



Data Visualisation for Business

ANL 201

Data Visualisation for Discrete and Continuous data
Study Unit 4

January 2024

Recap

- ▶ Four Stages of Data Visualisation:
 - ▶ data collection and storage, data pre-processing, graphics engine and human visual and cognitive processing
- ▶ Types of datasets, namely, record data, graph-based data, and ordered data
- ▶ Data collection and storage is the methodological process of gathering information about a specific subject and appropriately storing it in a secure
- ▶ Data preparation is the process of identifying data quality issues and treating these issues so that data can be used for business intelligence (BI), analytics and data visualisation applications
- ▶ As part of data preparation process, previous unit also discussed different data types, dataset architecture and exploratory analysis techniques



Visualisation of Categorical Data

Visualisation of Categorical Data

Best practices for visualising categorical data

- ▶ **Frequency tables/ bar charts-** Each rectangle represents a category. The longer the rectangle is, the greater the value it represents.
- ▶ **Pie-charts-** When we put categorical data together, the sum of the parts can be equal to the whole. A full circle represents 100%, and each wedge is a portion of that 100%.
- ▶ **Tree map-** used to visualise the sub-categorical data within the categorical data. Higher variability and more interesting insights to visualise.

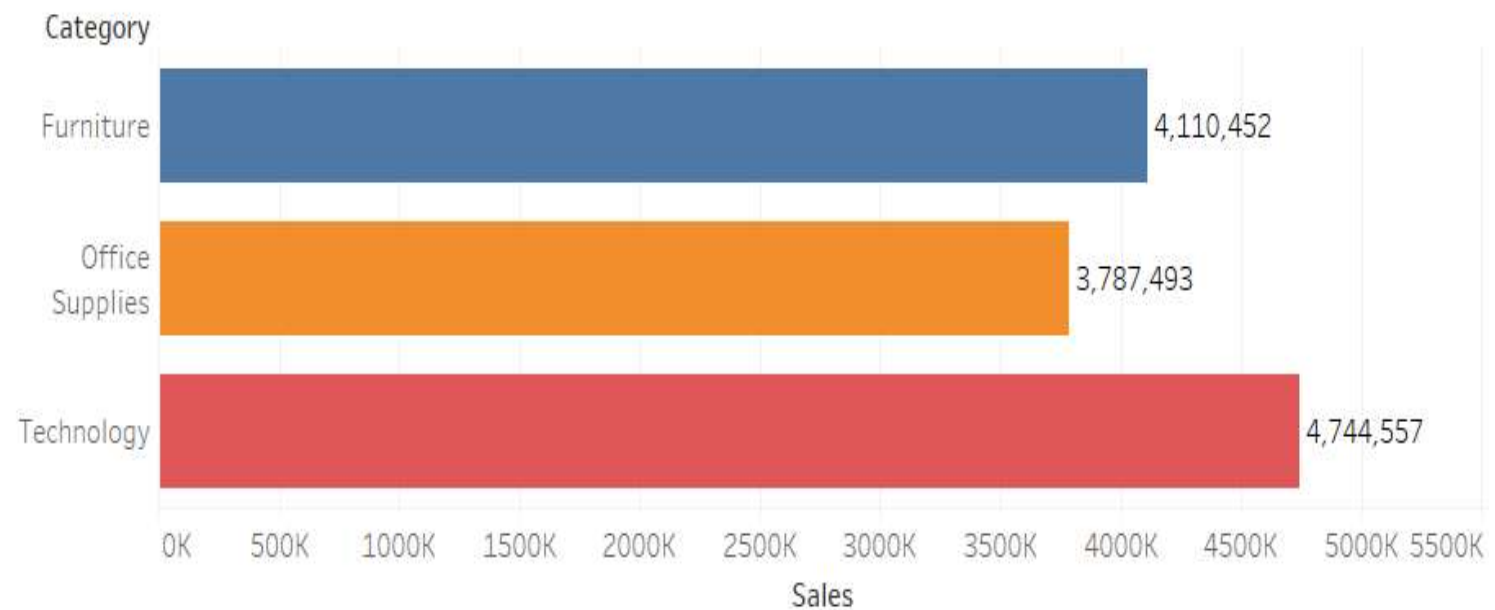
Visualisation of Categorical Data

Bar Chart- sales by category

Visualisation task:

Let us say the sales manager of a company, like global superstore, would like to visualise how much sales happened for each product category sold. He/she can use a horizontal bar chart or a vertical bar chart in Tableau.

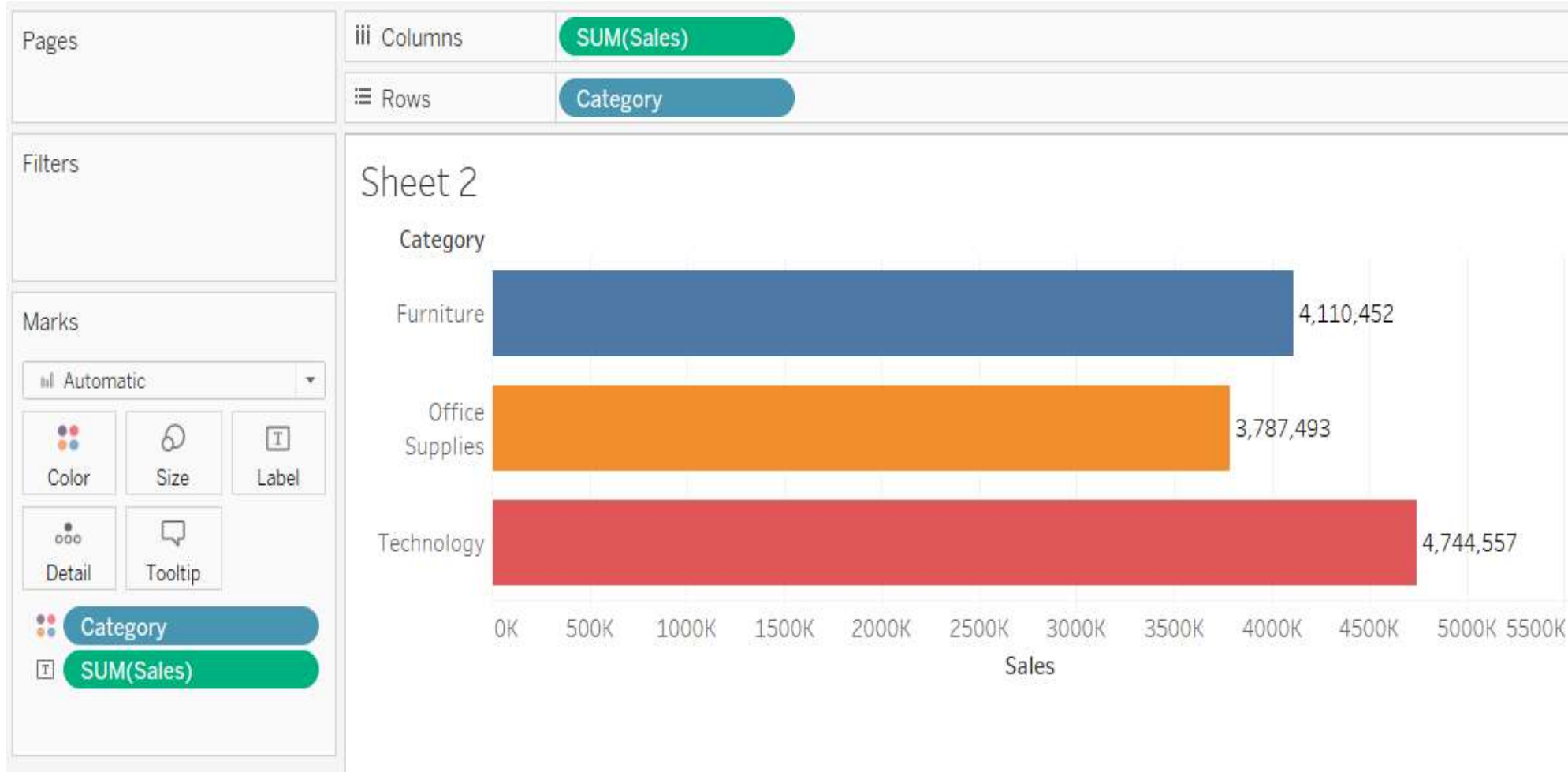
Chart type



These charts facilitate one-to-many comparisons. Bar chart is the most effective way to compare values across dimensions due to their linear nature that makes precise comparisons easy

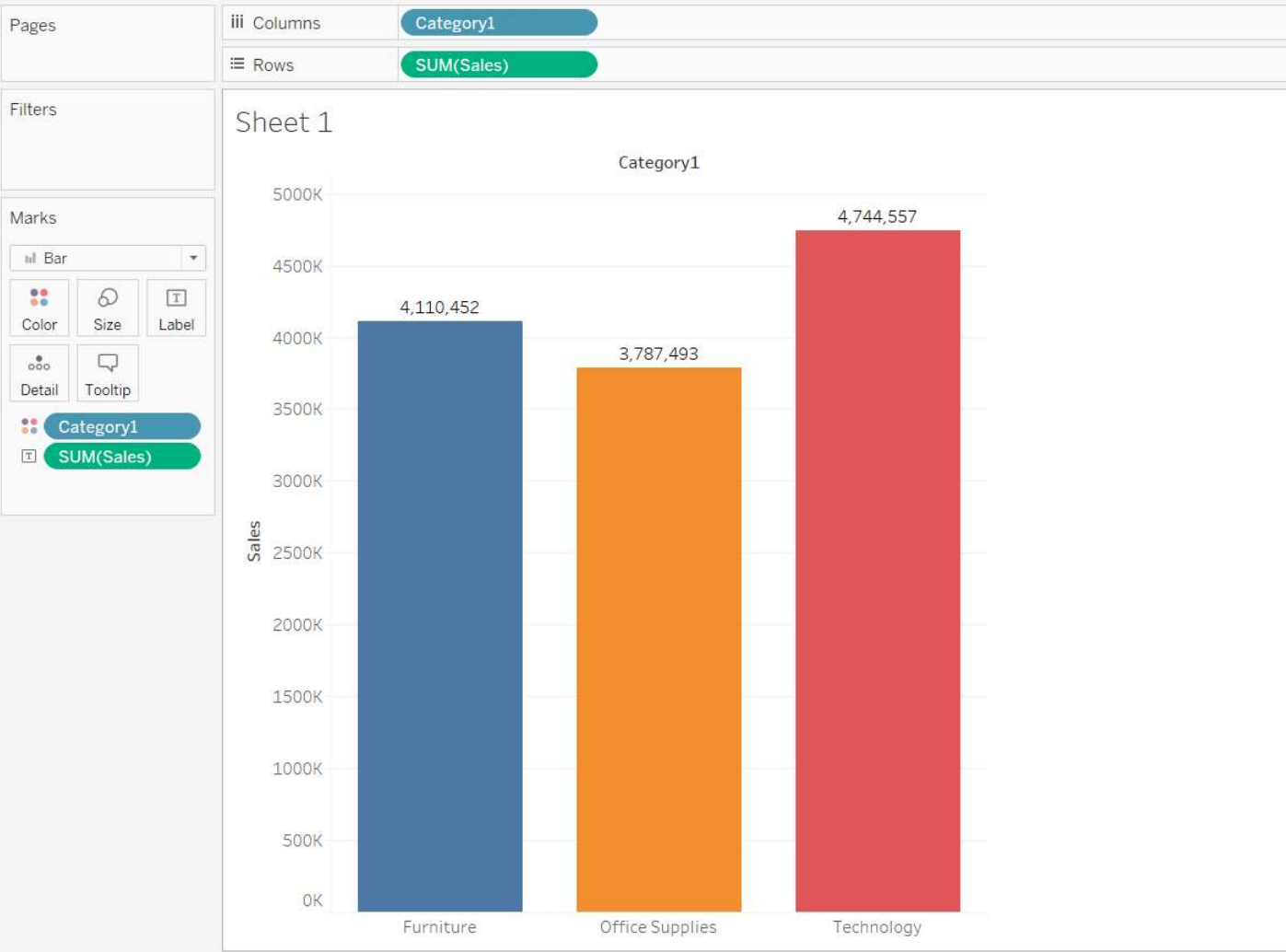
Visualisation of Categorical Data

Horizontal bar Chart- sales by category



Visualisation of Categorical Data

Vertical bar Chart- sales by category



Visualisation of Categorical Data

Bar Chart- sales by category

Below are the steps to creating a bar chart in Tableau:

1. Create a new worksheet.
2. For horizontal bar charts, drag “Category” dimension into the worksheet’s rows (for vertical bar charts, drag “Category” dimension into the worksheet’s columns).
3. For horizontal bar charts, drag one or more measures (“Sales” in this case) into the worksheet’s columns (for vertical bar charts, drag one or more measures (“Sales” in this case) into the worksheet’s rows).
4. Drag and drop “Category” dimension on the “Color” marks card shown on the left.
5. Drag and drop “Sales” measure on the “Label” marks card.

