SQL Lab

Estimated Time Needed: 60 min

In this challenge, you will create a database, import data from three sources to populate tables, and perform database operations in the MySQL database service using the phpMyAdmin graphical user interface (GUI) tool.

Objectives

After completing this lab, you will be able to use phpMyAdmin with MySQL to:

- Create and populate a database and tables.
- Execute Structured Query Language (SQL) commands to perform basic database operations.
- Retrieve data from tables using SELECT statements.
- Filter the data output using WHERE statements.
- Aggregate data get ordered results using functions like SUM, MIN, MAX, ORDER BY.
- Use window functions to get specified output.
- Retrieve data from two or more tables using SQL JOINS.

Software used in this lab

You will use MySQL to complete this lab. MySQL is a free, open-source relational database system that offers a command line interface (MySQL) and a third-party web interface (phpMyAdmin) to efficiently store, manipulate, and retrieve data.

MySQL is a service available on Skills Network Labs (SN Labs) Cloud IDE, a virtual lab environment used in this course. SN Labs Cloud IDE is great way to do projects without downloading, installing, configuring, and integrating software on your own computer.

Two Components of the SN Labs Cloud IDE:

- The instructions that you will follow to complete this lab are displayed on the left side of the screen.
- The area on the right side of the screen is where you will use the menus, terminals, and tools to complete the lab exercises.

Dataset used in This Lab

The datasets used in this lab are three SQL files called **Salary Data**, **Sales Data** and **Employee Data**. To complete the exercises in this lab, you will be instructed to save and upload the datasets to your local machine and use a locally installed database or the tool provided in the course.

Import the given SQL input files into a database.

