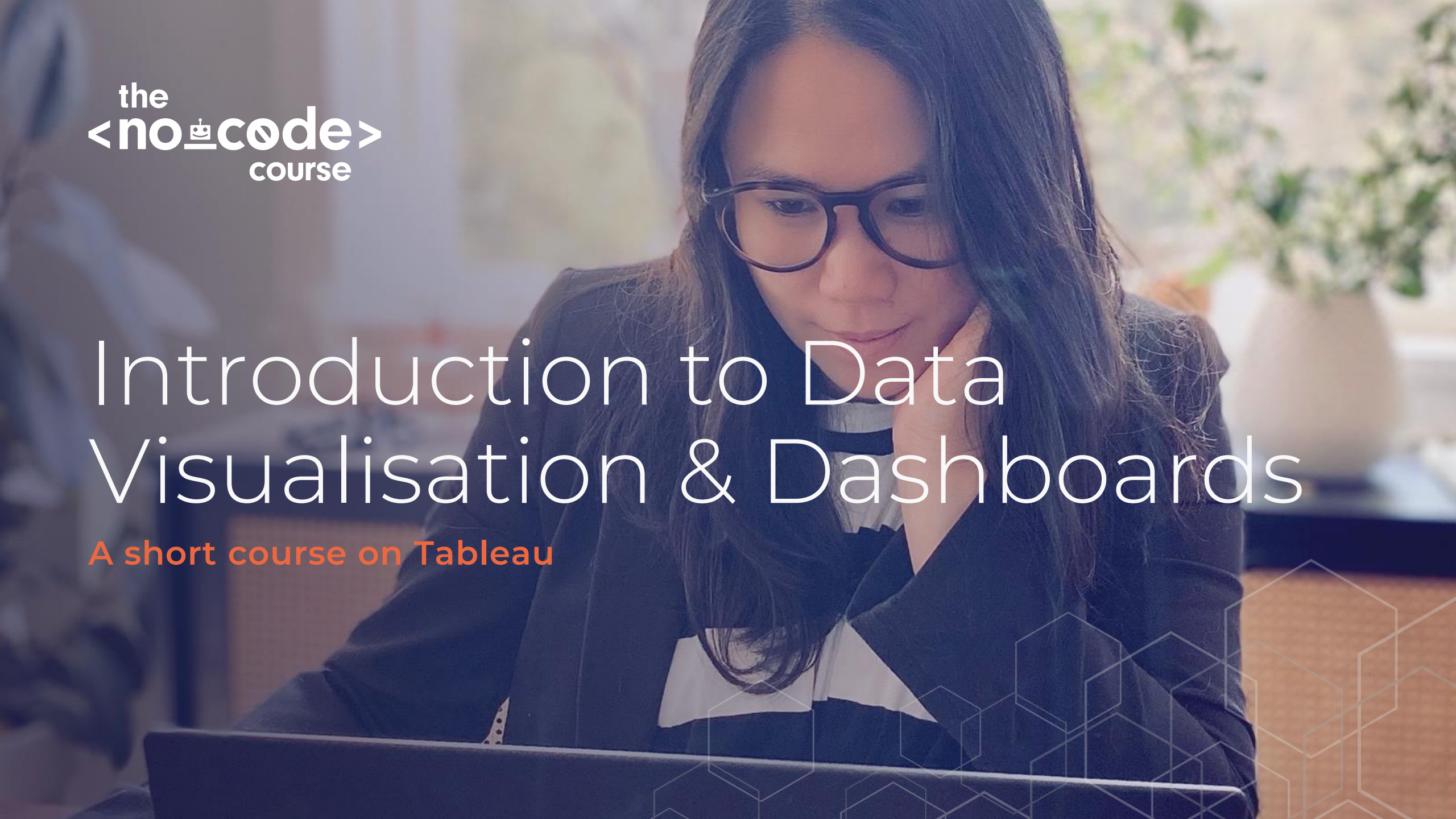


Introduction to Data Visualisation & Dashboards

A short course on Tableau

Table of Contents

You will need Tableau Desktop 2021 installed for this seminar

the <no code> course

Welcome: Learning Objectives & Introductions

About the Instructor

Munish Kumar, PhD, B Eng

Staff Geoscientist (Petrophysicist) – ERCE

- 12 years in oil & gas; worked at ExxonMobil and Total S.A. in strategy, portfolio, planning, upstream exploration, field development and production optimization
- In 2020, joined expert consultancy firm ERCE which provides energy consulting/auditing services (including greenhouse gas evaluation) to oil & gas, law firms, banks, institutional investors etc.

Adjunct Faculty – SUSS

- Data Visualisation for Business
- Python for Business
- Applied Project for Business Analytics

Tools of the trade: Python, Tableau, Power BI, Excel

INTRODUCE YOURSELF

Name

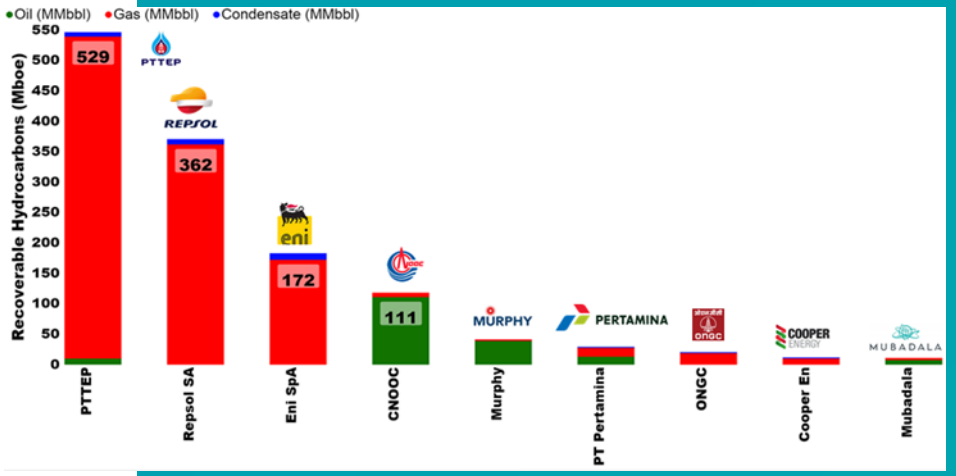
Profession/ What you do/ What you like

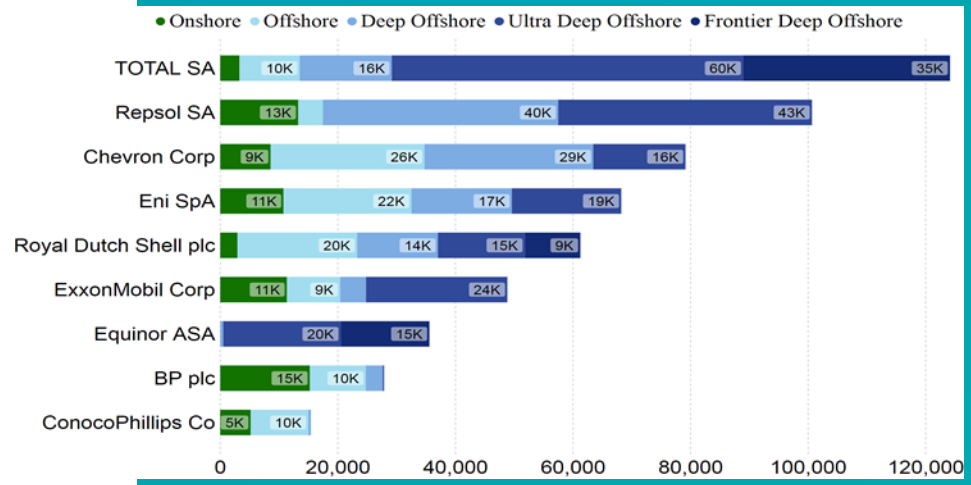
Do you currently use visualisation?

What would be one thing you want to take away from this class?

Learning Objectives

- Understand concepts of Data and Data Visualization
- Stages of Data Visualization Process
- Data Visualization Components
- Design Principles of a Business Performance Dashboard
- Appropriateness of Data Visualization techniques based on available data
- Appropriates of Business Performance Measure against Business Strategy of Organization





Some Textbooks:

1. Yau, N,.(2013). Data Points: Visualization that Means Something
2. Murray, D. G. (2013). Tableau Your Data

the <no code> course

Part 1 – Understanding Data

Understanding Data

The four measurement levels of data quality attribute

1. *Nominal* measurements measure items based on their labels or categories or other qualitative classification the items belong to with no implied order

2. *Ordinal* measurements arise from the operation of rank ordering

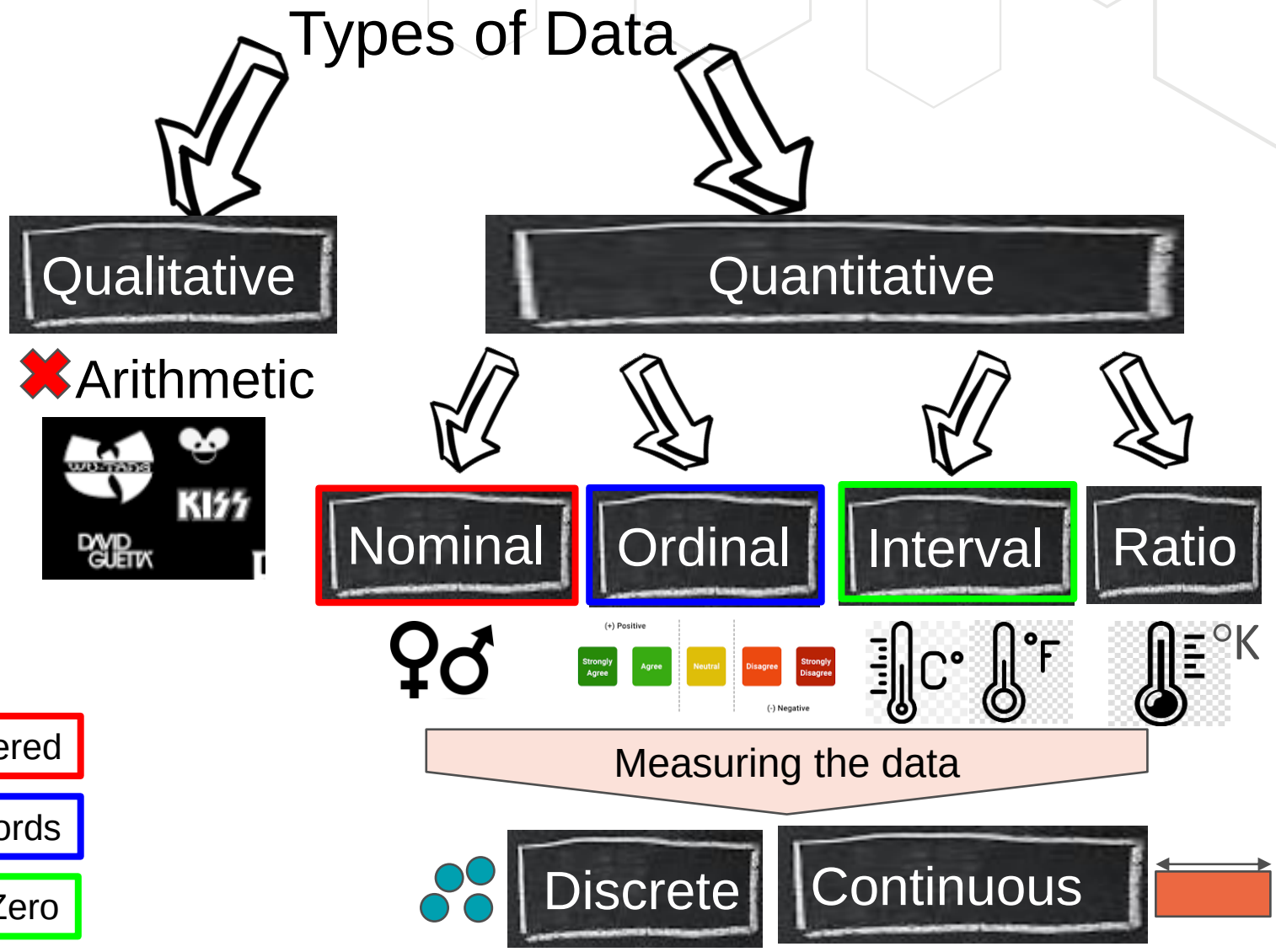
3. *Interval* measurements allow us to measure the degree of difference between items, but not the ratio between them

4. *Ratio* measurements estimate the ratio between a magnitude of a continuous quantity and a unit magnitude of the same kind

Understanding Types of Data

Tableau refers to interval or ratio data as “measures”.

All others are known as “dimensions”



How Tableau Views Types of Data

- Dimensions contain qualitative values (such as names, dates, or geographical data).
  - Dimensions can categorize, segment, and reveal the details in your data.
  - Dimensions affect the level of detail in the view.
- Measures contain numeric, quantitative values that you can measure.
  - Measures can be aggregated.
  - When you drag a measure into the view, Tableau applies an aggregation to that measure (by default)

- Green measures SUM(Profit) and dimensions YEAR(Order Date) are continuous. Continuous field values are treated as an infinite range. Generally, continuous fields add axes to the view.
- Blue measures SUM(Profit) and dimensions Product Name are discrete. Discrete values are treated as finite. Generally, discrete fields add headers to the view.

Understanding Data

Metadata

- Metadata is structured information that explain, describe or locate the original (i.e. also known as primary data), otherwise make the using of original data more efficient

Table Name	Column Name	Column Null	Column Datatype	Primary Key	Foreign Key
SOURCE_AUTO_LOAN_BY_WEB	AUTO_LN_ID	NOT NULL	VARCHAR2(20)	Yes	No
	AUTO_LN_AMT	NOT NULL	NUMBER(11)	No	No
	AUTO_VIN_ID	NOT NULL	VARCHAR2(16)	No	No
	BOR_FST_NAME	NOT NULL	VARCHAR2(20)	No	No
	BOR_LAST_NAME	NOT NULL	VARCHAR2(20)	No	No
	DATETIMESTMP	NOT NULL	DATE	No	No
TARGET_AUTO_LOAN_BY_WEB	AUTO_LN_ID	NOT NULL	VARCHAR2(20)	Yes	No
	AUTO_LN_AMT	NOT NULL	NUMBER(11)	No	No
	AUTO_LN_BRK_COMIS_AMT	NOT NULL	NUMBER	No	No
	AUTO_VIN_ID	NOT NULL	VARCHAR2(20)	No	No
	BOR_FULL_NAME	NOT NULL	VARCHAR2(40)	No	No
	DATETIMESTMP	NOT NULL	DATE	No	No

Source: http://learndatamodeling.com/wp-content/uploads/2015/10/tmeta.gif

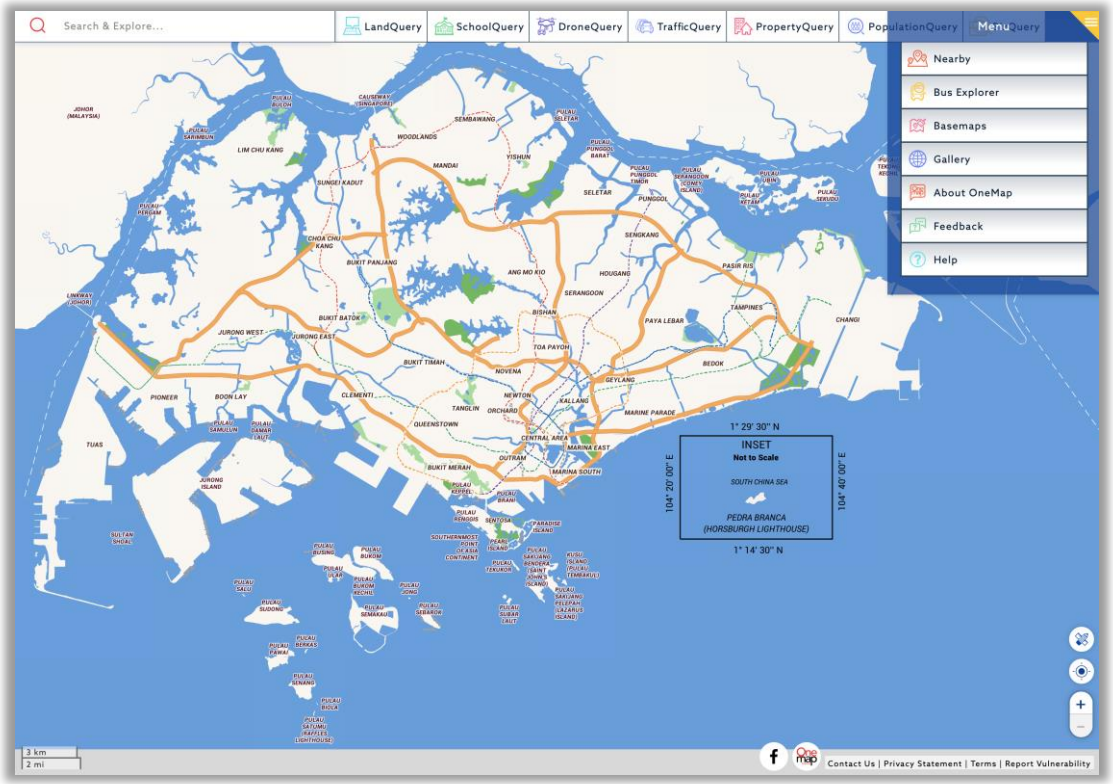
Part 2 – An Introduction to Data Visualisation

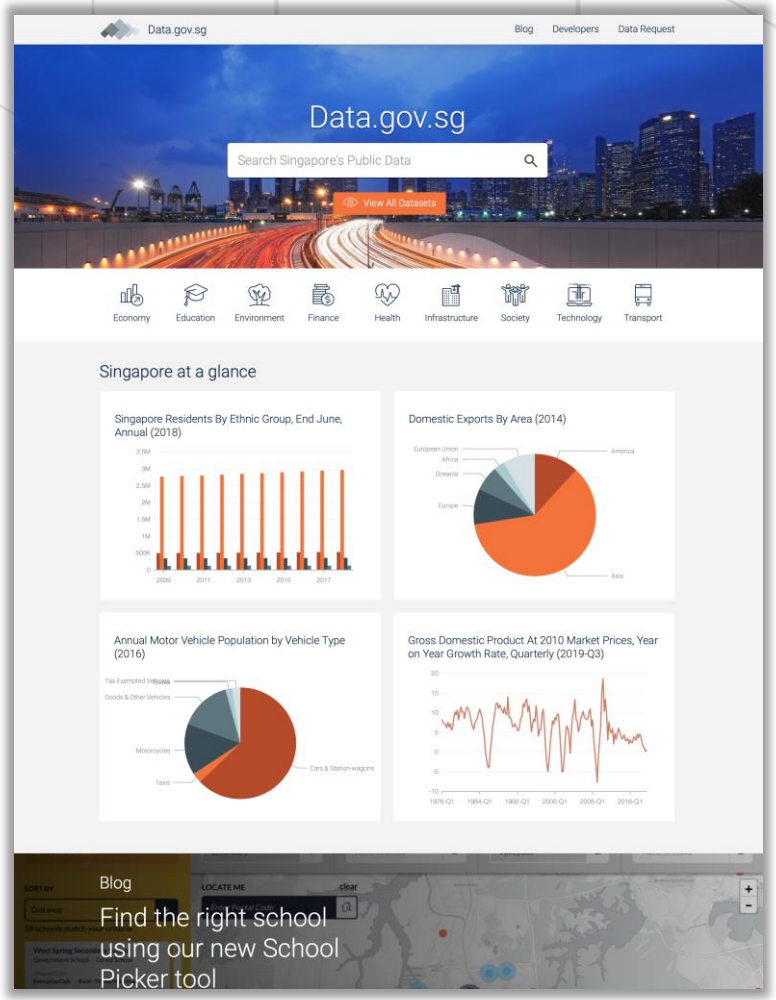
Data Visualisation

Overwhelming amount of data available today

A coloured map is essentially millions of data points/ measurements.

