CMT426/CMM426/CMT423 –Business Intelligence and Analytics

School of Computer Sciences

Academic Session 2022/2023

Semester II

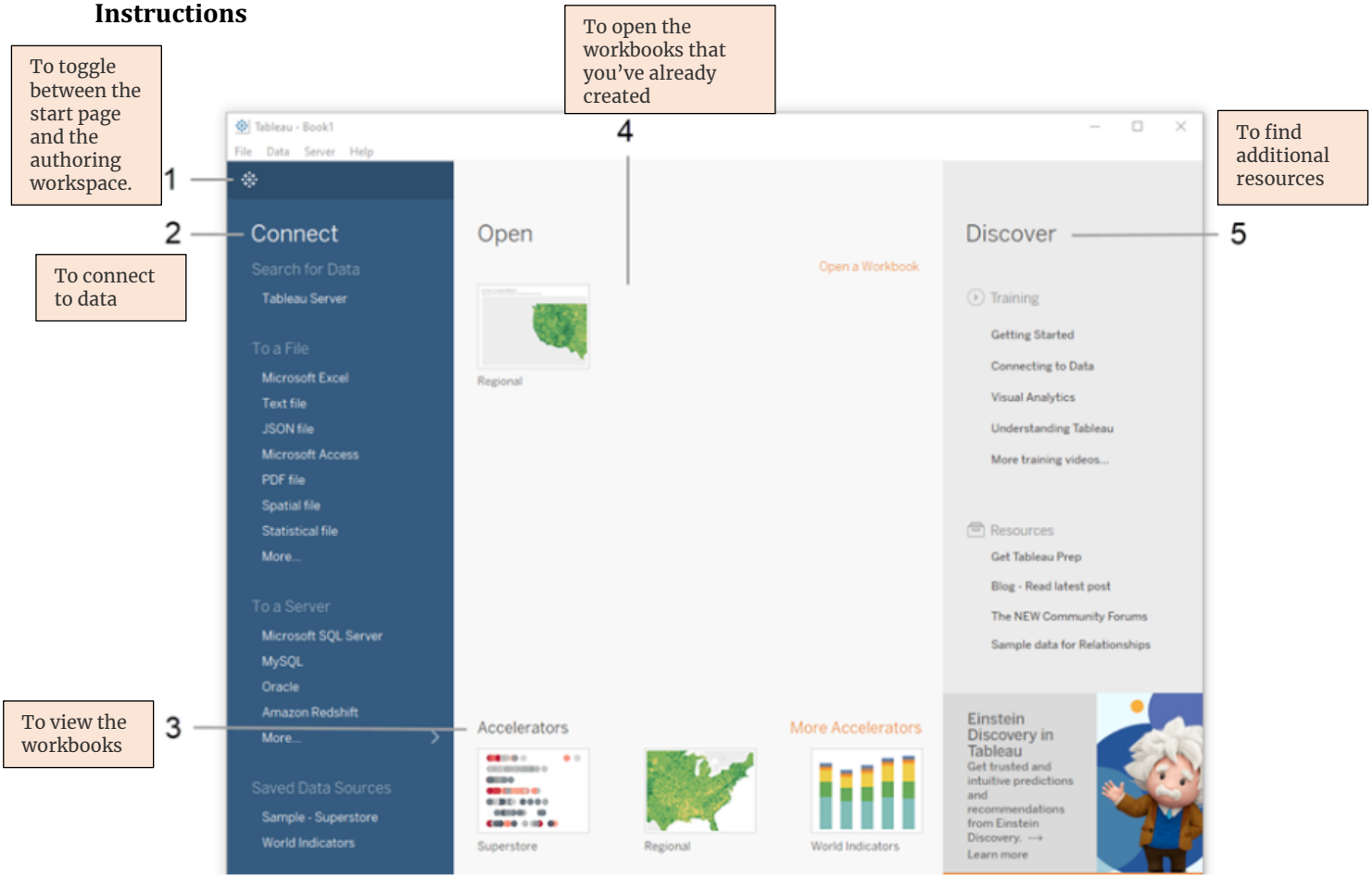
Lab 1

Date: 18 April 2023
Place: Online Webex
Provided by: Dr Noor Farizah Ibrahim

Description

What is Tableau? Tableau is a Data Visualisation tool widely used for Business Intelligence. It helps create interactive graphs and charts in the form of dashboards and worksheets to gain business insights. And all of this is made possible with gestures as simple as drag and drop!

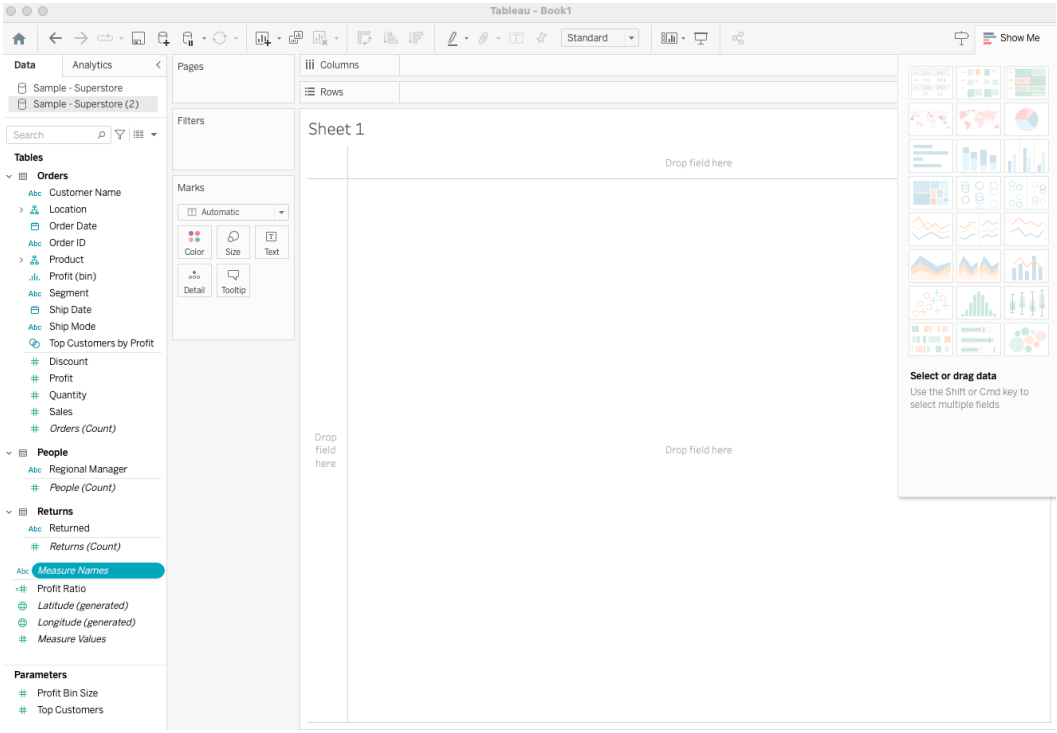
Instructions



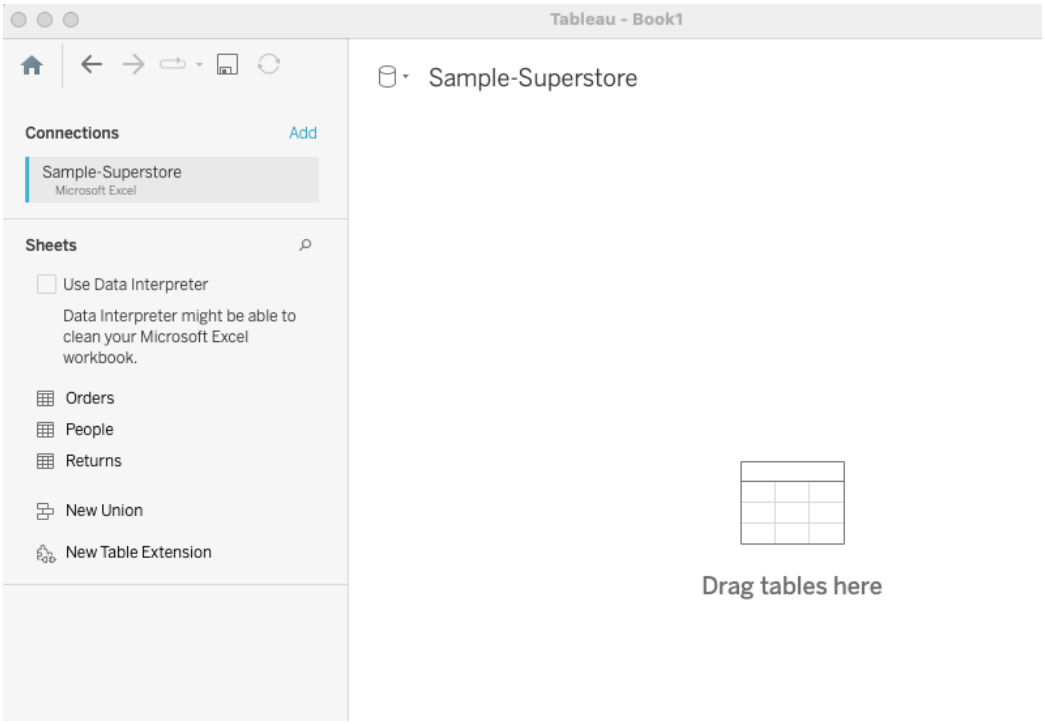
Let's get started!