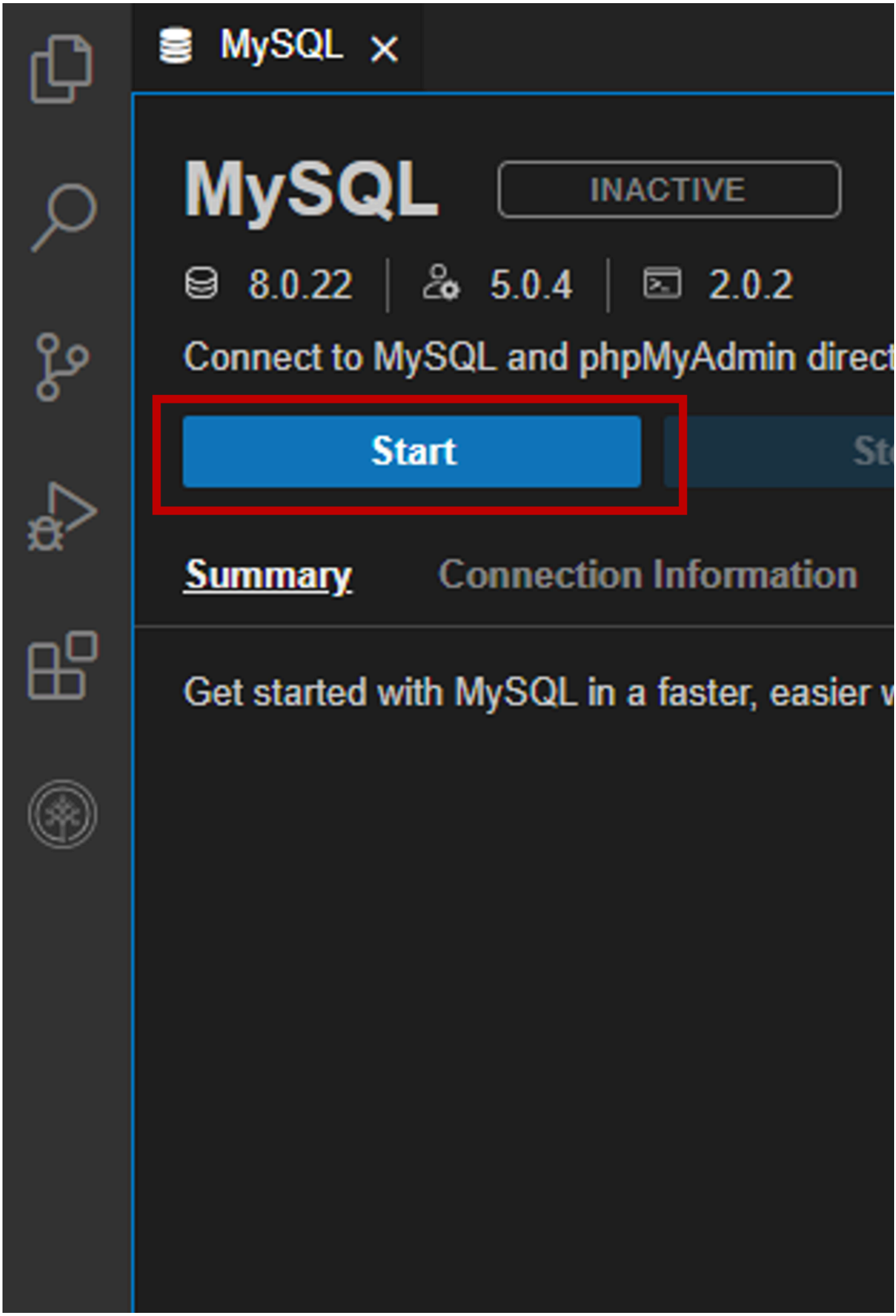
Prework - Create and populate database

TASK A: Create a Database

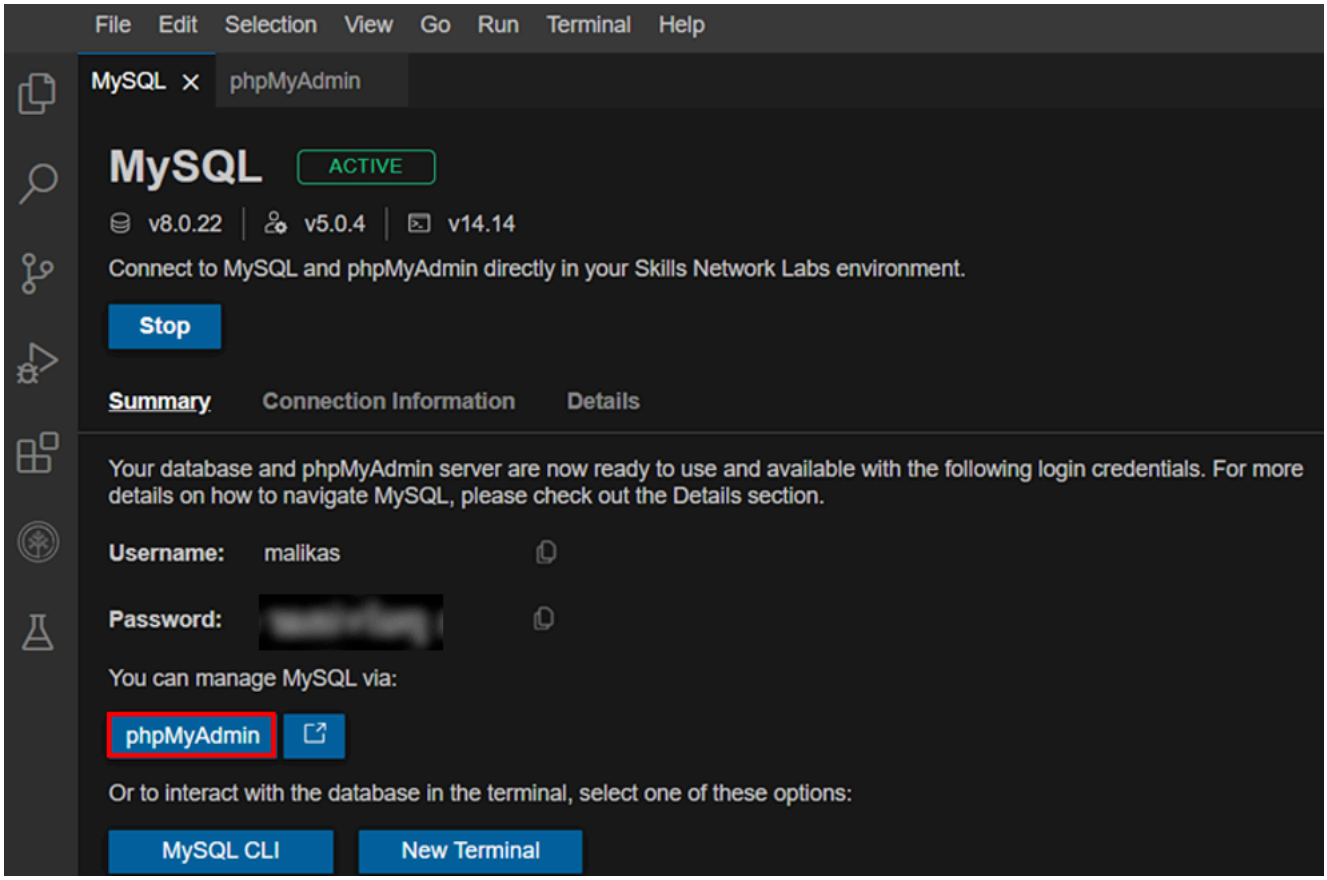
1. Start the MySQL service session using the `Open MySQL Page in IDE` button.