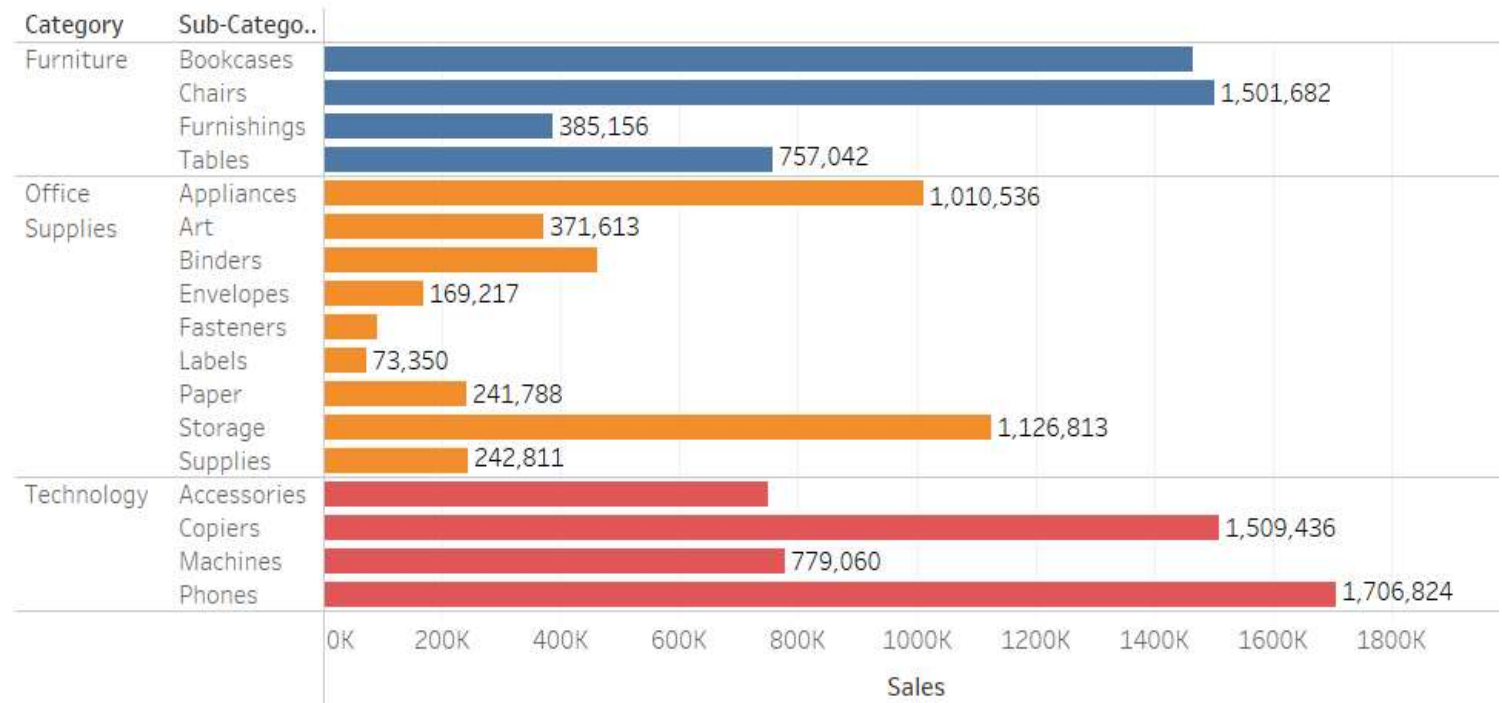
Visualisation of Categorical Data

Side-by-side Bar Chart- sales by sub-category

Visualisation task:

The same sales manager would now like to visualise the sales for each sub-category within the product category. He/she can use the horizontal side-by-side bar chart or the vertical side-by-side bar chart in Tableau.

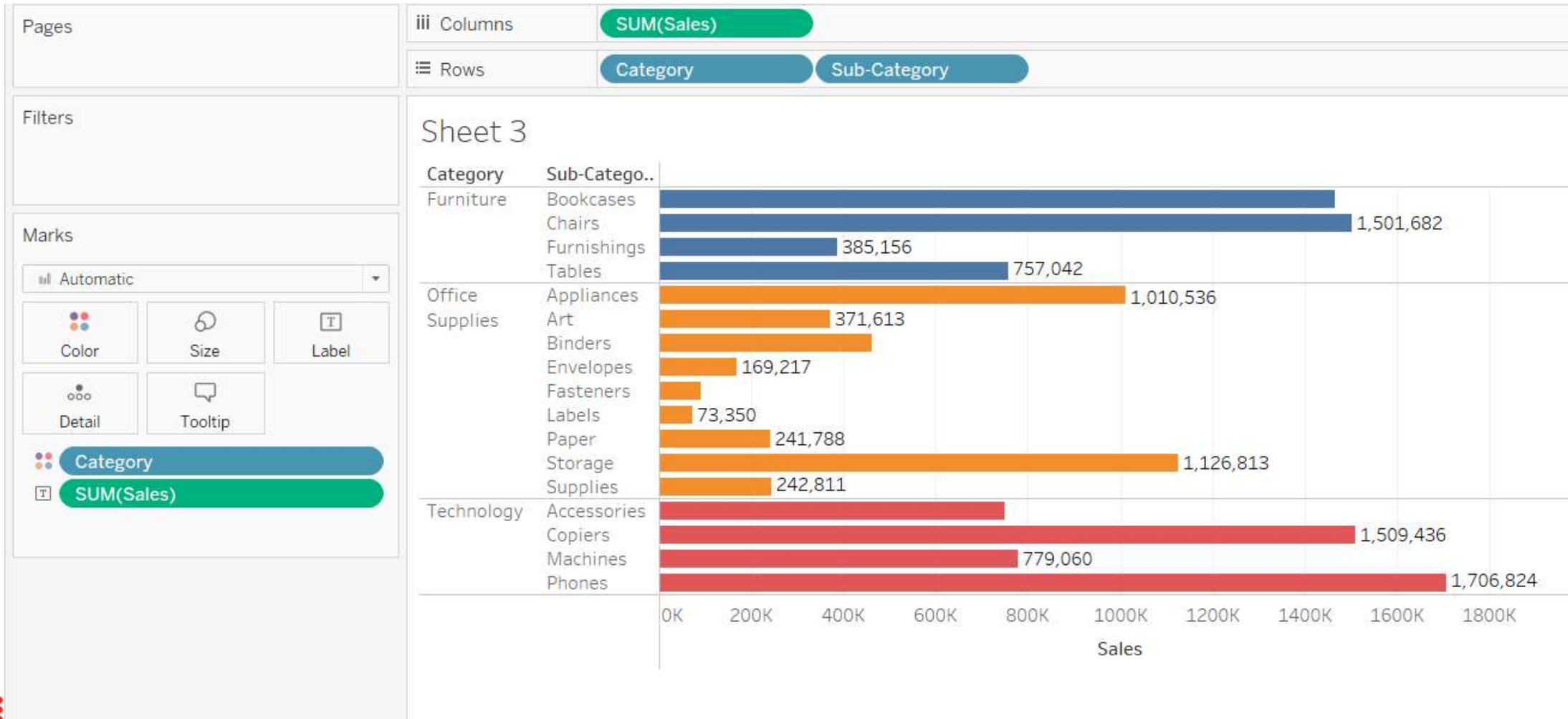
Chart type



The side-by-side bar chart provides a way to compare measure across dimensions on a single axis.

Visualisation of Categorical Data

Side-by-side Horizontal Bar Chart- sales by sub-category



Visualisation of Categorical Data

Side-by-side Vertical Bar Chart- sales by sub-category



Visualisation of Categorical Data

Side-by-side Bar Chart- sales by sub-category

Below are the steps to creating a side-by-side bar chart in Tableau:

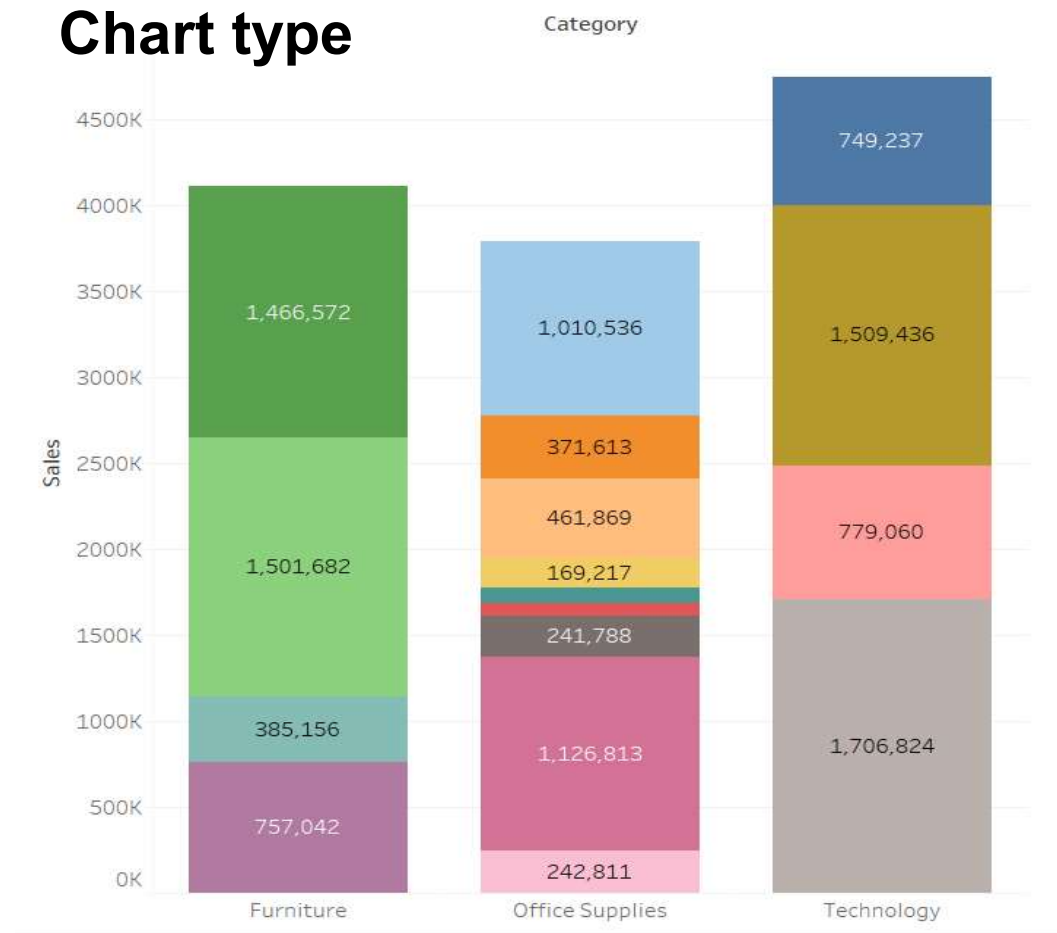
1. Create a new worksheet.
2. Drag the “Sales” measure into either the worksheet’s columns (for side-by-side horizontal bar chart) or rows (for side-by-side vertical bar charts).
3. For horizontal side-by-side bar charts, drag “Category” and “Sub-Category” dimensions into the worksheet’s rows (for vertical side-by-side bar charts, drag the “Category” and “Sub-Category” dimensions into the worksheet’s columns).
4. Drag and drop the “Category” dimension on the “Color” marks card on the left.
5. Drag and drop “Sales” measure on the “Label” marks card.

Visualisation of Categorical Data

Stacked Bar Chart- sales by sub-category

Visualisation task:

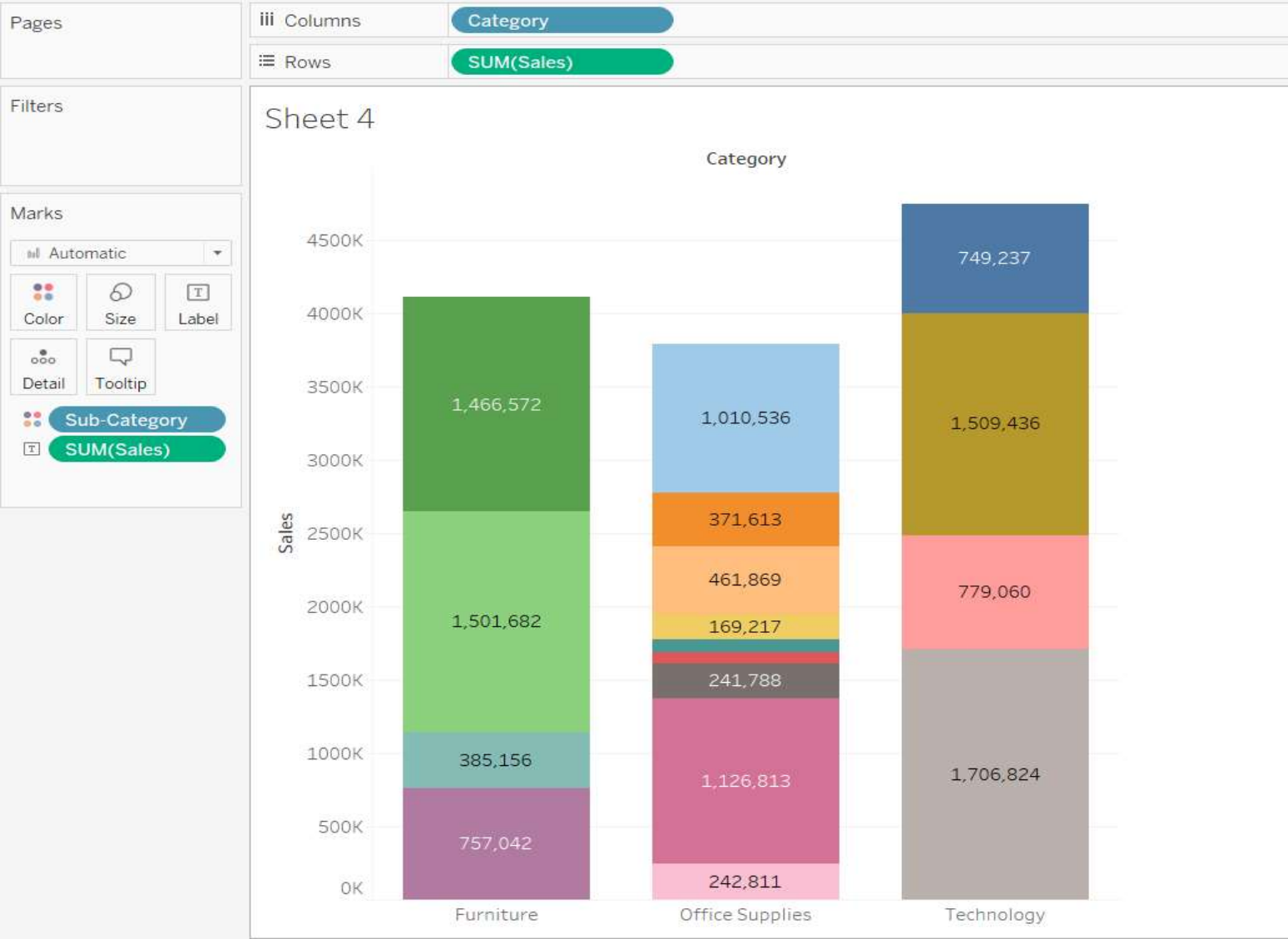
The same sales manager would like to visualise how much each sub-category contributed to the total sales of the product category. He/she can use horizontal stacked bar charts or vertical stacked bar charts in Tableau.



