

CSCE 5320 Scientific Data Visualization ICE-2

Creating Visualizations in Tableau

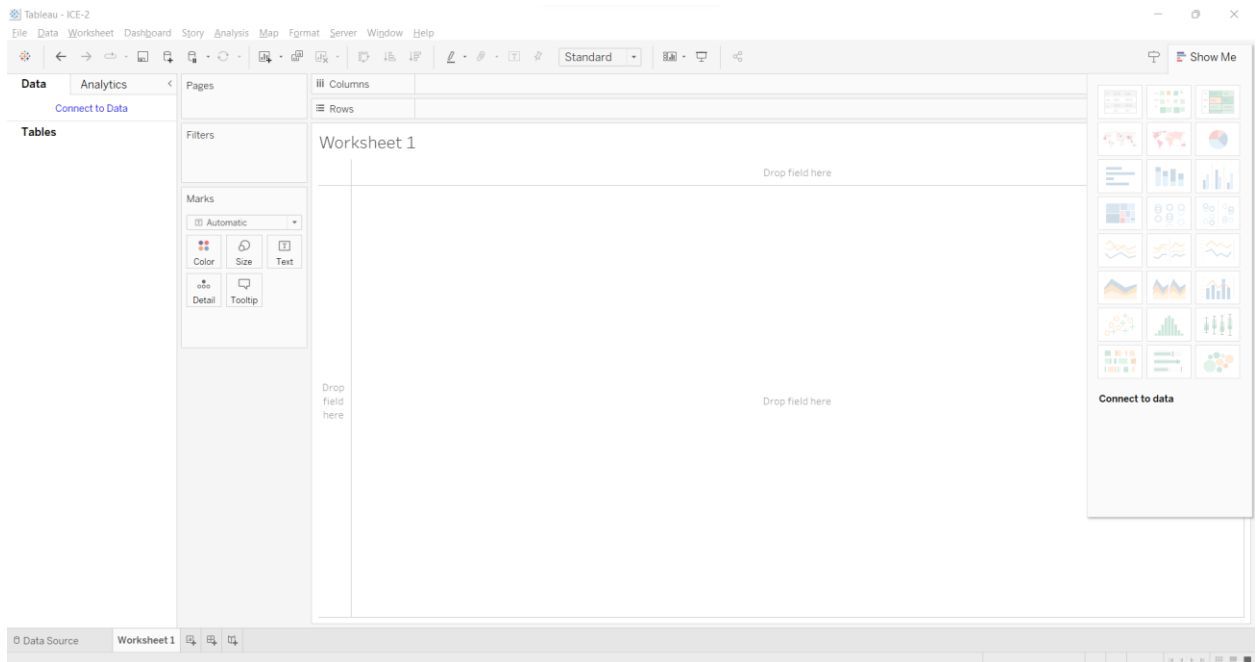
Introduction:

The SuperStore data consist of order details of sales throughout the United States of America between the years 2014-2017. The country is divided into 4 main regions Central, East, West, and South. The products for sale are categorized and sub-categorized. The Data contains order details, product details, shipping details, sales details, and customer details. Each item consists of attributes like Order Id, Order Date, Ship Date, Ship Mode, Customer ID, Customer Name, Segment, Country, City, State, Postal Code, Region, Product Id, Category, Sub-Category, Product Name, Sales, Quantity, Discount, and Profit. We shall be using all the data items to visualize the data for better understanding.

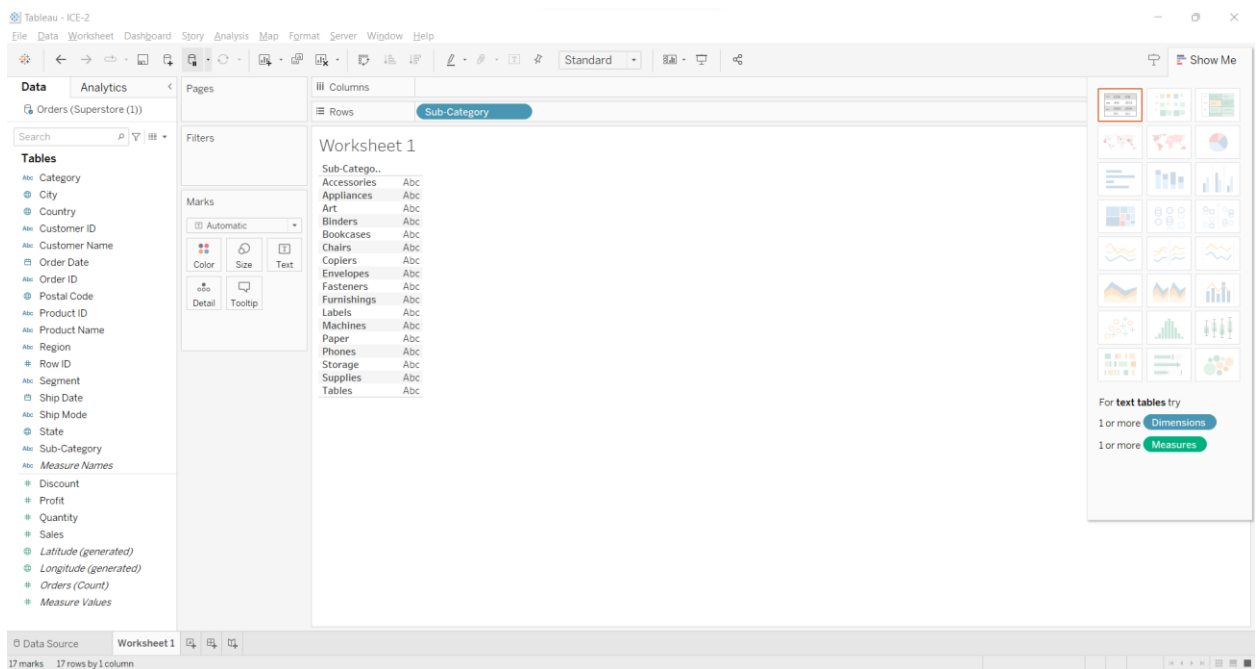
Methodology:

The provided data is huge and it might be tedious to analyze each and every data item individually and provide conclusions manually. So, we shall be using Tableau to visualize the given data for a better understanding of the data and to get conclusions about the data. Firstly, we load the data by connecting the data source to the tableau application. Then we use major attributes like Quantity, Sales, Region, State, Sales Dimensions, and Sub-categories to visualize the sales of different products across the country.

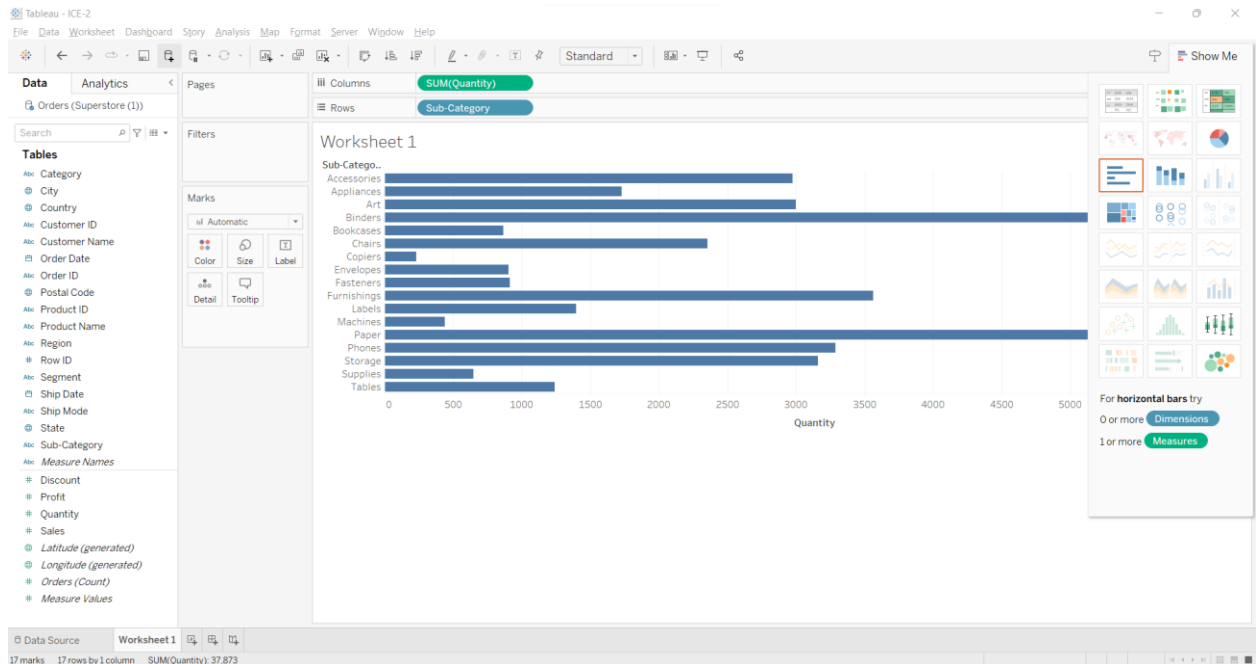
1. Open the Tableau Application and create a worksheet as Worksheet 1.



