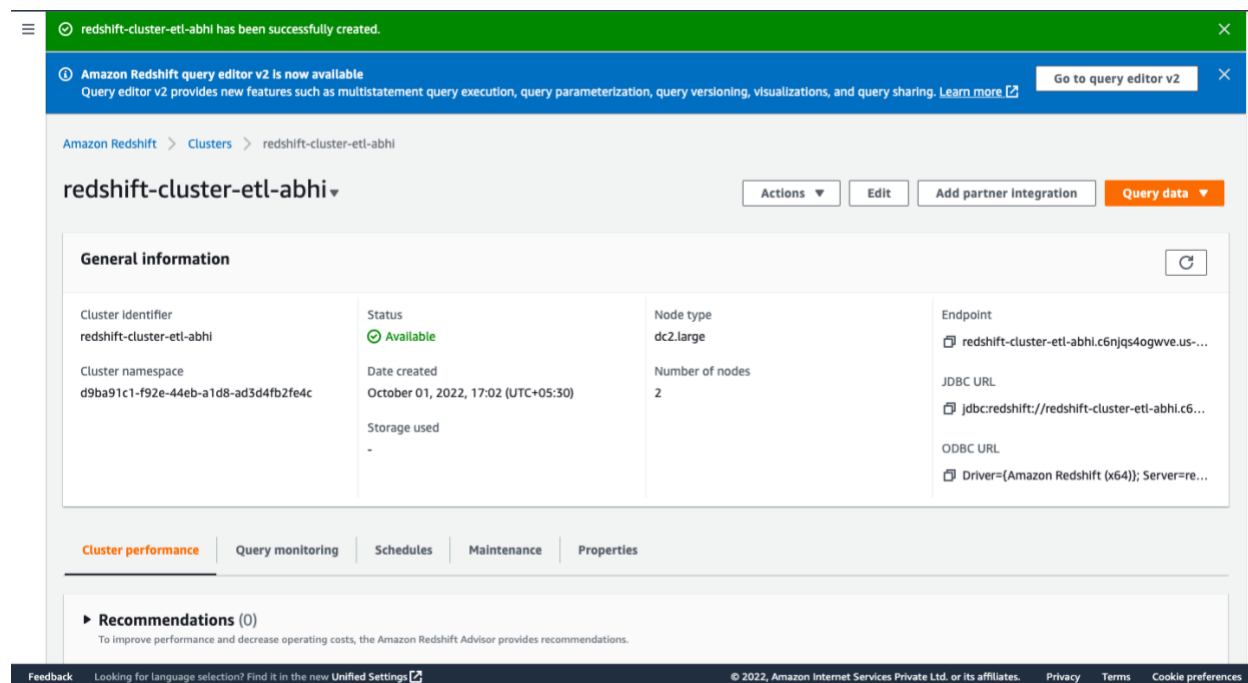


Creation of a Redshift Cluster

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



redshift-cluster-etl-abhi has been successfully created.

Amazon Redshift query editor v2 is now available
Query editor v2 provides new features such as multistatement query execution, query parameterization, query versioning, visualizations, and query sharing. [Learn more](#) [Go to query editor v2](#)

Amazon Redshift > Clusters > redshift-cluster-etl-abhi

redshift-cluster-etl-abhi

Actions Edit Add partner integration Query data

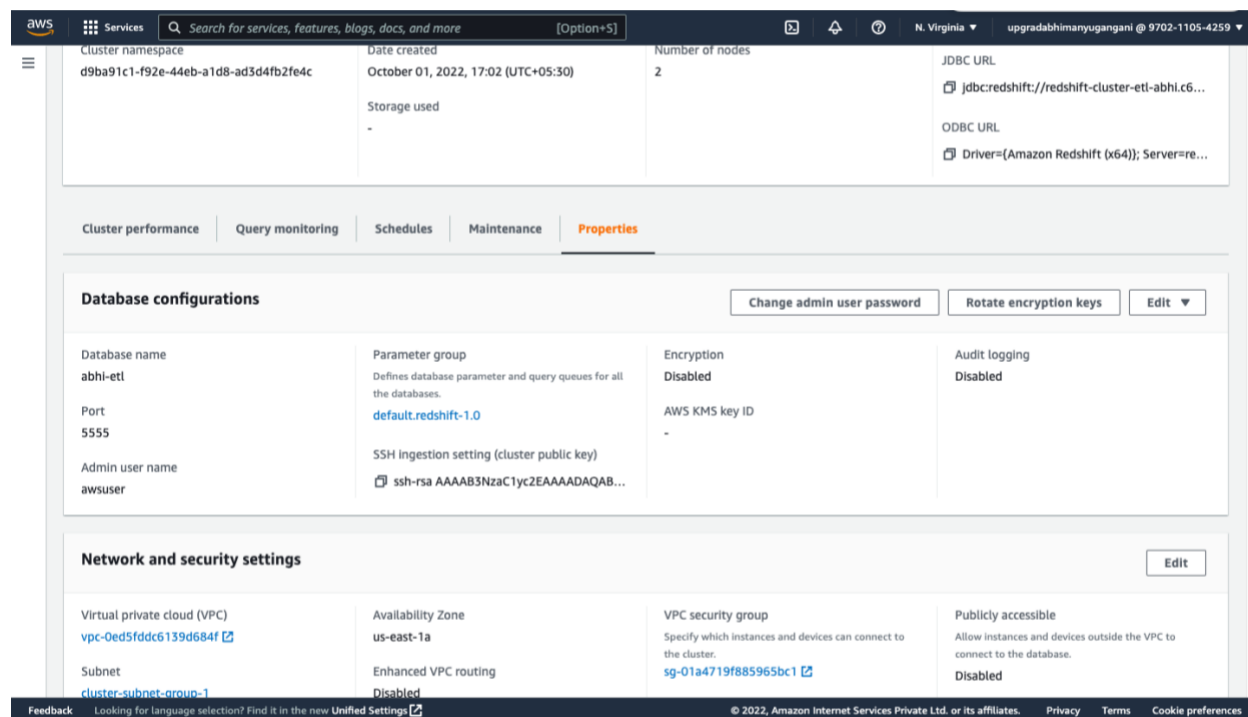
General information

Cluster identifier redshift-cluster-etl-abhi	Status Available	Node type dc2.large	Endpoint redshift-cluster-etl-abhi.c6njq54ogwve.us...
Cluster namespace d9ba91c1-f92e-44eb-a1d8-ad3d4fb2fe4c	Date created October 01, 2022, 17:02 (UTC+05:30)	Number of nodes 2	JDBC URL jdbc:redshift://redshift-cluster-etl-abhi.c6...
	Storage used -		ODBC URL Driver={Amazon Redshift (x64)}; Server=re...

Cluster performance Query monitoring Schedules Maintenance Properties

Recommendations (0)
To improve performance and decrease operating costs, the Amazon Redshift Advisor provides recommendations.

Feedback Looking for language selection? Find it in the new Unified Settings © 2022, Amazon Internet Services Private Ltd. or its affiliates. Privacy Terms Cookie preferences



aws Services Search for services, features, blogs, docs, and more [Option+S] N. Virginia upgradabhimanyugangani @ 9702-1105-4259

Cluster namespace
d9ba91c1-f92e-44eb-a1d8-ad3d4fb2fe4c

Date created
October 01, 2022, 17:02 (UTC+05:30)

Number of nodes
2

JDBC URL
jdbc:redshift://redshift-cluster-etl-abhi.c6...