The stacked bar chart is an efficient way to show the share of each sub-category in the category total.

Visualisation of Categorical Data

Stacked Bar Chart- sales by sub-category



Visualisation of Categorical Data

Stacked Bar Chart- sales by sub-category

Below are the steps to creating a vertical stacked bar chart in Tableau:

1. Create a new worksheet.
2. For a vertical stacked bar chart, drag the "Category" dimension into the worksheet's columns.
3. For a vertical stacked bar chart, drag the "Sales" measure into worksheet's rows.
4. Drag and drop "Sub-Category" dimension on the "Color" marks card on left.
5. Drag and drop "Sales" measure on the "Label" marks card.

In cases where there are too many sub-categories, too many colours will make the bar look very cluttered and difficult to interpret. As a rough rule of thumb, a stacked bar-chart is ideal for sub-categories below 7 or 8.

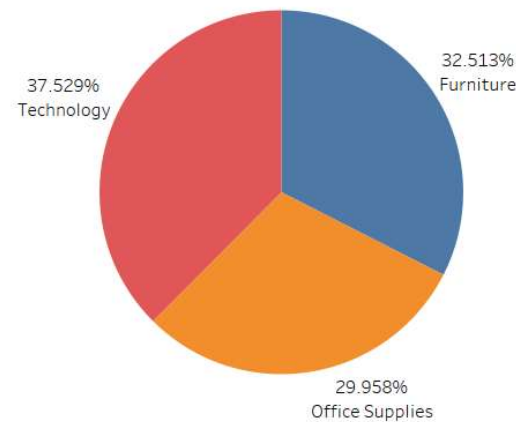
Visualisation of Categorical Data

Pie Chart – share of each category in total sales

Visualisation task:

When the sales manager would like to visualise the contribution of each product category sales to total sales of the company. He/she can use a pie chart in Tableau

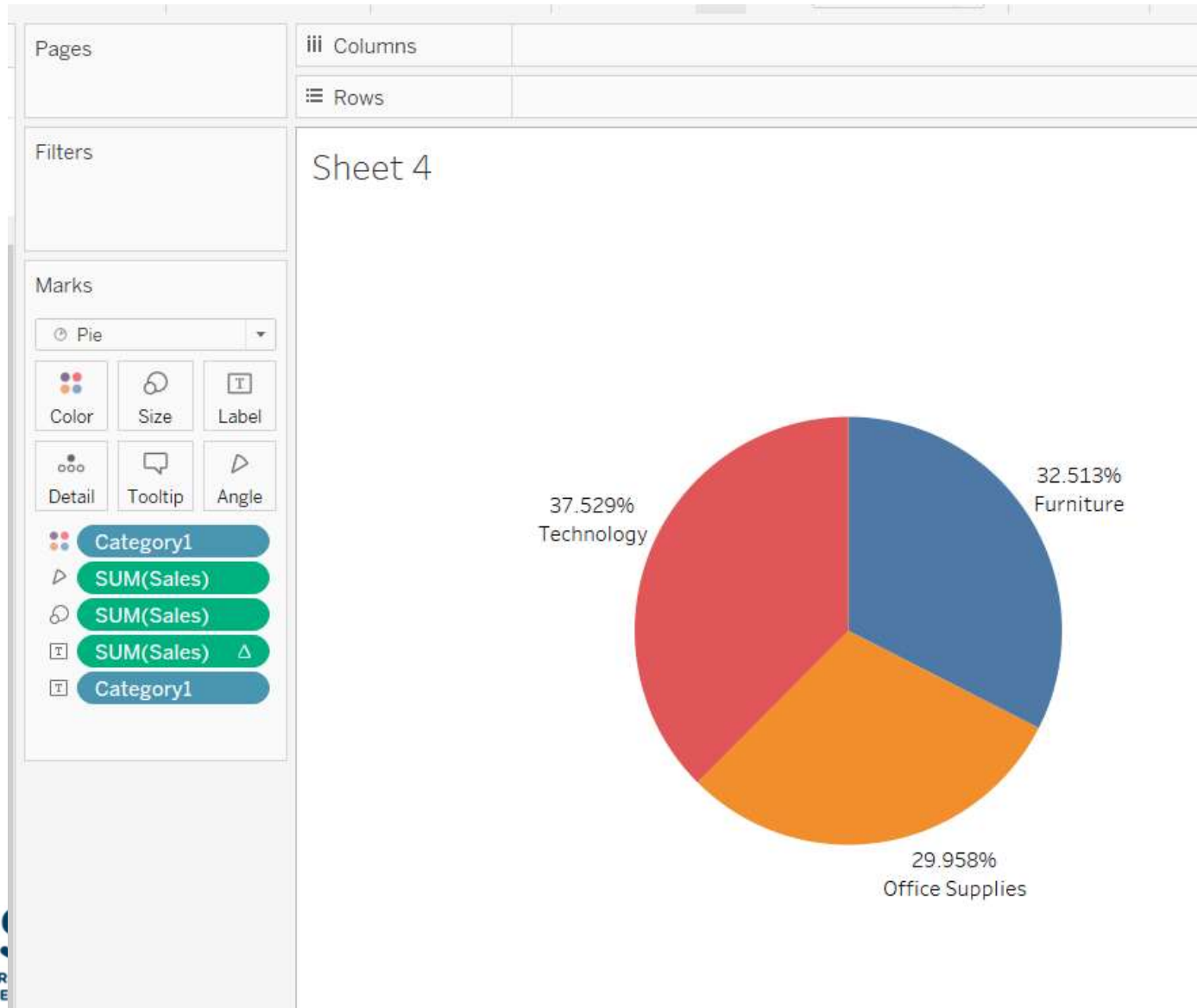
Chart type



Pie Charts should be used when you want to visualise the contribution/share of each part to the total. It should be used to get a general sense of magnitude, but not for precise comparisons.

Visualisation of Categorical Data

Pie Chart – share of each category in total sales



Pie Chart is a static chart, meaning it gives the share of the part in the total only for a specific time frame.

If you want to visualise how this share changes across years, one may have to plot multiple pie charts - one for each year.

For such purposes, one may also consider stacked bar charts placed side by side for each year or one of the other visualisations discussed in the section ahead for time-series data.

Visualisation of Categorical Data

Pie Chart – share of each category in total sales

Below are the steps to creating a pie chart in Tableau:

1. Create a new worksheet.
2. Drag and drop "Category" dimension on the "Color" marks card.
3. Drag and drop "Sales" measure on the "Size" marks card.
4. Choose "Pie Chart" as the chart type.
5. Drag and drop "Sales" measure on the "Angle" marks card.
6. Drag and drop "Category" dimension on "Label" marks card.
7. Drag and drop "Sales" measure on the "Label" marks card.
8. Click the drop down arrow on the "Sales" tab added to "Label" marks card in above step and select "Quick table calculation" -> "Percent of Total" from the drop down menu.
9. The result is a rather small pie. To make the chart bigger, hold down Ctrl + Shift (hold down ⌘ + z on a Mac) and press B several times.

Visualisation of Categorical Data

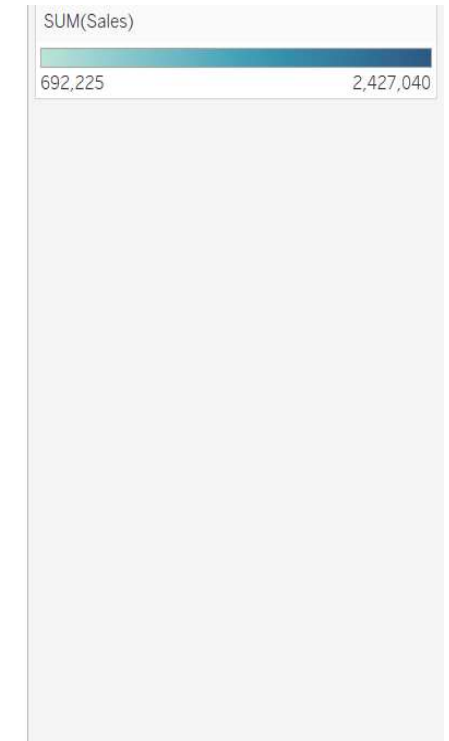
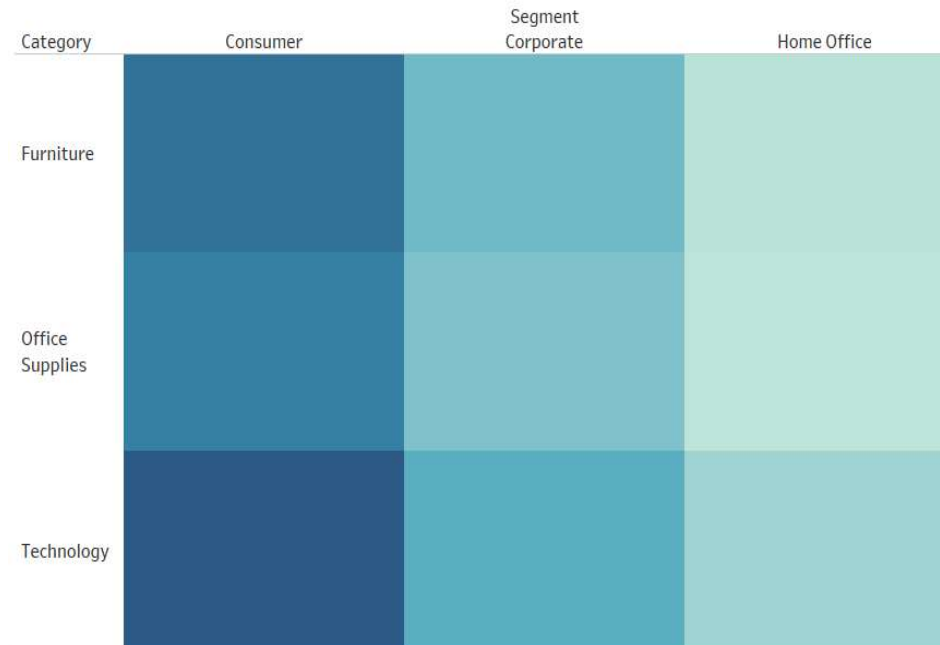
Heat Map- sales by category and consumer segment

Visualisation task:

When the sales manager would like to visualise the distribution of each product category sales across consumer segments using colours and sizes, he can use a heat map in Tableau.

