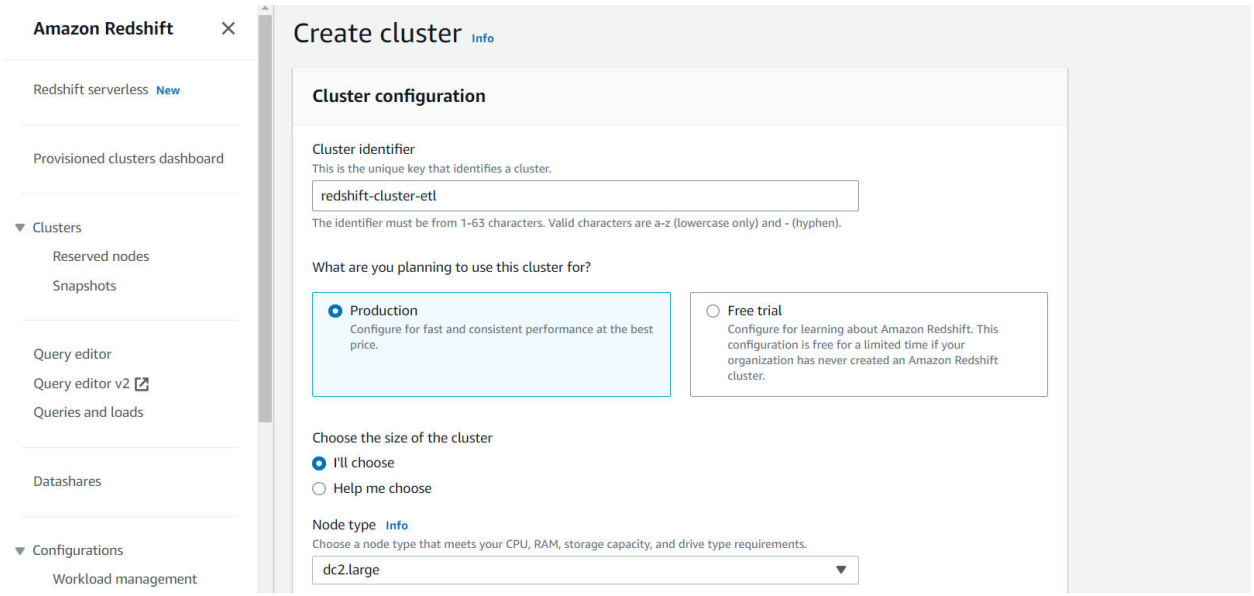
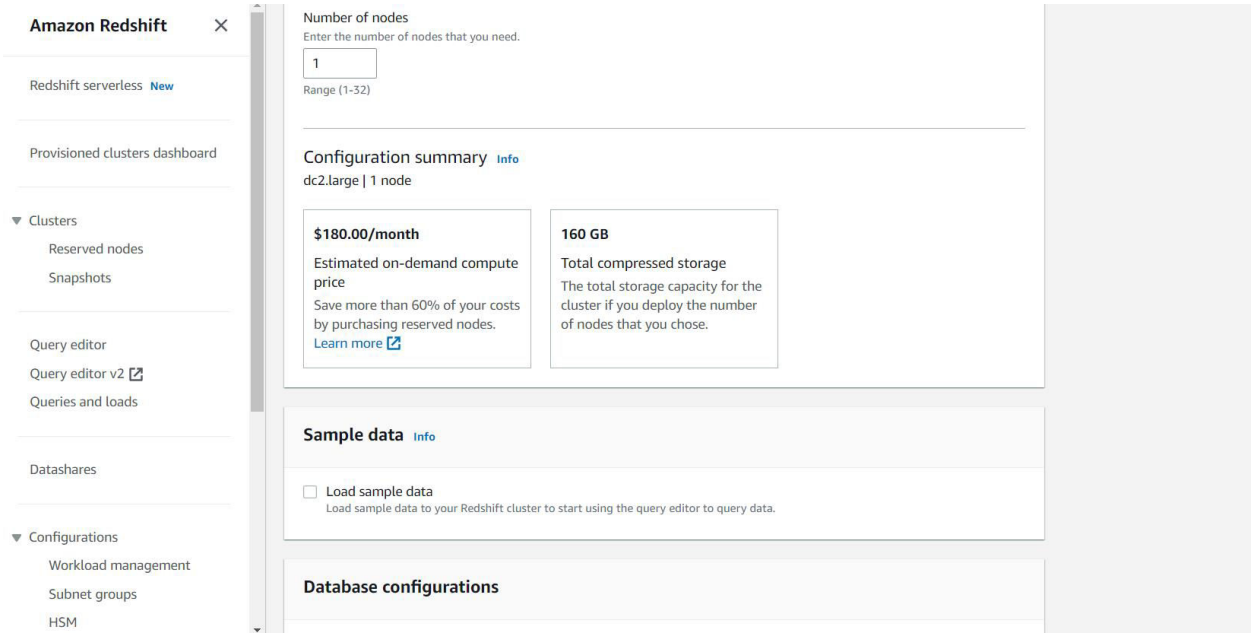