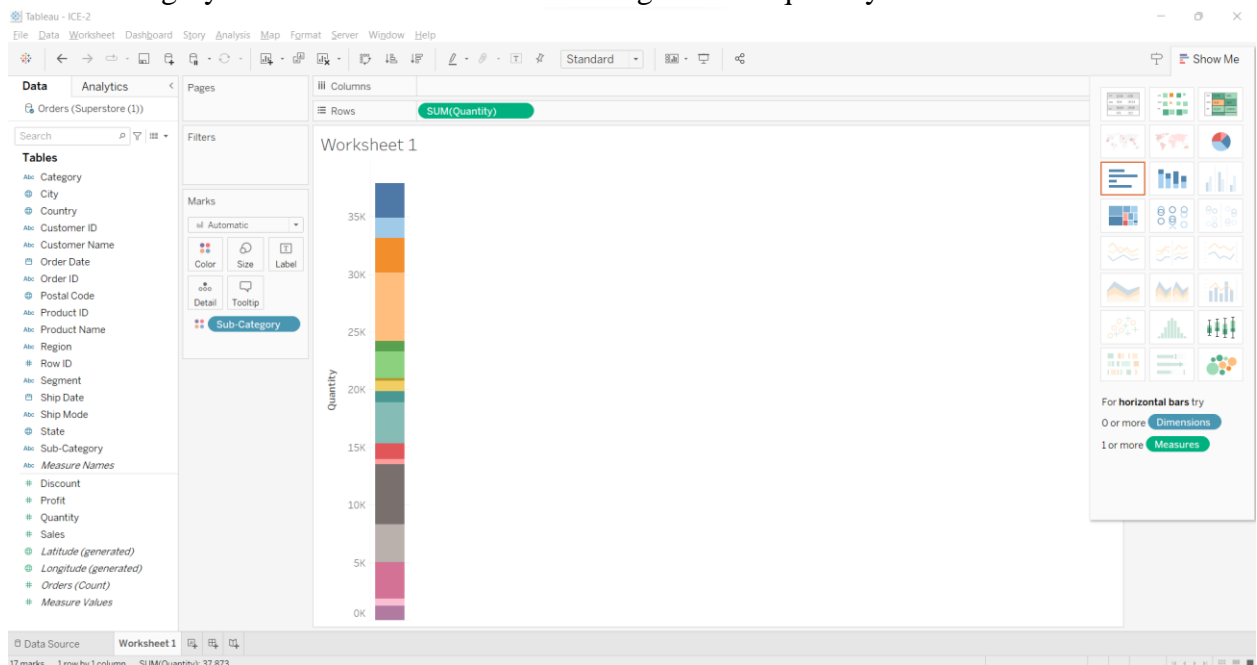
2. Connect to the given database and add the Sub-Category attribute in the rows shelf. The different sub-categories (Accessories, Appliances, Art, Binders, Bookcases, etc.) are shown below.



3. Now we add the Quantity attribute in the Column shelf. This shows the number of different subcategories based on the quantity data as shown below on the horizontal bar chart. From the below chart, we can observe that the quantity of Accessories is around 3000 and for furnishings is around 3500.

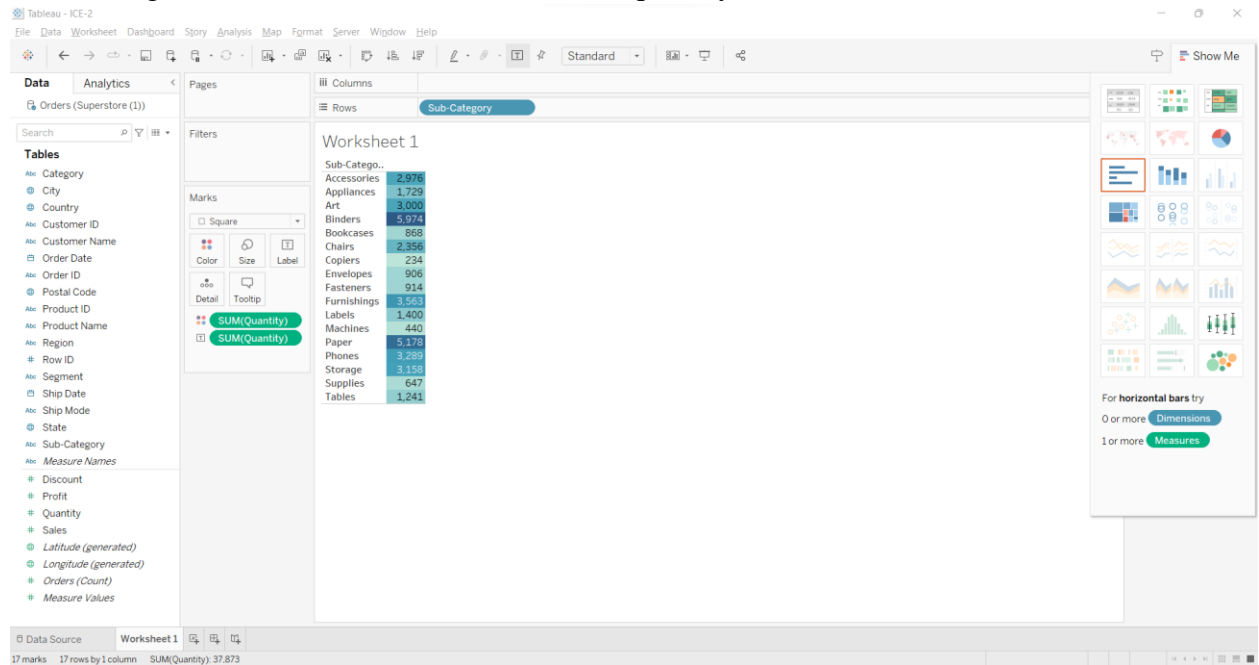


- The “Show me” button helps to select different visualizations. I have selected a stacked bar chart for visualization. In the below chart, all the sub-categories are color marked and are stacked one over the other showing the quantity of each sub-category. For example, the sub-category Art is color coded as Dark Orange and the quantity is 3000.

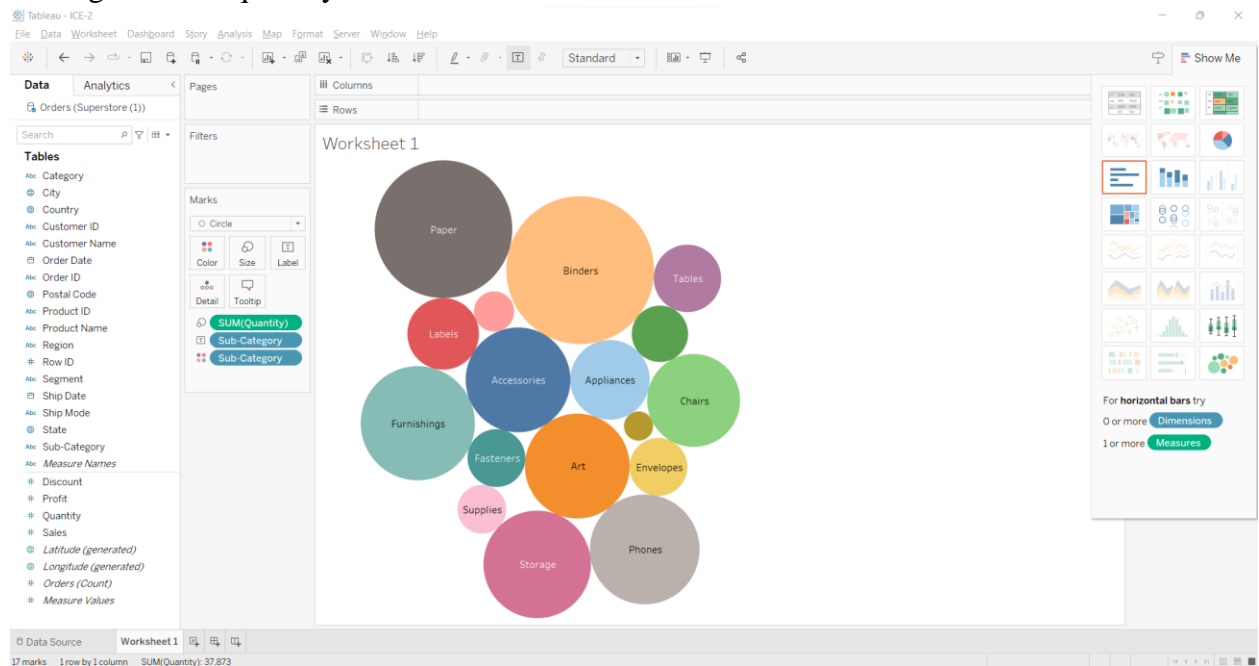


I have selected highlight tables for visualization. In the below chart, all the sub-categories are displaying the quantity of each sub-category. The quantities of each sub-category are color-coded based on the total quantity of each category, as the quantity keeps increasing the color gets darker. For example, the sub-category Binders is color coded as Dark Blue

