



# Data Visualisation for Business

## ANL 201

*Science and Art of Data Visualisation*  
*Study Unit 2*

January 2024

# Recap- Overview of Business Performance Measurement

- ▶ Business Performance Measurement- Definition, Performance Metrics, and importance of metrics
- ▶ *The Four Balanced Scorecard Perspectives*
- ▶ *Strategy Map*
  - ▶ *Vision, Mission, Strategic Themes, Strategic Objectives, Measure, Target, Initiatives*

# Data Visualisation

# What is Data Visualisation?

- ▶ Data - facts and statistics used for reference or analysis
- ▶ Visualisation - the graphical representation of data

**Data visualization**” refers to transforming figures and raw data into visual objects: points, bars, “line plots, maps, etc. “Data visualization is the art of depicting data in a fun and creative way, beyond the possibilities of Excel tables.

*Charles Miglietti, an expert in data visualization and co-founder of Toucan Toco*

<https://youtu.be/VyhLRJV0lrl>

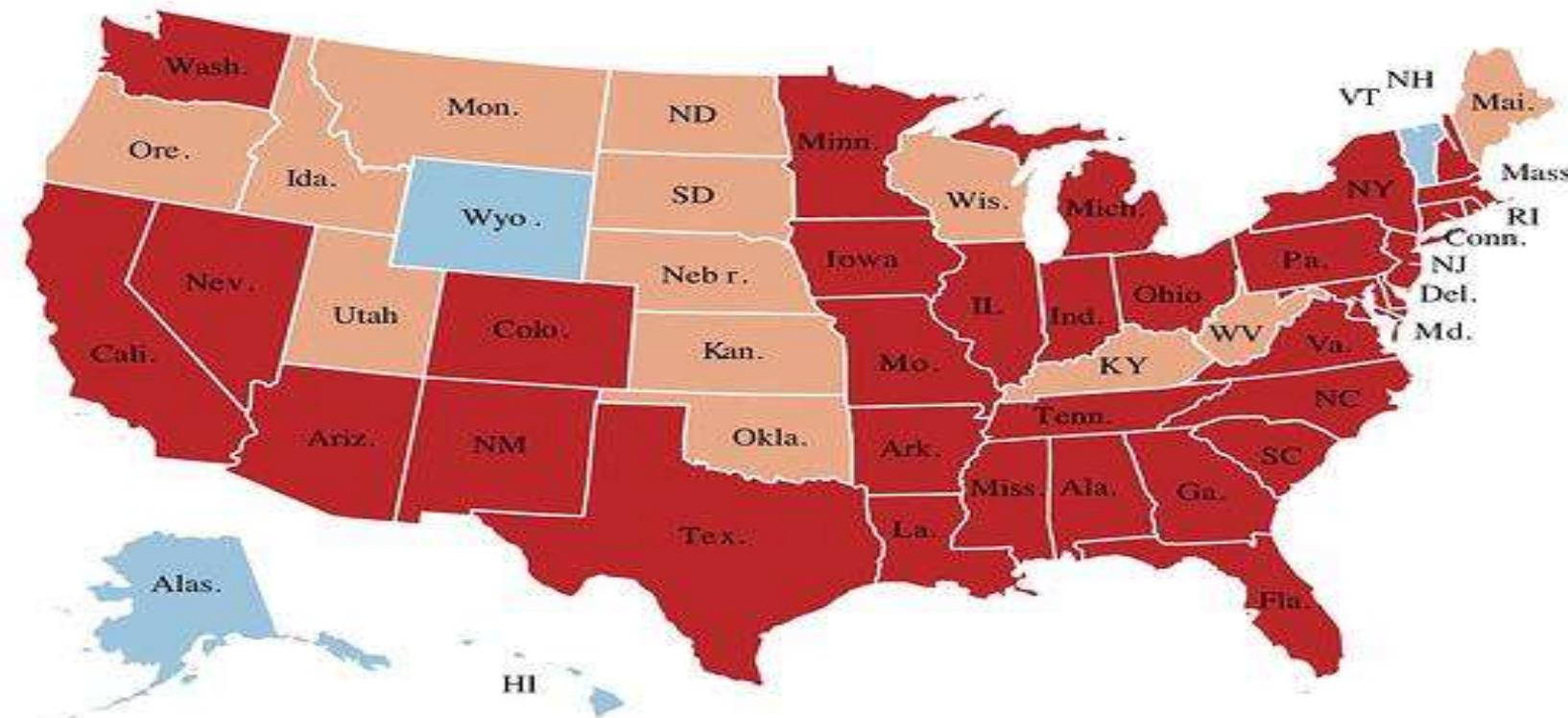
# Benefits of Data Visualisation

- ▶ Provides us the ability to comprehend huge amounts of data
- ▶ Allows the perception of emergent properties that are not anticipated
- ▶ Facilitates hypothesis formation
- ▶ Often enable problems with data to become immediately apparent

