SmartBridge Data Analytics 2023

Assessment – 1

Name - Mohamed Zaid
Reg No. - 20BCE2589

Course Name - VIT Data Analytics 2023
Institution - Vellore Institute of Technology
Campus - Vellore

Date of Submission - 30th May 2023

Vellore Institute of Technology

Questions

1) Download tableau desktop & MySQL in your PC or laptop.

2) Create a new schema in your MySQL workbench and upload your data.

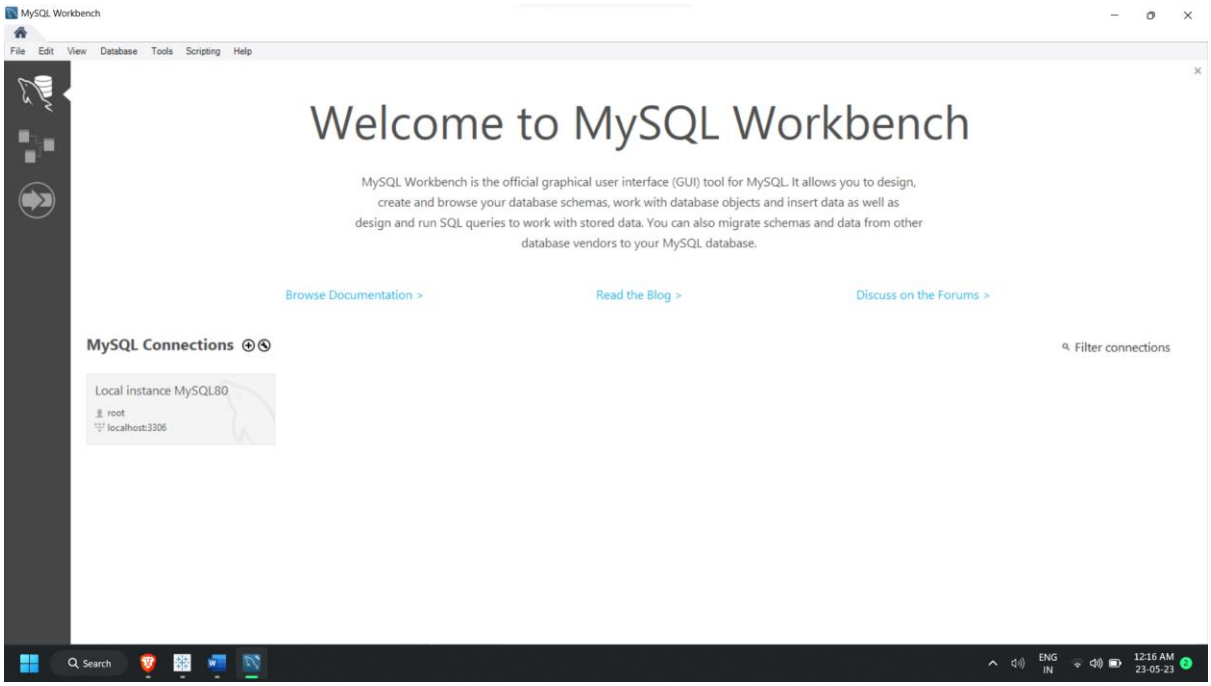
3) Connect your MySQL and Tableau

4) Create the below 5 charts/plots.
   - Bar chart
   - Tree map
   - Heatmap
   - Pie chart
   - Horizontal bar chart

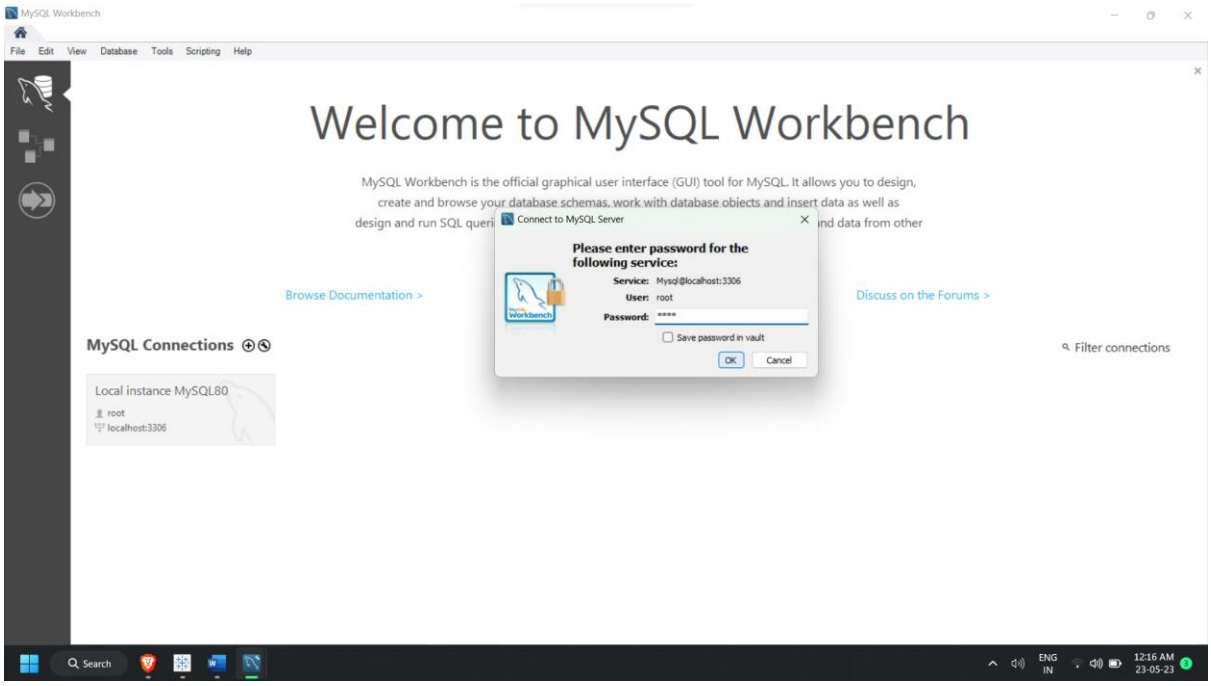
You can use the superstore data set.

Question 2

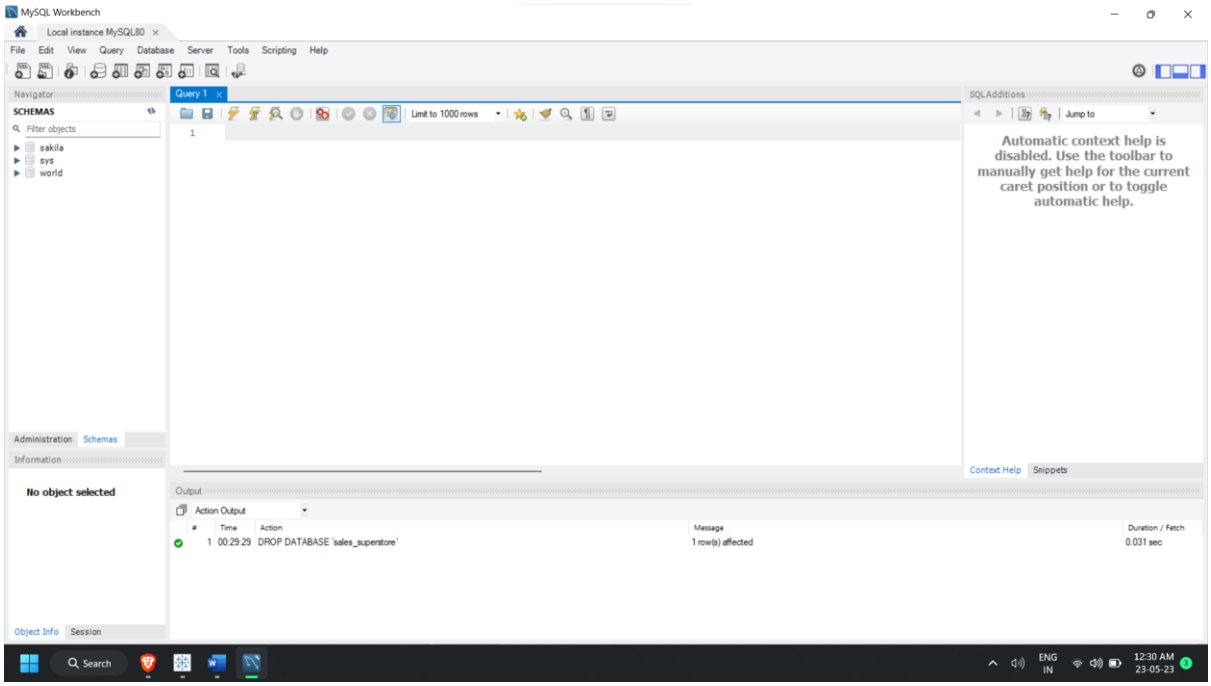
Step 1: Open MySQL Workbench.



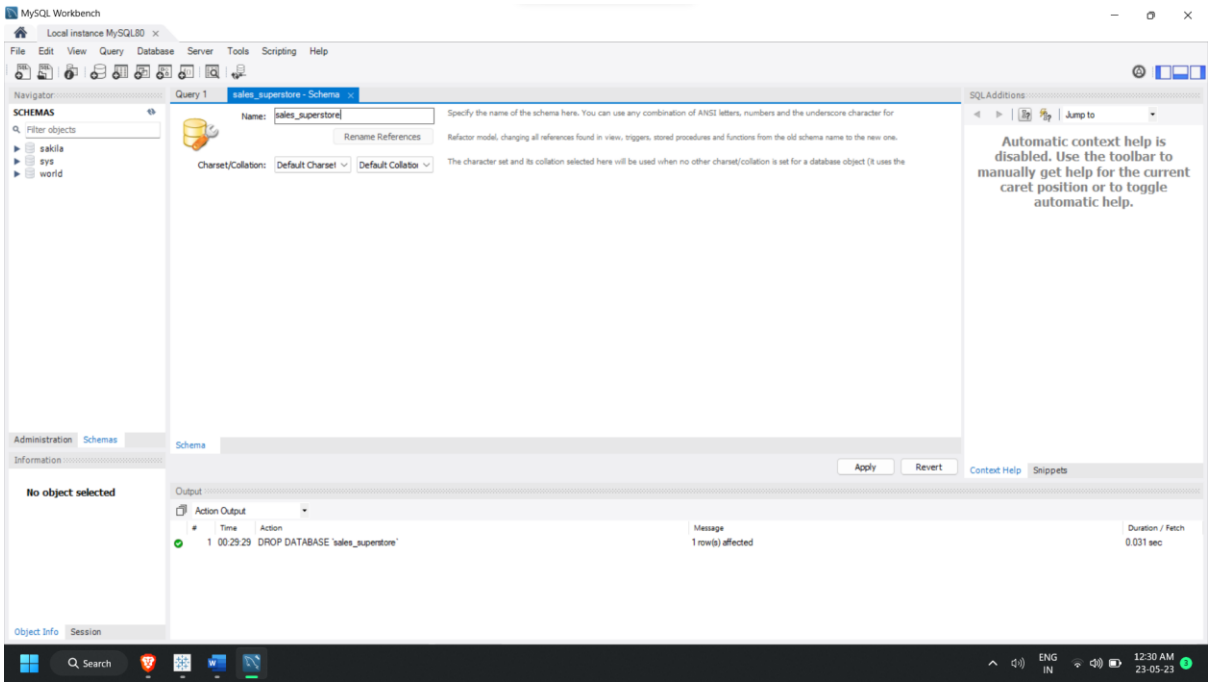
Step 2: Select the connection you have created while downloading MySQL and enter in the password.