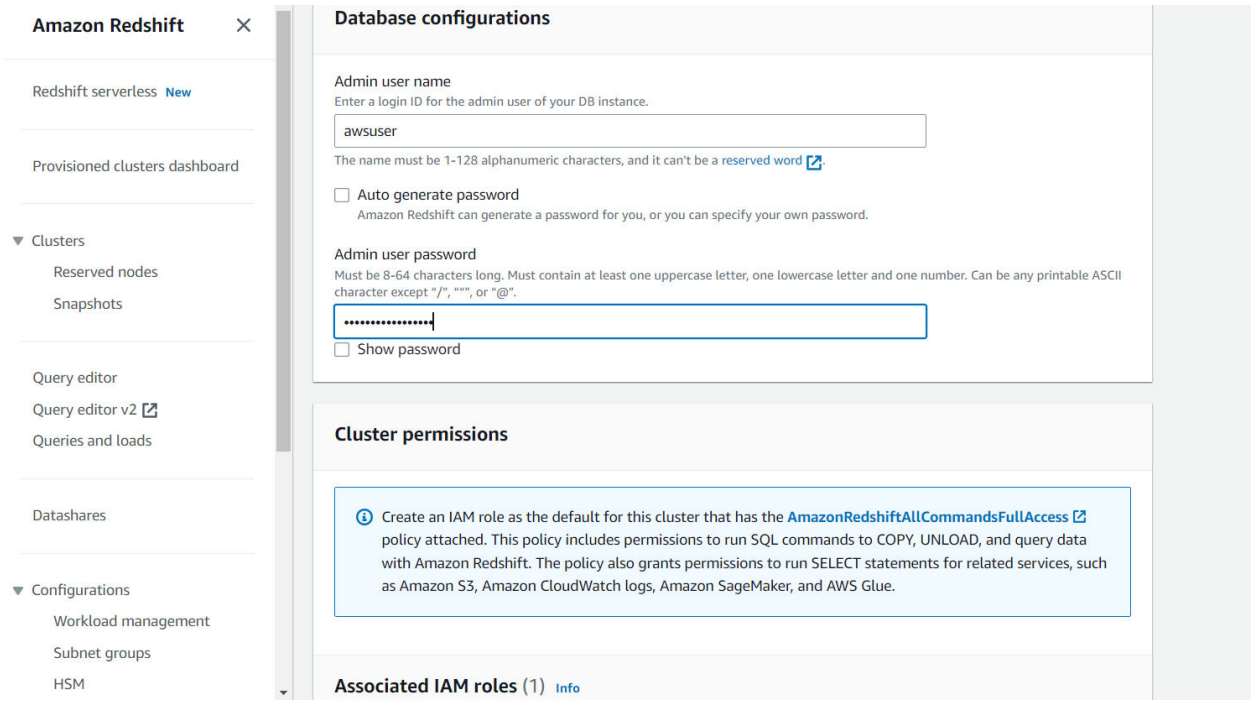
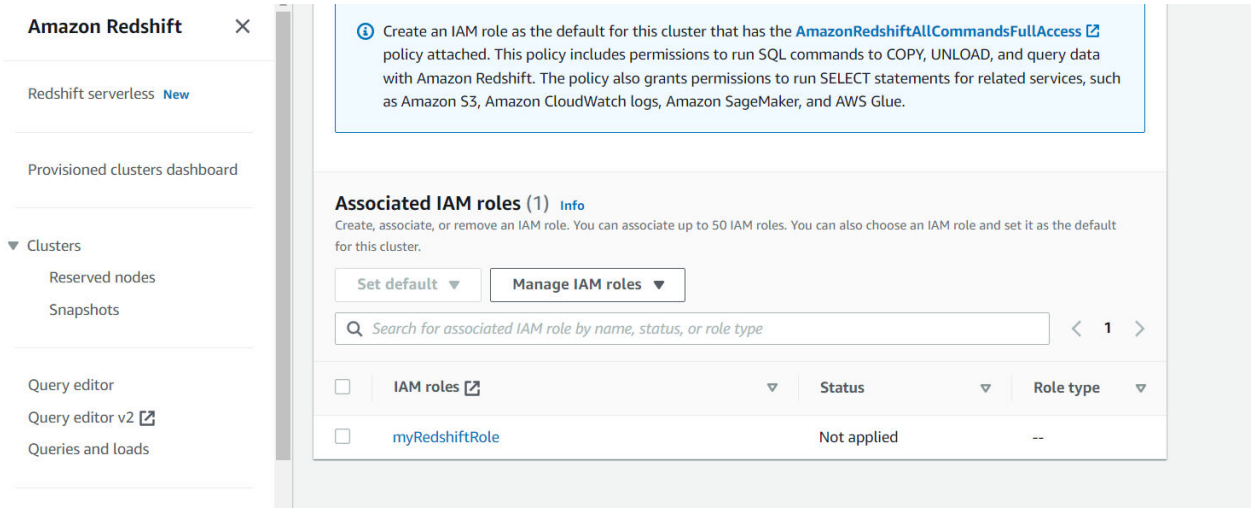
Creation of a Redshift Cluster