Chart type

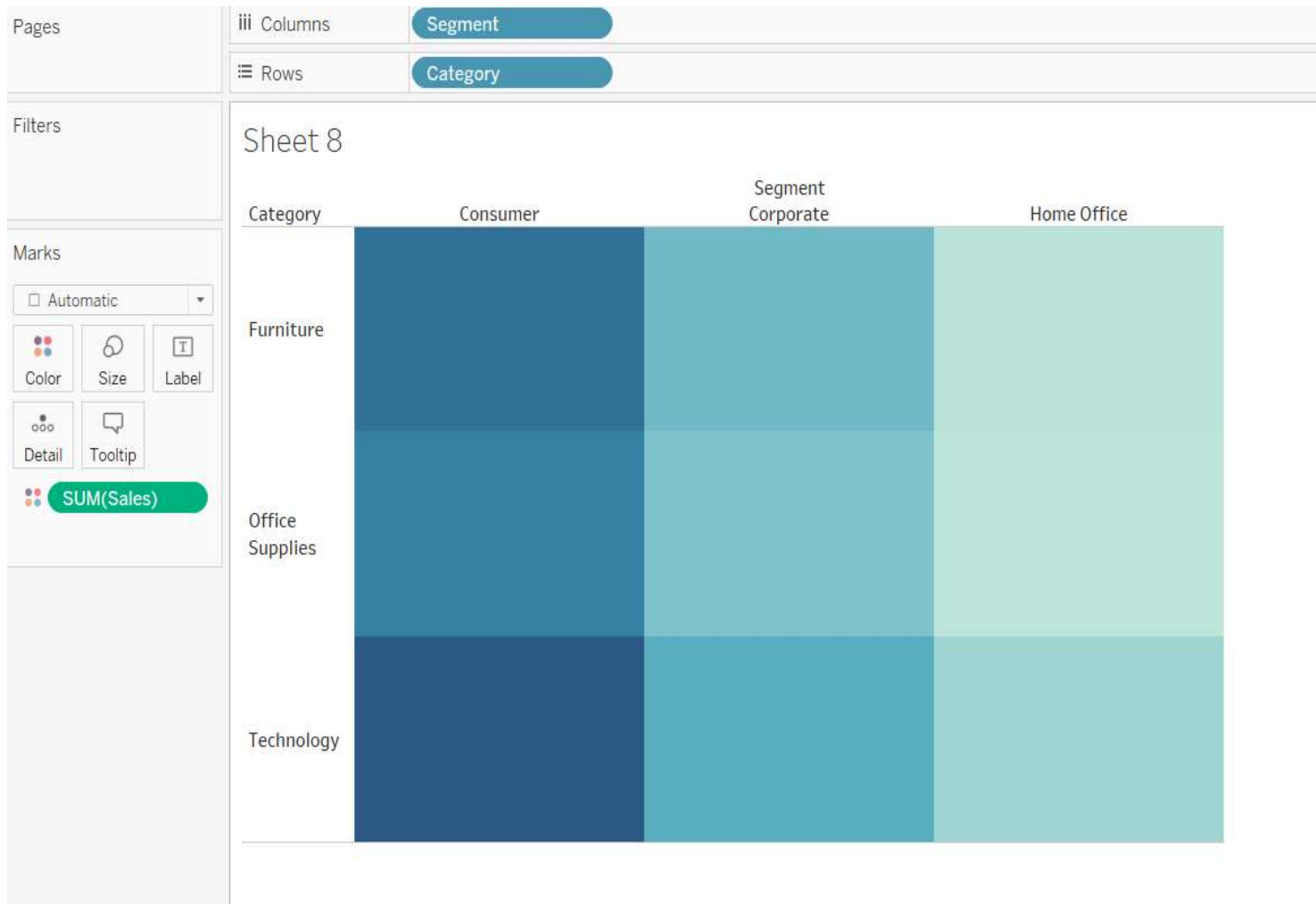
Sheet 8



A heat map is a data visualisation technique that shows the magnitude of a phenomenon as a colour in two dimensions. The variation in colour may be by hue or intensity, giving obvious visual cues to the reader about how the phenomenon is clustered or varies over space.

Visualisation of Categorical Data

Heat Map- sales by category and consumer segment



From the heatmap, we can see that the office supplies category sales lagged behind furniture and technology sales across consumer segments. But this could be due to the category size itself— hence for more conclusive evidence, it is recommended that one looks at the profitability of each of the categories across segments rather than absolute sales numbers.

Visualisation of Categorical Data

Heat Map- sales by category and consumer segment

Below are the steps to creating a heat map in Tableau:

1. Create a new worksheet.
2. Drag "Category" dimensions into worksheet's rows.
3. Drag "Segment" into worksheet's columns.
4. Drag and drop "Sales" measure on the "Color" marks card.

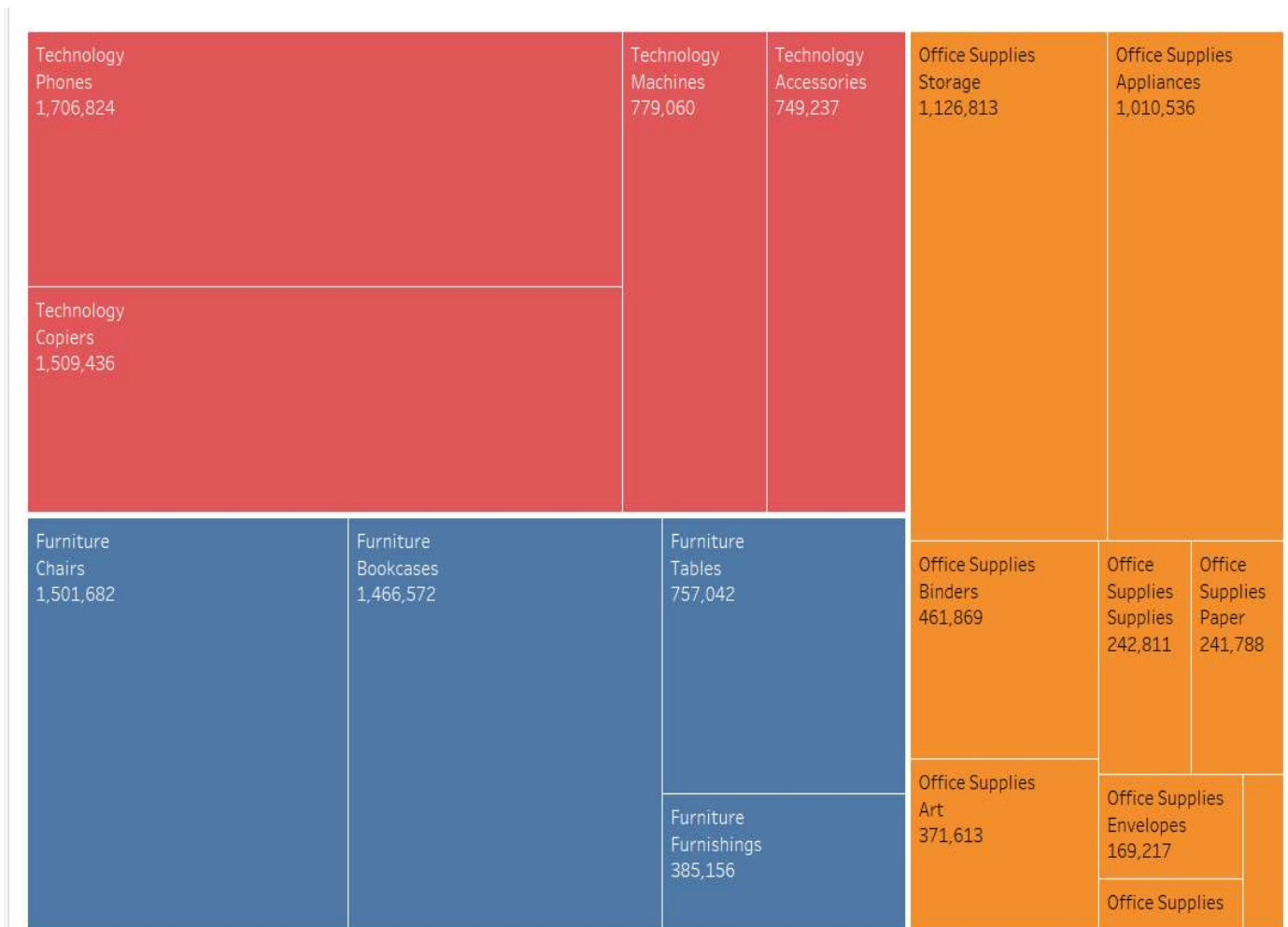
Visualisation of Categorical Data

Tree Map- Sales by category, sub-category

Visualisation task:

When the sales manager of the company would like to visualise the contribution of each product category sales to total sales and also each sub-category sales to category sales using colours and sizes, he/she can use a treemap in Tableau.

Chart type



Visualisation of Categorical Data

Tree Map- Sales by category, sub-category



Treemaps are ideal for displaying large amounts of hierarchically structured (tree-structured) data. The space in the visualisation is split up into rectangles that are sized and ordered by a quantitative variable.

Visualisation of Categorical Data

Tree Map- Sales by category, sub-category

Below are the steps to creating a treemap in Tableau:

1. Create a new worksheet.
2. Drag "Category" and "Sub-Category" dimensions into the worksheet's rows.
3. Drag "Sales" measure into the worksheet's columns.
4. Choose Treemaps as the chart type.
5. Drag and drop "Category" dimension" on the "Color" marks card.
6. Drag and drop "Sales" measure on the "Label" marks card.

Visualisation of Categorical Data

Simple Report- Profit by region and segment

Format the numbers:

1. Right click the table
2. Select format
3. Click Fields to select the field for formatting
4. Change the format of numbers to currency

The screenshot shows the Tableau interface with a pivot table titled "Profit Table by Region and Segment". The table has columns for Region, Segment, and Profit. The data is as follows:

Region	Segment	Profit
Central	Consumer	\$8,564
	Corporate	\$18,704
	Home Office	\$12,438
East	Consumer	\$41,191
	Corporate	\$23,623
	Home Office	\$26,709
South	Consumer	\$26,914
	Corporate	\$15,215
	Home Office	\$4,621
West	Consumer	\$57,451
	Corporate	\$34,437
	Home Office	\$16,530

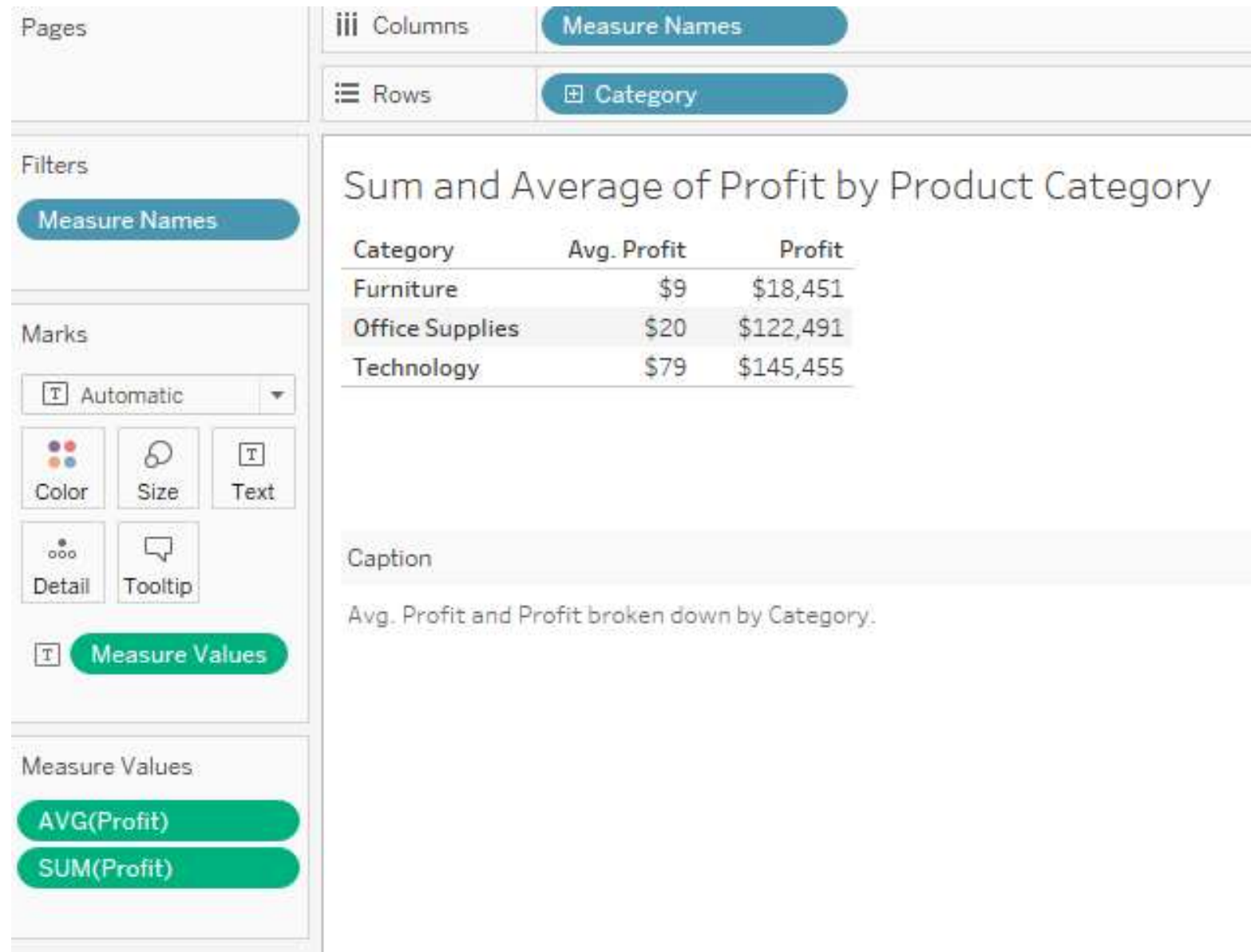
The interface also shows the "Columns" shelf with "Region" and "Segment", and the "Rows" shelf with "SUM(Profit)". The "Marks" shelf is set to "Automatic".

The screenshot shows the "Format SUM(Profit)" dialog box in Tableau. The "Fields" tab is selected, and the "Numbers" field is highlighted. The "Currency (Standard)" option is selected under "Numbers". The "Locale" is set to "English (United States)".