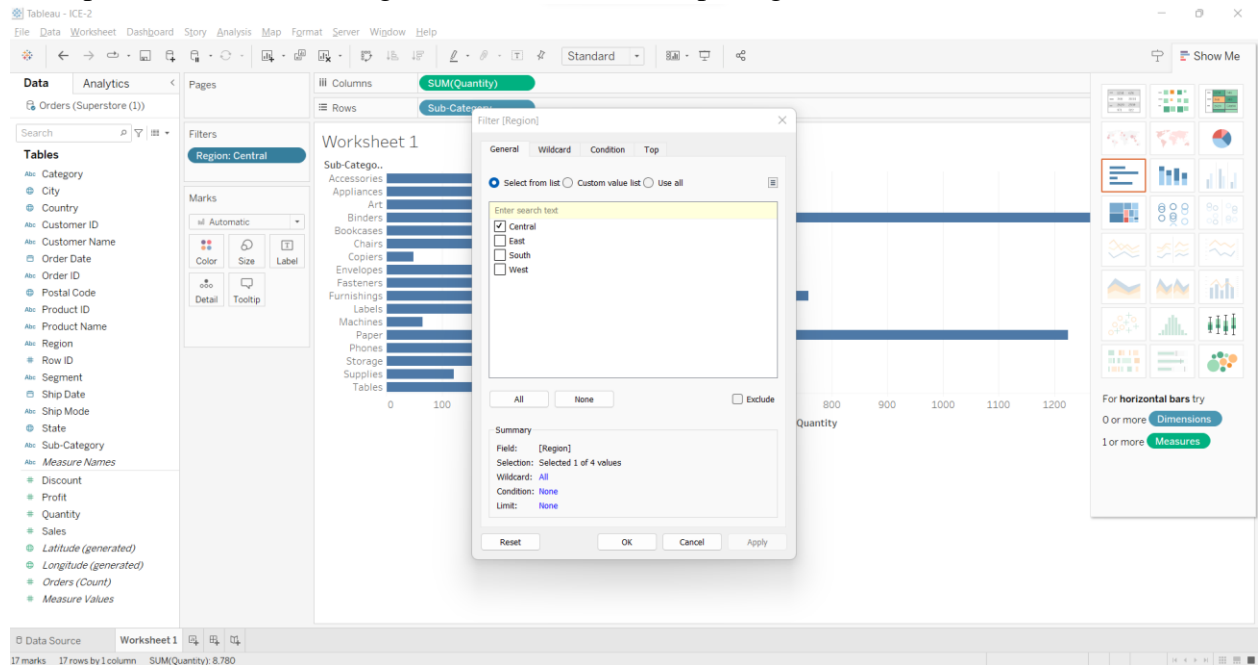
and the quantity is 5974, the color for Accessories, Art, Furnishings, Phones, and Storage is a little lighter than that of the Binders as their quantity is around 3000.



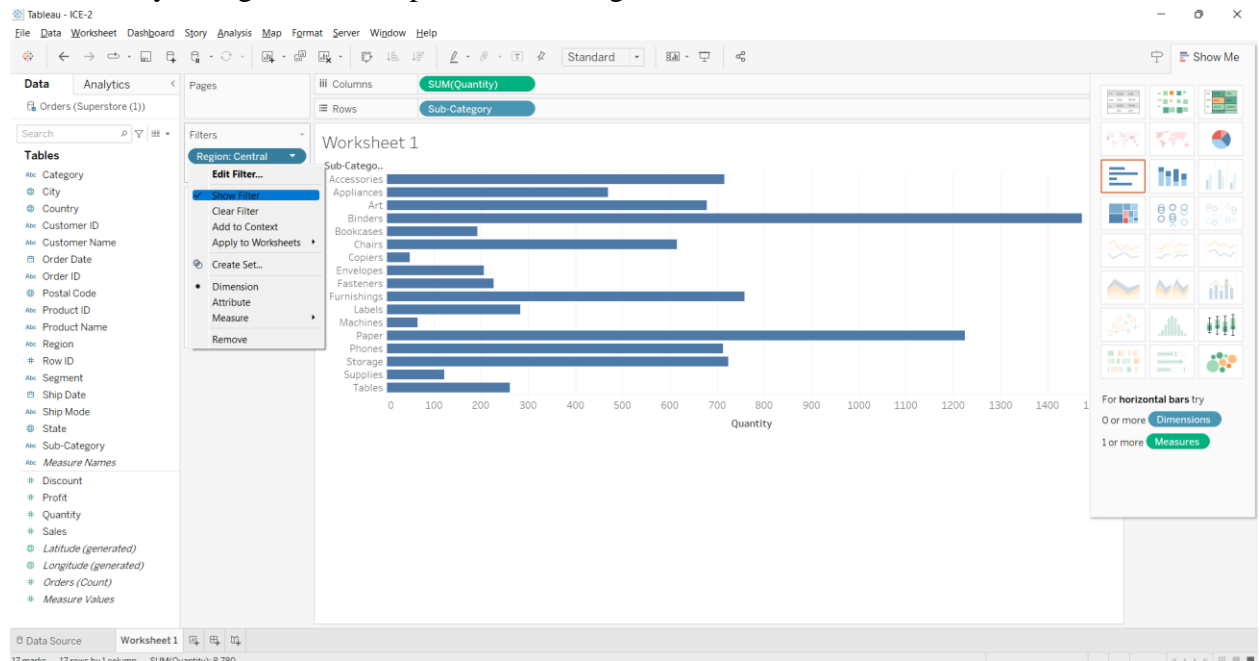
Now I have selected the packed bubbles for visualization. The different sub-categories are represented using different colors in the form of circles, and the quantity determines the size of the circle. As the quantity increases the size of the circle also increases, big the circle greater the quantity.



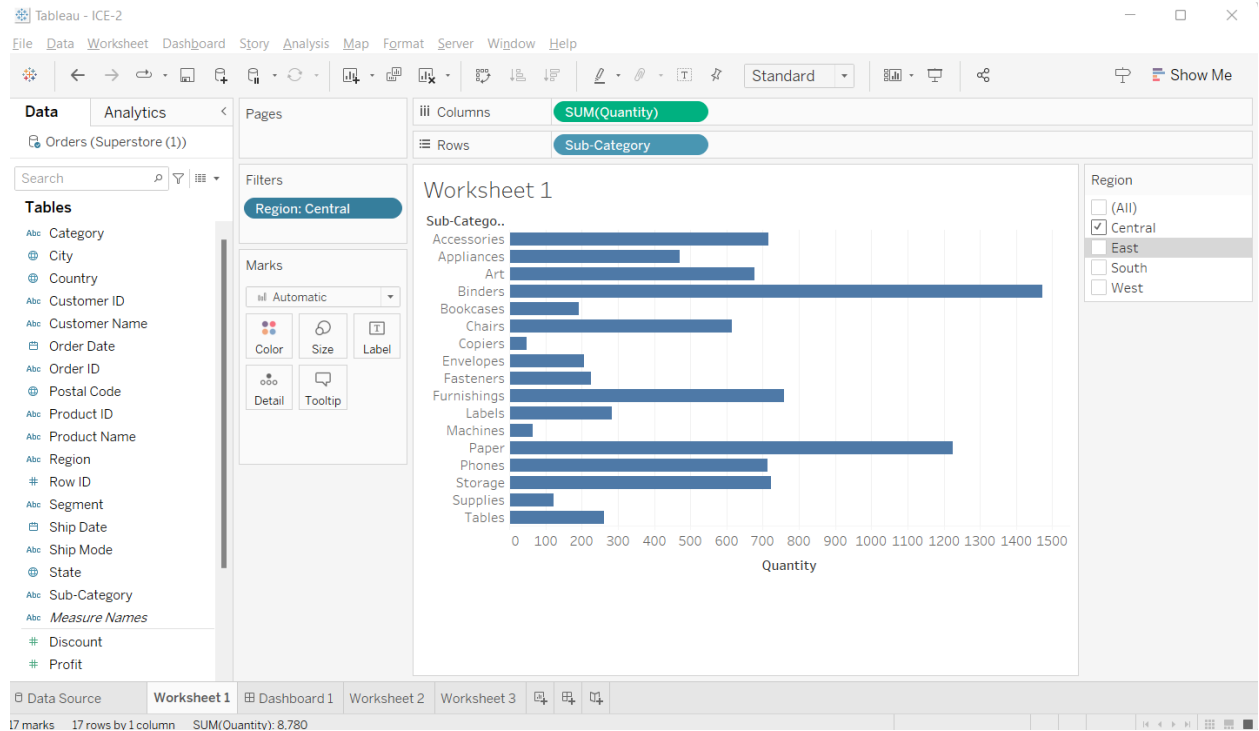
5. Now we add the region attribute to the filters, I have selected the central region, to get data specific to the central region, we can select multiple regions as well.