Storage used
-

ODBC URL
Driver={Amazon Redshift (x64)}; Server=re...

Cluster performance Query monitoring Schedules Maintenance Properties

Database configurations

Change admin user password Rotate encryption keys Edit

Database name abhi-etl	Parameter group Defines database parameter and query queues for all the databases. default.redshift-1.0	Encryption Disabled	Audit logging Disabled
Port 5555	SSH ingestion setting (cluster public key) ssh-rsa AAAAB3NzaC1yc2EAAAADAQAB...	AWS KMS key ID -	
Admin user name awsuser			

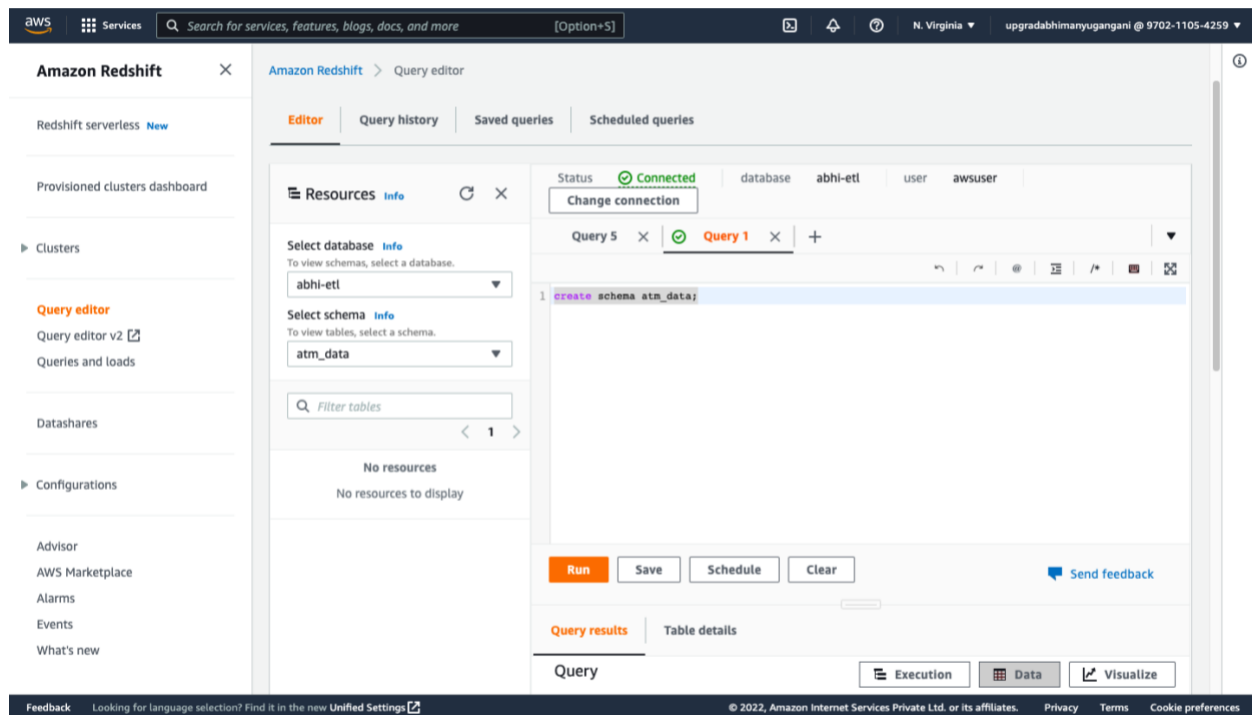
Network and security settings

Edit

Virtual private cloud (VPC) vpc-0ed5fddc6139d684f	Availability Zone us-east-1a	VPC security group Specify which instances and devices can connect to the cluster. sg-01a4719f885965bc1	Publicly accessible Allow instances and devices outside the VPC to connect to the database. Disabled
Subnet subnet-subnet-group-1	Enhanced VPC routing Disabled		

Feedback Looking for language selection? Find it in the new Unified Settings © 2022, Amazon Internet Services Private Ltd. or its affiliates. Privacy Terms Cookie preferences

Creating Schema atm_data



The screenshot displays the Amazon Redshift Query Editor interface. The left sidebar shows the navigation menu with options like Redshift serverless, Provisioned clusters dashboard, Clusters, Query editor (selected), Query editor v2, Queries and loads, Datashares, Configurations, Advisor, AWS Marketplace, Alarms, Events, and What's new.

The main panel is titled "Amazon Redshift > Query editor" and has tabs for Editor, Query history, Saved queries, and Scheduled queries. The "Editor" tab is active, showing a "Resources" panel on the left with "Select database" (abhi-etl) and "Select schema" (atm_data) dropdowns. Below these is a "Filter tables" search bar and a message "No resources to display".

The central area shows the query editor with a status bar indicating "Connected" and a "Change connection" button. The query editor has tabs for "Query 5" and "Query 1" (selected). The query text in the editor is:


```
1 create schema atm_data;
```

 Below the query editor are buttons for "Run", "Save", "Schedule", and "Clear". A "Send feedback" link is also present.

At the bottom, there are tabs for "Query results" and "Table details". The "Query" tab is selected, showing options for "Execution", "Data", and "Visualize".

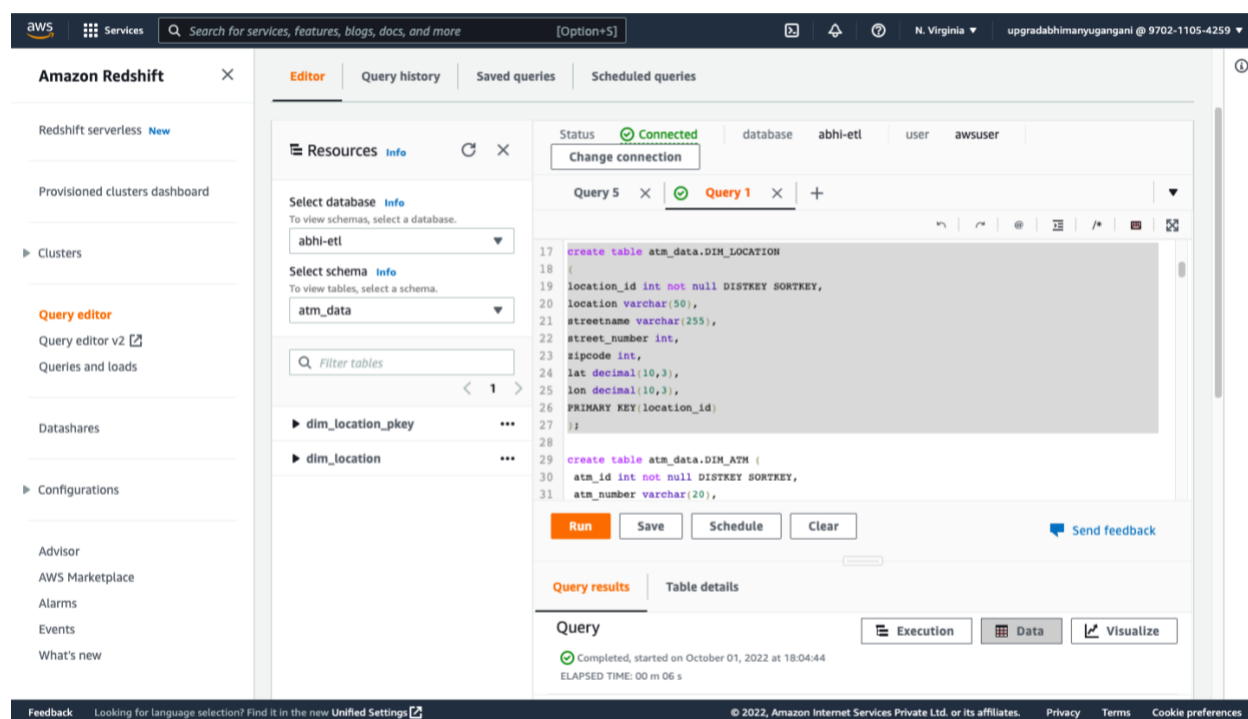
The footer of the interface includes a "Feedback" link, a language selection prompt, a copyright notice for 2022 Amazon Internet Services Private Ltd., and links for Privacy, Terms, and Cookie preferences.