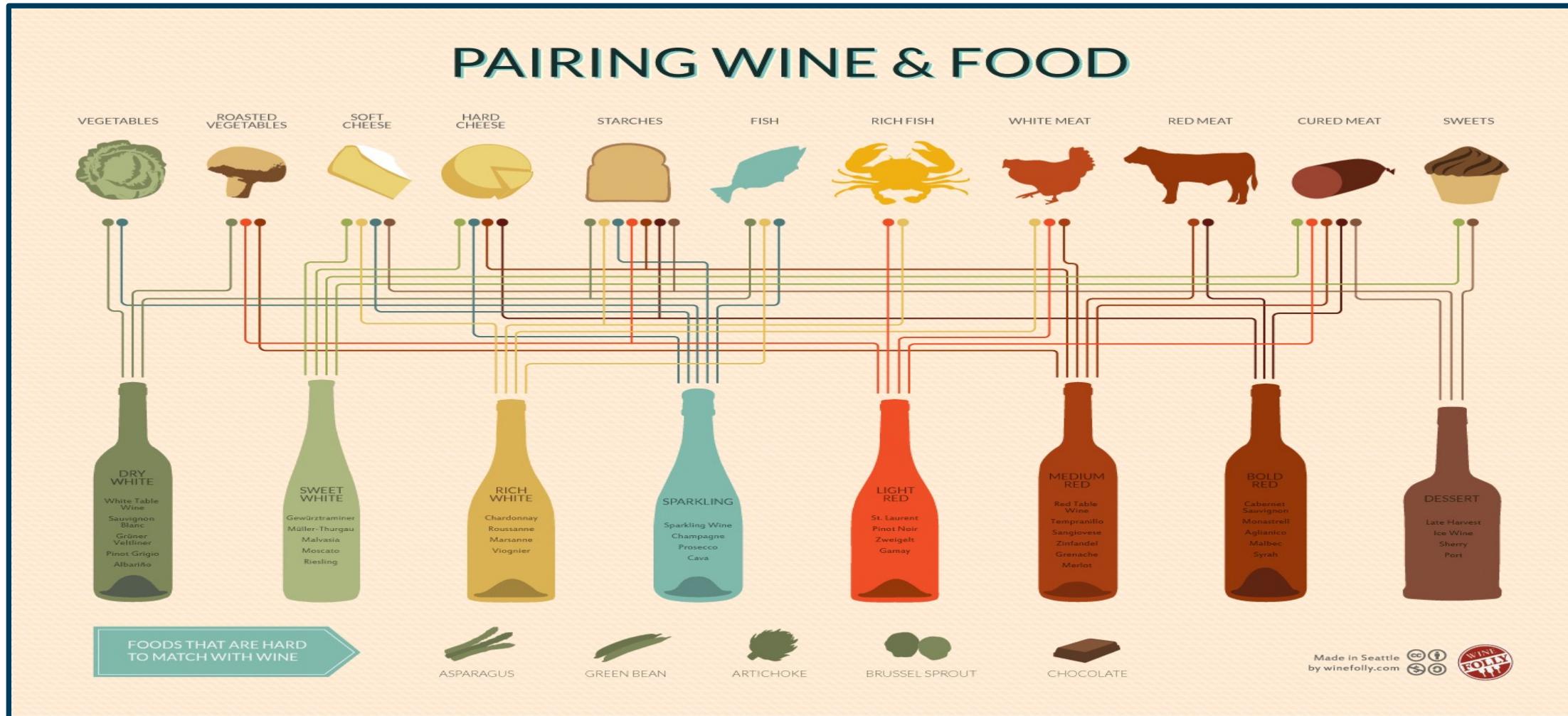
# Data Visualisation- Data visualisation in everyday life

**Number of confirmed Covid-19 deaths per 100,000 Americans**

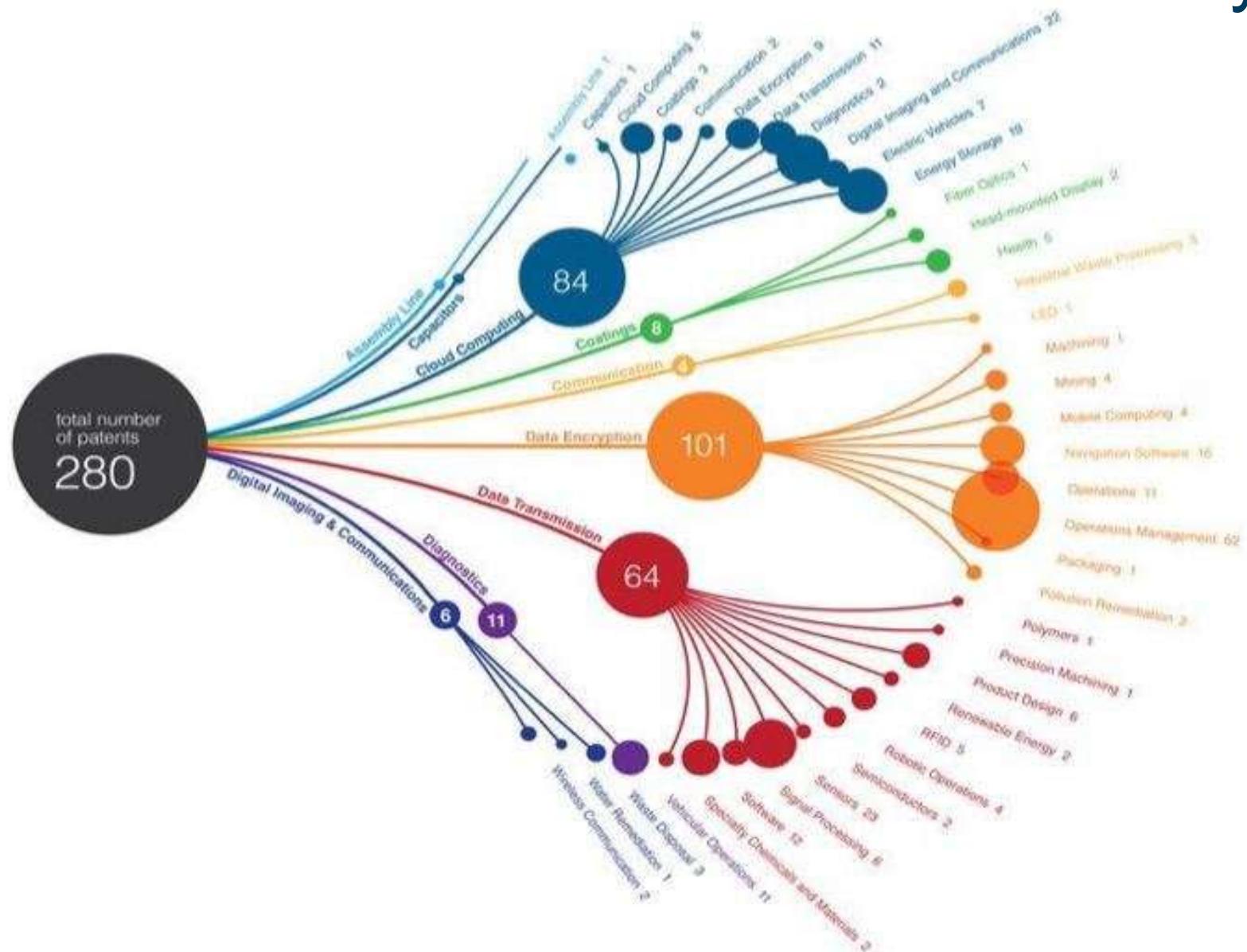
- █ Fewer than 5
- █ At least 5 per 100k
- █ At least 10 per 100k
- █ At least 25 per 100k