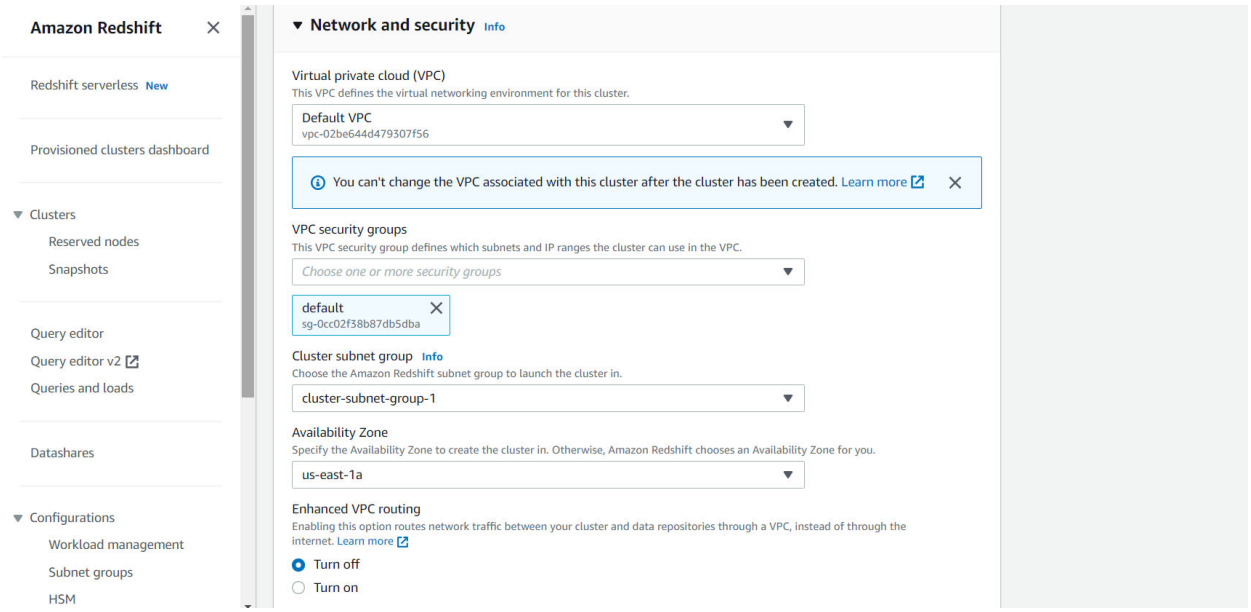
Screenshots of the configuration of the Redshift cluster that you have created:









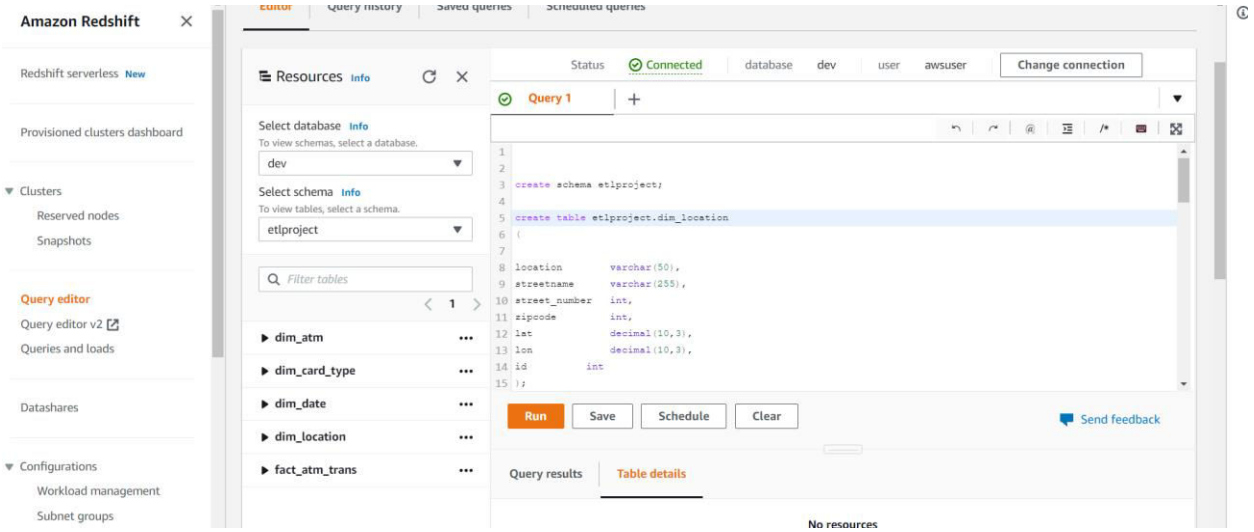


Setting up a database in the Redshift cluster and running queries to create the dimension and fact tables

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

```
create schema etlproject;

create table etlproject.dim_location
(
        location        varchar(50),
        streetname      varchar(255),
        street_number   int,
        zipcode         int,
        lat             decimal(10,3),
        lon             decimal(10,3),
        location_id     int
);


create table etlproject.dim_atm
(
        atm_number          int,
        atm_manufacturer    varchar(20),
        atm_location_id     varchar(50),
        atm_id              int
);


create table etlproject.dim_card_type
(
        card_type       varchar(30),
        card_type_id    int
);
```

```
create table etlproject.dim_date
(
        year            int,
        month           varchar(20) ,
        day             int,
        weekday         varchar(20) ,
        hour            int,
        full_date_time  varchar(30),
        date_id         int
);


create table etlproject.fact_atm_trans
(

        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),
        card_type_id            int,
        date_id                 int,
        atm_id                  int,
        weather_loc_id          int,
        trans_id                bigint

);
```



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 etlproject.dim_atm from 's3://etlprojectnew/data/dim_atm/part-00000-e0053268-59dc-47ab-bf23-3cdd69e01f23-c000.csv' iam_role 'arn:aws:iam::274560308997:role/myRedshiftRole' delimiter ',' region 'us-east-1' CSV IGNOREHEADER as 1
;

ELAPSED TIME: 00 m 08 s

Rows returned (113)