Region	Segment	Profit
Central	Consumer	\$8,564.05
	Corporate	\$18,703.90
	Home Office	\$12,438.41
East	Consumer	\$41,190.98
	Corporate	\$23,622.58
	Home Office	\$26,709.22
South	Consumer	\$26,913.57
	Corporate	\$15,215.22
	Home Office	\$4,620.63
West	Consumer	\$57,450.60
	Corporate	\$34,437.43
	Home Office	\$16,530.41

Visualisation of Categorical Data

Two-Measures Report- Sum and average of profit by product category





Visualisation of Time Series Data

Visualisation of Time Series Data

Discrete (bucketed) time series data

Time series data can be categorised as either discrete (bucketed) or continuous (unbroken)

- ▶ Discrete time series data values are from specific points or blocks of time, and there is a finite number of possible values
- ▶ A line chart is the most effective way to display time series data. A line chart for discrete time series data places breaks between time units like year, quarter, month and day
- ▶ Bar, stacked bar charts can also be used to display time series data

Visualisation of Time Series Data

Continuous (unbroken) time series data

- ▶ The difference between visualising continuous time series data from visualising discrete time series data is in what they represent in the physical world. Continuous time series data represents a constantly changing phenomena
- ▶ Line chart for continuous time series data is presented as unbroken lines
- ▶ Step charts are appropriate if the measure stays at a value for a long time and all of a sudden declines or inclines
- ▶ Trendlines are useful in the presence of noisy data or large amounts of data

Visualisation of Time Series Data

Visualising time series data

- ▶ Line Chart, line chart with dual axis, stacked line chart
- ▶ Stacked area chart
- ▶ Gantt chart
- ▶ Trendline
- ▶ Reference Line

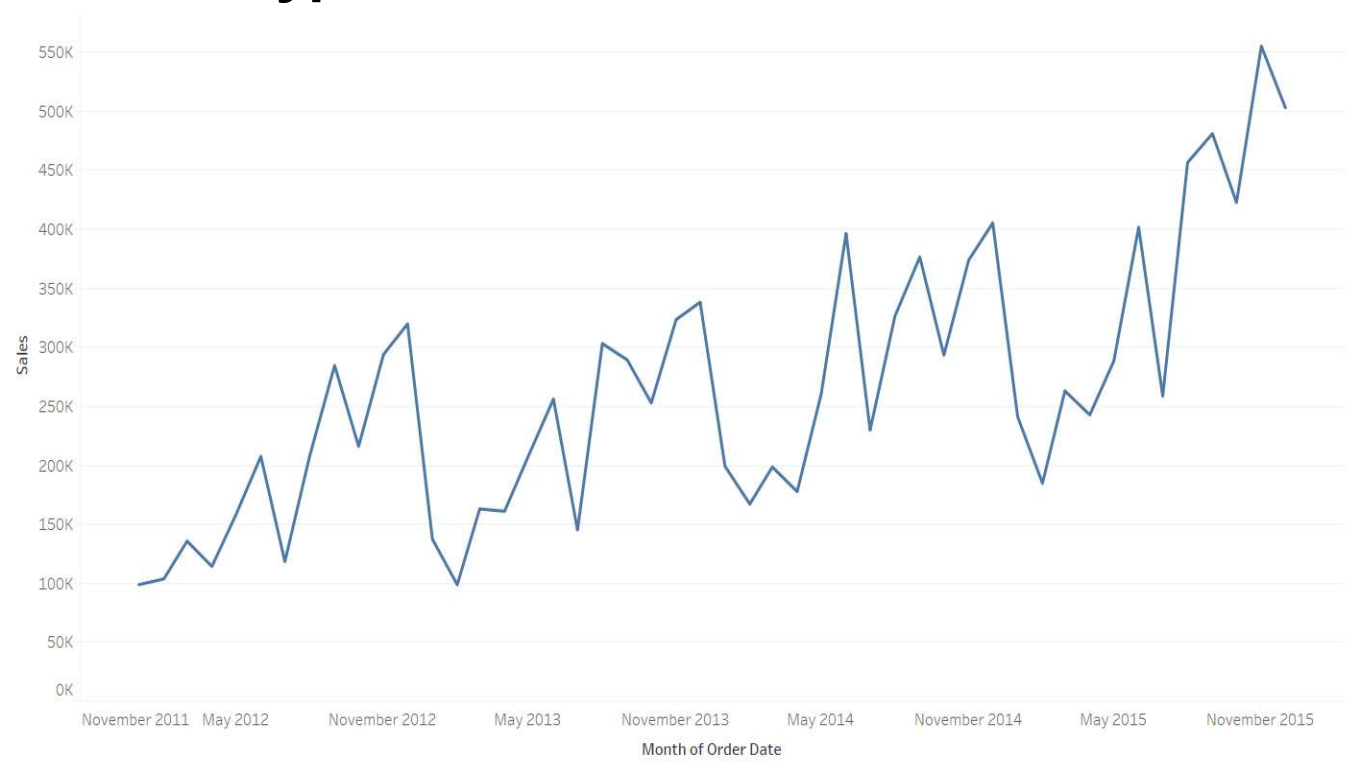
Visualisation of Time Series Data

Line chart – sales by month of the year

Visualisation task:

The sales vice president of the global superstore would like to visualise monthly sales trends between the years 2011 and 2015. He can use the line chart in Tableau.

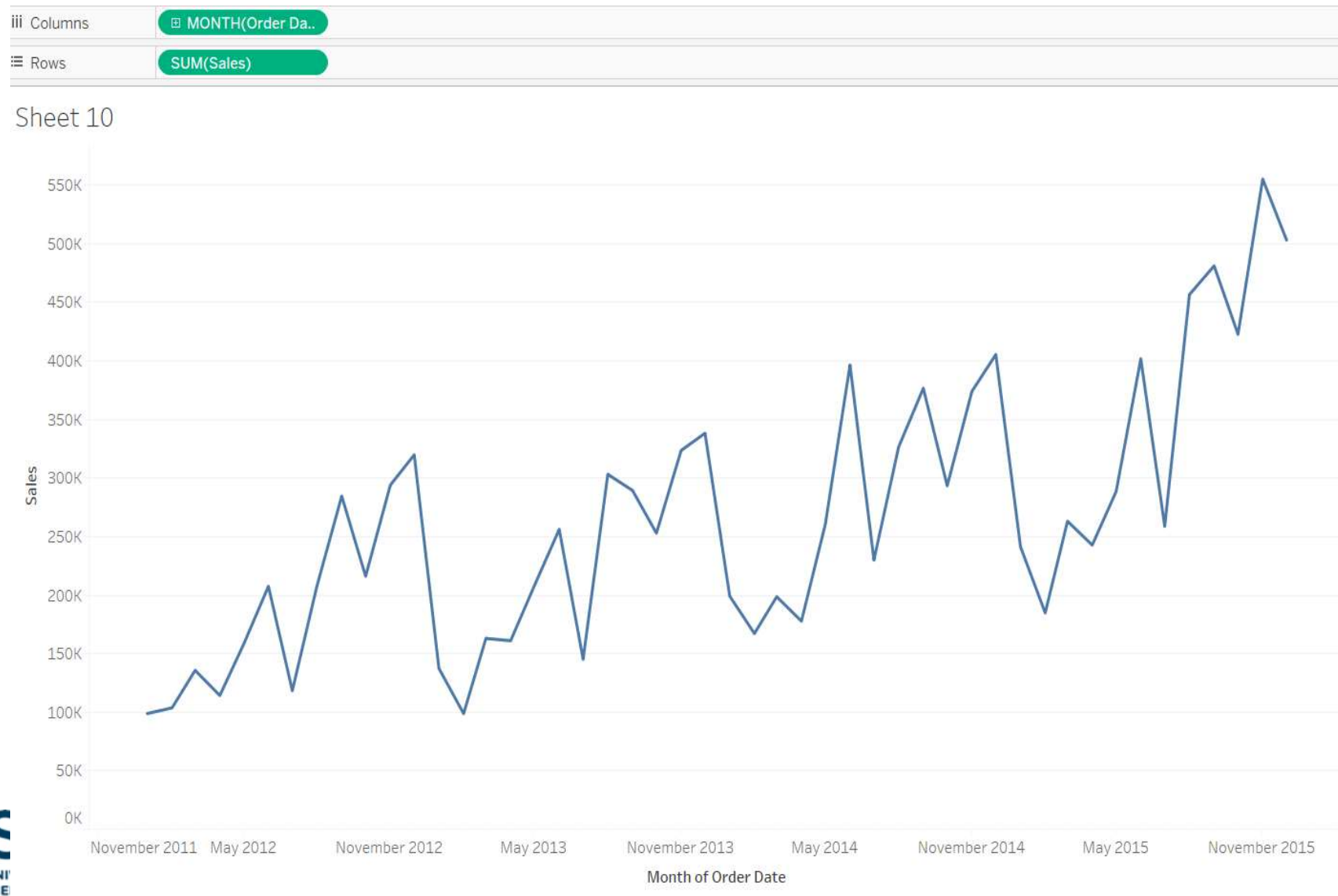
Chart type



The line chart is the most effective way to display time series data. Discrete time series data are presented with breaks between time units, while continuous time series data are presented in unbroken lines.

Visualisation of Time Series Data

Line chart – sales by month of the year



Visualisation of Time Series Data

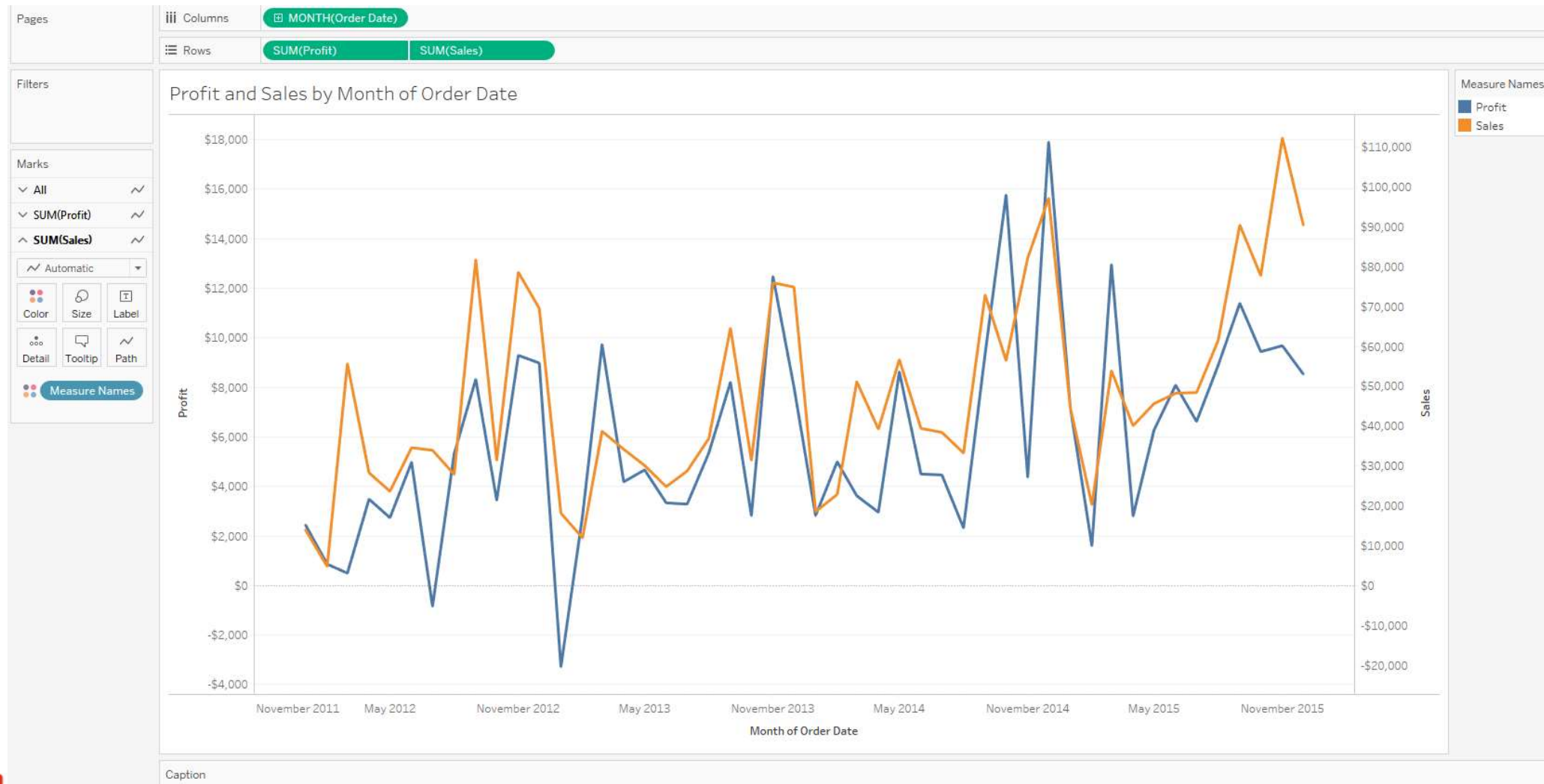
Line chart – sales by month of the year

Below are the steps to creating a line chart in Tableau:

1. Create a new worksheet.
2. Drag "Order Date" dimension into the worksheet's columns.
3. Expand the x-axis into month and year by clicking on the "Order Date" dimension in the columns and selecting "Month year" option from the drop down menu.
4. Drag "Sales" measure into the worksheet's rows.