Overwhelming in an excel sheet, but easy to see if visualized properly





the <nocode> course

Data Visualisation

The four stages of the data visualisation process

1. *Data Collection and Storage:* the collection and storage of data

2. *Data Pre-processing:* the pre-processing of data to transform it into something one can understand

3. *Graphics Engine:* the display hardware and the graphics algorithms to produce data visualisation on screen

4. *Human Visual and Cognitive Processing:* human perceptual and cognitive systems that are involved in interpreting the visualised data

Source: Information Visualization: Perception for Design, C. Ware, 2013

Data Visualisation

Benefits

- Provides us the ability to comprehend huge amounts of data
- Allows the perception of emergent properties that are not anticipated (e.g. red for danger)
- Often enable problems with data to become immediately apparent (missing data or blank (pixels))
- Facilitates the understanding of both large-scale and small-scale features of the data
- Facilitates hypothesis formation

5 mins

Part 3 – Tableau as a Tool for Business Analytics & Visualisation

Tableau Interface

A. Workbook - contains sheets (either a worksheet, dashboard, or story).

B. Cards and shelves - Drag fields to the cards and shelves in the workspace to add data to your view.

C. Toolbar - To access commands, analysis and navigation tools.

D. View - Workspace (Canvas) to create visualization ("viz").

E. Click this icon to go to the Start page to connect to data.

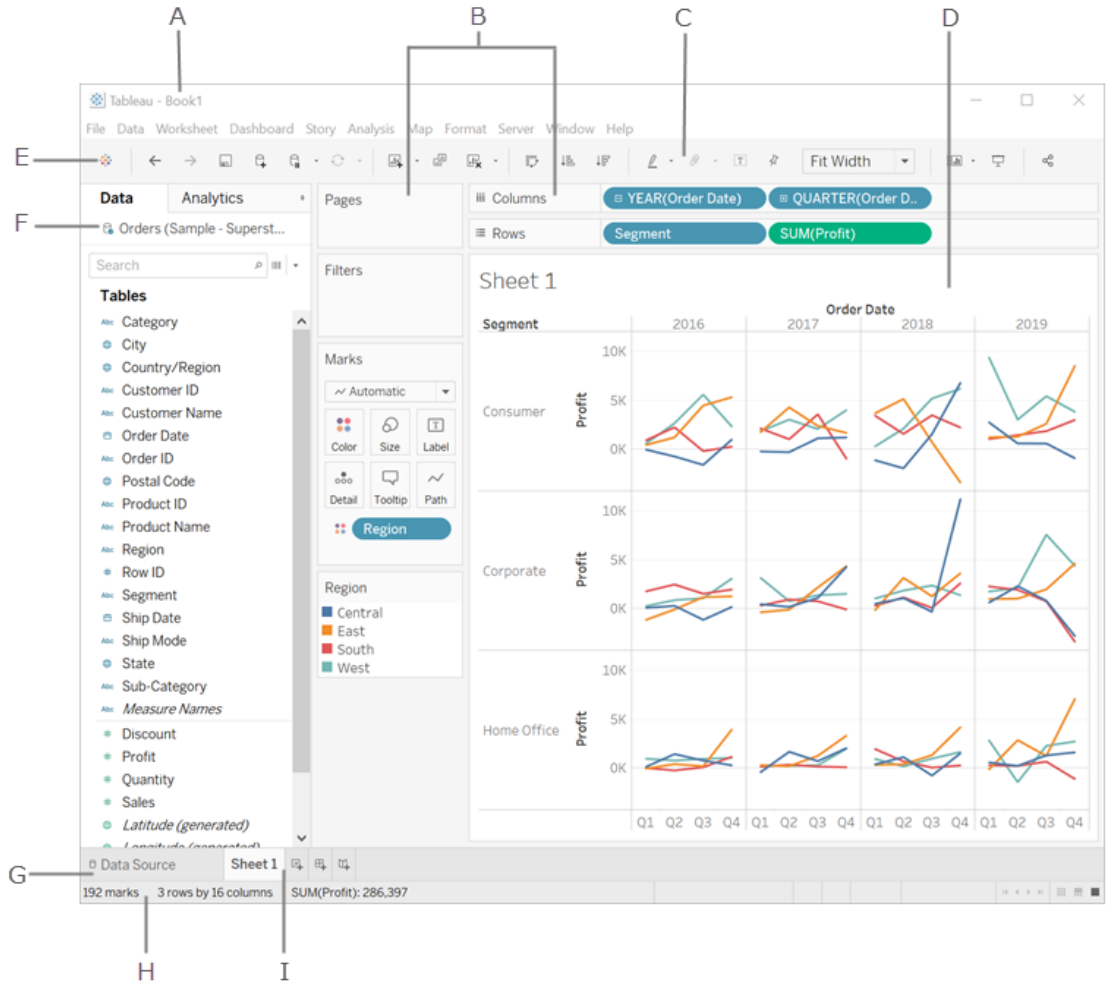
F. Side Bar- Contains the Data pane and the Analytics pane.

G. Data Source page - view and manipulate your data.

H. Status bar - Displays information about the current view.

I. Sheet tabs - Tabs represent each sheet (worksheets, dashboards, and stories) in your workbook.

Workspace area



Source: https://help.tableau.com/current/pro/desktop/en-us/environment_workspace.htm

Tableau File Extension

File Type	File Extension	Purpose
Tableau Workbook	.twb	It contains information on each sheet and dashboard that is present in a workbook. It has the details of the fields which are used in each view and the formula applied to the aggregation of the measures. It also has the formatting and styles applied. It also contains the data source connection information and any metadata information created for that connection.
Tableau Packaged Workbook	.twbx	This file format contains the details of workbook as well as the local data that is used in the analysis. Its purpose is to be share with other Tableau desktop or Tableau reader users assuming it does not need data from the server.
Tableau Data source	.tds	The details of the connection used to create the tableau report are stored in this file. In the connection details it stores the source type(excel/relational/sap etc.) as well as the data types of the columns.
Tableau Packaged Data source	.tdsx	This file is similar to the .tds file with the addition of data along with the connection details.
Tableau Data Extract	.tde	This file contains the data used in a .twb file in a highly compressed columnar data format. This helps in storage optimization. It also saves the aggregated calculations that are applied in the analysis. This file should be refreshed to get the updated data form the source.

Tableau Speak - Join, Union, Pivot, Relationships & Functions

Union vs Join

Union

Month	Price	Quantity
Jan	5	10

+

Month	Price	Quantity
Jan	10	20

=

Month	Price	Quantity
Jan	5	10
Jan	10	20

Adding rows to existing data

Join

Month	Price	Quantity
Jan	5	10

+

Month	Location
Jan	Bedok

=

Month	Price	Quantity	Location
Jan	5	10	Bedok

Adding Columns to existing data

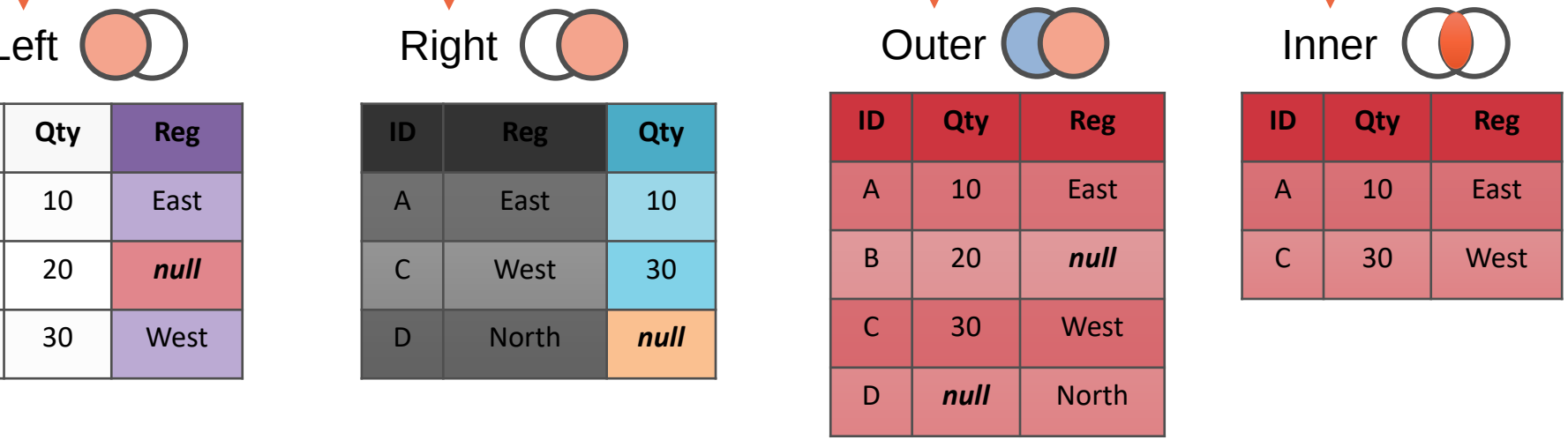
Tableau Speak – Join Types

Primary Table

ID	Qty
A	10
B	20
C	30

Secondary Table

ID	Reg
A	East
C	West
D	North



ID	Qty	Reg
A	10	East
B	20	*null*
C	30	West

ID	Reg	Qty
A	East	10
C	West	30
D	North	*null*

ID	Qty	Reg
A	10	East
B	20	*null*
C	30	West
D	*null*	North

ID	Qty	Reg
A	10	East
C	30	West

Primary = “Master”
Secondary = “Slave”

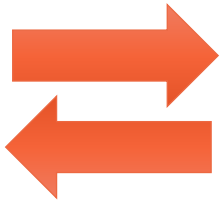
Primary = “Slave”
Secondary = “Master”

Equivalent to
Logical “OR”

Equivalent to
Logical “AND”

Tableau Speak – Pivot Data & Data Source Grouping

Year	0-4	5-9	10-14
1959	100	200	300
1960	400	0	500
1961	700	100	200



Year	Age Group	Population
1959	0-4	100
1960	0-4	400
1961	0-4	700
1959	5-9	200
1960	5-9	0
1961	5-9	100
1959	10-14	300
1960	10-14	500
1961	10-14	200

Pivoting reduces the number of columns, and increases the number of rows

Sometimes, by grouping data, the data structure and usability improves significantly.

Caveats:

1. Same Datasource
2. Same Connection
3. Excel, Textfile, Google Sheets, .pdf

Tableau Speak – Relationships & Data Blending

Single Data Source

Primary Table

ID	Qty
A	10
B	20

Secondary Table

ID	Reg
A	East
B	West

JOIN

Data Source 1

ID	Qty
A	10
B	20

Data Source 2

ID	Reg
A	East
B	West

RELATIONSHIP

Only Dimensions, not Measures

Tableau Speak – Building Blocks & Syntax for Data Source Manipulation with Calculated Fields

Roll	Student	Grade
1	Ali	80
2	Benny	40
3	Carrie	60
4	Daffy	50

```
IF [GRADE] > 50
THEN “PASS“
ELSE “FAIL“
END
```

Calculated
PAS
FAIL
PASS
FAIL

Functions

Fields

Operators

Literal Expressions

Number
ABS(-3) = 3

Strings
Len(“Ali”) = 3

Type Conversion
INT(50.1) = 50

Logical
IF [Ali] THEN “Pass”

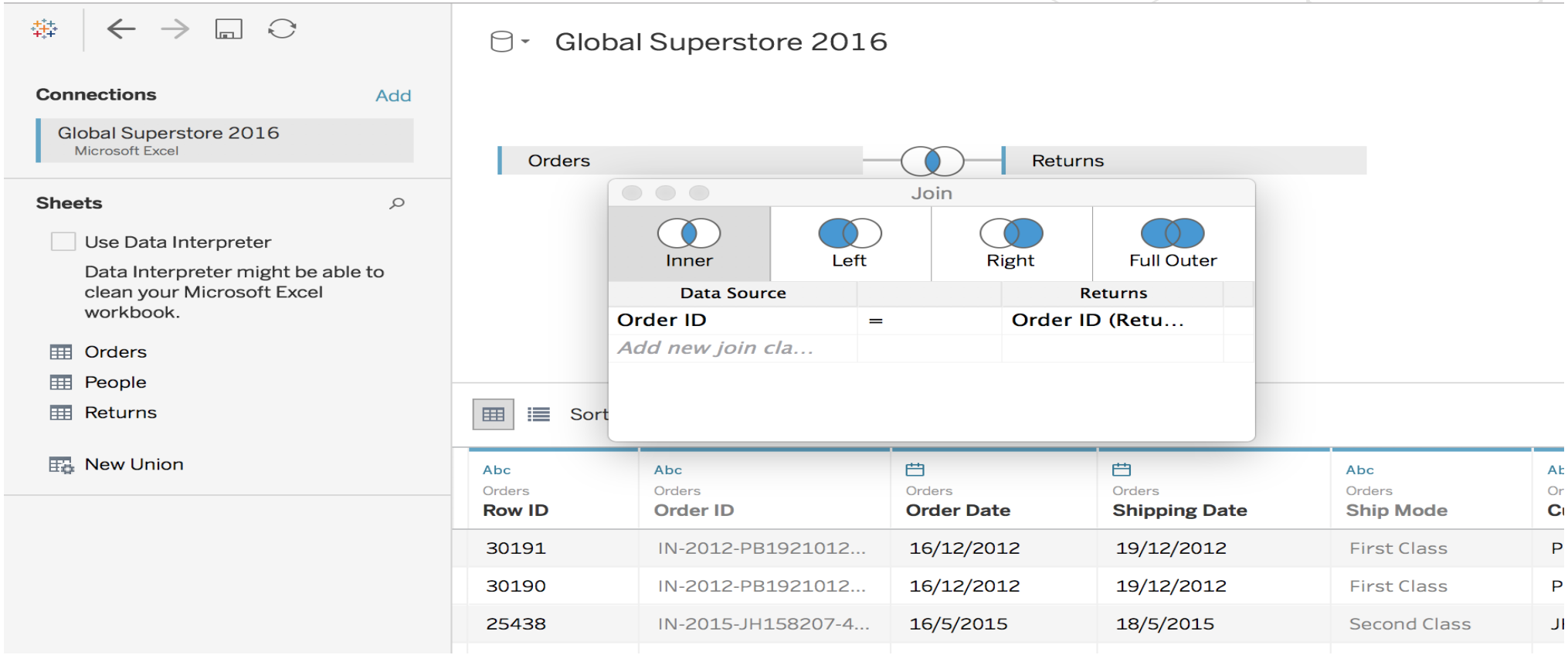
[GRADE]

>

“PASS”

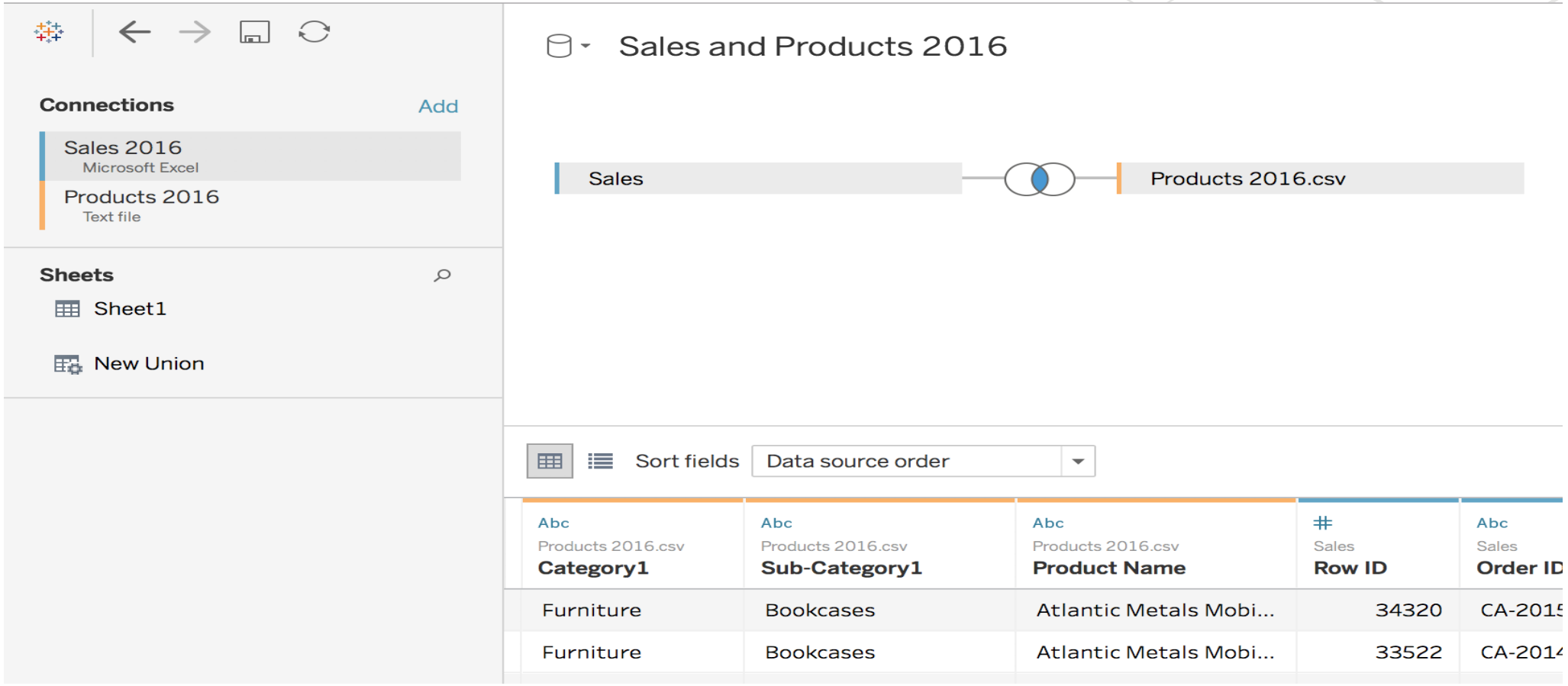
Source: https://help.tableau.com/current/pro/desktop/en-us/functions.htm

Tableau Speak – Table Join



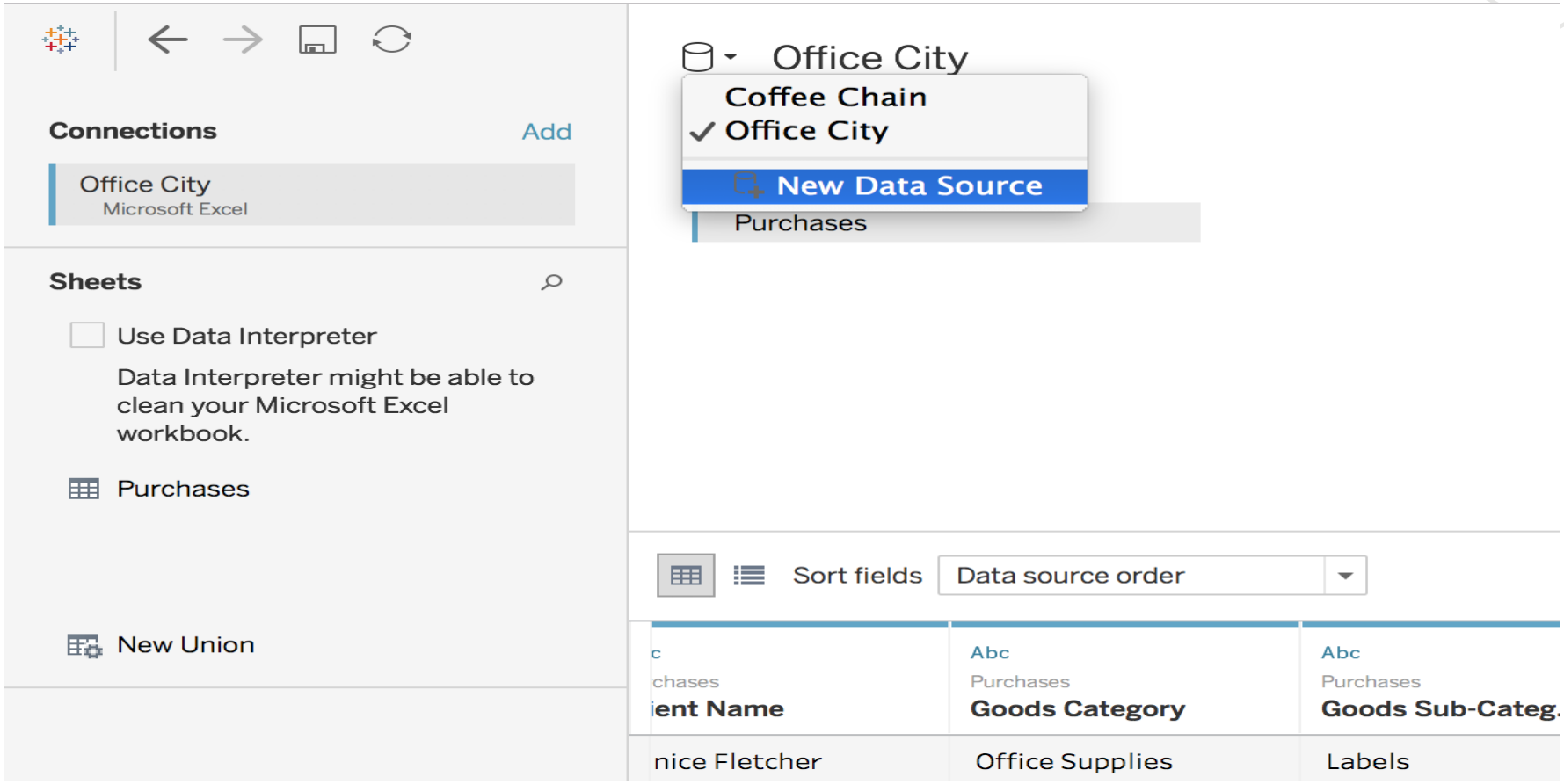
the <no code> course

Tableau Speak – Cross-database Join



the <nocode> course

Tableau Speak – Data Blending



the <no code> course

Tableau Speak – Pivot Data from Columns to Rows (or v.v.)

