CSCE 5320 Scientific Data Visualization

Activity 3

Task 1:

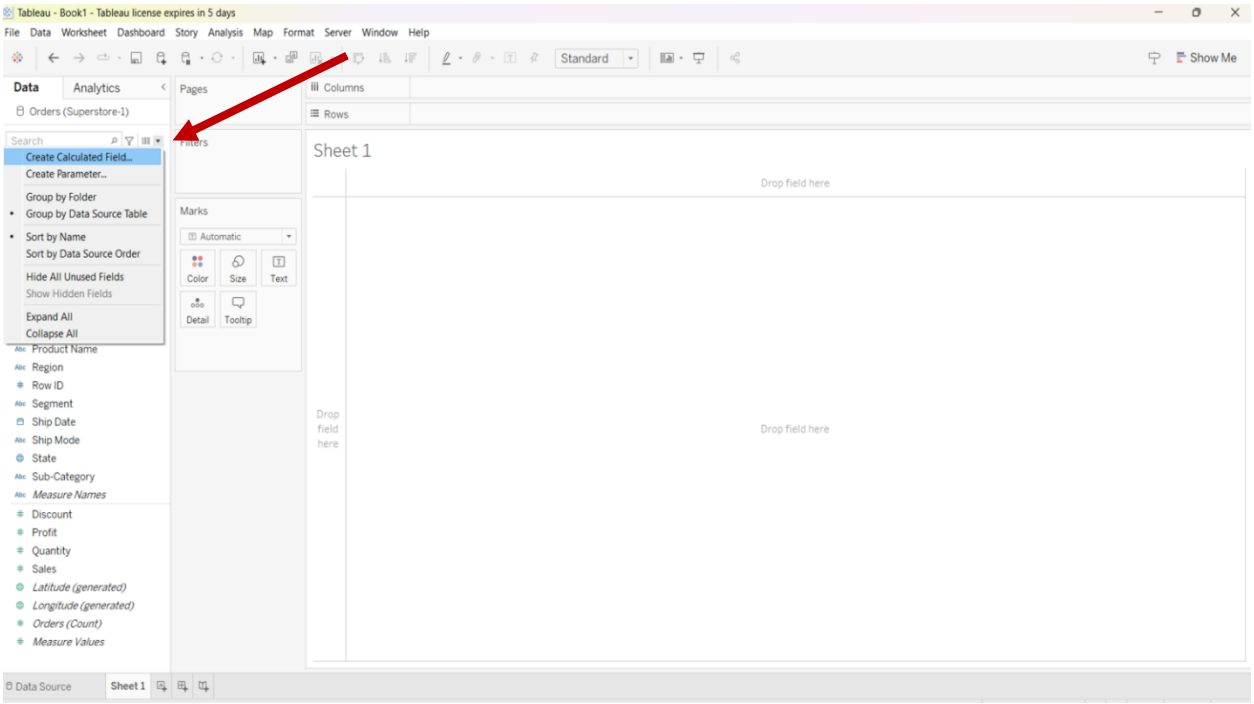
Tutorial 1

Create and Use the data calculated fields:

Step 1:

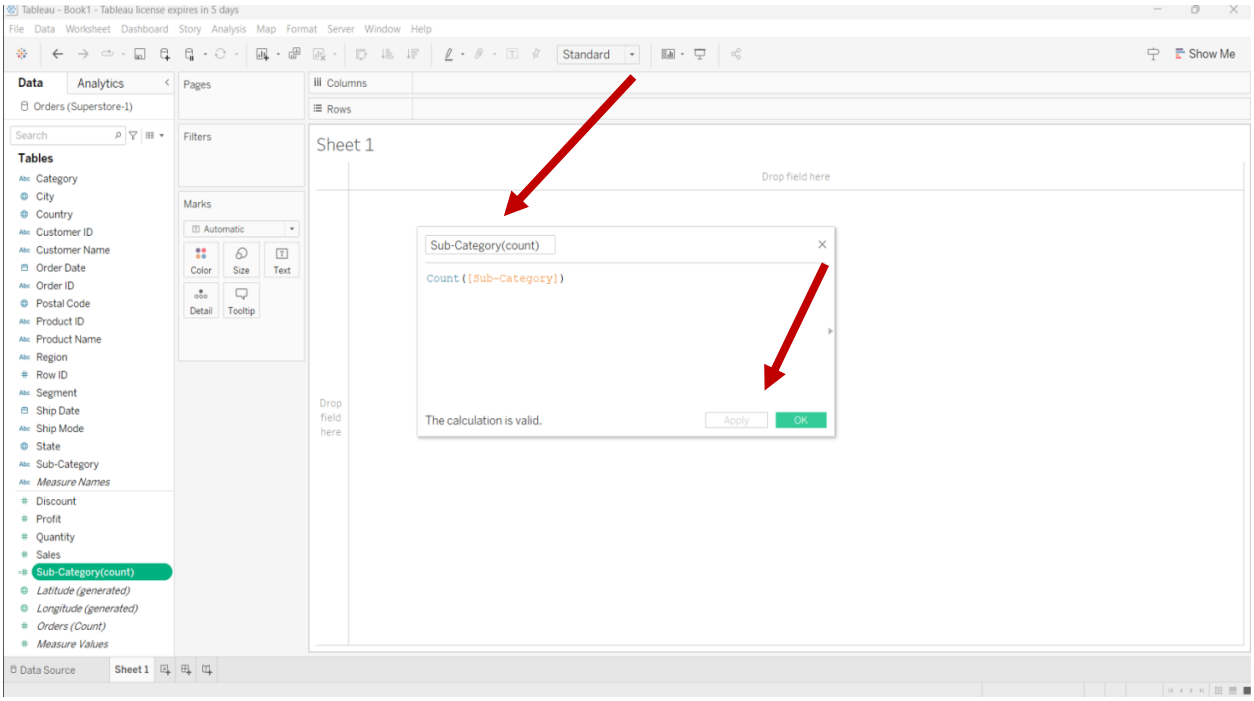
Download the excel file named “Superstore-1”. Open Tableau and then connect to the Excel data, the data source page shows the sheets or tables in your data.

Then create a new worksheet by clicking on “Sheet1”. After creating the new sheet, click on the downward arrow on the top left corner, then select “Create Calculated Filed”.



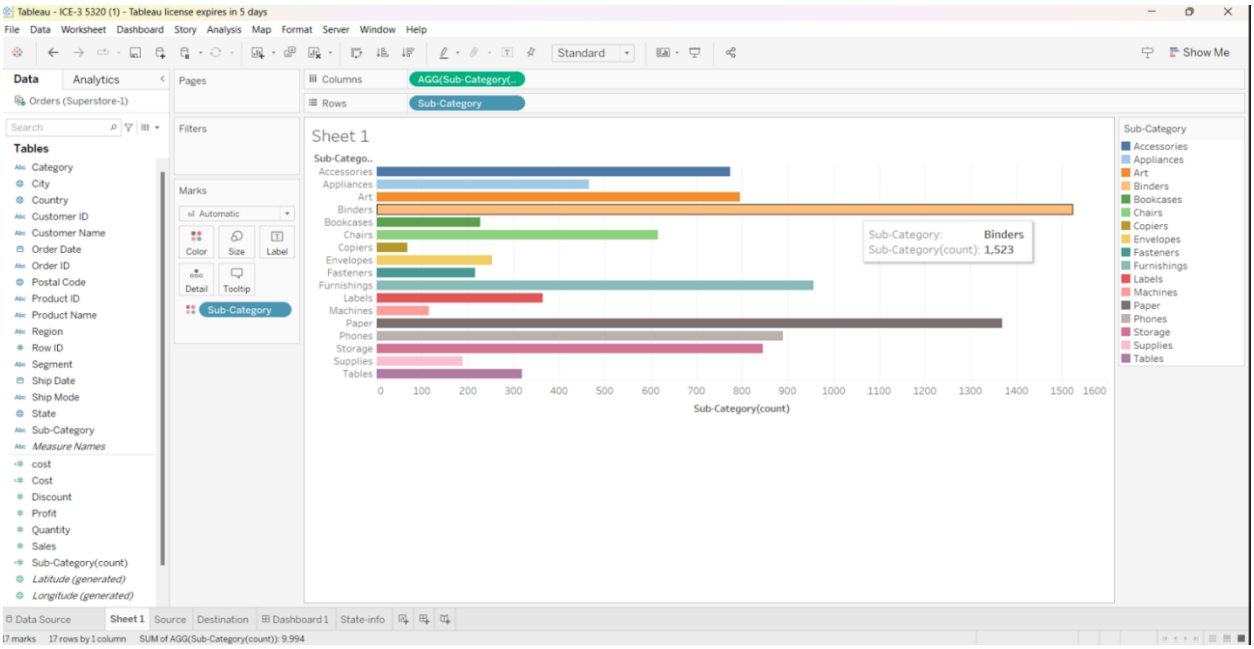
Step 2:

After selecting the “create calculate field”, a calculate field editor will be opened. Then Rename the calculated field as Sub-Category(count) and type the function Count (Sub-Category), Then click “Apply” and click ok.



Step 3:

Now, we can see the field Sub-Category(count) on the left side. Add Sub-Category(count) from the left side to the columns shelf and drag Sub-Category from the left to the rows shelf. To differentiate the Sub-Category's, add Sub- Category field to the color marks.



The above data used Horizontal bar graph visualization:

In general, horizontal bar graphs visualize categorical data which are usually represented in rectangle bars. The length of the bar is the value it consists of.

Types of Variables used in the graph:

Sub- Category: is categorical data, they are classified into different subcategories (i.e., Technology is classified as Accessories, Copiers, Machines, Phones etc.)

Sub-Category(count): It is numerical data. We created a new field as **Sub-Category(count)** which gives count of each Sub-Category.

Observations:

- We used **Sub-Category** in rows and **Sub-Category(count)** in columns, as we added sub-category to colors, so that it is visualized Horizontal bar graph so that each sub-category is visualized in detail and with different colors of each.

- Here, we can see sub-category “Binders” has count around 1,523. The highest count is for “Binders” sub-category and the least count is for “copiers” which is below 100.

Question 1 (15%)

Follow tutorial 1 and complete the below question for the given attributes.

1) Add a new sheet and Click on the downward arrow on the top left corner, choose Create Calculated Filed, open the calculate field editor.

2) Rename the calculated filed as **Profit Ratio** and type the function, SUM([Profit])/SUM([Sales]) and click Apply.

3) Use Profit Ratio, Category and Sub-Category from the left to create visualization. Choose anyone type of graphs and analyze the data and explain your understanding.

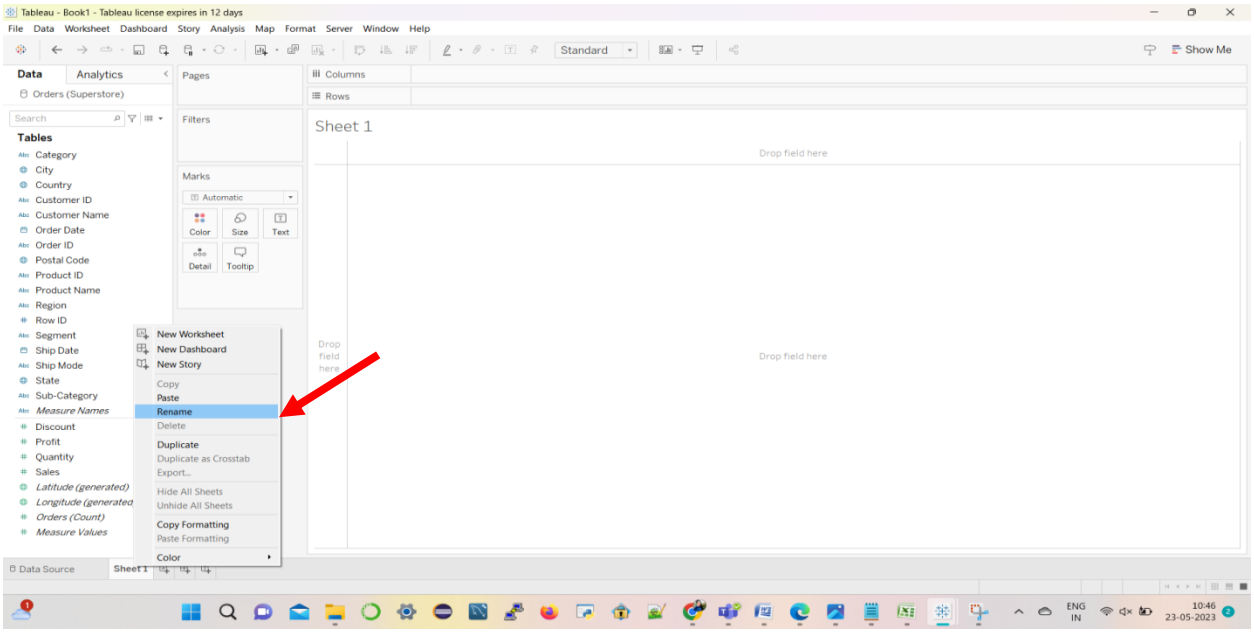
Task 2:

Tutorial 2

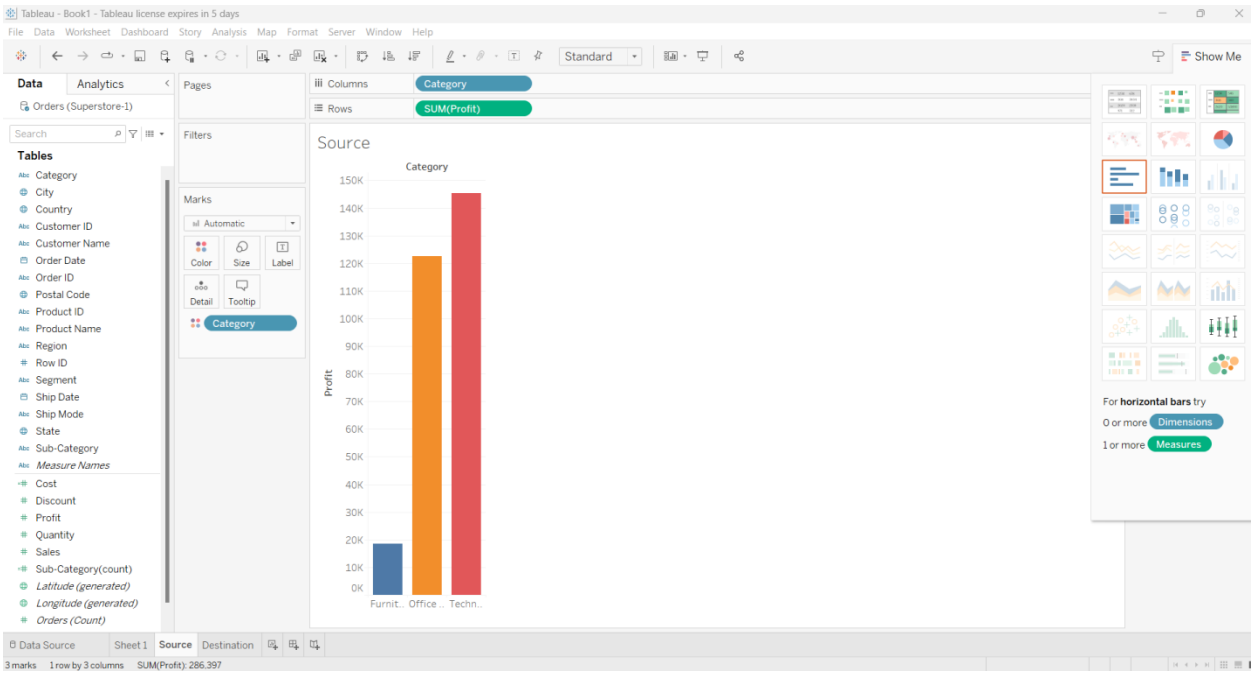
Interactivity with actions (filter):

Step 1:

First, we must create the two new worksheets name them as source and destination.

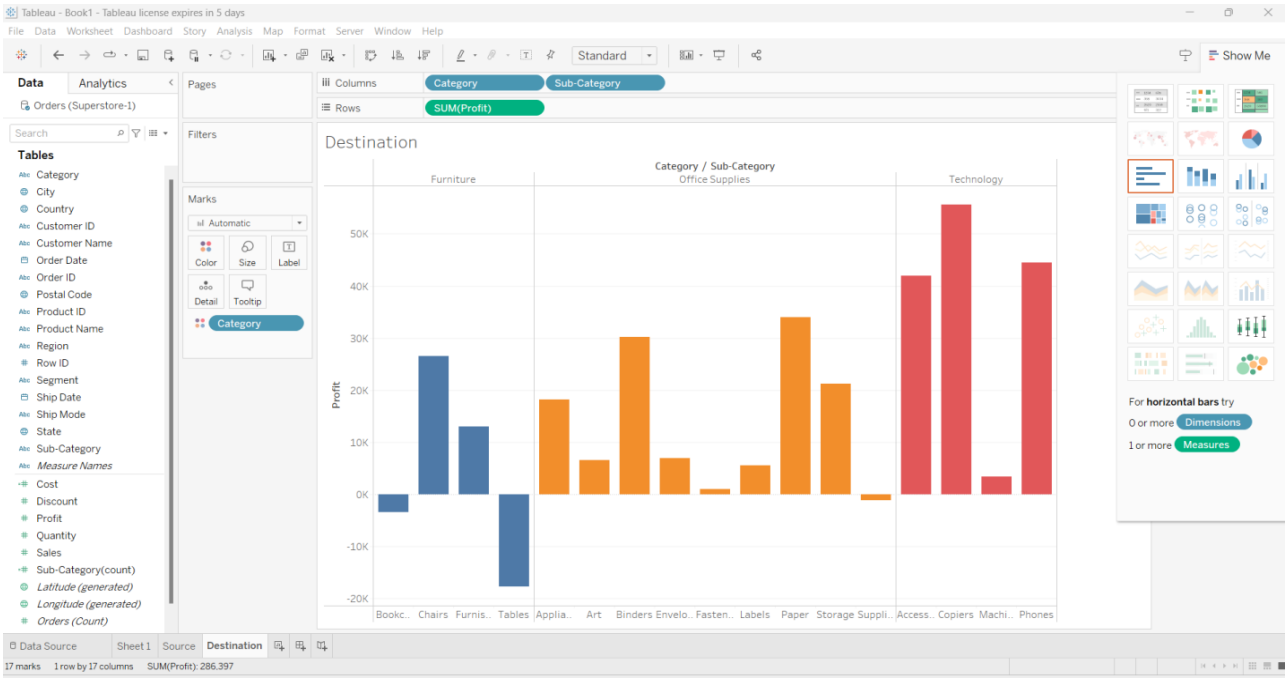


Click on the new worksheet and then rename it by giving right click and select rename, then change the name. Here, rename it has **source**, then the source worksheet will be created. After creating the source worksheet add category to the columns shelf and profit to the rows shelf and add category to the color marks, to differentiate category with different colors.



Step 2:

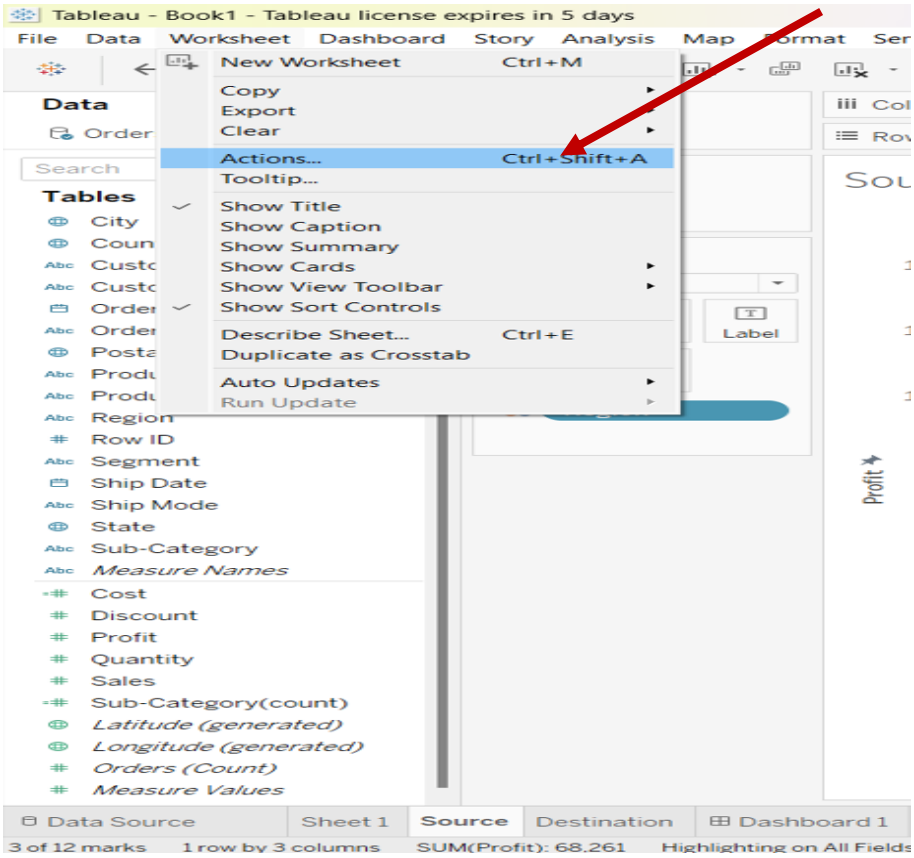
Similarly create a new worksheet and rename it has **destination** and in destination sheet add category and Sub-Category to the columns shelf and profit to the rows shelf and add category to the color marks, to differentiate category with different colors.



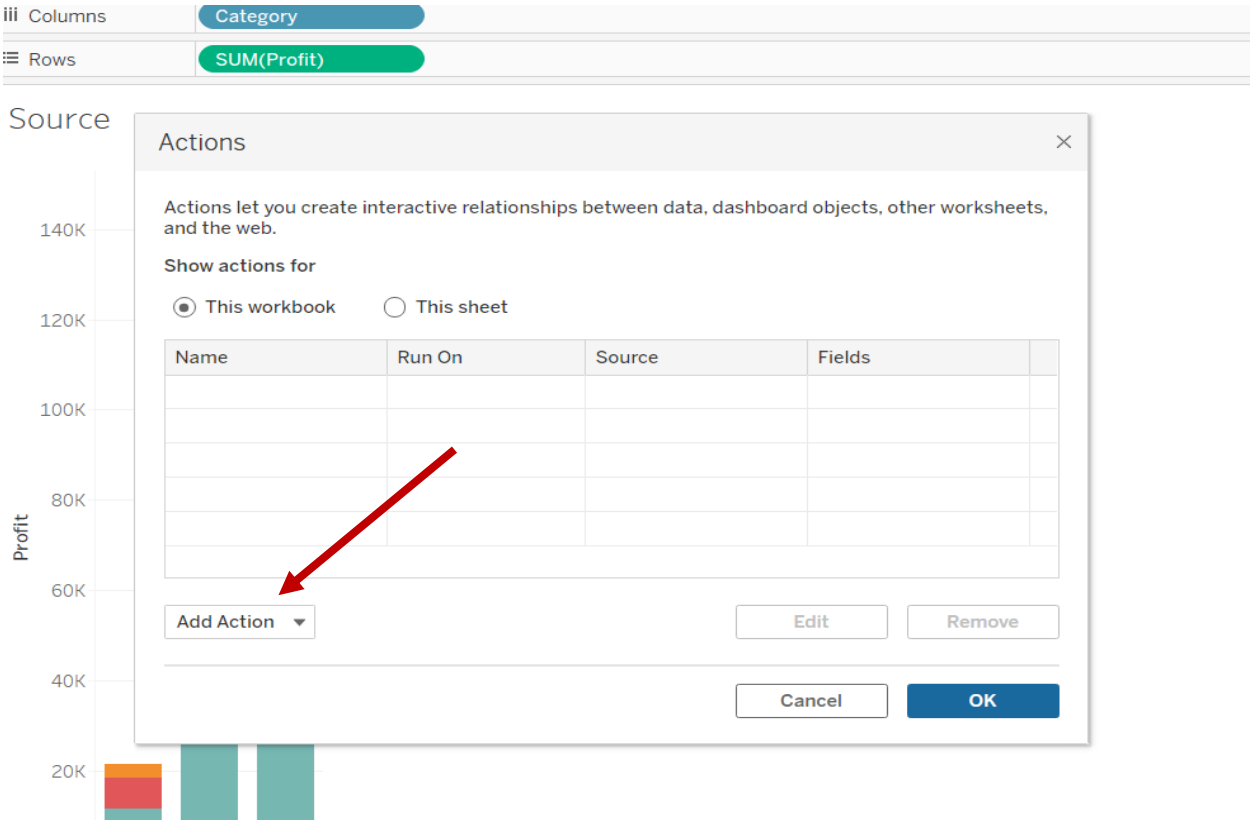
Step 3:

After creating Profit- category (Source) and Profit- subcategory (Destination) we should add Filter and make both sheets interactive.