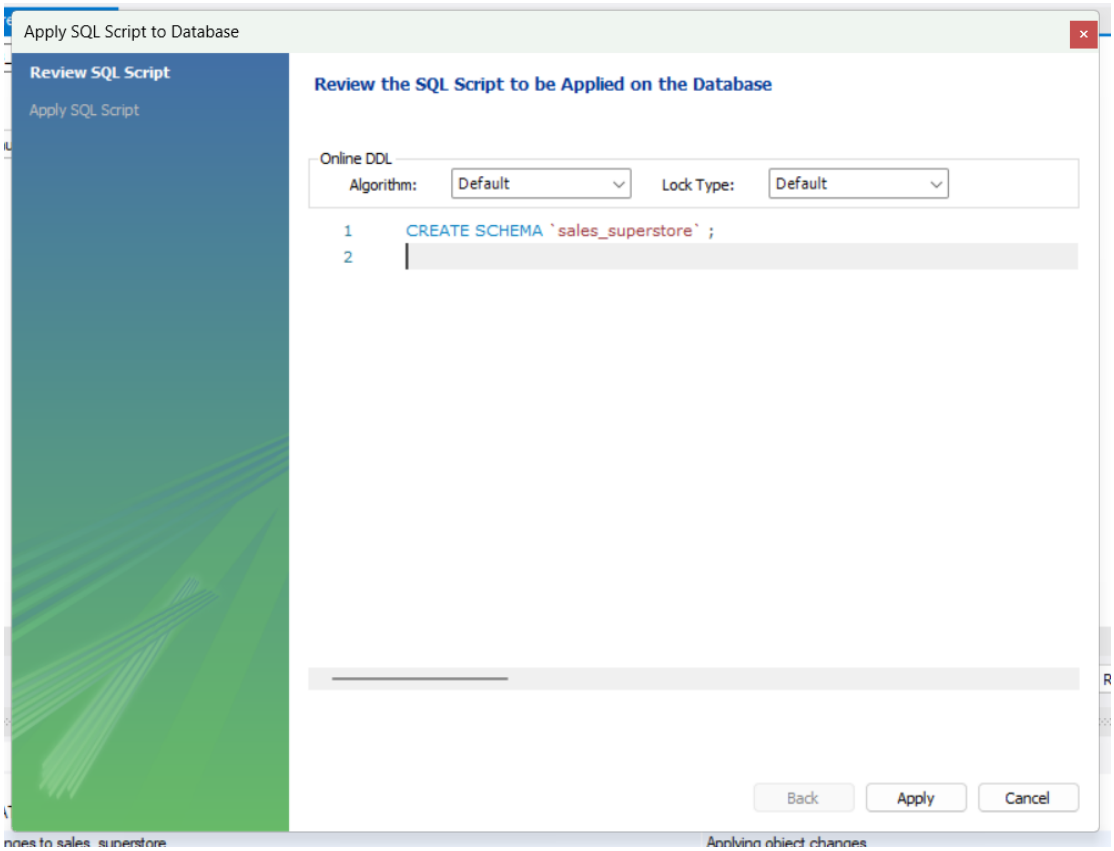
Step 3: You will be greeted with the MySQL Workbench homepage, once you are there you will need to select the new schema button that is under the Query tab in the main ribbon.

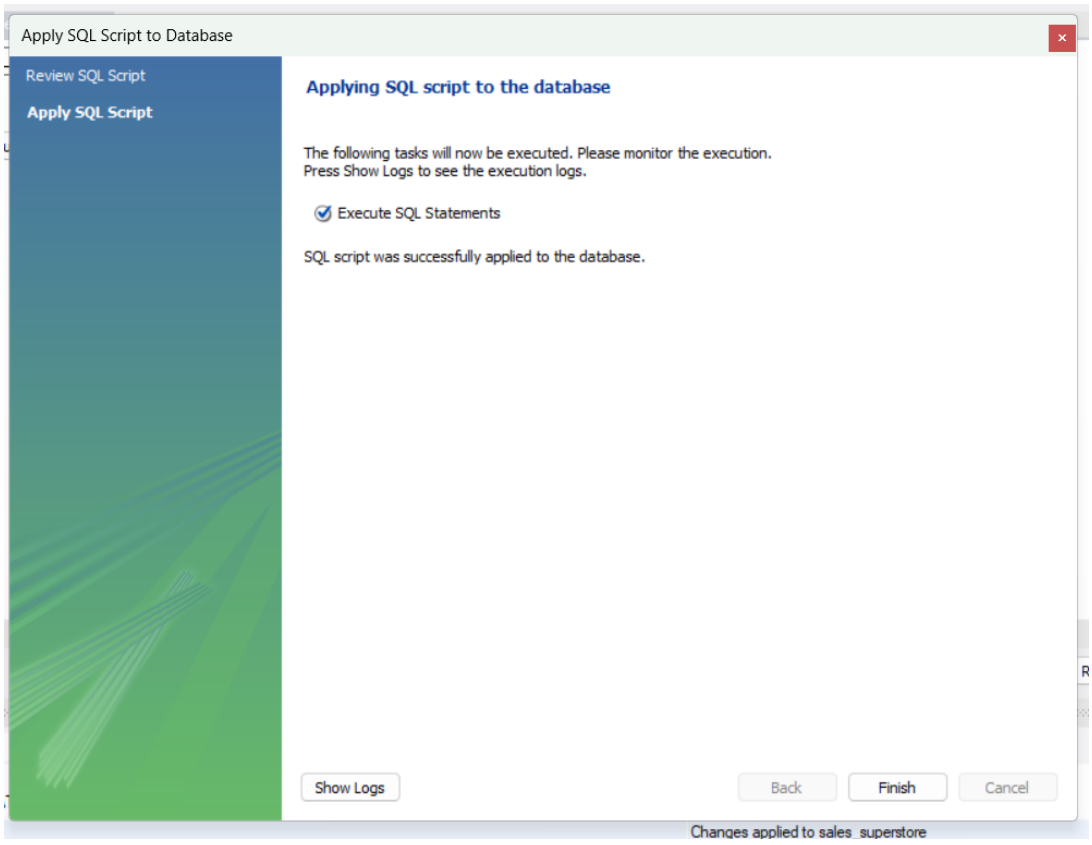


Step 4: Once you have selected the new schema button, you will be asked to end the name of the new schema, after that click apply.



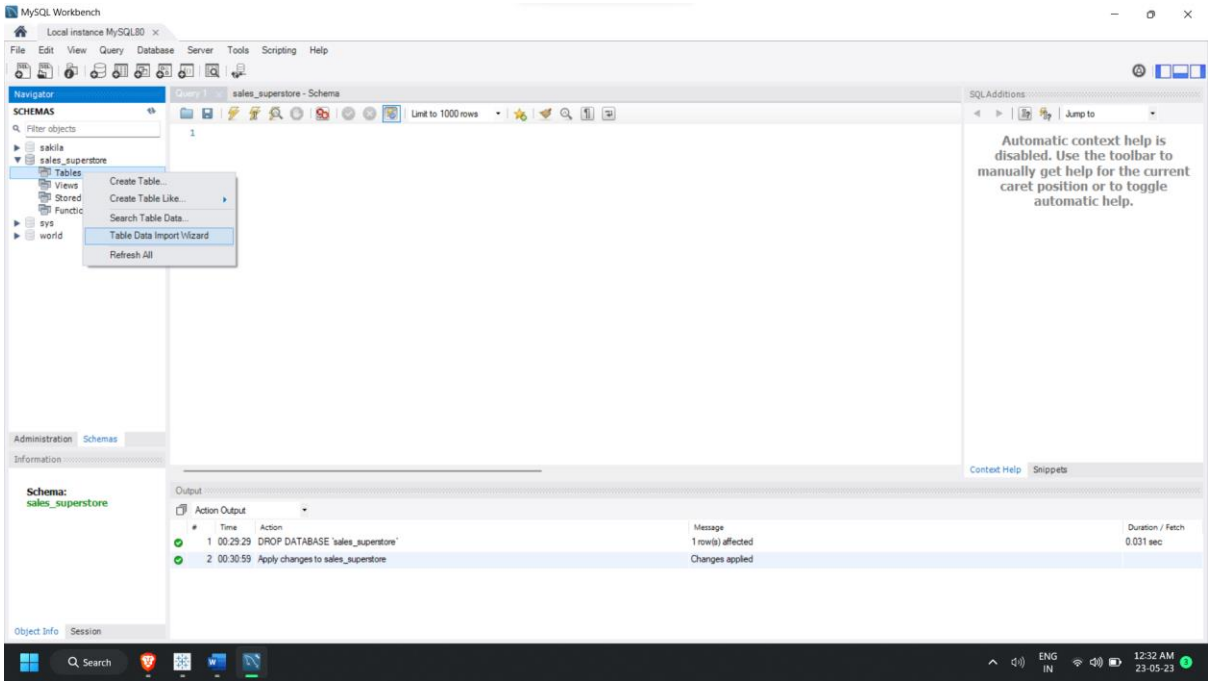
Step 5: A pop-up window will open and tell you some basic information, all you need to do is click apply.



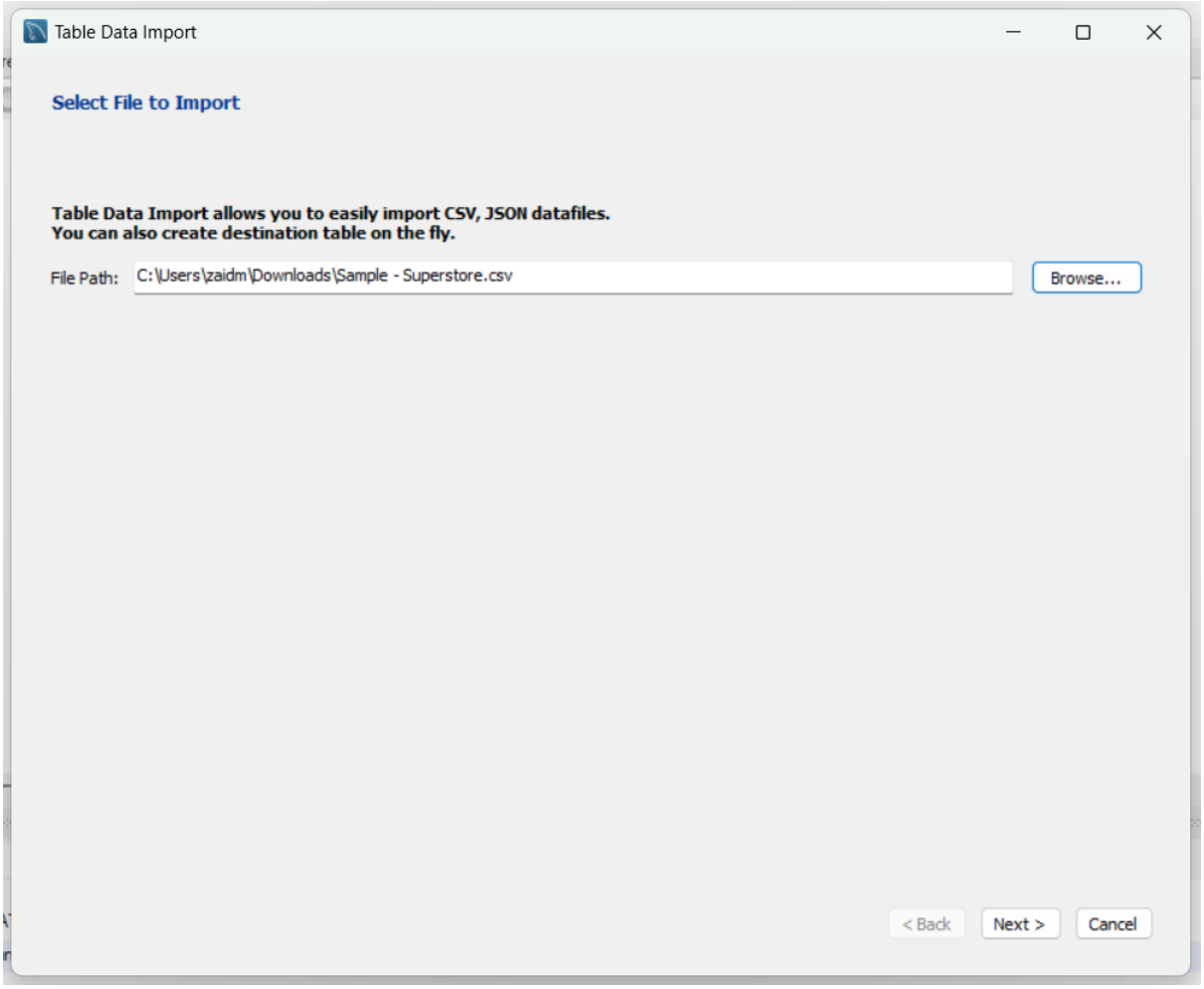


Step 6: After you click finish the pop window will close, you will have to click the refresh button that is on top of the schemas tab beside the lab.

