113692286195:role/upgrad-etl-project-dsc18' csv IGNOREHEADER 1 delimiter ',' region 'us-east-1';

5) copy etl_project.fact_atm_trans from 's3://etl-project-dsc18/fact-atm-trans/part-00000-44acf61e-3e37-4a6e-8a4f-53376f75524d-c000.csv'
iam_role 'arn:aws:iam::113692286195:role/upgrad-etl-project-dsc18' csv IGNOREHEADER 1 delimiter ',' region 'us-east-1';