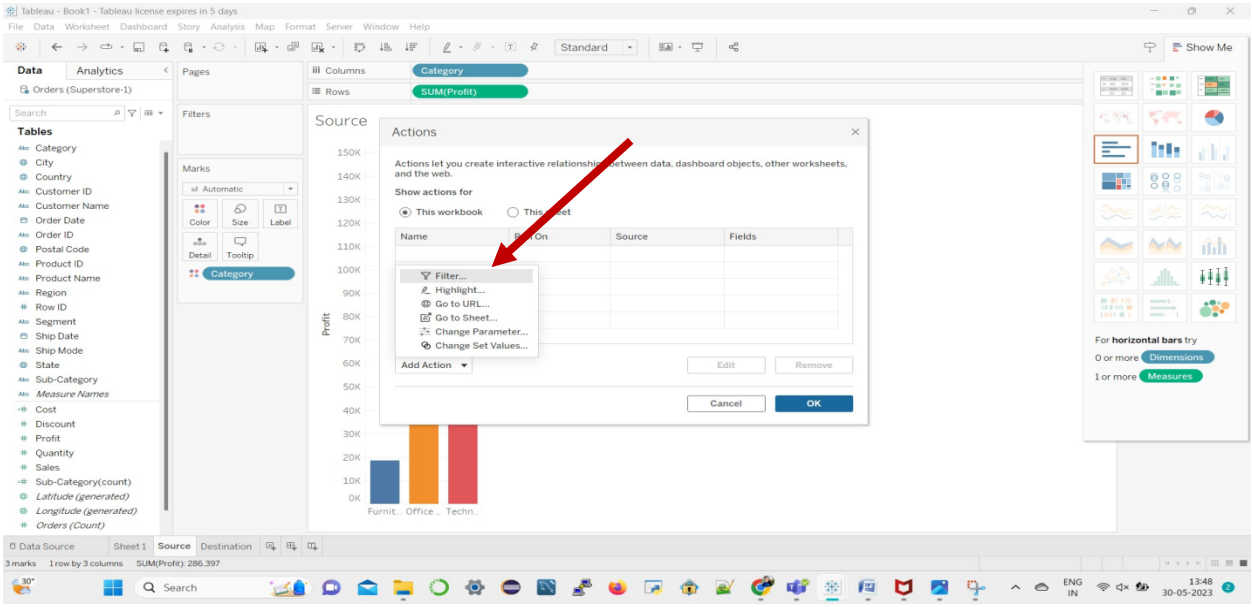
Click on the worksheet on the top and select actions.



Then click on add action.



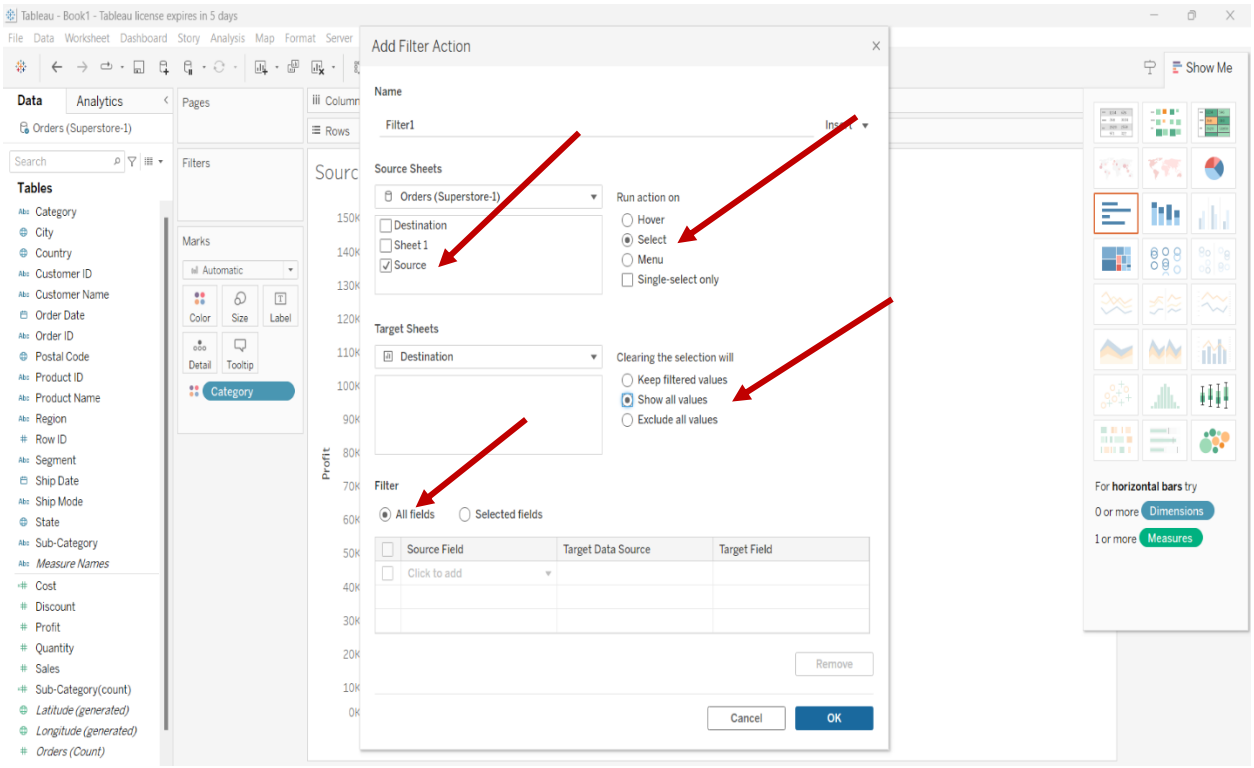
After clicking on add action, select filter to make interaction with filter.



Step 4:

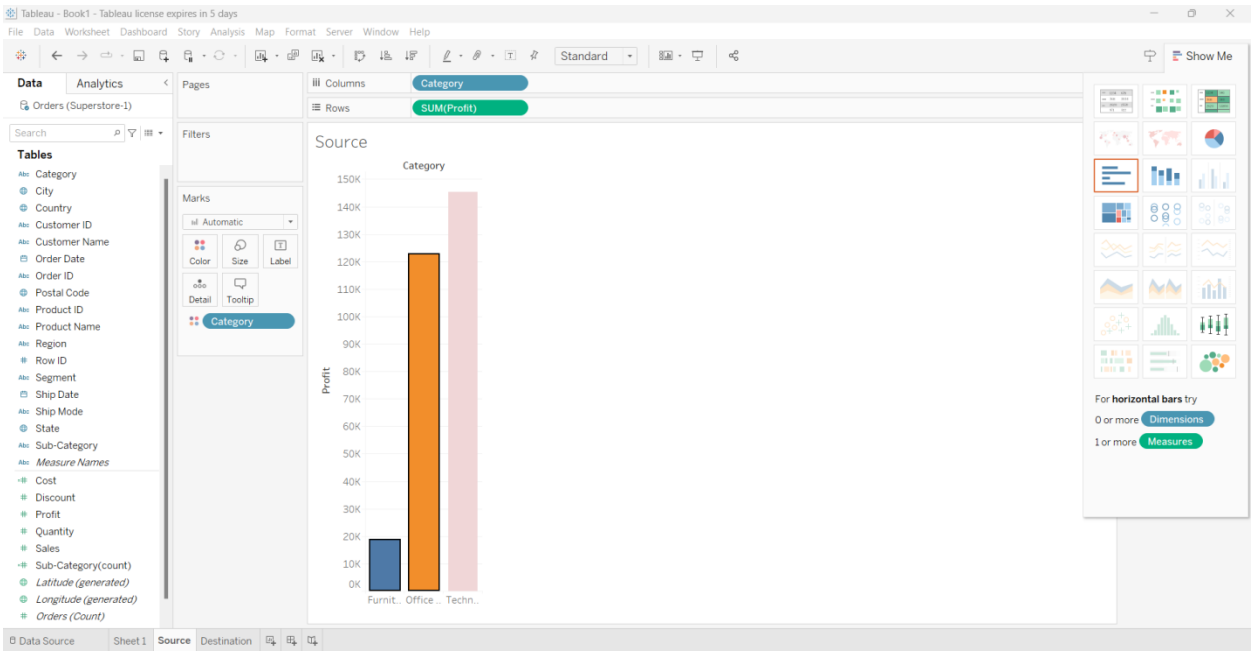
Now we should add a filter section. Select the source sheet as “Source” and Target sheet as “Destination” worksheet. Make sure that “Run action on” should have “**Select**” and clearing the selection will have selected “**Show all values**” and Filter is selected “**All fields**”. Then click ok.

Now connection between the source sheet and destination sheet is created.

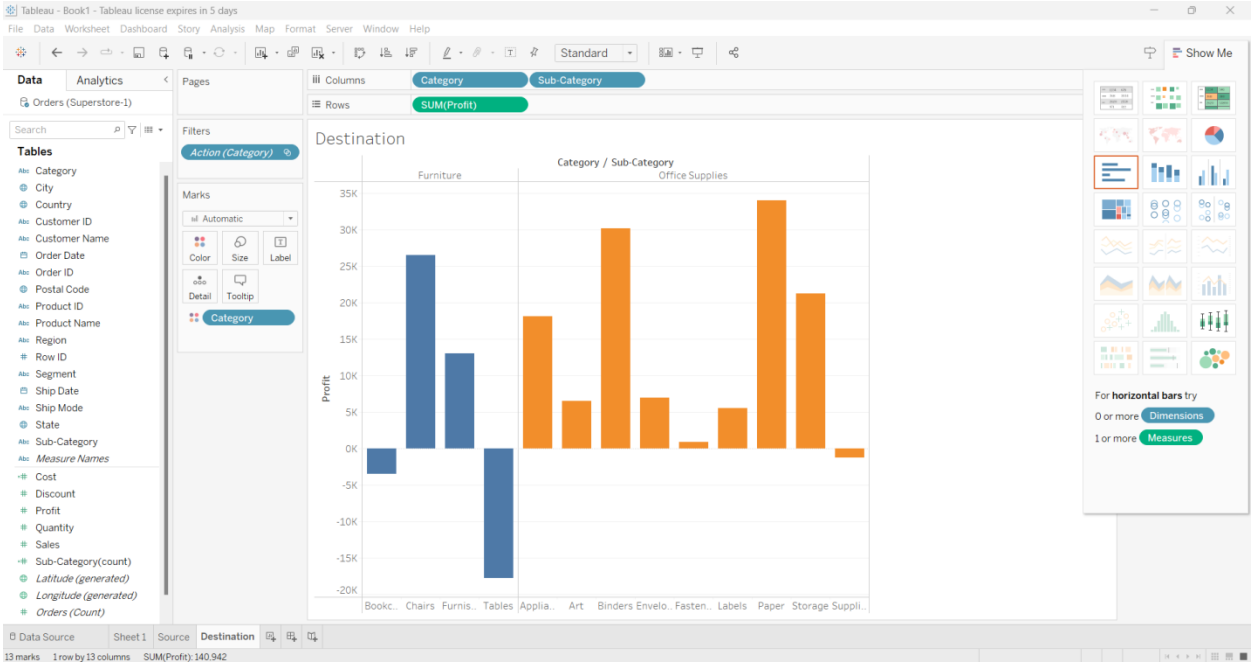


By using the control key, select any two categories in the source sheet simultaneously and then in destination sheet we will get the visualizations of only the categories we have selected in source.

Here in the source worksheet, we are selecting only furniture and office supplies categories.



In destination sheet we are getting the visualization only for furniture and office supplies category.



In the above graph, we can see profit for different subcategories of selected categories i.e. furniture and office supplies.

Question 2 (30%): -

Follow Tutorial 2 and complete the below Question.

1. Create a new sheet (Sales-Category) as the source sheet, put Sales in the rows and Category in the column.
2. Create another new sheet (Sales-Subcategory) as the destination sheet, put Sales in the rows, category and sub-category in the columns. Also apply color to category.
3. Go to Worksheet on the top, choose Actions, click Add Action, select filter, follow the tutorial and create connection between the source sheet and destination sheet.
4. Select any two categories in the source sheet simultaneously by using the control key, and give a screen shot on the destination sheet.