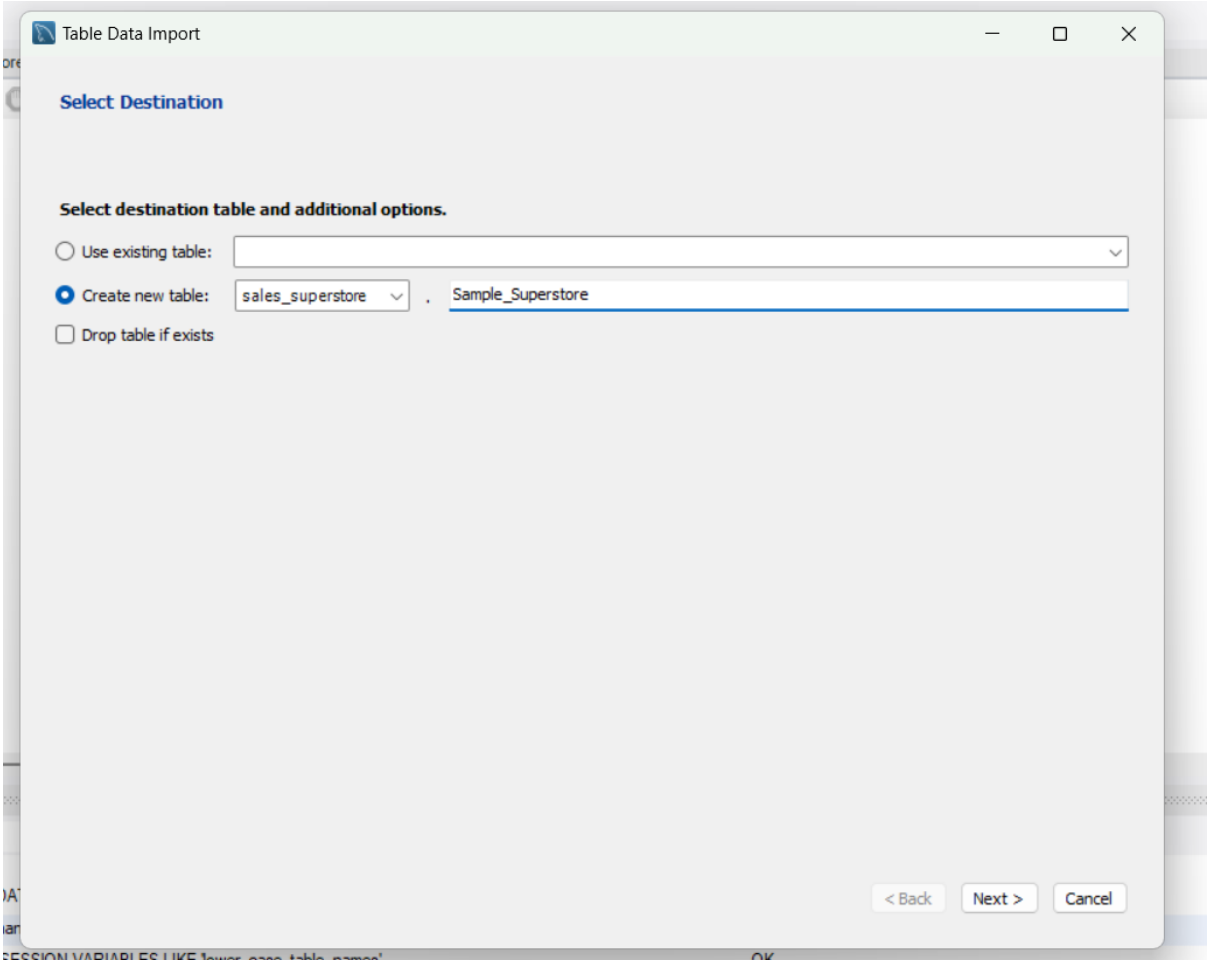
Step 7: Then you will see your schema name added to the list, double left click your schema, it will drop down a series of options, one of them will be “Tables”, right click the” Tables” option and select the “Table Data Import Wizard” option.



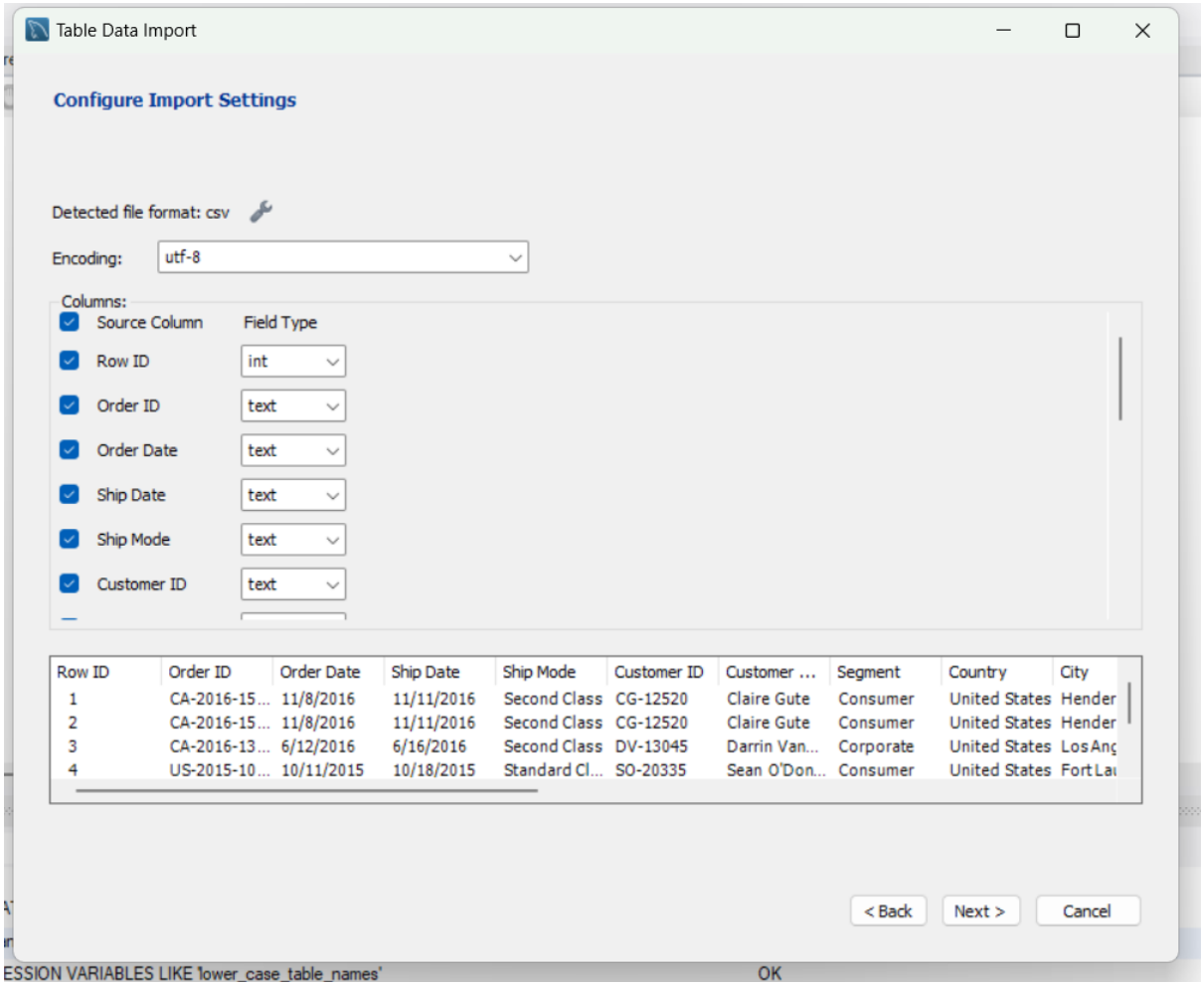
Step 8: A pop-up window will appear and ask you to locate the dataset that you want to store as a table in your schema.



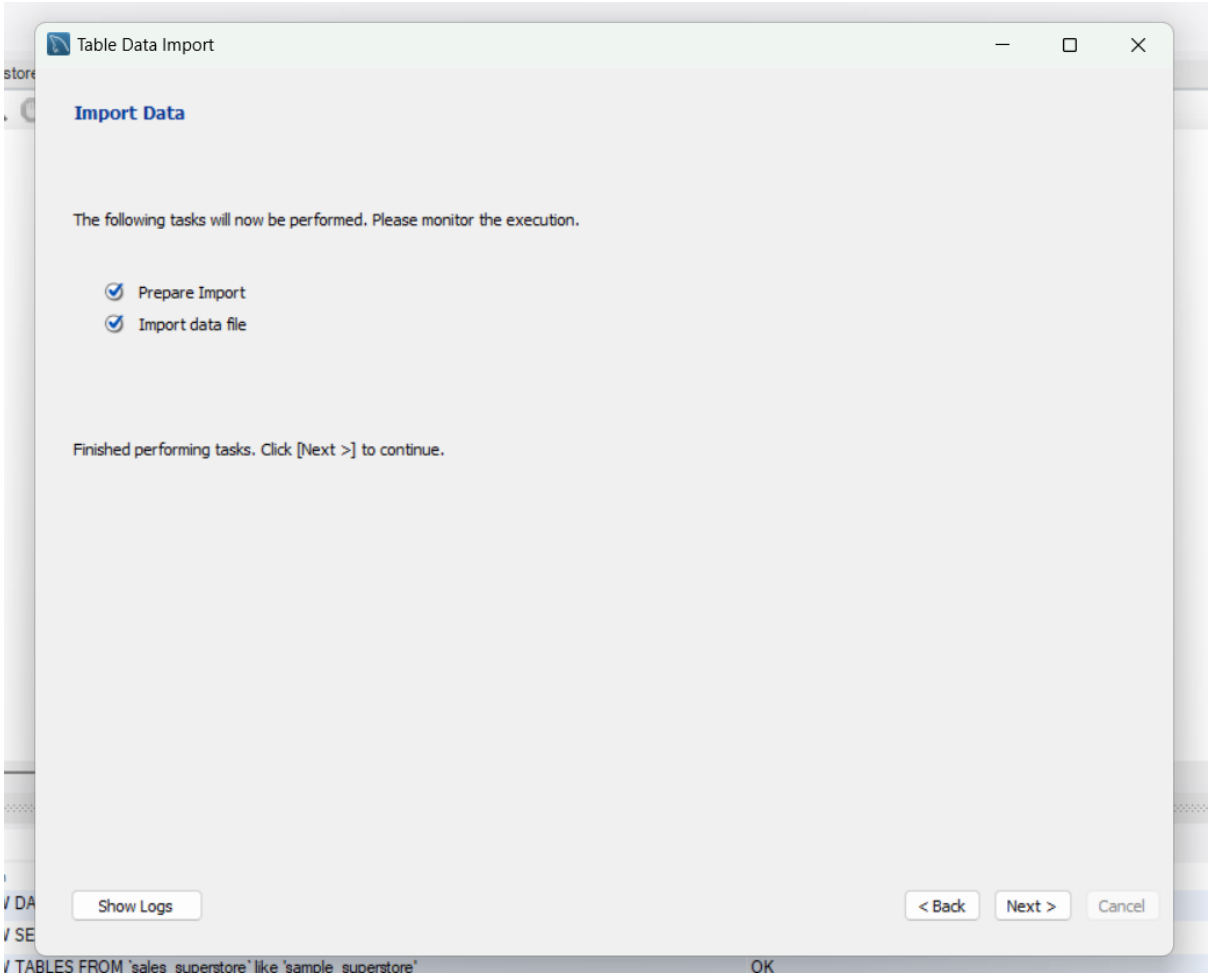
Step 9: In this step you will need to choose whether if you want to change the destination if the dataset or let it be in the same schema. You can also set a name for the dataset (the name of the table should not contain any space).