Visualisation of Time Series Data

Line chart with dual axis– sales and profit by month of the year



Caption

The trends of Profit and Sales for Order Date Month. Color shows details about Profit and Sales.

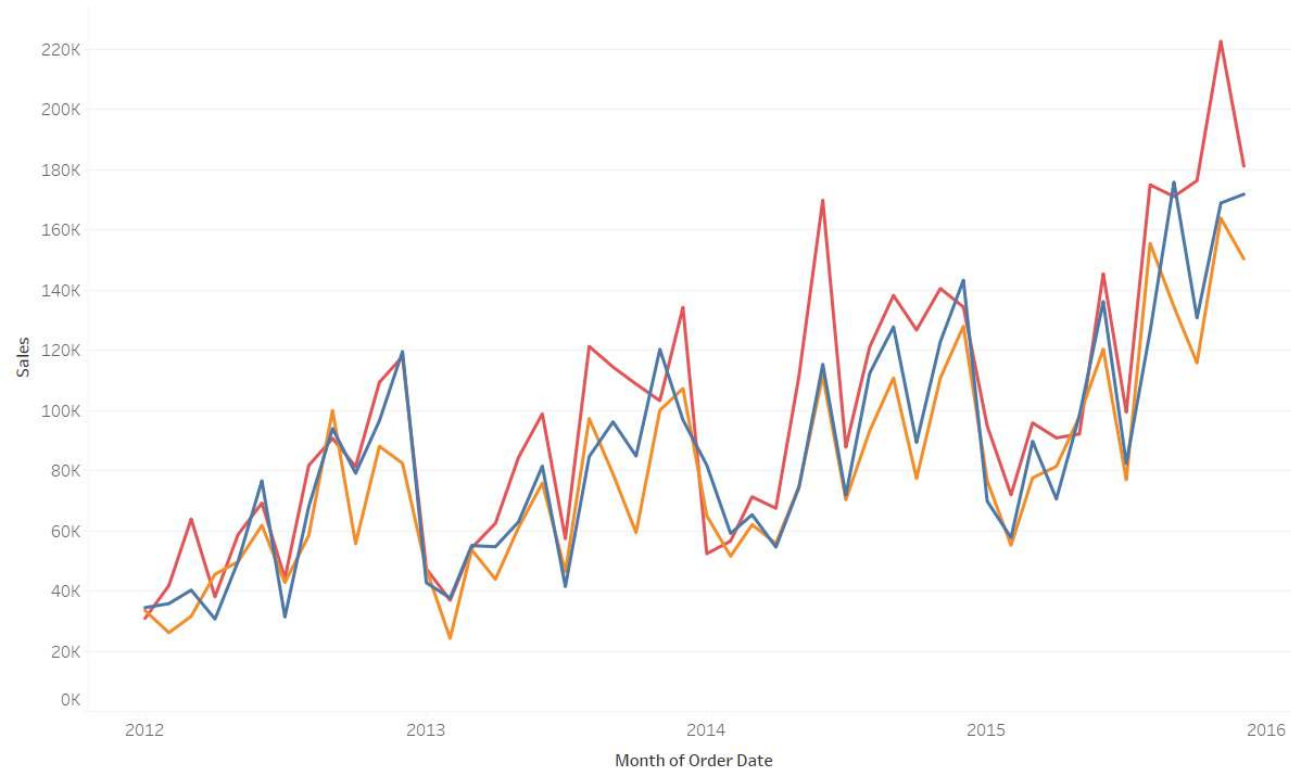
Visualisation of Time Series Data

Stacked Line chart – monthly sales by product category across years

Visualisation task:

To take a closer look at the sales trends, the sales vice president of the same global superstore would like to visualise monthly sales trends between the years 2011 and 2015 for each of the product category. This will also give him/her a fair idea of each category's performance.

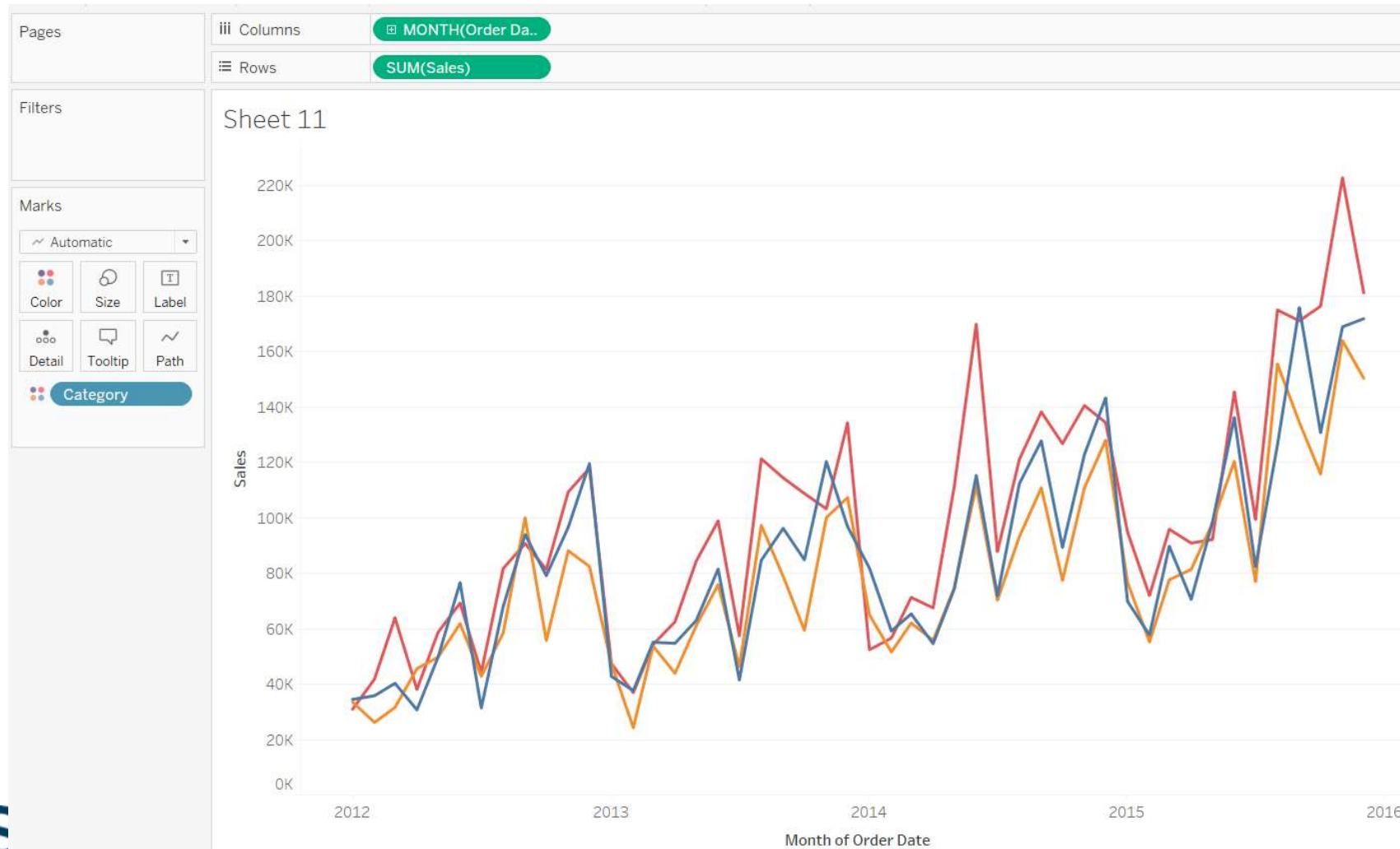
Chart type



One of the common ways to visualise this is by drawing a different coloured line for each product category

Visualisation of Time Series Data

Stacked Line chart – monthly sales by product category across years



Visualisation of Time Series Data

Stacked Line chart – monthly sales by product category across years

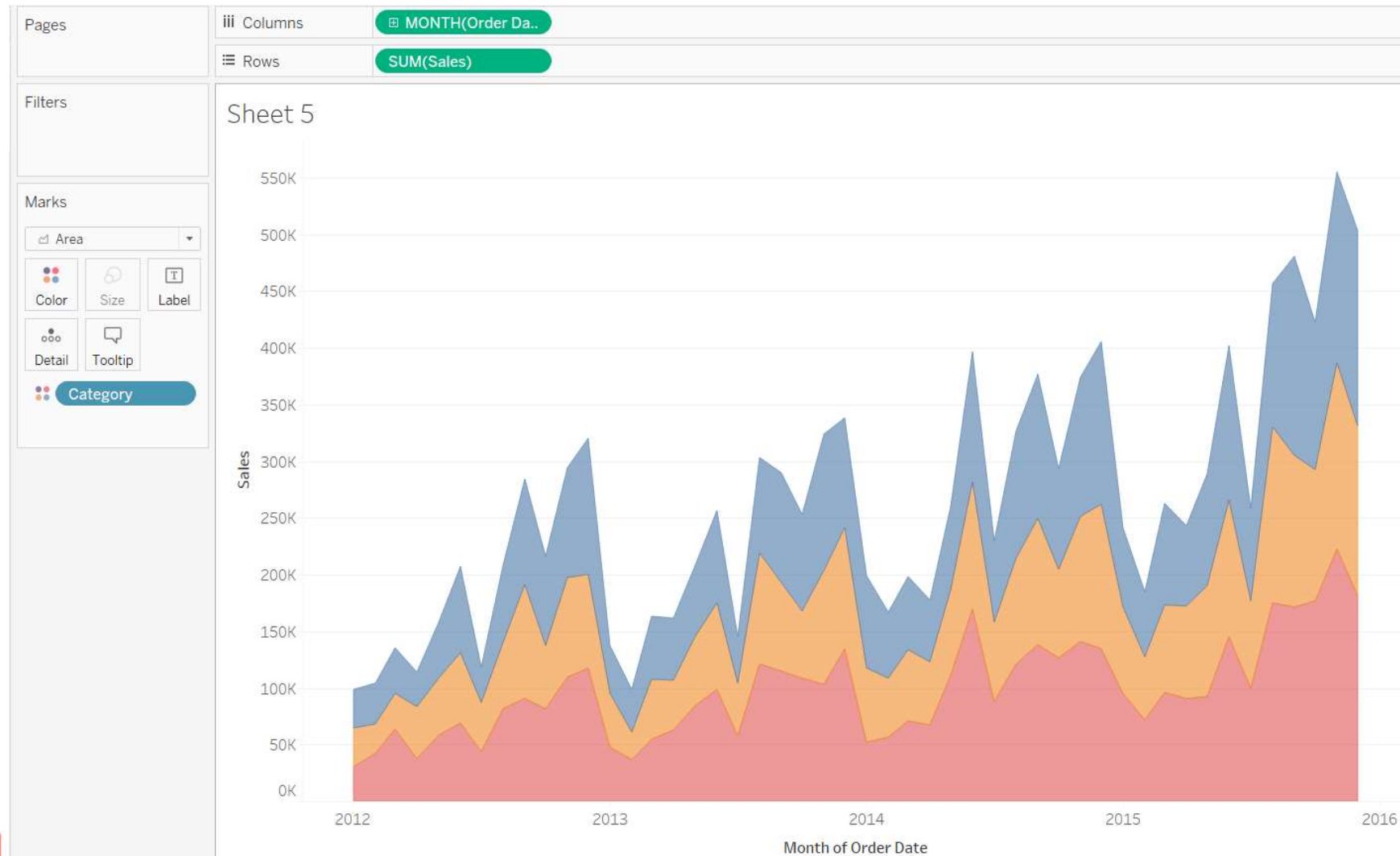
Below are the steps to creating a stacked line chart in Tableau:

1. Create a new worksheet.
2. Drag "Order Date" dimension into the worksheet's columns.
3. Expand the x-axis into month and year by clicking on the "Order Date" dimension in the columns and selecting "Month year" option from the drop down menu.
4. Drag "Sales" measure into the worksheet's rows.
5. Drag and drop "Category" dimension on the "Color" marks card on left.

It is difficult to contrast or compare the sales trend of each of the product category across months as the lines are overlapping and it looks visually cluttered. To overcome this drawback, another option is to use a stacked area chart. The stacked chart plots cumulative values of the measure on the vertical axis, taking one of the categories as the baseline and plotting every other category one above the next. This prevents overlapping of lines or data points and makes the chart visually easier to interpret.

Visualisation of Time Series Data

Stacked Area Chart – monthly sales by product category across years



The stacked area chart is a variation of the multi-line chart in that the lines do not cross each other and each coloured area is well demarcated from the other.