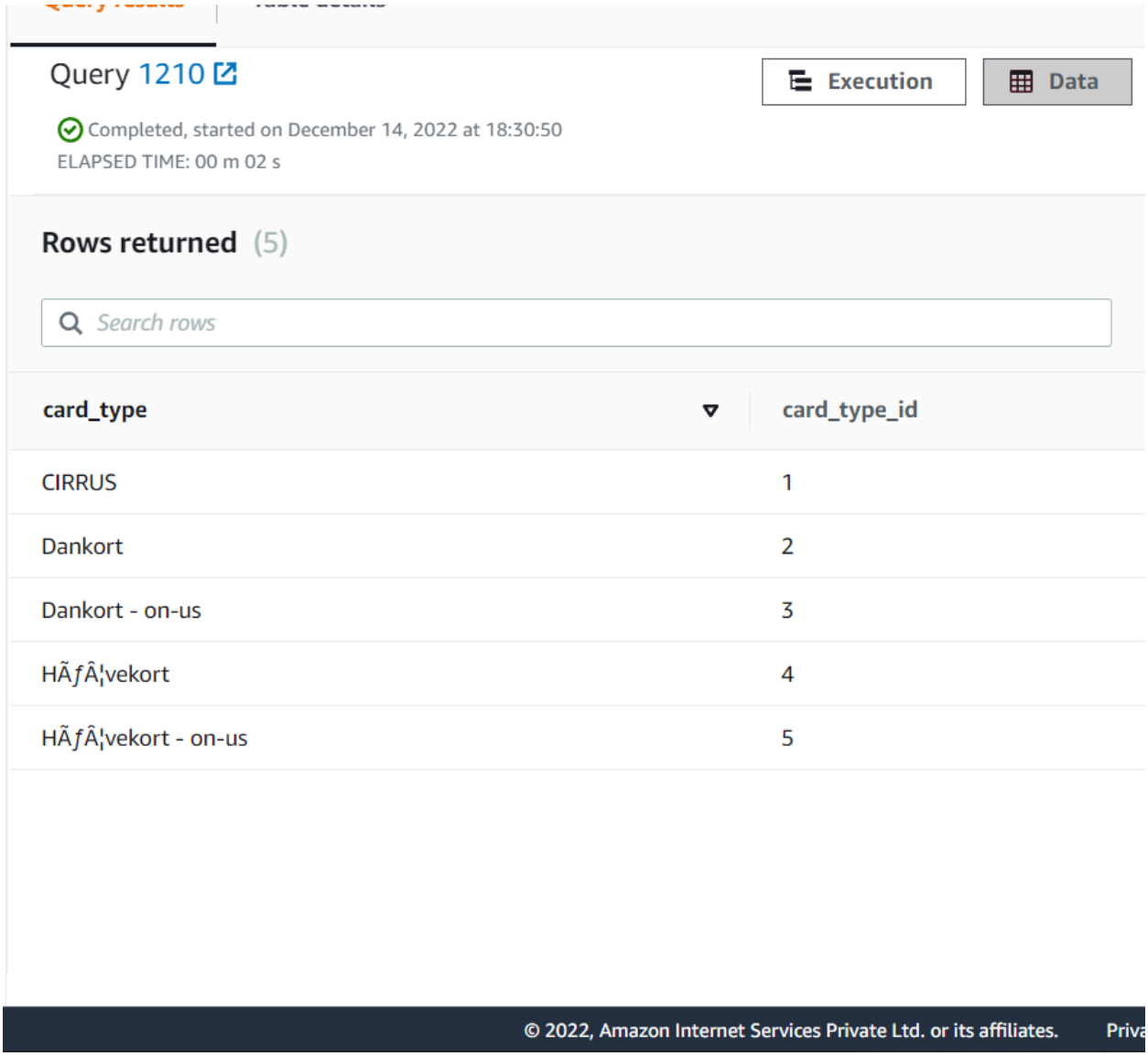
Export

Search rows

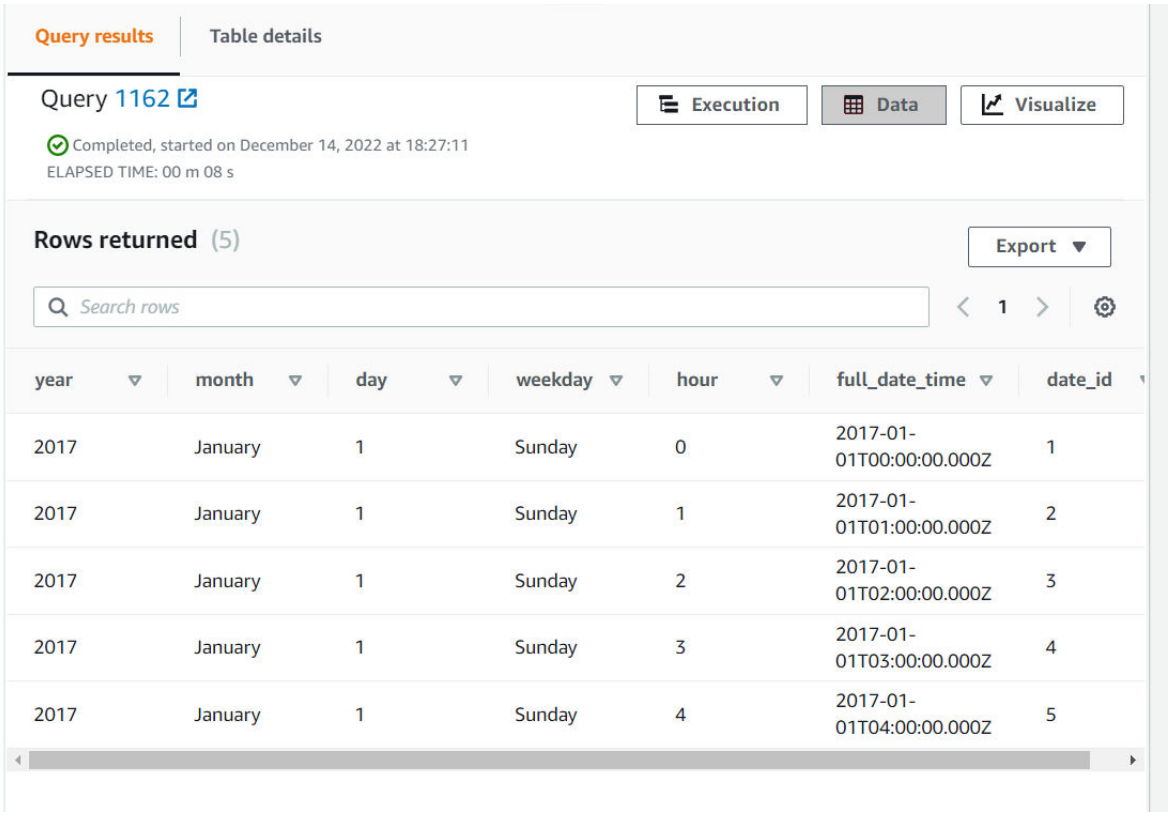
1 2 3 4 5 6 7 ... 12

atm_number	atm_manufacturer	atm_location	atm_id
1	NCR	74	1
10	NCR	76	2
100	NCR	56	3
101	NCR	17	4
102	NCR	3	5
103	Diebold Nixdorf	103	6
104	NCR	58	7
105	Diebold Nixdorf	76	8
106	NCR	55	9
107	Diebold Nixdorf	62	10

```
copy etlproject.dim_card_type from 's3://etlprojectnew/data/dim_card_type/part-00000-709454aa-f42c-41ec-b8f6-705ff18a208e-c000.csv' iam_role 'arn:aws:iam::274560308997:role/myRedshiftRole' delimiter ',' region 'us-east-1' CSV IGNOREHEADER as 1
;
```