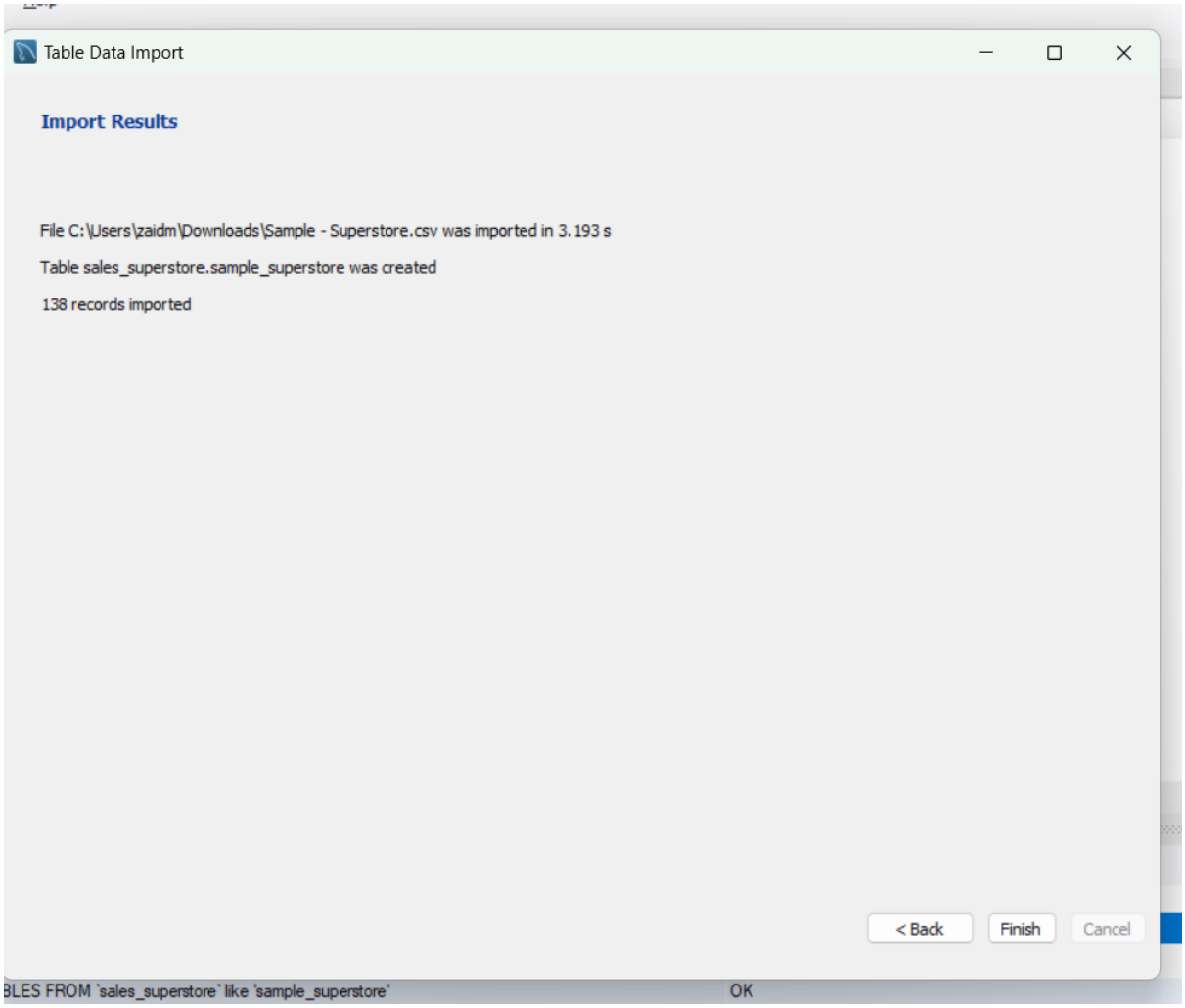
Step 10: In this step you can edit the fields that will be included in the table of your schema, if you want all the fields to be included, you can simply click next without changing any options.



Step 11: In this step, the pop-up window will be confirming you about the changes that is about to take place, all you need to do is to click next and wait for the changes that is taking place to complete.



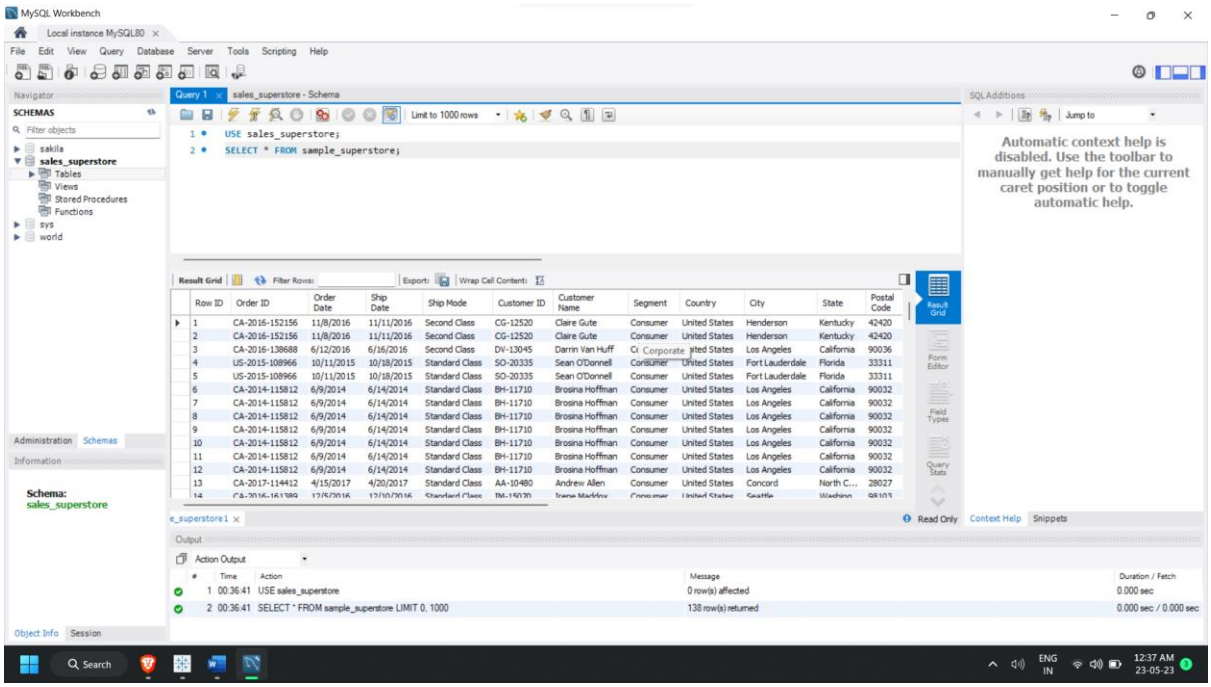
Step 12: After all the data from the file has been successfully imported into you table the pop-up window will inform you about the number of records that have been imported into your table, you can click finish.



Step 13: After the pop-up window closes, you can view the dataset you have imported and stored in the table by executing the query:

"USE sales_superstore;
SELECT * FROM sample_superstore;"

This query will select all the records present in your table and display it.



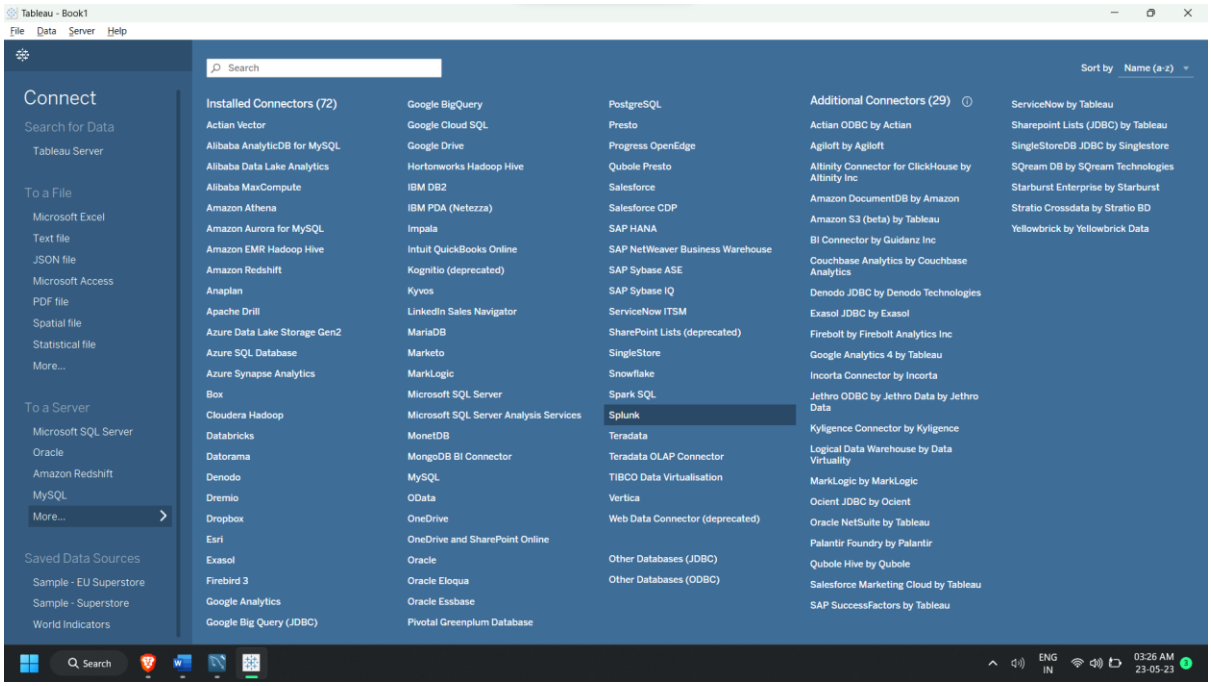
Congratulations, you have successfully imported your dataset file into your MySQL Workbench.

Question 3

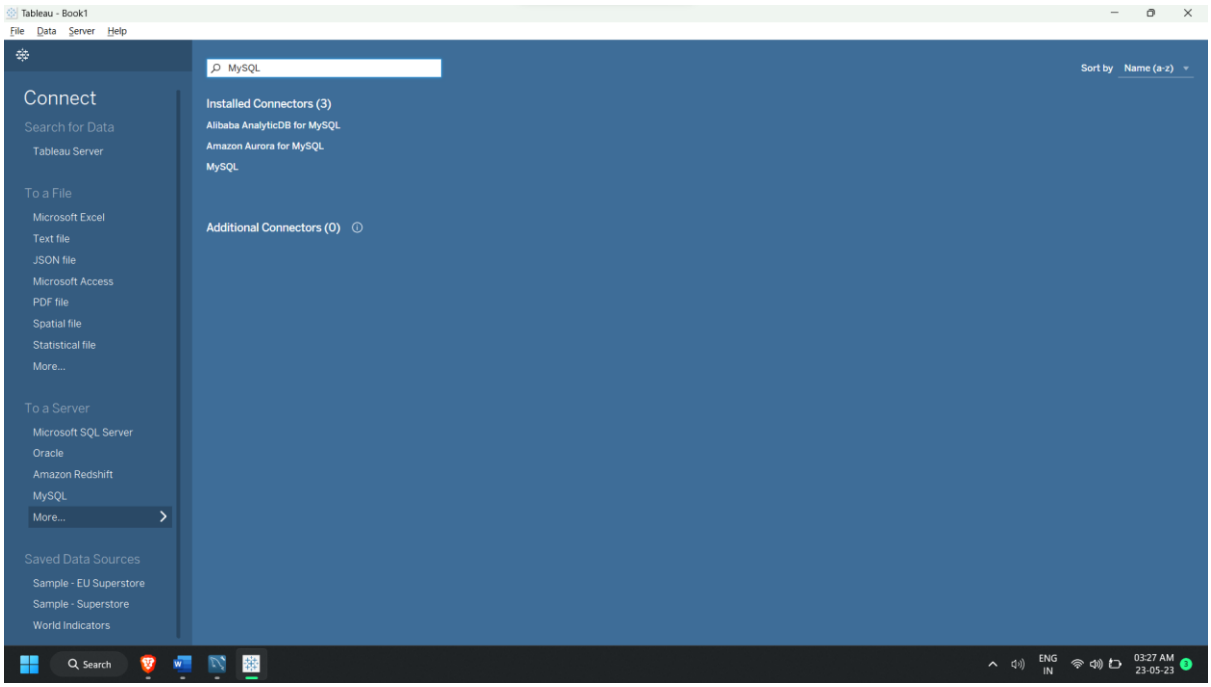
Step 1: Open Tableau Desktop.