Pivot from wide to long format

Employee	1/1/15	1/2/15	1/3/15	1/4/15	1/5/15	1/6/15	1/7/15	1/8/15	1/9/15	1/10/15	1/11/15	1/12/15
B-002	4	1	5	2	3	0	3	1	2	0	2	5
E-055	1	2	1	3	4	1	4	0	2	1	4	0
E-075	14	17	16	15	18	16	14	17	12	13	14	12
B-066	4	4	5	2	5	0	0	2	0	1	0	3
C-025	17	13	17	18	17	17	12	15	17	17	14	15
E-030	2	2	1	1	0	3	5	5	0	2	4	1
C-001	14	14	14	14	13	18	17	14	13	18	15	14
E-038	4	1	0	4	0	2	5	0	2	2	2	2
C-054	2	5	4	4	2	3	0	5	5	5	3	5
A-081	3	2	4	5	2	2	2	4	1	4	2	0
B-031	14	14	14	14	15	13	15	14	12	16	12	18
D-019	2	3	0	0	4	4	1	2	5	0	5	5
E-096	2	0	4	4	5	3	3	0	5	4	2	0
D-026	0	2	0	2	5	3	1	0	0	2	5	4
E-022	3	3	4	3	4	2	0	3	2	3	3	1
C-015	1	5	3	5	2	1	3	3	1	1	5	2
B-062	14	12	16	16	16	18	12	12	18	16	12	17
E-029	5	1	2	4	0	3	5	4	5	3	4	5
A-037	2	2	0	2	3	4	2	0	2	1	2	2
E-087	14	17	13	17	18	13	13	12	13	16	13	16
C-040	5	0	4	5	3	5	2	1	1	4	2	1
A-077	3	5	3	5	2	3	5	3	4	5	4	4
C-041	18	18	15	15	15	17	17	12	18	17	16	16
D-005	4	0	5	3	2	3	5	1	0	2	3	3
E-046	12	13	14	17	16	14	14	18	13	16	15	17
C-053	2	5	5	0	2	1	0	4	5	1	1	3

Pivot Date	Pivot Resolved Incidents	Resolved Incidents Employee
1/1/2015	4	B-002
1/1/2015	1	E-055
1/1/2015	14	E-075-II
1/1/2015	4	B-066
1/1/2015	17	C-025-II
1/1/2015	2	E-030
1/1/2015	14	C-001-II
1/1/2015	4	E-038
1/1/2015	2	C-054
1/1/2015	3	A-081
1/1/2015	14	B-031-II

Pivot from long to wide format

Tableau Speak – Split

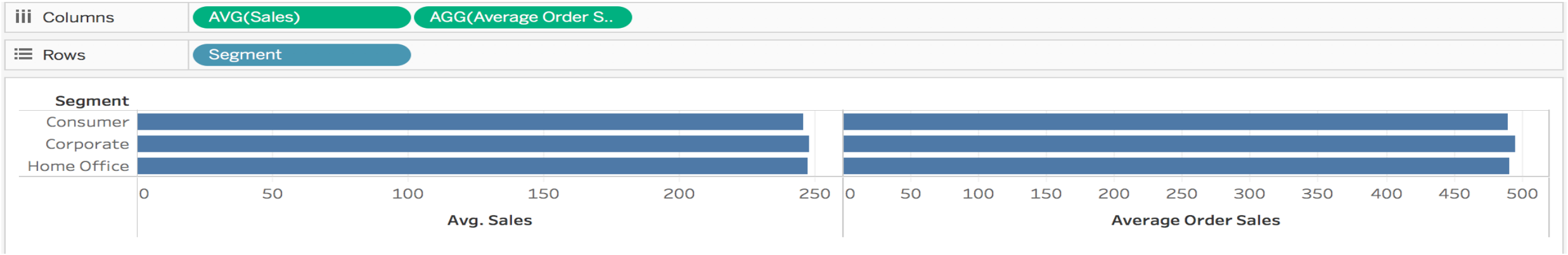
Split “Employee” column:

Abc Resolved Incidents **Employee**	=Abc Calculation **Location**	=Abc Calculation **Employee ID**	=Abc Calculation **Tier**
B-002	B	002	
E-055	E	055	
E-075-II	E	075	II
B-066	B	066	
C-025-II	C	025	II
E-030	E	030	
C-001-II	C	001	II
E-038	E	038	
C-054	C	054	
A-081	A	081	

Tableau Speak – Calculation: Aggregate VS Record-Level

Average Order Sales

```
SUM([Sales])/COUNTD([Order ID])
```



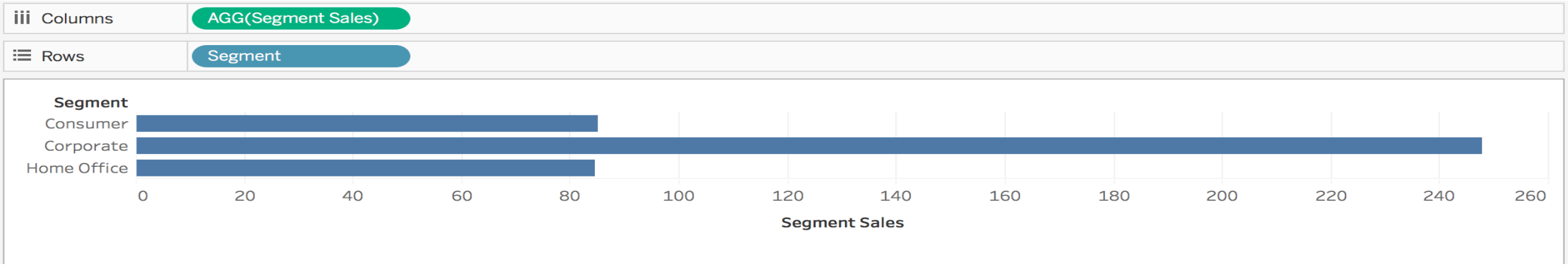
the <nocode> course

Tableau Speak –Aggregate Functions

1. Aggregation of a measure
2. Aggregation of a dimension

Segment Sales

```
IF ATTR([Segment]) = "Corporate"
THEN AVG([Sales])
ELSE MEDIAN([Sales])
END
```



the <nocode> course

Lets

with

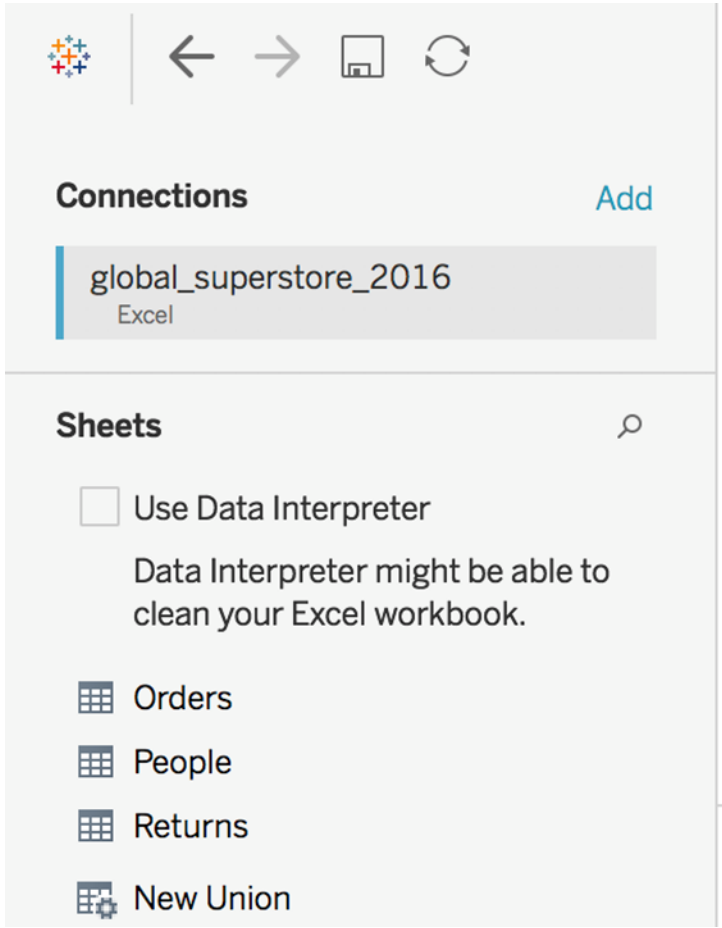
- Ensure that you have the following datasets downloaded onto your computer:
    1. global_superstore_2016.xlsx
    2. Sales 2016.xlsx
    3. Products 2016.csv
    4. Coffee Chain.xlsx
    5. Office City.xlsx

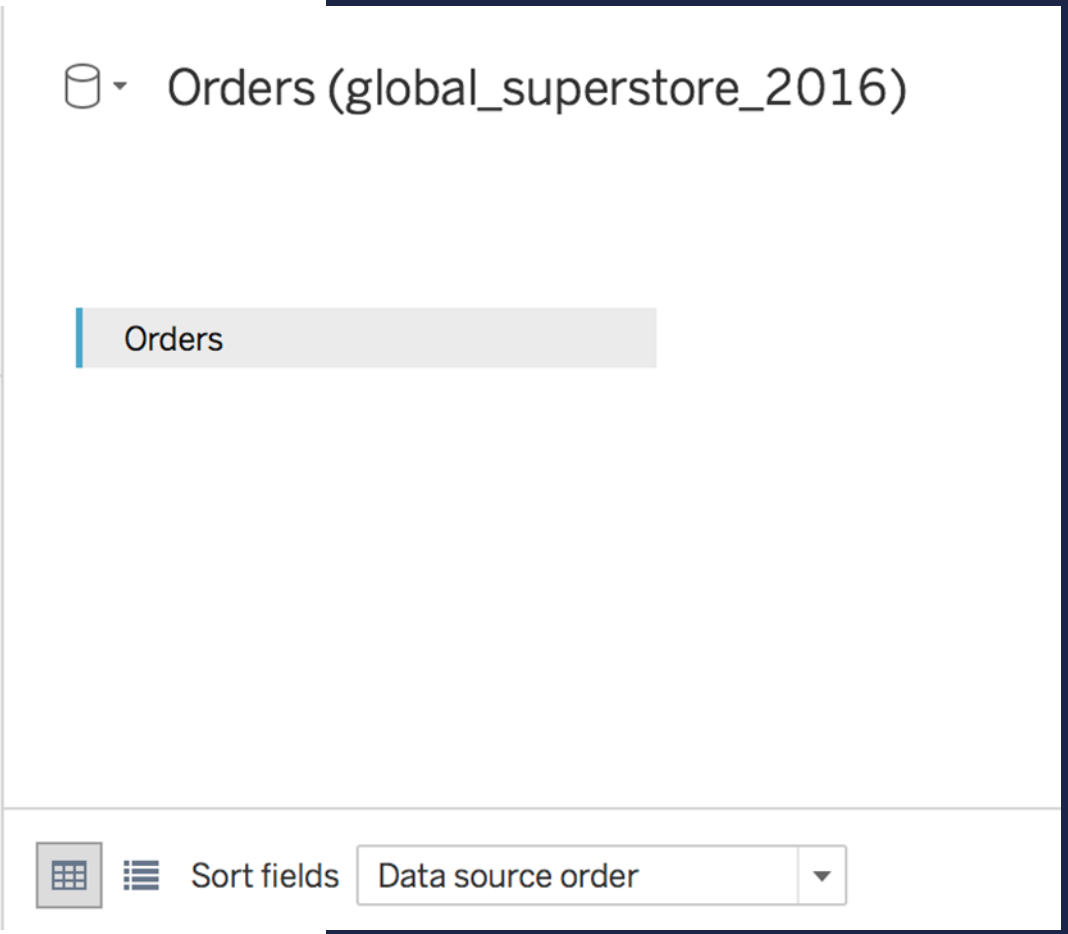
tableau

the <nocode> course

Tableau Class Activity

- Import Excel data: global_superstore_2016.xlsx (orders)
- Import office city.xlsx and coffee chain.xlsx
- Build a simple worksheet/ report

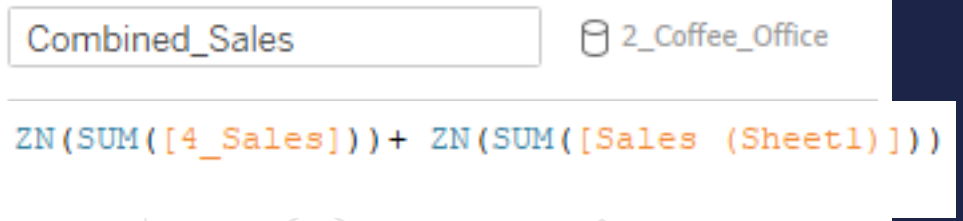




the <nocode> course

Tableau Class Activity

- Import Excel data: office city.xlsx and coffee chain.xlsx
- Create a calculated field:
  - For each state, calculate combined sales
  - Percent of Total



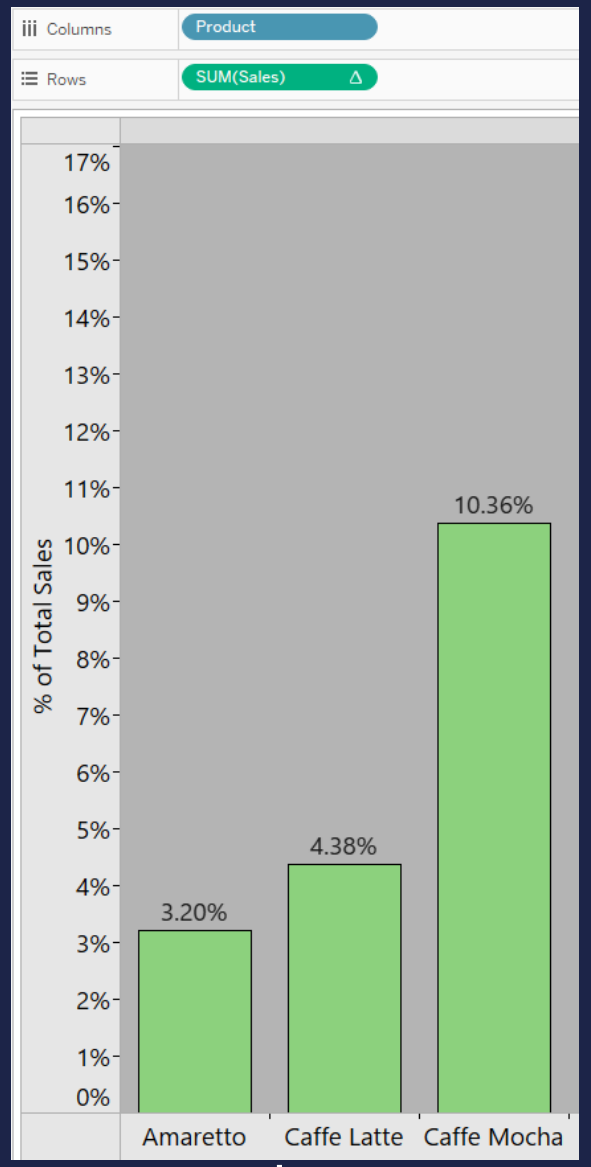
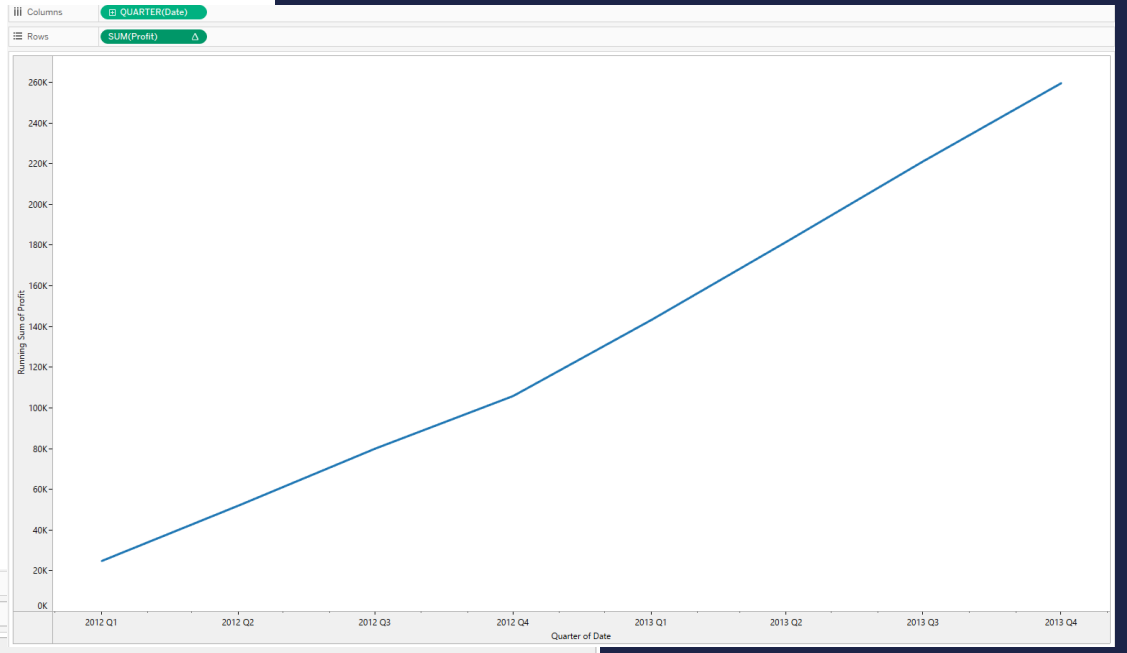
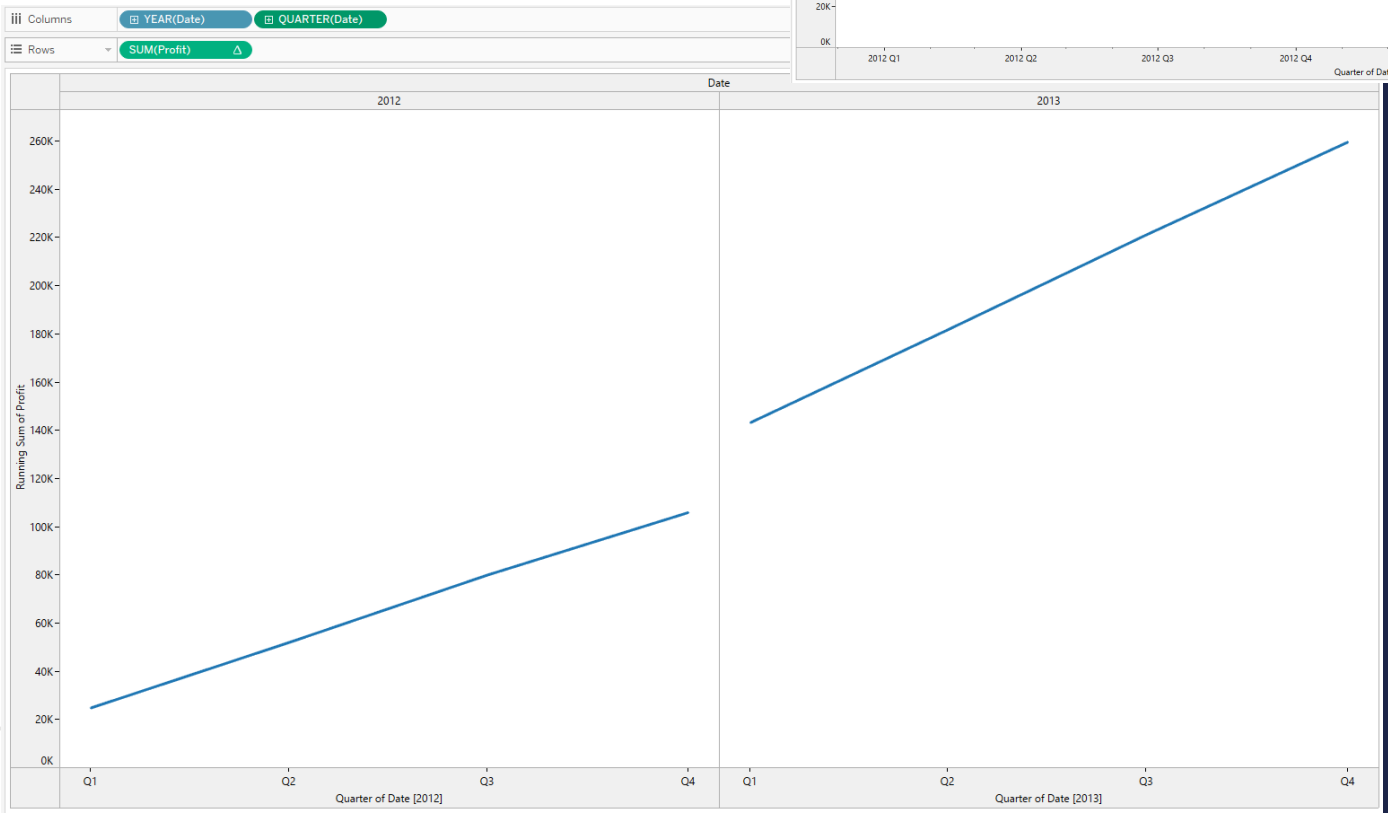


Tableau Class Activity

- Using coffee chain.xlsx
- Create a calculated field:
  - Running Total
  - Running total for each year (Is the chart correct?)

