DAT 375 Module Three Assignment Tutorial

Use this tutorial to assist with your completion of the module three assignment on Choosing the Right Scripts.

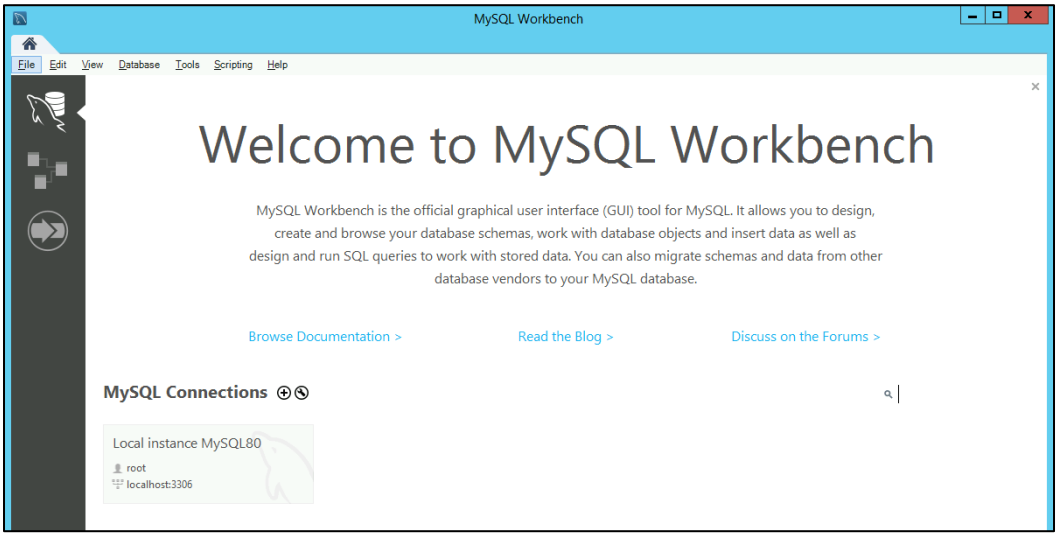
Directions

1. In the **Virtual Lab** module of this course, select click **Virtual Lab Access**.

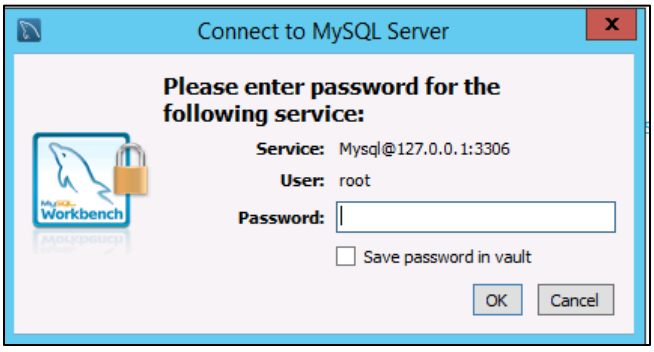
2. Use your SNHU login to access Apporto.

3. Navigate to “Data Analysis Techniques.”

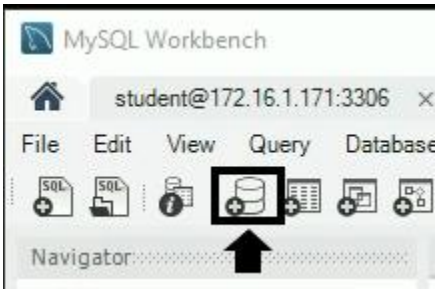
4. Open **MySQL Workbench** from the Desktop.



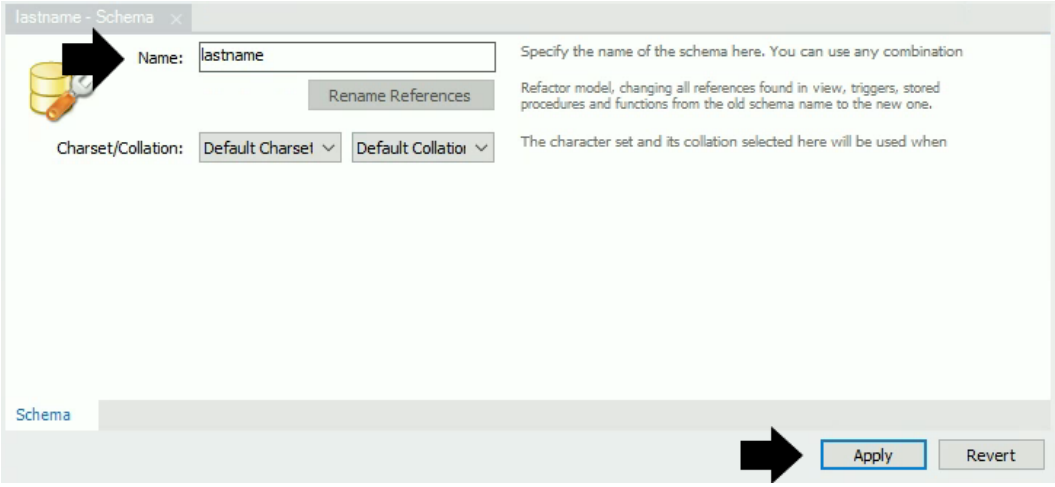
5. Enter the password: student.



