

## 1. Connect Source Data

Open the source data Excel Sample-Superstore.xls. It will be connected and all 3 different sheets will be showing, drag Order sheet to the up right space and the detailed information of it will be shown at bottom, and order sheet will be used as a source data during the following process.

In the left side bar, the categorically descriptive fields will be shown on top and quantitative fields are at bottom, which can be used to do calculation in further process.

Tableau Public - Book1

Connections: Sample-Superstore (Microsoft Excel)

Sheets: Orders, People, Returns

Need more data? Drag tables here to relate them. Learn more

Sort fields: Data source order

Show aliases Show hidden fields 1,000 + rows

| Row ID | Order ID       | Order Date | Ship Date | Ship Mode      | Customer ID | Customer Name    | Segment     | Country       | City         | State        | Postal Code | Region |
|--------|----------------|------------|-----------|----------------|-------------|------------------|-------------|---------------|--------------|--------------|-------------|--------|
| 7981   | CA-2011-103896 | 3/1/2013   | 7/1/2013  | Standard Class | DP-19000    | Darren Powers    | Consumer    | United States | Houston      | Texas        | 77095       | Co     |
| 740    | CA-2011-112326 | 4/1/2013   | 8/1/2013  | Standard Class | P0-19195    | Phillina Ober    | Home Office | United States | Naperville   | Illinois     | 60540       | Co     |
| 741    | CA-2011-112326 | 4/1/2013   | 8/1/2013  | Standard Class | P0-19195    | Phillina Ober    | Home Office | United States | Naperville   | Illinois     | 60540       | Co     |
| 742    | CA-2011-112326 | 4/1/2013   | 8/1/2013  | Standard Class | P0-19195    | Phillina Ober    | Home Office | United States | Naperville   | Illinois     | 60540       | Co     |
| 1760   | CA-2011-141817 | 5/1/2013   | 12/1/2013 | Standard Class | MB-18005    | Mick Brown       | Consumer    | United States | Philadelphia | Pennsylvania | 19143       | Ea     |
| 5328   | CA-2011-130813 | 6/1/2013   | 8/1/2013  | Second Class   | LS-17230    | Lycoris Saunders | Consumer    | United States | Los Angeles  | California   | 90049       | Wt     |
| 7181   | CA-2011-060603 | 6/1/2013   | 7/1/2013  | First Class    | JO-15145    | Jack O'Briant    | Corporate   | United States | Athens       | Georgia      | 30605       | So     |
| 7475   | CA-2011-167199 | 6/1/2013   | 10/1/2013 | Standard Class | ME-17320    | Maria Etezadi    | Home Office | United States | Henderson    | Kentucky     | 42420       | So     |
| 7476   | CA-2011-167199 | 6/1/2013   | 10/1/2013 | Standard Class | ME-17320    | Maria Etezadi    | Home Office | United States | Henderson    | Kentucky     | 42420       | So     |
| 7477   | CA-2011-167199 | 6/1/2013   | 10/1/2013 | Standard Class | ME-17320    | Maria Etezadi    | Home Office | United States | Henderson    | Kentucky     | 42420       | So     |
| 7478   | CA-2011-167199 | 6/1/2013   | 10/1/2013 | Standard Class | ME-17320    | Maria Etezadi    | Home Office | United States | Henderson    | Kentucky     | 42420       | So     |
| 7479   | CA-2011-167199 | 6/1/2013   | 10/1/2013 | Standard Class | ME-17320    | Maria Etezadi    | Home Office | United States | Henderson    | Kentucky     | 42420       | So     |
| 7480   | CA-2011-167199 | 6/1/2013   | 10/1/2013 | Standard Class | ME-17320    | Maria Etezadi    | Home Office | United States | Henderson    | Kentucky     | 42420       | So     |
| 7481   | CA-2011-167199 | 6/1/2013   | 10/1/2013 | Standard Class | ME-17320    | Maria Etezadi    | Home Office | United States | Henderson    | Kentucky     | 42420       | So     |

## 2. Sales by Region

All unique values of Region will be shown after dragging it into Columns and the sum for each region will be shown once simply double click on Quantity and Sales.

Drag Region into Colour to add different colours to different region.

Sheet 1

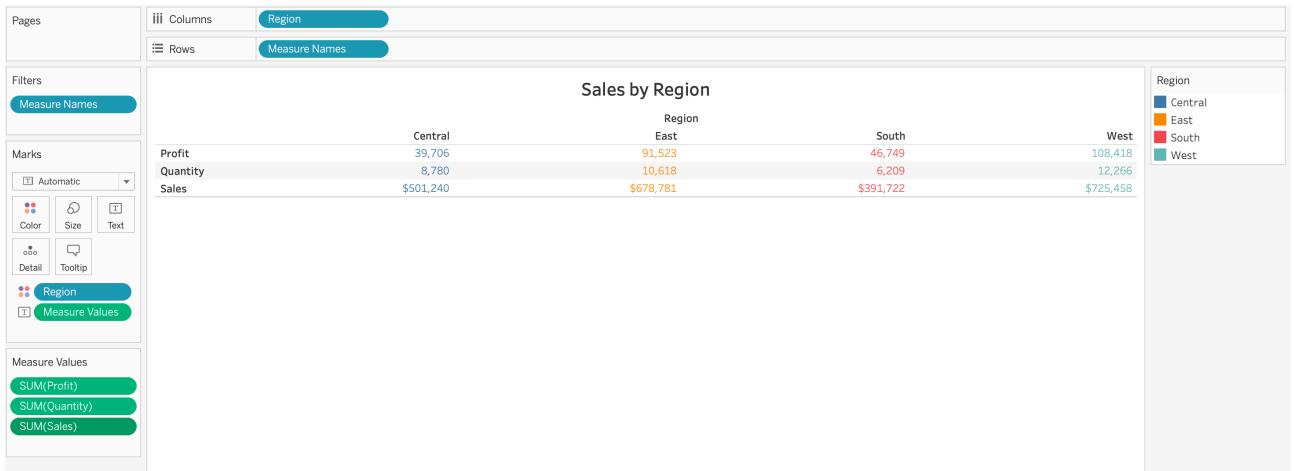
| Region  | Quantity | Sales   |
|---------|----------|---------|
| Central | 8,760    | 501,240 |
| East    | 10,618   | 678,781 |
| South   | 6,209    | 391,722 |
| West    | 12,766   | 725,458 |

Region: Central, East, South, West

Quantity: SUM(Quantity)

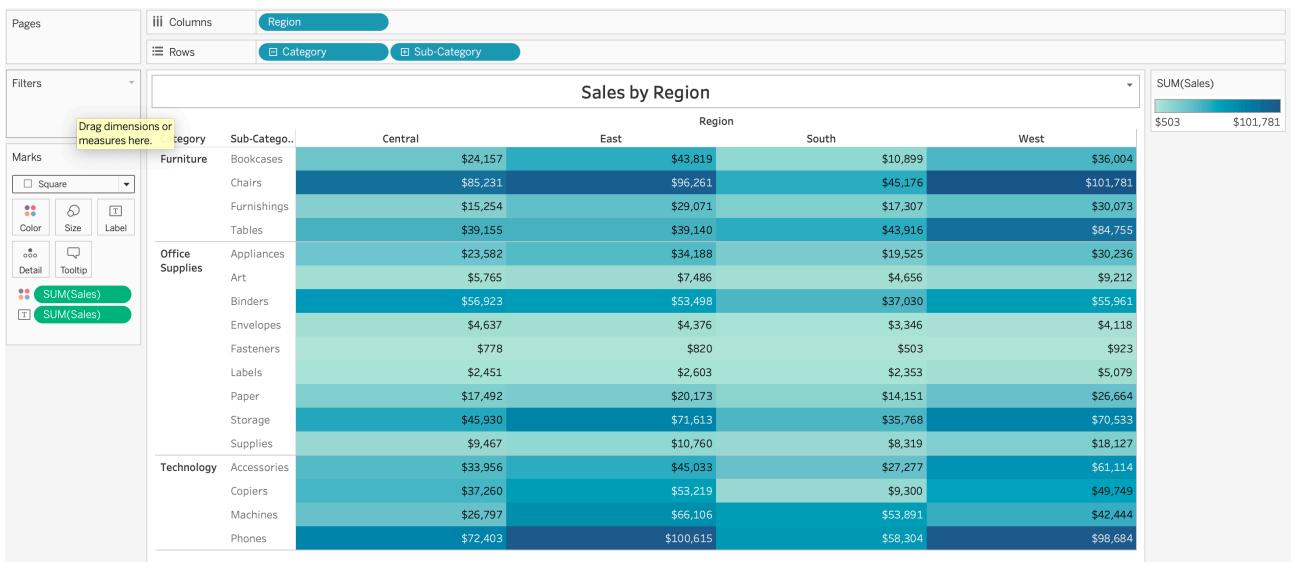
Sales: SUM(Sales)

Edit filters and more other quantitative fields can be chosen. Measure Values can be formatted, such as keep zero decimals and put a \$ sign before currency. Change Standard to Fit the width, edit the title.



Further process: add region hierarchy structure, the changes will be applied all the time after editing sale format with zero decimal with \$ sign at left measures names bar. Colour will fill in the whole square after choosing Square type.

The darkest square has the highest total sales and the sub-category can be folded up once clicking the “-” sign before category and only the detail of category will be showed.

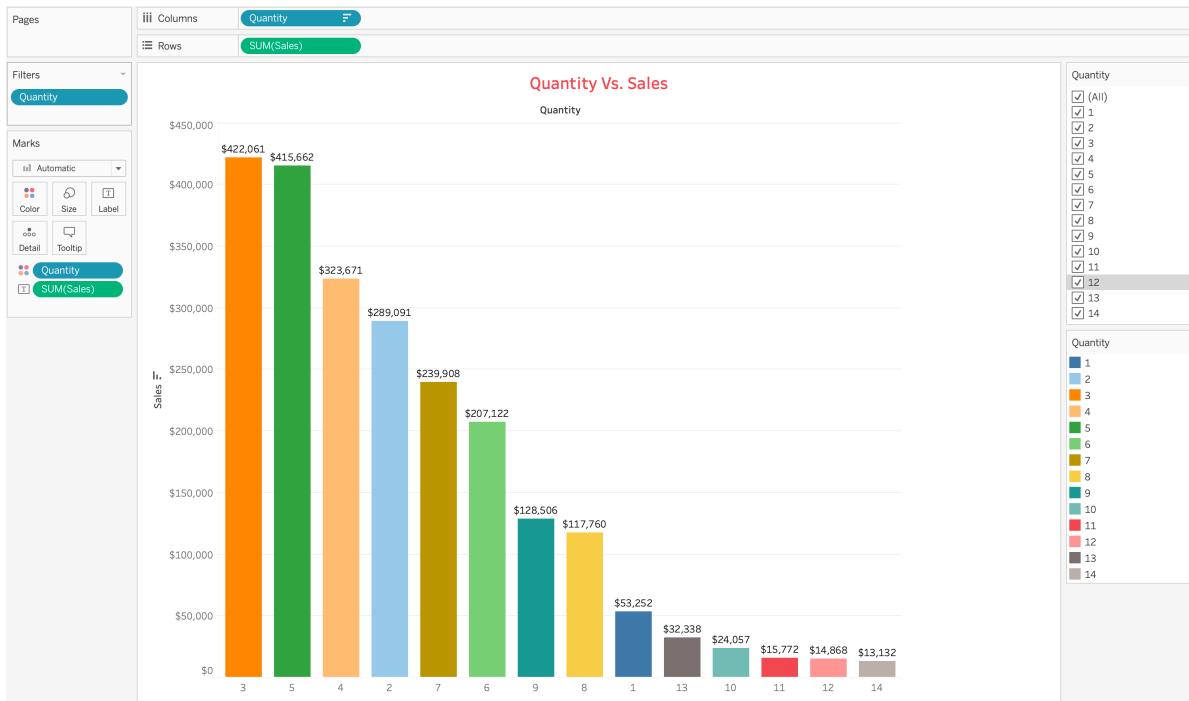


### 3. Quantity Vs. Sales

Since the quantity is finite number for the order, therefore, the correlation between quantity and sales can be visualised. Change quantity from continues into dimension and discrete type, then the sales for different quantities can be observed.