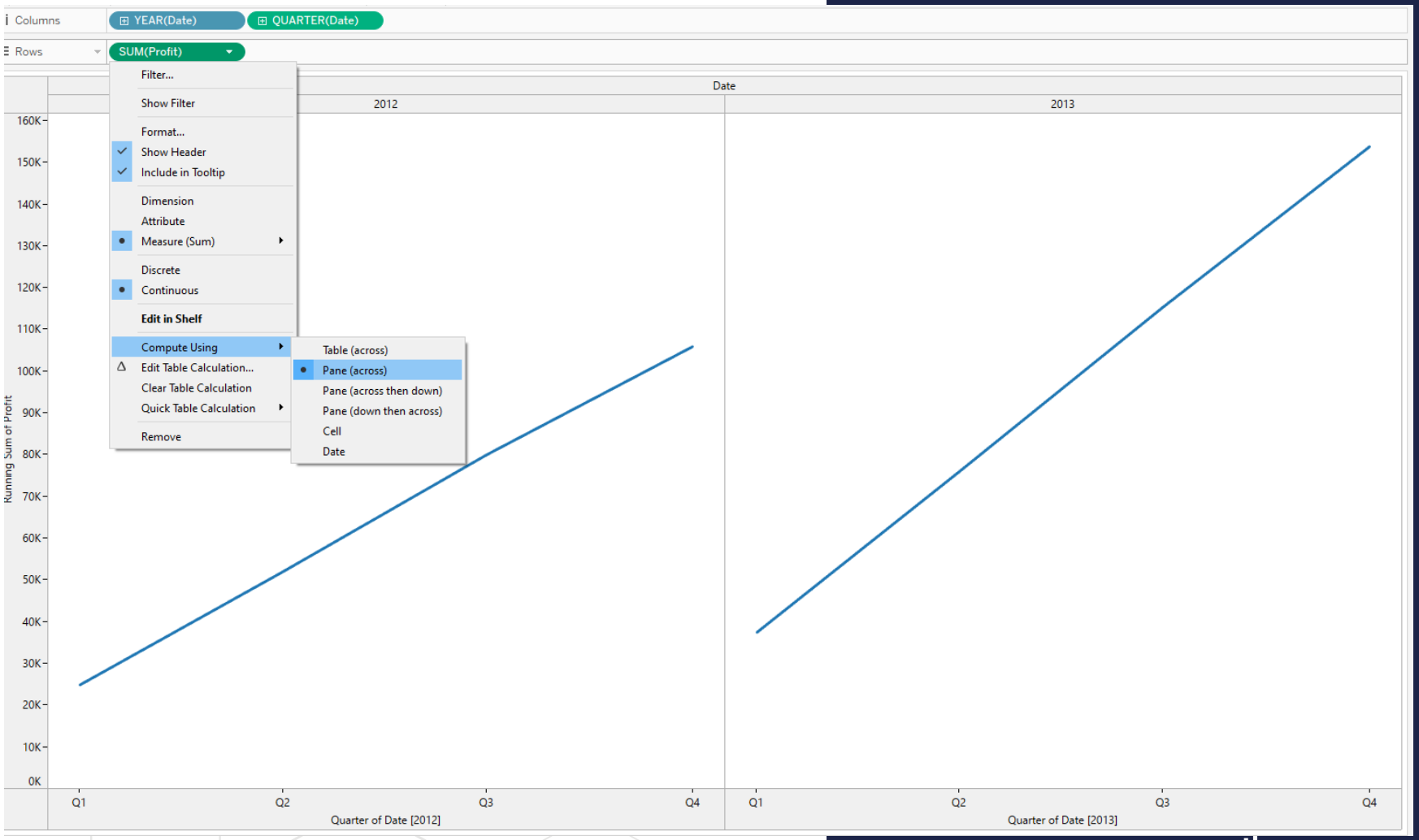




the <no code> course

Tableau Class Activity

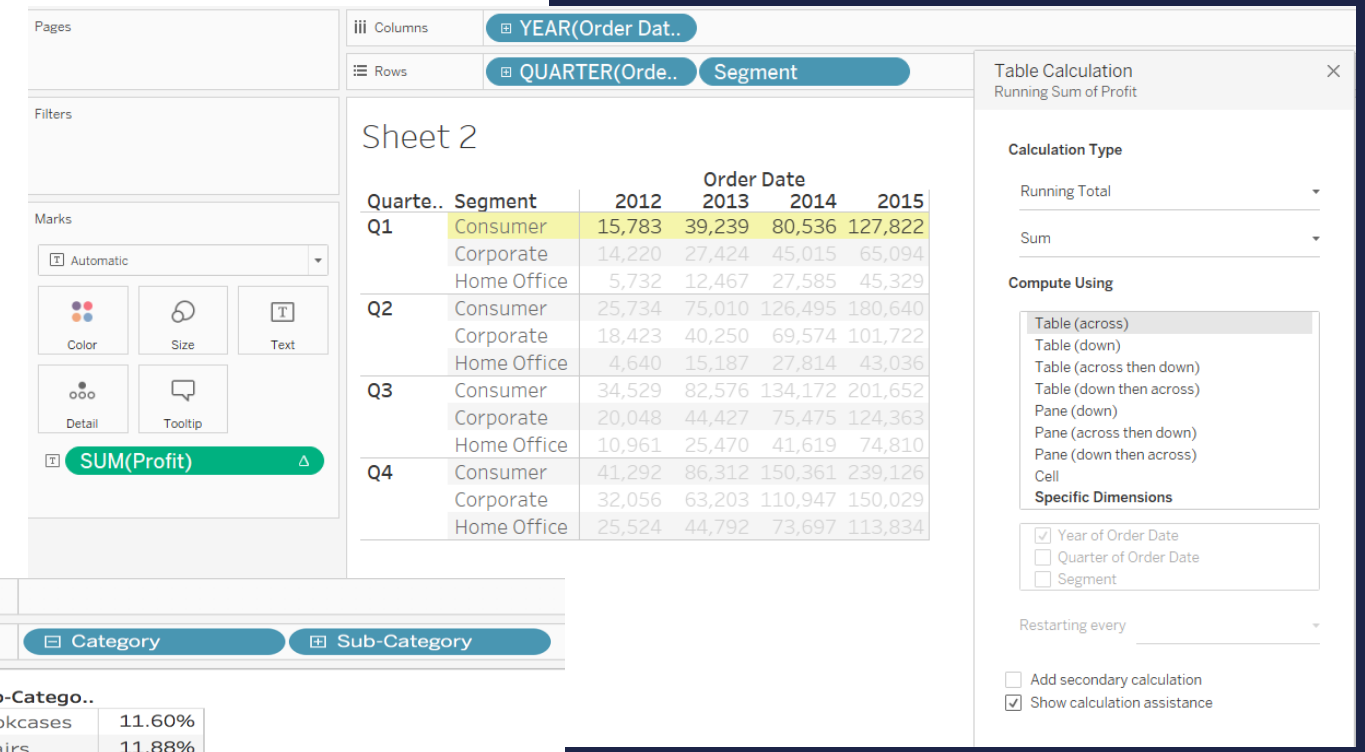
- Using coffee chain.xlsx
- Create a calculated field:
  - Running total for each year (done correctly)

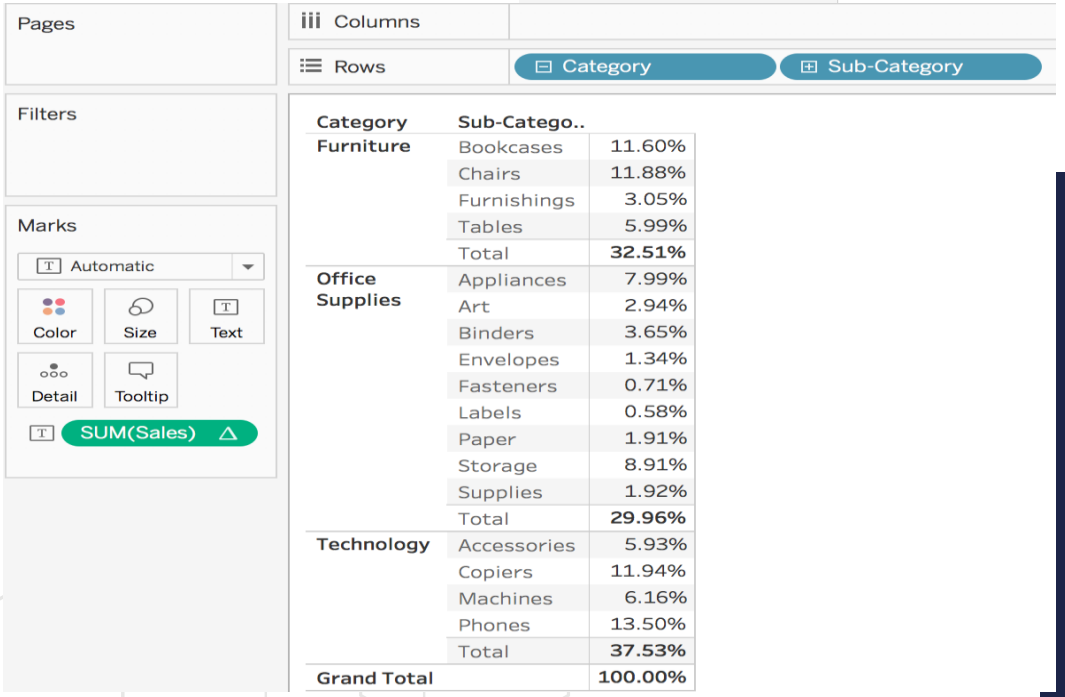


the <no code> course

Tableau Class Activity

- Using global_superstore_2016.xlsx
- Create some Quick Table Calculations:
  - Compare “Compute Using”





the <nocode> course

5 mins

Part 4 – Basic Data Visualisation

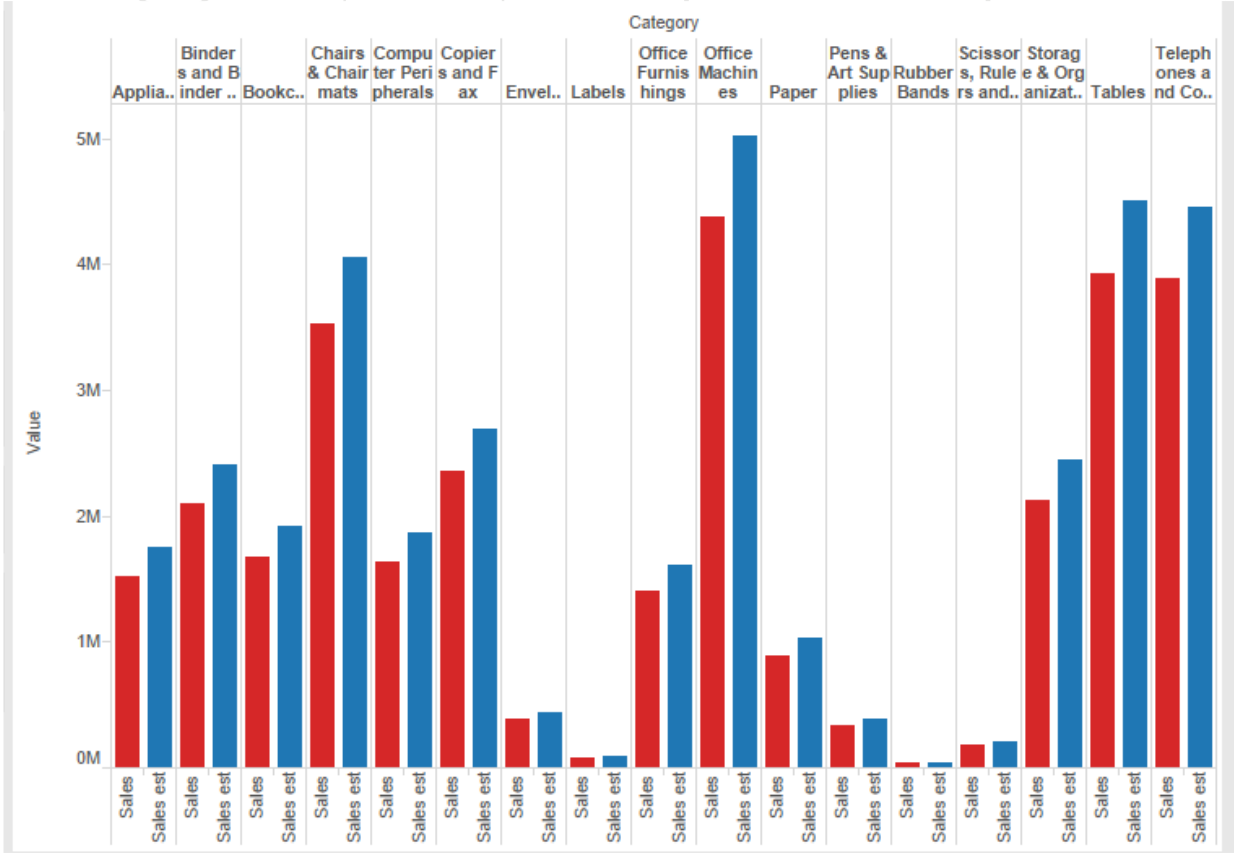
Visualisation of Categorical Data

- Bar / Stacked Bar Chart
- Side-by-side Bar Chart
- Pie Chart
- Area Chart
- Heat Map
- Tree Map

Visualisation of Categorical Data

Side-by-side (i.e., grouped) bar chart

1. Drag two or more measures into either the worksheet's columns or rows.
2. Measure Values and Measure Names will then be created automatically.
   - Measure Names and Measure Values can be used to express various measures in a data set or multiple measures on a single axis.
   - Measure Values contains the data while Measure Names is used to separate the bars used on the marks



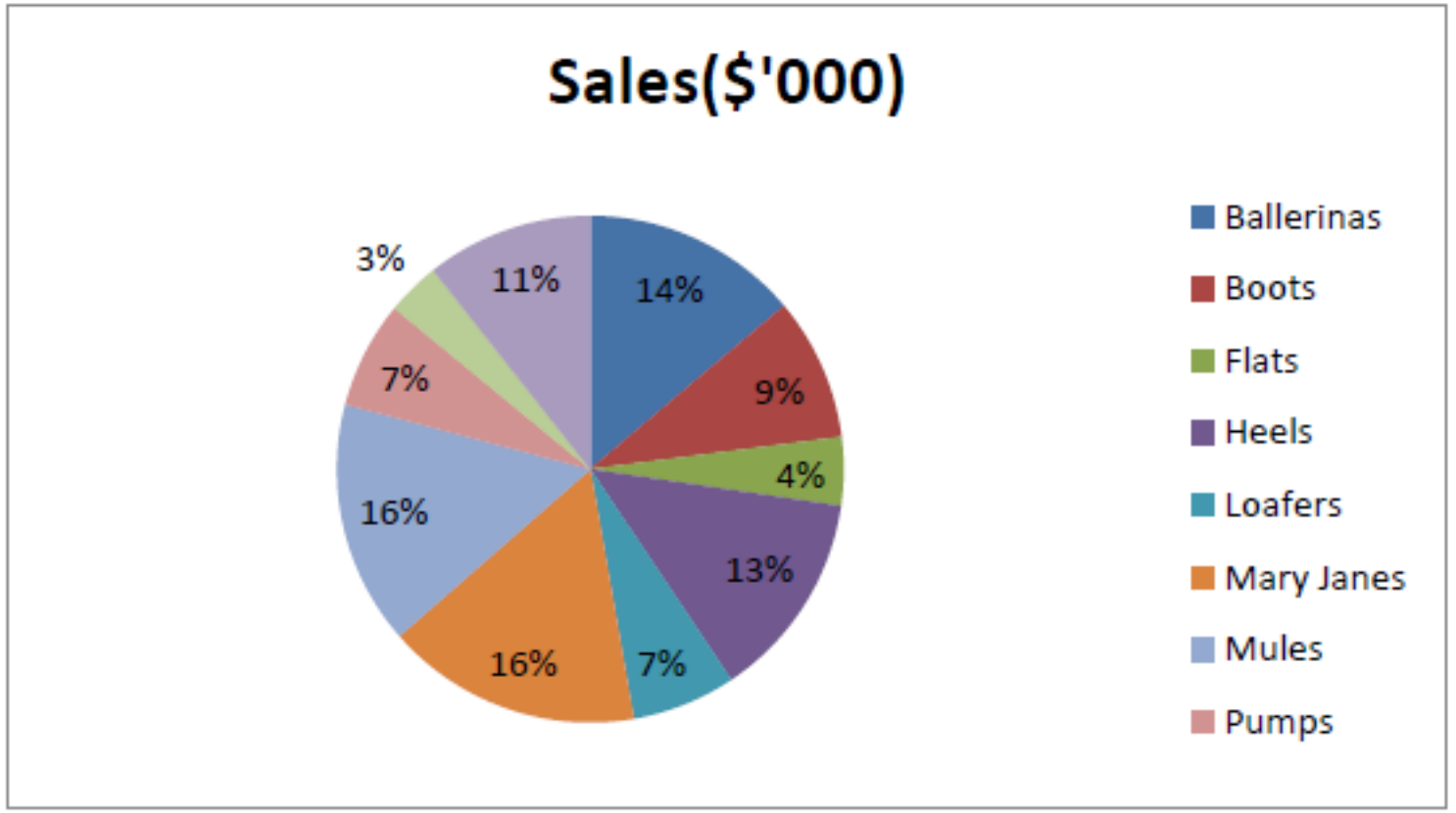
the <no code> course

Visualisation of Categorical Data

Pie chart

1. Pie Chart should be used to get a general sense of magnitude, but not for precise comparisons.

2. A pie chart is limited by its area meaning that if you have many categories, you have many slivers and its hard to compare.

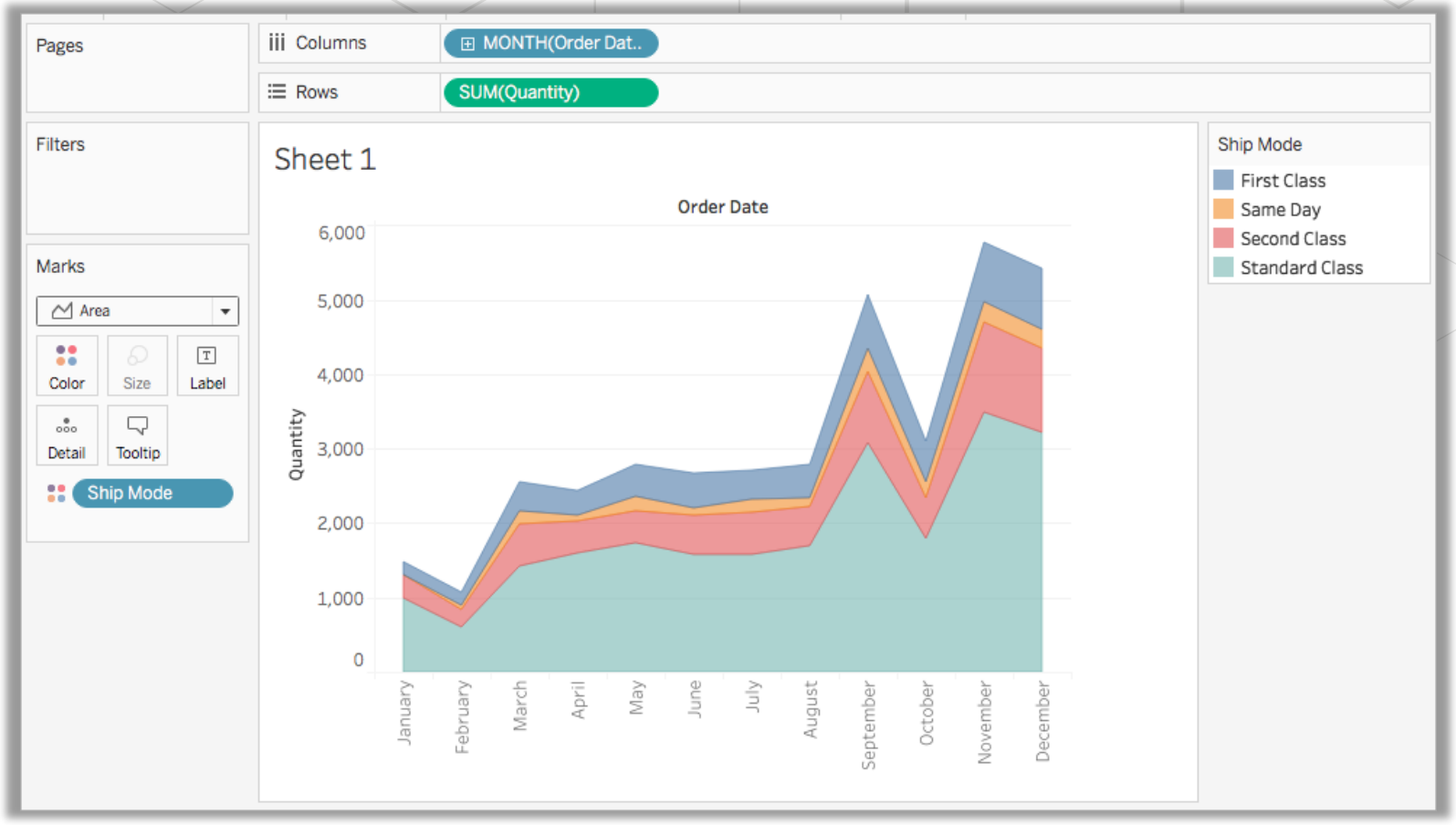


Visualisation of Categorical Data

Area chart

1. Area Fill chart plots values as bands, thus it is easy to misinterpret the top band as being the largest value in a data set.