6. This step will open the **Navigator** on the left. Create a new schema by selecting the **schema icon** (if you hover over the icon, it will read “create a new schema in the connected server.”):

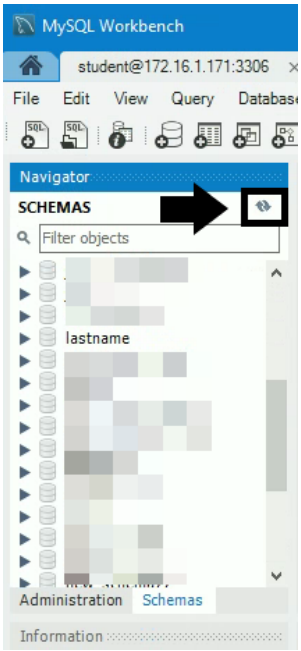


7. Enter your last name in lowercase as the schema name. Select **Apply.**

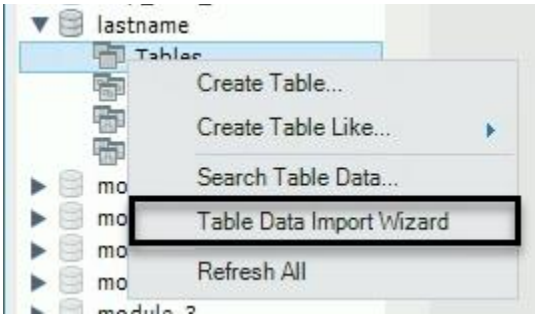


8. A window will pop up to Review the SQL Script. Select **Apply**. A second screen will appear to Apply SQL Script. Select **Finish**.

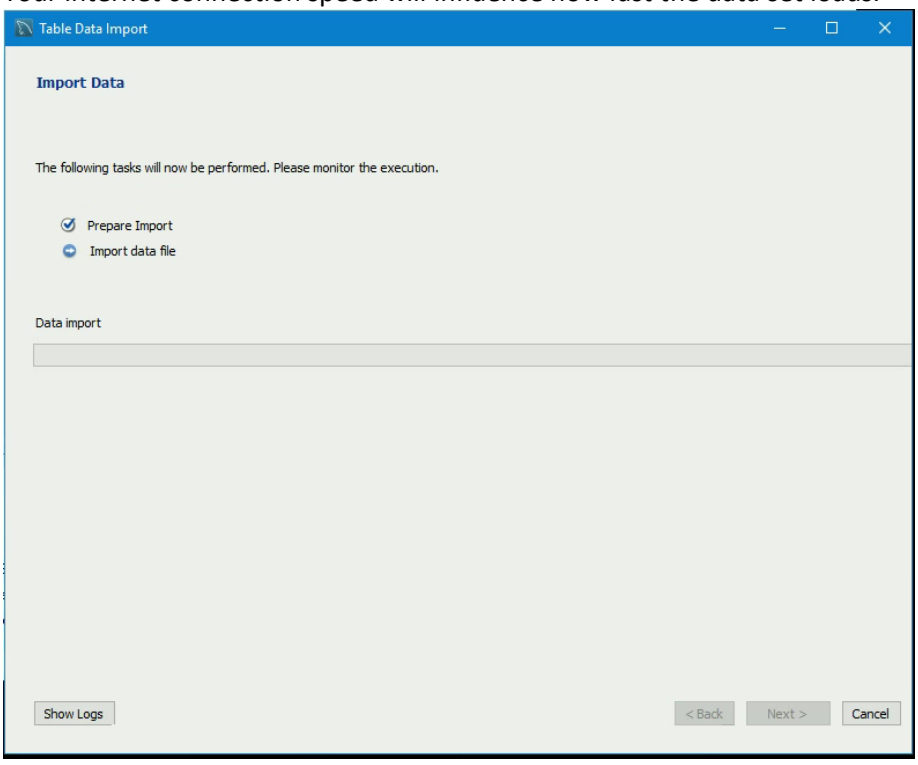
9. Select the **refresh** icon next to SCHEMAS. Locate your script alphabetically in the list.