Open MySQL Page in IDE

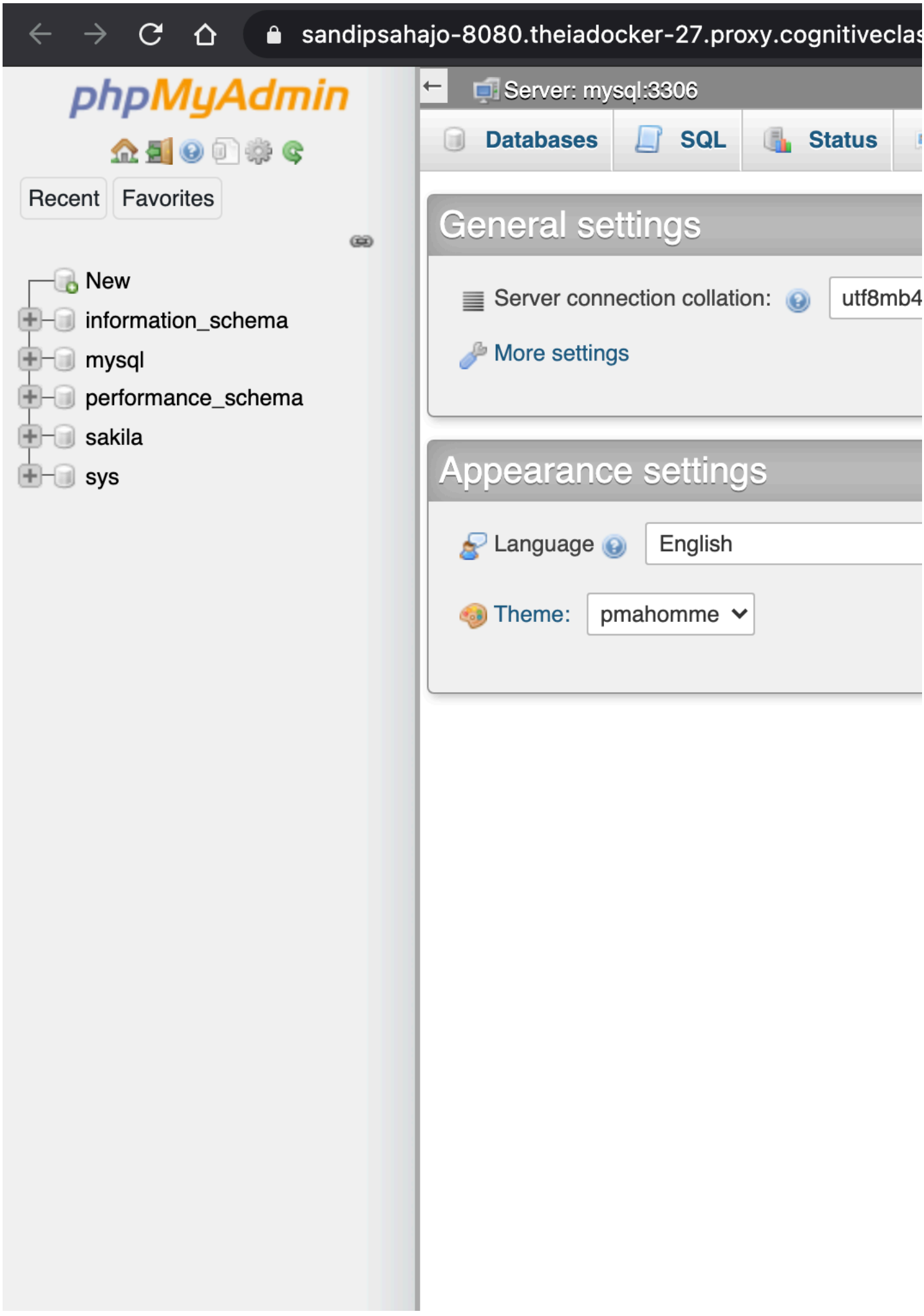
To start the MySQL, click `Start`.