2. The Area Fill Chart is best used to plot a single dimension to avoid such misinterpretation.

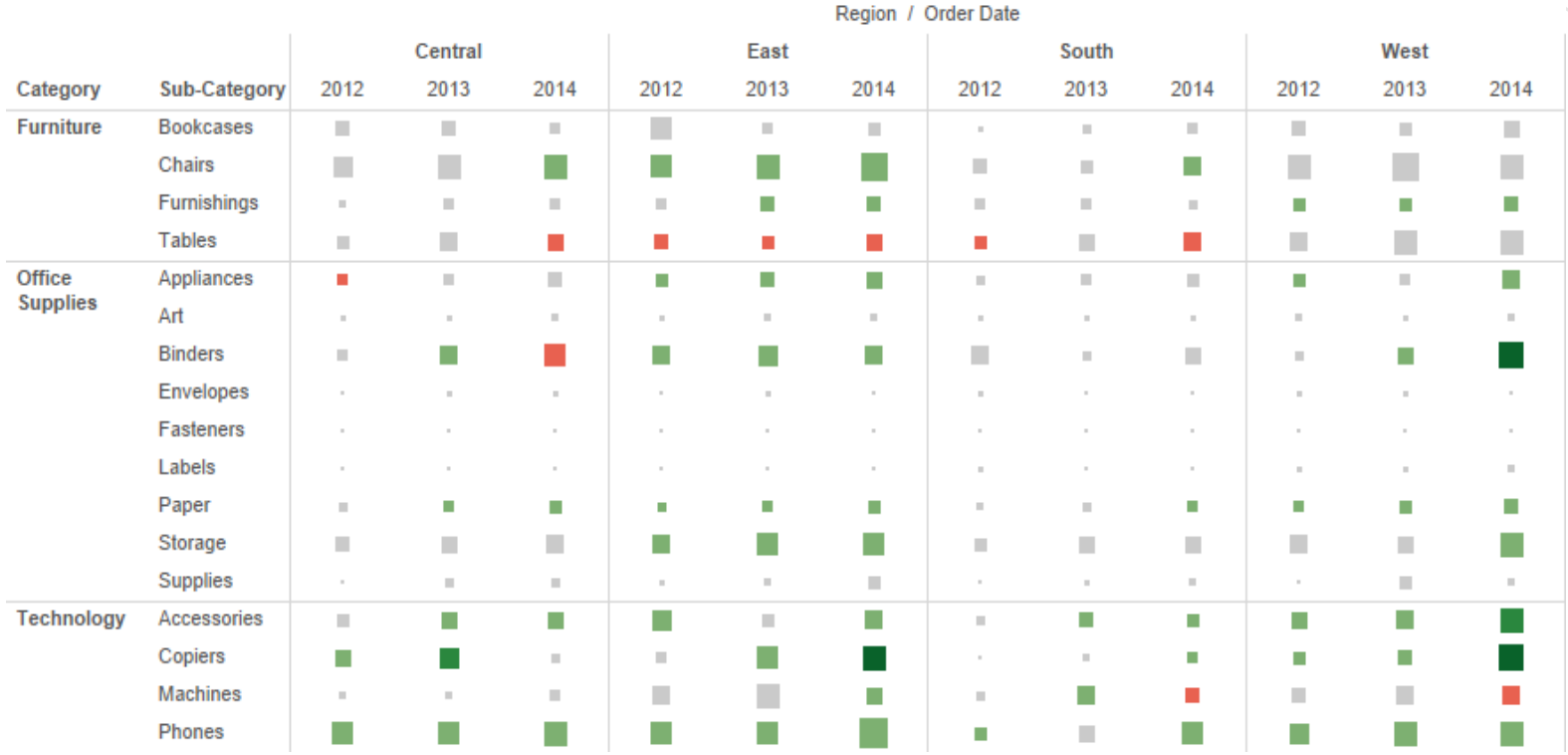


the <no code> course

Visualisation of Categorical Data

Heat maps

1. Heat Maps use colours and sizes to compare up to two measures.

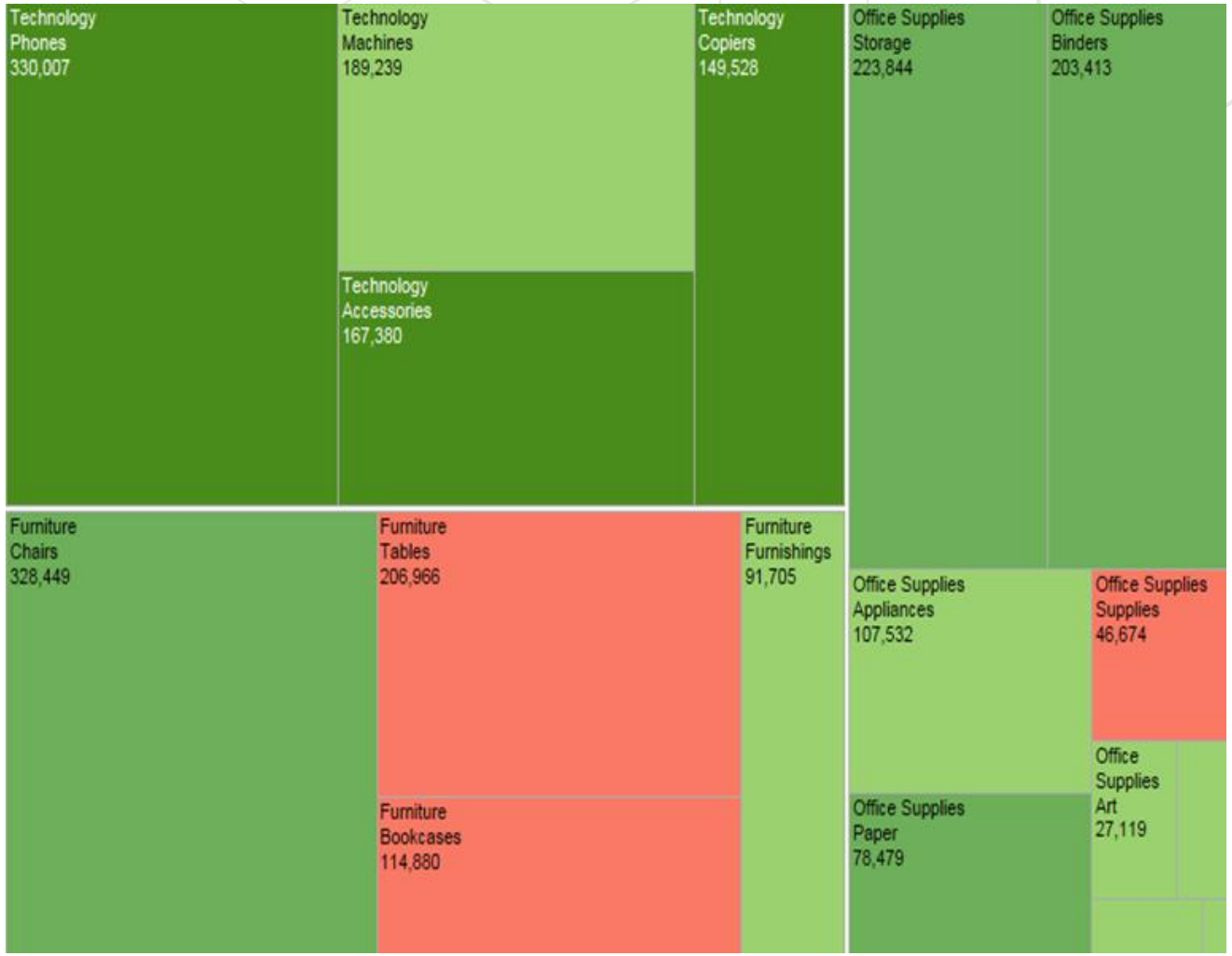


the <no code> course

Visualisation of Categorical Data

Tree maps

1. Tree maps effectively display larger dimension sets using colours and sizes to display one or more dimensions, and up to two measures.



the <no code> course

Visualisation of Time Series Data

Discrete (bucketed) time series data

- 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, and scatterplots can also be used to display time series data

The power of “Time Series Analysis” lies in looking for trends in temporal data and forecasting i.e., based on what I see in the past, can I predict the future? Is this pattern seasonal/cyclical?

A whole picture view is more important than the individual data points

Visualisation of Time Series Data

Continuous (unbroken) time series data

- Visualising continuous time series data is similar to visualising discrete time series data. We still have a discrete number of data points, even if the dataset is continuous
- **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

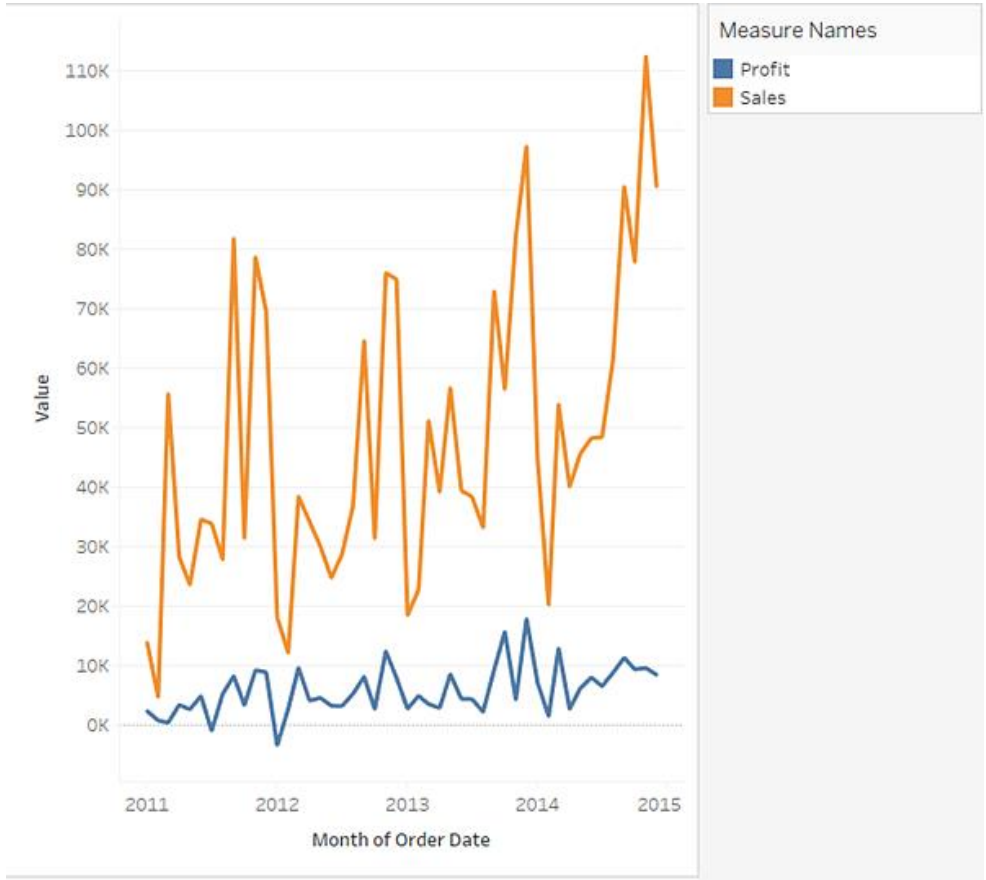
Visualising time series data

- Line Chart
- Gantt Chart
- Trendline
- Reference Line

Powerful Story Telling Tool – Crude Price & History

Visualisation of Time Series Data

Line Chart



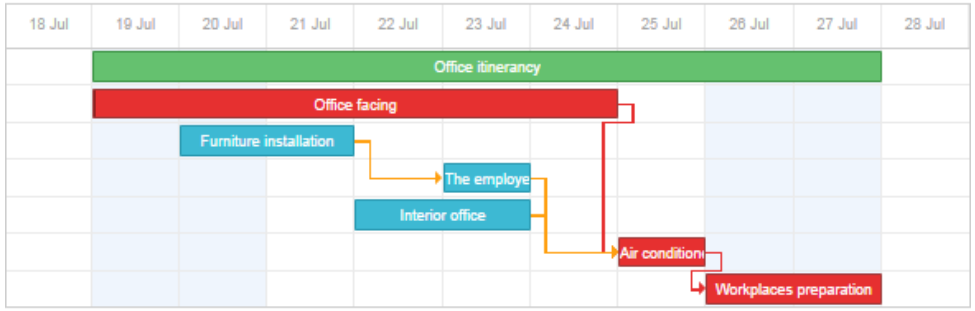
Useful for data that does not stay constant i.e., is variable.

Data that stays constant, then suddenly changes, then stays constant etc is better visualized with a “step” graph

Data that is TOO highly variable (noisy data) sometimes needs to be smoothed and/or a “trend line” can be applied to make sense of it

Visualisation of Time Series Data

Gantt Chart



Great project managing tool; has duration information and events/resources are visualized as bars. The physical position of the bars denotes start and end time; length is (obviously) duration.

Gantt charts also allow you to trace the “critical path”. Critical path analysis shows the sequence of scheduled tasks that determine the duration of a project. A critical path analysis identifies which tasks you must complete in order to meet your project deadline.

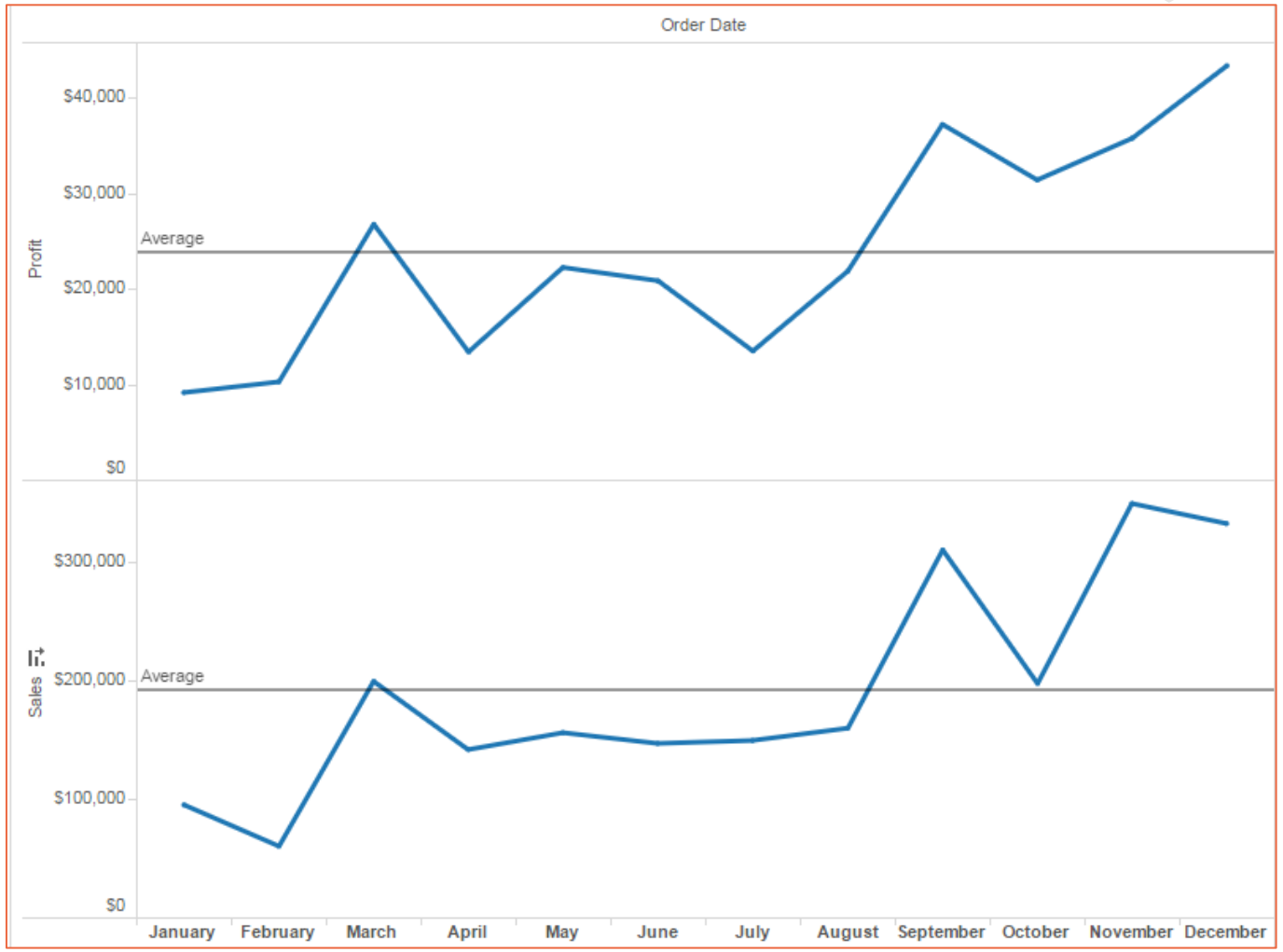
Visualisation of Time Series Data

Adding a trend line

the <no code> course

Visualisation of Time Series Data

Adding a reference line/band/distribution/boxes



the <no code> course

Lets

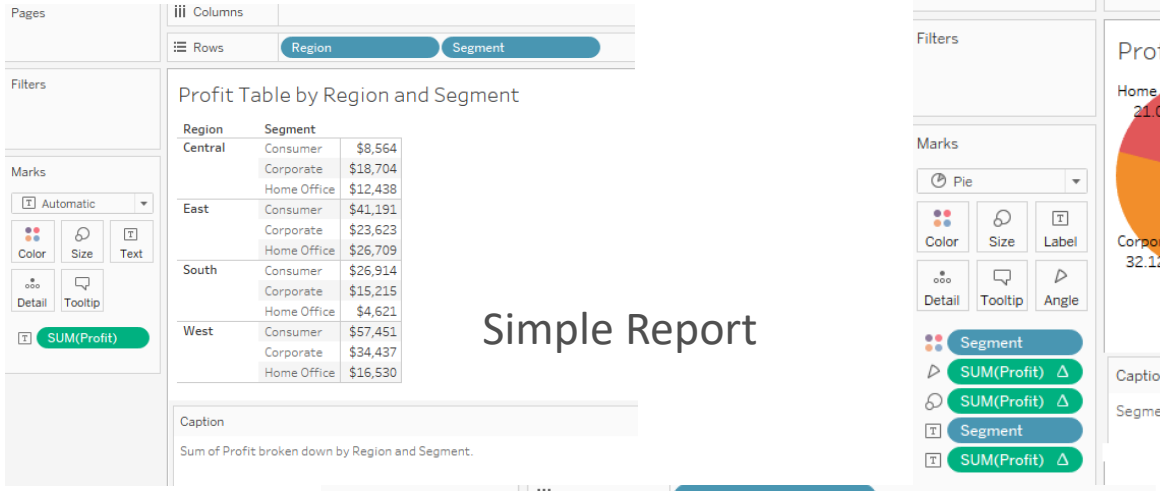
With

tableau

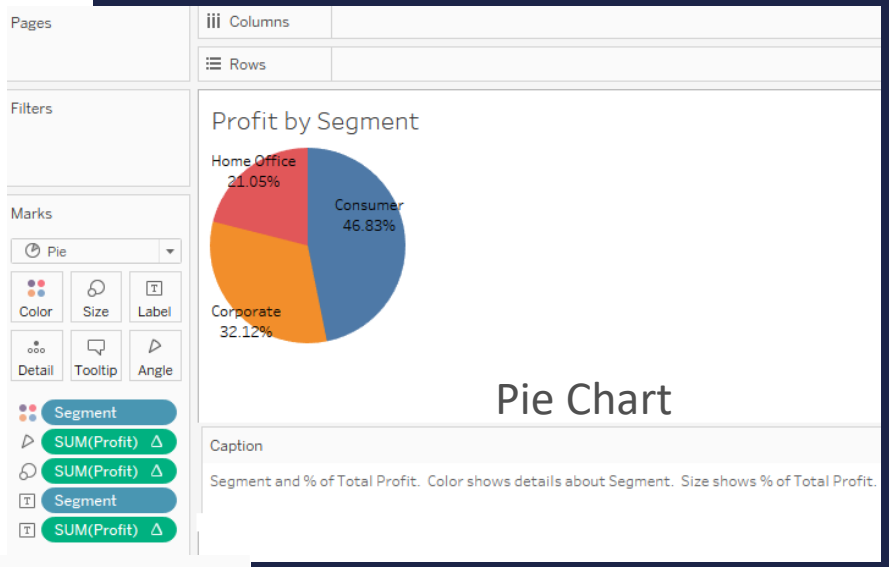
the <no code> course

Tableau Class Activity

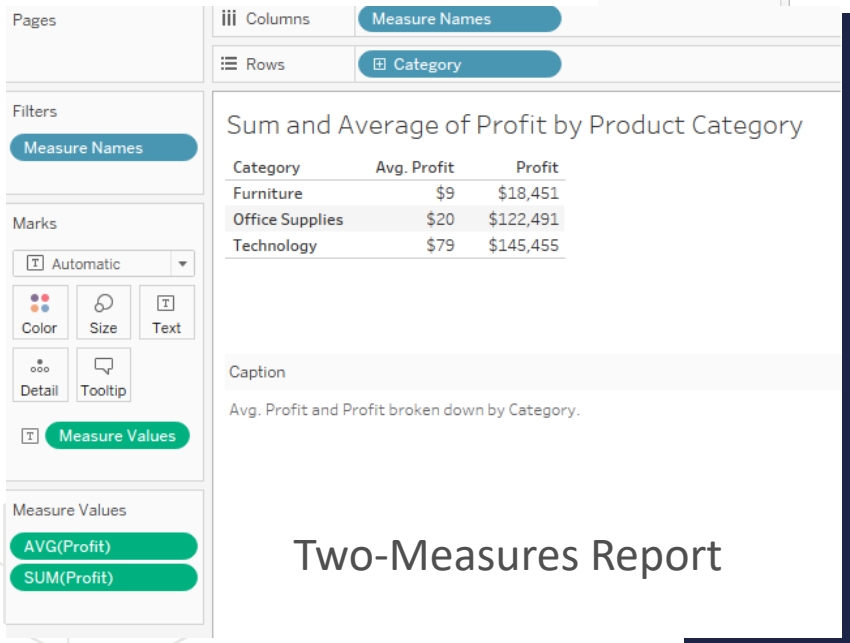
- Using Excel data: global_superstore_2016.xlsx (orders)
- Create the following
  - Simple report
  - Two measures report
  - Pie chart