Task 3:

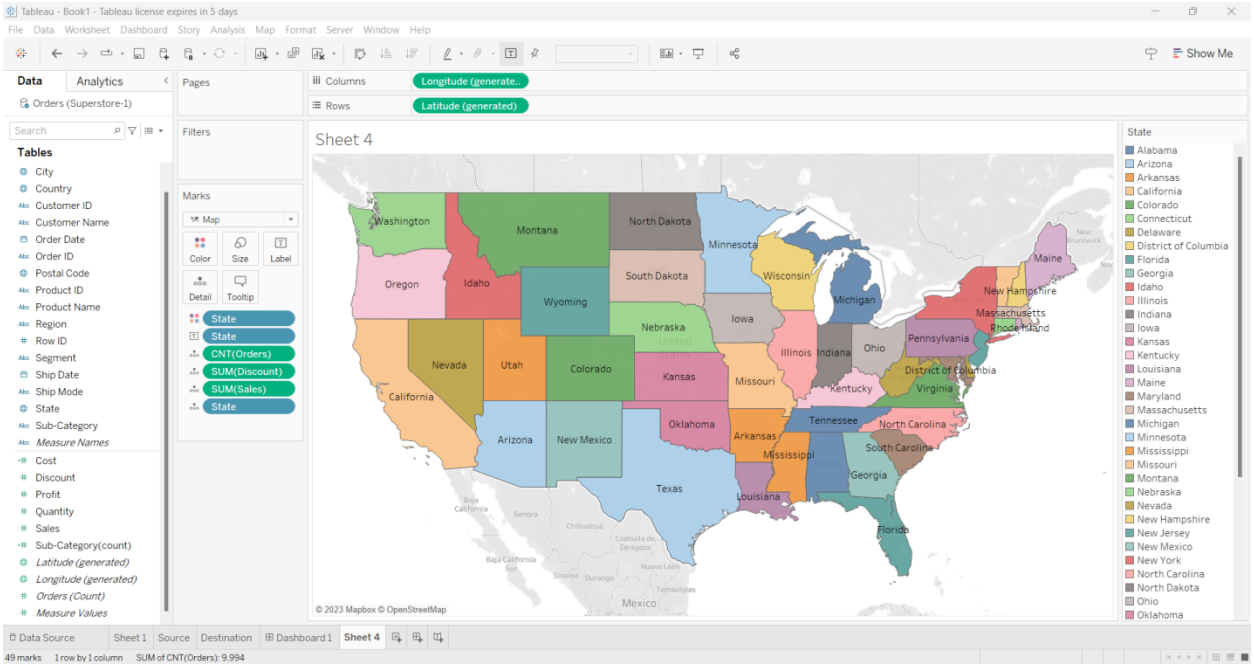
Tutorial 3

Interactivity with actions – URL:

Step 1:

Click on the new worksheet and name it as “StateSale” and add state to color marks from the left side and drag Latitude to the rows and longitude to the columns shelf.

Also add state to the label and detail. Drag Orders (Count),discount and sales fields from the left side to the detail.

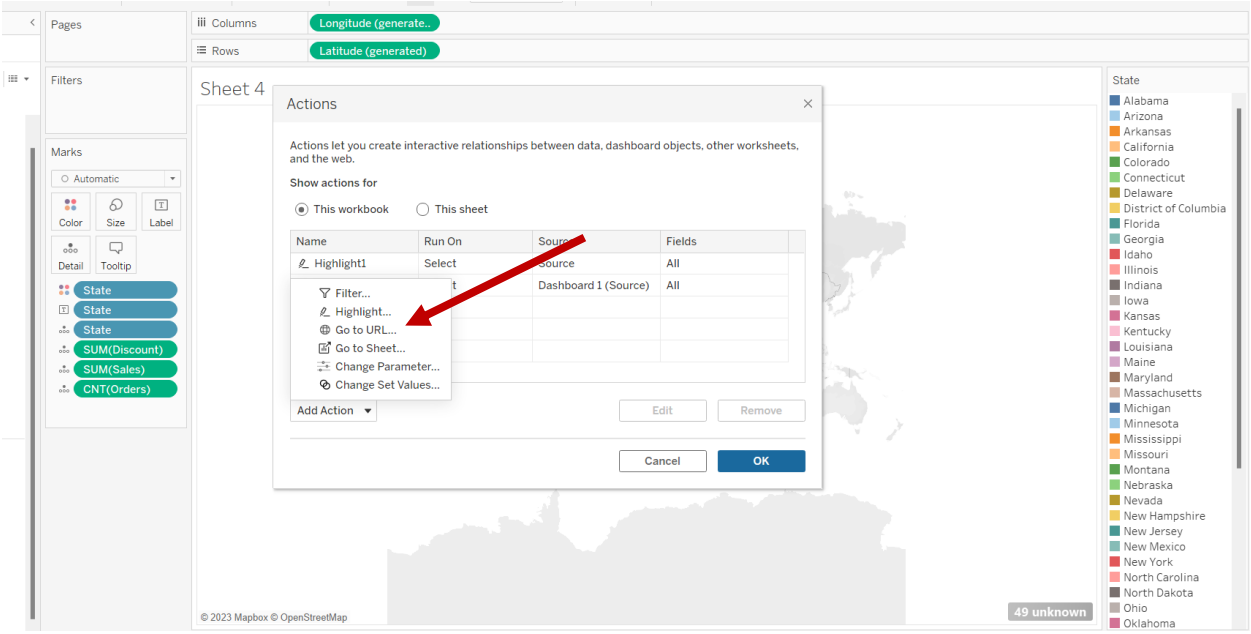


Above we can see the map-based visualization in which states are differentiated with different colors.

Step 2:

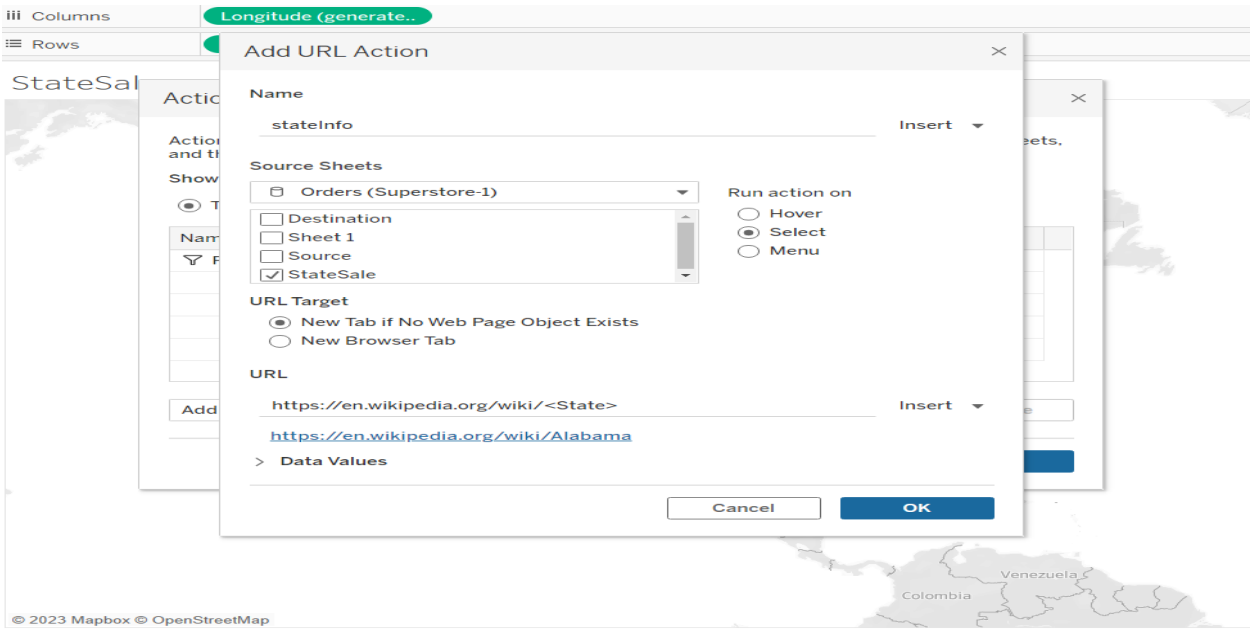
Adding URL action:

To add URL as an action, click on worksheet and select actions. Then click on add action and select "**Go to URL**". Then click ok.



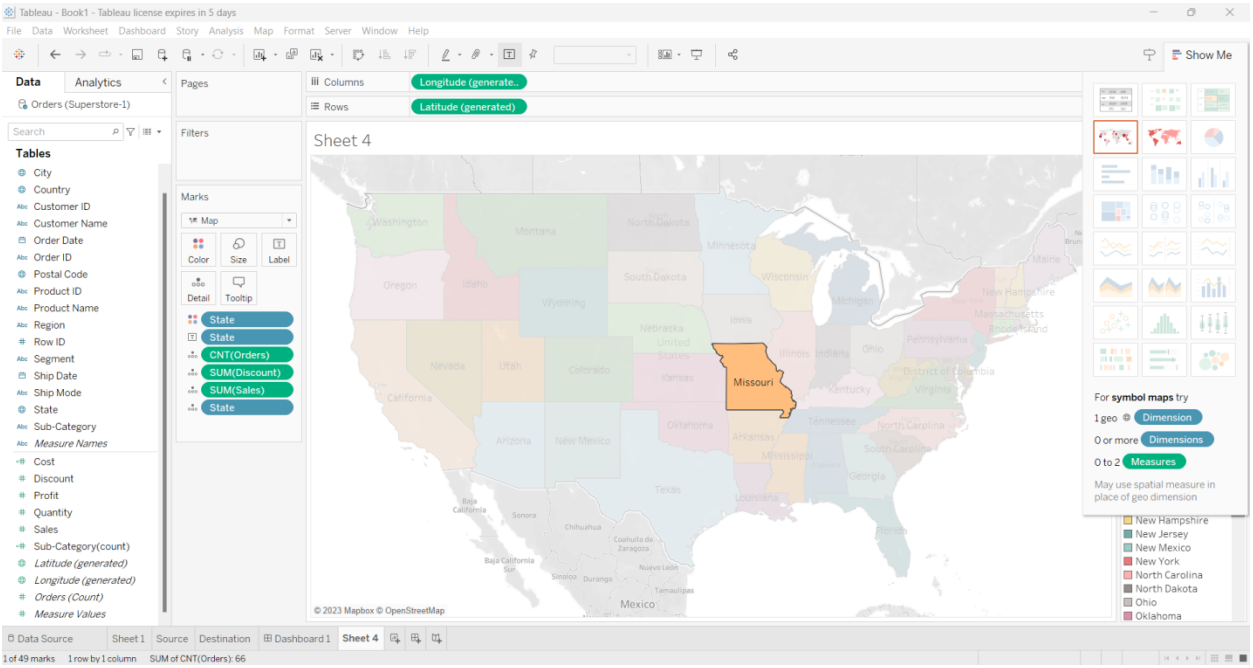
We should add an URL action to the StateSale worksheet. Give the name as "stateInfo" and select the source sheet as "StateSale". Make sure that "Run action on" should have "**Select**" and URL Target is selected as "**New Tab if no web page object Exists**". Add the given below URL and see whether URL has <state> at the end. Then click ok.