MySQL
MySQL
INACTIVE
8.0.22
5.0.4
2.0.2
Connect to MySQL and phpMyAdmin direct
Start
Summary
Connection Information
Get started with MySQL in a faster, easier

2. Once MySQL has started, click on `phpMyAdmin` button to open phpMyAdmin in the same window.

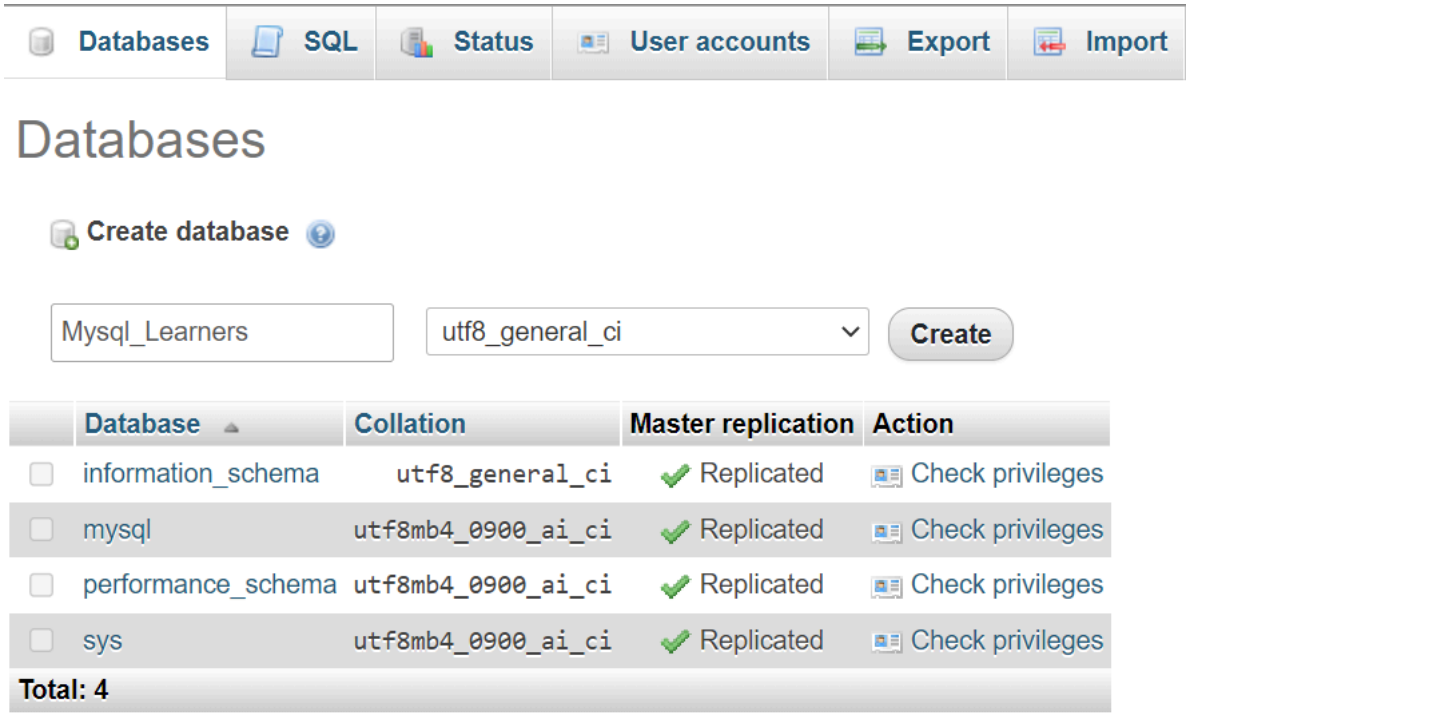


3. You will see the **phpMyAdmin GUI** tool.


sandipsahajo-8080.theiadocker-27.proxy.cognitiveclas
phpMyAdmin
Recent
Favorites
New
information_schema
mysql
performance_schema
sakila
sys
Server: mysql:3306
Databases
SQL
Status
General settings
Server connection collation:
utf8mb4
More settings
Appearance settings
Language
English
Theme:
pmahomme

4. In the tree-view, click `New` to create a new empty database. Then enter `Mysql_Learners` or any other name you desire, as the name of the database and select utf8_general_ci and click `Create`.