1. You can import your data in a flat file such as Excel or CSV or directly load it from data servers. For this tutorial, we'll use an example workbook from Tableau called **Superstore.** The data is about United States' Superstore which is deliberating over its expansion. It wishes to know the prospective regions of the country where it could and hence requires your help. There are two ways to get the sample workbook:
    a. Use **Superstore** sample data from the Tableau server by clicking 'Sample-Superstore' under the 'Saved Data Source' panel.
    OR
    b. Go to eLearn@USM and download **Superstore.xls**

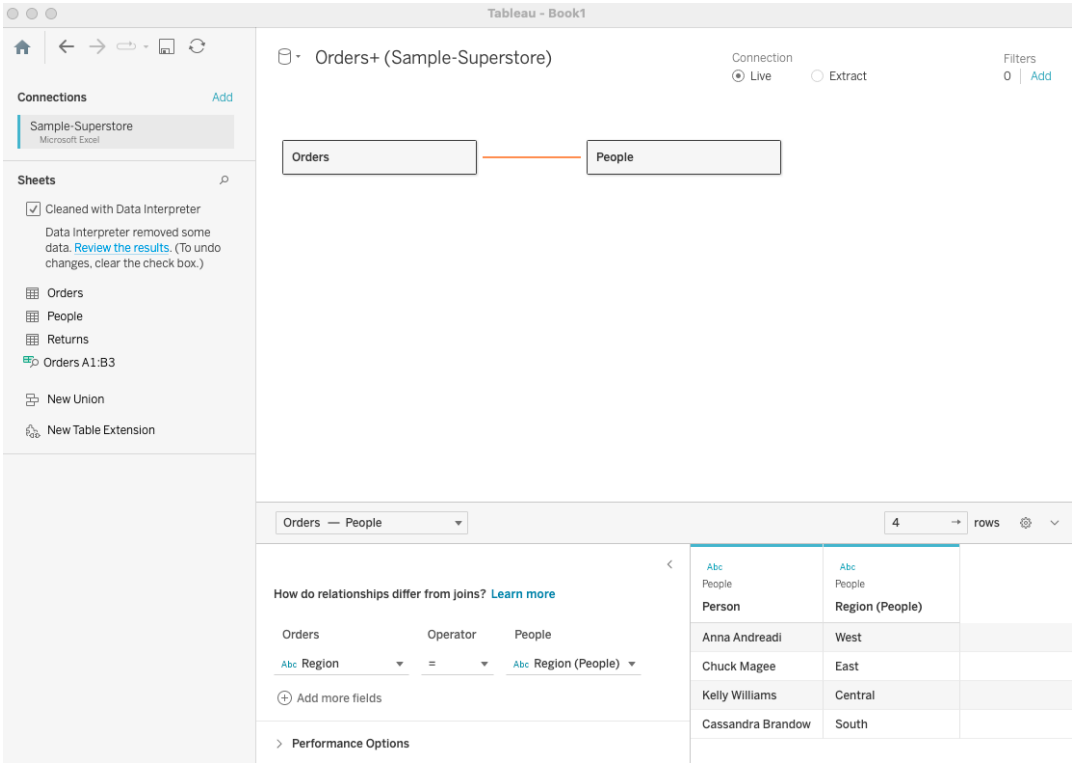
    #Hover around and look at the colour – blue (discrete) and green (continuous)
2. If you use data from Tableau server, click on the data and you will see something like this:



3. To upload your Excel file: Connect panel – > To a File – > Microsoft Excel – > Go to your location of Samplestore.xls, and you'll get:



4. Drag **Order** table to the sheet to view the table and then, click on **Data Interpreter** to clean your data. Then, drag **People** table onto the same sheet and you'll see the connection between these two tables through **Region** variable/feature/column

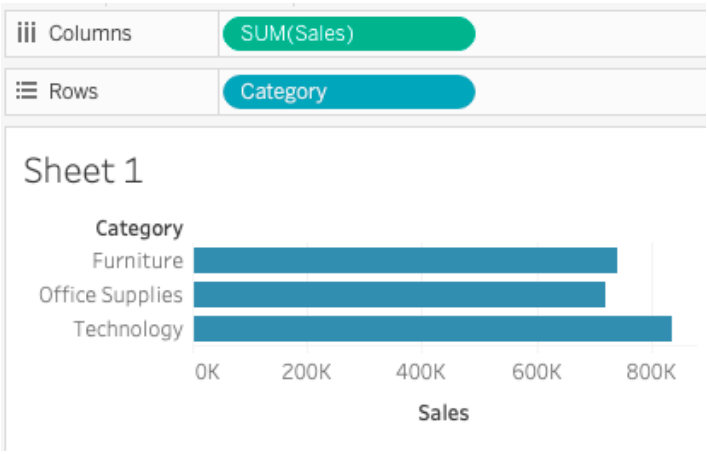


Now you can start plotting the data,

5. Firstly, look at different data type under **Orders** table.

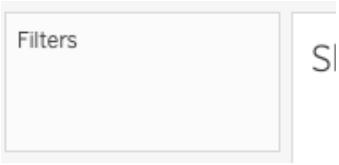
First plot (worksheet):

6. To plot the data, under Sheet 1 tab, take two different items: Drag the **Product category** onto the row shelf and **Sales** onto the column shelf.

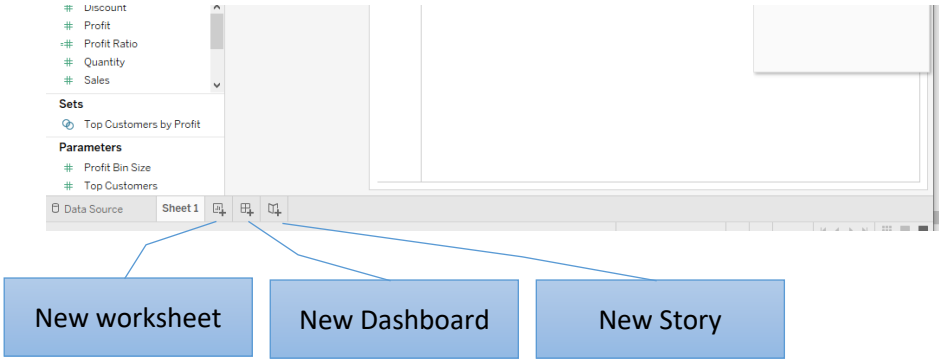


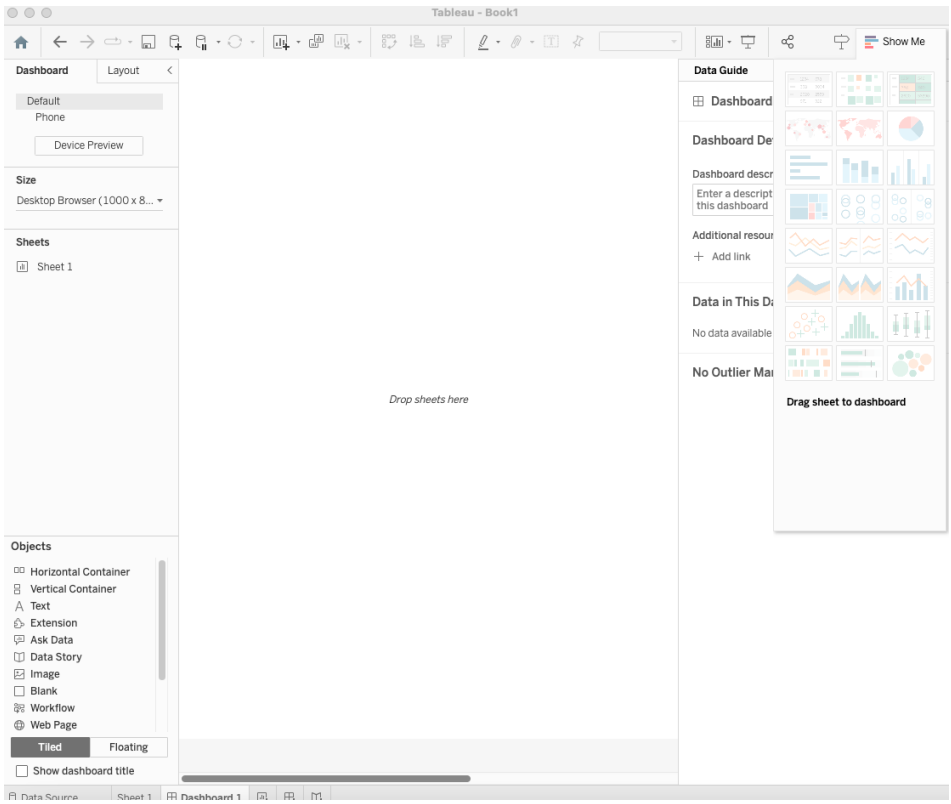
7. **Show Me function, the brain of Tableau!:** When you drag and drop fields onto the visualisation area, Tableau makes default graphs for you, as we shall see soon, but you can change these by referring to the Show Me option. You can play around and choose the ideal chart or graph suitable to your data.