URL: https://en.wikipedia.org/wiki/<State>

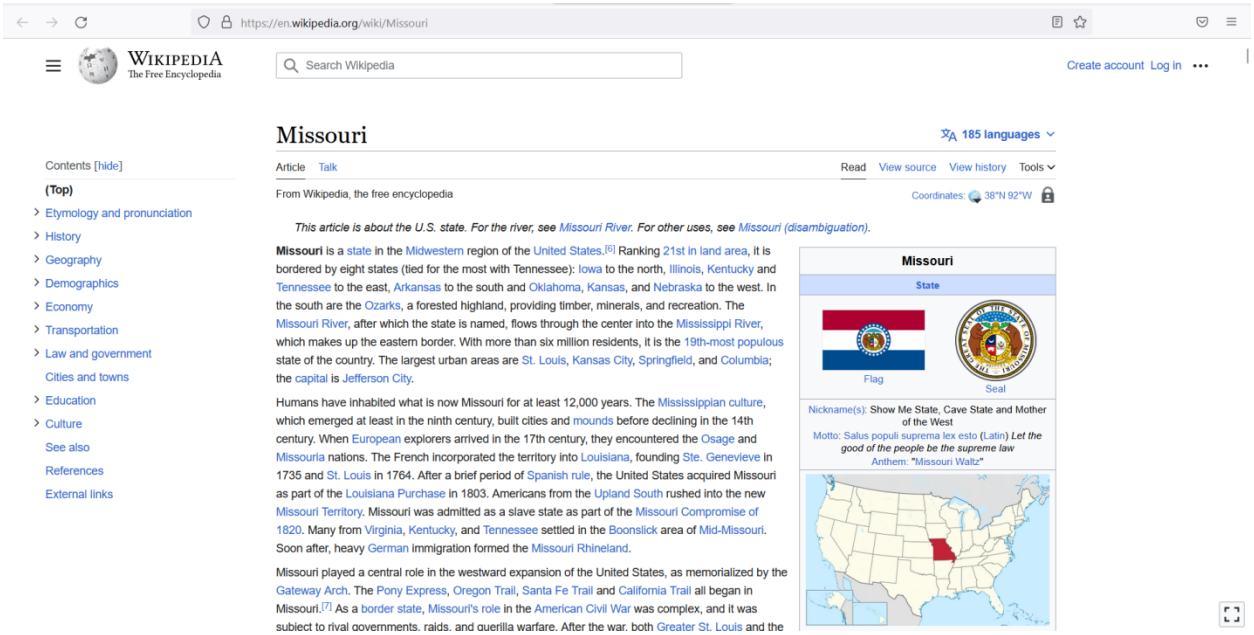


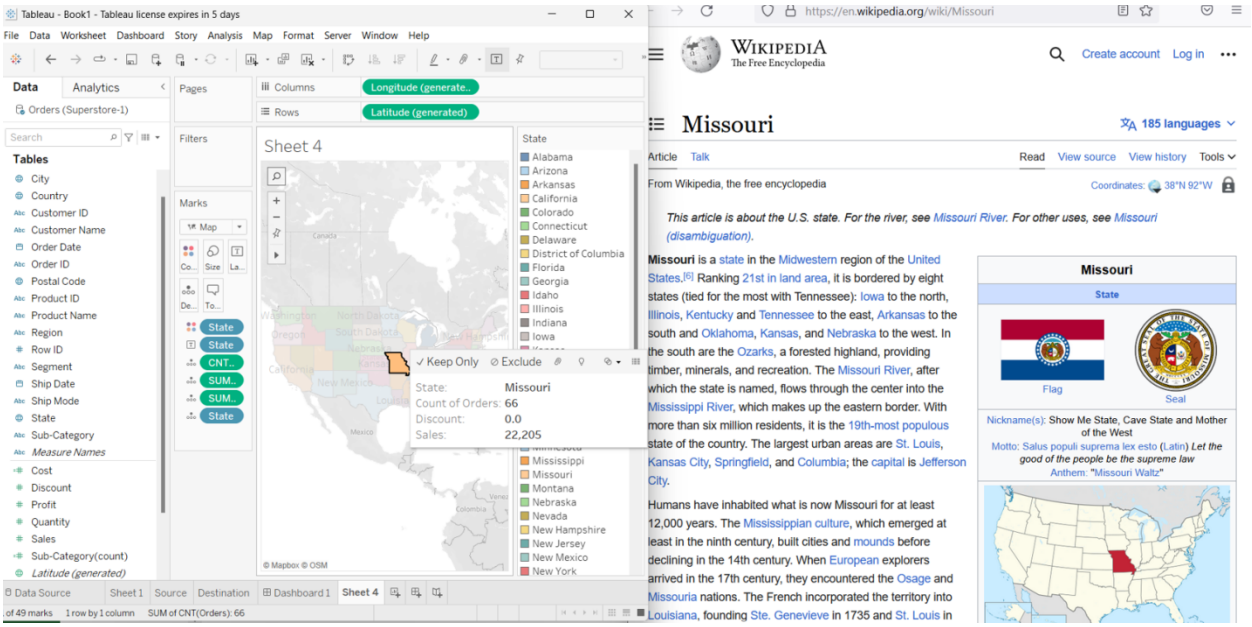
Step 3:

By using control key,when we select a state ,it redirects to the wikipedia page of the particular state.



Here I have selected the Missuori state, the wikipedia page of Missuori gets opened.Similarly we can select any state and we can get the information of the particular state.





Question 3 (20%): -

Follow Tutorial 3 and complete the below Question.

1. Create a new worksheet (City_Profit) and create a map-based visualization by adding city to the color mark, detail and label and add Profits to label.
2. Apply URL actions to the new worksheet. When you click on the city, the corresponding webpage (ex, Wikipedia) will be shown in the browser.

   URL: https://en.wikipedia.org/wiki/<City>

3. Select Two cities Indianapolis and Wichita and Provide the screenshots of Wikipedia of those two cities.
4. What is the profit gained by the above two cities.

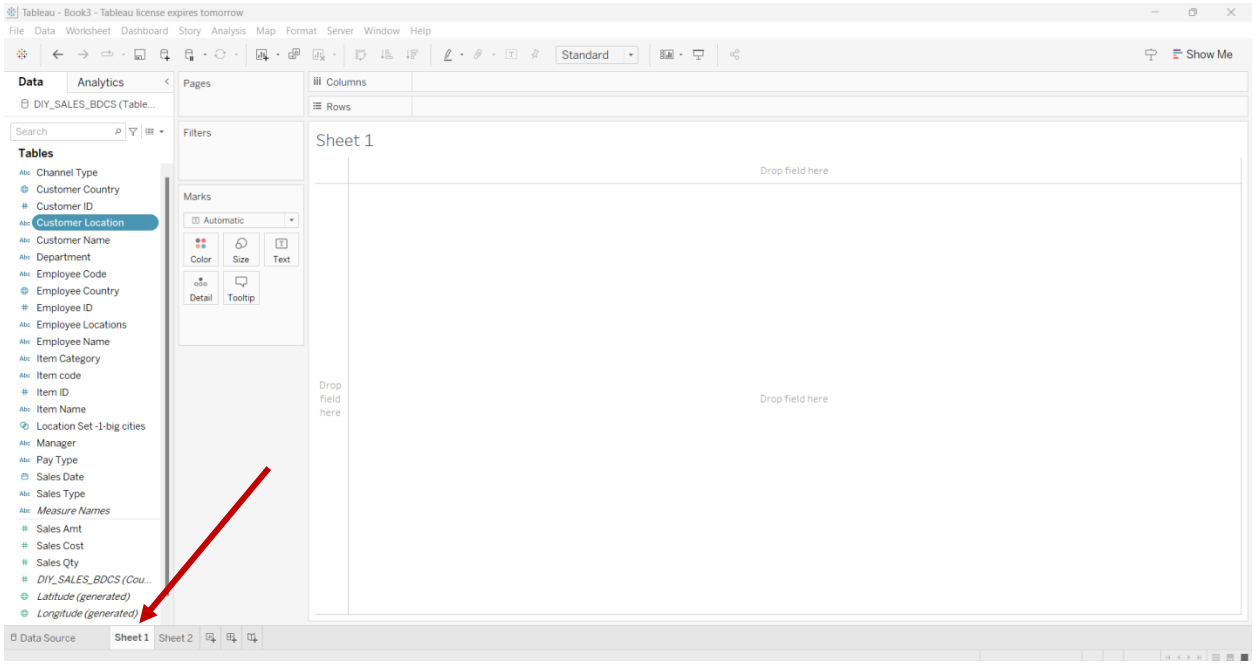
Task 4:

Tutorial 4

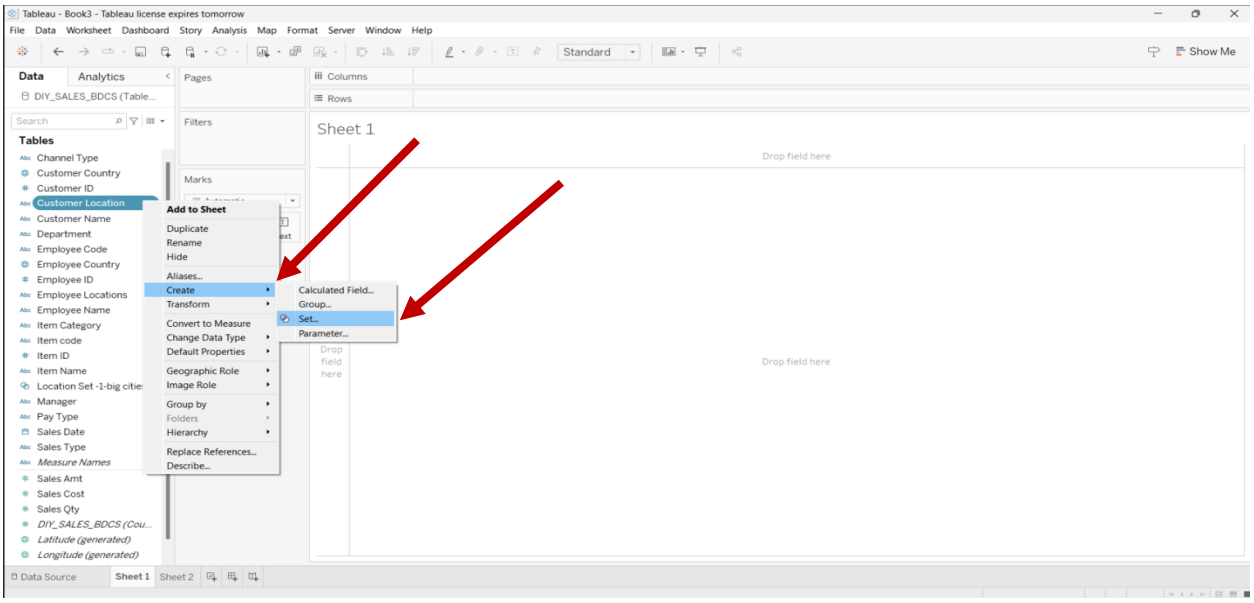
Creation of Sets:

Step 1:

Download the excel file named "Tableau_Data". Open Tableau and then connect to the Excel data, the data source page shows the sheets or tables in your data.

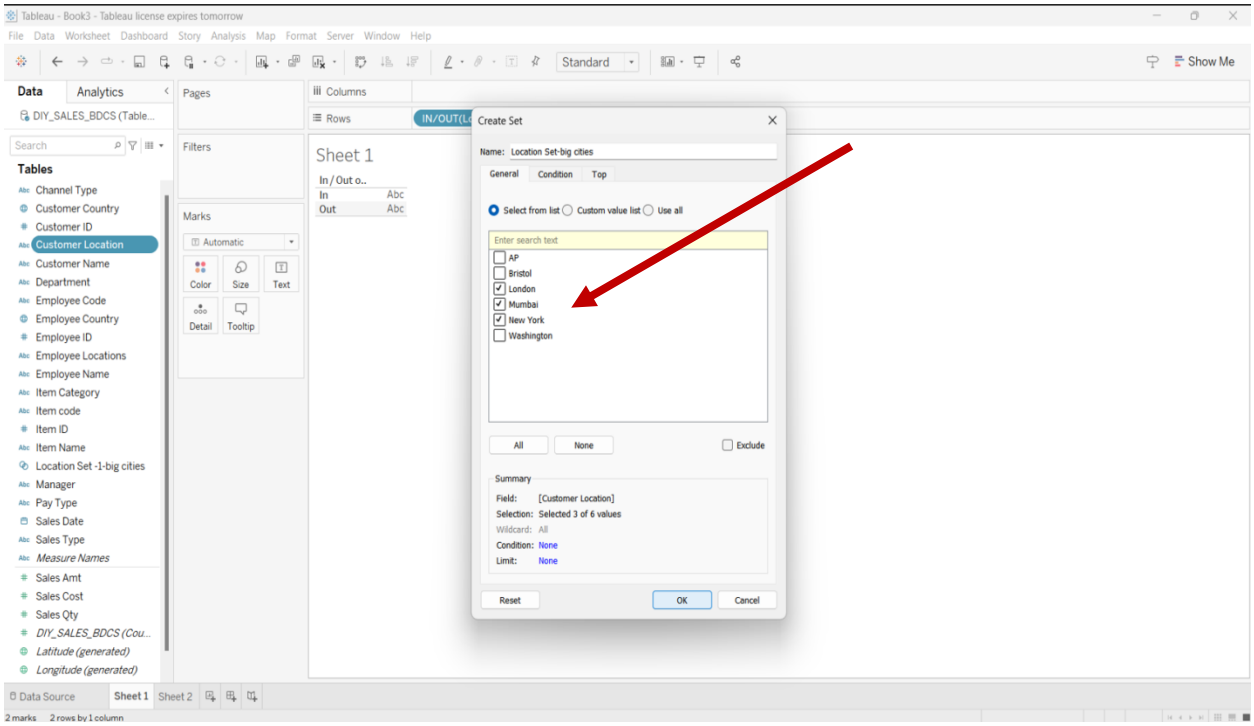


After creating the new sheet, right click on the Customer location attribute on the left side and select create ->Set.

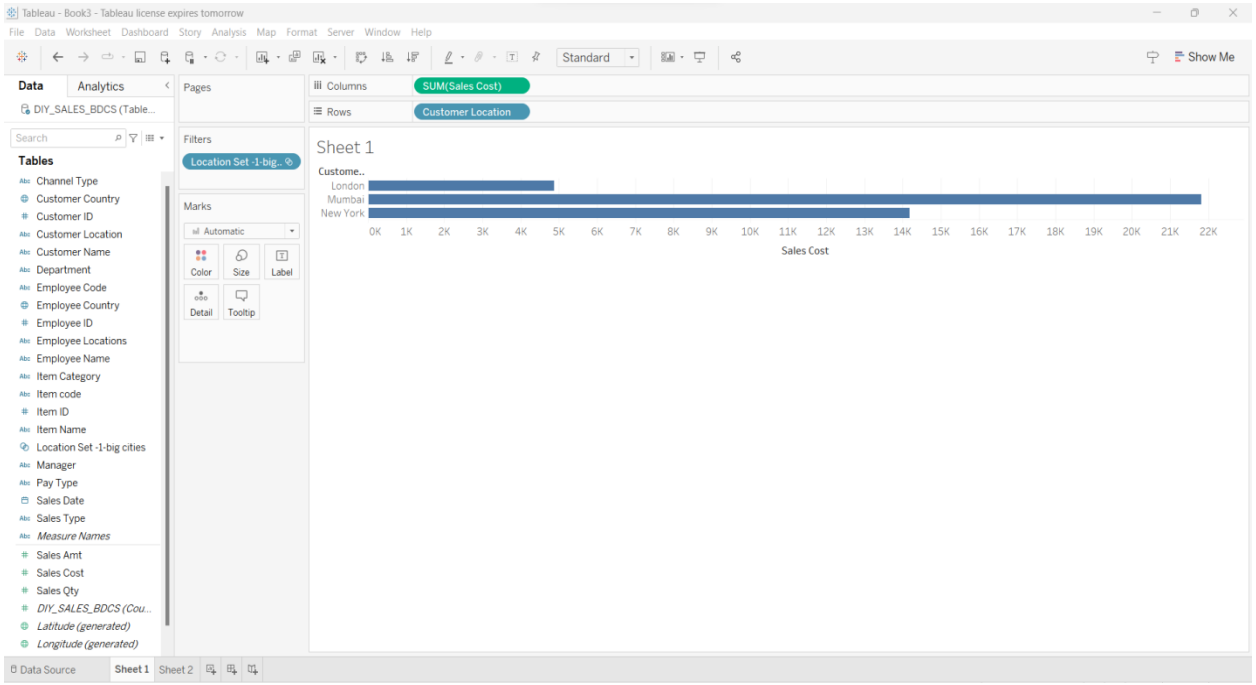


Step 2:

Then rename the set as “Location Set-big cities” and select the cities like Mumbai, NewYork, London by clicking on the respective check box to include them into a set. Then click ok.