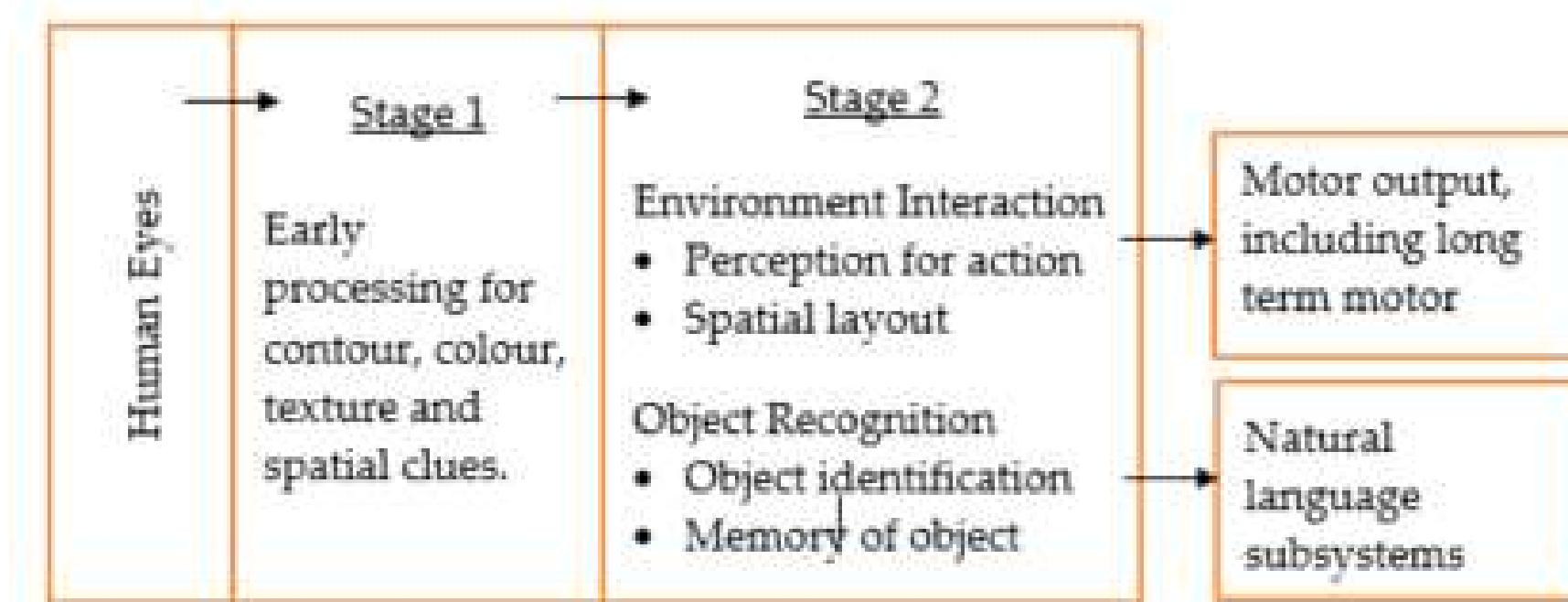
# Data Visualisation- Data visualisation in everyday life