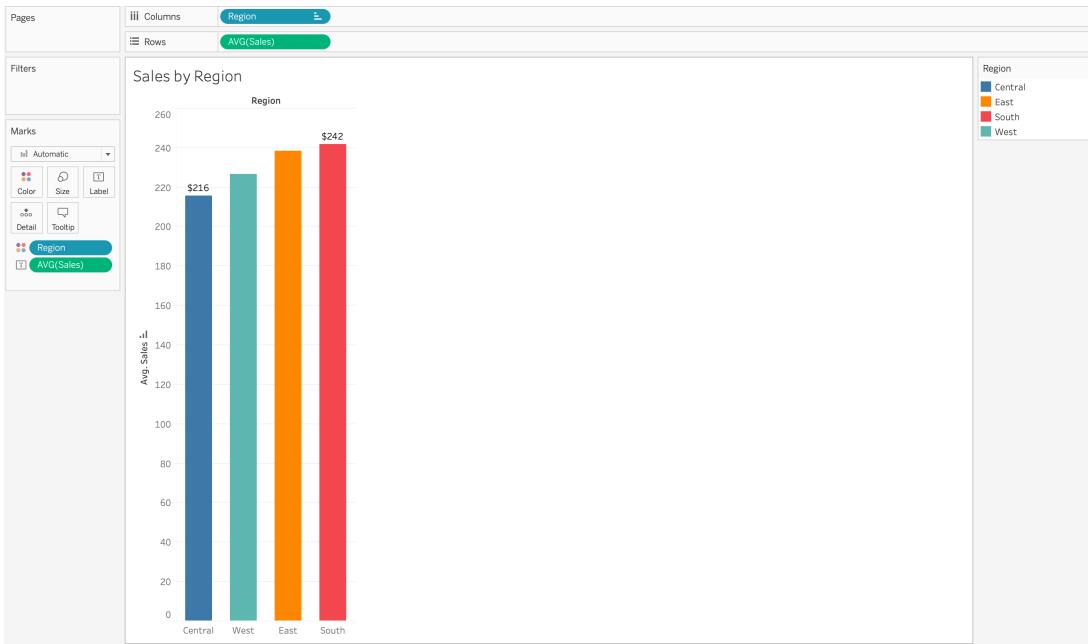
Filter and label features can be added to improve readability and interactive ability:



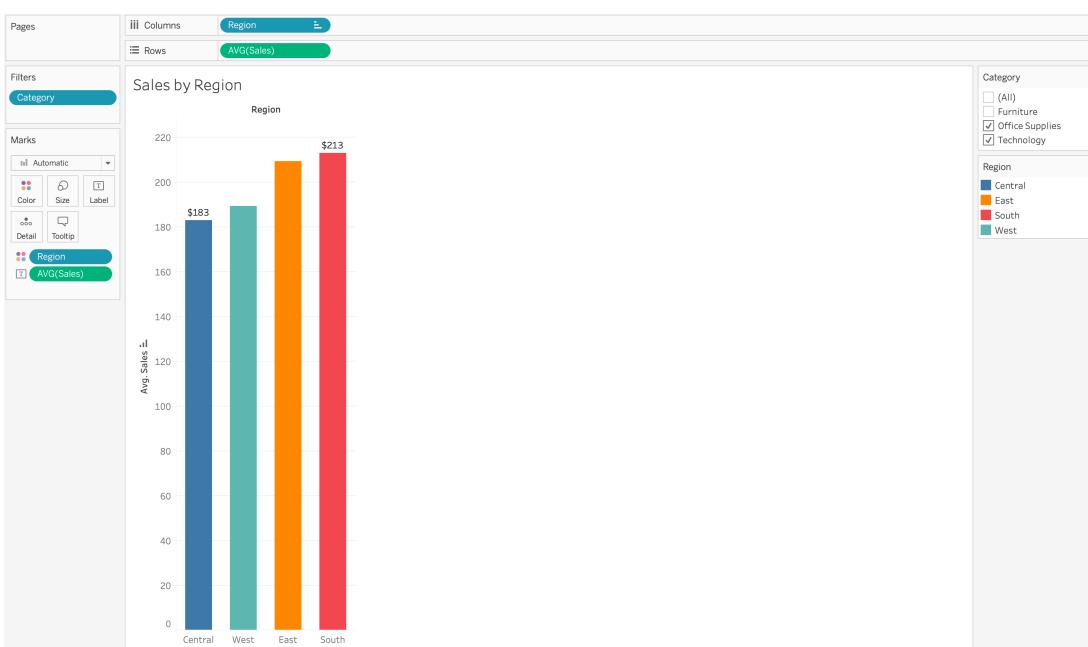
Therefore, the quantity of purchasing 3 or 5 products have a higher total sales number.

## 4. Sales by Region II

Different colour can be added by dragging region into colour or only click colour to change into one another single colour. The bar size can be adjusted through size mark. The detailed number of the sales can be added through label mark, and all labels for each bar or only some of them can be chosen through right clicking “Label” mark to choose.



Filter can be added added and show filter. The sales number will be actively changed once only some categories chose.

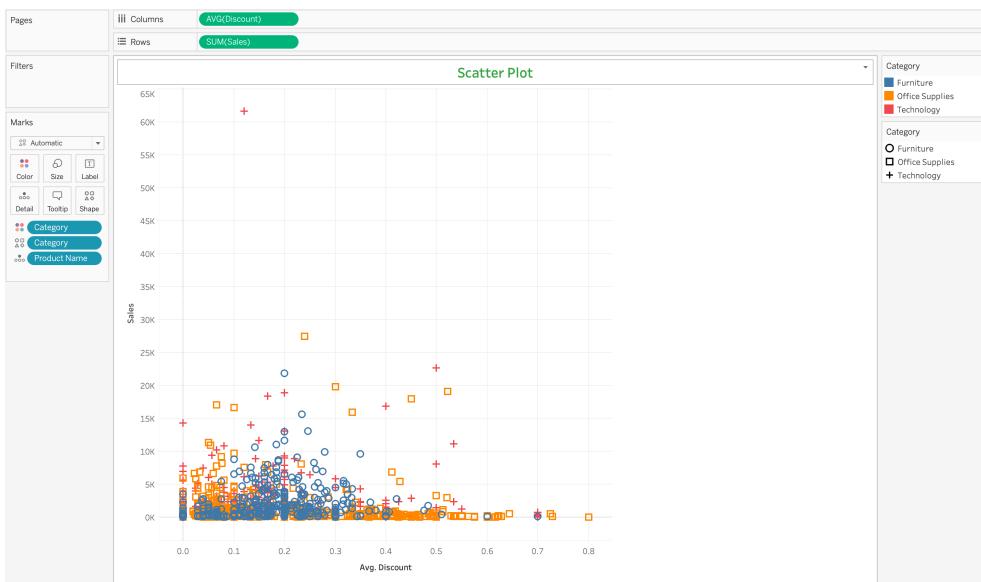


## 5. Scatter Plot

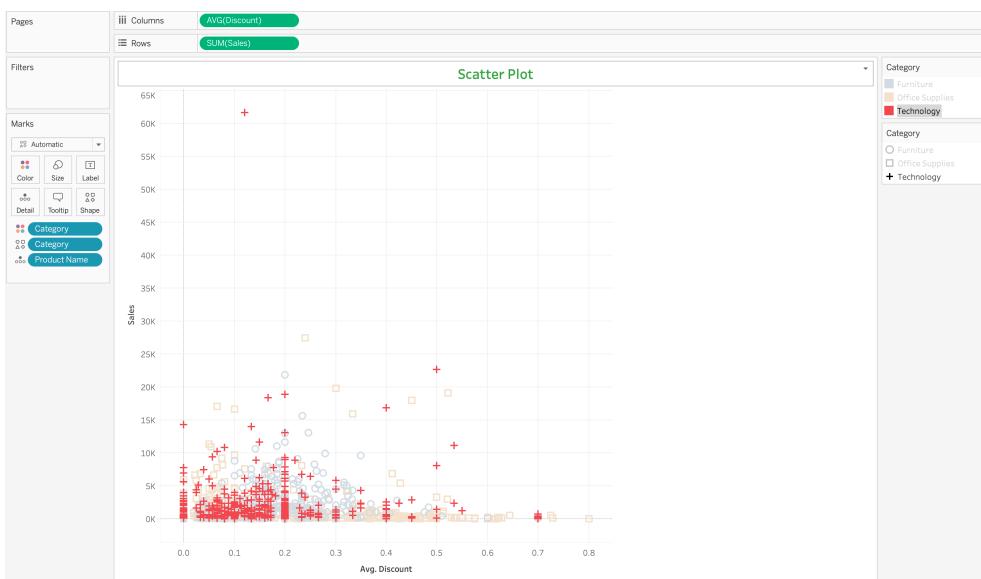
Scatter plot can be applied to study the correlation between discount and sales.

Put total sales in row and average discount in column, there is only one point showing, put product name into Detail Mark and drag category into colour and shape mark, then different category will show different colour and shape on the worksheet.

The reason why we did not choose to use product name to mark colour and shape is that there are too many different kinds of products though it can be applied.



Choose applying highlight in both category colour and shape legend, then both can be used to highlight one or more category.



## 6. Profit by Category

The application of hierarchy: there is a hierarchy structure between category, sub-category and product name. Therefore, drag subcategory and product to category and combine them into a hierarchy structure. Drag category hierarchy structure into columns and unfold it into category and sub-category.