   UTF-8 is the most commonly used character encoding for content or data.



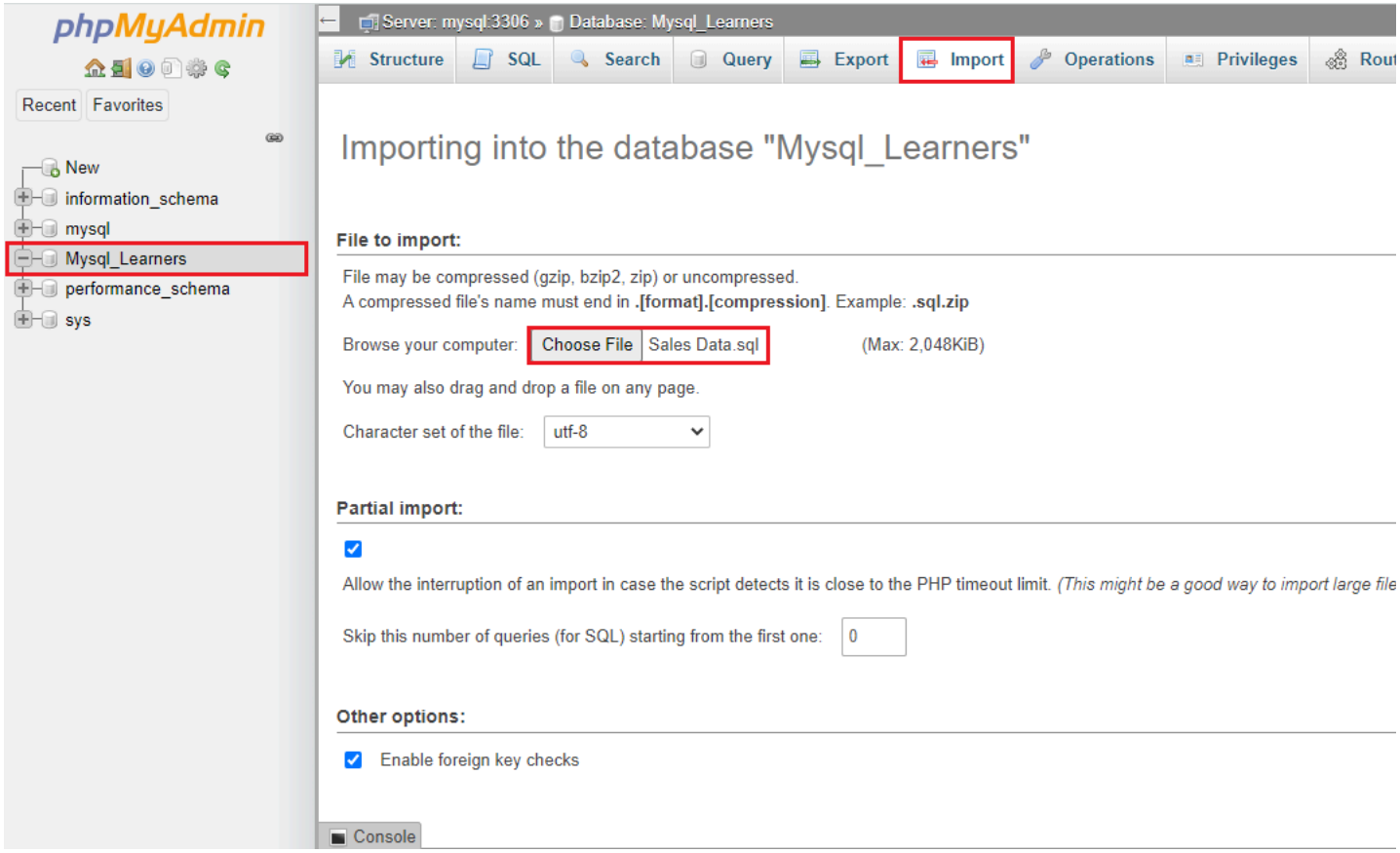
TASK B: Create and load tables using sql files.