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:

- Creating location 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) );
```



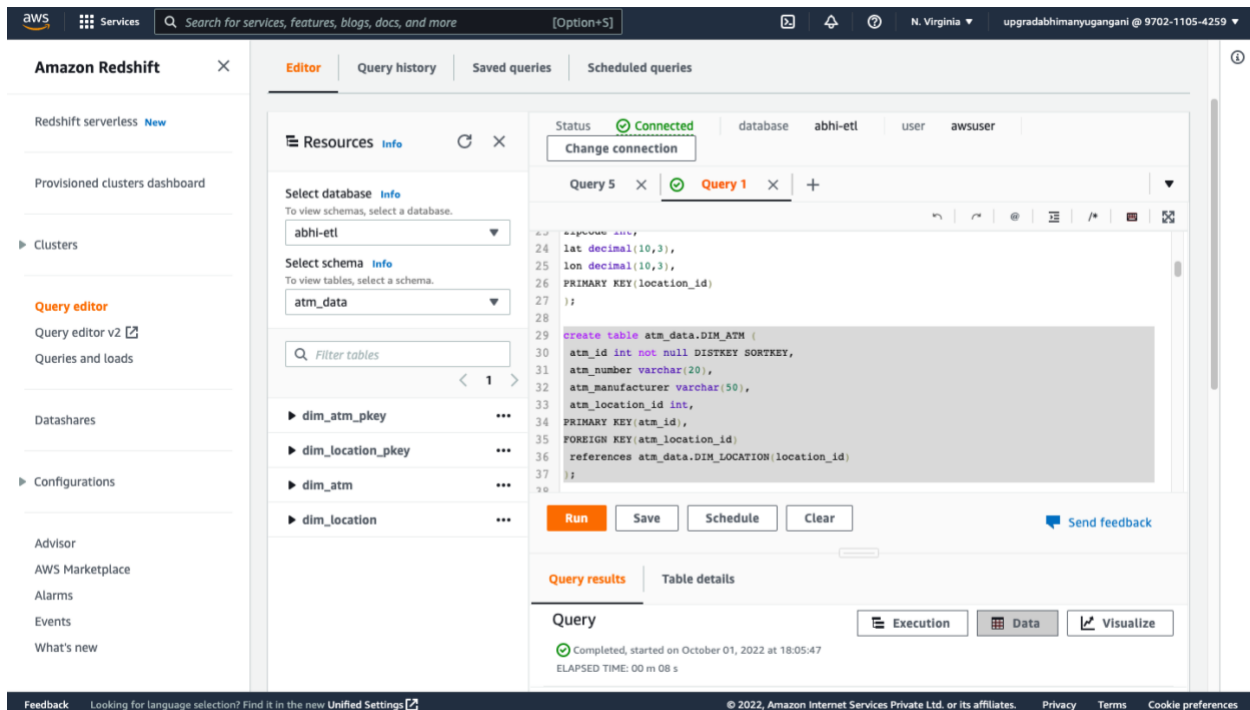
The screenshot shows the Amazon Redshift Query Editor interface. The left sidebar contains navigation options like 'Redshift serverless', 'Provisioned clusters dashboard', 'Clusters', 'Query editor', 'Databases', 'Configurations', 'Advisor', 'AWS Marketplace', 'Alarms', 'Events', and 'What's new'. The main editor area is titled 'Query 1' and shows the SQL code for creating the 'DIM_LOCATION' table. The code is as follows:

```
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) );
```

Below the code, there are buttons for 'Run', 'Save', 'Schedule', and 'Clear'. The 'Query results' section at the bottom shows the query status as 'Completed' and the elapsed time as '00 m 06 s'.

- 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)
);
```



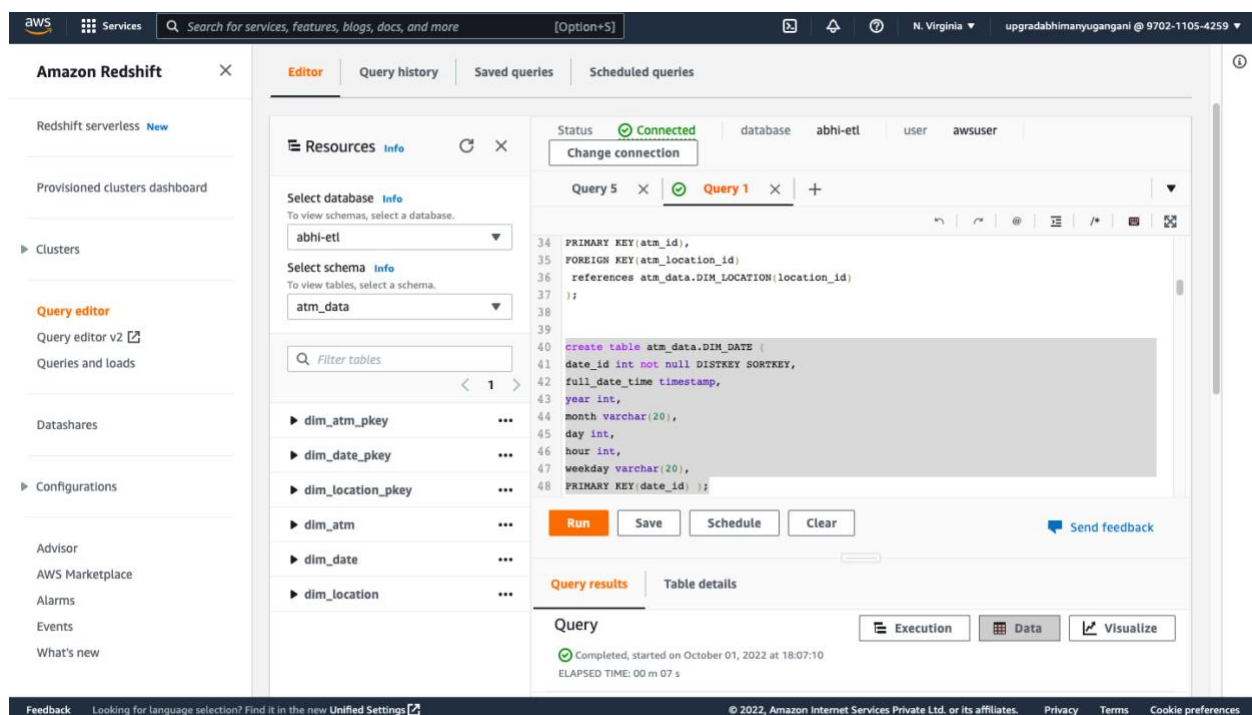
The screenshot displays the Amazon Redshift Query Editor interface. The left sidebar shows the navigation menu with options like 'Query editor', 'Query editor v2', and 'Queries and loads'. The main editor area is titled 'Query 1' and shows the SQL code for creating the 'DIM_ATM' table. The code is as follows:

```
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)
);
```

Below the code editor, there are buttons for 'Run', 'Save', 'Schedule', and 'Clear'. The 'Query results' section at the bottom shows the query execution status: 'Completed, started on October 01, 2022 at 18:05:47' with an 'ELAPSED TIME: 00 m 08 s'.