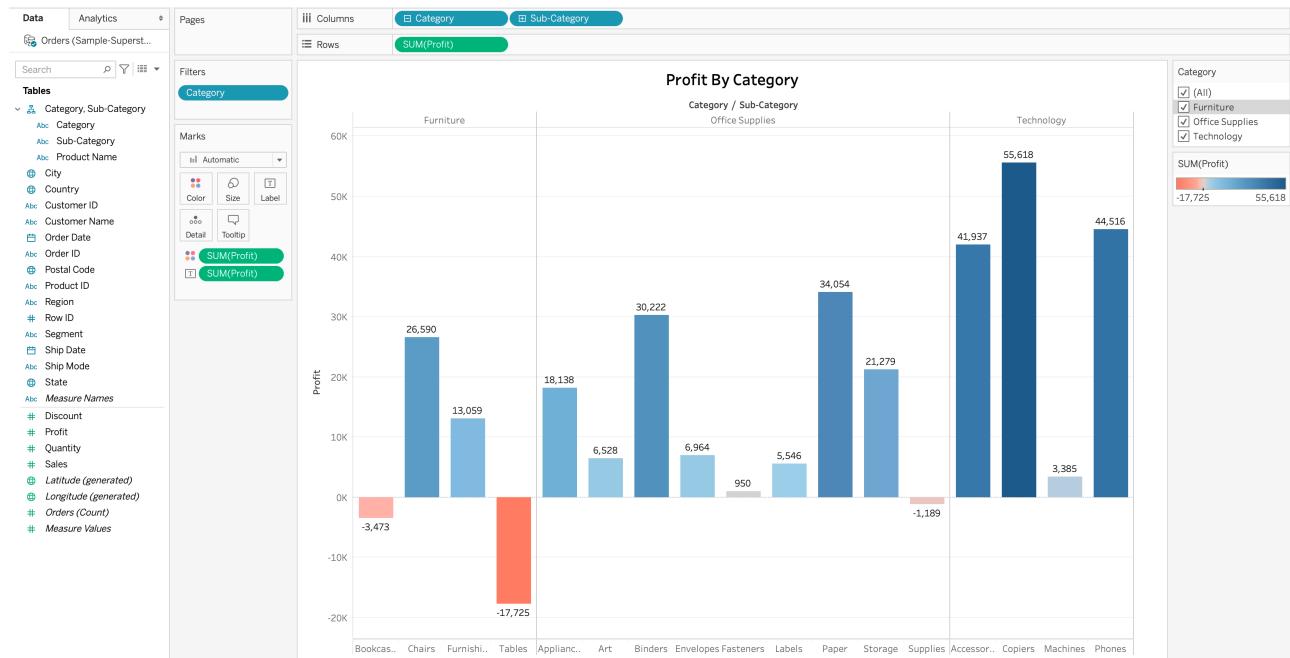
**Tables**

- Category, Sub-Category
  - Category
  - Sub-Category
  - Product Name**
- City
- Country

Columns Category Sub-Category

Rows SUM(Profit)

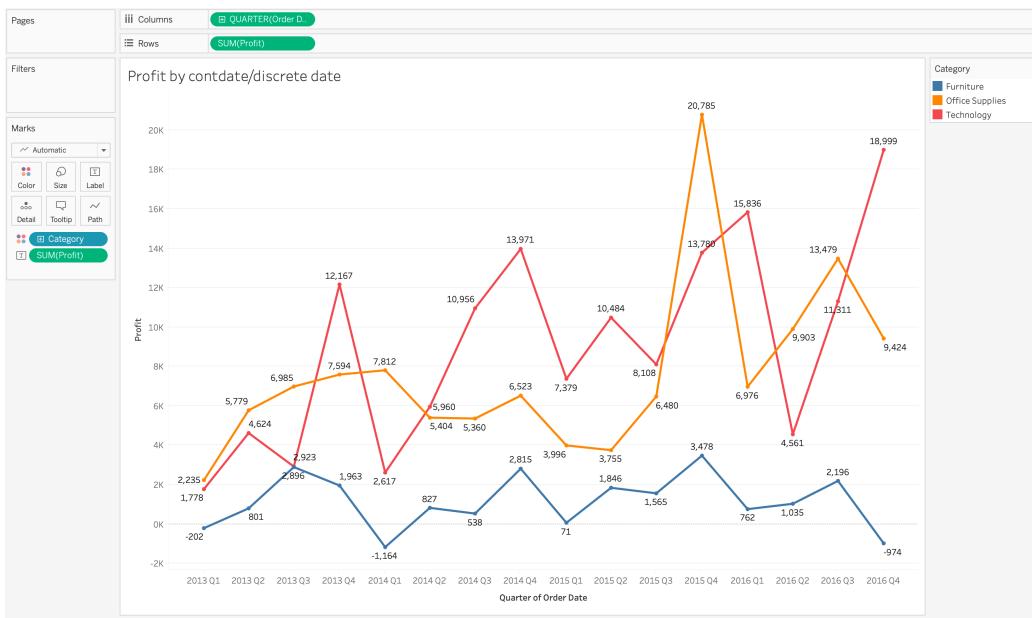
Filter can be added. And total sales can be applied to colour mark, the colour changes from dark red to dark blue according to number changes from small to big. Different colour can be used through edit colour.



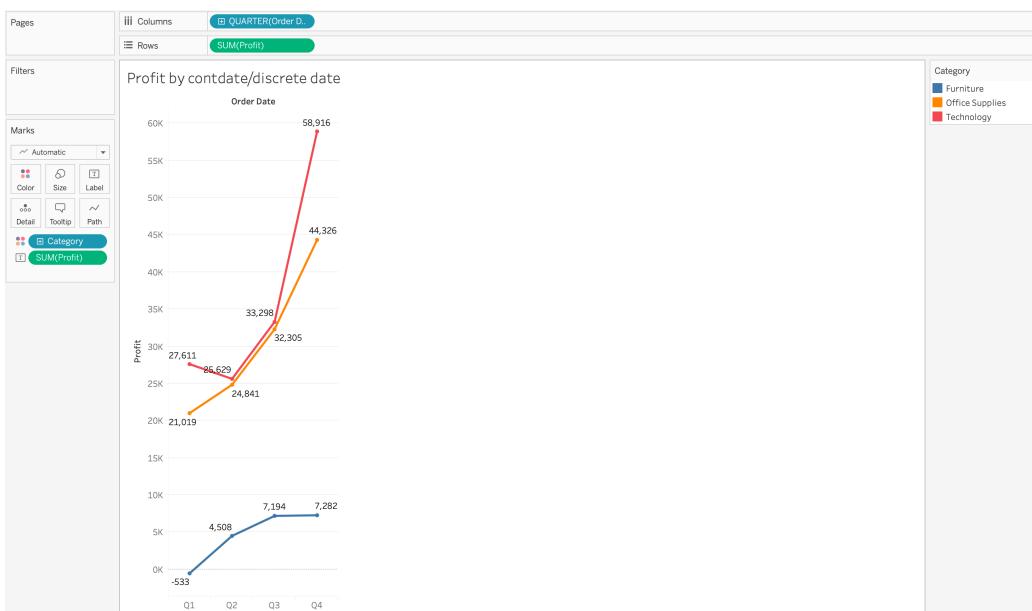
The reason why we did not unfold the hierarchy structure to product name level is that there are too many different kinds of product name and it will make the plot into too much detail though it can be applied.

## 7. Profit by Continuous Date/ Discrete Date

Order date can be dragged into columns and then different hierarchy level can be chosen by clicking the drop-down list. The quarter in columns is showing as green colour when continuous date applied.



Blue colour shows that the discrete date hierarchy level has been chosen. Only distinct value is shown for discrete column, therefore, only 4 distinct quarters are shown no matter how many years period from the source data. And this is difference between continues and discrete date in use.



## 8. Calculated Field

Drag order ID to rows and choose a specific order id through the filter and then add button from the pop-up window since there are too many different order ID.

The screenshot shows a Tableau interface. On the left, a 'Filter [Order ID]' dialog is open with the 'General' tab selected. It lists several order IDs with checkboxes, and 'CA-2011-100762' is checked. On the right, a data view is displayed with 'Order ID' in the rows shelf and 'Product Name' in the columns shelf. A filter for 'Order ID: CA-2011-100762' is applied. The 'Calculated field' section shows the formula: `Order ID  
CA-2011-100762`.

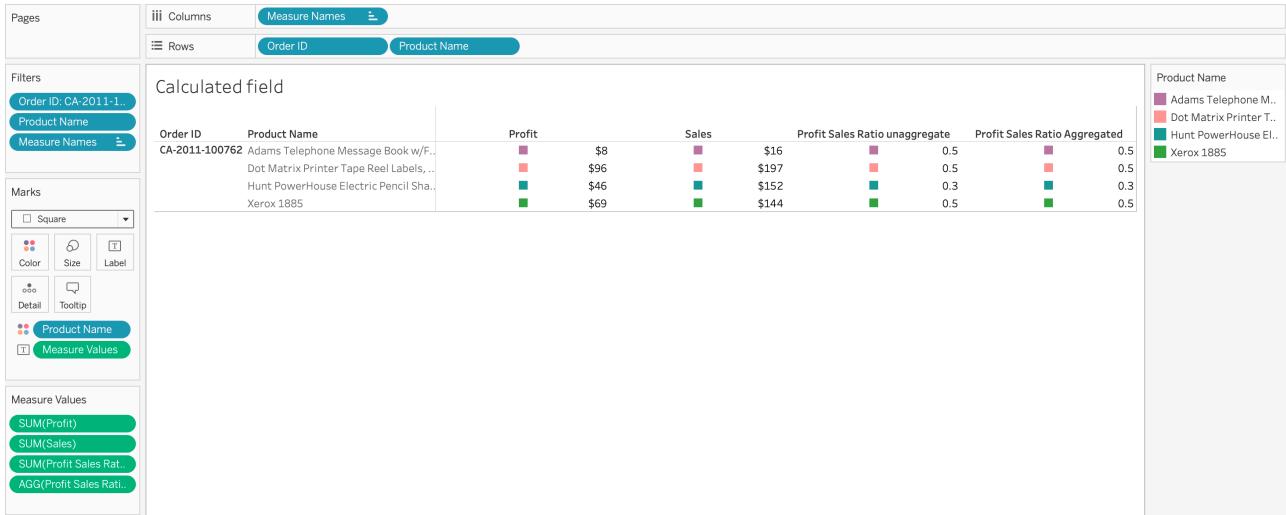
Similarly, drag the product name into rows and choose “use all” from filter and then add button from the pop-up window since there is limit number of product in one single order.

The screenshot shows a Tableau interface. On the left, a 'Filter [Product Name]' dialog is open with the 'General' tab selected. It has the 'Use all' option selected. On the right, a data view is displayed with 'Order ID' in the rows shelf and 'Product Name' in the columns shelf. A filter for 'Order ID: CA-2011-100762' and 'Product Name' is applied. The 'Calculated field' section shows the formula: `Order ID  
Product Name  
CA-2011-100762`.

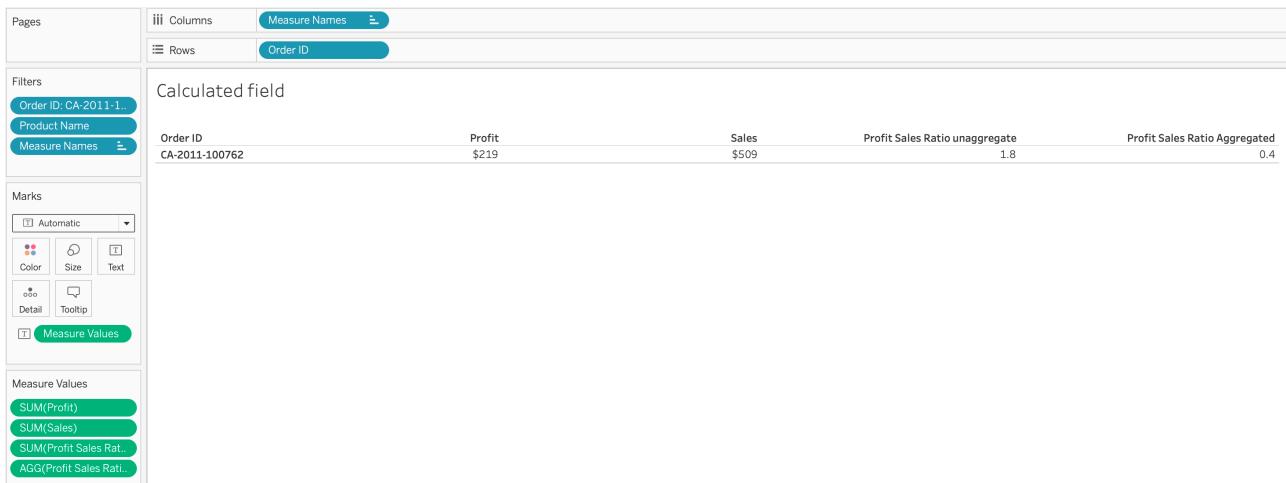
Then, choose Create Calculate Field from the drop-down list above Tables and create Profit Sales Ratio unaggregated formula and Profit Sales Ratio Aggregated formula.