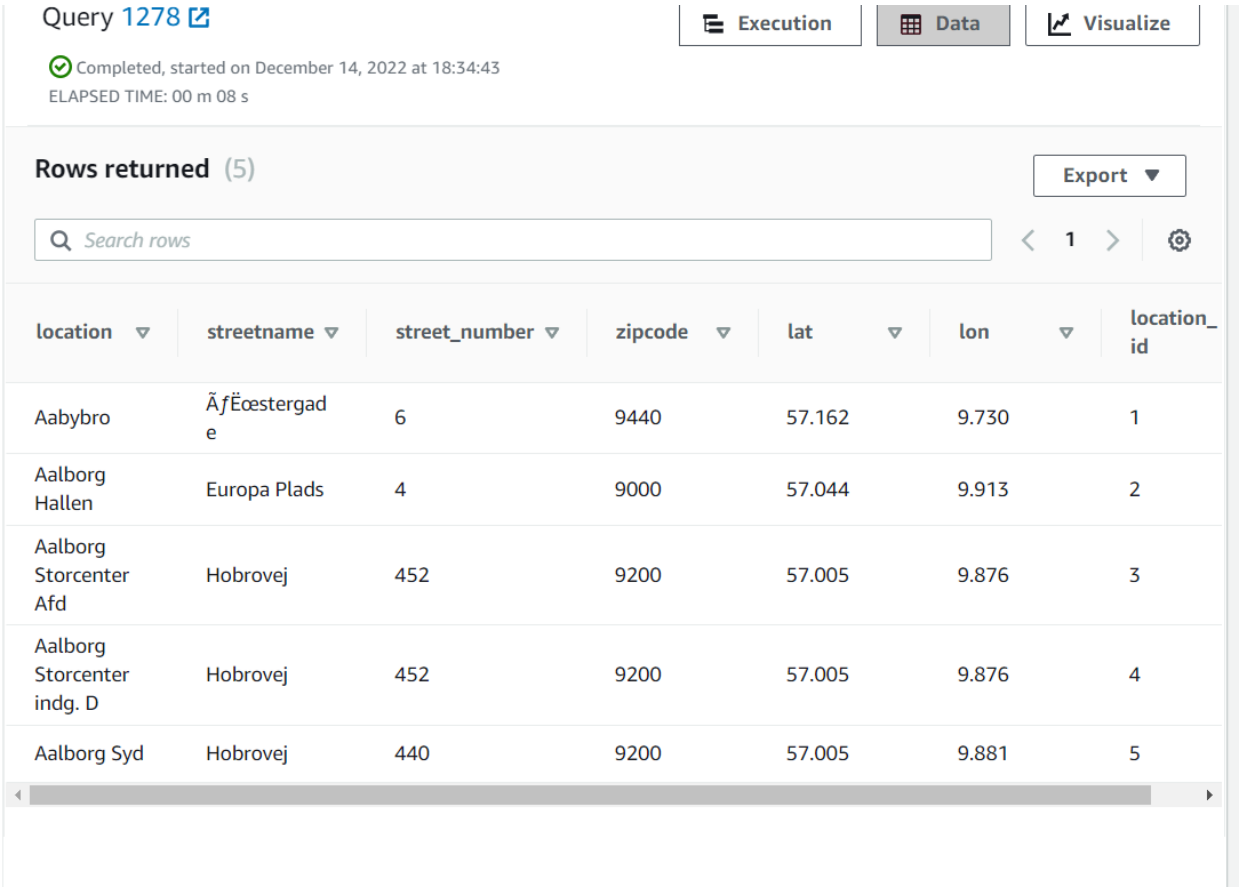


```
copy etlproject.dim_date from 's3://etlprojectnew/data/dim_date/part-00000-20bba22e-cd33-41d2-9c42-cb90c656e149-c000.csv' iam_role 'arn:aws:iam::274560308997:role/myRedshiftRole' delimiter ',' region 'us-east-1' CSV IGNOREHEADER as 1
;
```