- Creating date dimension table

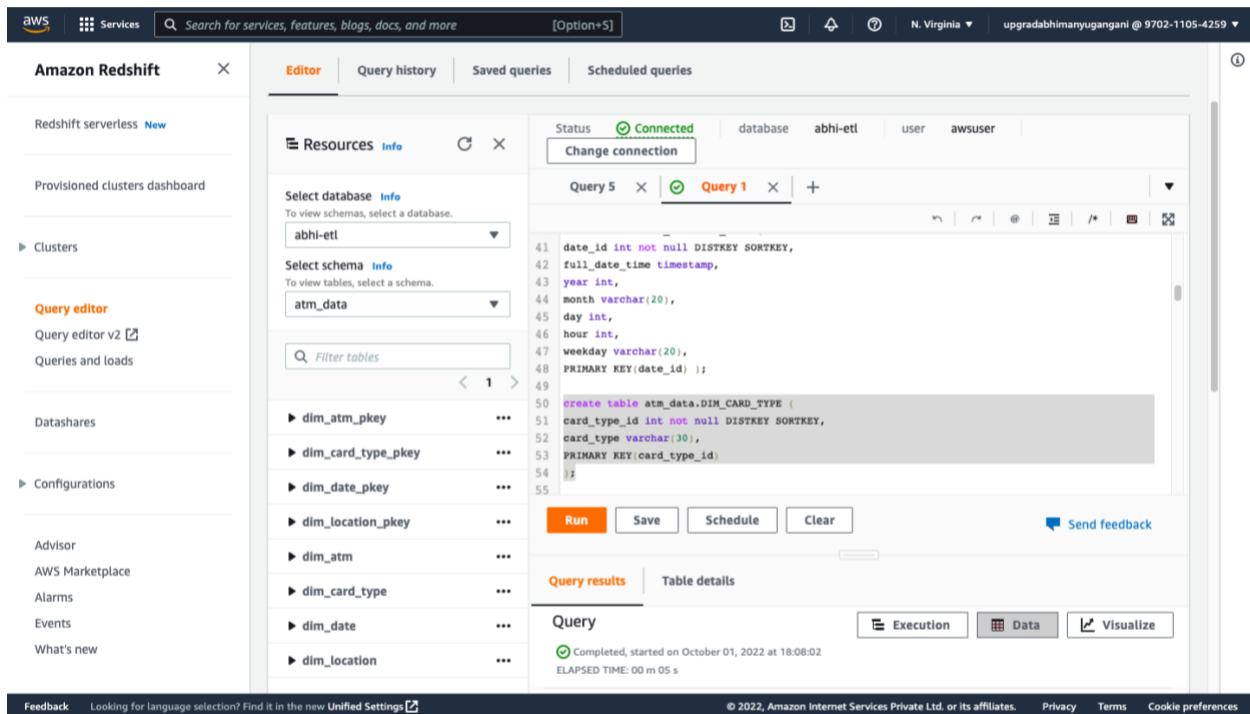
```
create table atm_data.DIM_DATE (
date_id int not null DISTKEY SORTKEY,
full_date_time timestamp,
year int,
month varchar(20),
day int,
hour int,
weekday varchar(20),
PRIMARY KEY(date_id) );
```



The screenshot shows the Amazon Redshift Query Editor interface. The left sidebar contains navigation options like 'Redshift serverless', 'Provisioned clusters dashboard', 'Clusters', 'Query editor', 'Datashares', 'Configurations', 'Advisor', 'AWS Marketplace', 'Alarms', 'Events', and 'What's new'. The main editor area is titled 'Editor' and shows a 'Resources' panel on the left with a tree view of databases and schemas. The 'Select database' dropdown is set to 'abhi-etl' and the 'Select schema' dropdown is set to 'atm_data'. The 'Query editor' tab is active, displaying a SQL query that creates a table named 'DIM_DATE' with columns: 'date_id' (int, not null, DISTKEY SORTKEY), 'full_date_time' (timestamp), 'year' (int), 'month' (varchar(20)), 'day' (int), 'hour' (int), 'weekday' (varchar(20)), and a primary key on 'date_id'. The query is highlighted in a light blue box. Below the query editor, there are buttons for 'Run', 'Save', 'Schedule', and 'Clear'. The 'Query results' panel at the bottom shows the query status as 'Completed, started on October 01, 2022 at 18:07:10' with an elapsed time of '00 m 07 s'.

- 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)
);
```



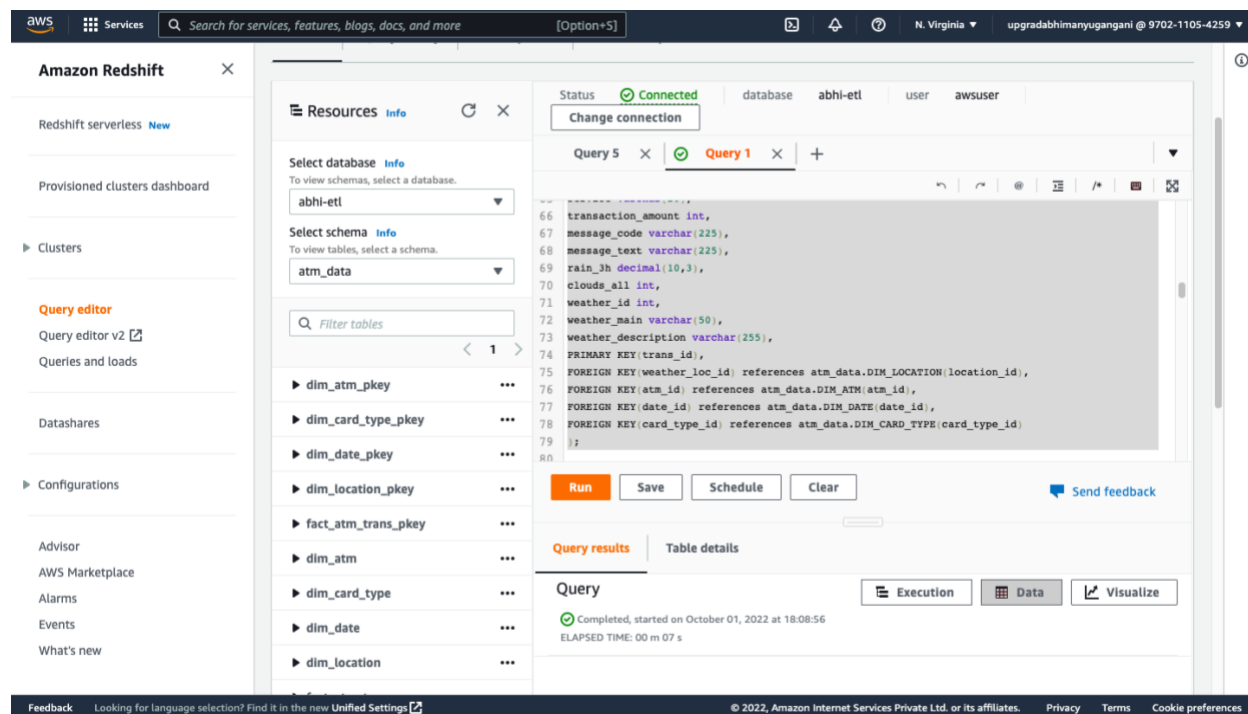
The screenshot shows the Amazon Redshift Query Editor interface. The left sidebar contains navigation options like 'Redshift serverless', 'Provisioned clusters dashboard', 'Clusters', 'Query editor', 'Datashares', 'Configurations', 'Advisor', 'AWS Marketplace', 'Alarms', 'Events', and 'What's new'. The main editor area is titled 'Amazon Redshift' and includes tabs for 'Editor', 'Query history', 'Saved queries', and 'Scheduled queries'. The 'Editor' tab is active, showing a SQL query editor with a 'Resources' panel on the left and a 'Query' panel on the right. The 'Resources' panel shows the selected database 'abhi-etl' and schema 'atm_data'. The 'Query' panel displays the SQL code for creating the 'DIM_CARD_TYPE' table. The query is as follows:

```
41 date_id int not null DISTKEY SORTKEY,
42 full_date_time timestamp,
43 year int,
44 month varchar(20),
45 day int,
46 hour int,
47 weekday varchar(20);
48 PRIMARY KEY(date_id) );
49
50 create table atm_data.DIM_CARD_TYPE (
51 card_type_id int not null DISTKEY SORTKEY,
52 card_type varchar(30),
53 PRIMARY KEY(card_type_id)
54 );
55
```

Below the query editor, there are buttons for 'Run', 'Save', 'Schedule', and 'Clear'. The 'Run' button is highlighted. The 'Query results' section shows the query status as 'Completed, started on October 01, 2022 at 18:08:02' with an 'ELAPSED TIME: 00 m 05 s'.

- 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)
);
```



The screenshot shows the Amazon Redshift Query Editor interface. The left sidebar contains navigation options like 'Resources', 'Query editor', 'Databases', and 'Configurations'. The main area displays the 'Query editor' with a 'Select database' dropdown set to 'abhi-etl' and a 'Select schema' dropdown set to 'atm_data'. Below these, a list of tables is shown, including 'dim_atm_pkey', 'dim_card_type_pkey', 'dim_date_pkey', 'dim_location_pkey', 'fact_atm_trans_pkey', 'dim_atm', 'dim_card_type', 'dim_date', and 'dim_location'. The 'Query editor' tab is active, showing the SQL code for creating the 'FACT_ATM_TRANS' table. The code is as follows:

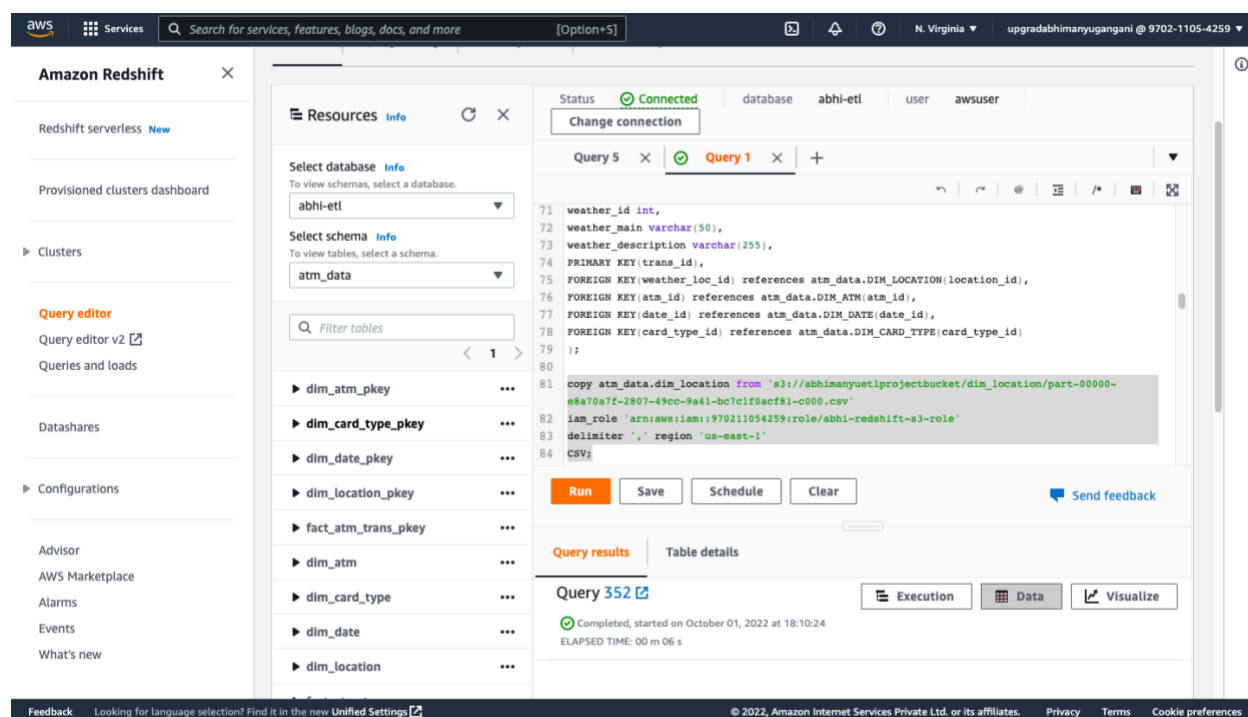
```
66 transaction_amount int,
67 message_code varchar(225),
68 message_text varchar(225),
69 rain_3h decimal(10,3),
70 clouds_all int,
71 weather_id int,
72 weather_main varchar(50),
73 weather_description varchar(255),
74 PRIMARY KEY(trans_id),
75 FOREIGN KEY(weather_loc_id) references atm_data.DIM_LOCATION(location_id),
76 FOREIGN KEY(atm_id) references atm_data.DIM_ATM(atm_id),
77 FOREIGN KEY(date_id) references atm_data.DIM_DATE(date_id),
78 FOREIGN KEY(card_type_id) references atm_data.DIM_CARD_TYPE(card_type_id)
79 );
```

Below the SQL code, there are buttons for 'Run', 'Save', 'Schedule', and 'Clear'. The 'Query results' tab is selected, showing a message: 'Query Completed, started on October 01, 2022 at 18:08:56 ELAPSED TIME: 00 m 07 s'.