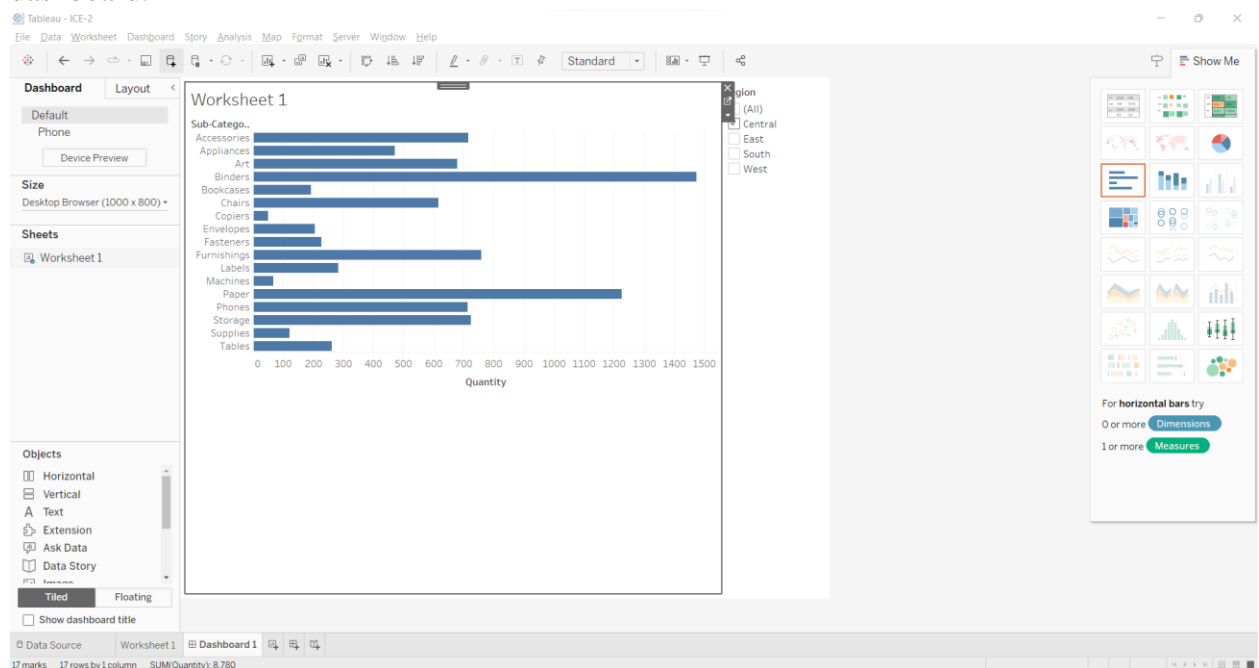
Now I have selected the show filter option to display the filters on the right to interactively change the filters specific to the region attribute.



As shown below, we can interactively change the regions to analyze the quantity data specific to the region. Here I have selected the Central region.

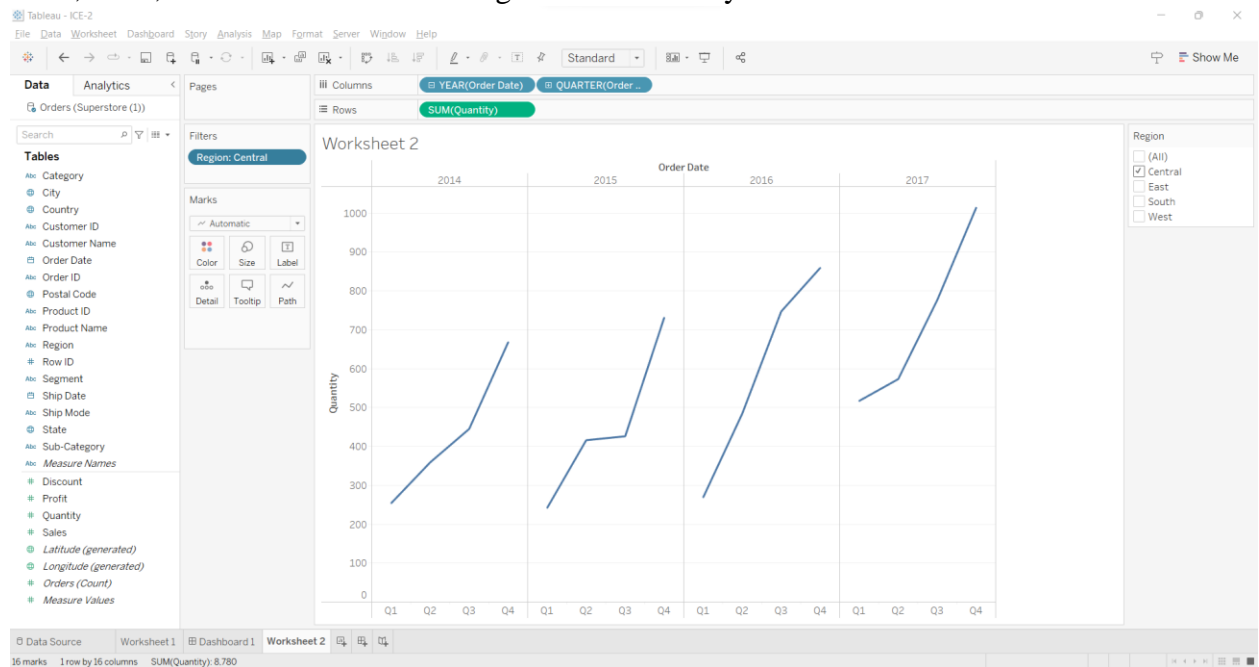


- Now, we create a new dashboard as Dashboard 1, from the Sheets section on the left, we can drag and drop worksheet 1 onto the dashboard to make the worksheet float on the dashboard.

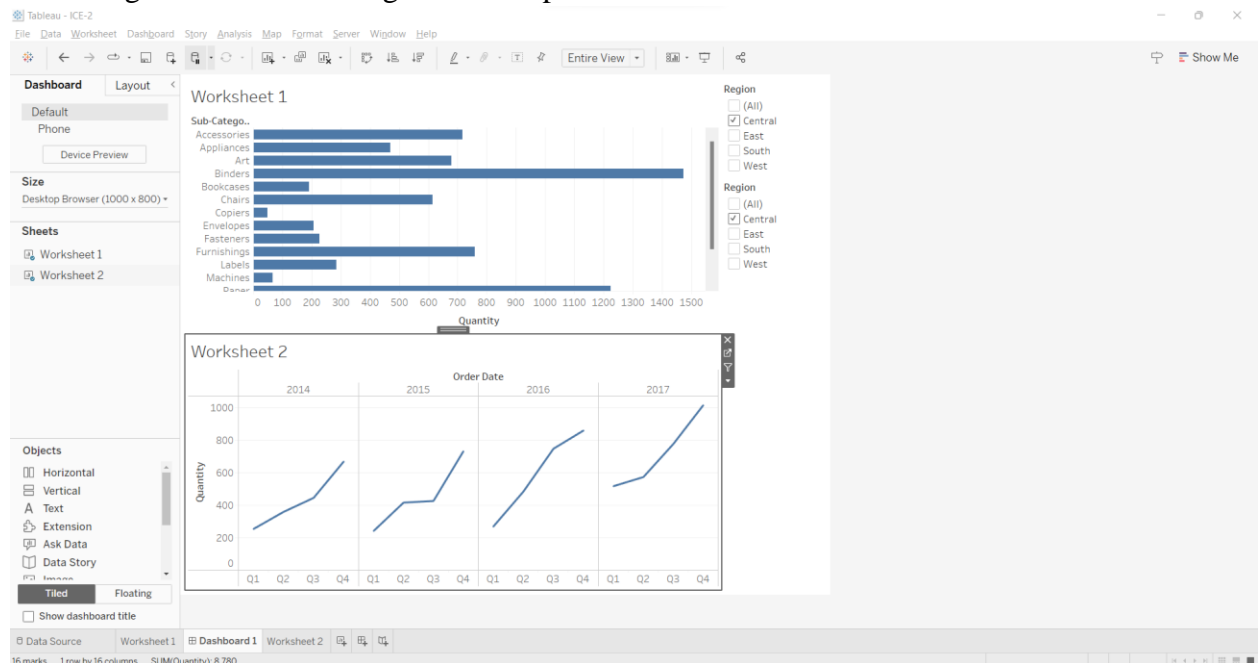


- Now we create another worksheet as Worksheet 2, we add the Quantity attribute to the rows and the Order Date attribute to the Column. This shows the graph of Quantity v/s