Visualisation of Time Series Data

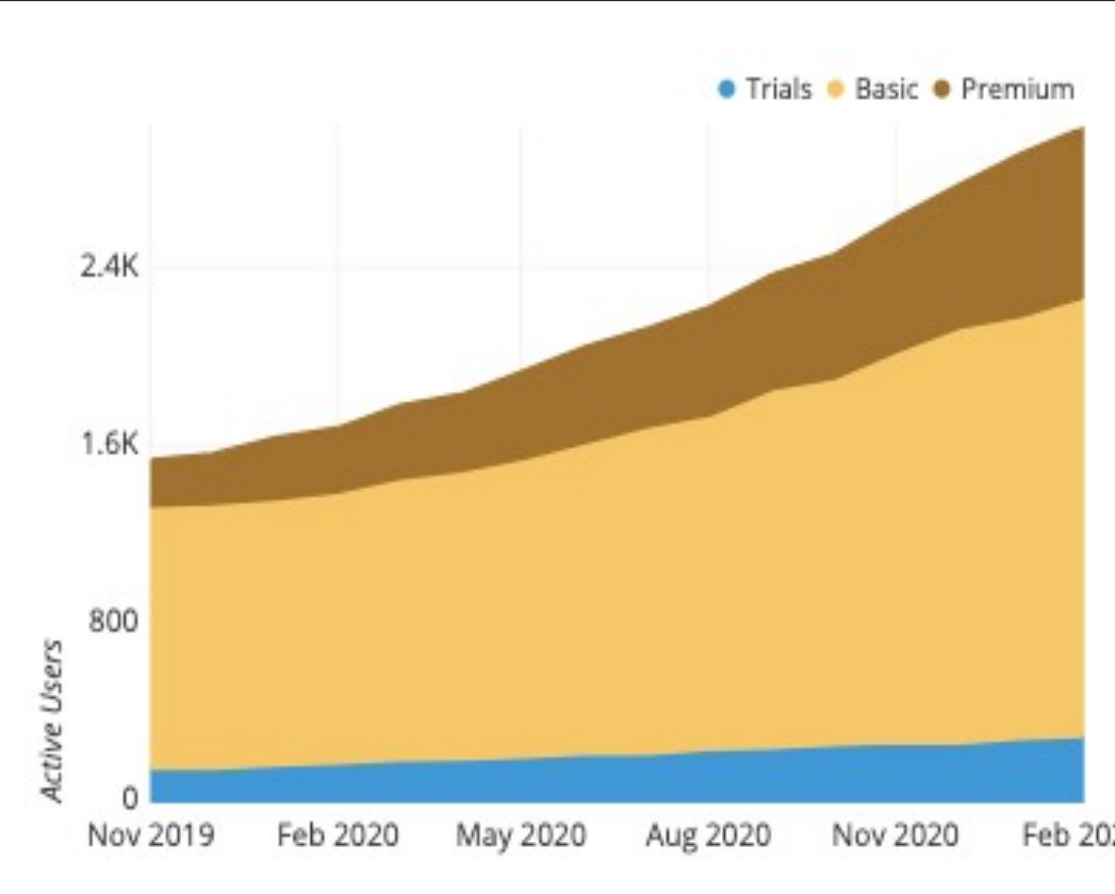
Stacked Area Chart – monthly sales by product category across years

Below are the steps to create a stacked area chart in Tableau:

1. Create a new worksheet.
2. Drag “Order Date” dimension into the worksheet’s columns.
3. Expand the x-axis into month and year by clicking on the “Order Date” dimension in the columns and selecting “Month year” option from the drop down menu.
4. Drag “Sales” measure into the worksheet’s rows.
5. Drag and drop “Category” dimension on the “Color” marks card on the left.
6. Choose “Area Chart” as chart type.

Visualisation of Time Series Data

Stacked Area Chart – How to interpret stacked area charts?



MONTH	TRIALS	BASIC	PREMIUM
2019-11	154	1180	201
2019-12	157	1186	219
2020-01	170	1195	270
2020-02	180	1213	285

MONTH	TRIALS	+ BASIC	+ PREMIUM
2019-11	154	1334	1535
2019-12	157	1343	1562
2020-01	170	1365	1635
2020-02	180	1393	1678

Figure 4.15: Number of active users per month basis subscription type

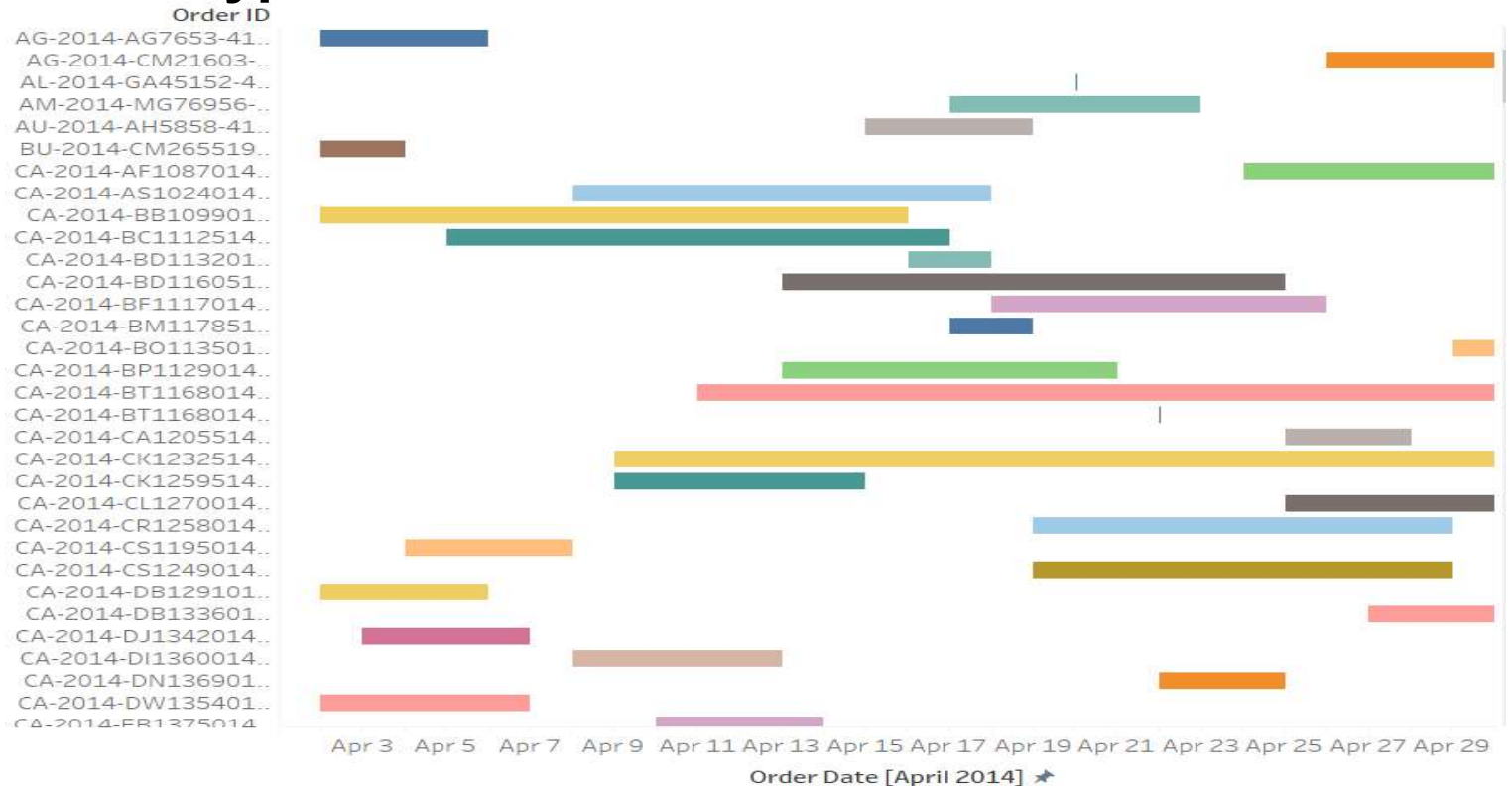
Visualisation of Time Series Data

Gantt chart- Shipping lead time for orders in April

Visualisation task:

Let us say that the supply chain manager of the global superstores would like to visualise the shipping lead time (time between order and shipping date) for each order in the month of April. He/she can use a Gantt chart in Tableau.

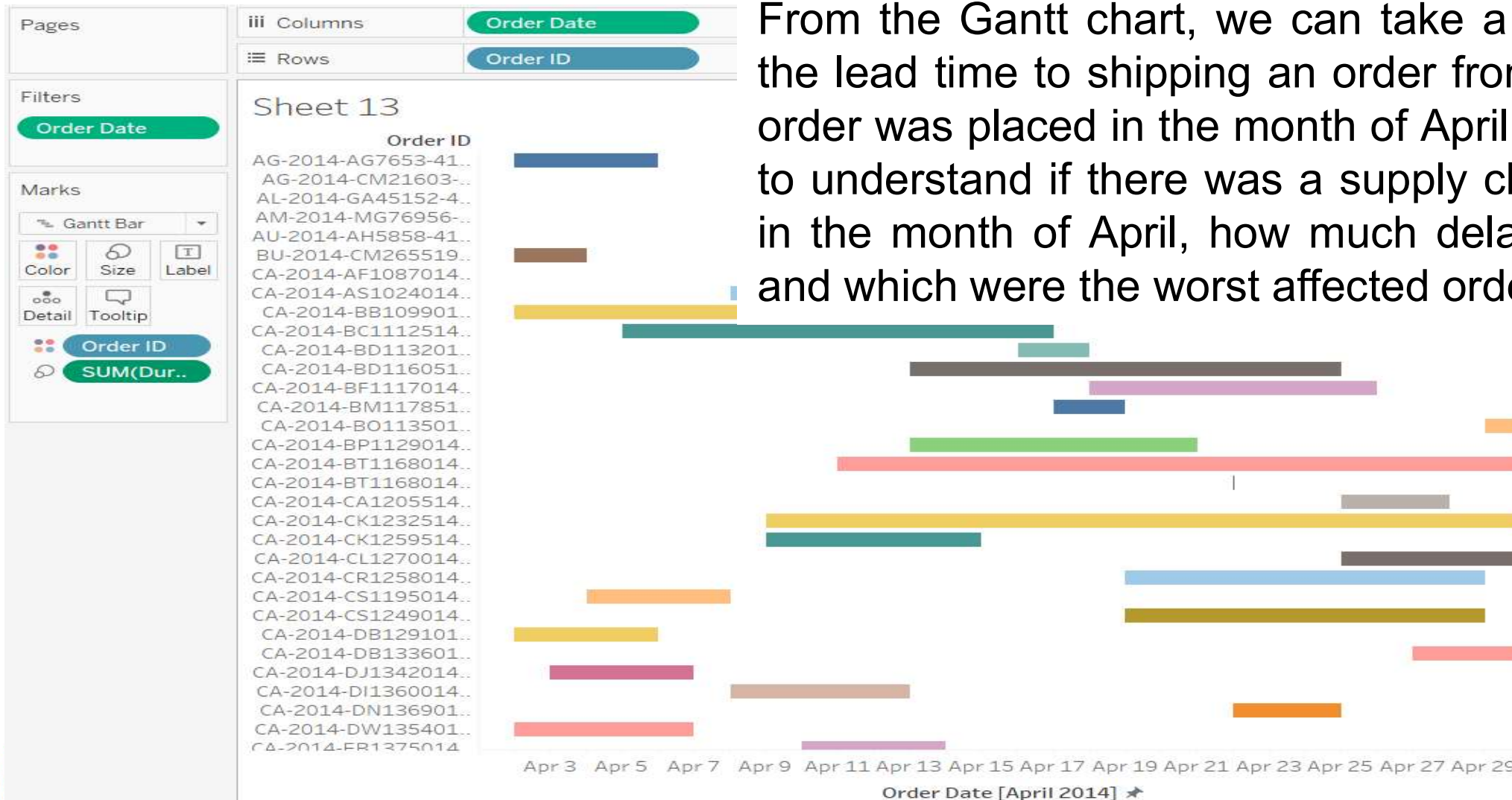
Chart type



A Gantt chart is commonly being used for project planning. The activity or event start time is visualised by the bar's horizontal position, and the duration of each activity or event is visualised by the individual bar length

Visualisation of Time Series Data

Gantt chart- Shipping lead time for orders in April



Visualisation of Time Series Data

Gantt chart- Shipping lead time for orders in April

Below are the steps to creating a Gantt chart in Tableau:

1. Create a new worksheet.
2. Select “Analysis”, then “Create Calculated Field”.
3. In the “Calculated Field” dialogue box, type “Duration in Days” in the name textbox and type “DATEDIFF” ('second', [Order Date], [Ship Date])/86400.
4. Drag the “Order Date” field into the worksheet’s columns.
5. Right click on the “Order Date” field and select “Exact Date”.
6. Drag the “Order ID” field into the worksheet’s rows and into the “Colour Mark”.
7. On the Mark Card’s drop down menu, select “Gantt Bar”.
8. Drag the “Duration in Days” field to “Size”.
9. Right click on the “Order Date” field and select “Edit Axis”.
10. In the “Edit Axis” dialogue box, under “Range”, select “Fixed”. Change the date range, and click “Apply”.
11. On the “Tick Marks” tab, fix the major and minor tick marks to every one day. Change the origin date to April 1, 2014. When finished, click “OK”.
12. Drag “Order Date” field into Filters Cards to filter the “Order Date” from April 1, 2014 to April 30, 2014.