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 atm_data.dim_location from 's3://abhimanyuetlprojectbucket/dim_location/part-00000-e8a70a7f-2807-49cc-9a41-bc7c1f0acf81-c000.csv'
iam_role 'arn:aws:iam::970211054259:role/abhi-redshift-s3-role'
delimiter ',' region 'us-east-1'
CSV;



The screenshot shows the Amazon Redshift Query Editor interface. On the left, there's a sidebar with navigation options like 'Resources', 'Select database', 'Select schema', and 'Filter tables'. The main area displays a SQL query for copying data from an S3 bucket into a Redshift table. The query is as follows:

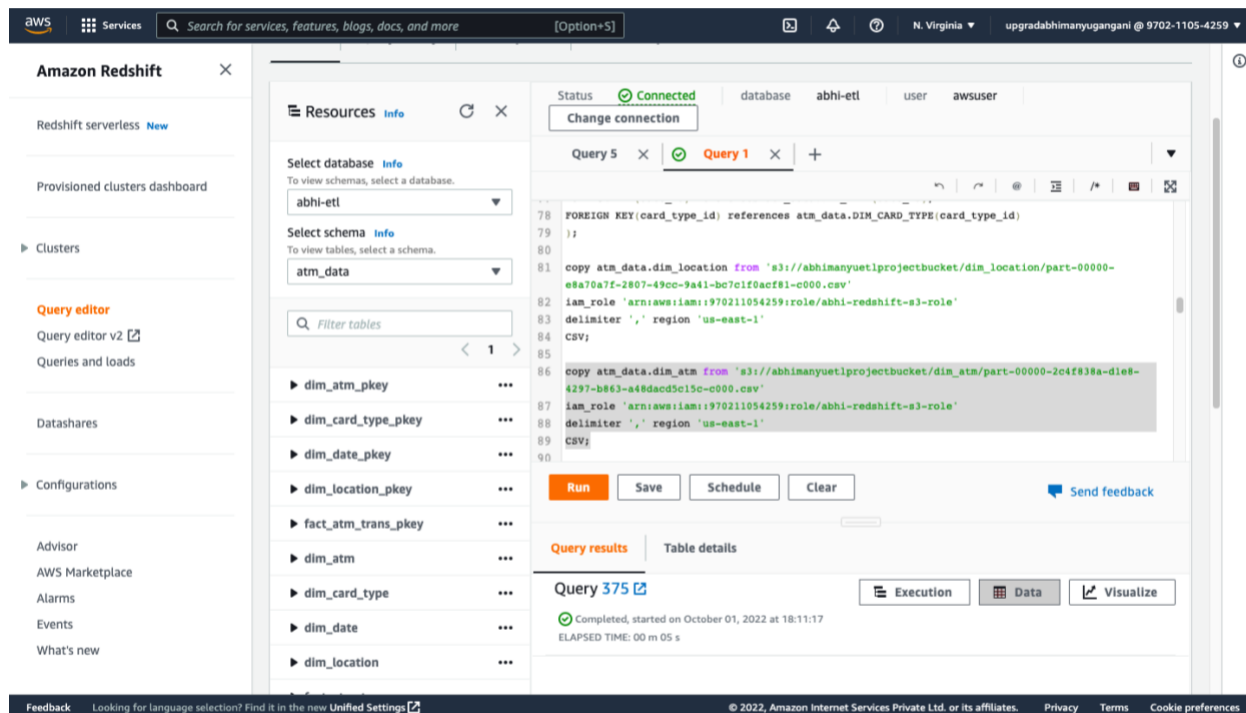
```

71 weather_id int,
72 weather_main varchar(50),
73 weather_description varchar(255),
74 PRIMARY KEY(trans_id),
75 FOREIGN KEY(weather_loc_id) references atm_data.DIM_LOCATION(location_id),
76 FOREIGN KEY(atm_id) references atm_data.DIM_ATM(atm_id),
77 FOREIGN KEY(date_id) references atm_data.DIM_DATE(date_id),
78 FOREIGN KEY(card_type_id) references atm_data.DIM_CARD_TYPE(card_type_id)
79 ;;
80
81 copy atm_data.dim_location from 's3://abhimanyuetlprojectbucket/dim_location/part-00000-
82 e8a70a7f-2807-49cc-9a41-bc7c1f0acf81-c000.csv'
83 iam_role 'arn:aws:iam::970211054259:role/abhi-redshift-s3-role'
84 delimiter ',' region 'us-east-1'
85 CSV;

```

Below the query, there are buttons for 'Run', 'Save', 'Schedule', and 'Clear'. The 'Run' button is highlighted. The 'Query results' tab is selected, showing the query execution status: 'Completed, started on October 01, 2022 at 18:10:24' and 'ELAPSED TIME: 00 m 06 s'.

copy atm_data.dim_atm from 's3://abhimanyuetlprojectbucket/dim_atm/part-00000-2c4f838a-d1e8-4297-b863-a48dacd5c15c-c000.csv'
iam_role 'arn:aws:iam::970211054259:role/abhi-redshift-s3-role'
delimiter ',' region 'us-east-1'
CSV;



The screenshot shows the Amazon Redshift Query Editor interface. The left sidebar contains navigation options like 'Redshift serverless', 'Provisioned clusters dashboard', 'Clusters', 'Query editor', 'Datashares', 'Configurations', 'Advisor', 'AWS Marketplace', 'Alarms', 'Events', and 'What's new'. The main area is divided into three sections: 'Resources', 'Query editor', and 'Query results'. The 'Resources' section shows the selected database 'abhi-etl' and schema 'atm_data'. The 'Query editor' section displays a SQL query that copies data from an S3 bucket into the 'dim_atm' table. The 'Query results' section shows the query status as 'Completed' and the elapsed time as '00 m 05 s'.

Query 375

```

78 FOREIGN KEY (card_type_id) references atm_data.DIM_CARD_TYPE (card_type_id)
79 ;;
80
81 copy atm_data.dim_location from 's3://abhimanyuetlprojectbucket/dim_location/part-00000-
82 e8a70a7f-2807-49cc-9a41-bc7c1f0acf81-c000.csv'
83 iam_role 'arn:aws:iam::970211054259:role/abhi-redshift-s3-role'
84 delimiter ',' region 'us-east-1'
85 CSV;
86
87 copy atm_data.dim_atm from 's3://abhimanyuetlprojectbucket/dim_atm/part-00000-2c4f838a-d1e8-
88 4297-b863-a48dacd5c15c-c000.csv'
89 iam_role 'arn:aws:iam::970211054259:role/abhi-redshift-s3-role'
90 delimiter ',' region 'us-east-1'
91 CSV;