Step 2: On the left side of the screen in the connect pane, under the “To a Server” section, there will be a “More…” option, select the said option.

Step 3: The selected option will list all the possible servers that can be connected to Tableau.



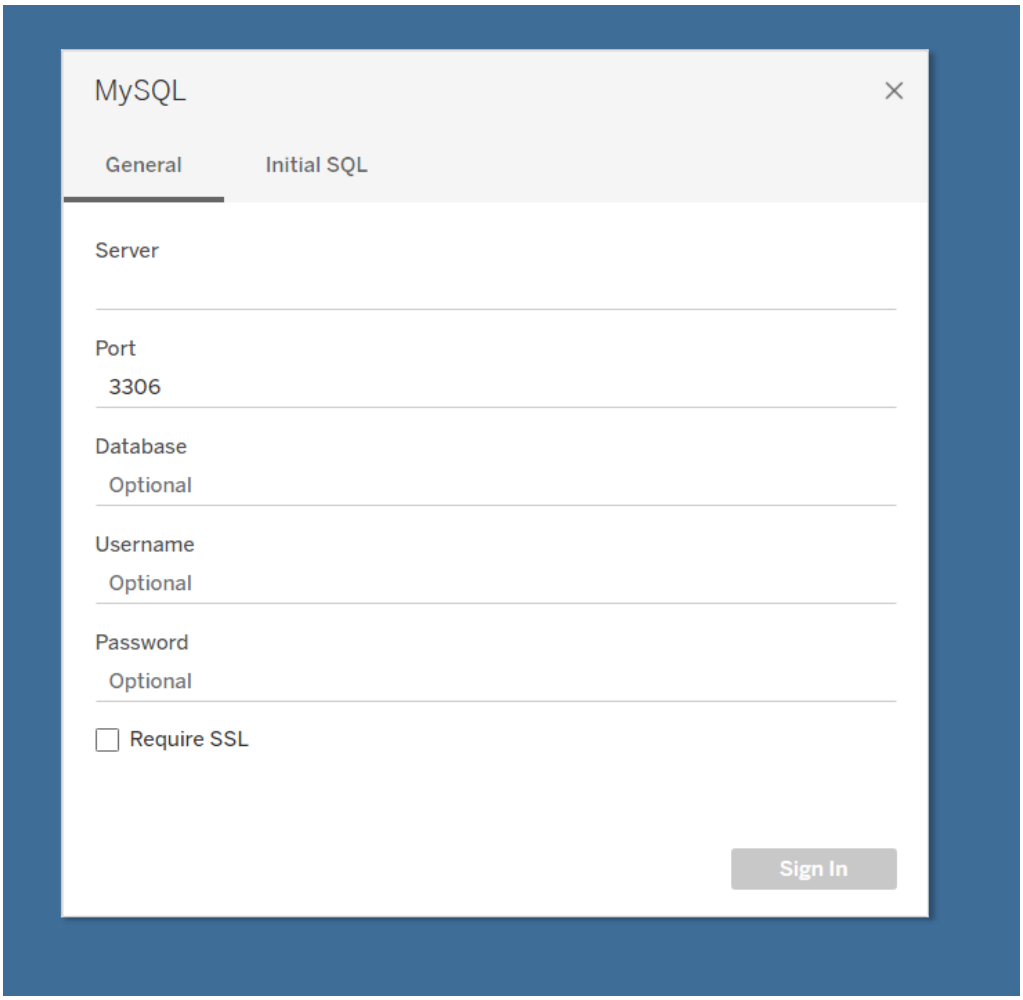
Step 4: In the top there will be a search box, we type “MySQL” in the search box, the search box will filter out all the present servers to show you “MySQL”. Then we choose the shown option that says “MySQL”.



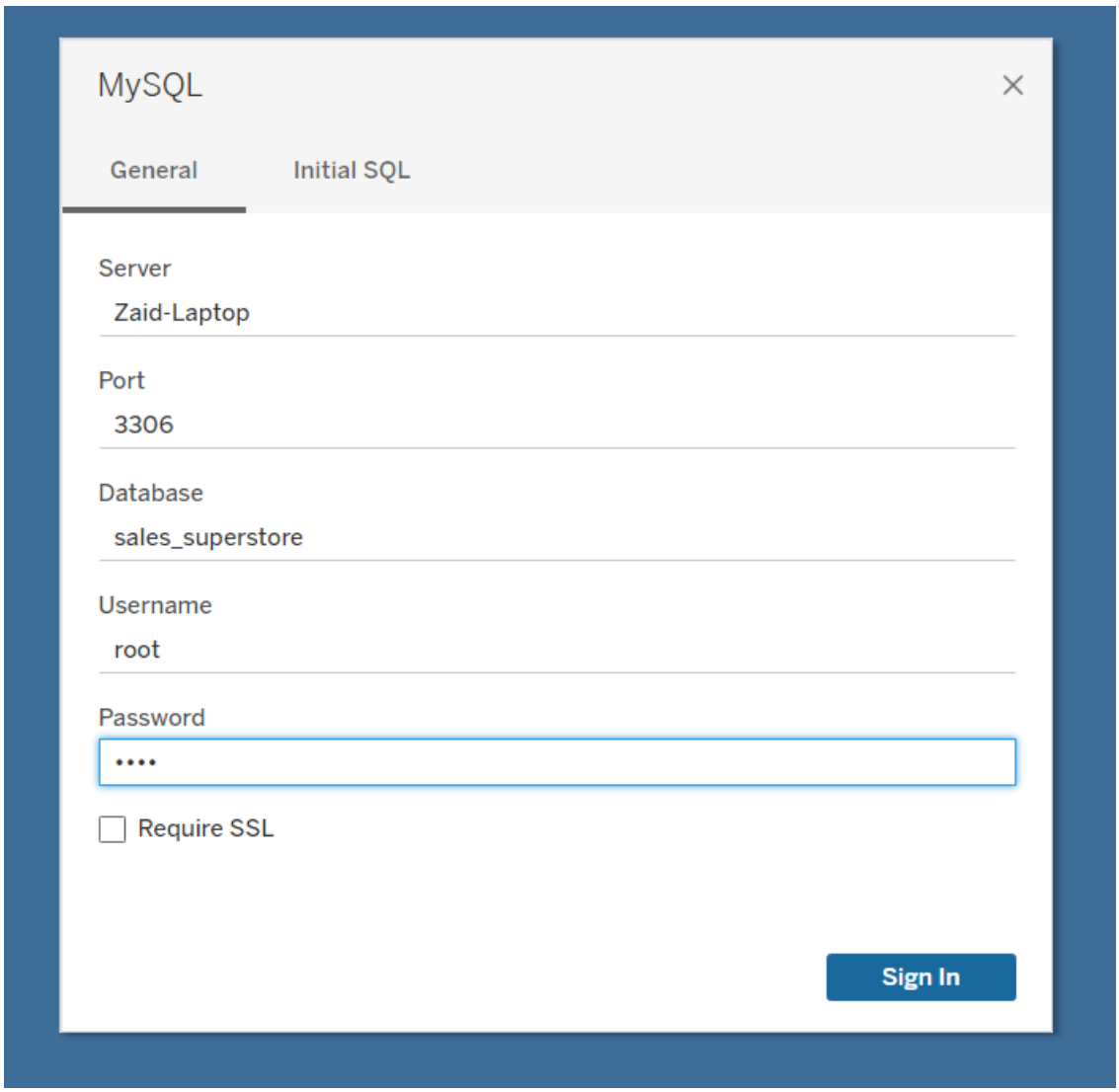
Step 5: After selecting the “MySQL” option, a pop-up window will open, asking you some basic information regarding the server for which you wish to connect.

Before entering the asked information, you would need to open the “MySQL Workbench” Application and start up the connection you wish to use in Tableau.

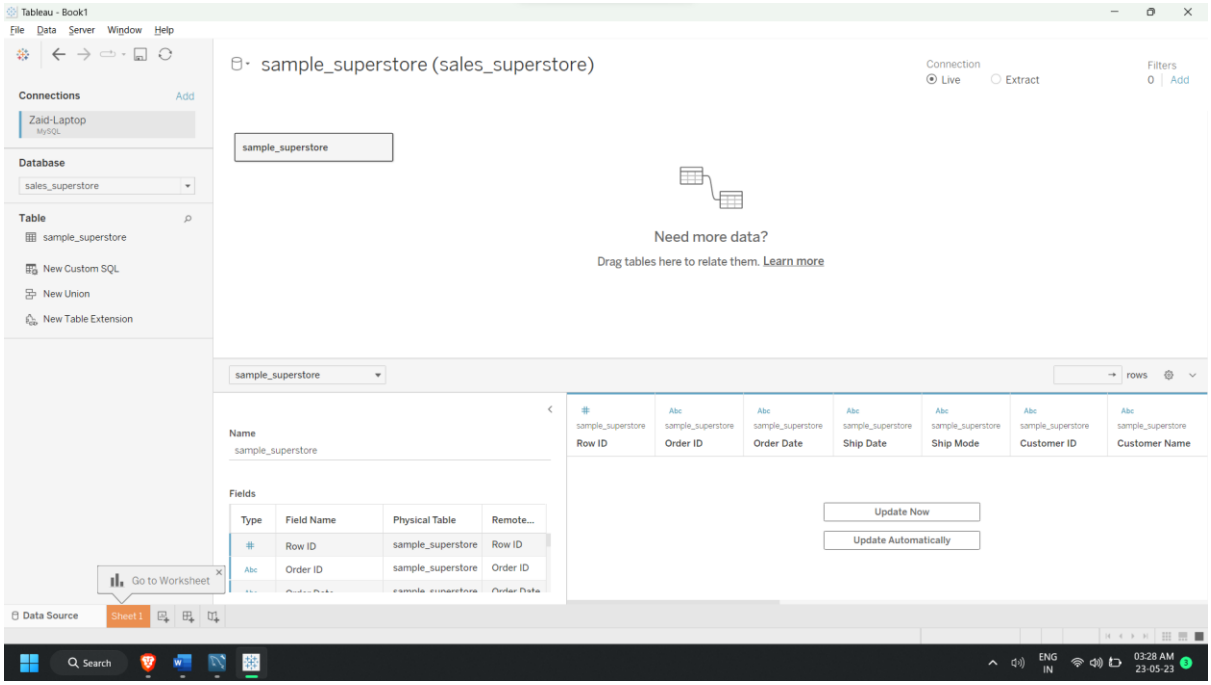
- Under the “Server” filed, you would need to type in the name of the connection you wish to connect, usually it will be “localhost” unless you configure the name of the connection while installing MySQL.
- Under the “Port” field, it will already be filled with “3306”, do not change this value until the specified value of port in your connection in MySQL Workbench is different.
- Under the “Database” field, this is an optional field, you can enter the database you wish to use after connection, this can be changed later to.
- Under the “Username” field, this is an optional field, you will need to enter the username which is used to enter the connection.
- Under the “Password”, this is an optional field, you will need to enter the password corresponding to the above mentioned “Username”.