1. Download the 3 sql files below to your local computer:

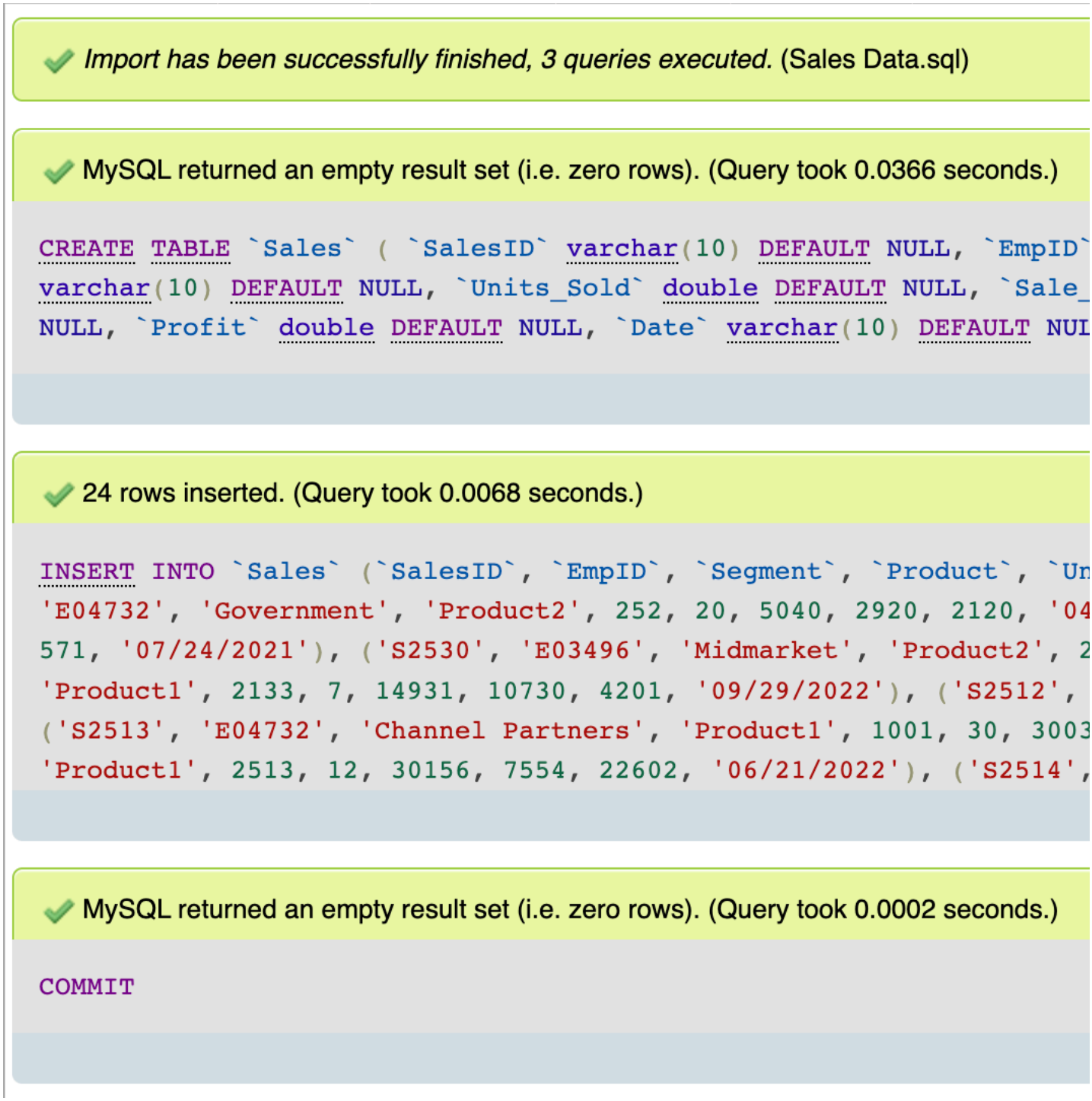
   - Employee Data.sql
   - Salary Data.sql
   - Sales Data.sql

2. To load each sql file do the following steps.

   - Select your database, in the case of the example shown if it is **Mysql_Learners** and click on **Import** tab.
   - Click on `Choose File`. Browse for the file and upload it .
   - Later scroll down and click the `Go` button.

