Hands-on Lab: Setting up a staging area

Estimated time needed: 30 minutes

Objectives

In this lab you will:

- Setup a staging server for a data warehouse
- Create the schema to store the data
- Load the data into the tables
- Run a sample query

About Skills Network Cloud IDE

Skills Network Cloud IDE (based on Theia and Docker) provides an environment for hands on labs for course and project related labs. Theia is an open source IDE (Integrated Development Environment), that can be run on desktop or on the cloud. To complete this lab, we will be using the Cloud IDE based on Theia running in a Docker container.

Important Notice about this lab environment

Please be aware that sessions for this lab environment are not persistent. A new environment is created for you every time you connect to this lab. Any data you may have saved in an earlier session will get lost. To avoid losing your data, please plan to complete these labs in a single session.

Exercise 1 - Start the PostgreSQL server

We will be using the PostgreSQL server as our staging server.

Start the PostgreSQL server.

Open a new terminal, by clicking on the menu bar and selecting **Terminal**->**New Terminal**.

This will open a new terminal at the bottom of the screen.

Run the commands below on the newly opened terminal. (You can copy the code by clicking on the little copy button on the bottom right of the codeblock below and then paste it, wherever you wish.)

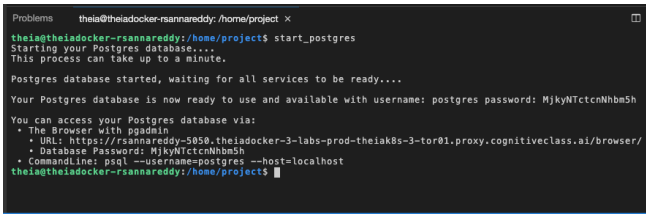
Start the PostgreSQL server, by running the command below:

```
1
```

```
start_postgres
```

Copied! Executed!

You should see an output similar to the one below.



Exercise 2 - Create Database

Create the database on the data warehouse.

Using the createdb command of the PostgreSQL server, we can directly create the database from the terminal.

Run the command below to create a database named billingDW.

```
1
```

```
createdb -h localhost -U postgres -p 5432 billingDW
```

Copied! Executed!

In the above command

- -h mentions that the database server is running on the localhost
- -U mentions that we are using the user name postgres to log into the database
- -p mentions that the database server is running on port number 5432

You should see an output like this.

Exercise 3 - Create data warehouse schema

Step 1:
Download the schema files.

The commands to create the schema are available in the file below.

https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DB0260EN-SkillsNetwork/labs/Setting%20up%20a%20staging%20area/billing-datawarehouse.tgz
Run the commands below to download and extract the schema files.

```
1
2
3
4
```

```
wget https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DB0260EN-SkillsNetwork/labs/Setting%20up%20a%20staging%20area/billing-datawarehouse.tgz

tar -xvzf billing-datawarehouse.tgz
ls *.sql
```

Copied! Executed!

You should see 4 .sql files listed in the output

Step 2: Create the schema

Run the command below to create the schema in the `billingDW` database.

```
1. 1
```

```
1. psql  -h localhost -U postgres -p 5432 billingDW < star-schema.sql
```

Copied! Executed!

You should see an output similar to the one below.

```
theia@theiadocker-rsannareddy:/home/project$ psql  -h localhost -U postgres -p 5432 billingDW < star-schema.sql
BEGIN
CREATE TABLE
CREATE TABLE
CREATE TABLE
ALTER TABLE
ALTER TABLE
COMMIT
theia@theiadocker-rsannareddy:/home/project$
```

Exercise 4 - Load data into Dimension tables

When we load data into the tables, it is a good practice to load the data into dimension tables first.

Step 1: Load data into DimCustomer table

Run the command below to load the data into DimCustomer table in `billingDW` database.

```
1. 1
```

```
1. psql  -h localhost -U postgres -p 5432 billingDW < DimCustomer.sql
```

Copied! Executed!

Step 2: Load data into DimMonth table

Run the command below to load the data into DimMonth table in `billingDW` database.

```
1. 1
```

```
1. psql  -h localhost -U postgres -p 5432 billingDW < DimMonth.sql
```

Copied! Executed!

Exercise 5 - Load data into Fact table

Load data into FactBilling table

Run the command below to load the data into FactBilling table in `billingDW` database.

```
1. 1
```

```
1. psql  -h localhost -U postgres -p 5432 billingDW < FactBilling.sql
```

Copied! Executed!

Exercise 6 - Run a sample query

Run the command below to check the number of rows in all the tables in the `billingDW` database.

```
1. 1
```

```
1. psql  -h localhost -U postgres -p 5432 billingDW < verify.sql
```

Copied! Executed!

You should see an output similar to the one below.

```
theia@theiadocker-rsannareddy:/home/project$ psql  -h localhost -U postgres -p 5432 billingDW < verify.sql
"Checking row in DimMonth Table"
 count
-------
   132
(1 row)

"Checking row in DimCustomer Table"
 count
-------
  1000
(1 row)

"Checking row in FactBilling Table"
 count
--------
 132000
(1 row)
```

Your data warehouse staging area is now ready.

Practice exercises

In this practice session, you will create a database named `practice` and load the data into it.

1. Problem:

   Create a database named `practice`.

▸ Click here for Hint
▾ Click here for Solution

```
1. 1
```

```
1. createdb -h localhost -U postgres -p 5432 practice
```

Copied! Executed!

2. Problem:

   In the `practice` database, create a schema using `star-schema.sql`.

▾ Click here for Hint

   Use the psql command, make sure you are setting the database to `practice`.

▾ Click here for Solution
```bash psql -h localhost -U postgres -p 5432 practice < star-schema.sql ```

3. Problem:

   In the `practice` database, load the data into all tables using the DimMonth.sql, DimCustomer.sql and FactBilling.sql.

▸ Click here for Hint
▾ Click here for Solution

```
1. 1
2. 2
3. 3
```

```
1. psql  -h localhost -U postgres -p 5432 practice < DimMonth.sql
2. psql  -h localhost -U postgres -p 5432 practice < DimCustomer.sql
3. psql  -h localhost -U postgres -p 5432 practice < FactBilling.sql
```

Copied! Executed!

4. Problem:

   _Verify that you have correctly loaded the data into the `practice` database.

▸ Click here for Hint
▾ Click here for Solution

```
1. 1
```

```
1. psql  -h localhost -U postgres -p 5432 practice < verify.sql
```

Copied! Executed!

Congratulations!! You have successfully finished the Setting up a staging server lab.

Authors

Ramesh Sannareddy

Other Contributors

Rav Ahuja

Change Log

Date (YYYY-MM-DD)	Version	Changed By	Change Description
```

Date (YYYY-MM-DD)	Version	Changed By	Change Description
2023-05-11	0.6	Eric Hao & Vladislav Boyko	Updated Page Frames
2023-05-10	0.5	Eric Hao & Vladislav Boyko	Updated Page Frames
2023-05-10	0.4	Eric Hao & Vladislav Boyko	Updated Page Frames
2021-09-4	0.1	Ramesh Sannareddy	Created initial version of the lab
2022-07-29	0.2	Lakshmi Holla	Updated markdown
2023-05-05	0.3	Vladislav Boyko	Updated markdown, removed create terminal image