Simple Report



Pie Chart

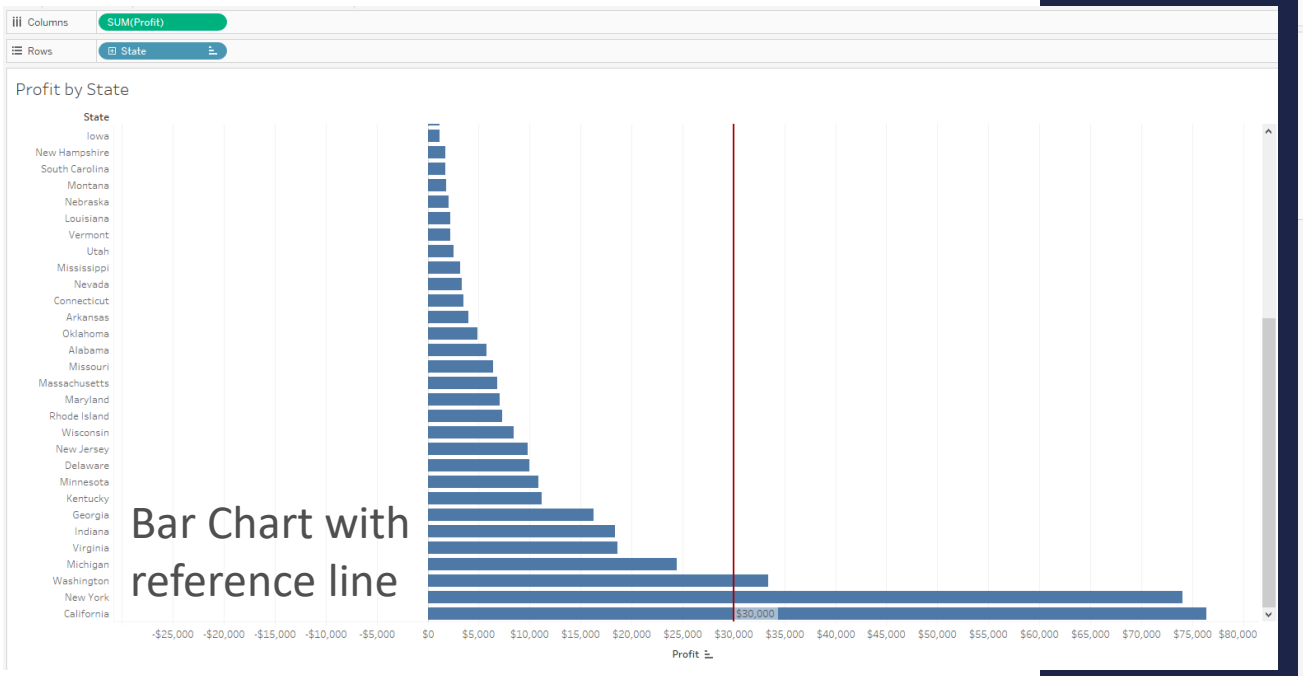


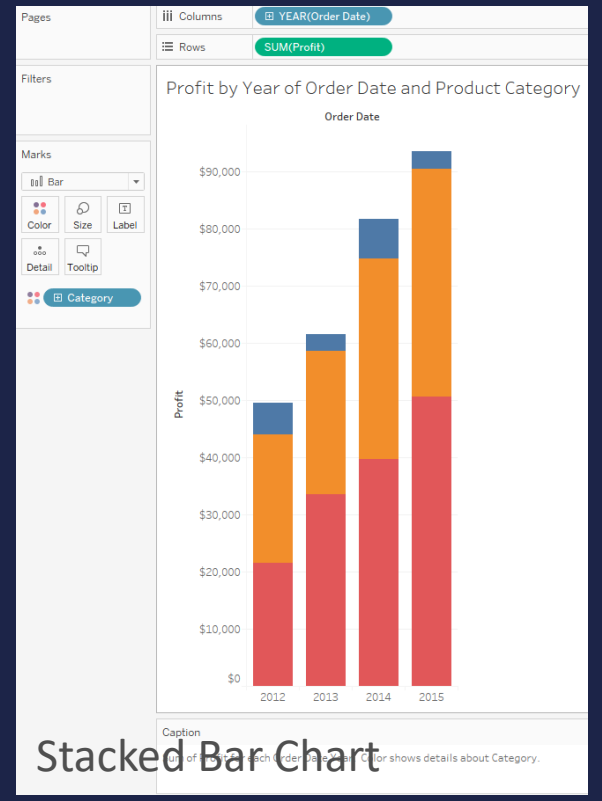
Two-Measures Report

the <nocode> course

Tableau Class Activity

- Using Excel data: global_superstore_2016.xlsx (orders)
- Create the following
  - Bar chart with reference line
  - Stacked bar chart

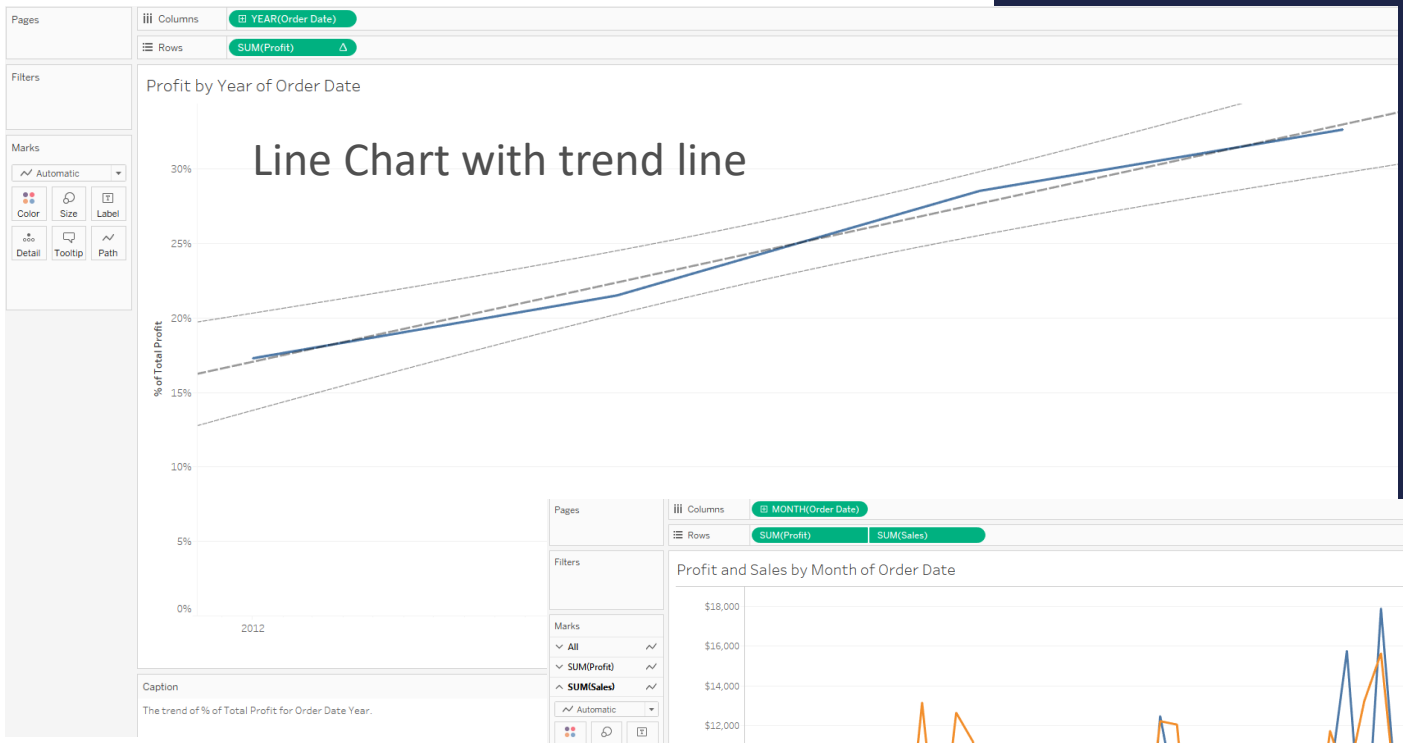




the <nocode> course

Tableau Class Activity

- Using Excel data: global_superstore_2016.xlsx (orders)
- Create the following
  - Line chart with trend line
  - Line chart with 2 axis



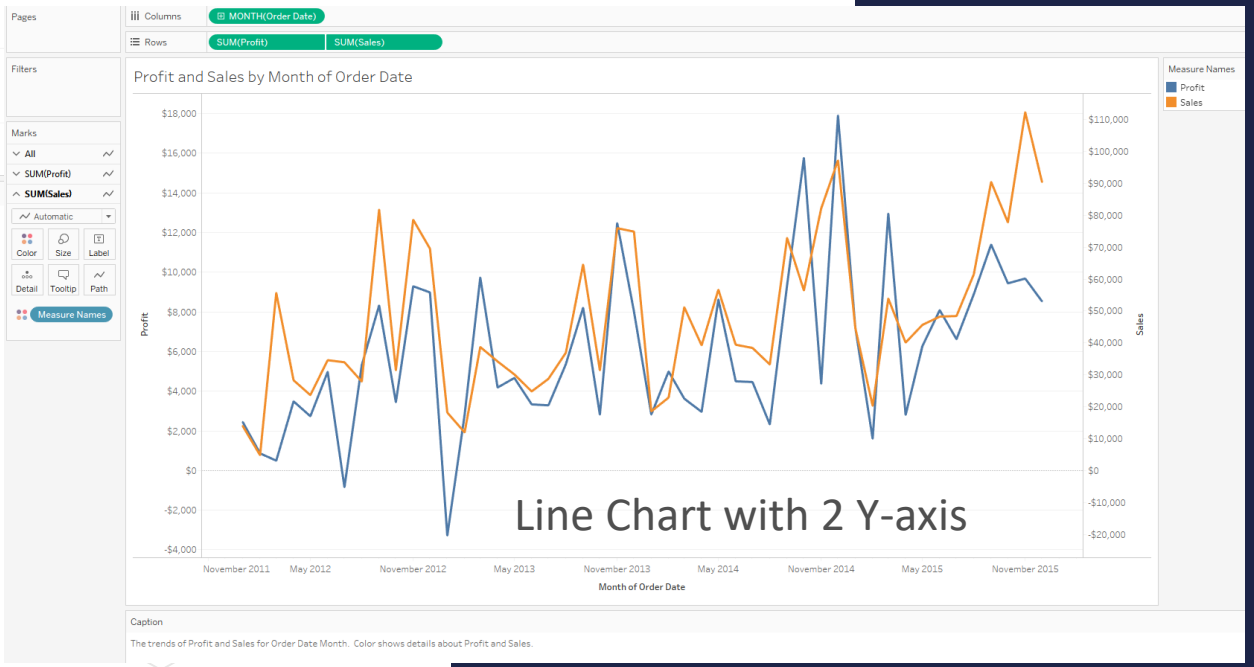
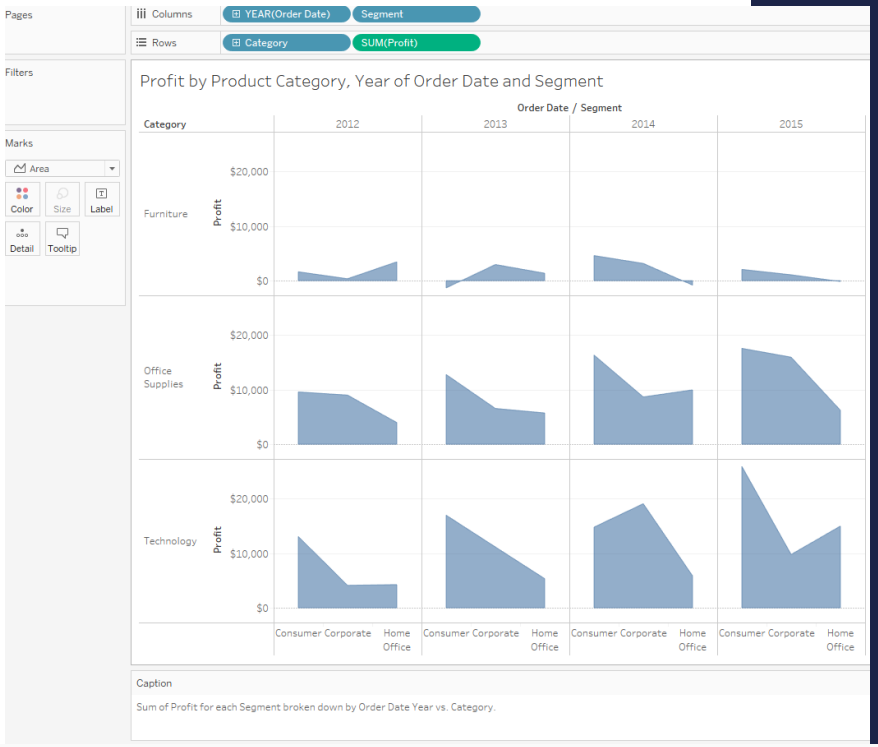
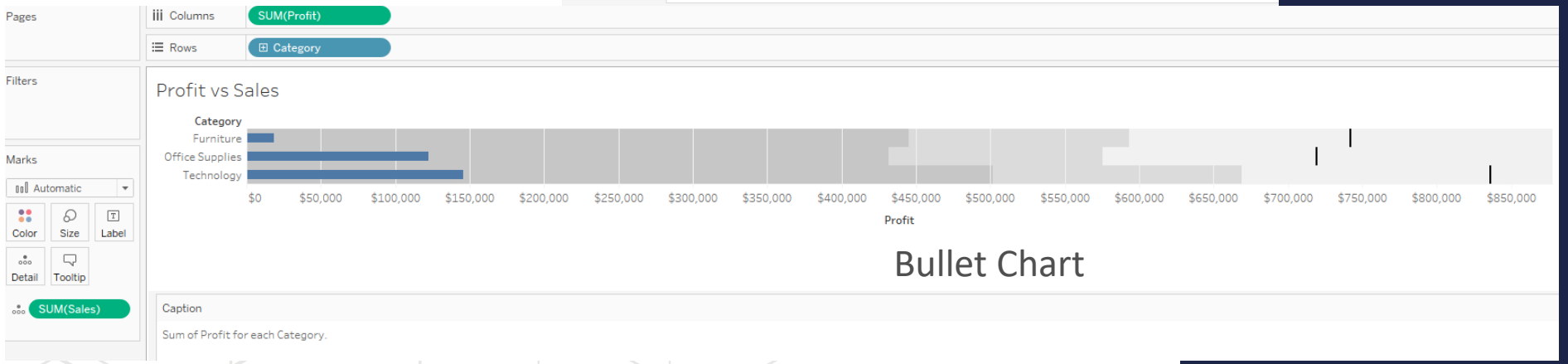


Tableau Class Activity

- Using Excel data: global_superstore_2016.xlsx (orders)
- Create the following
  - Area chart
  - Bullet chart

Area Chart





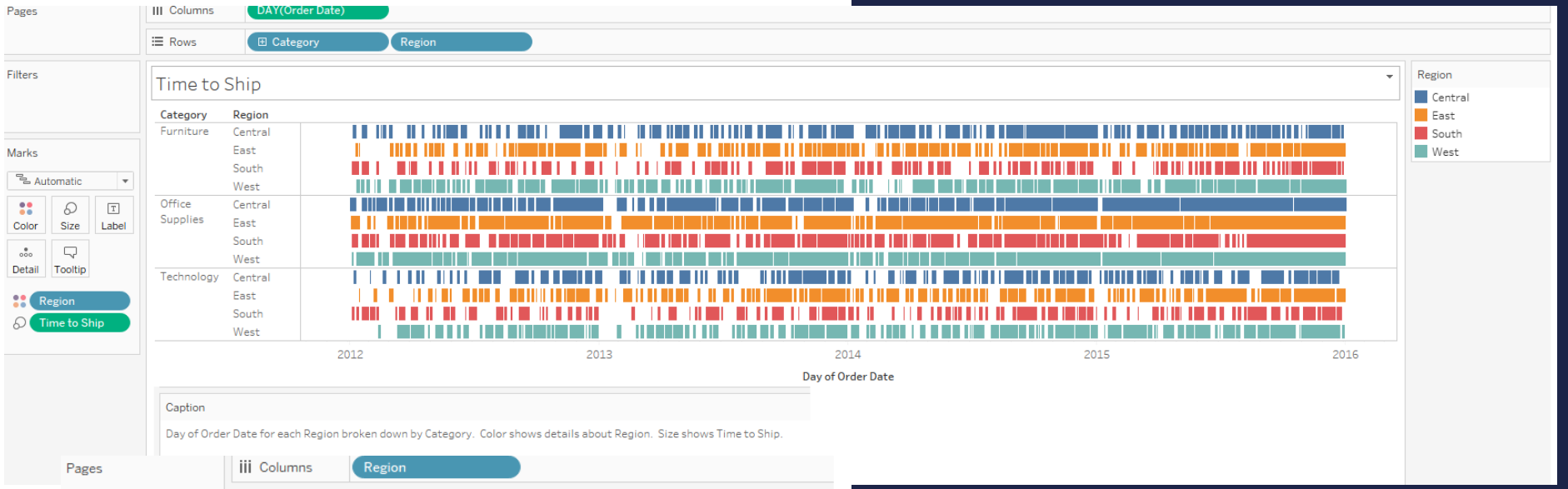
Bullet Chart

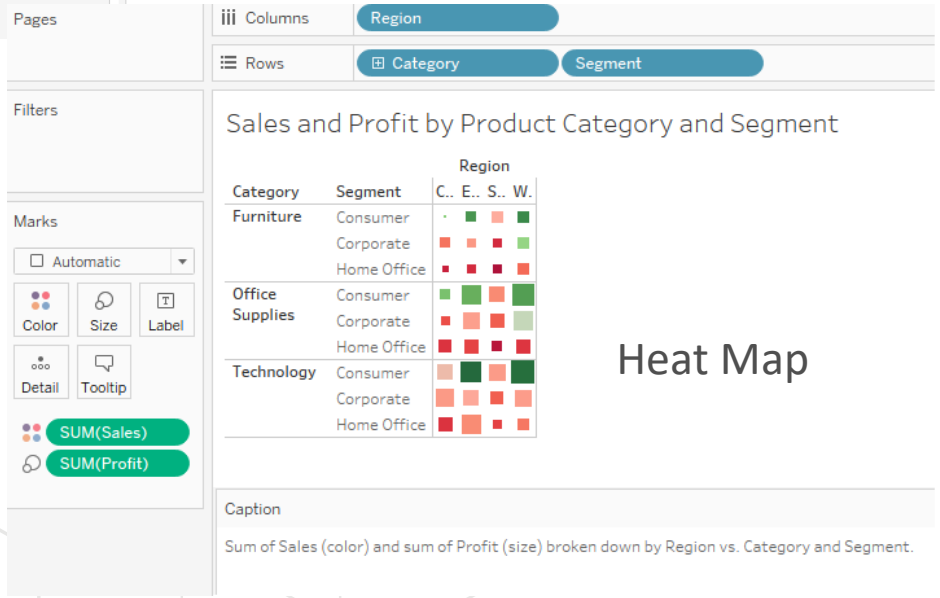
the <nocode> course

Tableau Class Activity

- Using Excel data: global_superstore_2016.xlsx (orders)
- Create the following
  - Gantt chart
  - Heat Map

Gantt Chart





Heat Map

the <nocode> course

15 mins

Part 5 – Advanced Data Visualisation

Visualisation of Spatial Data

Best practices for visualising spatial data

- The most common way to visualise spatial data is with maps that place values within a geographic coordinate. We can visualise the geographic coordinate of a location by mapping the latitude and longitude coordinates to two-dimensional space, and draw a point on the space.

- When the density of individual locations across a region is more informative than the overlapping points on a map, we may want to colour code the region based on the density scale, or use lines to show data continuously over geography.

Visualisation of Spatial Data

Best practices for visualising spatial data

- We can also size the regions by the data and ignore the physical area, so that the regions with high density data will appear bigger than the regions with low density data

- If we want to explore relationships between entities, we can plot each entity on a map, and draw lines to connect each with the others they are associated with

A practical tip - there is a “local” context component to spatial data mapping. For example, would you expect someone who has never been to Singapore to know where Clementi is on a map, unless its explicitly pointed out?