8. **Filter function:** Filters help you view a strained version of your data. For example, instead of seeing the combined Sales of all the Categories, you can look at a specific one, such as just Furniture. Apply a filter to look at one specific region. Right-click on the filter shelf to filter according to region.



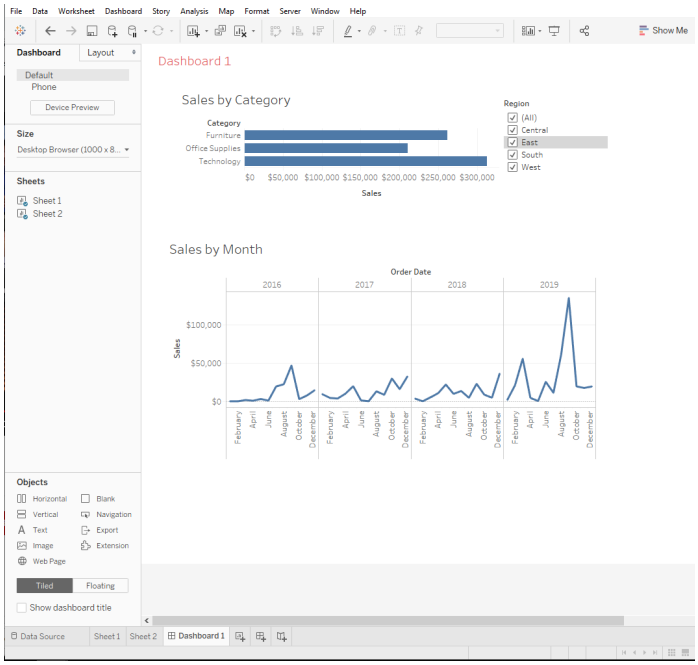
Dashboard: Now it's time to create a simple dashboard.





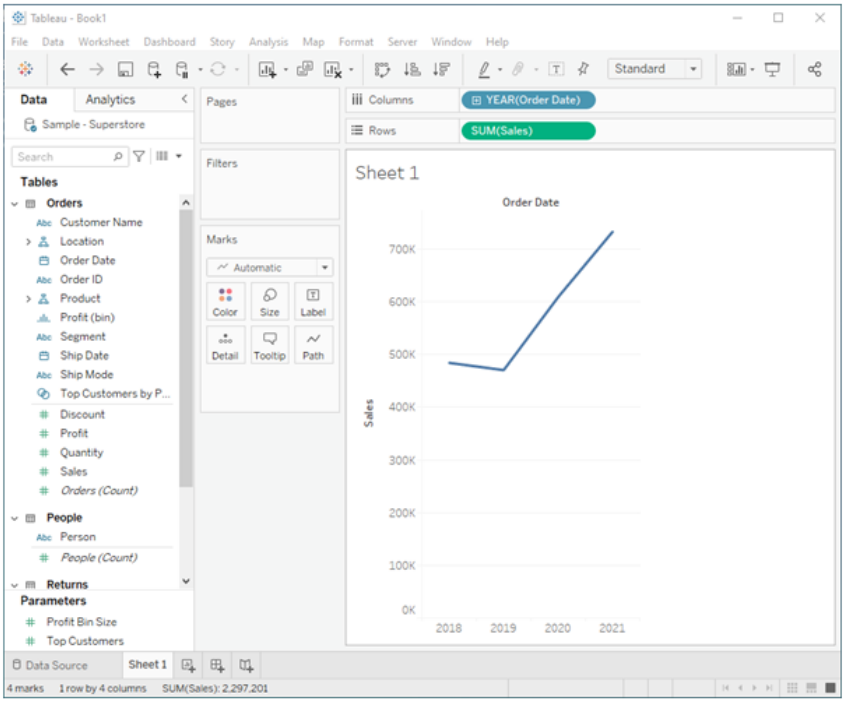
Dashboard panel

9. Drag the sheet created to the dashboard. To get a floating layout, hold **drag + shift** key to lock the sheet. Design your dashboard on floating layout.



<u>Second plot:</u>

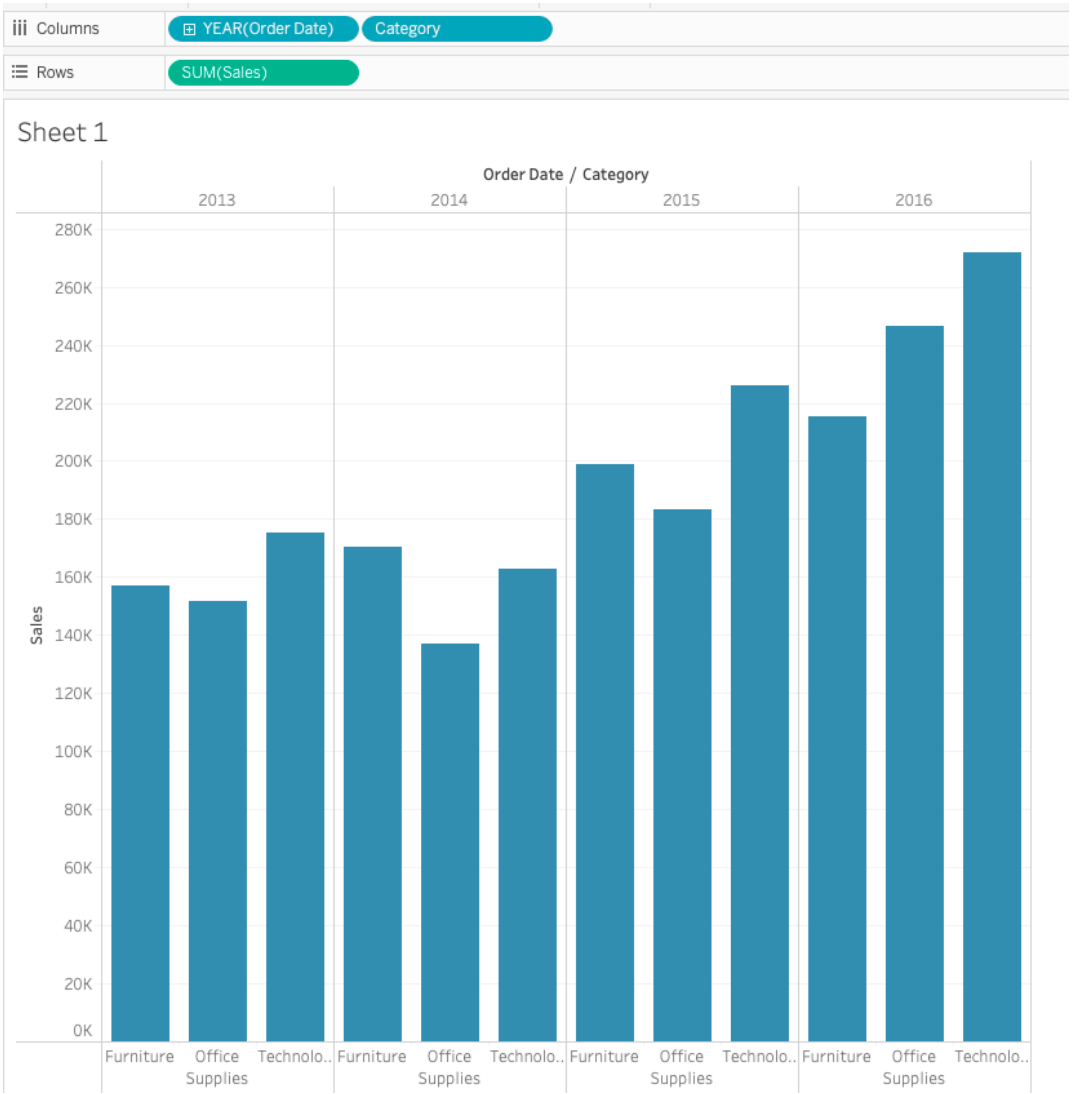
In a new sheet, drag the **Order Date** onto columns shelf and **Sales** onto rows shelf. Observe the change in charts.



When you first create a view that includes time (in this case Order Date), Tableau automatically generates a line chart.

This line chart shows that sales look pretty good and seem to be increasing over time. This is good information, but it doesn't really tell you much about which products have the strongest sales and if there are some products that might be performing better than others.