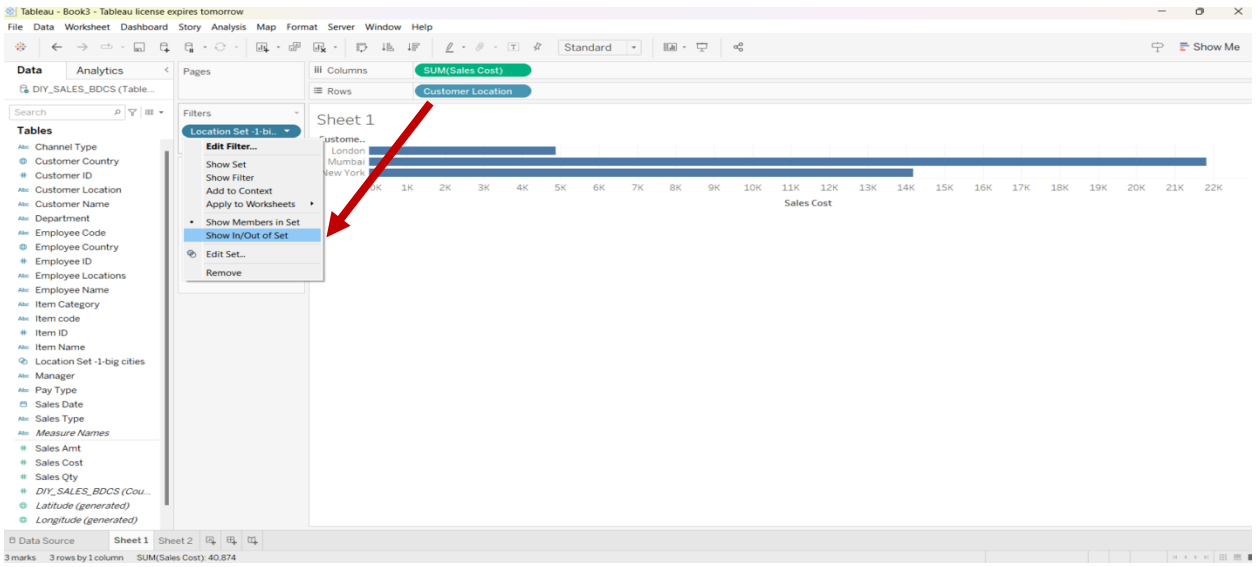


Location Set - big cities set will be created and it will display on the left side. Add Location Set - big cities to the Filters shelf and add attributes “Sales Cost” to the columns and “Customer Location” to the rows.



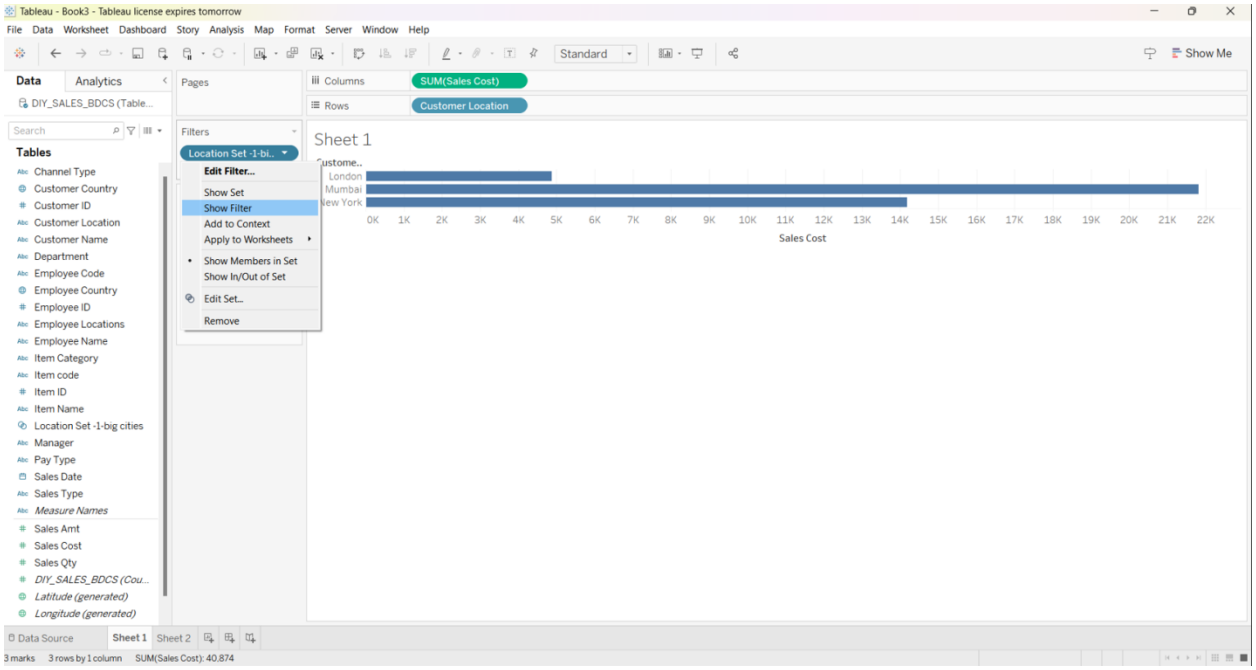
The above data used Horizontal bar graph visualization. In the above screenshot we can see Sales cost of cities London, New York and Mumbai only, since we have added “Location Set - big cities” to the filters shelf.

Click on Location Set - big cities in filters shelf and Select Show In/out of set.



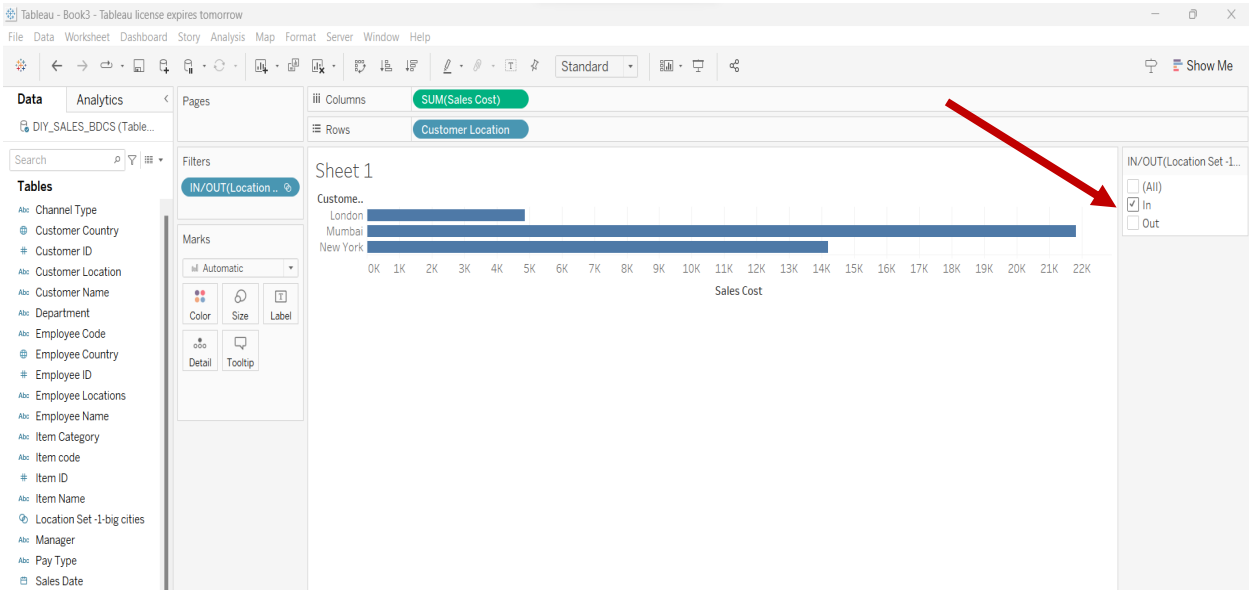
Step 3:

Then Click on Show filter to show the filters on the right side of the screen. By Changing the filters on right side, we can see the different types of visualizations.



Selected “In” of set:

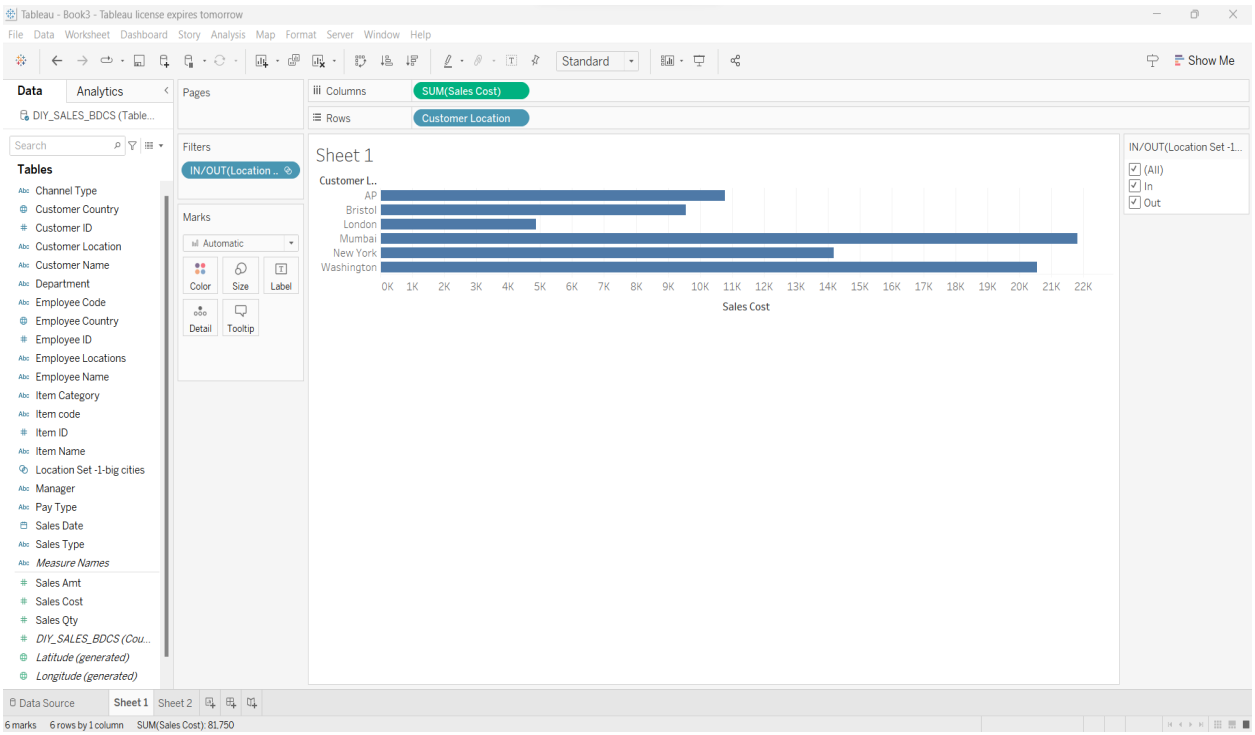
When we Select “In” of set, it gives the visualization of Customer locations which are selected in the set.



In the above graph, we can see only Sales cost of London , Mumbai and NewYork cities as they are inside the set.

Selected “In” and “Out” of set:

When we Select both “In” and “Out” of set from the right side, it gives the visualization of Customer locations of both inside and outside of the set.