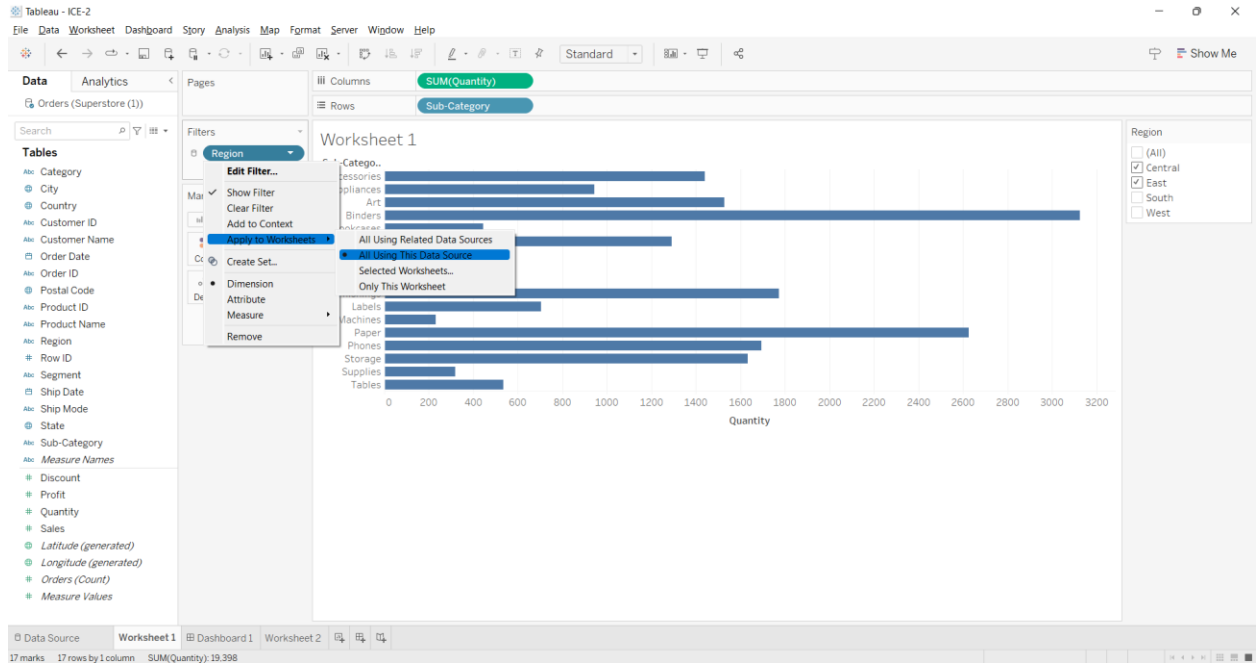
Order Date(year), it shows the increase in quantity from the order date in 2014, 2015, 2016, and 2017 years. We need to click on the “+” on the Year of order date attribute to get the Quarter of the Order Date, now the data shows the quantity of products with order dates in years 2014, 2015, 2016, and 2017. Also, each year is divided into four quarters. Now, we add the region attribute to the filters and select the show filter. This will show the data specific to the selected region i.e., either Central, East, West, or South. I have selected the filter Central, which shows the quantity of products sold in the year 2014, 2015, 2016, and 2017 in the central region of the country.



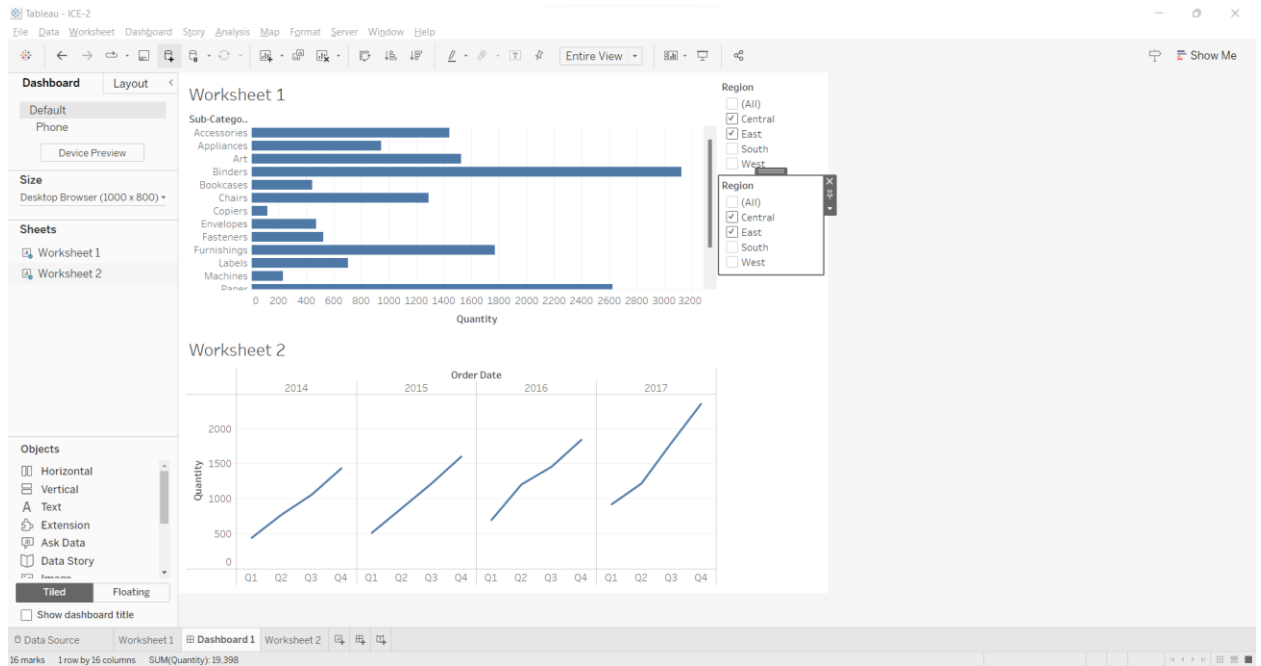
Now we go to Dashboard 1 again and drop Worksheet 2 into it.



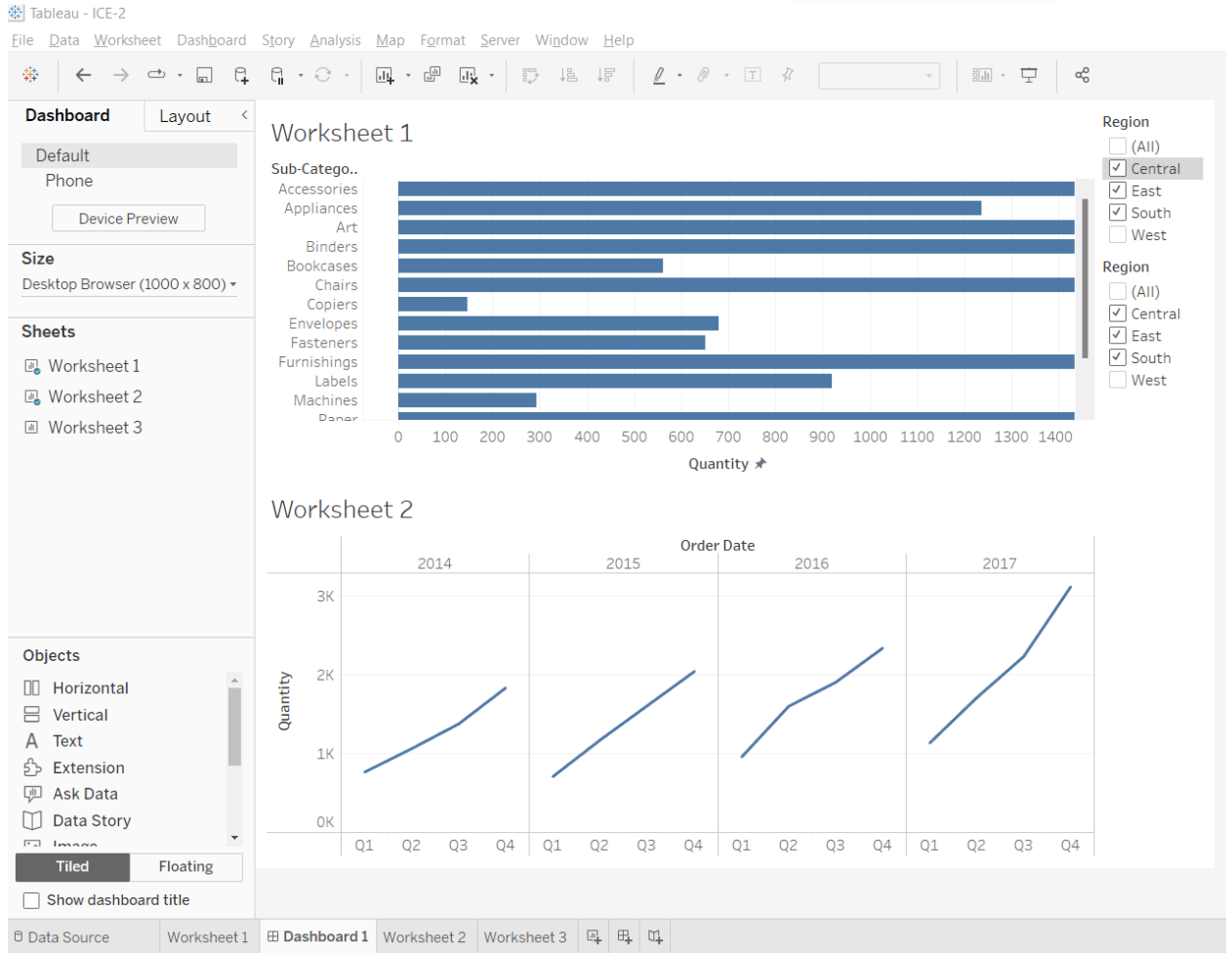
8. On worksheet 1, the region is already added to the filters, now right click on the Region in the Filters and choose the option “Apply to worksheet” and from that choose the option “All using this Data Sources”. By doing this we apply the region filter to the worksheet and to all the other places where this worksheet will be used.