# Data Visualisation- Data visualisation in everyday life



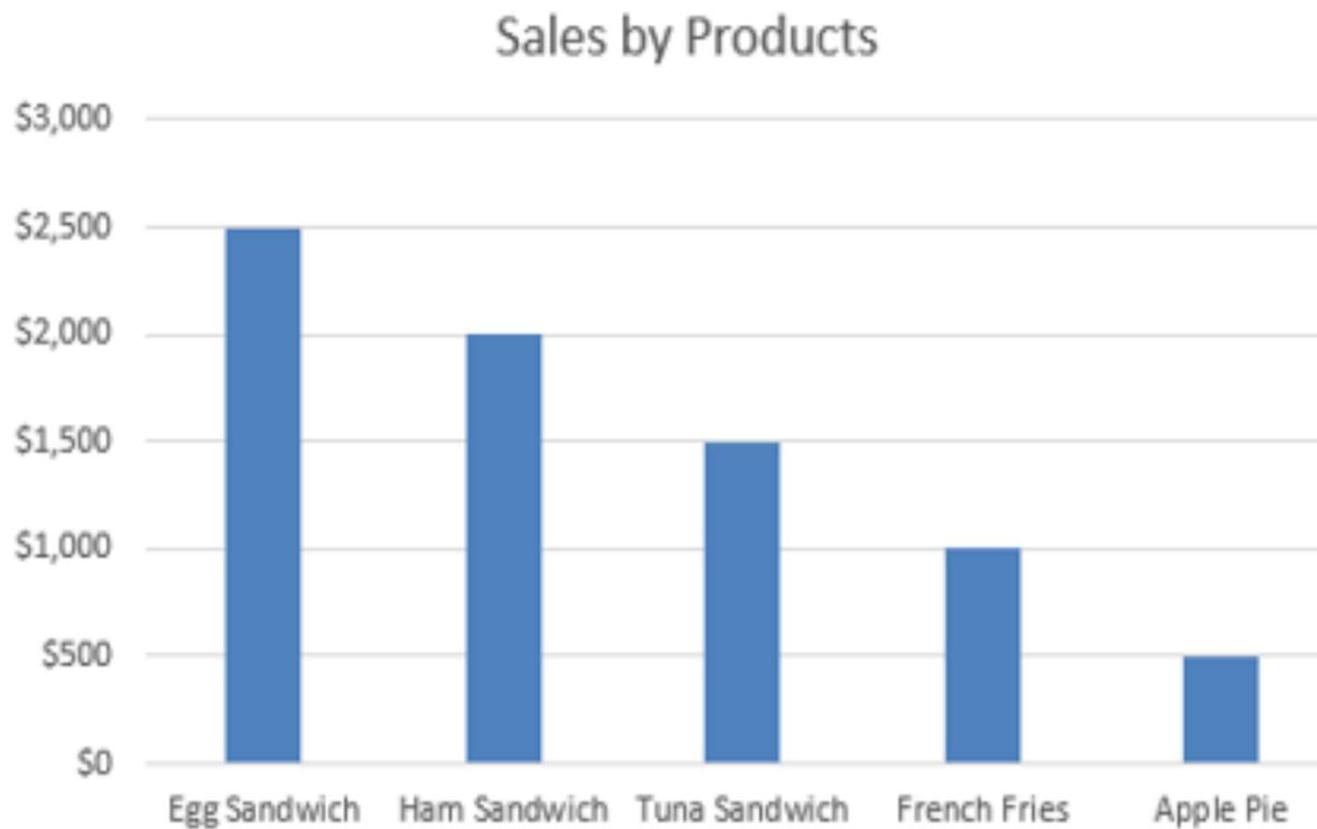
# Perceptual Processing Model



*Overview of a two-stage model of human visual information processing (Ware, 2013)*