The screenshot shows two 'Create Calculated Field' dialog boxes. The left box is titled 'Profit Sales Ratio unaggregate' and contains the formula `[Profit]/[Sales]`. The right box is titled 'Profit Sales Ratio Aggregated' and contains the formula `sum([Profit])/sum([Sales])`. Both dialogs show a validation message 'The calculation is valid.' at the bottom.

The difference between these two ratio can be observed by the formulae firstly, and then also have different meaning during the process: they have same ratio number when they under product name which is the lowest level currently.



When remove the product name field from rows, the ratio numbers will be different, the Profit Sales Ratio aggregated is the sum(profit)/sum(sales) = $219/509= 0.4$ , however, the Profit Sales Ratio unaggregated = sum of sub-ratio of the each product =  $0.5 + 0.5 + 0.3 + 0.5 = 1.8$ . Obviously, the unaggregated one is not very reasonable, therefore, be noted in use.



## 9. Sales per Customer

Total number of customers can be calculated by count distinct number of customer ID, and then Sales per Customer can be created through sum(sales)/Total number of customers.

Total Number of Customers

```
COUNTD([Customer ID])
```

The calculation is valid.

2 Dependencies ▾

Apply OK

Sales per Customer

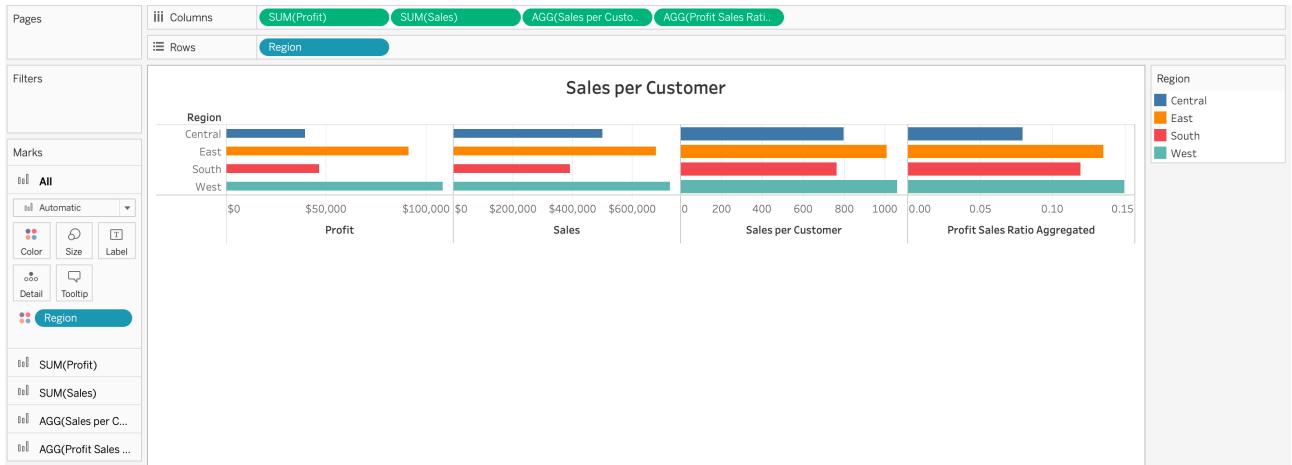
```
sum([Sales])/[Total Number of Customers]
```

The calculation is valid.

1 Dependency ▾

Apply OK

It will applied to all sub-plot when adjusting the the Marks in All, and it only applied to each single one plot when adjusting the each individual Mark. For example, it applied to Profit and Sales only when adjusting the size only for Profit and Sales.

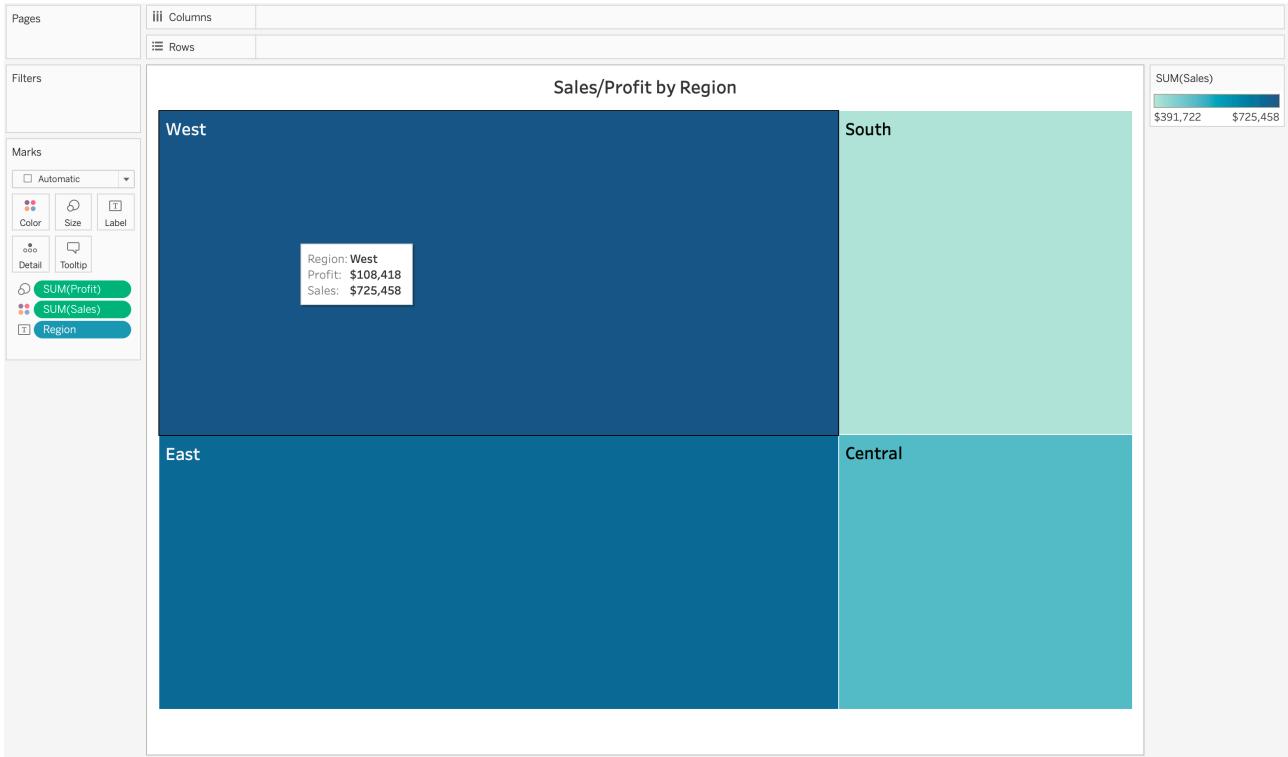


## 10. Sales/Profit by Region - Treemap

There are existing models that can be applied through show me sign on the top right corner. It shows that a Tree-maps view can be created by 1 or more dimensions and 1 or 2 measures when moving mouse to the Tree-maps sign.



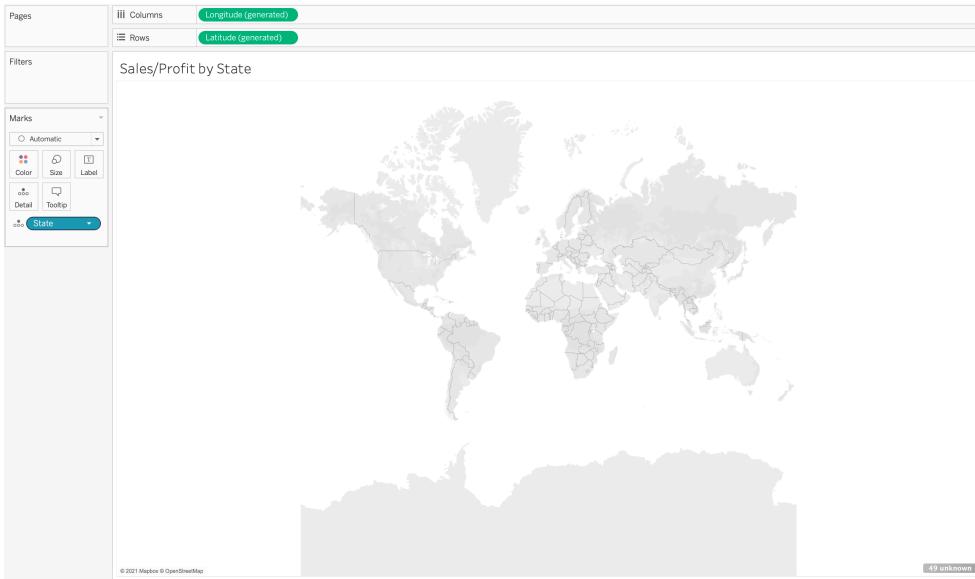
Selected 1 dimension Region and 2 measures Sales and Profit at the same time and click the tree-maps sign, a tree-maps plot can be shown as below snapshots.



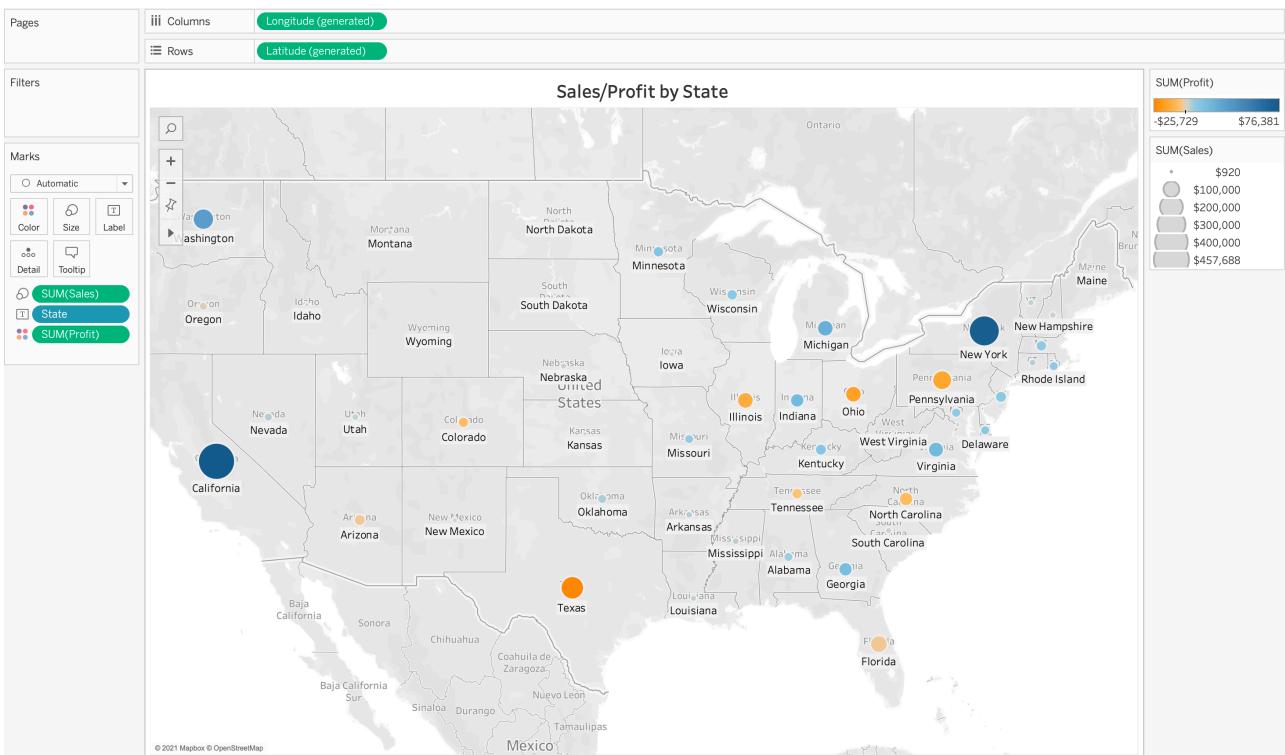
There are 2 measures in this plot. Profit is marked as size, which means that bigger size bigger number for these 4 regions. Sales in marked in colour, which means the darker the colour the bigger number based on the legend showing. Therefore, West has the biggest both sales and profit, then it's East.