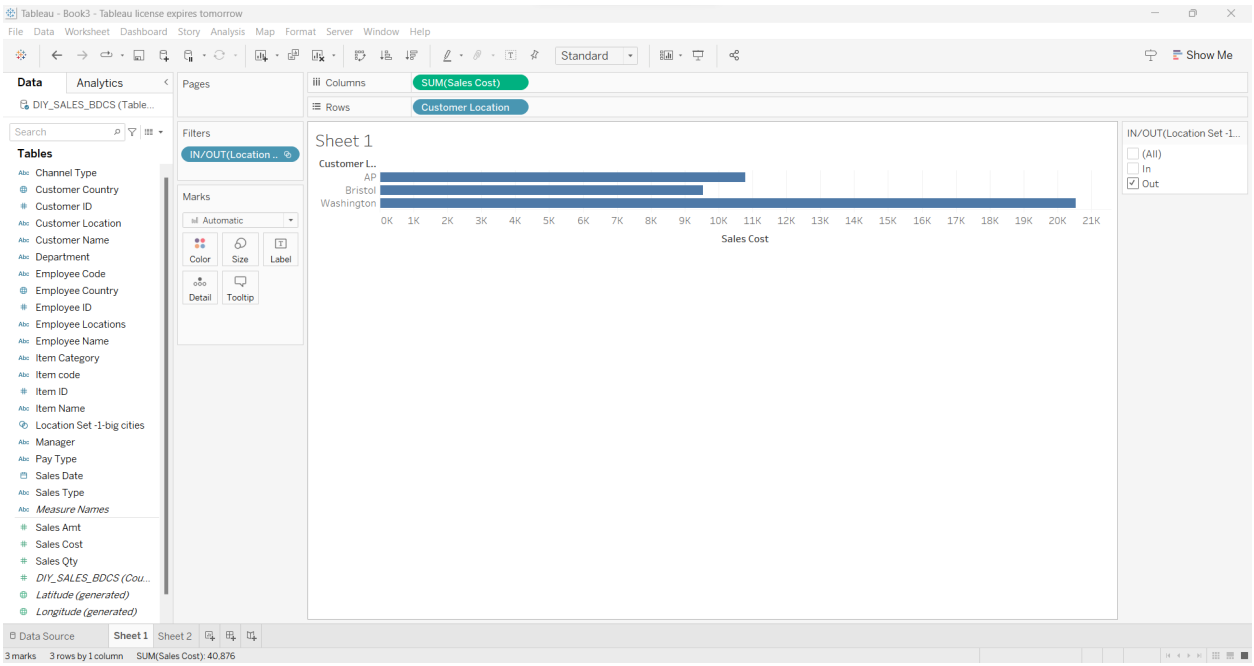


In the above graph, we can see “Sales cost” of all the cities like London, NewYork, Mumbai, AP, Washington, Bristol.

Selected “Out” of set:

When we Select “out” of set from right side, it gives the visualization of Customer locations which are not selected in the set.

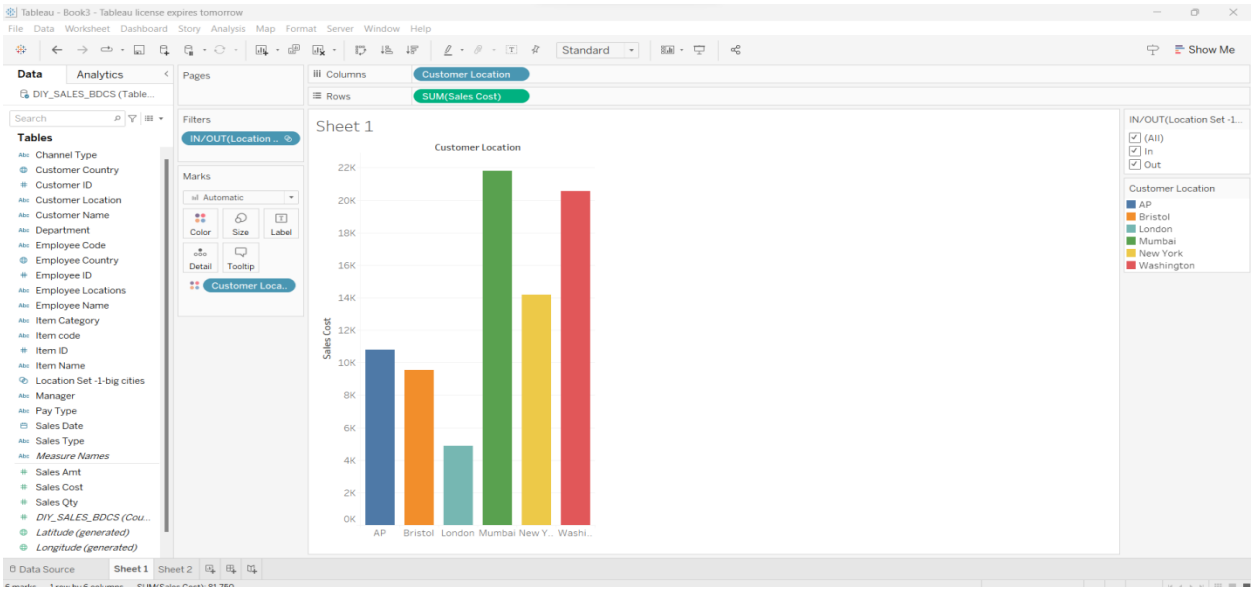
In below graph, we can see only Sales cost of AP, Washington and Bristol cities as they are outside the set.



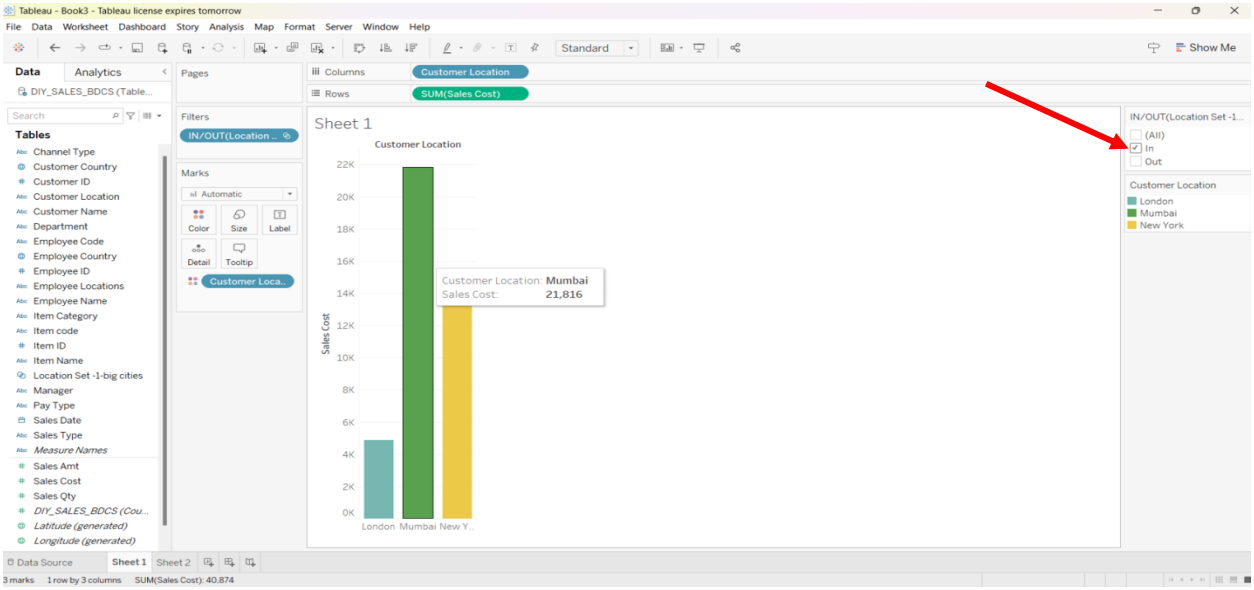
Step 4:

Analyze the different Visualizations:

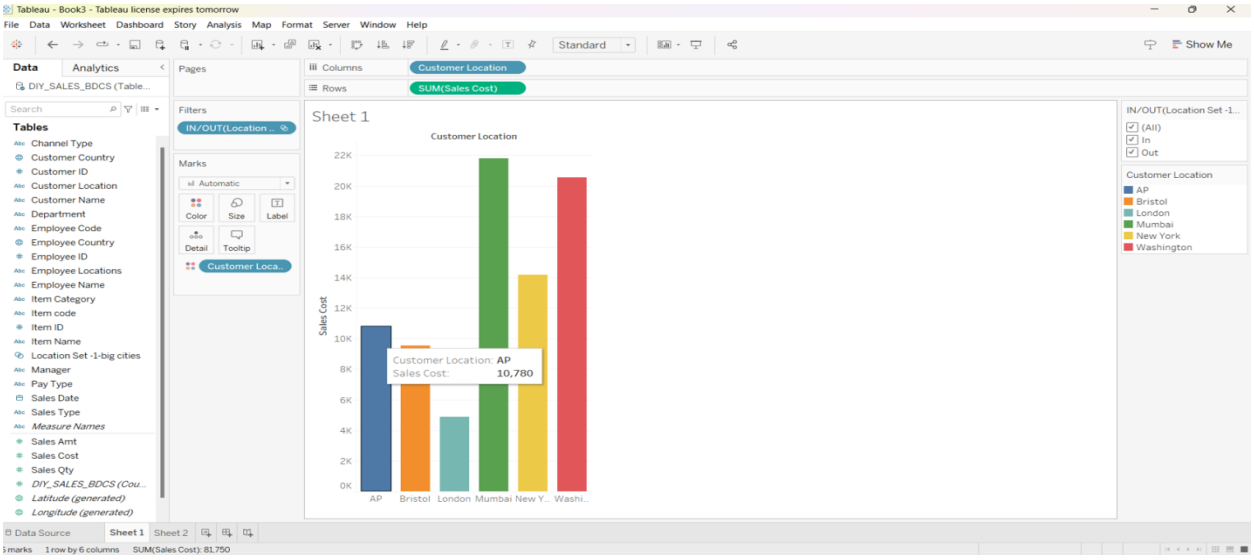
Add Sales Count to the rows shelf and add Customer location to the columns shelf and add Customer Location to the color mark to differentiate the locations with different colors.



The above graph represents the vertical bar graph visualization. We can see the sales cost of different locations which are differentiated with different colors and in this graph, it shows all the locations as both “In” and “Out” of set is selected.



Above, we have selected only “In” of set. So, we can see the sales cost of locations which are inside the set. When we place cursor on the graph, we get details like Mumbai location has sales cost around “21816”.



Above, we have selected both “In” and “Out” of set. So, we can see the sales cost of locations of both which is inside and outside the set. When we place cursor on the graph, we get details like AP location has sales cost around “10780”.

Question 4 (20%): -

Follow Tutorial 4 and complete the below Question by using “Superstore -1” dataset.

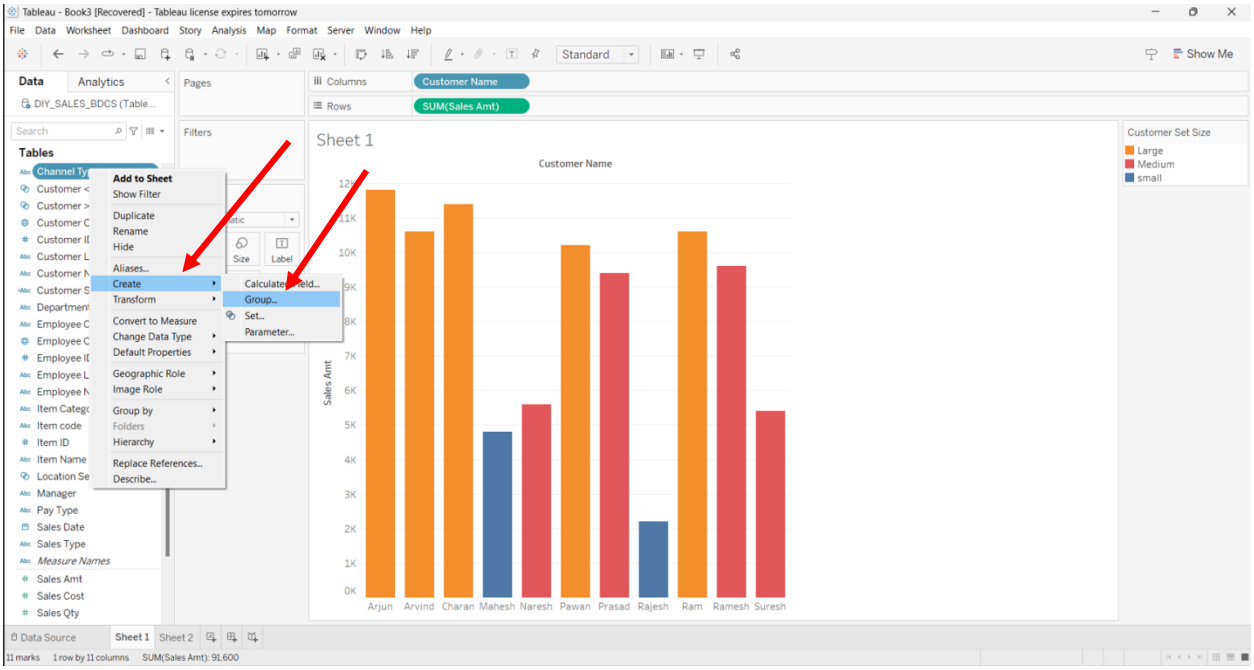
1. Click on the Sub Category from the left side and select Create -> Set .
2. Rename set has “SetofSubcategory” and Select sub categories like Accessories, Appliances, Binders, Bookcases and chairs into a new set and click ok.
3. Add Sub Category to the rows shelf and add Quantity in the columns shelf and add the newly created set to the filter shelf.
4. Click on SetofSubcategory in the Filters shelf and select show In/Out of set.
5. Click on the show filter. By changing the filters analyze the different visualizations. Analyze quantities of any two subcategories inside and outside of the set.