# Four Components of Data Visualisation

# Four Components of Data Visualisation



Visual



Scale



Coordinate Systems

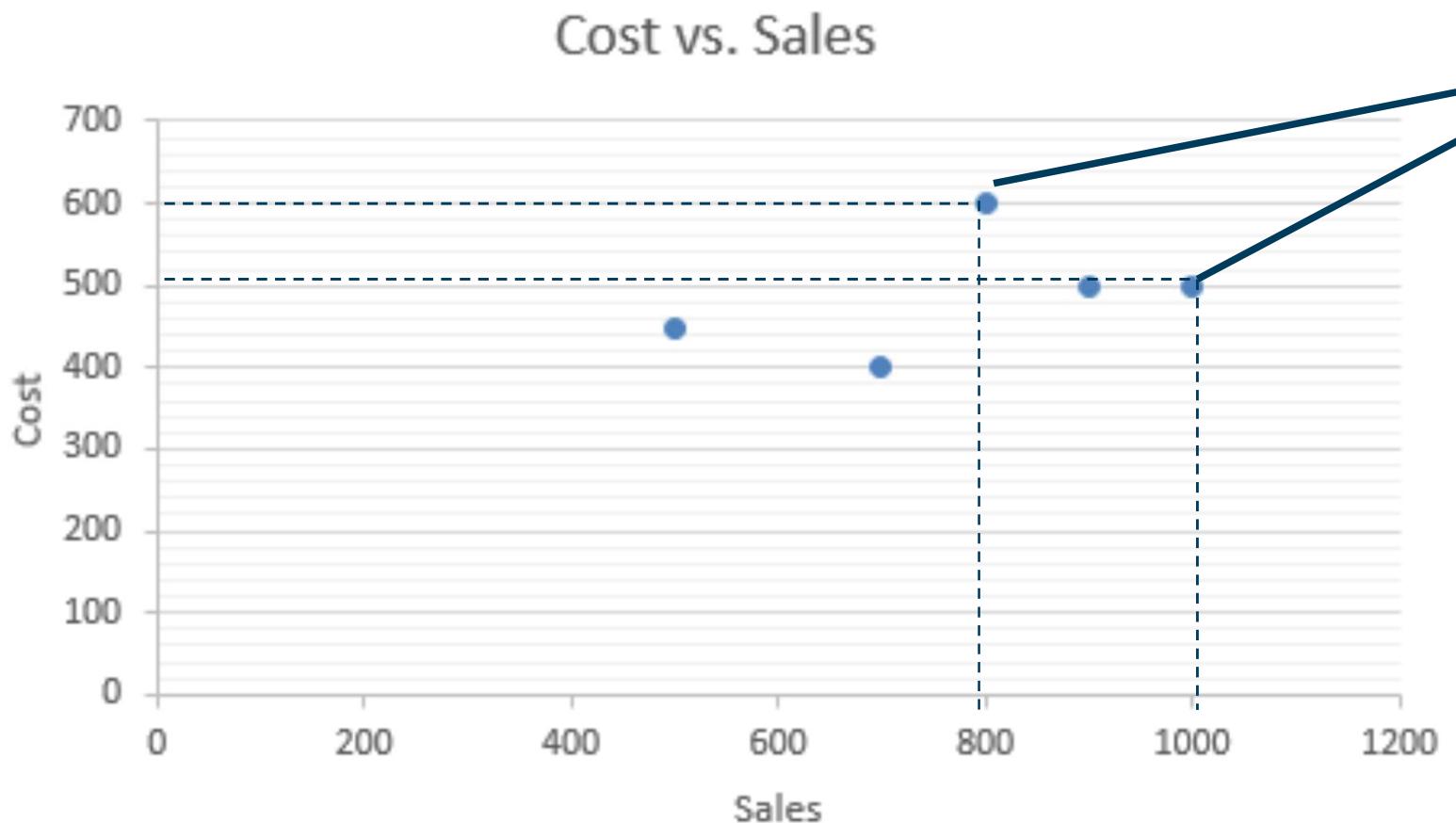
Context



# Visual Cues

# Visual Cues- Types

## 1. Position (e.g., scatterplot)

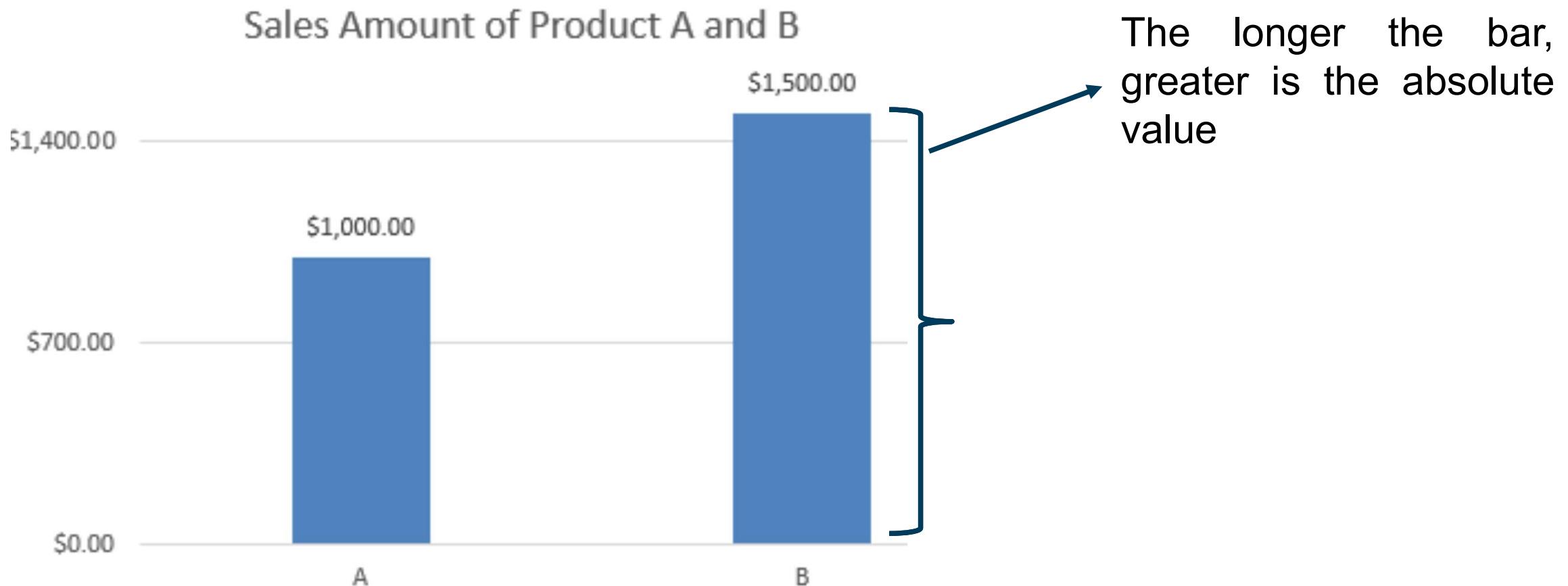


Position defines the position of a data point in space as compared to other data points in the data

Helps you compare the sales and cost

# Visual Cues- Types

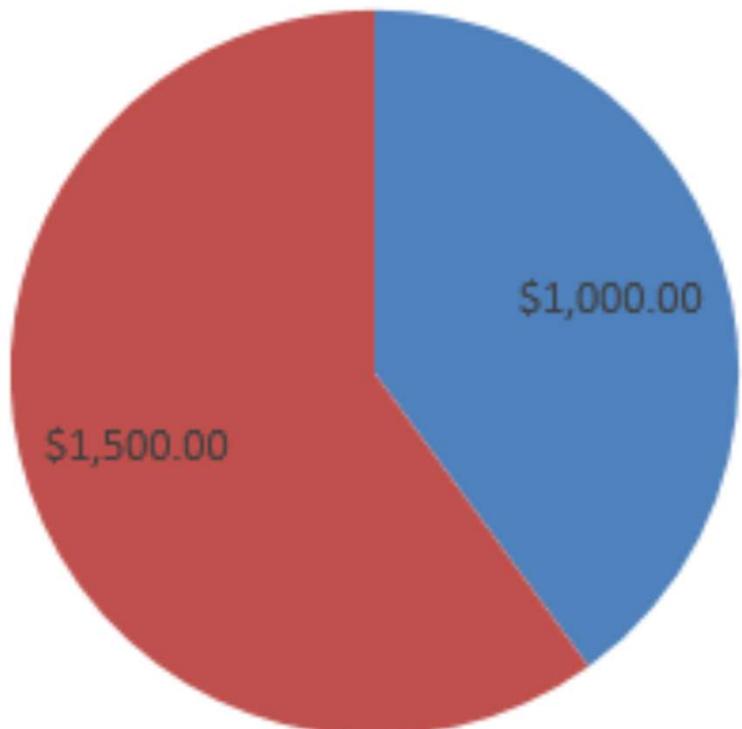
## 2. Length (e.g., bar chart)



# Visual Cues- Types

3. Angle (e.g., pie chart)

Sales Amount of Product A and B



Angles are commonly used to represent parts of a whole

# Visual Cues- Types

## 4. Direction (e.g., line graph)



Line graphs help us see increase, decrease and fluctuations

# Visual Cues- Types

## 5. Shape (e.g., scatterplot)



Shape or symbol is commonly used as a visual cue to differentiate different products/categories/objects in data

# Visual Cues- Types

## 6. Area and Volume- Area charts



Bigger objects represent greater values- larger area means profit made by the grocery chain in yellow is higher than the one in grey and the one colored in blue makes the least profit

# Visual Cues- Types

## 7. Colour SPOT THE BEAR!



Color as a visual cue helps us break camouflage and differentiate things visually from their surroundings

# Class Discussion 1

Can you identify the components of visual cues on the dashboard ?