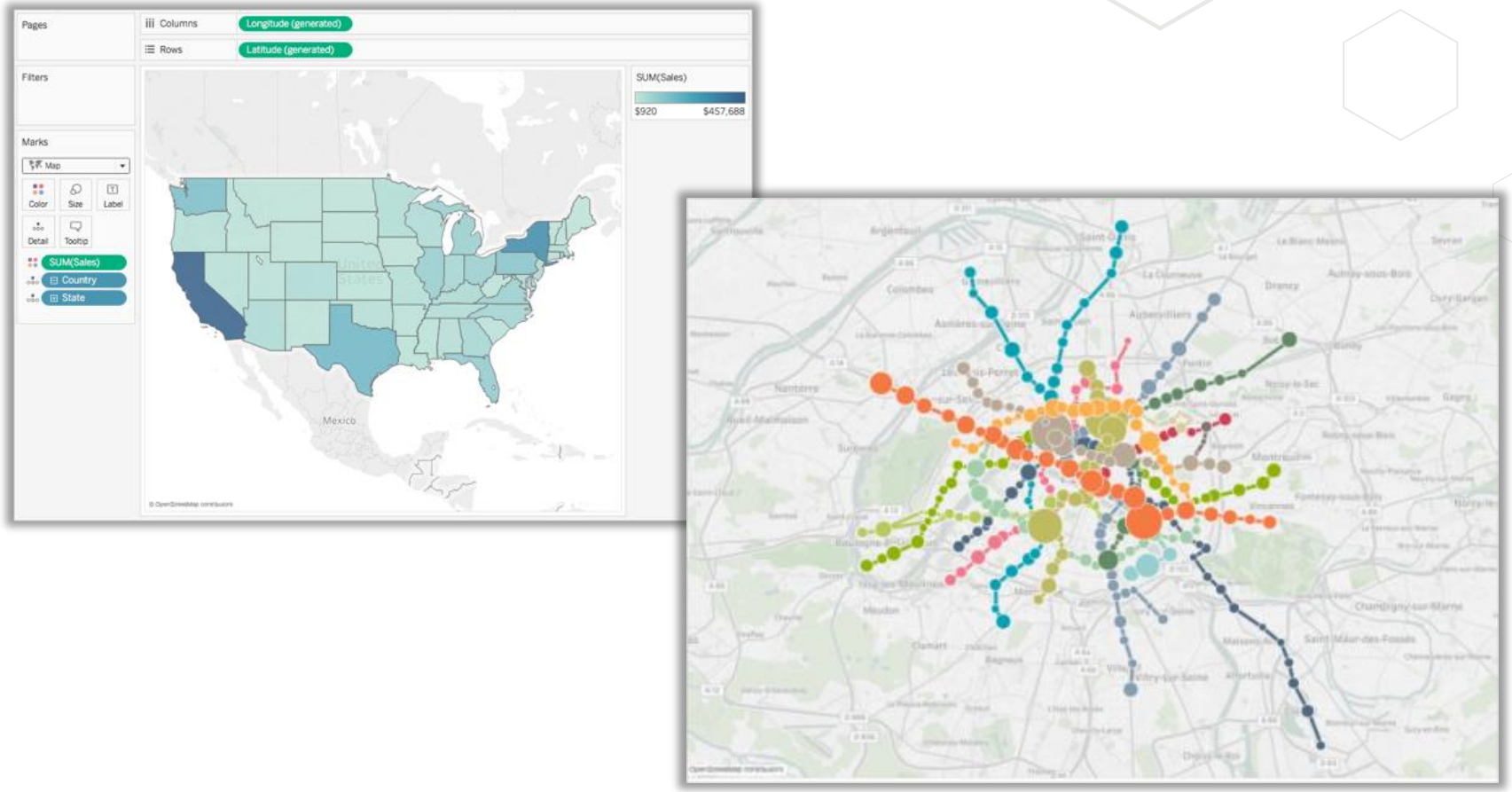
Maps require that you have some background knowledge of the locale; therefore, if you want to use maps, know your audience!

Visualisation of Spatial Data

Best practices for visualising spatial data

- Map
- Mapping Point-to-Point

  Details on Map



the <nocode> course

Visualisation of Spatial Data

Map

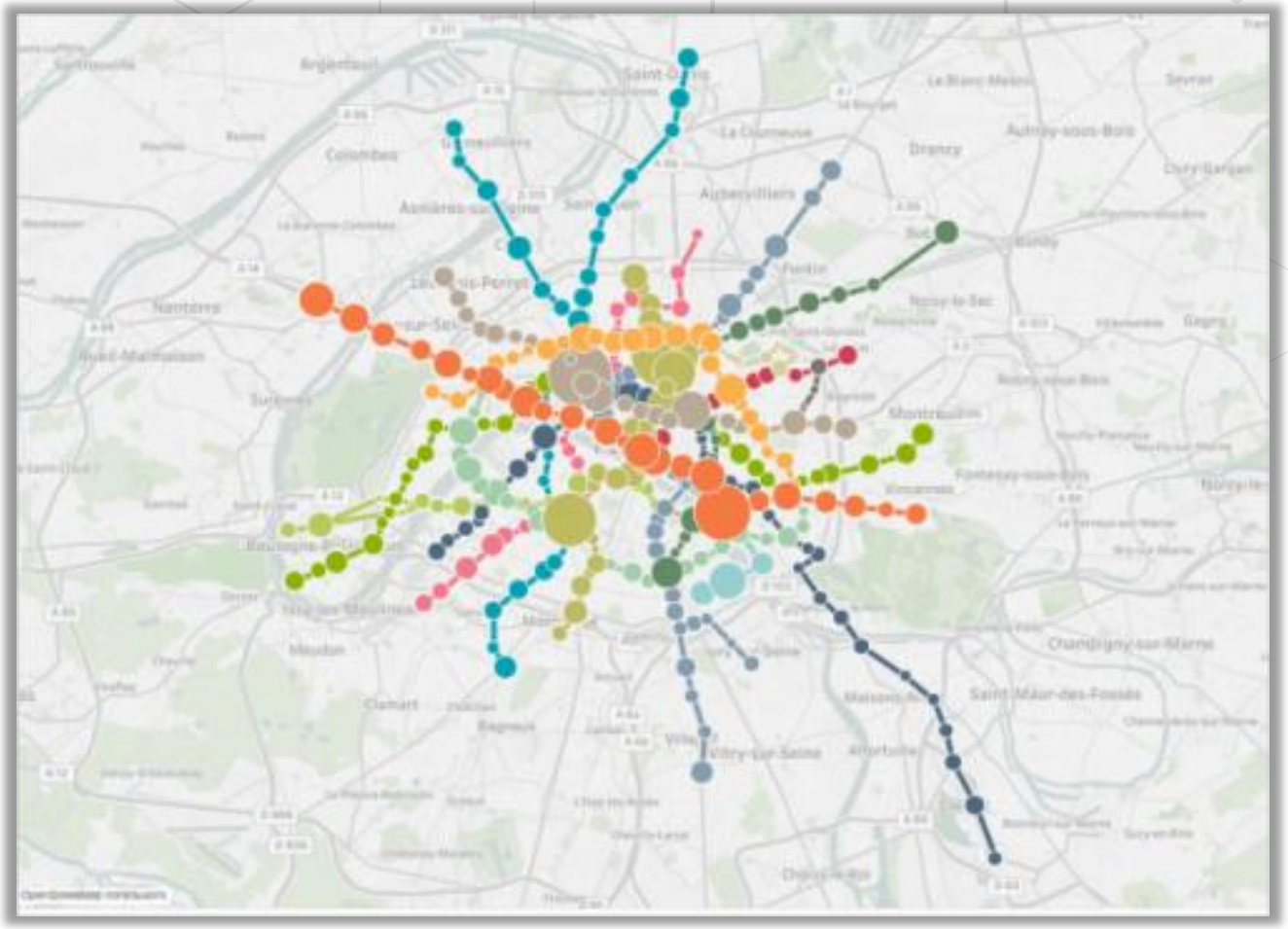
- Assign “Geographic Role” to fields that are Geographic data.
- Double click the “Region” field under Dimension of the Data window. The Tableau will automatically add Longitude and Latitude coordinates to the columns and rows shelves. The “Region” field is automatically placed on the Level of Detail shelf.
- Drag the “Postal Code” field to the Level of Detail shelf to specify more details.

the <no code> course

Visualisation of Spatial Data

Point-to-point details on map

- Used to see linkages between data points e.g., point of origin of goods to where customers are based to answer questions related to “zone of influence”. Could be used to map transportation or logistic routes as well.
- Note that it can be a bit overwhelming; sometimes in visualization, less is more!



Visualisation of Multi Variable Data

Best practices for visualising multi variable data

- To visualise multi-variable data, we can fit all data onto a screen and display the relationships amongst variables or trends in each variables
- We can use scatterplots to visualise the relationships amongst variables
- However, if the relationships amongst variables are not so straightforward, we should consider using multiple views through more straightforward charts

Correlation → how one variable's value tends to change in a certain way as the other variable's value changes
Causation → Extent of how one variable will impact another
Easier to prove correlation than causation
Correlation DOES NOT IMPLY Causation

Visualisation of Multi Variable Data

Correlation vs Causation – An (absurd) example

	Big Foot Sightings	US Netflix growth	UK Netflix growth
2010	1688390	84.06%	5.58%
2011	1310773	58.47%	80.69%
2012	444190	94.32%	5.14%
2013	1840845	53.54%	69.57%
2014	1813040	92.03%	11.70%
2015	67017	4.61%	68.90%

Correlations	
BFS - US Netflix growth	0.50
BFS - UK Netflix growth	-0.12
UK growth - US growth	-0.78

A very good book on the topic is: 'Spurious correlations' by Tyler Vigen.
Here's a link to his personal website: http://www.tylervigen.com/spurious-correlations

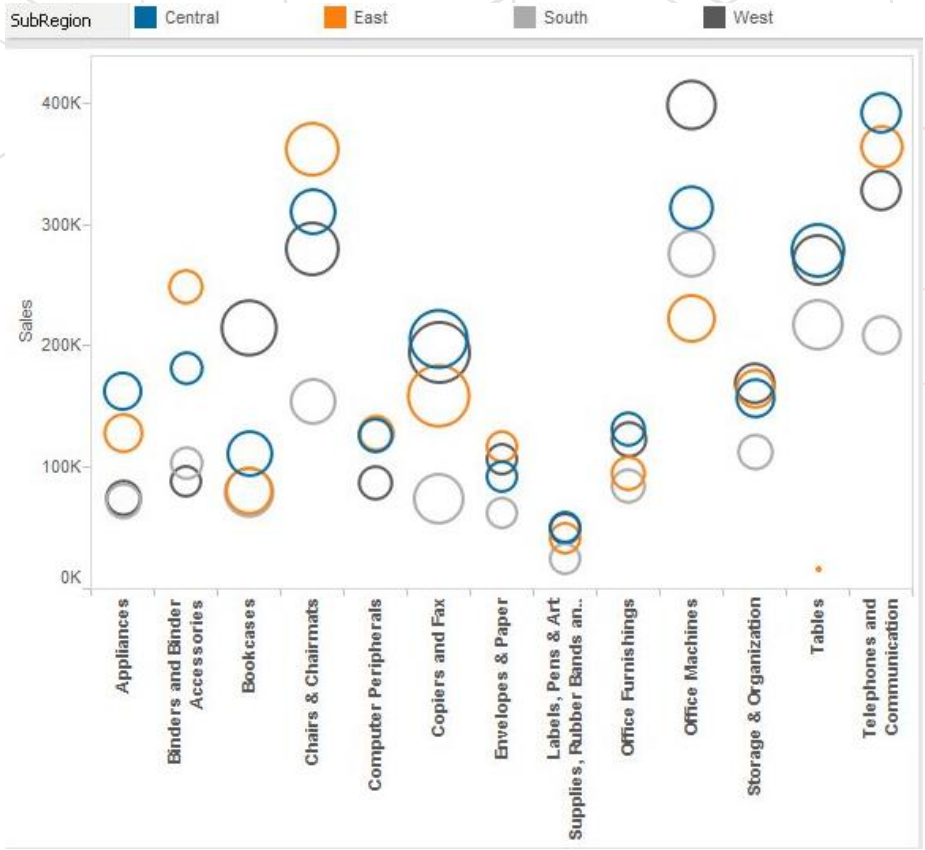
Correlation is in the range of +/- 1, the larger the absolute number, the stronger the relationship between the 2 variables.

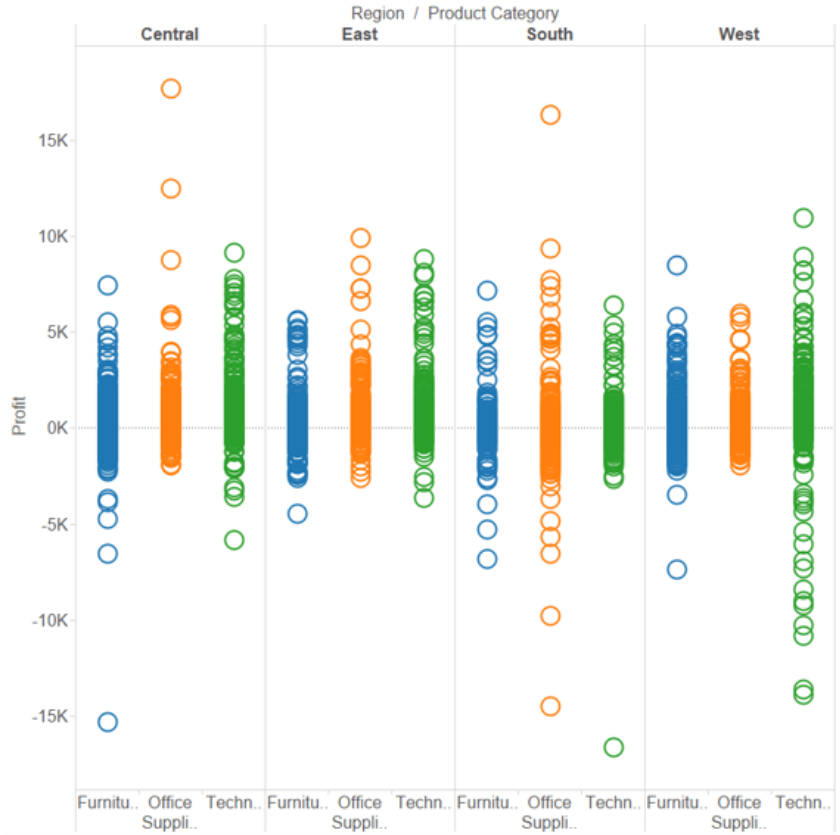
In the above example, I would say there is a (fairly) strong correlation between Big Foot Sightings and Netflix growth in the US. But would it be logical if I said big foot sightings cause the growth in US Netflix users?

Visualisation of Spatial Data

Visualising multi variable data

- Circle View
- Side-by-Side Circle Plots
- Scatterplot



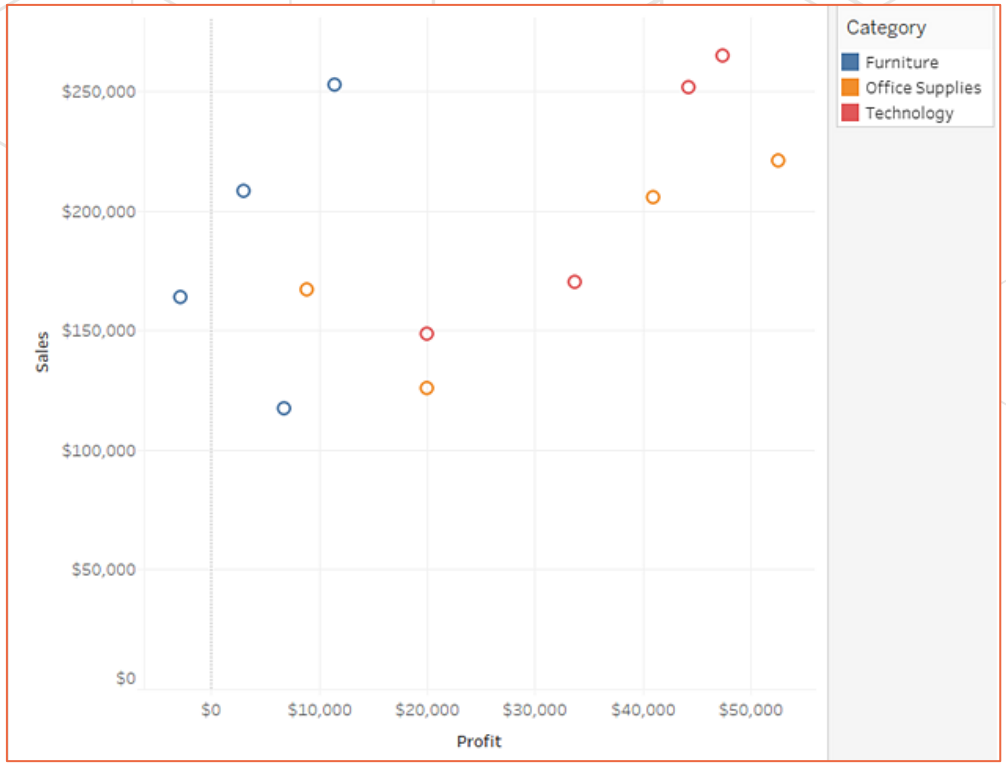


the <nocode> course

Visualisation of Spatial Data

Scatterplot

- Uses Cartesian coordinates to display values for typically two variables for a set of data.
- Dots not only report the values of individual data points, but also patterns when data are taken together i.e., how closely points are clustered together. Also useful to identify gaps in data.
- Identification of correlational relationships are common with scatter plots. Variable on horizontal axis denoted an independent variable, and the variable on the vertical axis the dependent variable. Relationships between variables can be described in many ways: positive or negative, strong or weak, linear or nonlinear.



Overplotting

Can be difficult to tell how densely-packed data points are when many of them are in a small area.

- Subset/Averaging
- Optics - Adding transparency or reducing point size

Visualisation of Data Distributions

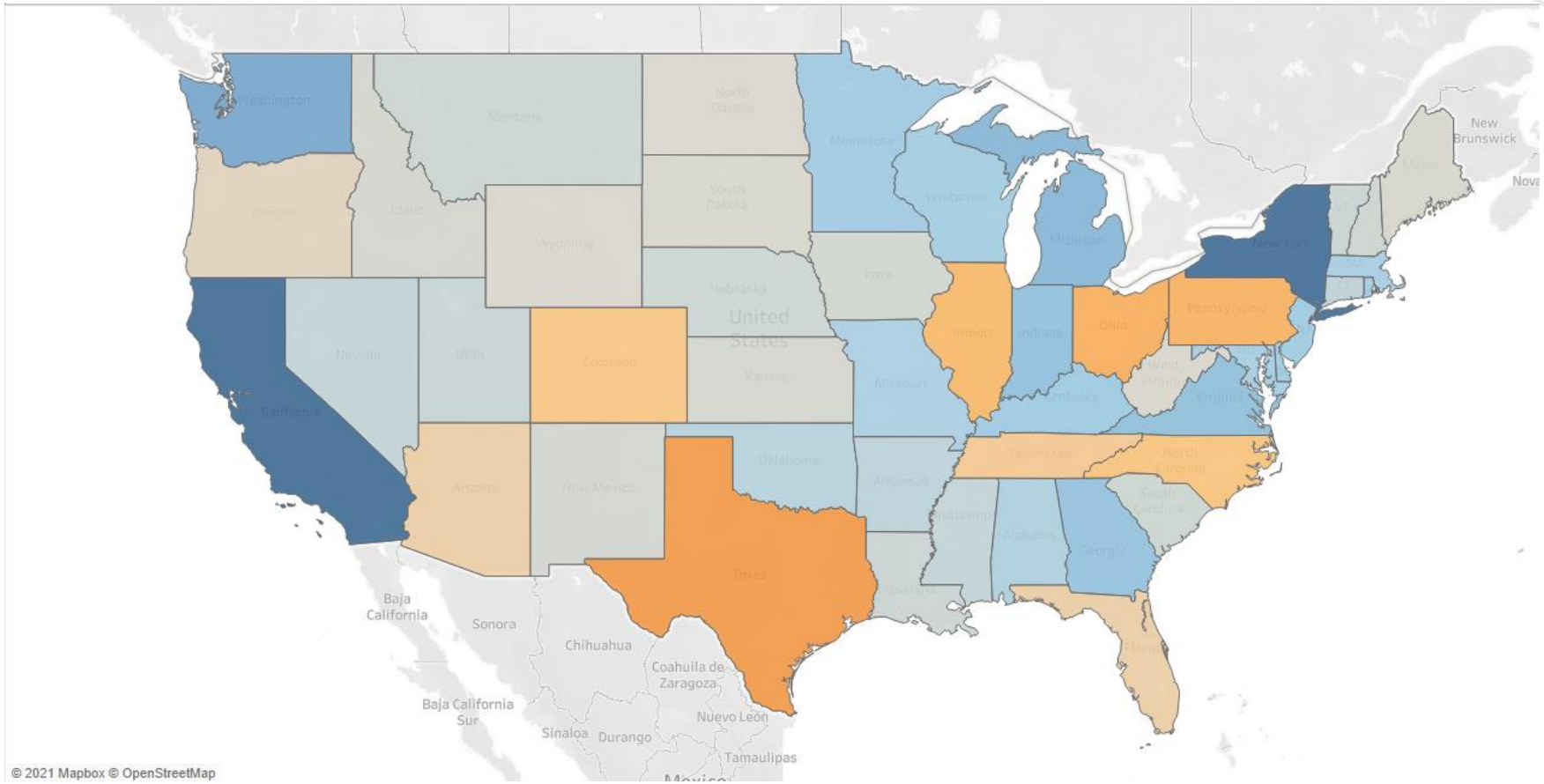
- We can visualise the distribution of data at different granularities with bullet charts, bubble charts, and histograms
- For single-variable distributions, using a histogram will enable us to see where the data is clustered and any outliers, by keeping track of where the outliers sit on the vertical axis
- For multi-variable distributions, sometimes values come as pairs, so it makes more sense to visualise both values at the same time

Visualisation of Data Distributions

- Map View
- Scatterplot
- Bubble Chart
- Histogram
- Box Plot
- Interactive Plot
- Forecasting