10. From the Data pane, drag **Category** to the Columns shelf and place it to the right of **YEAR (Order Date).**

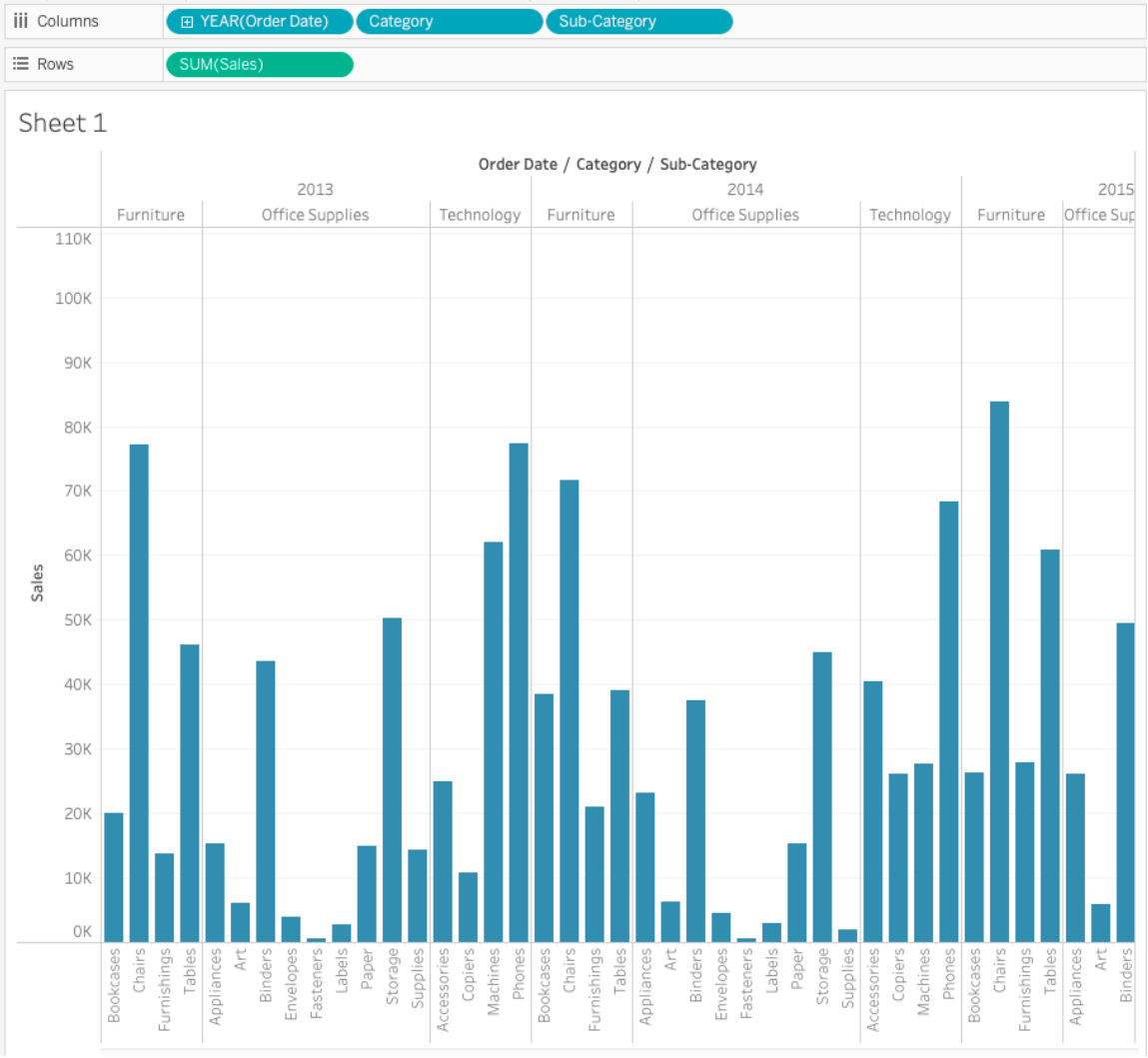


From this view, you can see that sales for furniture are growing faster than sales for office supplies, even though Office Supplies had a really good year in 2021.
Insights: Perhaps you can recommend that your company focus sales efforts on furniture instead of office supplies? Your company sells a lot of different products in those categories, so you'll need more information before you can make a recommendation.

To help answer that question, you decide to look at products by sub-category to see which items are the big sellers. For example, for the **Furniture** category, you want to see details

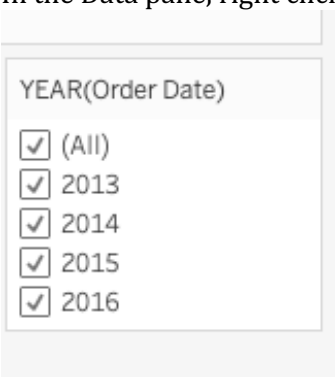
about bookcases, chairs, furnishings, and tables. Looking at this data might help you gain insights into sales and later on, overall profitability, so add sub-categories to your bar chart.