Source: <https://dapresy.com/wp-content/uploads/2016/07/Image-4.png;>  
[http://www.sas-sr.com/intro\\_sas/mercredi/major.PNG](http://www.sas-sr.com/intro_sas/mercredi/major.PNG)

# Class Discussion 1

Can you identify the components of visual cues on the dashboard ?

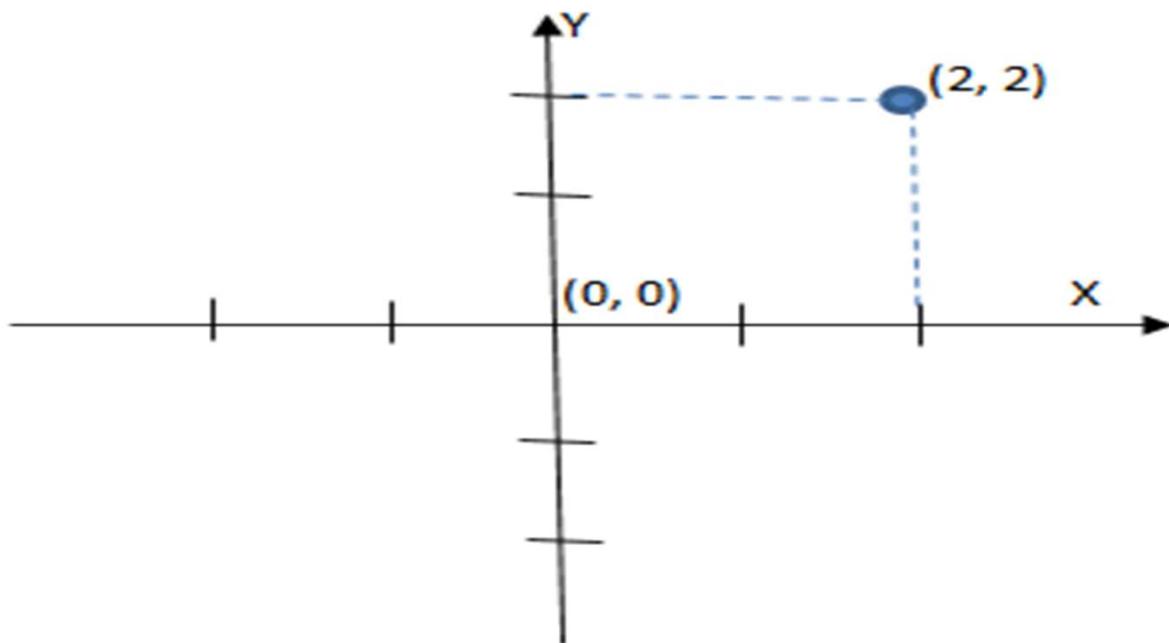




# Coordinate Systems

# Coordinate Systems- Types

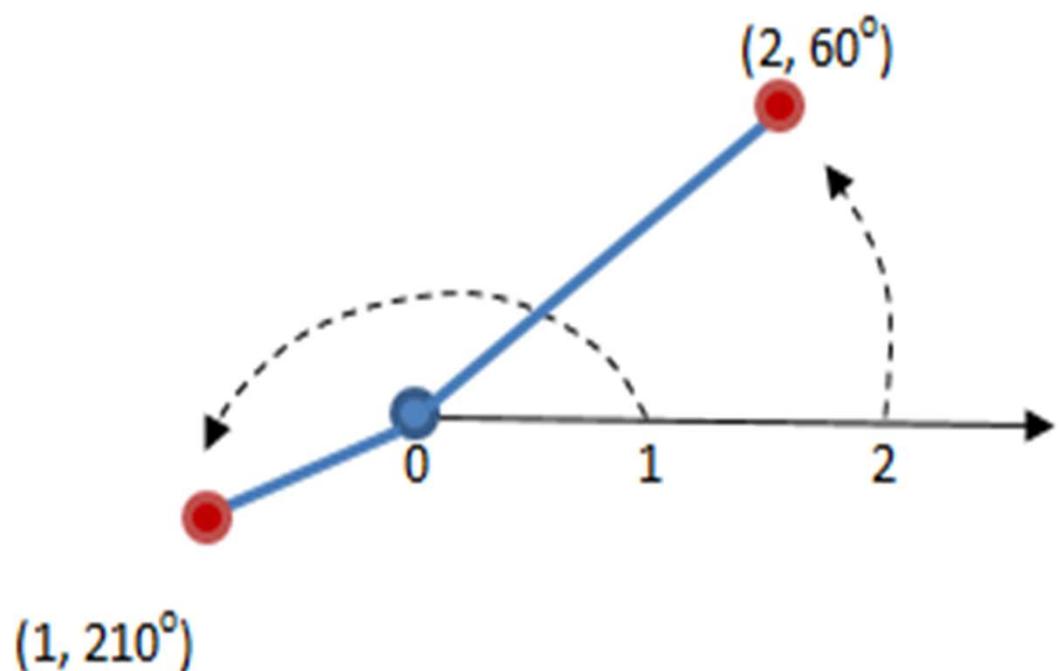
The cartesian coordinate system



- ▶ Two fixed perpendicular reference lines- x-axis and y-axis
- ▶ Both axes meet at a point, called the origin (0,0)
- ▶ The numerical coordinates are the signed distance from the origin