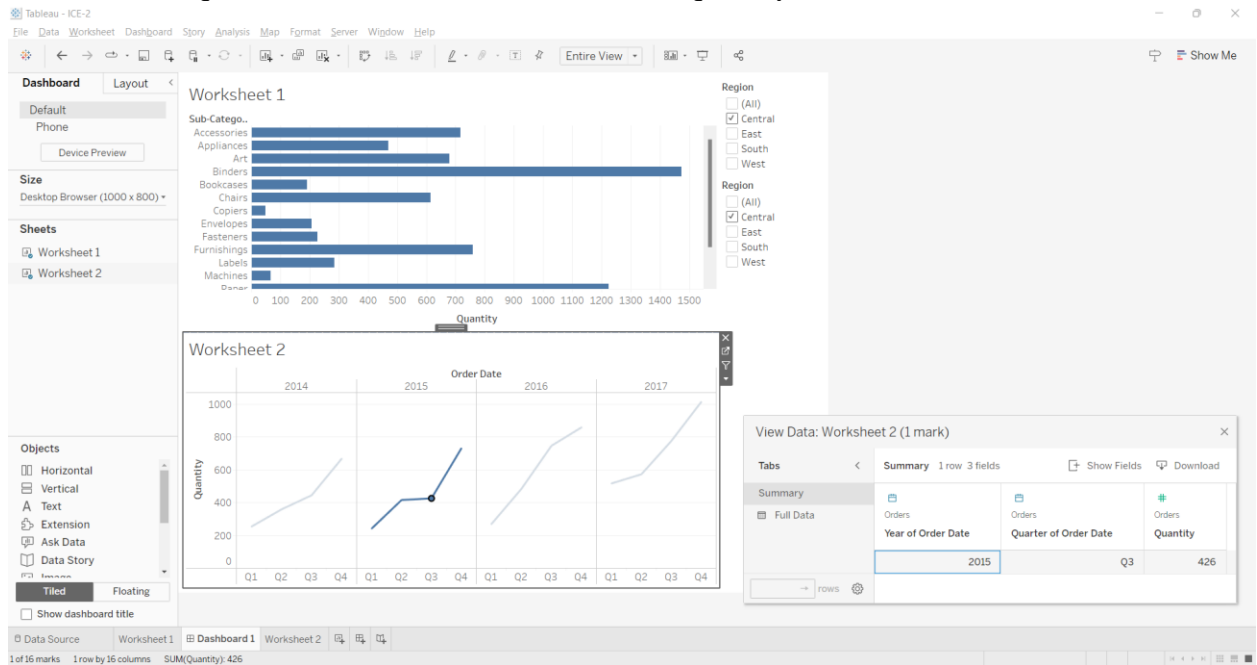
Now, we can see that the two worksheets are interactive, changing the filters in the first one changes the data in worksheet 2 as well. Here I have selected two regions Central and East, which shows the quantity of each sub-category of products sold in the Central and East region in worksheet 1 as well as the quantity of products sold in the year 2014, 2015, 2016, and 2017 in the central and east region of the country in worksheet 2.



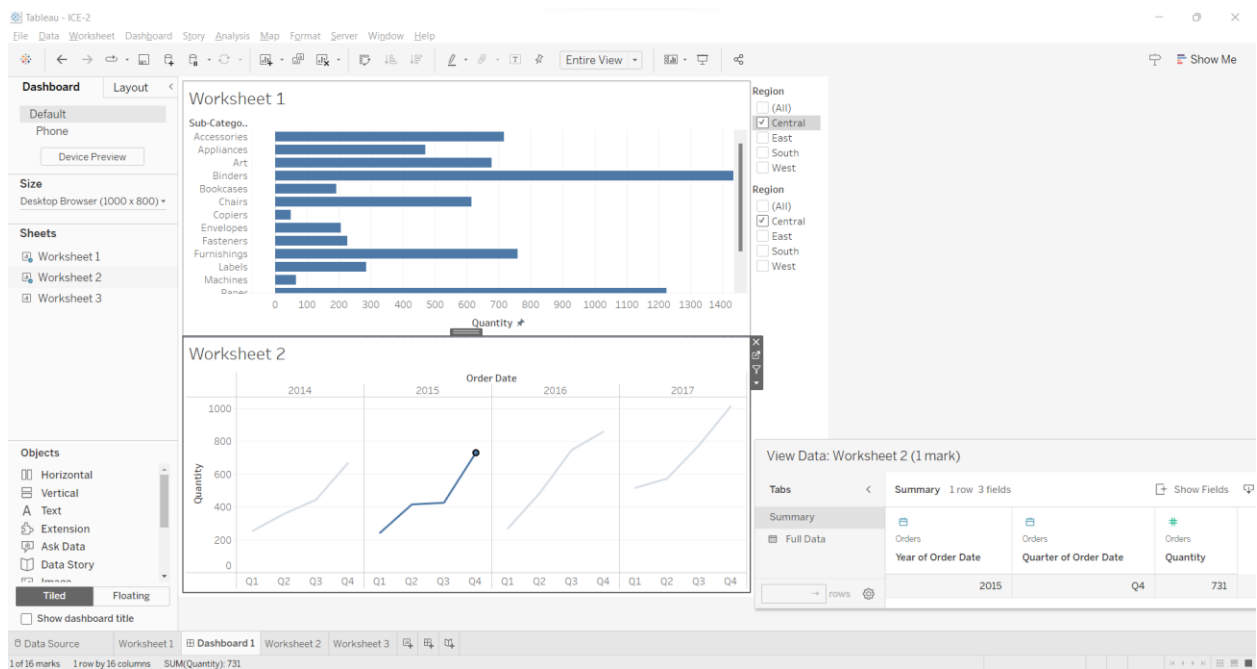
Now I change the filter by checking the South region. It changes the data in both worksheets as shown below.



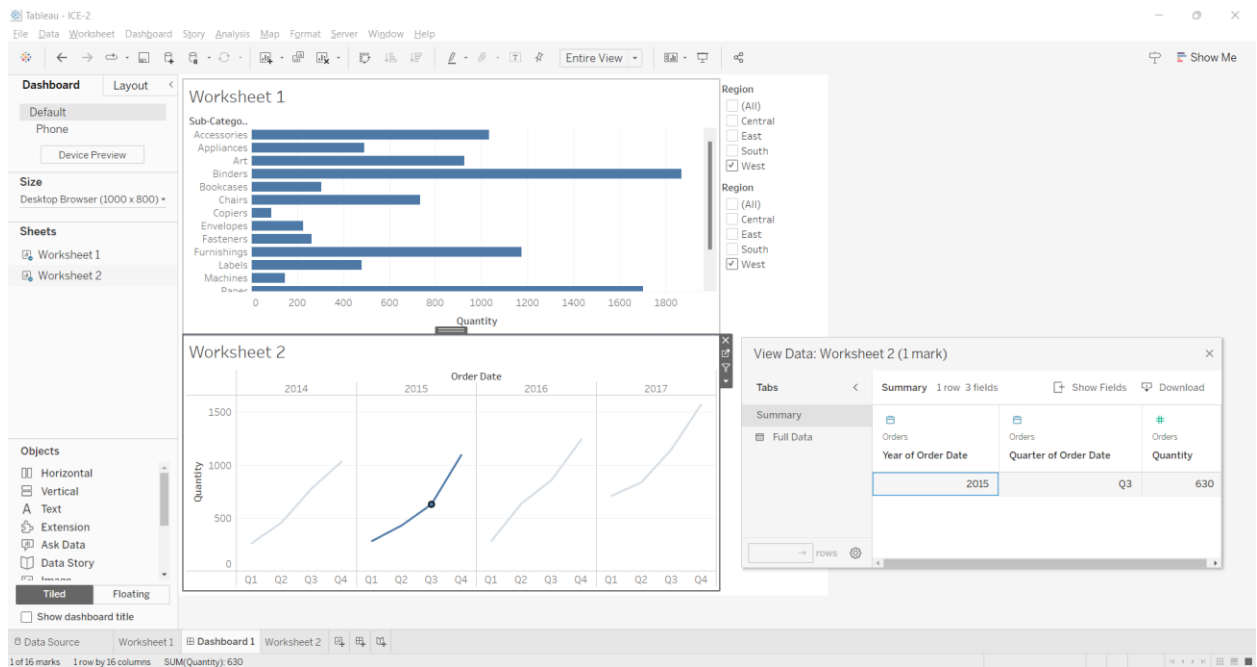
9. Now on worksheet 2 of dashboard 1, I have selected the Central region and a data point from the third quarter of order date in 2015, here the quantity is 425.