## 11. Sales/Profit by State - Geographic map

State is a geographic dimension, a map as below can be shown when double clicking the state dimension.

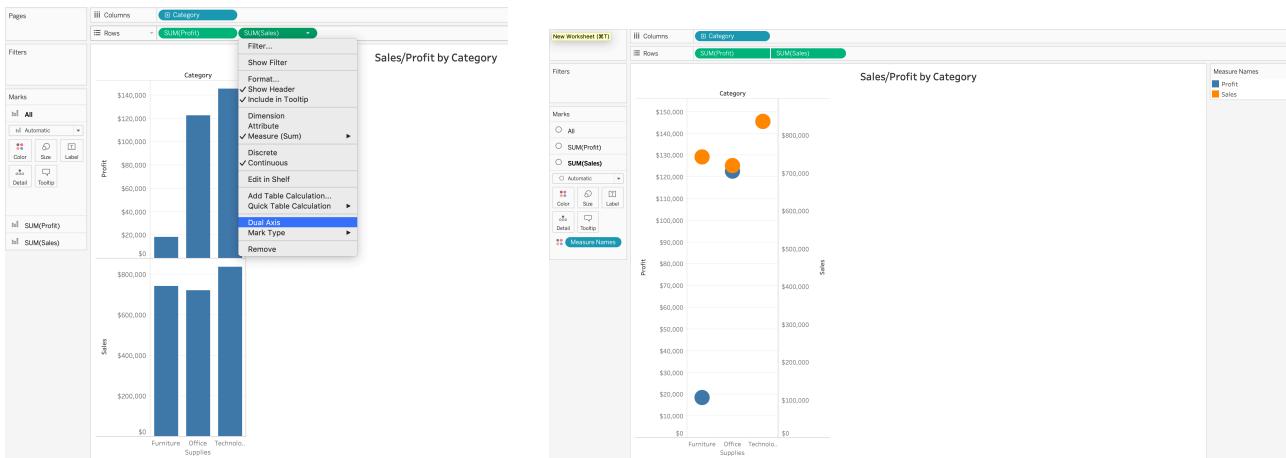


Edit the location as United States by clicking the 49 unknown at the bottom right corner, the regions can be recognised as below. Similarly, the size and colour will stand for the the number of Sales and Profit by legend.

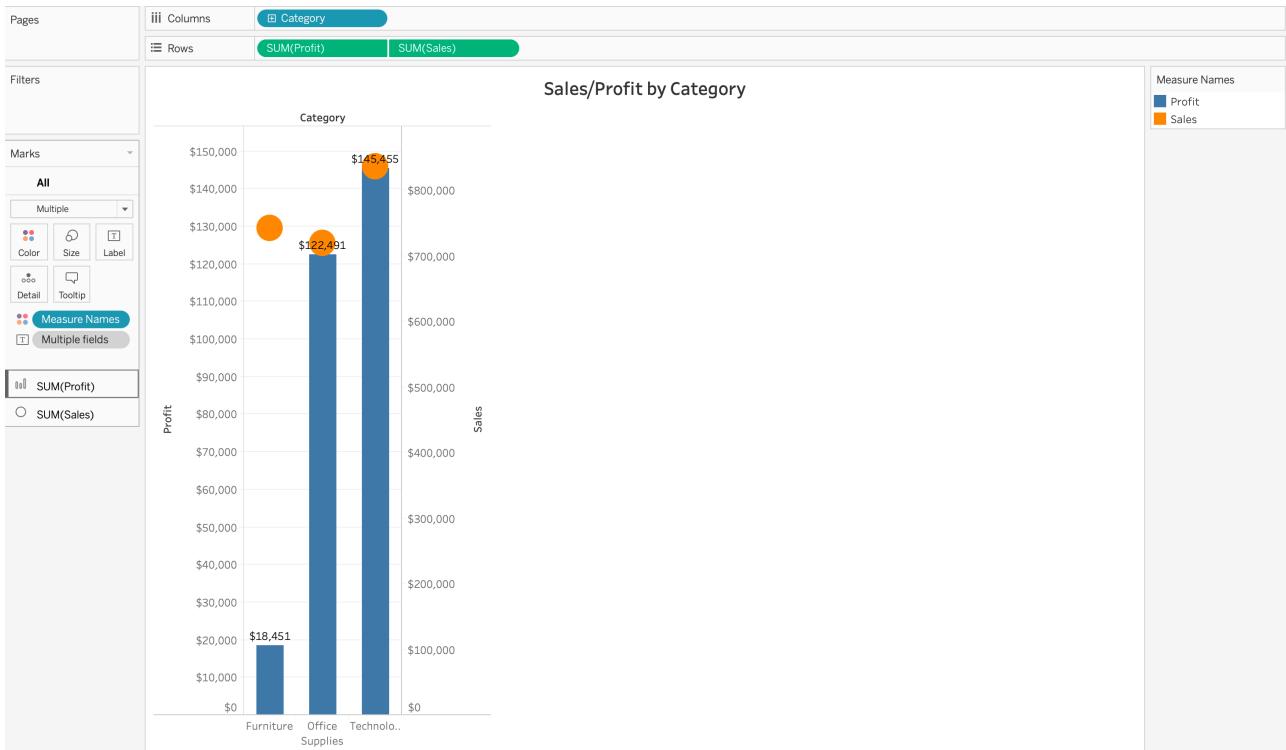


## 12. Sales/Profit by Category - Dual-Axis

By dragging category, sales and profit, a double-bar plot will be shown, choose Dual-Axis by right clicking the drop-down list of one measure, e.g. sales, the plot will be changed into the second one with two colour circle plot.



Change one of them from circle to bar chart, the Dual-Axis plot is showing as below: now the profit is showing by bar chart and sales is showing as circle in orange.



## 13. Sales/Profit Details

A bar chart can be created through dragging the dimension and measures into rows and columns as done before.

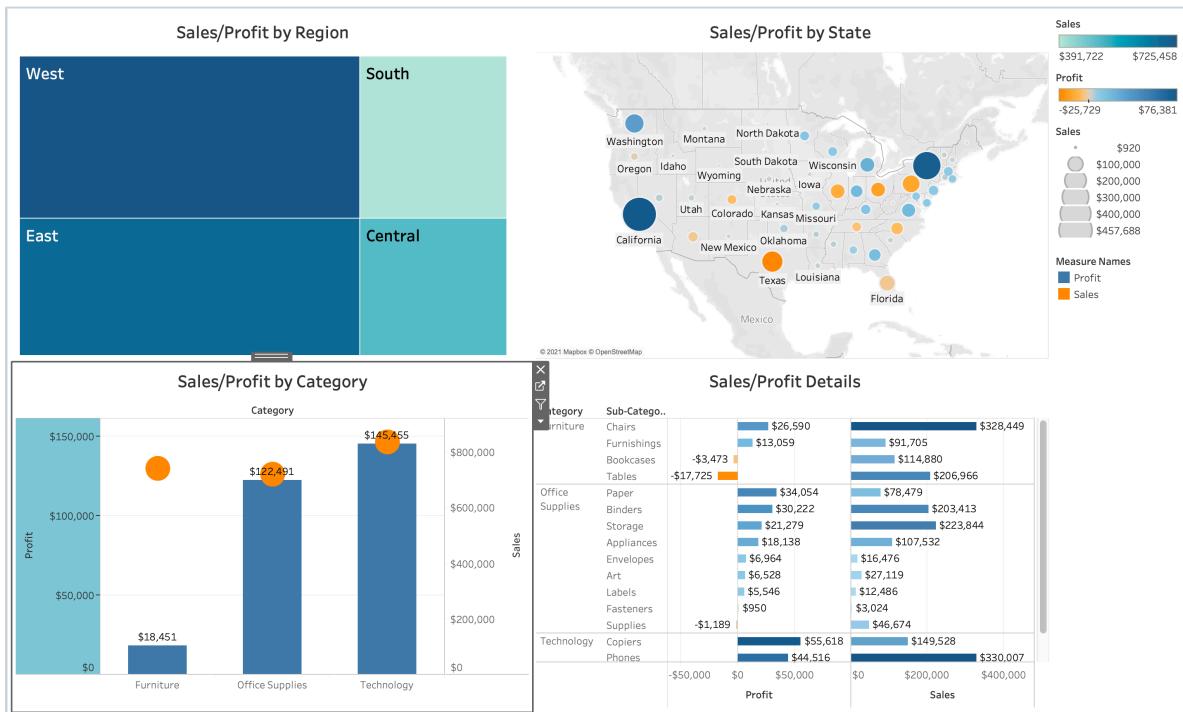
Do noted that the plot can only be sorted by only one measure(Profit) as shown below, but it can not be applied to sales as well at the same time.

There are 2 legends for Profit and Sales separately.