# Coordinate Systems- Types

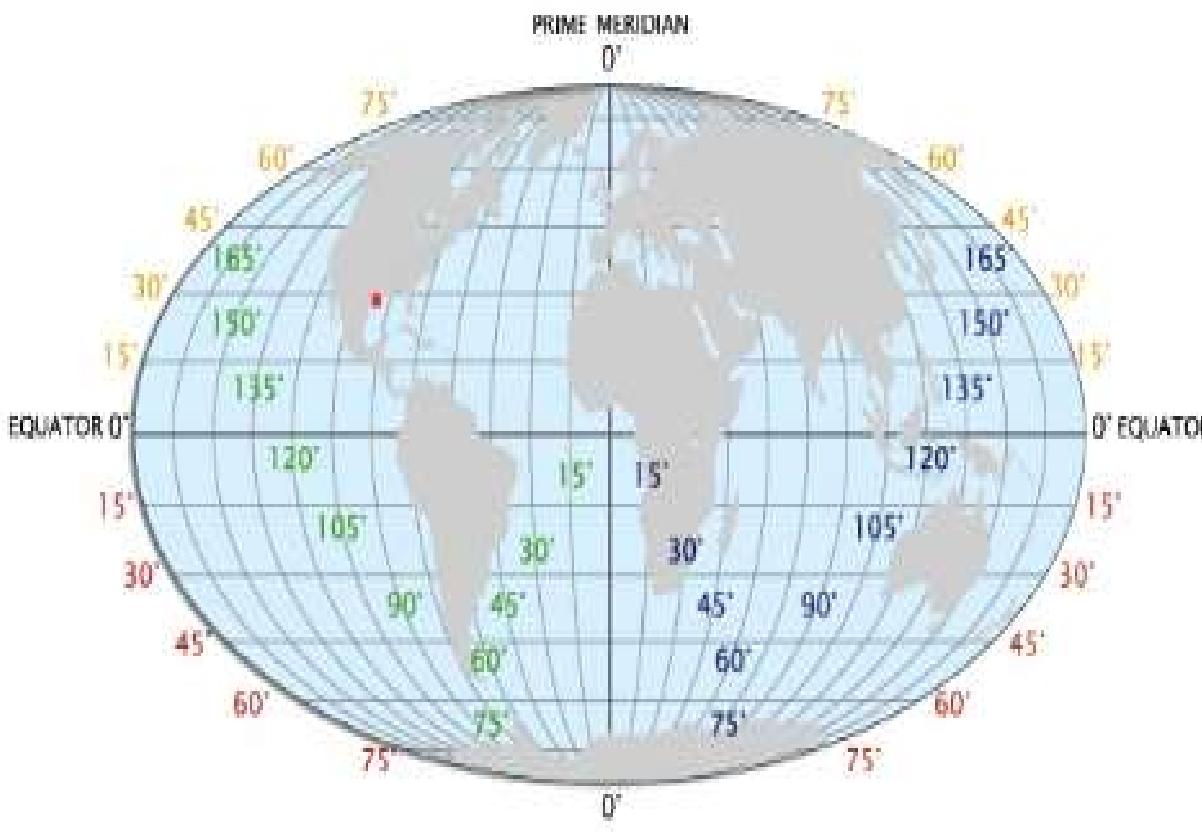
The polar coordinate system



- ▶ The fixed point  $(0,0)$  is called the pole
- ▶ The ray or half-line from the pole in the fixed direction is called the polar axis
- ▶ Each data point is determined by the distance and an angle

# Coordinate Systems- Types

The geographic coordinate system



- ▶ Represents every location on the earth using latitude and longitude
  - Latitude lines run east and west
  - Longitude lines run north and south
  - Elevation can be thought of as a third dimension (height above sea level)