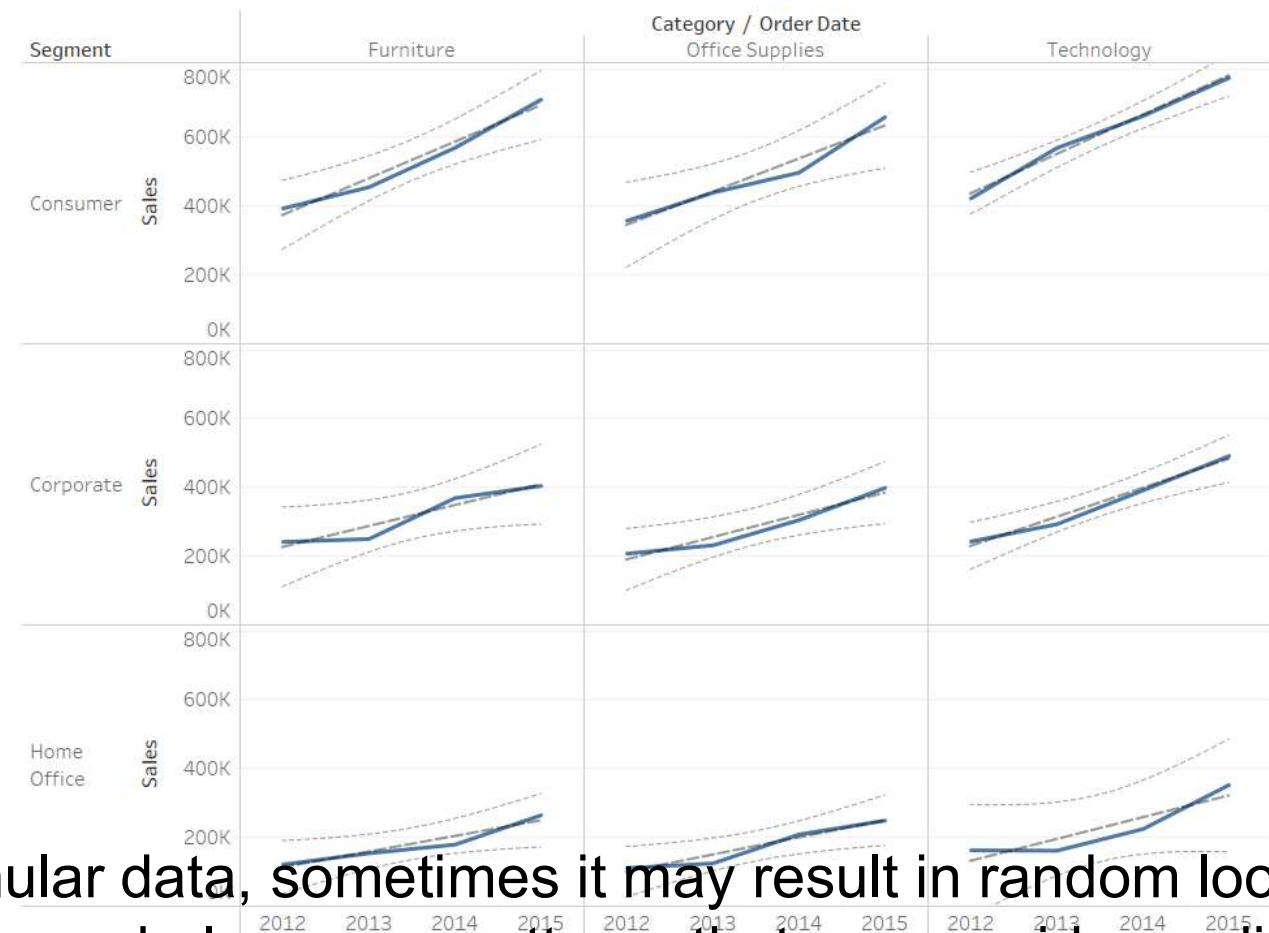
Visualisation of Time Series Data

Adding a trend line to the yearly sales graphs for each segment and category

Visualisation task:

When the VP of sales of the company would like to see the year-on-year sales trend for each product category across consumer segments, he/she can use trend lines together with a line chart in Tableau.

Chart type



When we try to visualise granular data, sometimes it may result in random looking data visualisation. Trend lines can help us see patterns that can provide predictive value, by drawing a line that best fits the values in the visualisation.

Visualisation of Time Series Data

Adding a trend line to the yearly sales graphs for each segment and category



From the graph, we can see that: Trend line has the steepest slope for the “Consumer” segment as compared to the “Corporate” segment, followed by the “Home office” segment, for all three product categories.

We can say that the rise in year-on-year sales for the “Consumer” segment is the highest, followed by the “Corporate” segment, and sales remained at the same level or showed marginal increment for the “Home office” segment.

Visualisation of Time Series Data

Adding a trend line to the yearly sales graphs for each segment and category

Below are the steps to creating trend lines in Tableau:

1. Create a new worksheet.
2. Drag “Category” and ”Order Date” dimension into the worksheet’s columns. From the drop down menu for “Order Date” dimension select “Year”.
3. Drag “Segment” dimension and “Sales” measure into the worksheet’s rows.
4. Right click on the chart and select “Show Trend Lines”.
5. For more the “Trend Lines” option, point at the “Trend Line”, right click, and select “Edit Trend Lines”. This will expose the Trend Line Menu. Select “Show Confidence Bands”.

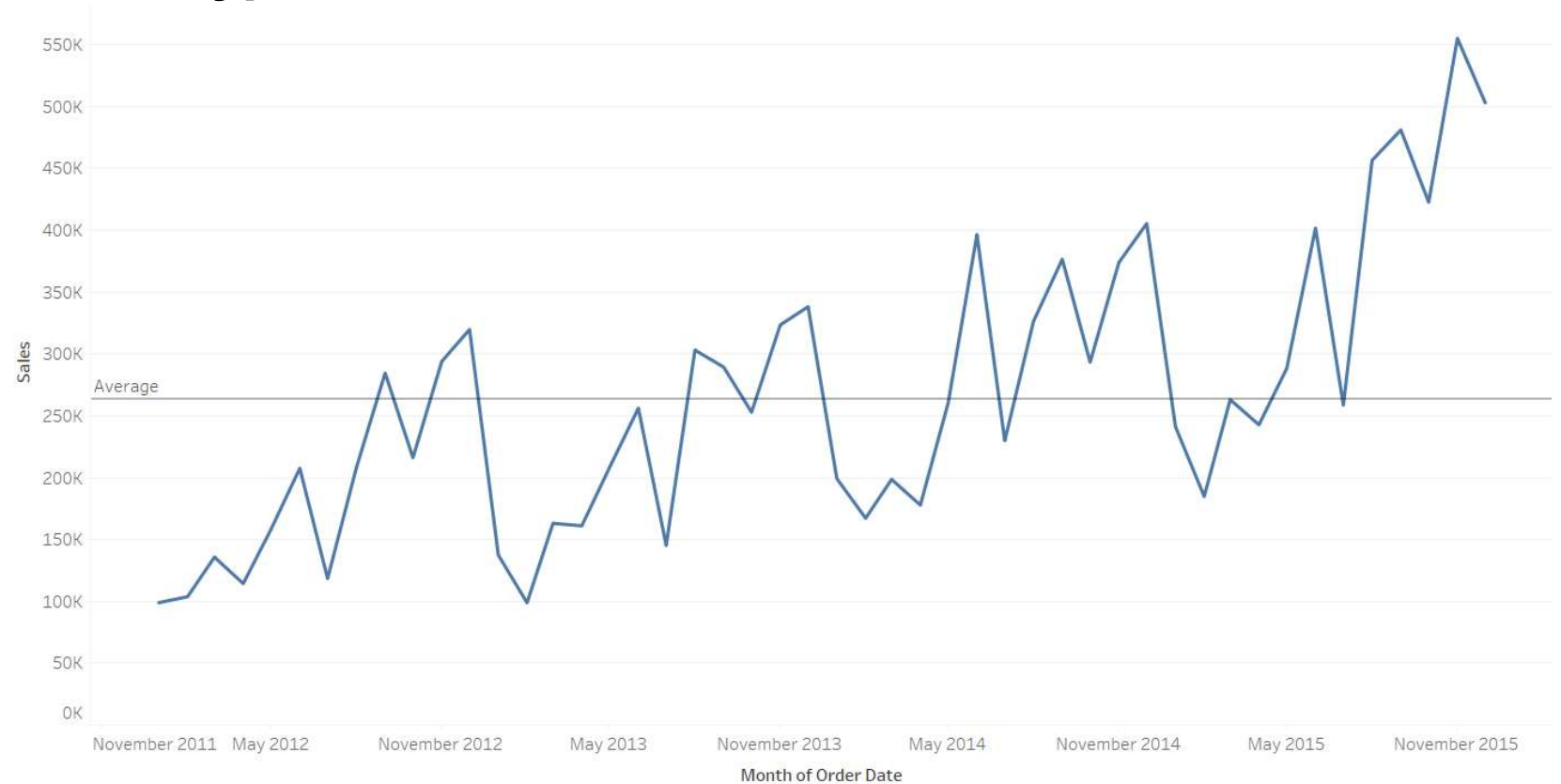
Visualisation of Time Series Data

Adding a reference line

Visualisation task:

When the VP of sales of the company would like to compare the actual sales against average sales in between the years 2012 and 2015, he/she can use reference lines together with a line chart in Tableau.

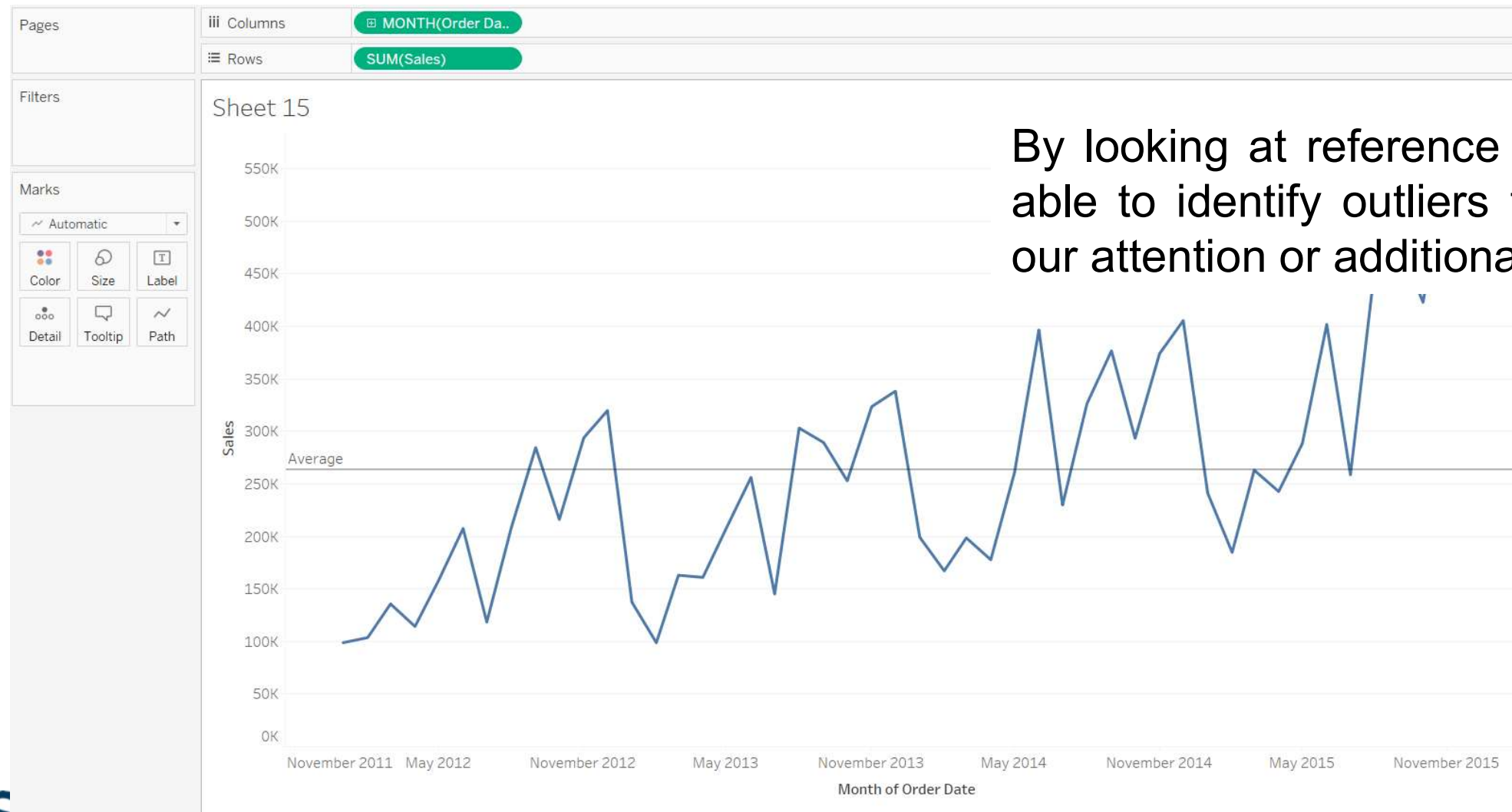
Chart type



Reference lines can help us compare the actual data presented in the visualisation against targets

Visualisation of Time Series Data

Adding a reference line



Visualisation of Time Series Data

Adding a reference line

Below are the steps to creating reference lines in Tableau:

1. Create a chart, for example, a line chart, by following the steps that was described in the line chart section previously.
2. Right click on the axis on which we want to apply the reference line, and select “Add Reference Line”, “Band”, or “Box”.
3. Explore the line, band, and distribution buttons in conjunction with the computation value’s dropdown menu to see all the available options for reference line types.



Tableau Exercises

Homework



Create the following in Tableau:

Use sample_superstore.xls

1. Pie chart
2. Stacked bar chart
3. Line chart with trend line
4. Area chart
5. Gantt chart
6. Heat Map