11. Double-click or drag **Sub-Category** to the Columns shelf.

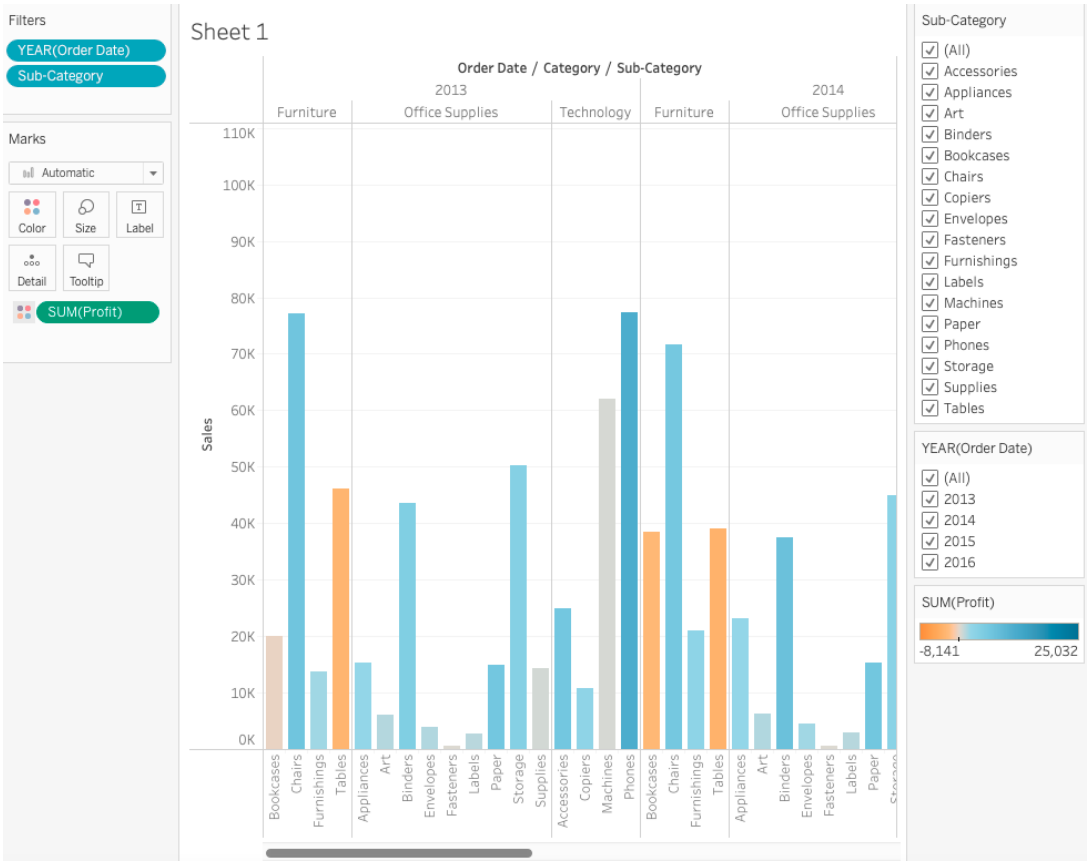


<u>Add filters and colours</u>

12. In the Data pane, right click **Order Date** and select **Show Filter**



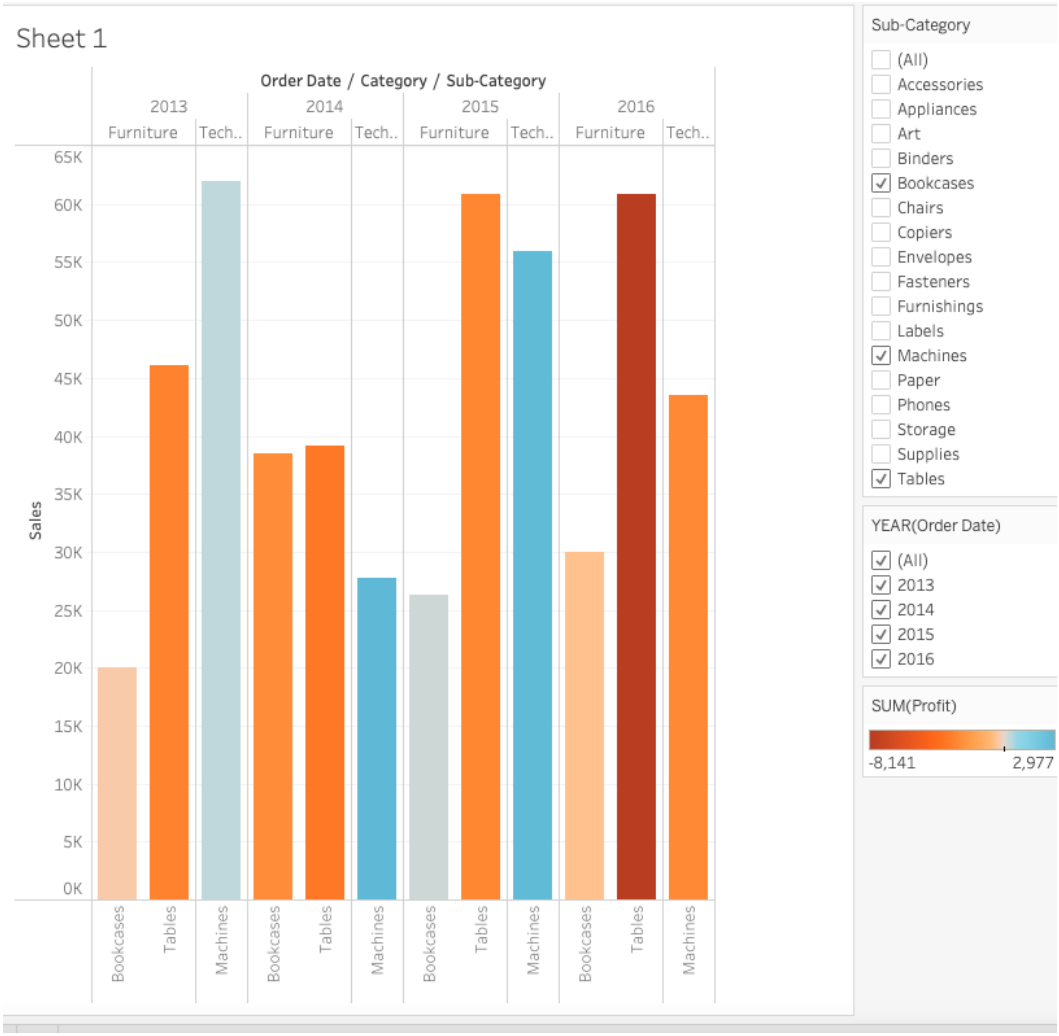
13. Repeat the step above for the **Sub-Category**
14. From the **Data** pane, drag **Profit** to **Color** on the Marks card.



By dragging profit to color, you now see that you have negative profit in Tables, Bookcases, and even Machines. Another insight is revealed!

Find key insights

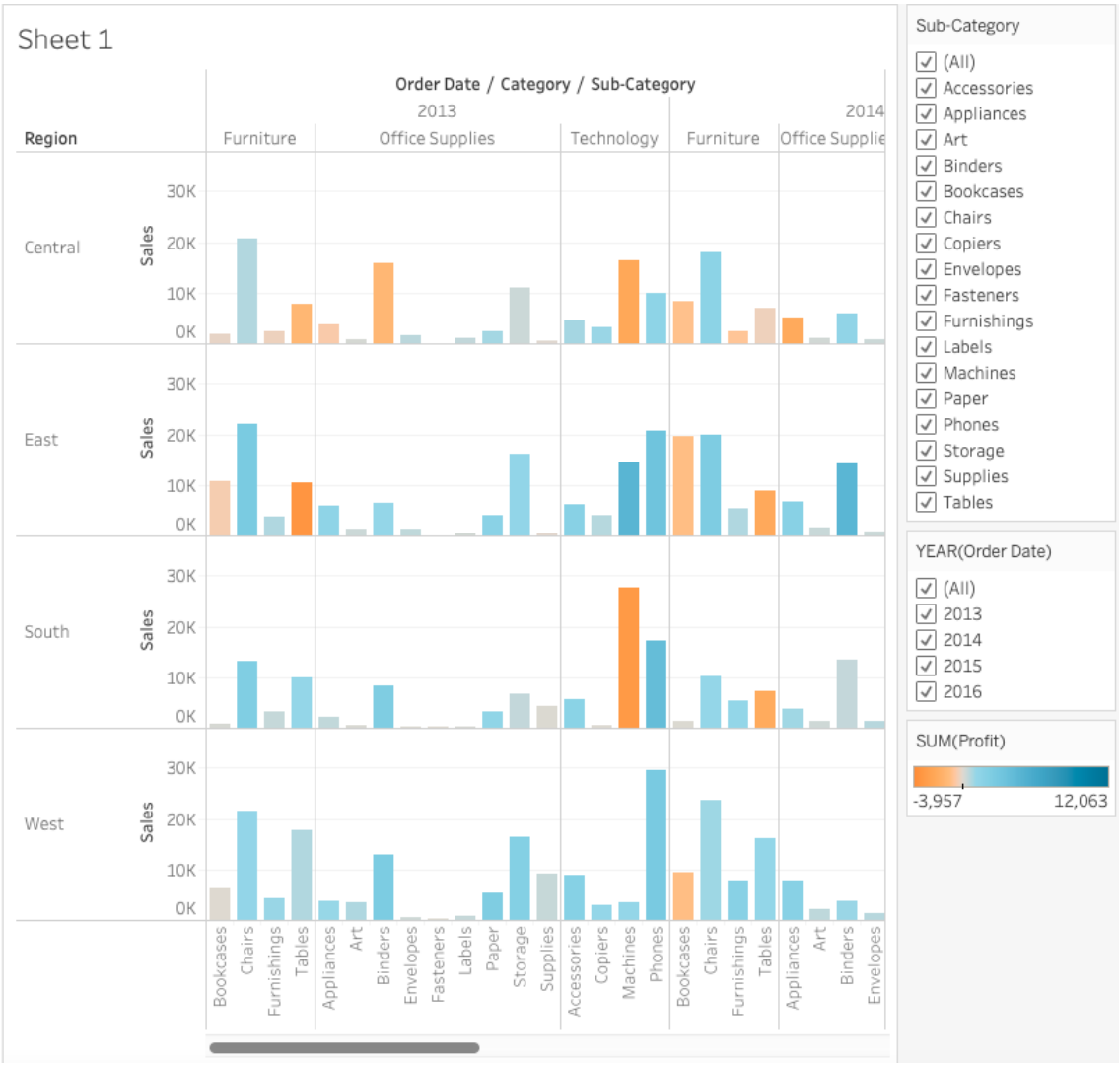
15. Looking at your view, you saw that you had some unprofitable products, but now you want to see if these products have been unprofitable year over year.
16. In the view, in the **Sub-Category** filter card, clear all of the check boxes except Bookcases, Machines, and Tables.



17. Now you can see that, in some years, Bookcases and Machines were profitable. However, recently Machines are unprofitable. While you've made an important discovery, you want to gather more information before proposing any action items to your boss.

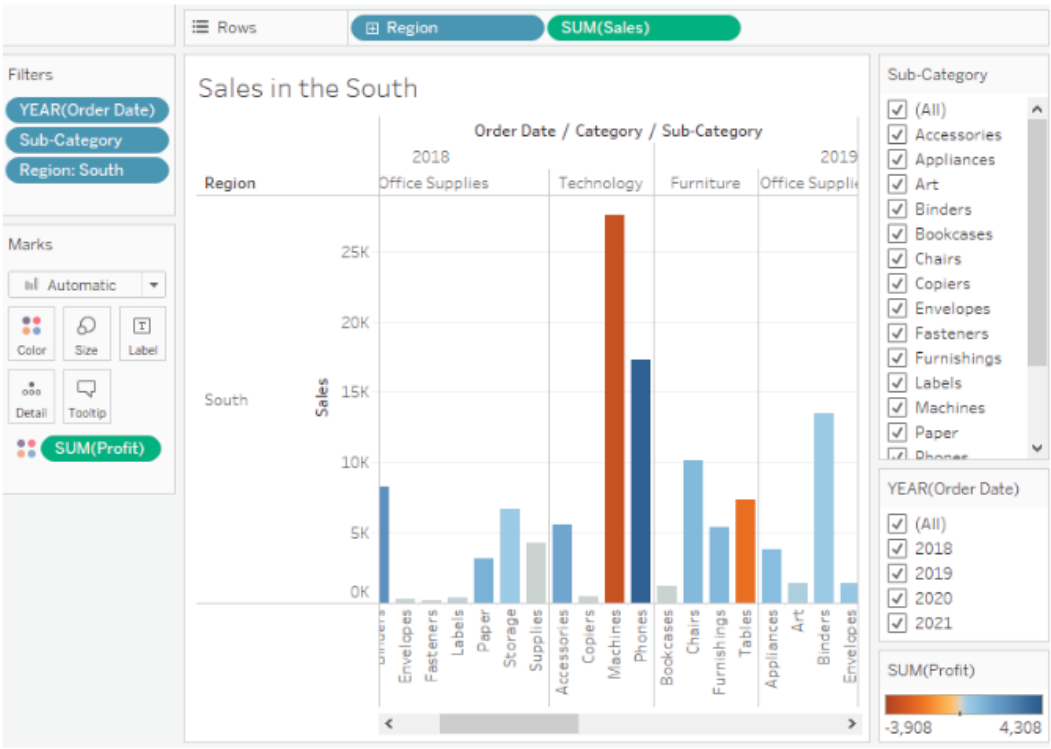
Break up your view by region:

18. Select All in the **Sub-Category** filter card to show all sub-categories again.
19. From the Data pane, drag **Region** to the Rows shelf and place it to the left of Sum(Sales).
20. Tableau creates a view with multiple axes broken down by region as below.