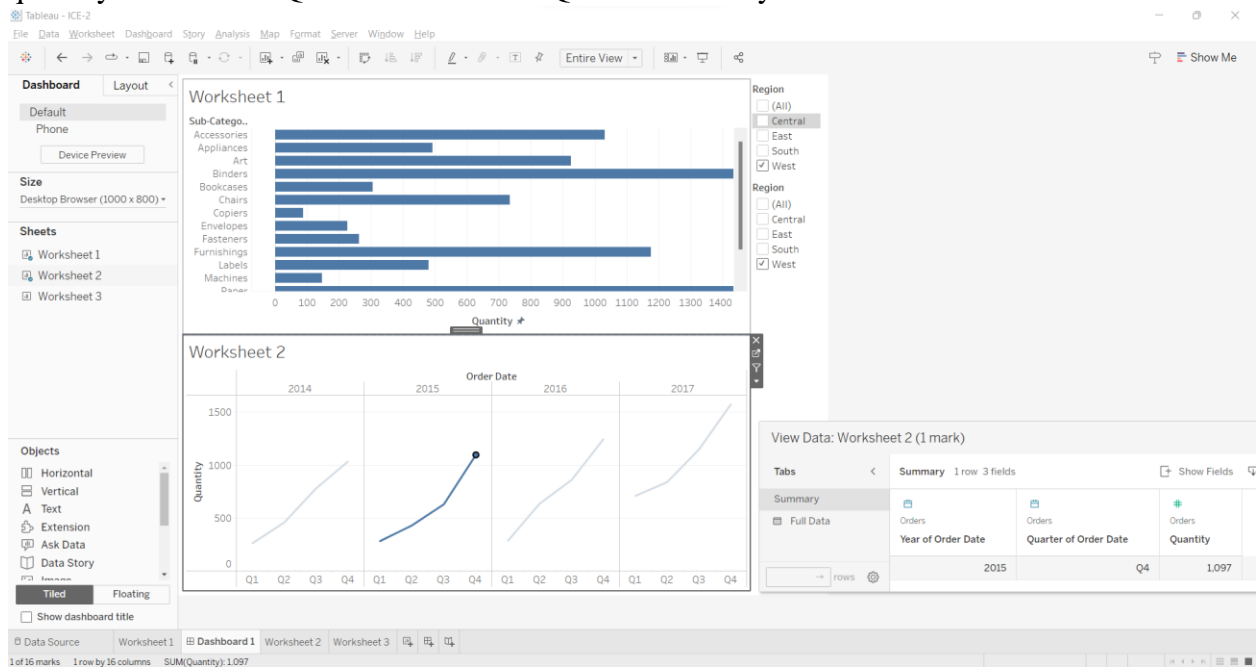
I have selected the same Central region and a different data point from the fourth quarter of the order date in 2015, here the quantity is 731. We can observe the increase in quantity in 2015 from Quarter 3 to Quarter 4.



Now, I have selected a different region i.e. West. I have selected the West region and a data point from the third quarter of order date in 2015, here the quantity is 630.

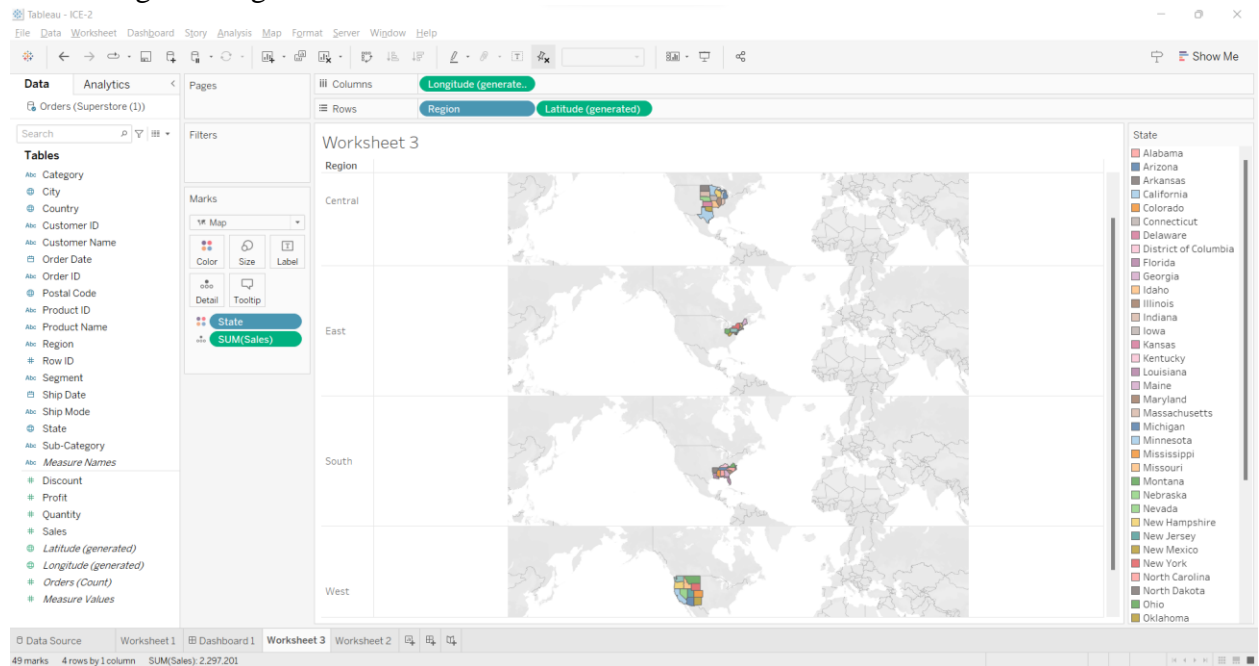


I have selected the same Central region and a different data point from the fourth quarter of the order date in 2015, here the quantity is 1097. We can observe the increase in quantity from 630 in Quarter 3 to 1097 in Quarter 4 in the year 2015.

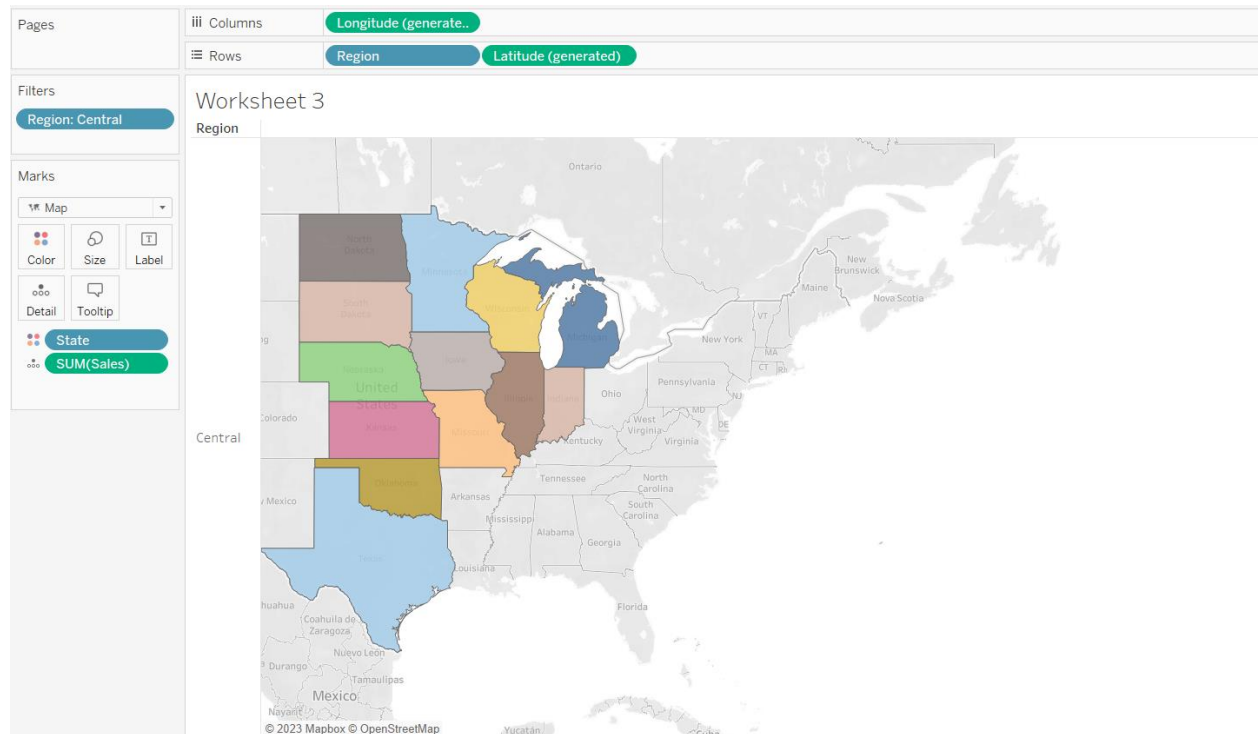


- Finally we create another worksheet as Worksheet 3, Now we assign the Region attribute to the row field and Latitude attribute of the Sales Dimension to the row field and the Longitude attribute of the Sales Dimension to the column field and then finally we add the state attribute to the color marks and the sales attribute to the details. Once all the attributes are added, I created a map-based visualization. Firstly, the Longitude and