```

Query results

Query 375

Completed, started on October 01, 2022 at 18:11:17
ELAPSED TIME: 00 m 05 s

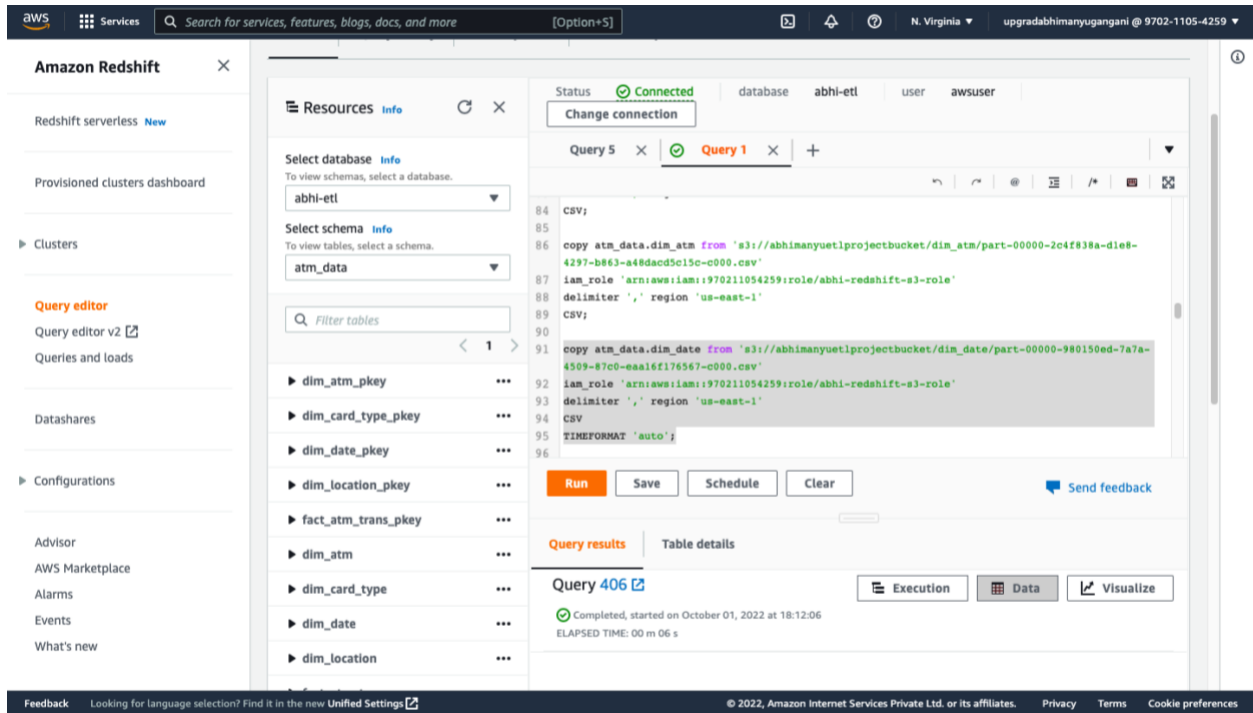
copy atm_data.dim_date from 's3://abhimanyuetlprojectbucket/dim_date/part-00000-980150ed-7a7a-4509-87c0-eaa16f176567-c000.csv'

iam_role 'arn:aws:iam::970211054259:role/abhi-redshift-s3-role'

delimiter ',' region 'us-east-1'

CSV

TIMEFORMAT 'auto';



The screenshot shows the Amazon Redshift Query Editor interface. The left sidebar contains navigation options: Amazon Redshift, Redshift serverless, Provisioned clusters dashboard, Clusters, Query editor (selected), Databases, Configurations, Advisor, AWS Marketplace, Alarms, Events, and What's new. The main panel displays the 'Query editor' with a 'Select database' dropdown set to 'abhi-etl' and a 'Select schema' dropdown set to 'atm_data'. Below these are search filters and a list of tables: dim_atm_pkey, dim_card_type_pkey, dim_date_pkey, dim_location_pkey, fact_atm_trans_pkey, dim_atm, dim_card_type, dim_date, and dim_location. The SQL query editor shows the following code:

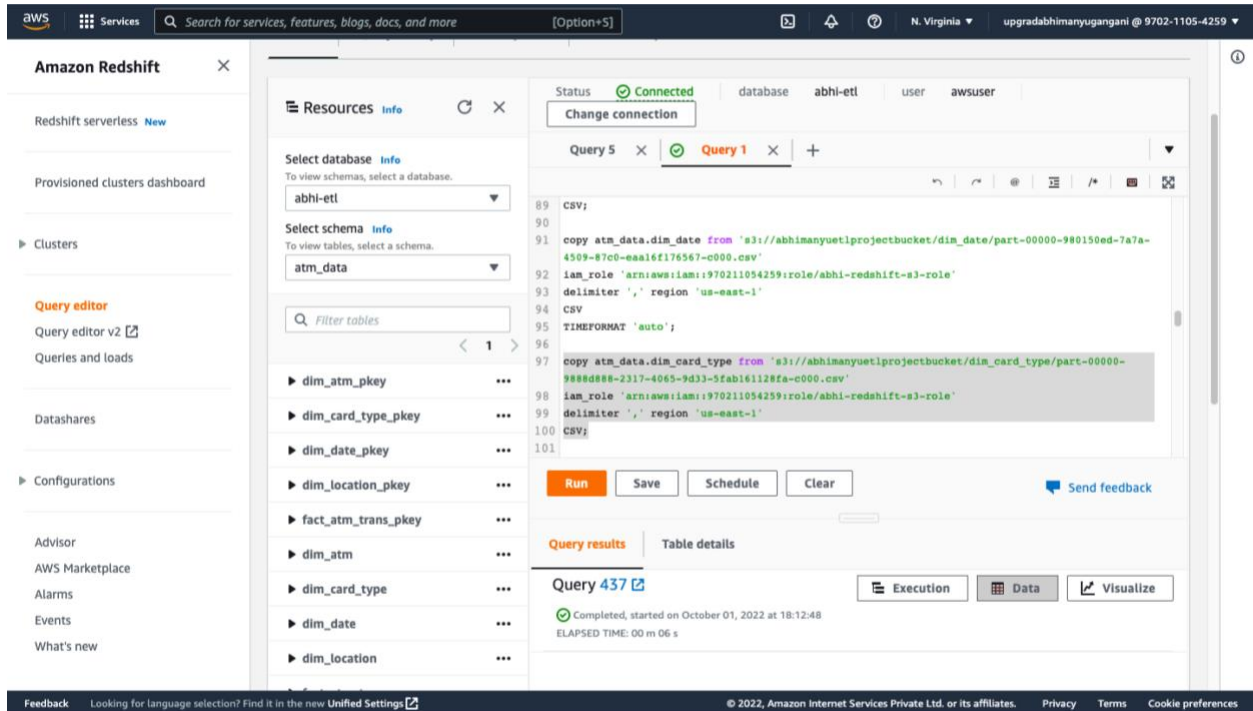
```

84 CSV;
85
86 copy atm_data.dim_atm from 's3://abhimanyuetlprojectbucket/dim_atm/part-00000-2c4f838a-d1e8-
87 4297-b863-a48dadc5c15c-c000.csv'
88 iam_role 'arn:aws:iam::970211054259:role/abhi-redshift-s3-role'
89 delimiter ',' region 'us-east-1'
90 CSV;
91
92 copy atm_data.dim_date from 's3://abhimanyuetlprojectbucket/dim_date/part-00000-980150ed-7a7a-
93 4509-87c0-eaa16f176567-c000.csv'
94 iam_role 'arn:aws:iam::970211054259:role/abhi-redshift-s3-role'
95 delimiter ',' region 'us-east-1'
96 CSV
97 TIMEFORMAT 'auto';

```

The query is labeled 'Query 406' and has been executed successfully. The status bar at the bottom indicates 'Completed, started on October 01, 2022 at 18:12:06' with an 'ELAPSED TIME: 00 m 06 s'.

copy atm_data.dim_card_type from 's3://abhimanyuetlprojectbucket/dim_card_type/part-00000-9888d888-2317-4065-9d33-5fab161128fa-c000.csv'
iam_role 'arn:aws:iam::970211054259:role/abhi-redshift-s3-role'
delimiter ',' region 'us-east-1'
CSV;



The screenshot shows the Amazon Redshift Query Editor interface. The left sidebar contains navigation options like 'Redshift serverless', 'Provisioned clusters dashboard', 'Clusters', 'Query editor', 'Datashares', 'Configurations', 'Advisor', 'AWS Marketplace', 'Alarms', 'Events', and 'What's new'. The main area is divided into three sections: 'Resources', 'Query editor', and 'Query results'.