Task 5:

Tutorial 5

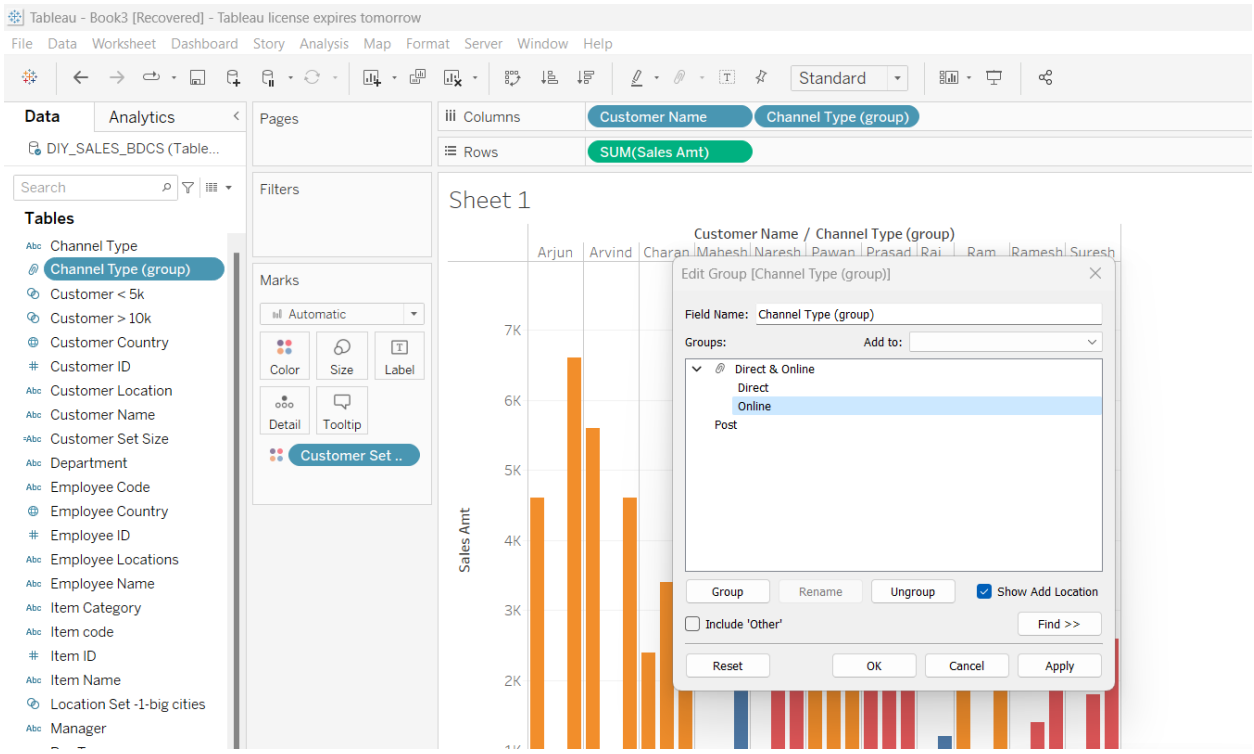
Creation of Groups:

Step 1: Use Tableau_Data to perform the below task



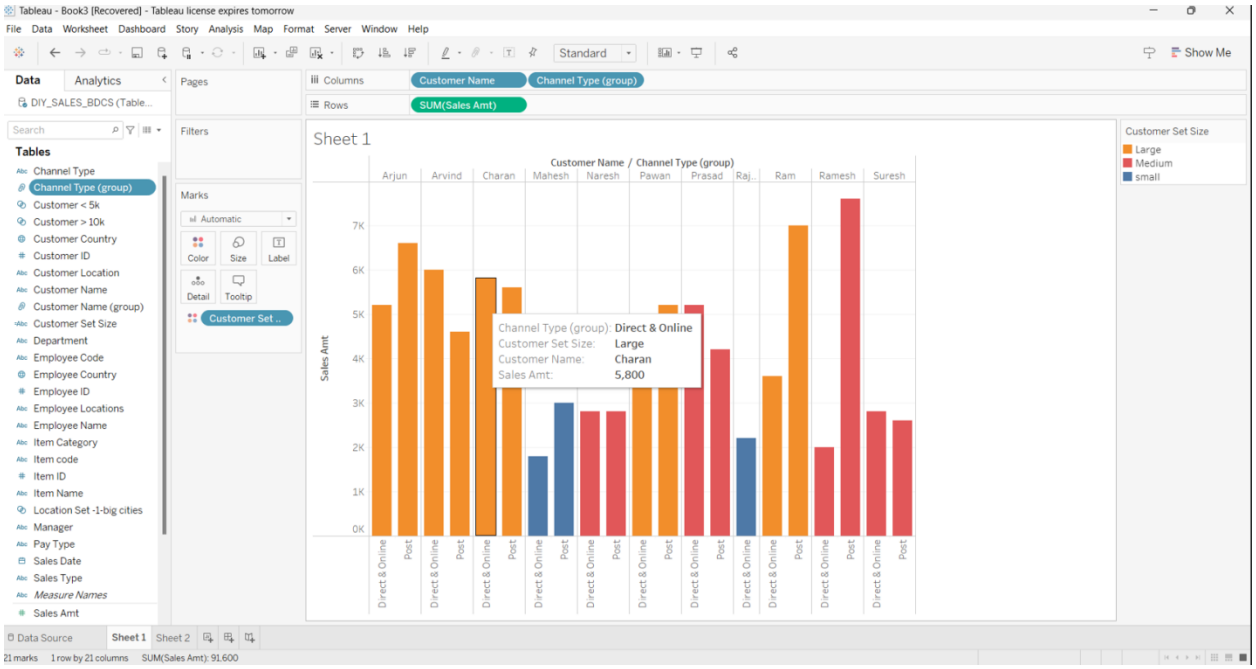
Right click on channel type and select create ->Group, a new group will be created.

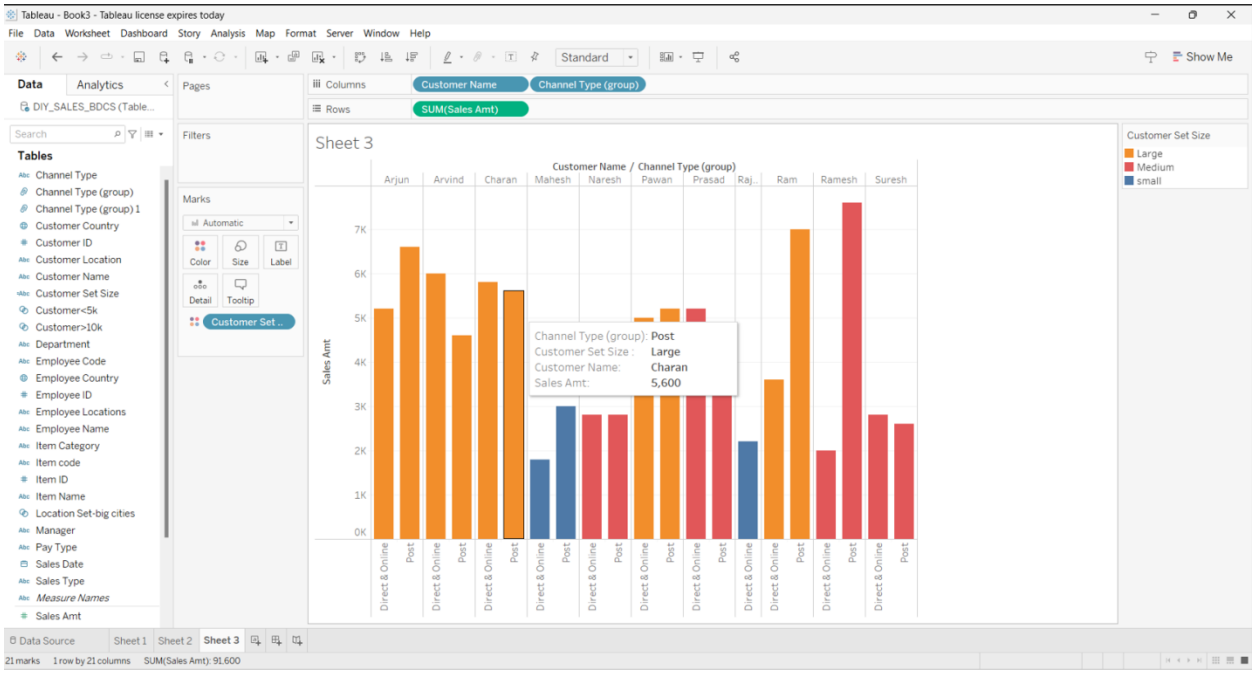
Step 2:



Then click on group and add direct and online into one group. Then click ok. A attribute called ChannelType(group) will be create on left side.

Add ChannelType(group) to columns along with Customer Name.





From above visualizations, we can see Channel type Direct and online as one group and post as another group. When we place cursor on the graph we can see Customer name “Charan” in Direct and online Channel Type has sales amount of 5,800. In post channel Type “Charan” has sales amount of 5,600.

Total Sales amount of Charan is 5600+5800 = 11400.

As Sales amount is greater than 10,000 customer set size will be large and differentiated with orange color.

Question 5 (15%): -

Follow Tutorial 5 and complete the below Question by using "Superstore -1" dataset.

1. Right click on the Category from the left side and select Create -> Group.
2. Click on group and make Office Supplies and Technology into one group, then click ok and Category(group) attribute will be created.
3. Add Category(group) and Sub category from left side to the columns shelf and Profit to the rows shelf.
4. Analyze the different Visualizations and explain your understanding

Rubric:

Questions	Ratings		Pts
Question 1: 1) Add a new sheet and Click on the downward arrow on the top left corner, choose Create Calculated Filed, open the calculate field editor. 2) Rename the calculated filed as **Profit Ratio** and type the function, SUM([Profit])/SUM([Sales]) and click Apply. 3) Use Profit Ratio, Category and Sub-Category from the left to create visualization. Choose anyone type of graphs and analyze the data and explain your understanding.	**1.5 pts** **Full Marks**	**0 pts** **No Marks**	1 pt for the every screenshot of visualization and 0.5 pts for Explanation
Question 2: 1. Create a new sheet (Sales-Category) as the source sheet, put Sales in the rows and Category in the column. 2. Create another new sheet (profit-Subcategory) as the destination sheet, put Sales in the rows, category and sub-category in the columns. Also apply color to category. 3. Go to Worksheet on the top, choose Actions, click Add Action, select filter, follow the tutorial and create connection between the source sheet and destination sheet.	**3 pts** **Full Marks**	**0 pts** **No Marks**	1.75 pts for the every screenshot of visual images of data and 1.25pts for Explanation

4. Select any two categories in the source sheet simultaneously by using the control key, and give a screen shot on the destination sheet.			
Question 3: 1. Create a new worksheet (City_Profit) and create a map-based visualization by adding city to the color mark, detail and label and add Profits to label. 2. Apply URL actions to the new worksheet. When you click on the city, the corresponding webpage (ex, Wikipedia) will be shown in the browser. URL: https://en.wikipedia.org/wiki/<City> 3. Select Two cities Indianapolis and Wichita and Provide the screenshots of Wikipedia of those two cities. 4.What is the profit gained by the above two cities.	**2 pts** **Full Marks**	**0 pts** **No Marks**	1pt for every screenshot of visual images of data and 1 pt for explanation

Question 4: 1. Click on the Sub Category from the left side and select Create -> Set .	**2 pts** **Full Marks**	**0 pts** **No Marks**	1 pt for the every screenshot of visual images of data and 1 pt for explanation

2. Rename set has "SetofSubcategory" and Select sub categories like Accessories, Appliances, Binders, Bookcases and chairs into a new set and click ok. 3. Add Sub Category to the rows shelf and add Quantity in the columns shelf and add the newly created set to the filter shelf. 4. Click on SetofSubcategory in the Filters shelf and select show In/Out of set. 5. Click on the show filter. By changing the filters analyze the different visualizations. Analyze quantities of any two subcategories inside and outside of the set.			
Question 5: 1. Right click on the Category from the left side and select Create -> Group. 2. Click on group and make Office Supplies and Technology into one group, then click ok and Category(group) attribute will be created. 3. Add Category(group) and Sub category from left side to the	**1.5 pts**	**0 pts**	1 pts for screenshot of visual images of data and 0.5 pts for the explanation

columns shelf and Profit to the rows shelf. 4. Analyze the different Visualizations and explain your understanding			