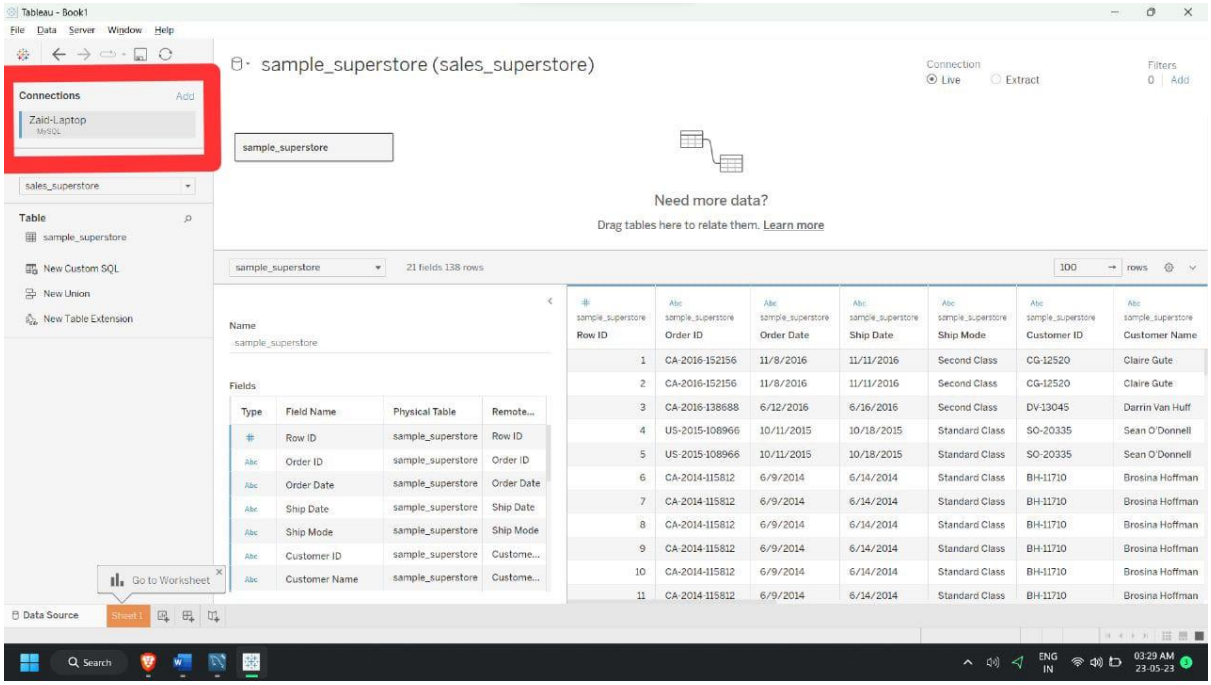
Step 6: After entering all the relevant details, you need to press the “Sign In” button, the connection process will take place, this will usually take a few seconds.



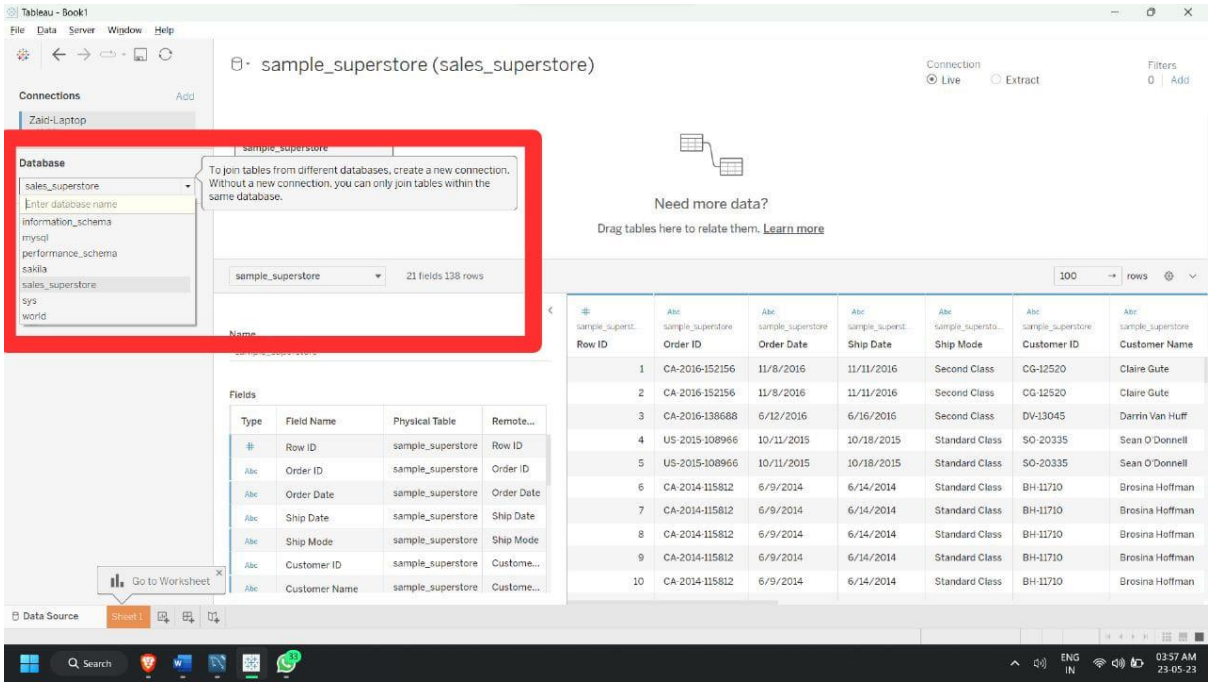
Step 7: If the connection is successful, you will be faced with the data source page that looks like the image below.



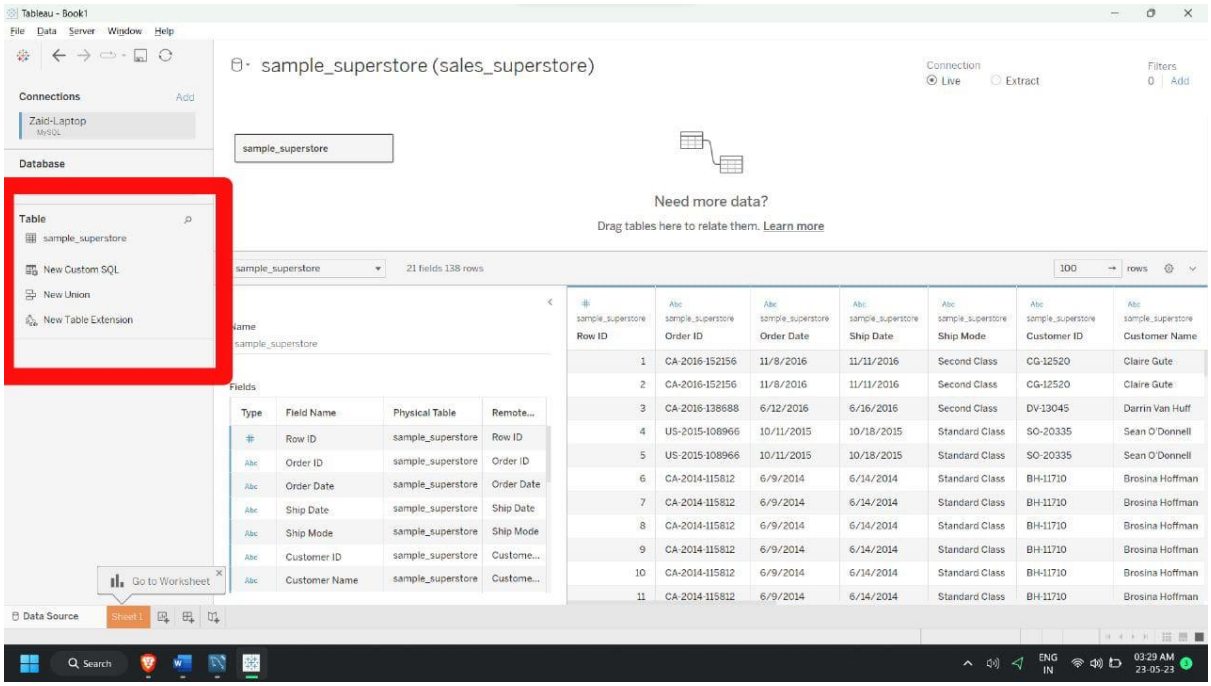
In the left side of the data source page there is side panel that has three sections, one of which is the “Connection” section, this indicates the server that you are connected to.



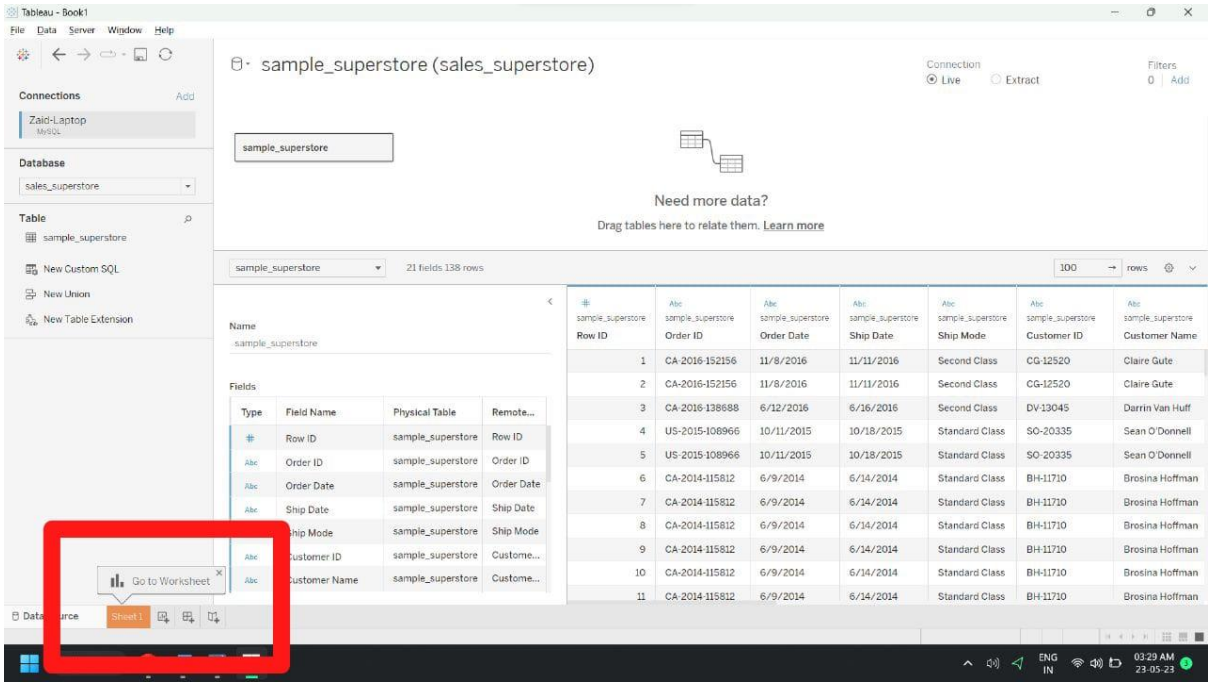
Below the “Connection” section is the “Database” section, this is a drop-down list, from which we can choose the various number of databases in that said connection, choosing one database will allow you to use the set tables available in that database.



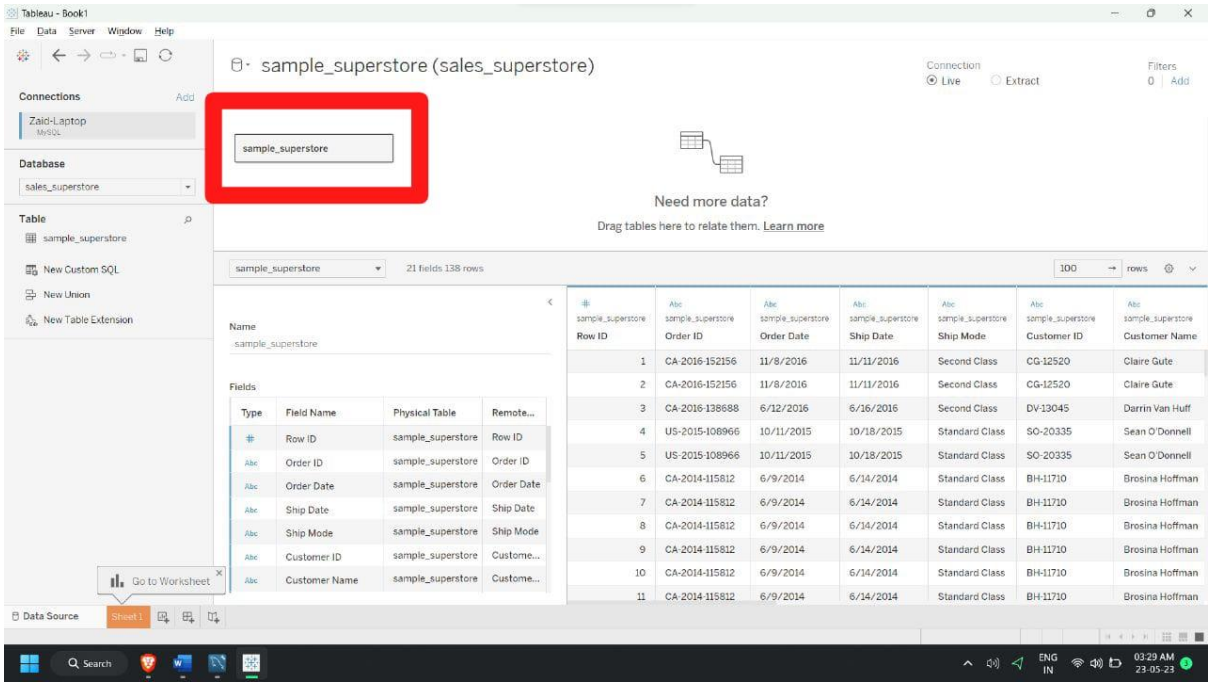
Below the “Database” section is the “Table” section, this section will contain all the tables that have been created in that database, you can choose the various tables available to use throughout Tableau.