In the 'Resources' section, the 'Select database' dropdown is set to 'abhi-etl' and the 'Select schema' dropdown is set to 'atm_data'. Below these, a list of tables is shown, including 'dim_atm_pkey', 'dim_card_type_pkey', 'dim_date_pkey', 'dim_location_pkey', 'fact_atm_trans_pkey', 'dim_atm', 'dim_card_type', 'dim_date', and 'dim_location'.

The 'Query editor' section displays a SQL query (Query 1) with the following content:

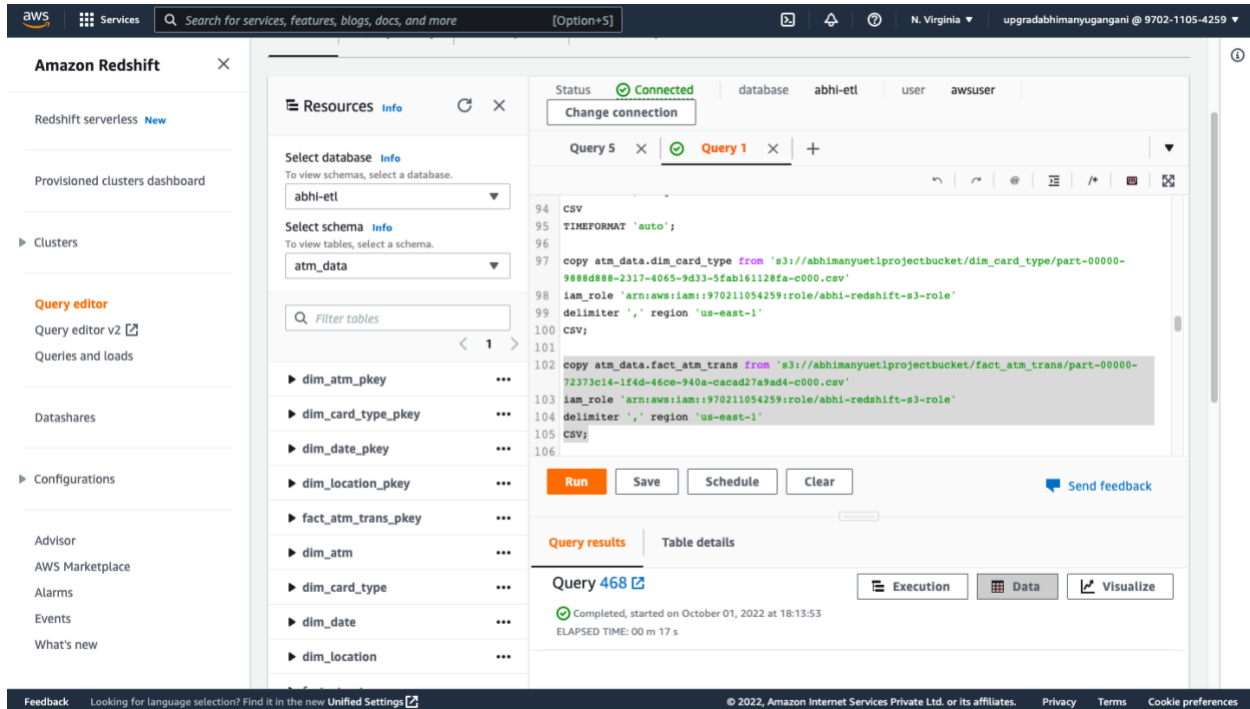
```

89 CSV;
90
91 copy atm_data.dim_date from 's3://abhimanyuetlprojectbucket/dim_date/part-00000-980150ed-7a7a-
4509-87c0-eaa16f176567-c000.csv'
92 iam_role 'arn:aws:iam::970211054259:role/abhi-redshift-s3-role'
93 delimiter ',' region 'us-east-1'
94 CSV
95 TIMEFORMAT 'auto';
96
97 copy atm_data.dim_card_type from 's3://abhimanyuetlprojectbucket/dim_card_type/part-00000-
9888d888-2317-4065-9d33-5fab161128fa-c000.csv'
98 iam_role 'arn:aws:iam::970211054259:role/abhi-redshift-s3-role'
99 delimiter ',' region 'us-east-1'
100 CSV;
101

```

Below the query editor, there are buttons for 'Run', 'Save', 'Schedule', and 'Clear'. The 'Query results' section shows the status of the query (Query 437) as 'Completed, started on October 01, 2022 at 18:12:48' with an 'ELAPSED TIME: 00 m 06 s'. There are also buttons for 'Execution', 'Data', and 'Visualize'.

```
copy atm_data.fact_atm_trans from 's3://abhimanyuetlprojectbucket/fact_atm_trans/part-00000-72373c14-1f4d-46ce-940a-cacad27a9ad4-c000.csv'
iam_role 'arn:aws:iam::970211054259:role/abhi-redshift-s3-role'
delimiter ',' region 'us-east-1'
CSV;
```



The screenshot shows the Amazon Redshift Query Editor interface. The left sidebar contains navigation options like 'Redshift serverless', 'Provisioned clusters dashboard', 'Clusters', 'Query editor', 'Databases', 'Configurations', 'Advisor', 'AWS Marketplace', 'Alarms', 'Events', and 'What's new'. The main area is divided into 'Resources' and 'Query editor'. The 'Resources' panel shows the selected database 'abhi-etl' and schema 'atm_data'. The 'Query editor' panel displays a SQL query (Query 1) that copies data from an S3 bucket into a Redshift table. The query is as follows:

```
94 CSV
95 TIMEFORMAT 'auto';
96
97 copy atm_data.dim_card_type from 's3://abhimanyuetlprojectbucket/dim_card_type/part-00000-9888d888-2317-4065-9d33-5fab161128fa-c000.csv'
98 iam_role 'arn:aws:iam::970211054259:role/abhi-redshift-s3-role'
99 delimiter ',' region 'us-east-1'
100 CSV;
101
102 copy atm_data.fact_atm_trans from 's3://abhimanyuetlprojectbucket/fact_atm_trans/part-00000-72373c14-1f4d-46ce-940a-cacad27a9ad4-c000.csv'
103 iam_role 'arn:aws:iam::970211054259:role/abhi-redshift-s3-role'
104 delimiter ',' region 'us-east-1'
105 CSV;
106
```

Below the query editor, there are buttons for 'Run', 'Save', 'Schedule', and 'Clear'. The 'Query results' section shows that the query (Query 468) has been completed successfully on October 01, 2022, at 18:13:53, with an elapsed time of 00 m 17 s. The bottom of the interface includes a footer with 'Feedback', 'Looking for language selection? Find it in the new Unified Settings', and copyright information for Amazon Internet Services Private Ltd. or its affiliates.