Try filter the Sub-Category for **Machines** only, you notice that machines in the South are reporting a higher negative profit overall than in your other regions. You've discovered a hidden insight!

21. Select All in the **Sub-Category** filter card (if you changed your filter) to show all sub-categories again, name the worksheet, and add a title.
22. At the bottom-left of the workspace, double-click Sheet 1 and type **Sales by Product/Region**.
23. In your workbook, right-click the Sales by Product/Region sheet and select Duplicate.
24. Rename the duplicated sheet to **Sales in the South**.
25. In your new worksheet, from the Data pane, drag **Region** to the Filters shelf to add it as a filter in the view.
26. In the Filter **Region** dialog box, clear all check boxes except South and then click OK.



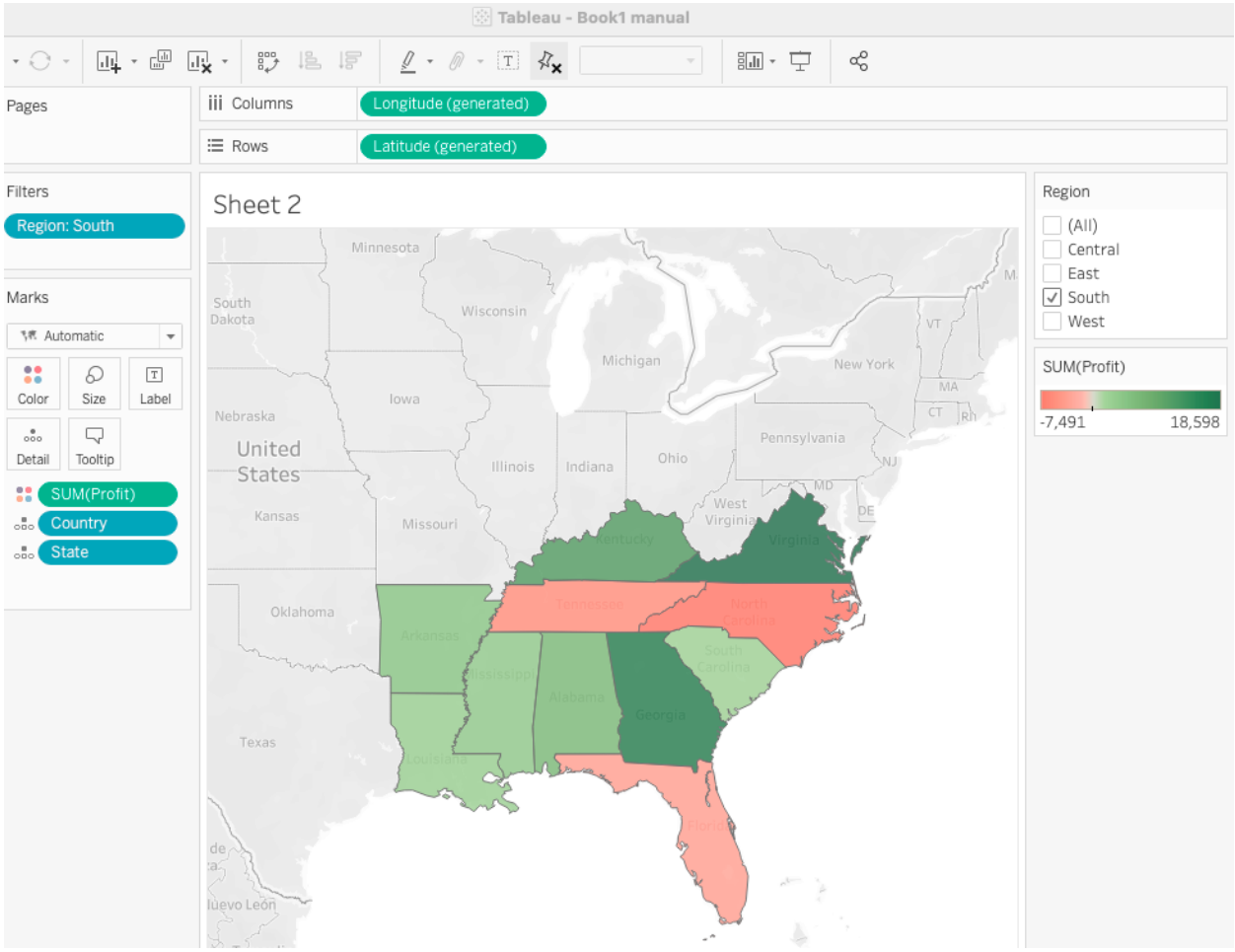
Now you can focus on sales and profit in the South. You immediately see that machine sales had negative profit in 2018 and again in 2021. This is definitely something to investigate!

27. Save your work by selecting **File > Save As**. Give your workbook a name, such as Regional Sales and Profits.

Level 2

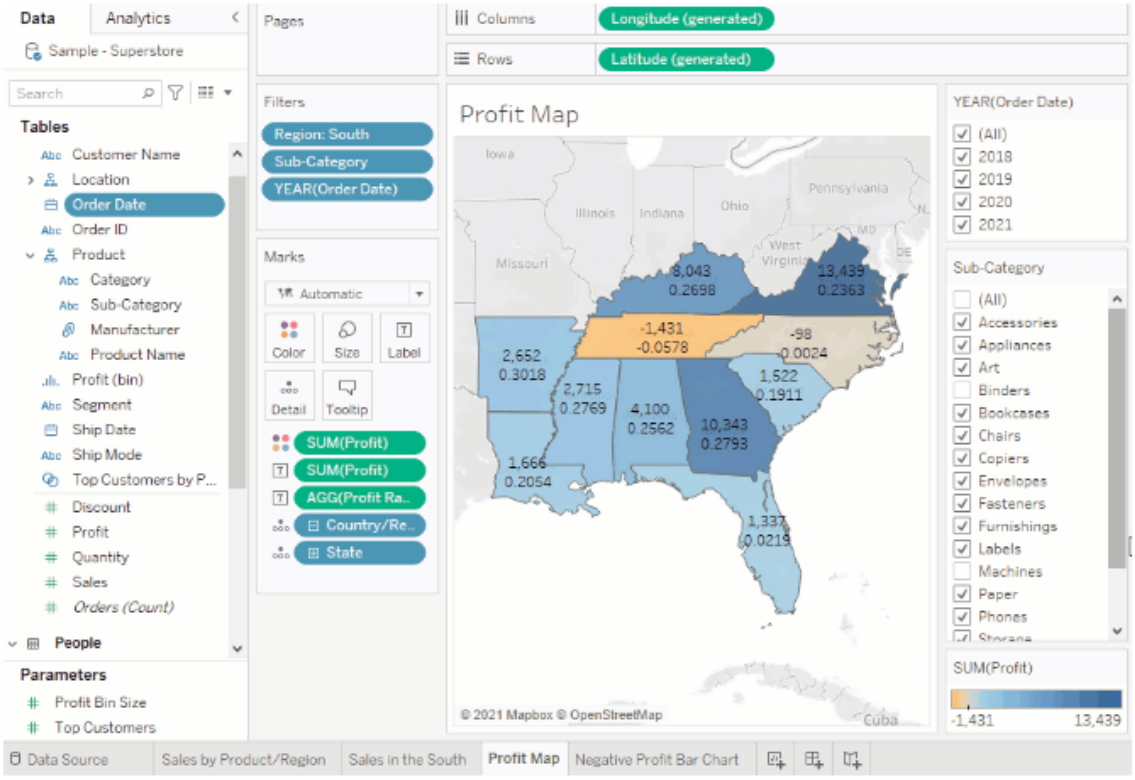
1. In new worksheet in the **Data** pane, double-click **State** to add it to **Detail** on the Marks card. You can see the map now

2. Drag **Region** to the Filters shelf, and then filter down to the **South** only. The map view zooms in to the South region, and there is a mark for each state (11 total).
3. Drag the **Sales** measure to Color on the Marks card.
4. Add **Country** field is also added to Marks card.
5. Continue drag/drop until you get the same screen as below