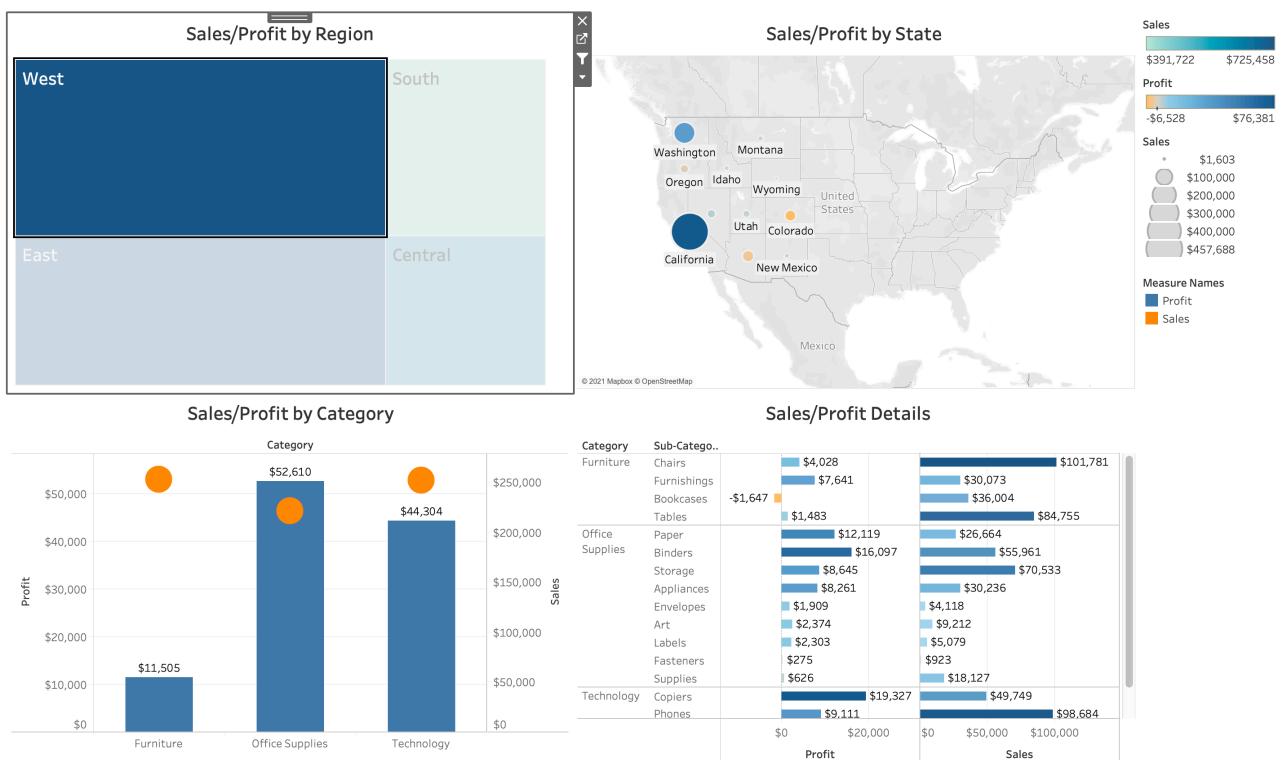


## 14. Sales/Profit Dashboard

The four created worksheets can be used to establish a Sales/Profit Dashboard as below.

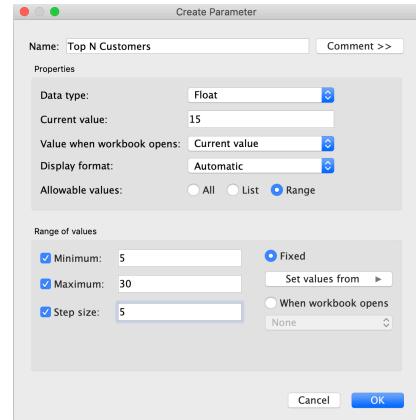
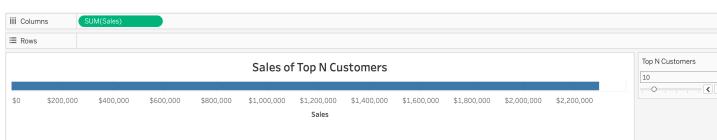


One sub-plot can be chosen as a filter by clicking the plot and choose the filter sign, for example, choose the first sub-plot as filter. Only one related west region sales/profit details will be shown after clicking the West in the first sub-plot.

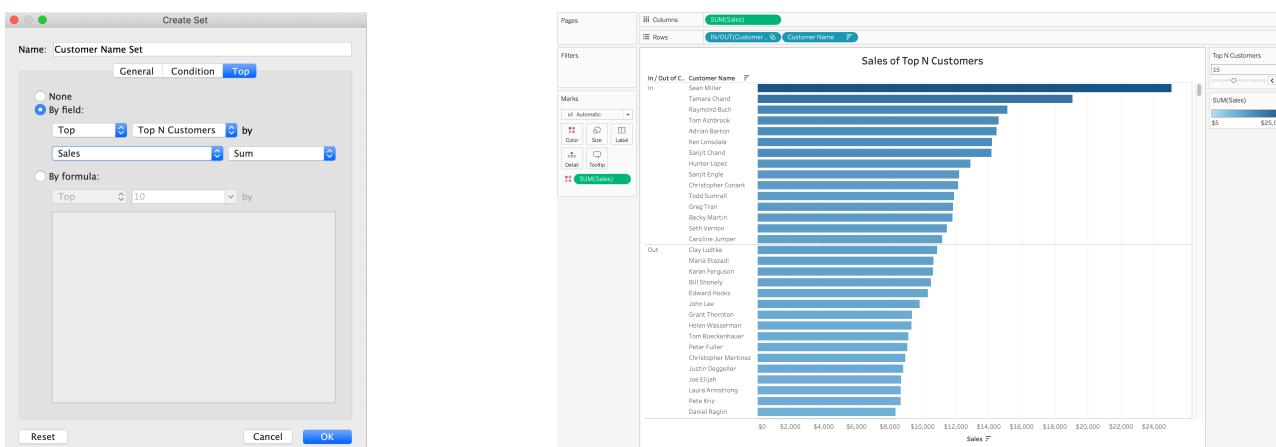


## 15.Sales of Top N customers - using parameter and set

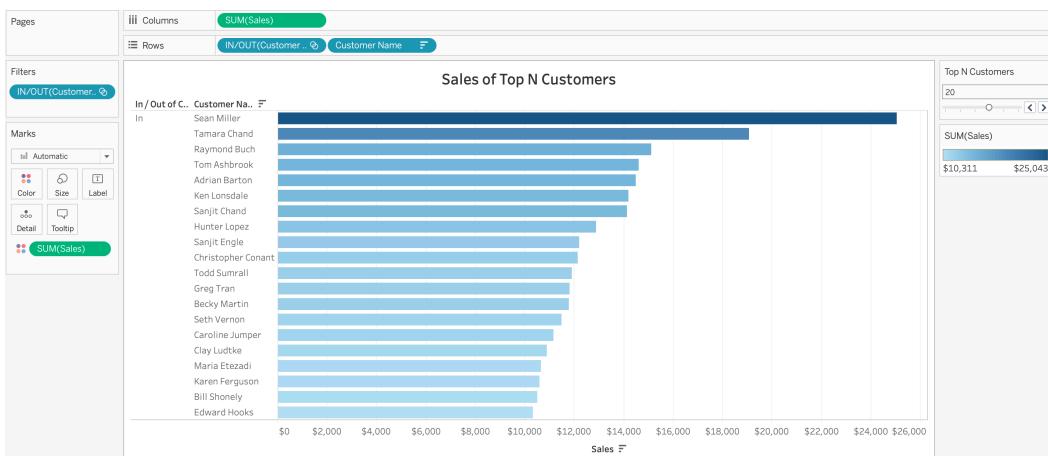
First of all, a parameter of Top N Customers can be created as below, top 15 customers is needed. Range can be chosen for the available values, like from 5 to 30 and each step is 5, then we can choose to check top 5, 10, 15 ... 30 customers by this parameter. Clicking show parameter, Top N Customers is showing at the top right corner. It can be changed by adding or subtracting by step.



A customer name set should be created for the Top N customers which is ranked by sum of sales. The In/Out(Customer name set) means within or outside of the customer set.

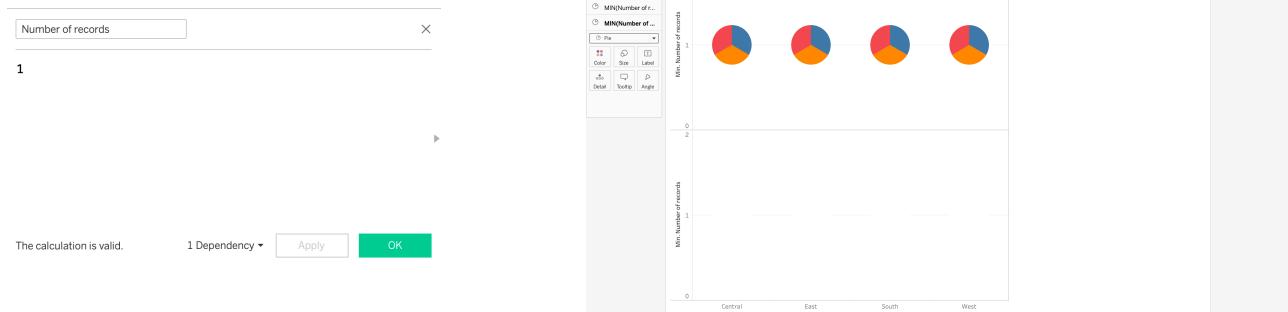


The top N customers outside of the customer set is not cared about, so it can be removed.

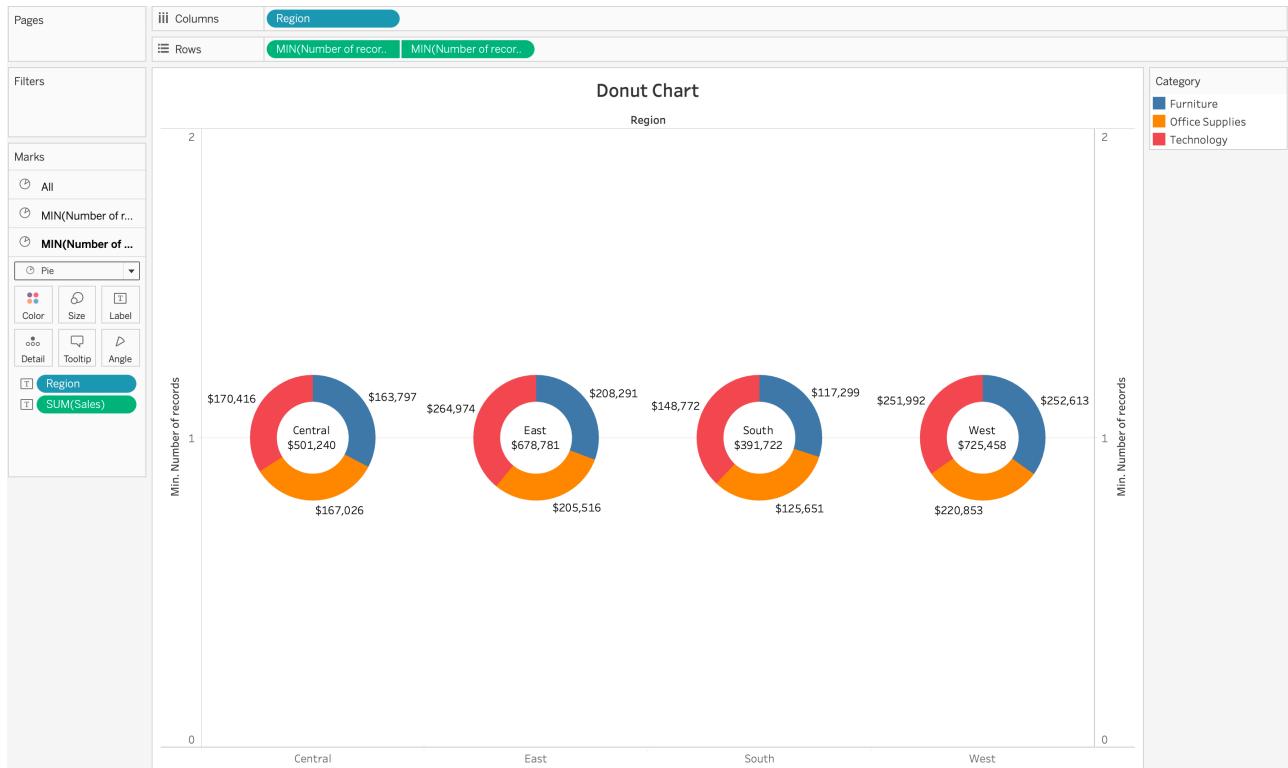


## 16. Donut Chart - using two pie charts

Create a new calculated field which is set as 1 for further use. Then chose region and sales and click pie chart at the same time, then drag region into columns and Number of Records into rows twice with measure type as minimum. The minimum number is still 1 since the number of records is constant 1. And drag category into colour mark of first one and mark the second one as white colour.