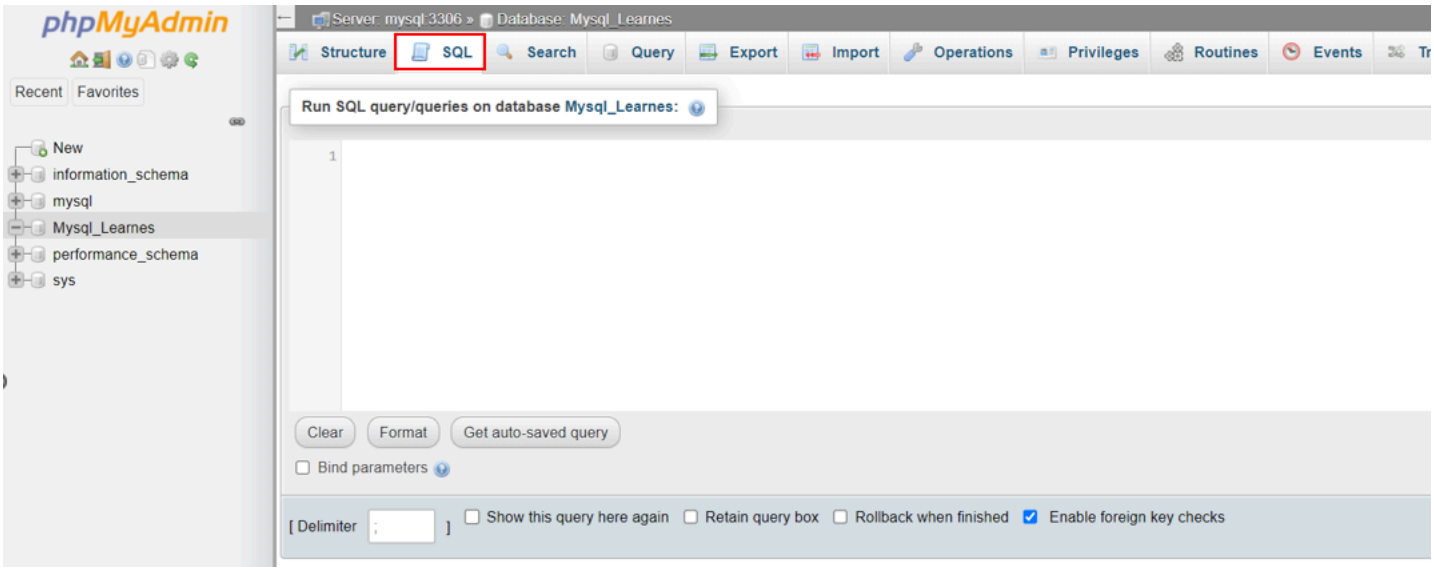


- Once the scripts are loaded, you will get a message that, it is imported successfully.



You can import the other sql files in the same way.

3. To run the SQL queries you need to copy the given codes and paste it to the text area of the SQL page and click on Go.



Data Engineering

In this section you will perform data cleansing (removing duplicates) and data transformation (change column name) operations on the data.

1. Identify the duplicate entry for employees in the employee table using GROUP BY and HAVING statements.

▶ Solution syntax
▶ Output

2. Select the duplicate entry for employees and delete the row with the higher EMPID.

▶ Solution syntax
▶ Output

3. Data Transformation – Change the column Salary in the Salary table to “Annual_Income”

▶ Solution syntax
▶ Output

SQL JOINS, Aggregations

Use SQL JOINS, Aggregations where needed, to derive metrics from the database tables.

1. Using the tables given, find out the Total number of men and women employees in the company who are aged below 50 yrs.

▶ Solution syntax
▶ Output

2. Using the tables, find the employees whose salary is greater than $150000.

   Note – Rename column Annual_Income back to Salary in the table Salary

▶ Solution syntax
▶ Output

SQL GROUP BY, HAVING

Use SQL GROUP BY and HAVING statements to get some count metrics from database tables.

1. Display products grouped by segments with total Sales greater than $100,000.

▶ Solution syntax
▶ Output

Formatted Output

Show output result ordered in a certain way (Use window functions row_number() or rank() and order by statements).

1. Show an output table of Sales generated by employees ordered highest to lowest.

▶ Solution syntax
▶ Output

2. Show an output table of Sales(ordered highest to lowest) generated by employees in different segments and rank them for each employee.

▶ Solution syntax
▶ Output

Explain the following:

- When you would use COALESCE function
- What is the difference between Union and Union
- What is the difference between clustered and non-clustered indexes

Author

Roopa Raghavan

© IBM Corporation. All rights reserved.