In the bottom of the data source page is where you can create worksheets, when we establish the connection to a MySQL database, Tableau automatically creates a Workbook, a workbook is where you create all your visualizations specific to dataset, a workbook contains multiple worksheets, where each worksheet has its own visualization.



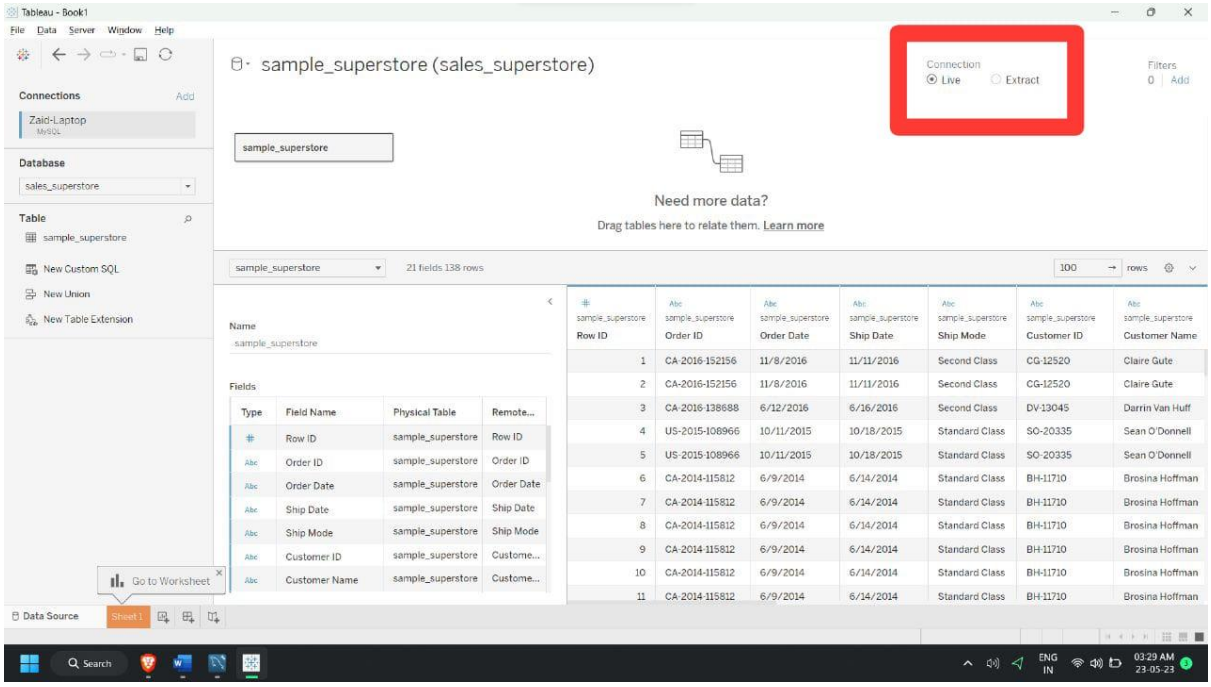
In the upper middle part of the data source page are tables active throughout your Tableau workbook, you can have multiple tables present with unique connections between each table.



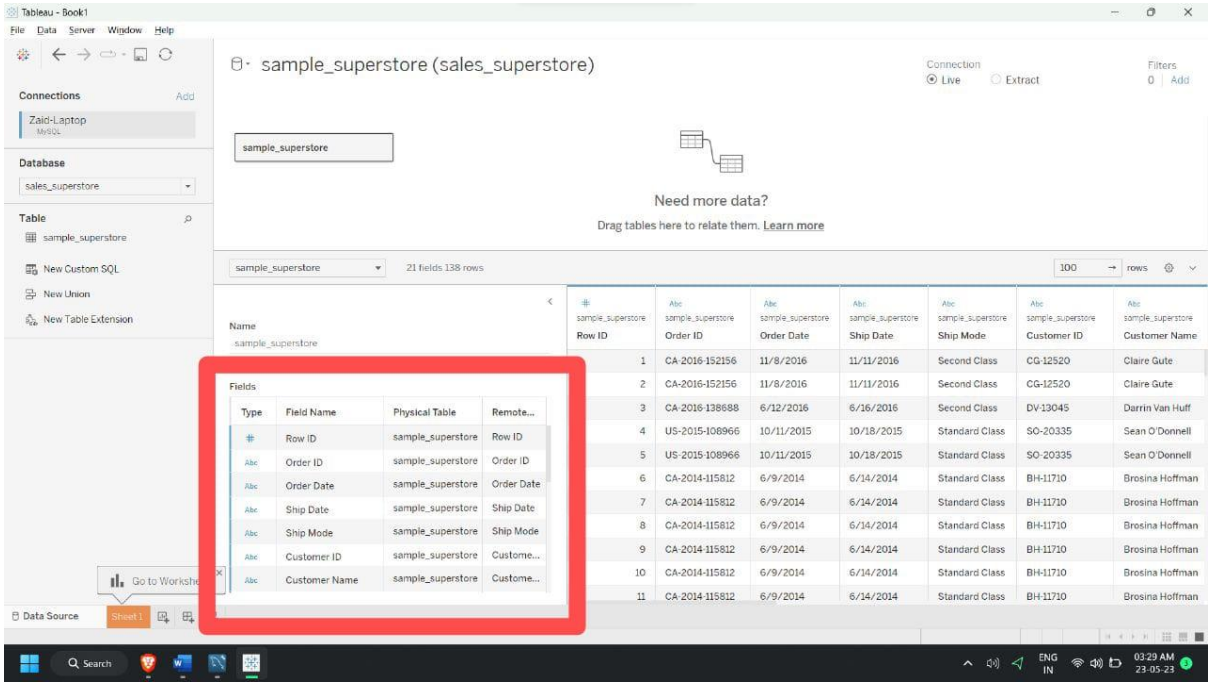
Towards the top right corner of the data source page are these two radio buttons under the title labelled “Connection”, these radio buttons signify the type of connection between you Tableau workbook and the MySQL.

There are two types:

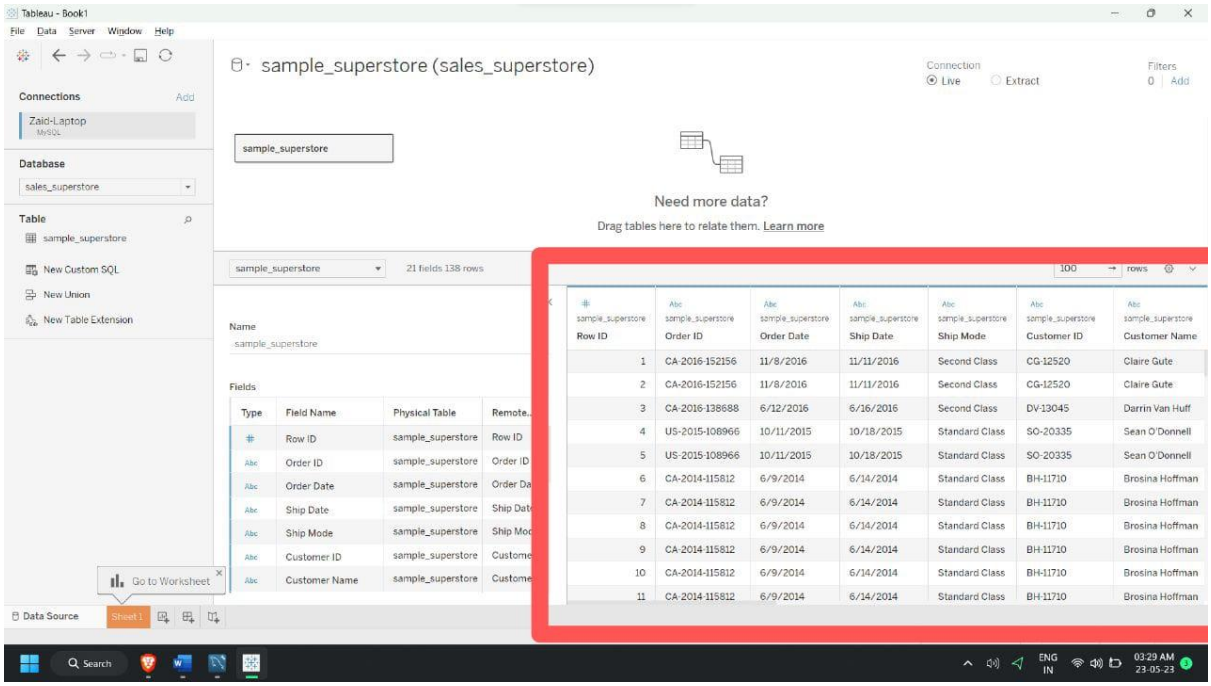
- Live: This type of connection is where the changes made to database in Tableau is directly reflected directly in the database itself.
- Extract: This type of connection is where the dataset that is being used from the MySQL database is locally extracted and stored within Tableau, that dataset is used for all the Tableau needs, so changes made in that dataset won’t be reflected to the MySQL database.



Towards the lower middle of the data source page is the “Fields” section, this section indicates the fields present in the tables that is active in you Tableau workbook.



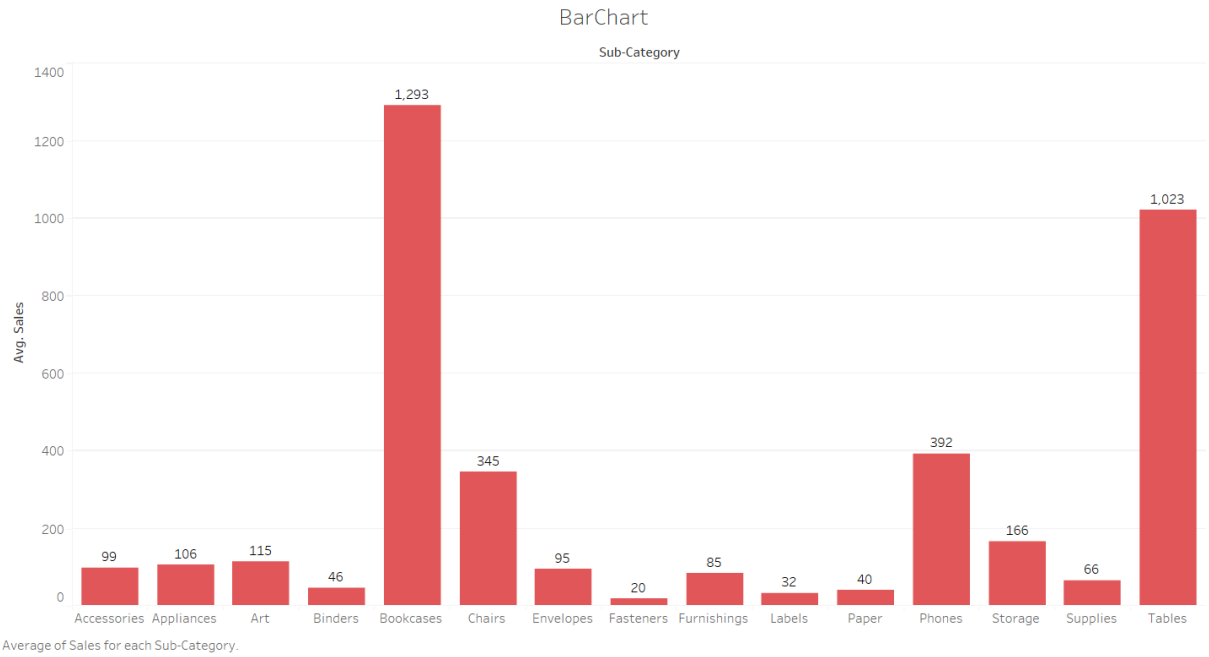
Towards the right of the “Fields” section is the actual preview of the dataset that is being used in your workbook, this contains all the records that is stored in the specified table.