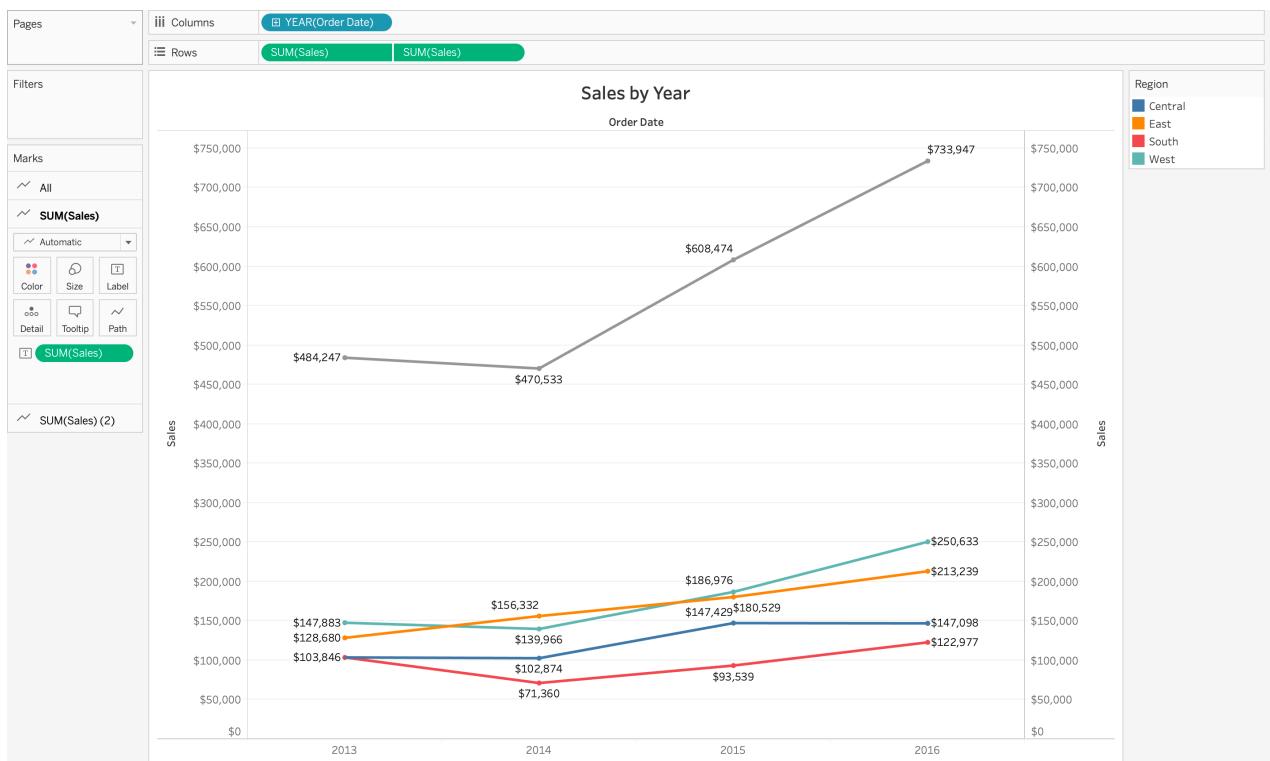


Adjust the size of the two rows of pie charts and using the Dual-Axis method to make a donut chart. Adjust the detailed info, put region and sum of sales with alignment in in middle of vertical in the second pie chart, and the donut chart is showing as below:

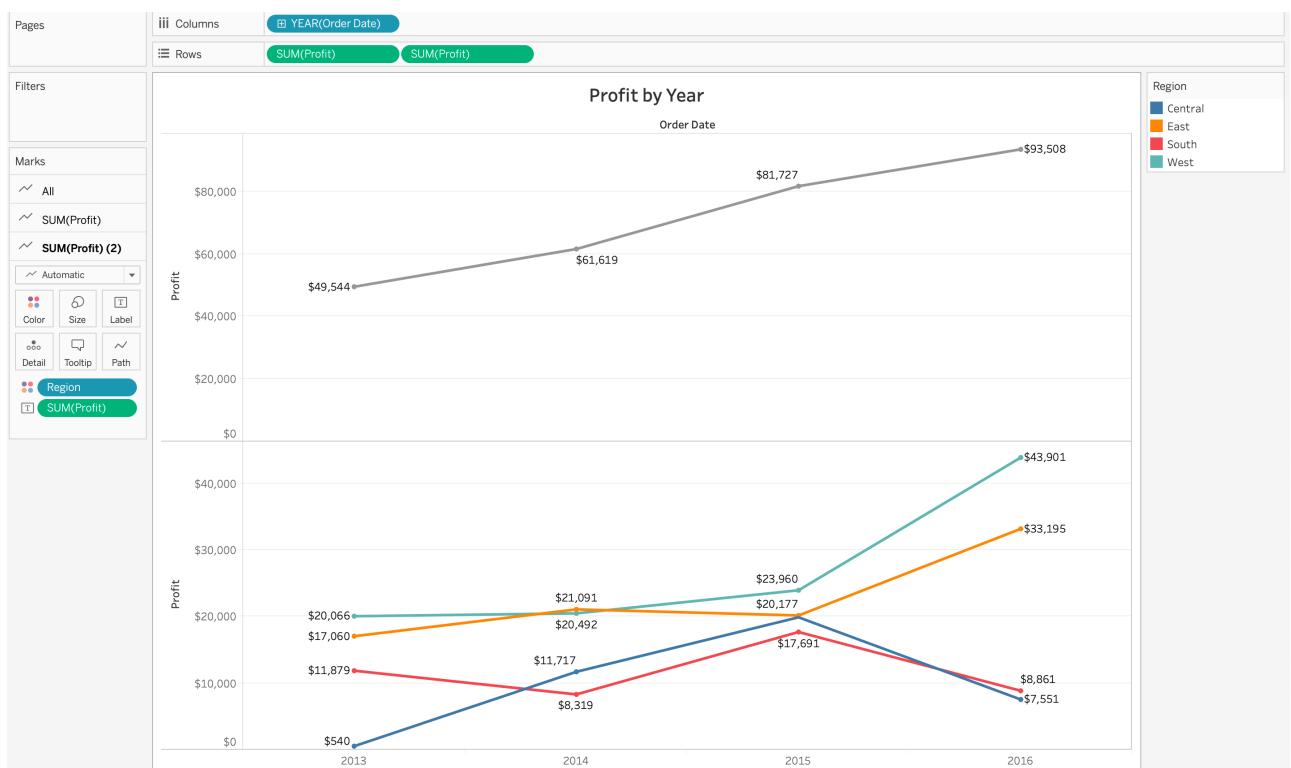


## 17. Sales/Profit by Year - Dual-Axis line chart

Create two line charts for the year sales and then use dual-axis. The first one is sum of the sales and the second is sales by region. It's important to synchronised axes to make sure they show the correct sales number.

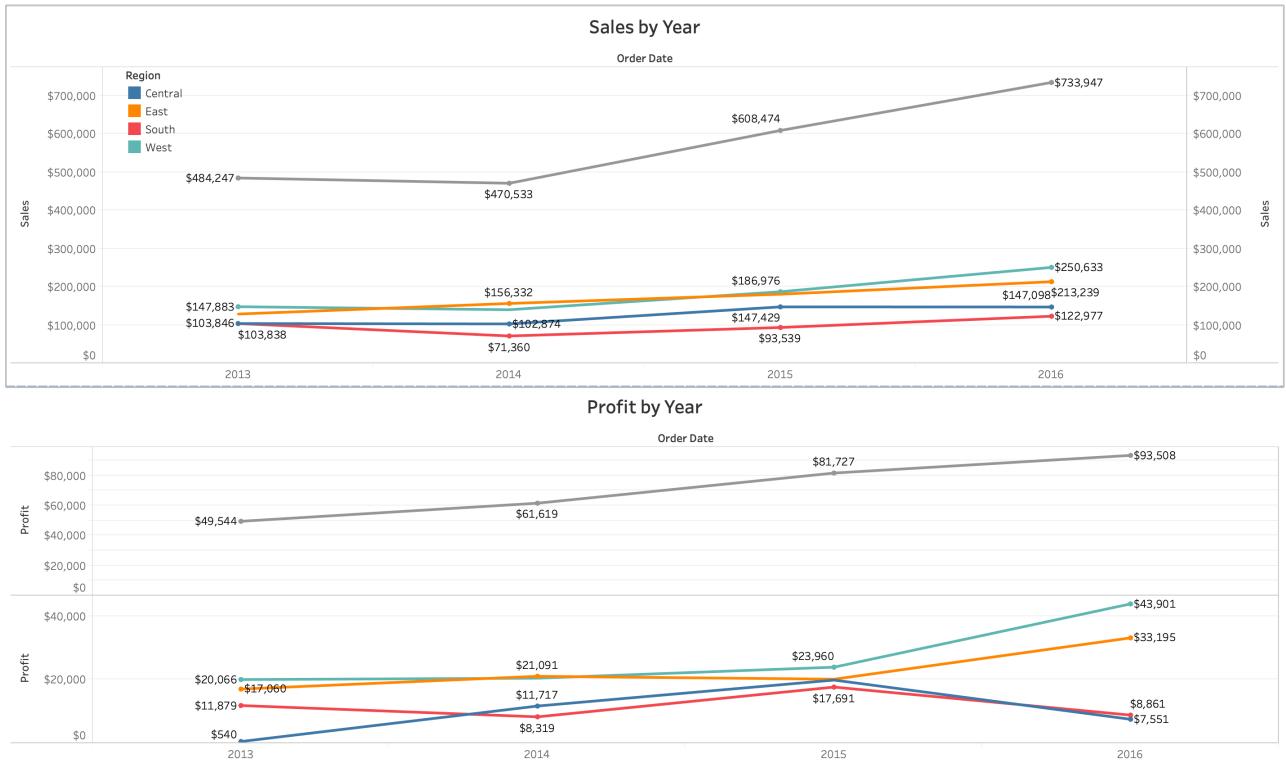


Duplicate the above worksheet and drag profit on sales to quick create profit by year.