Visualisation of Data Distributions

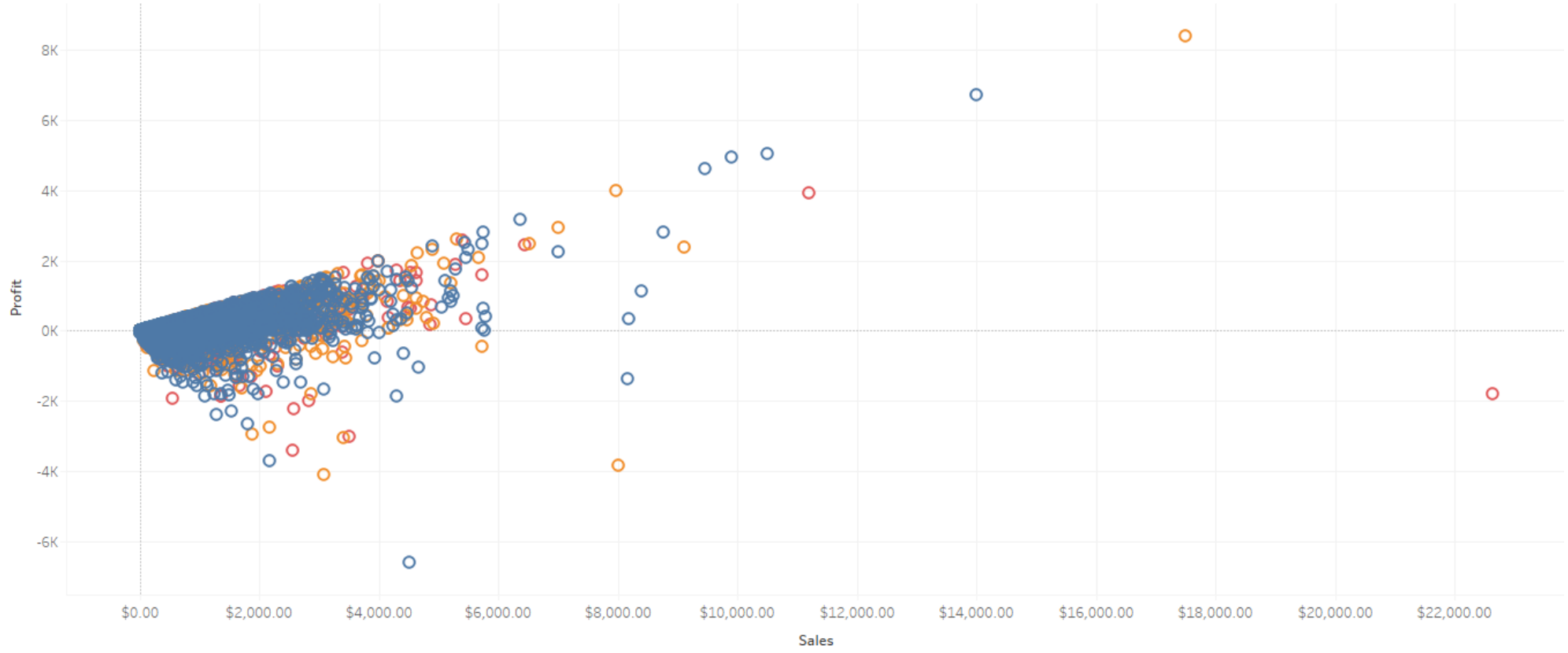
Map View



the <no code> course

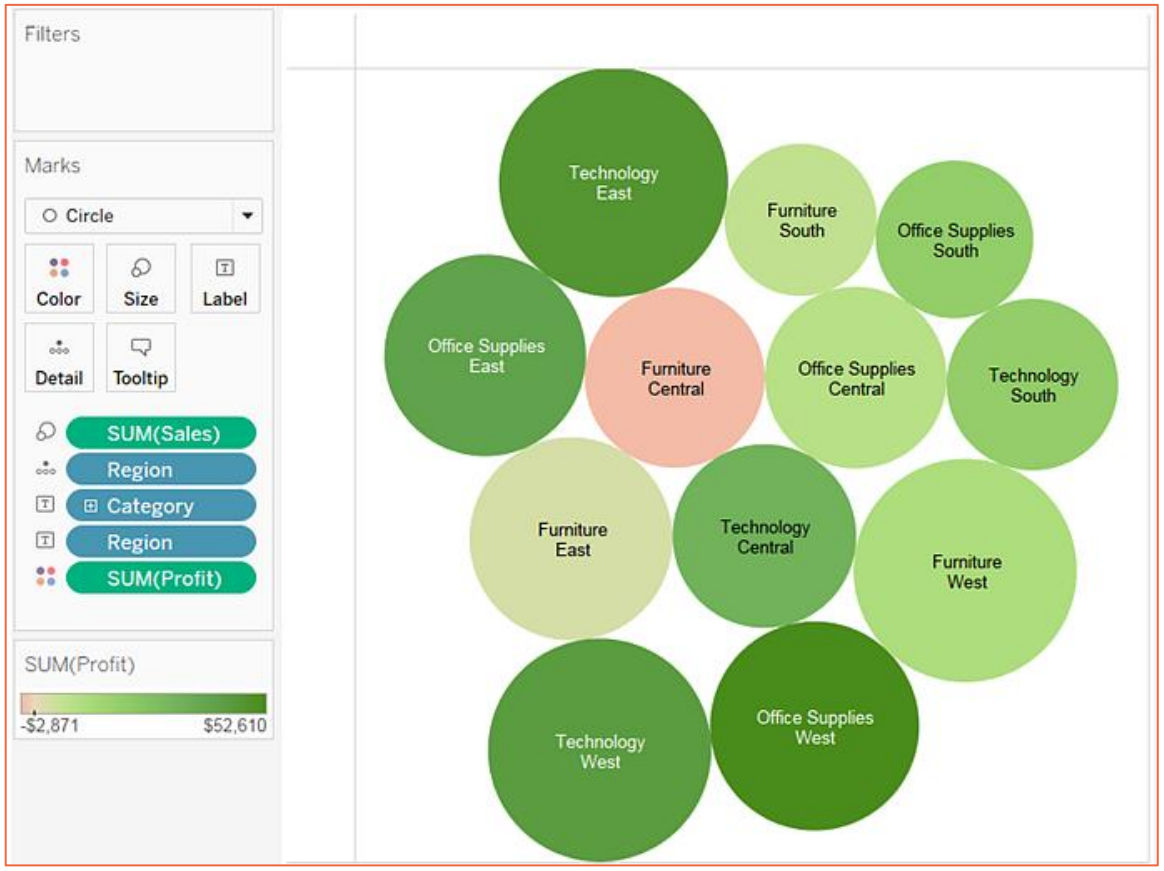
Visualisation of Data Distributions

Scatterplot



Visualisation of Data Distributions

Bubble Chart



Use packed bubble charts to display data in a cluster of circles.

Dimensions define the individual bubbles, and measures define the size and color of the individual circles.

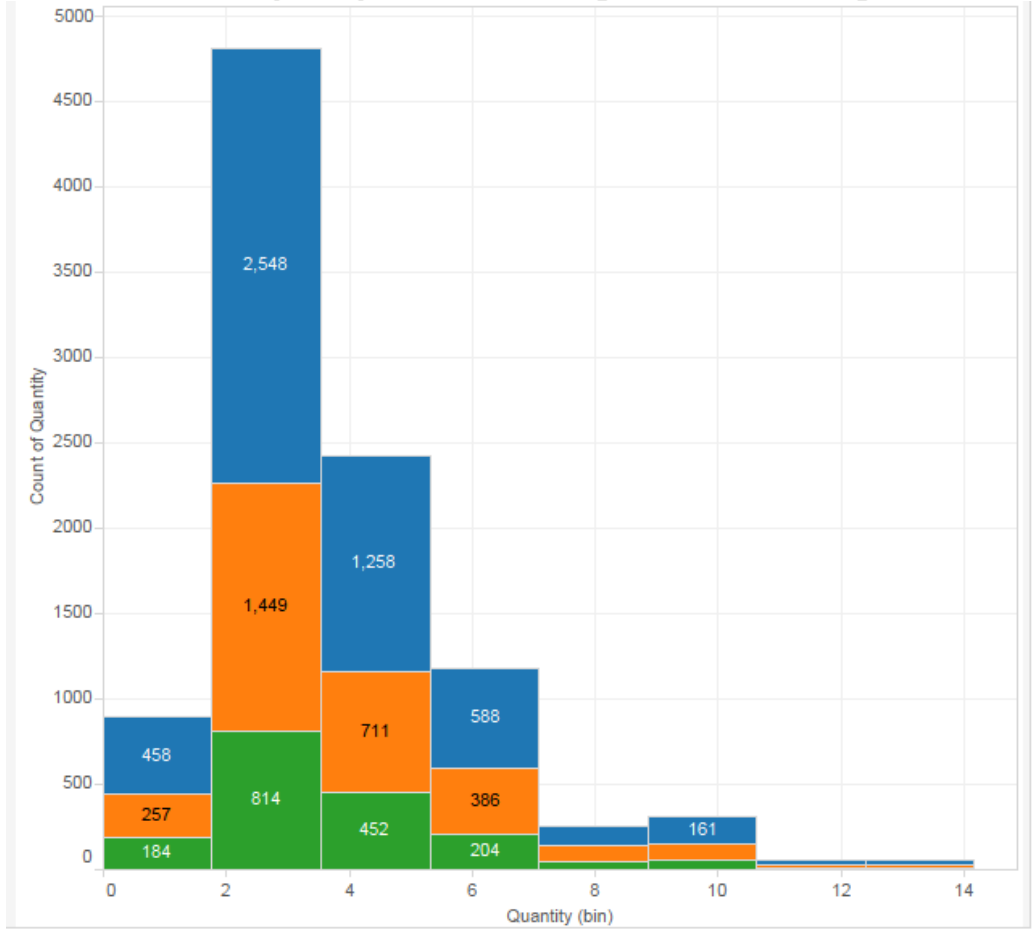
Note that its not very useful for precise comparisons

Visualisation of Data Distributions

Histogram

Histograms are used to show distributions of variables while bar charts are used to compare variables.

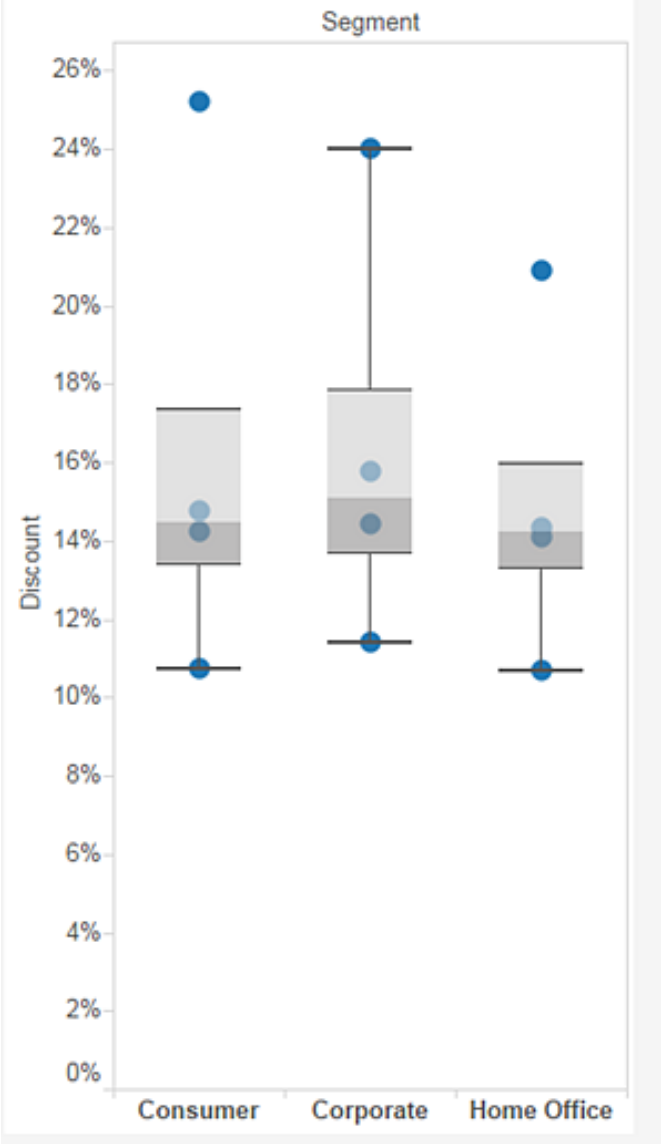
Histograms plot quantitative data with ranges of the data grouped into bins or intervals while bar charts plot categorical data.



the <no code> course

Visualisation of Data Distributions

Box plot



Forecasting

- Forecasting is the act of predicting future values based on historical values
- Parameters are variables that allow data visualisation users to alter the content of a formula or change a dimension or measure contained in the view
- Using parameters, data visualisation users can change normally static values into dynamic entities that facilitate ad-hoc analysis without the need to change the design of the data visualisation

Forecasting

- Trend and season
- Trend only
- Season only
- No trend or season
- Parameters: Reference Line, Bin Size (Histogram), Ranking (in-value comparison)

Please read these two webpages to understand how to perform forecasting in Tableau…

https://help.tableau.com/current/pro/desktop/en-us/forecast_options.htm

https://help.tableau.com/current/pro/desktop/en-us/forecast_how_it_works.htm

Tableau recommends the best forecasting method for the data but note that “predicting” the future is always imprecise. Even simple trend lines have an implied “Ceteris Paribus” (all things being equal) assumption.

Don’t try to suggest/utilize forecasting data unless you are familiar with the methods, especially pros and cons.

Lets

with

tableau

10 mins

Part 6 – Dashboard & Storyboards

Dashboard Design Principles

- Keep it simple
- Don’t display everything
- Keep to a single page
- Avoid fancy formatting
- Use layout and placement
- Format numbers effectively
- Use titles and labels effectively

Ask the fundamental questions:

- Who (is my audience)?
- What (is my objective – information sharing or KPI or …)?
- Where (will I need to focus on – visuals, data, data collection)?
- When (was this data collected, when do I intend for it to be actioned, is it still relevant)?
- Why (is a dashboard necessary, can I just use pen and paper)?
- How (will I explain what I created)?

Keep it Super Simple

Tableau Class Activity

1. Follow along to create a storyboard using dashboards for Amazing Mart.

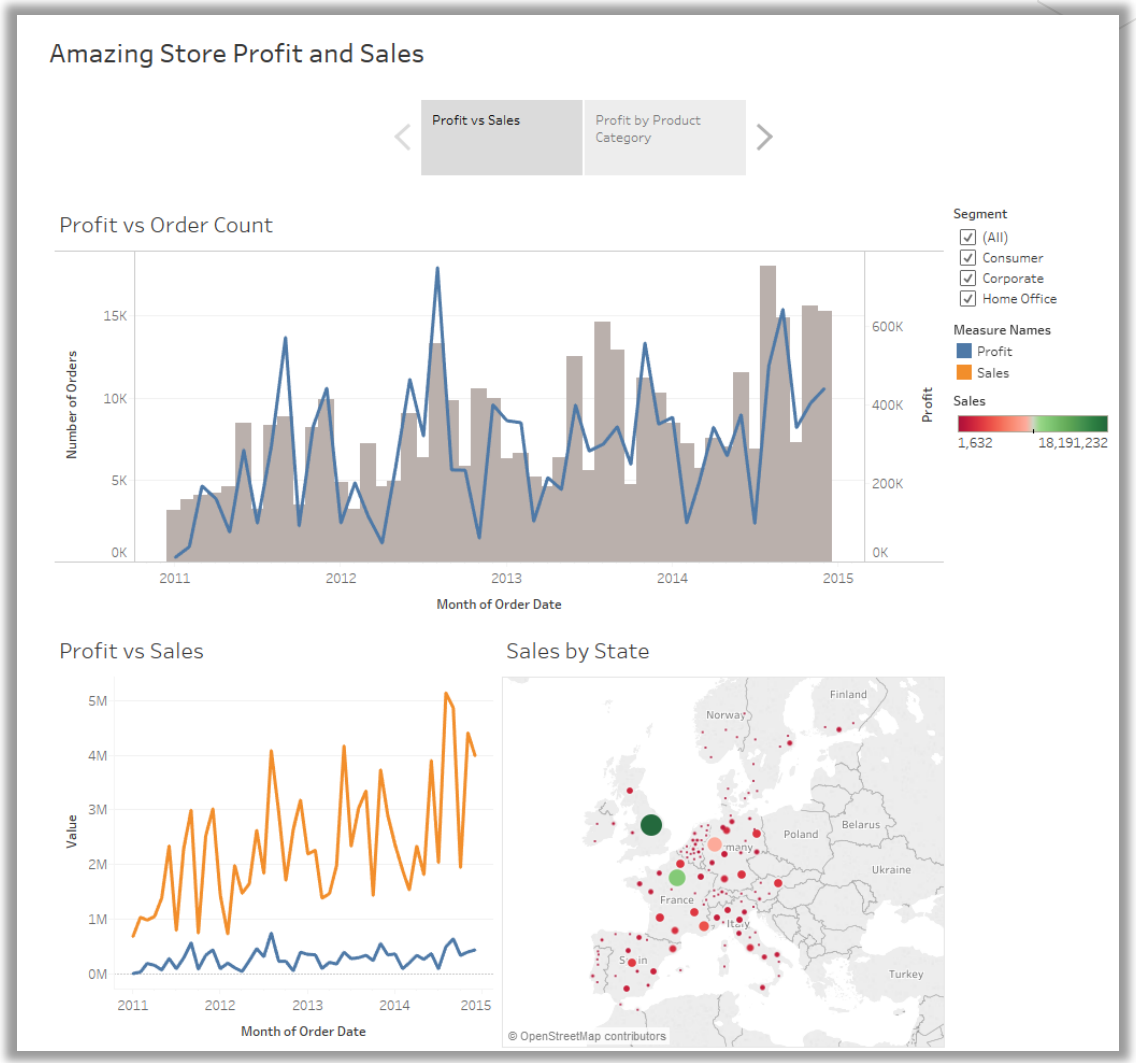
- Amazing Mart's Sales Manager would like to have a dashboard that **tracks the sales and profits** of the products they sell.
- The collected data is stored in 2 entities, **List of Orders** and **Order Breakdown**.
- **List of Orders** contains information about the customer, such as customer name, the segment they are in, order date, shipping date, and mode.
- **Order Breakdown** contains information about the product they purchased, the sales price and the profit that was made from the sales.

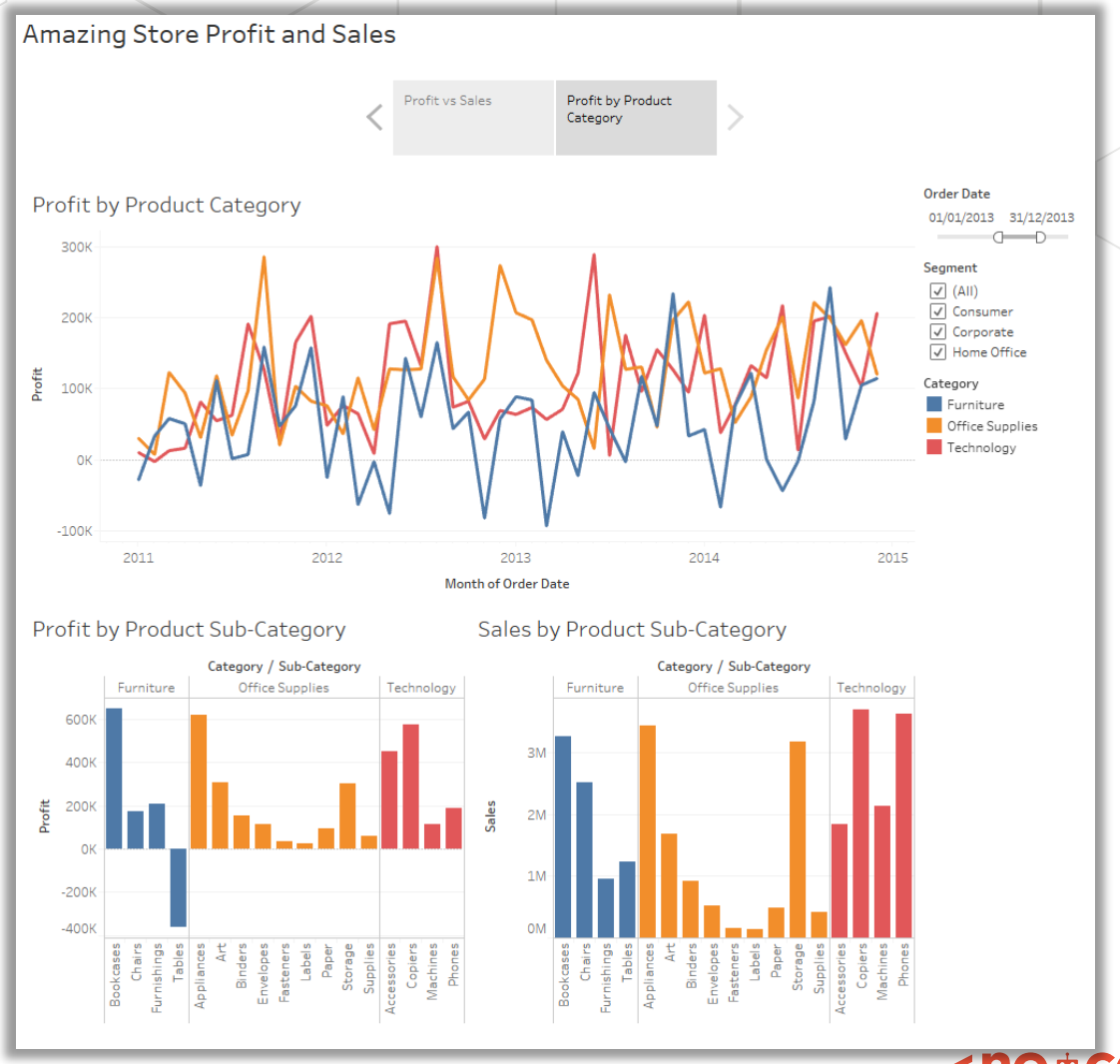
Lets

with

tableau

Tableau Class Activity






Thank you!
the
<no code>
course