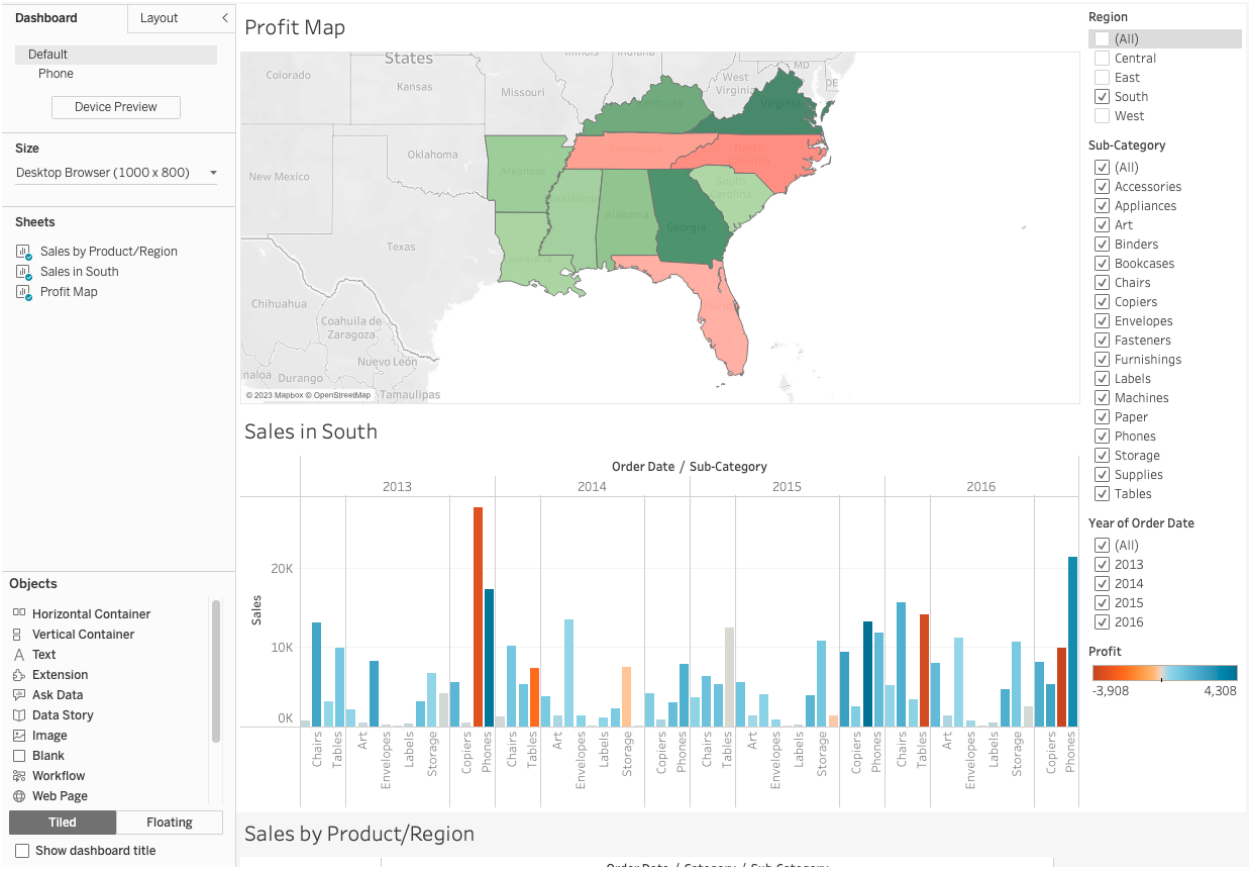


6. To edit color: Click Color on the Marks card and select Edit Colors.
7. In the Palette drop-down list, select **Red-Green Diverging** and click OK. This allows you to see quickly the low performers and the high performers. In this example, you can see which states are doing well, and which states are doing poorly in sales.

You can also try to build the worksheet as shown below



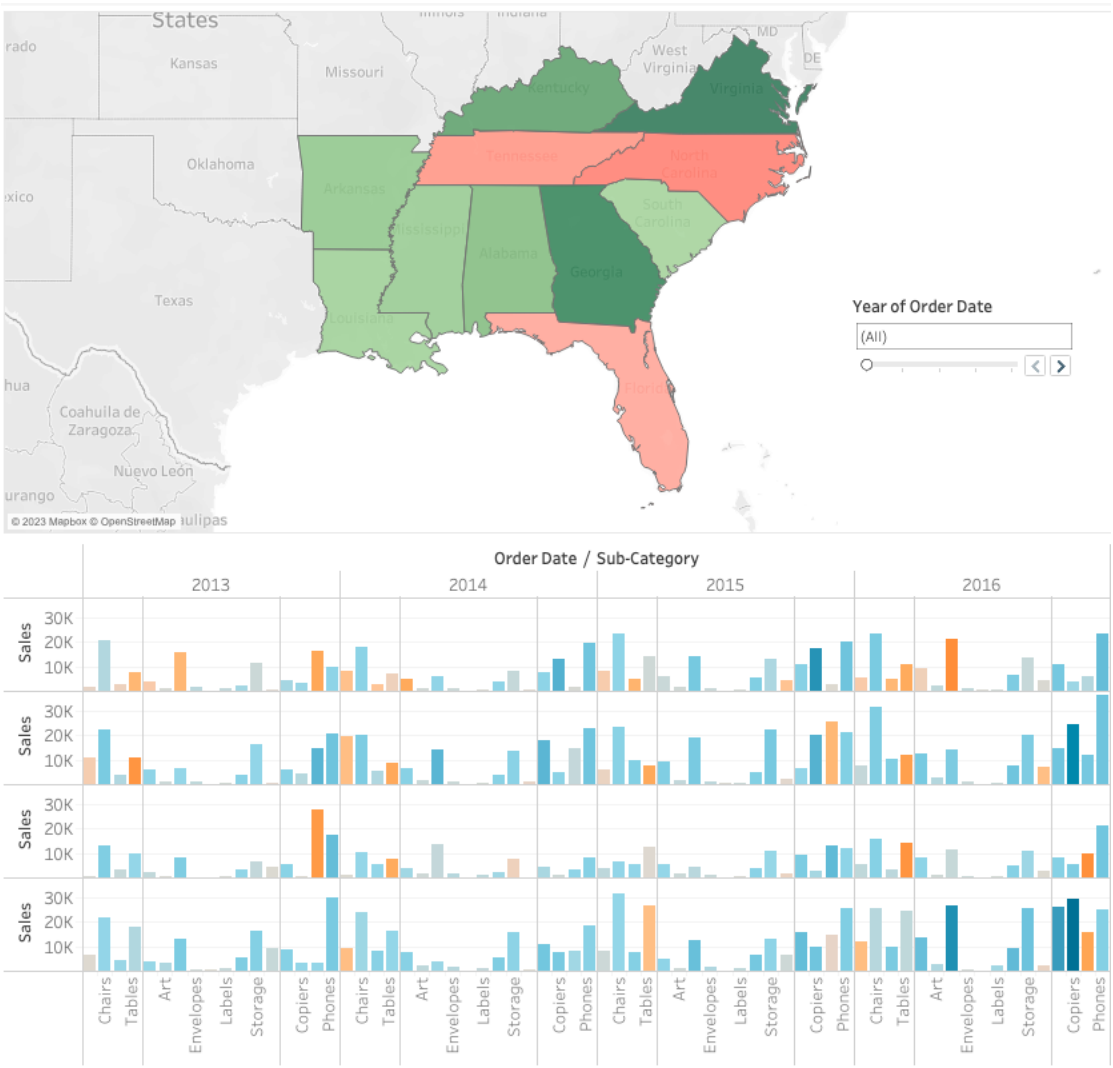
Now, create a new dashboard using those worksheets that you have created.



Make-up to tidy up the Dashboard.