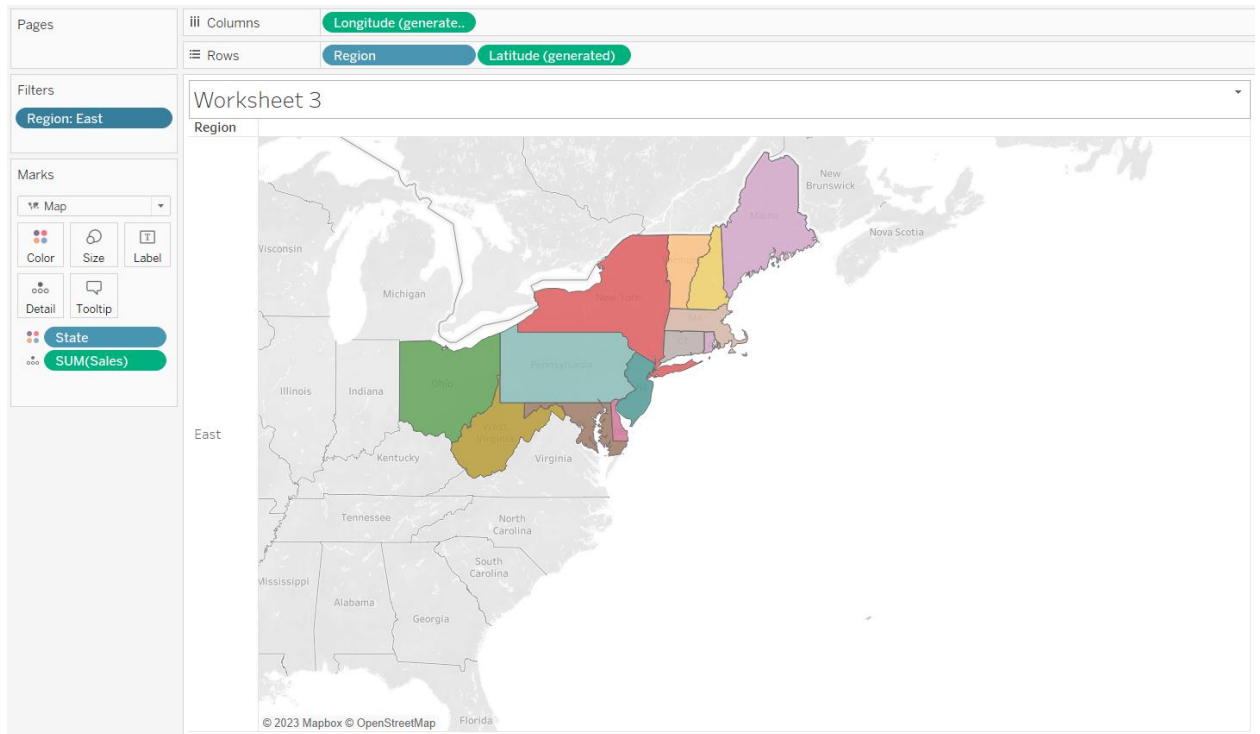
Latitude of each state of the given regions are marked on the map. Each State is color marked with different colors. The details of total sales in each state can be seen just by hovering over each state. The below map shows the total sales of each state color marked concerning each region.



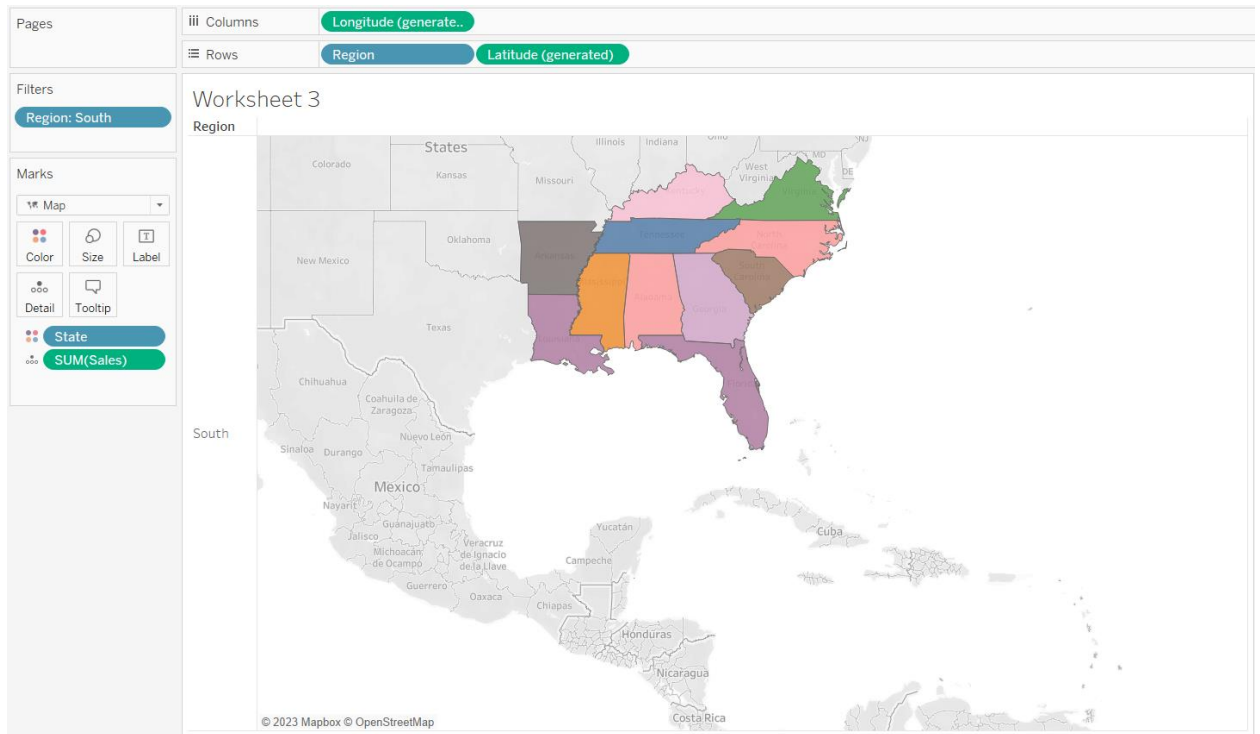
Below is the map visualization for Central region sales.



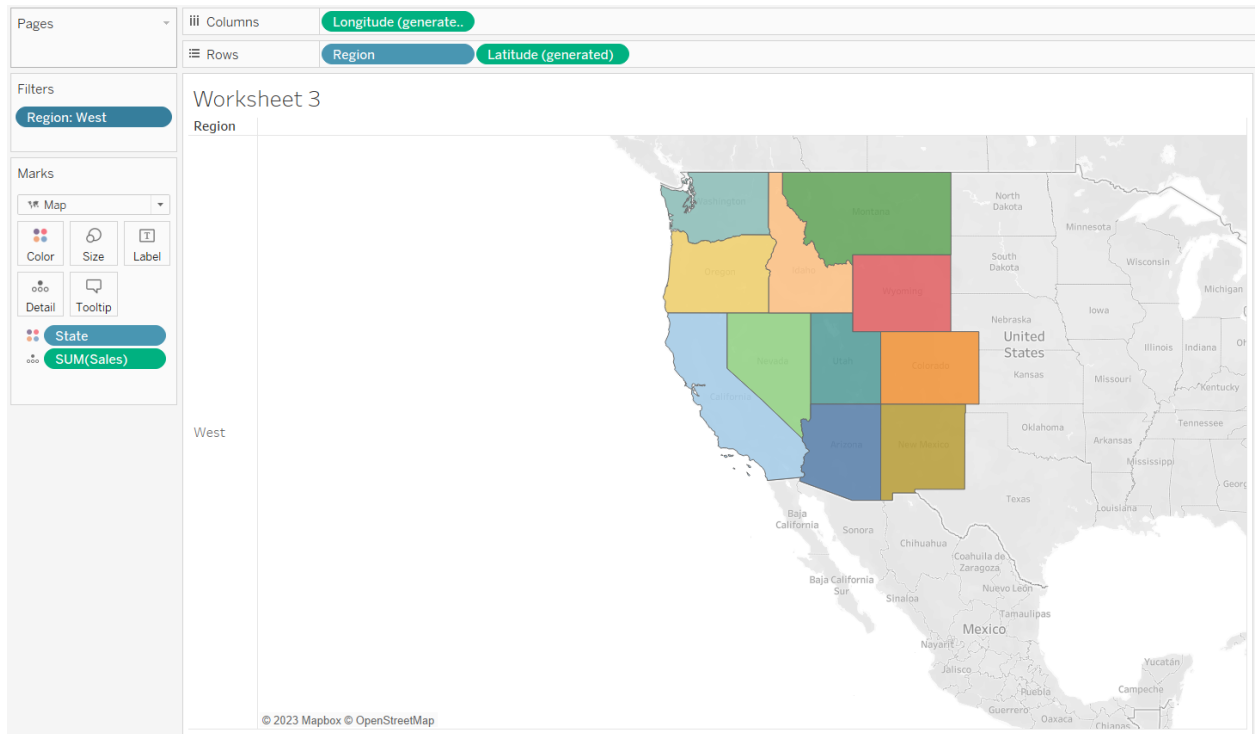
Below is the map visualization for East region sales.



Below is the map visualization for South region sales.



Below is the map visualization for West region sales.



Conclusion:

After visualizing the data, we can provide different views of the data. Worksheet 1 shows the quantity of different sub-categories of products that have been sold between the years 2014-2017 inclusive. For example, the quantity of Machines sold in 4 years is 440, Copiers sold is 234, Appliances sold is 1729, etc. Worksheet 2 shows the quantity of products sold in each quarterly year separately and how the sales increased throughout the year. For example, for the whole data, In the quantity of products sold irrespective of the category is 2871, 3144, 3587, and 4696 in the years 2014, 2015, 2016, and 2017 respectively. Worksheet 3 displays the map-based visualization, it shows the sales data concerning each state in respective regions. For example, Sales in the Texas State in Central Region is 170188. Dashboard 1, displays both Worksheet 1 and Worksheet 2 Data by interactively changing the regions.