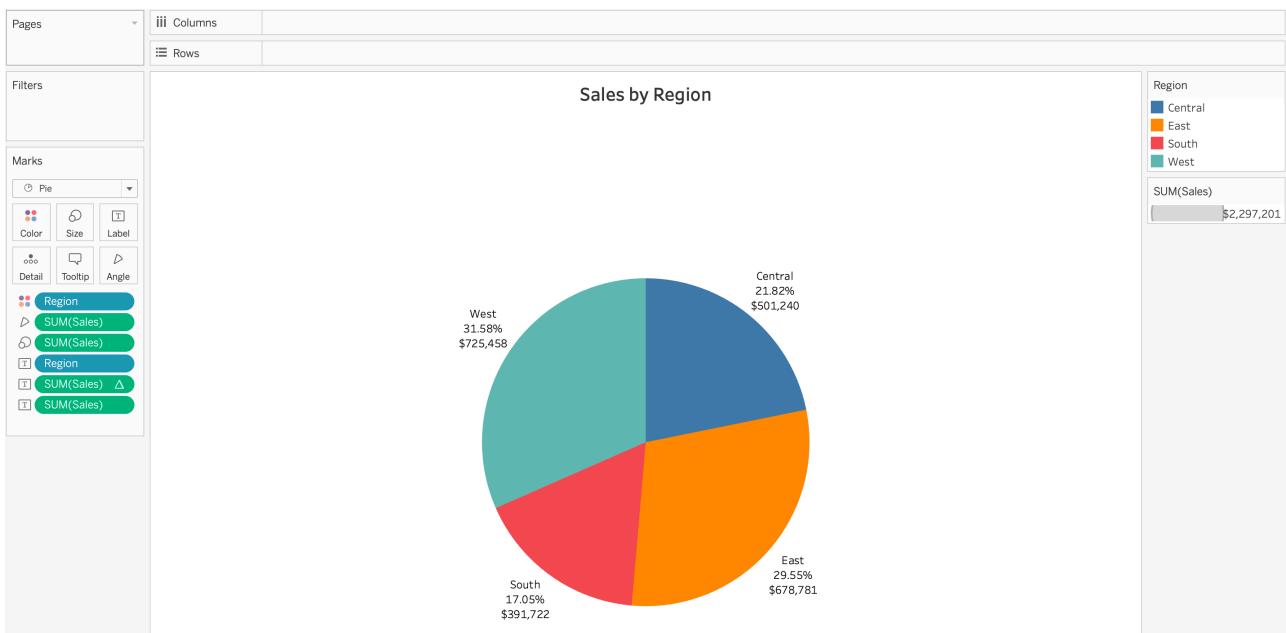
## 18. Sales/Profit by Year Dashboard

The dashboard for Sales/Profit by Year is as below, the legend can be moved and floating on the plot by choosing floating from the drop-down list.



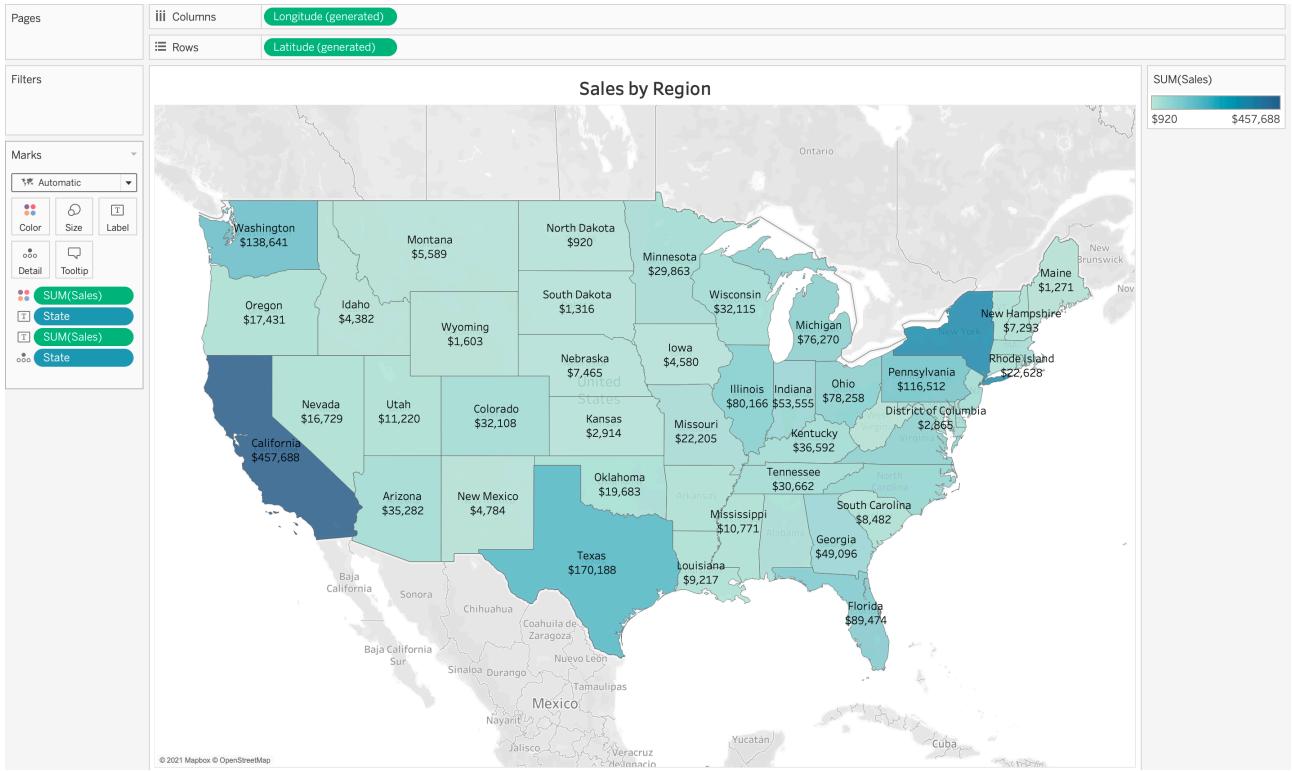
## 19. Sales by Region - Pie Chart

Choose quick table calculation to show the percentage of total sales from label mark.

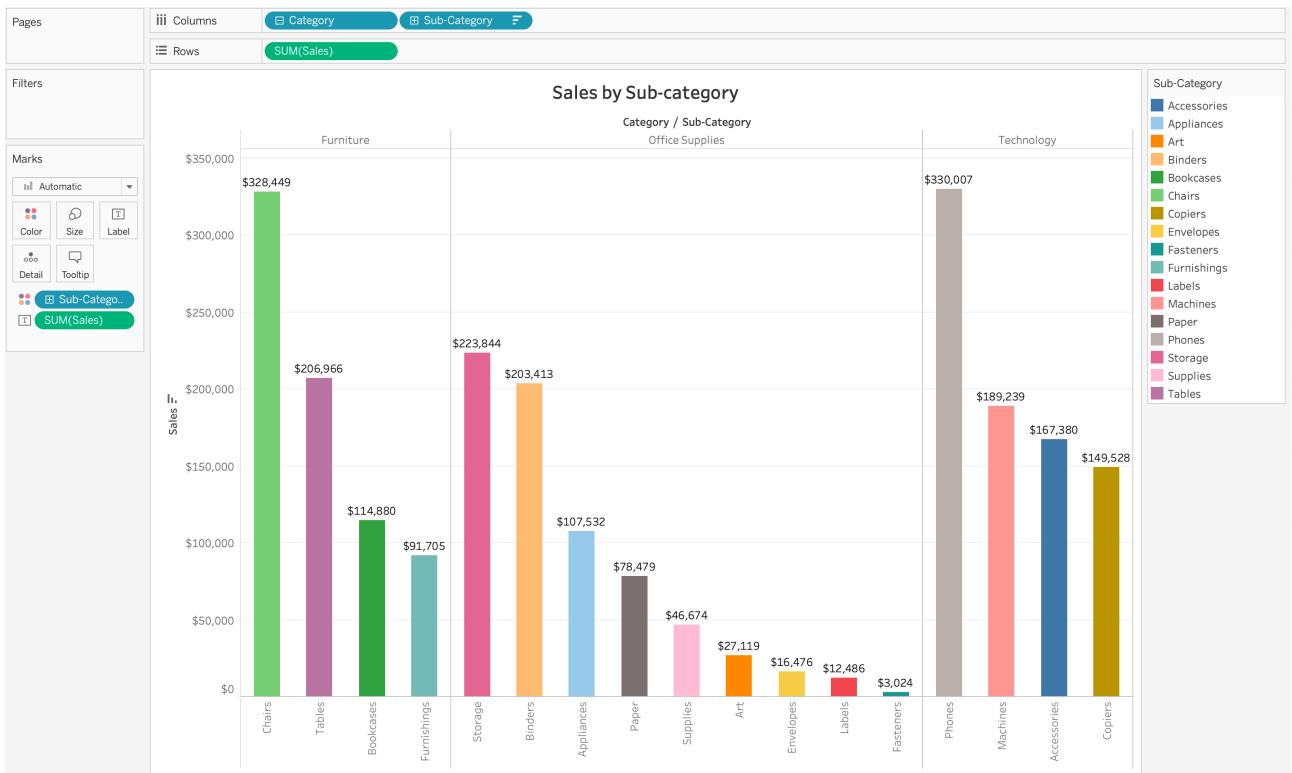


## 20. Sales by State - Geographic map II

The colour stands for the sum of sales once the sum of sales is dragged into colour mark.

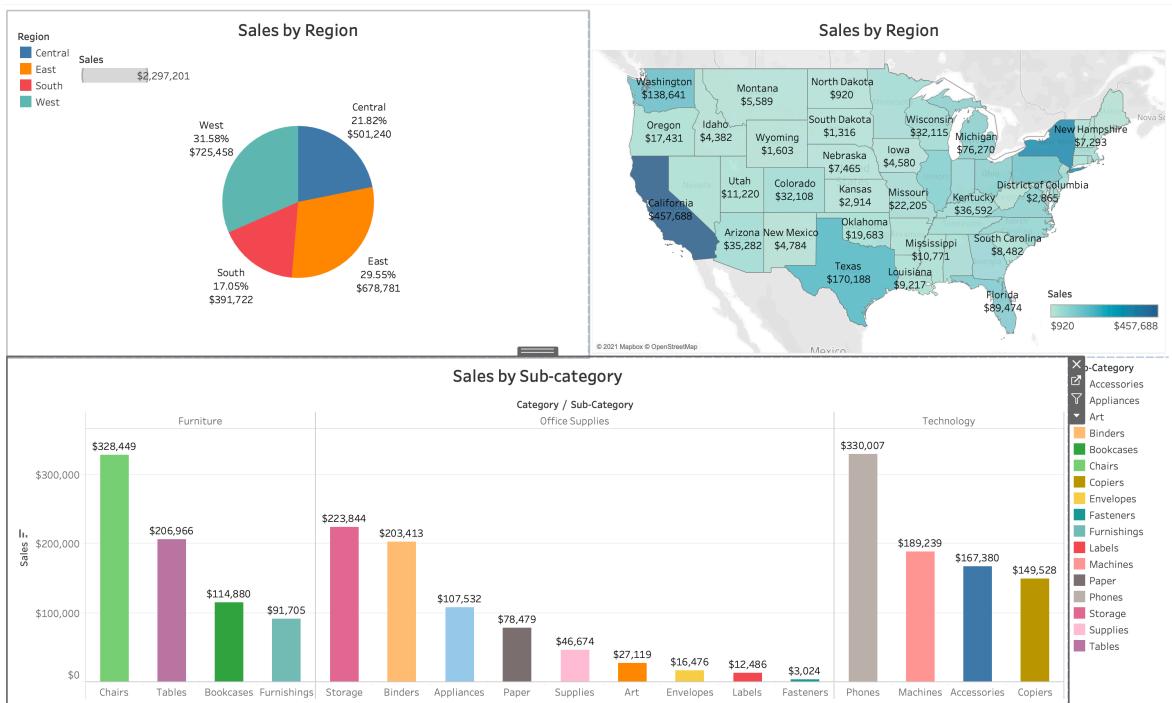


## 21. Sales by Sub-category



## 22. Sales Dashboard

By applying floating to move the legend to the proper location for each sub-plot.



## 23. Profit Dashboard

Filter can be applied to all worksheet from drop-down list of the dashboard.