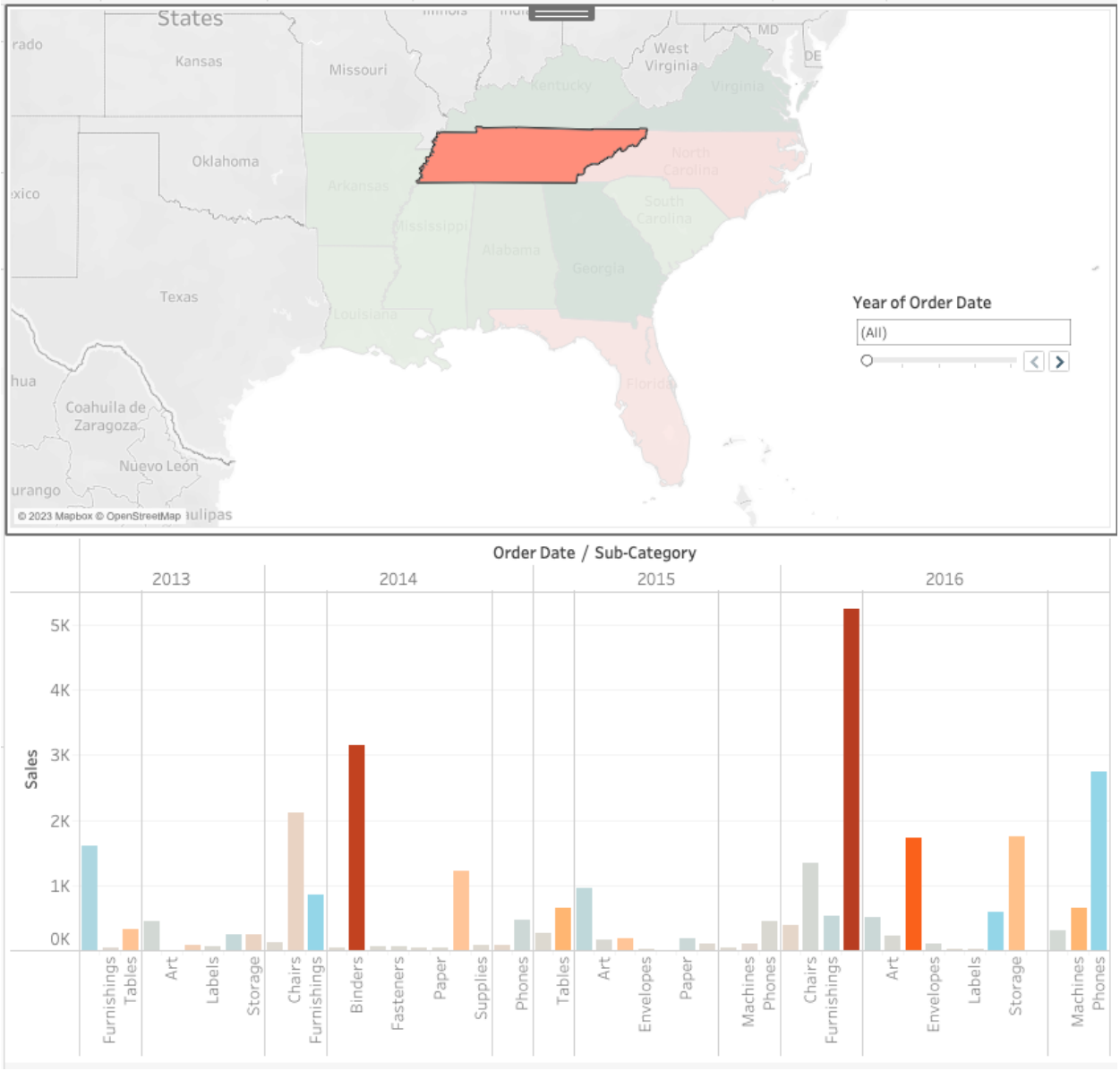
1. Drag all the related worksheets to the dashboard page as the example above.
2. Right-click the **Profit Map** title and select **Hide Title** and repeat.
3. Select the first **Sub-Category** filter card on the right side of your view, and at the top of the card, click the Remove icon X From Dashboard icon.
4. Click the drop-down arrow at the top of the Year of Order Date filter, and select Single Value (Slider).
5. In the Palette drop-down list, select Red-Green Diverging and click OK. This allows you to see quickly the low performers and the high performers.
6. Finally, select the remaining Year of Order Date filter, click its drop-down arrow, and then select **Floating.** Move it to the white space in the map view.
7. Try selecting different years on the Year of Order Date filter. Your data is quickly filtered to show that state performance varies year by year. That's nice, but it could be made even easier to compare.

Your dashboard should look like this.



8. Select **Profit Map** in the dashboard, and click the **Use as filter** icon  in the upper right corner.
9. To make it **interactive**, Select the **Year of Order Date** filter, click its drop-down arrow, and select Apply to Worksheets > All using this data source

Your dashboard should look like this.



10. Rename the dashboard to **Regional Sales and Profit**, and you do the same by double-clicking the Dashboard 1 tab and typing Regional Sales and Profit.

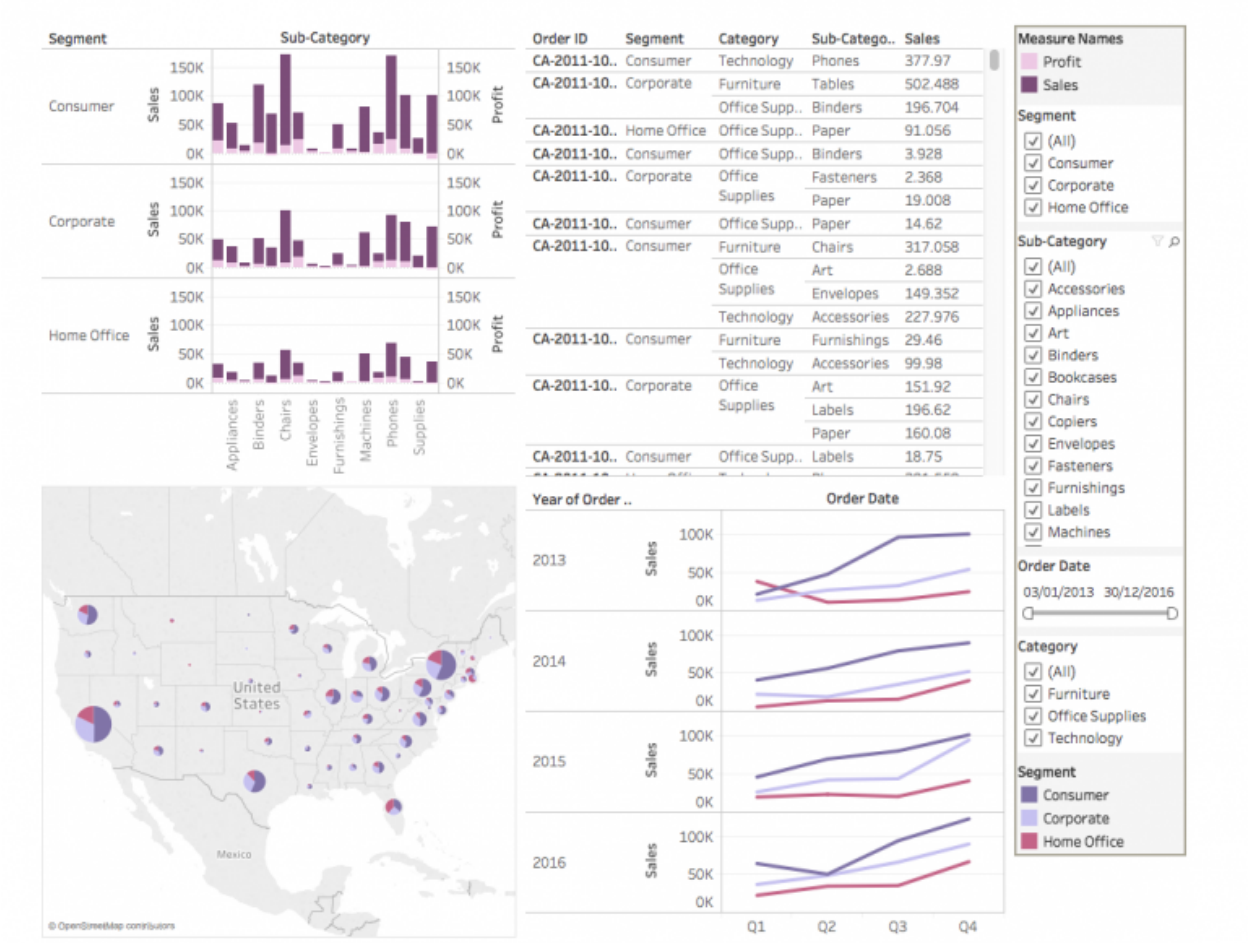
The end..

Reference:

https://help.tableau.com/current/guides/get-started-tutorial/en-us/get-started-tutorial-connect.htm

Task: Individual Lab Report (25 marks)

1. Using the same data, develop and design your second dashboard like the sample below.



2. Use your own data (if you have any) **OR** publicly available data **OR** use one of the data sets that comes with the tool (such tools usually have one or more data sets for demonstration purposes). Study the data and come up with:
    a. Simple business problems
    b. Use data visualization to analyze, visualize, and potentially solve those problems (using Tableau)
    c. Summarise what can you observe from the visualisation and how it helps to solve those problems stated in (a).

The end….

Reference: Textbook

Submission:

Online submission at eLearn@USM: Saturday (29 April 2023), before 12 a.m (midnight).

Submission: **Tableau file (for dashboard), dataset** and **report in PDF format**

Late submissions will be penalized.

CO-PO Mapping: Note: You will construct models relevant to business intelligence and analytics using appropriate technology and software (CO2, PO2).

Read Me: Policy

- All assignments and lab exercises MUST be submitted **before or on** the specified date. Late submissions without any reasons and without permission from the lecturer(s) will not be accepted. The grade for late submission (even with permission) will be reduced as determined by the lecturer(s).
- Students who copied or **plagiarized** other's work or let their work be copied or plagiarized will be given F grade for the work, test or the whole coursework component as determined by the lecturer(s). The said student may be barred from sitting for final exam and reported to the university's disciplinary board.
- **Plagiarism** (using other people's ideas and text without proper acknowledgment and using them as your own) is a serious academic offence. The consequences for plagiarism are severe.