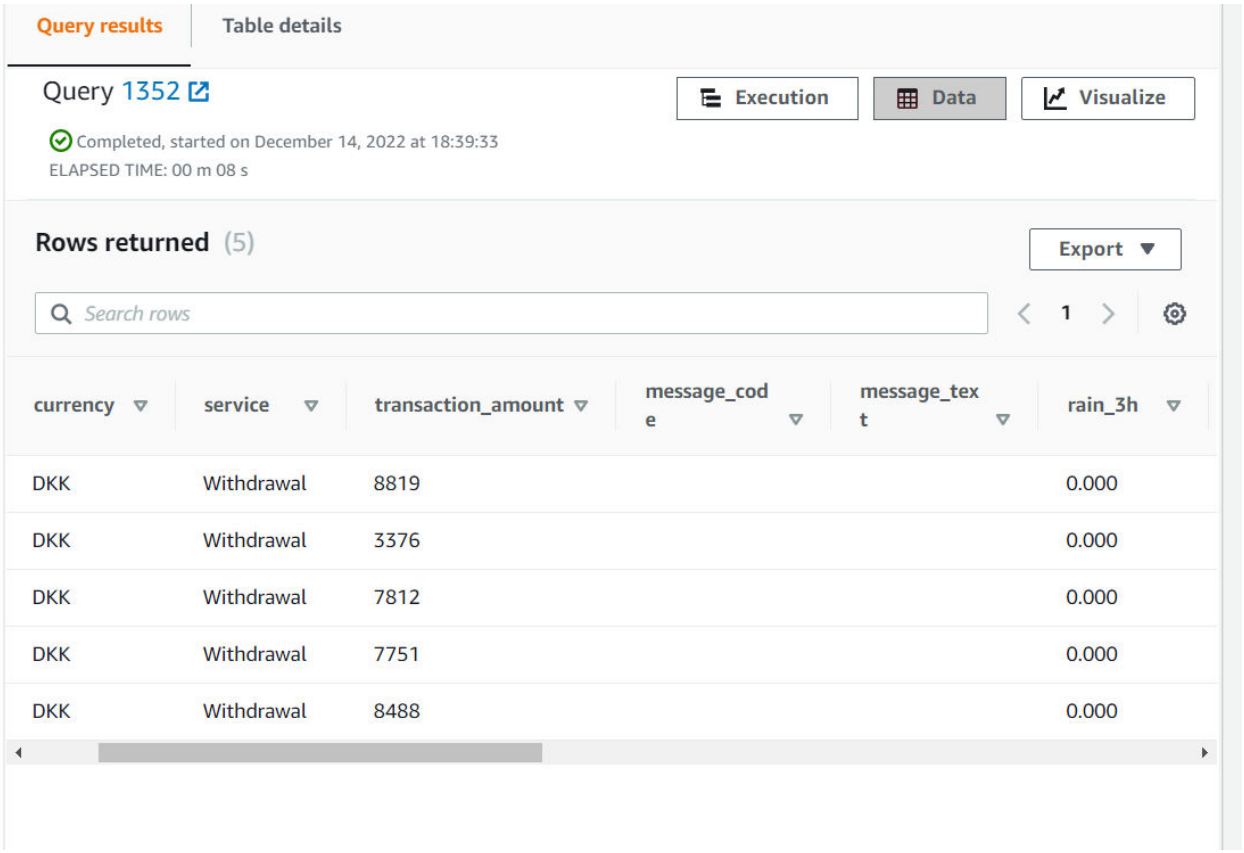
year	month	day	weekday	hour	full_date_time	date_id
2017	January	1	Sunday	0	2017-01-01T00:00:00.000Z	1
2017	January	1	Sunday	1	2017-01-01T01:00:00.000Z	2
2017	January	1	Sunday	2	2017-01-01T02:00:00.000Z	3
2017	January	1	Sunday	3	2017-01-01T03:00:00.000Z	4
2017	January	1	Sunday	4	2017-01-01T04:00:00.000Z	5

copy etlproject.dim_location from 's3://etlprojectnew/data/dim_location/part-00000-26c1b703-2b27-424b-ad58-0dc814f08346-c000.csv' iam_role 'arn:aws:iam::274560308997:role/myRedshiftRole' delimiter ',' region 'us-east-1' CSV IGNOREHEADER as 1
;



location	streetname	street_number	zipcode	lat	lon	location_id
Aabybro	ÃƒËœstergade	6	9440	57.162	9.730	1
Aalborg Hallen	Europa Plads	4	9000	57.044	9.913	2
Aalborg Storcenter Afd	Hobrovej	452	9200	57.005	9.876	3
Aalborg Storcenter indg. D	Hobrovej	452	9200	57.005	9.876	4
Aalborg Syd	Hobrovej	440	9200	57.005	9.881	5

copy etlproject.fact_atm_trans from 's3://etlprojectnew/data/fact_atm_trans/part-00000-c4a91d36-6042-45c5-b144-81cef9fafff1-c000.csv' iam_role 'arn:aws:iam::274560308997:role/myRedshiftRole' delimiter ',' region 'us-east-1' CSV IGNOREHEADER as 1
;



Query results | Table details

Query 1352

Execution | Data | Visualize

Completed, started on December 14, 2022 at 18:39:33
ELAPSED TIME: 00 m 08 s

Rows returned (5)

Export

Search rows

1

currency	service	transaction_amount	message_code	message_text	rain_3h
DKK	Withdrawal	8819			0.000
DKK	Withdrawal	3376			0.000
DKK	Withdrawal	7812			0.000
DKK	Withdrawal	7751			0.000
DKK	Withdrawal	8488			0.000