Congratulations, you have successfully connected you MySQL database to Tableau.

Question 4

I. Bar chart

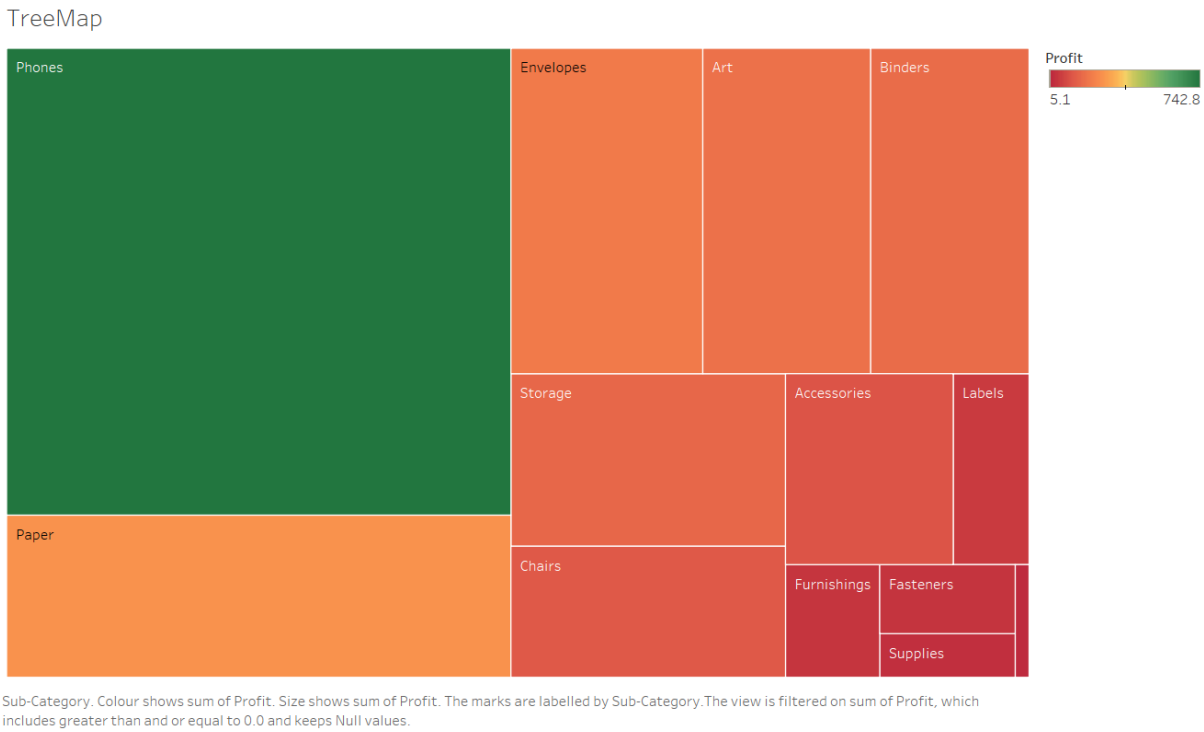


This a bar chart that is plotted with Sub-Category of products against average sales.

This chart shows us the average sales that took place in each sub-category.

from this chart we can say that 4 categories are having higher average sales, Book Cases, Chairs, Phones and Tables.

II. Tree map



This is a Tree map, that shows the sales(profit) of each subcategory of products.

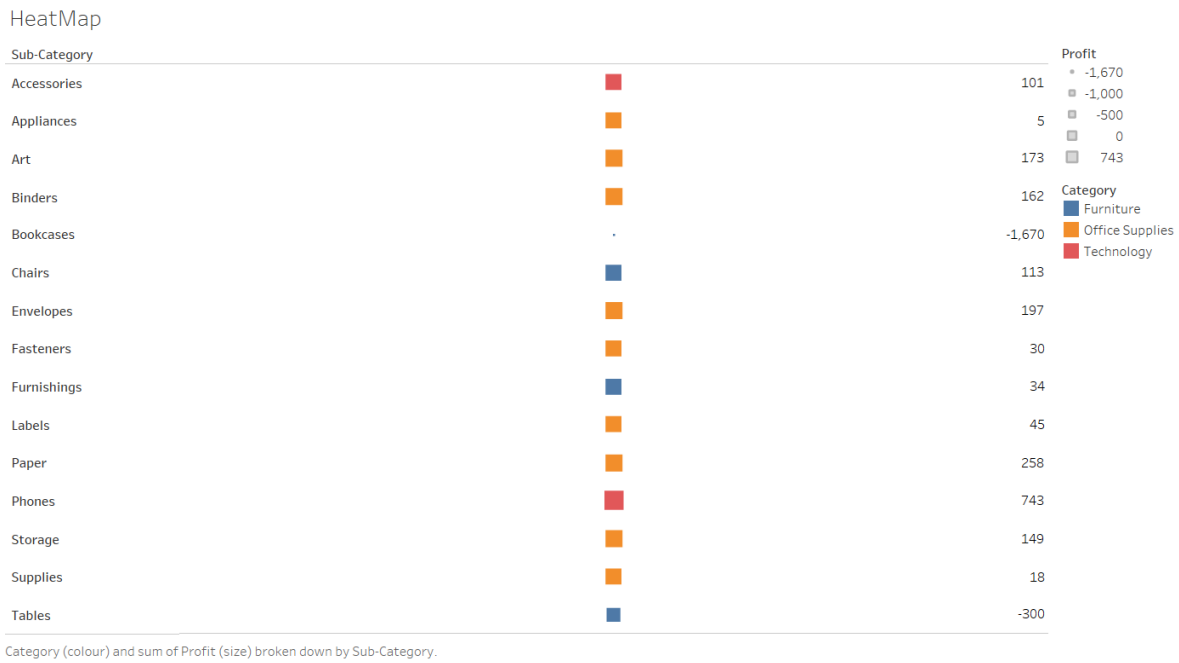
There are two components to this map that is important:

- Size
- Colour

The larger the size of each section, higher the sales, lower the size, lower the sales.

From the legend, we can figure out that if each section ranges from the colour, green to red,
If a section is closer to green the section has higher sales(profit),
If a section is closer to red the section has lower sales(profit).

III. Heatmap



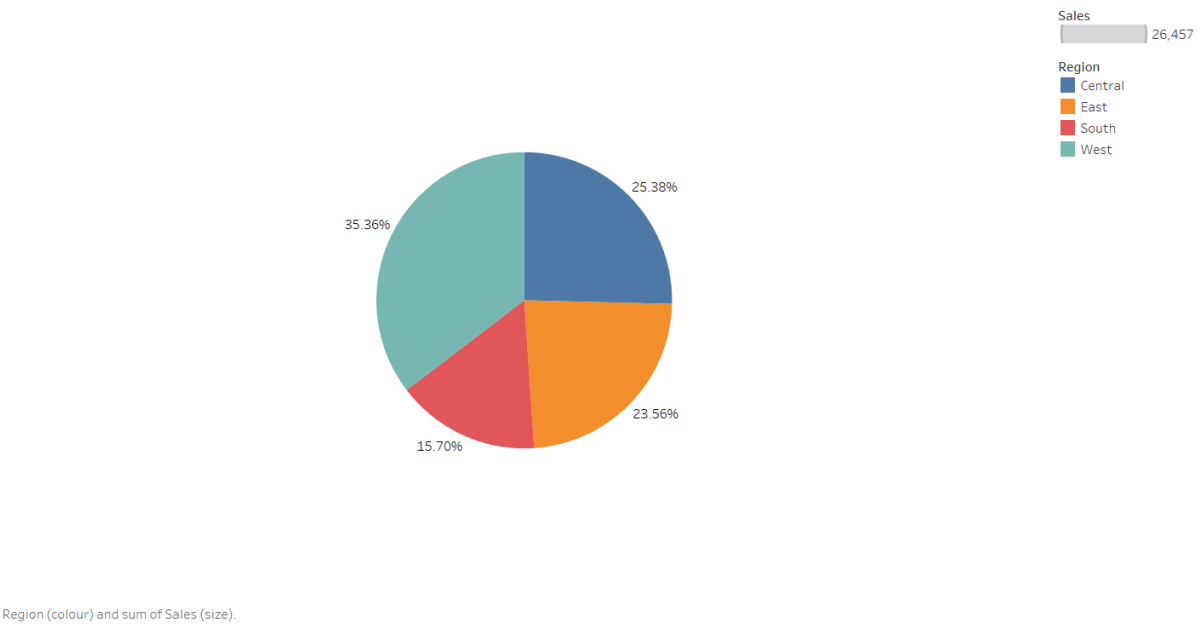
This chart is plotted with subcategory against sales(profit), further more each subcategory is colour coded with its respective category.

Sales of each subcategory is represented by the size of the square. Each subcategory belongs to a certain category, which is distinguished by the colour of the square.

From this chart we can see that, there are more number of subcategories within Office supplies than any other categories,

IV. Pie chart

PieChart

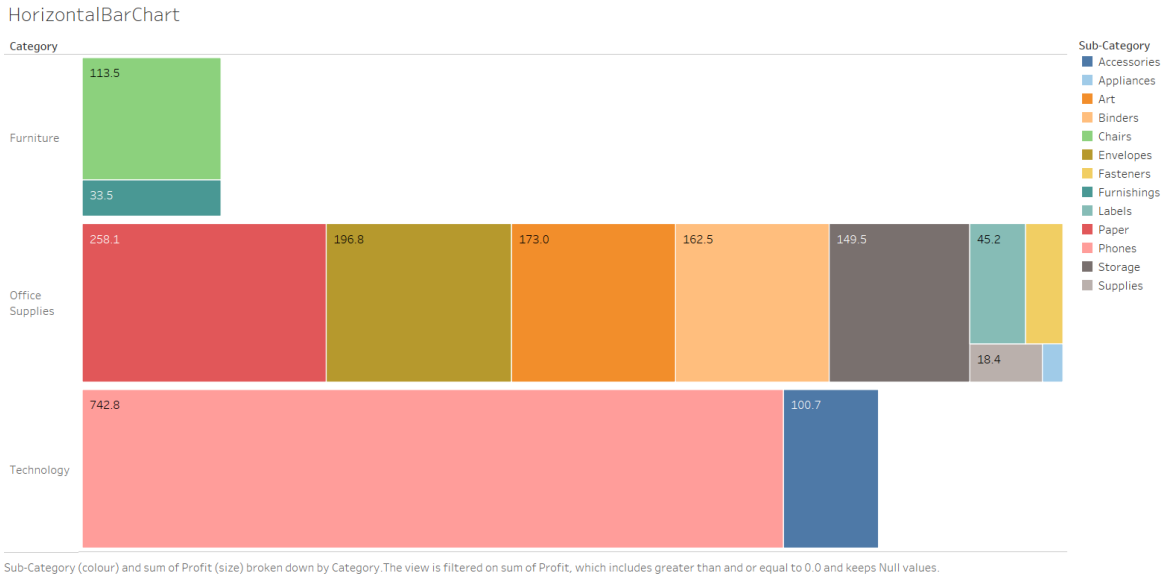


Region (colour) and sum of Sales (size).

This chart represents the Sales (profit) that has been made by the organisation specific to each region across the country.

We can see from the graph that the west region constituted to more sales than any other region, central and east region branches have almost similar sales, finally the south branches of the organisation is the one with least sales.

V. Horizontal bar chart



This graph represents the Sales(profit) made by each subcategory segregated by category.

Each horizontal line represents each category, then each square/rectangle section represents the respective sub-categories within each category which is color coded.

The size of each section in a line depends on the sales(profit) made by each subcategory.

- From the graph we can make out that the furniture category has the least sales(profit) out of all the other categories.
- Even though the technology category has the same number of subcategories as furniture category, the sales(profit) made by the technology category is similar to office supplies category.