10. Select the dropdown arrow in front of your schema to open the **Tables** option. Right-click **Tables** and select **Table Data Import Wizard**.



11. The **Table Data Import** window will open. Select **Browse** to select a file to import.
    a. Select **Desktop** and select **DAT 375 Module Three Assignment Dataset.** Select **Open**, and then **Next**.
    b. The **Select Destination** screen appears. It defaults to: **Create new table**. Select **Next.**
    c. The **Configure Import Settings** screen appears. Keep default settings and select **Next.**
    d. The Import Data screen appears. Select **Next.**
        i. **Note:** The import data step will take a considerable time (10–30 minutes or more). The **Data import** bar will show loading progress by filling with green bars. Your internet connection speed will influence how fast the data set loads.



    e. Once it loads, select **Next**, then **Complete**.

12. For your data set to show up under your schema, you will need to right-click on the **schema** and select **Refresh All**.

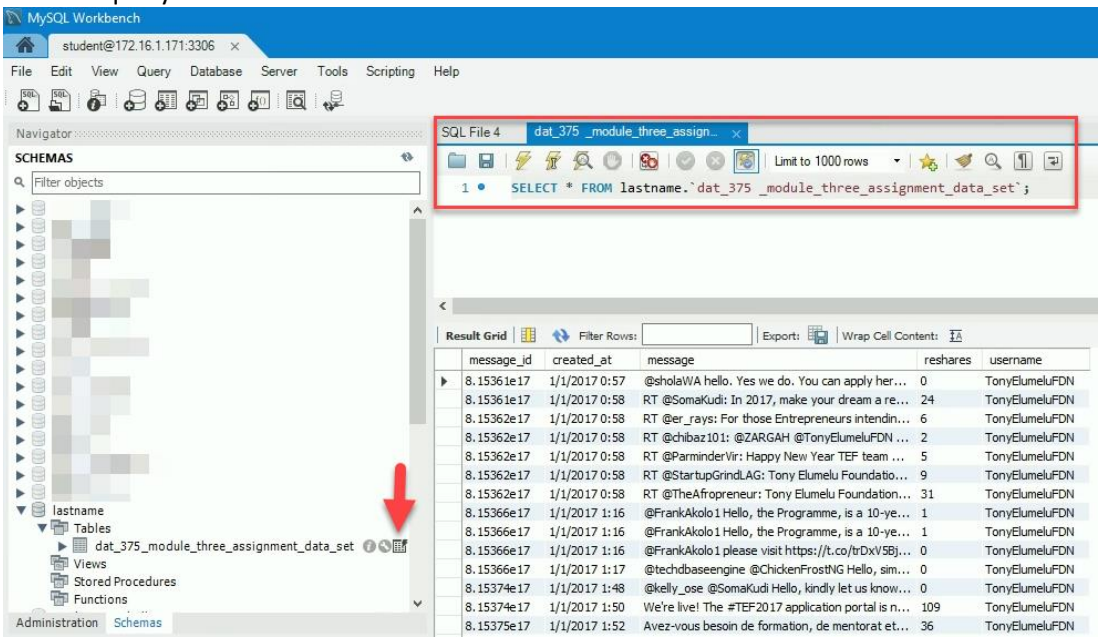
13. Select **File**, then **New Query Tab** to access the query screen.

14. For your analysis, refer to your new table using the following command:

    SELECT * FROM [lastname].[tablename];

    Note: Replace [lastname] with the name of your schema and [tablename] with the name of your table.

    If you select the **Select *** icon next to your table name, MySQL will copy the name of your table into the query in the correct format.



15. Now you can perform some analysis on the data. Experiment with the following scripts. Consider SQL Commands: The Complete List (w/ Examples) by Dataquest for a comprehensive list of SQL commands. Make sure to replace [lastname] with the name of your schema and [tablename] with the name of your table..

    - To find the count by username:

      SELECT username, count(*) as count
      FROM [lastname].[tablename]
      GROUP BY username DESC
      ORDER BY count DESC

    - To find the average number of messages by a single username:

```
SELECT AVG(count)
FROM (
SELECT username, count(*) as count
FROM [lastname]. [tablename]
GROUP BY username DESC
ORDER BY count DESC
) as counts;
```

- To find the average number of reshares by a single username:

```
SELECT AVG(reshares)
FROM (
SELECT username, sum(reshares) as reshares
FROM [lastname]. [tablename]
GROUP BY username
) as counts;
```

16. **Important:** After you finish your analysis, delete your schema by right-clicking on the **schema** and selecting **Drop Schema**. A pop-up window will confirm your selection. Select **Drop Now**.