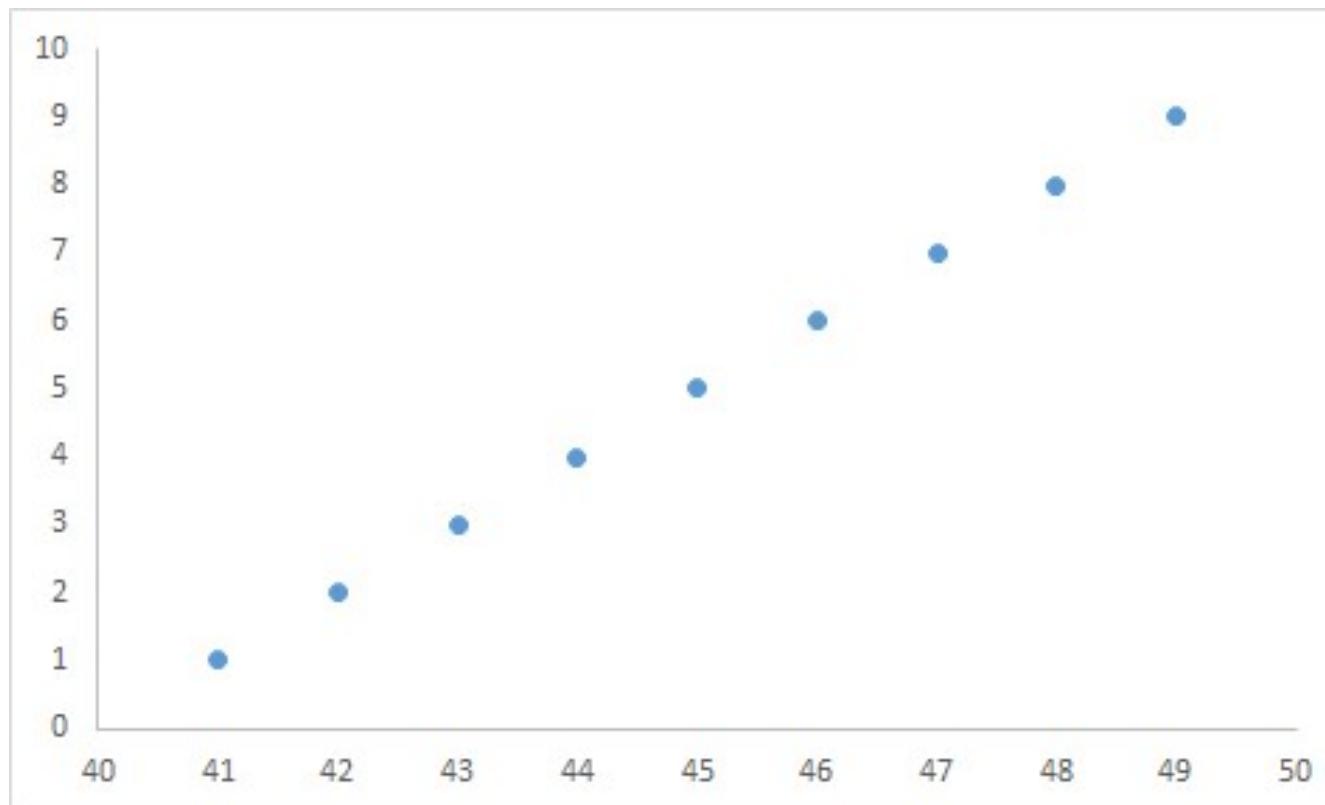


# Scales

# Scales- Types

## Linear Scale

- ▶ Visual spacing between each of the data points is the same regardless where the data points are on the axis



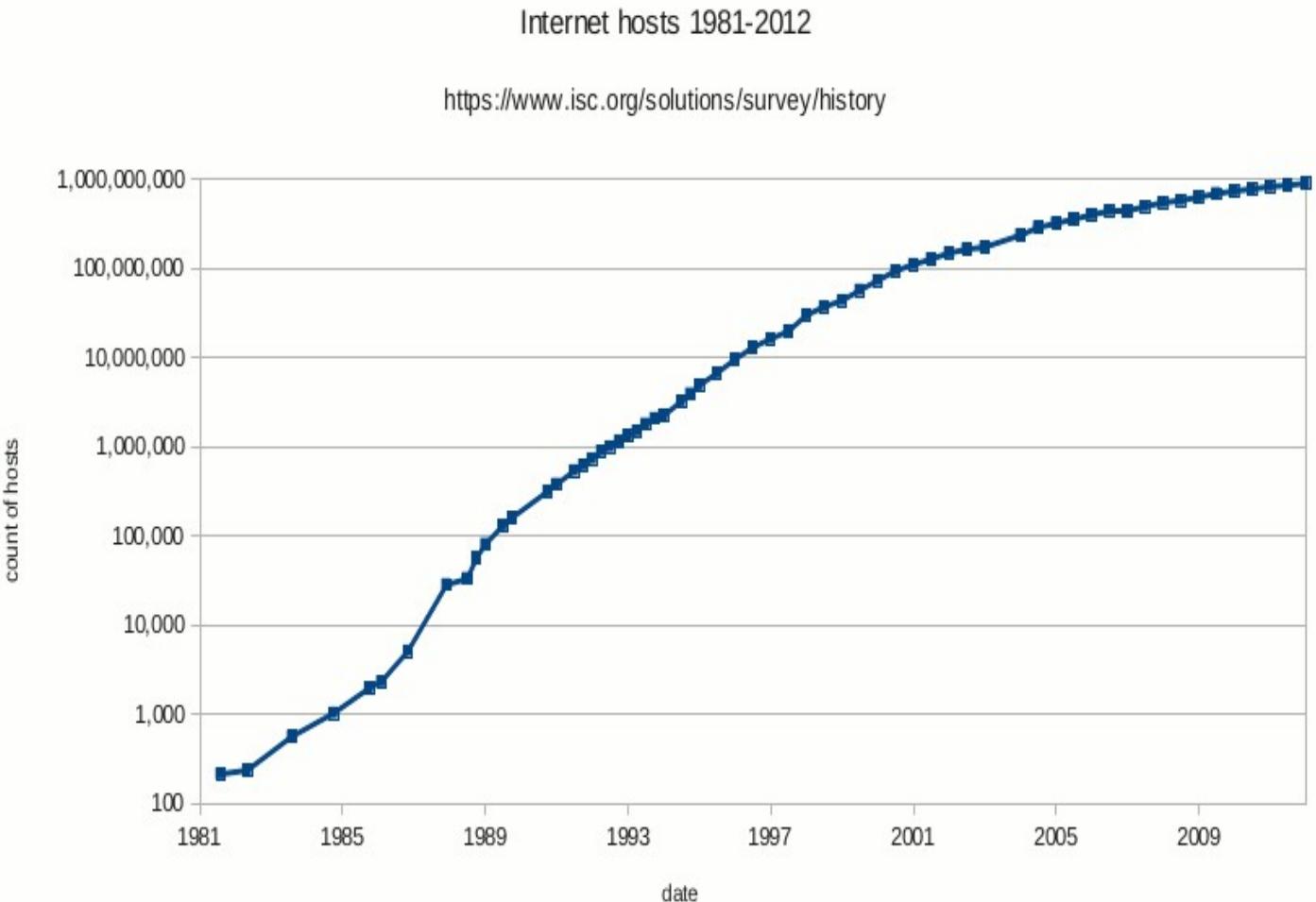
Source:

[https://www.google.com/search?q=examples+of+linear+scales&rlz=1C1GCEA\\_enSG969SG969](https://www.google.com/search?q=examples+of+linear+scales&rlz=1C1GCEA_enSG969SG969)

# Scales- Types

## Logarithmic scale

- ▶ Takes very large, exponentially growing numbers and displays them in a way that is easier for the brain to understand
- ▶ Is used when numbers multiply by a factor larger than 2 from one time interval to another



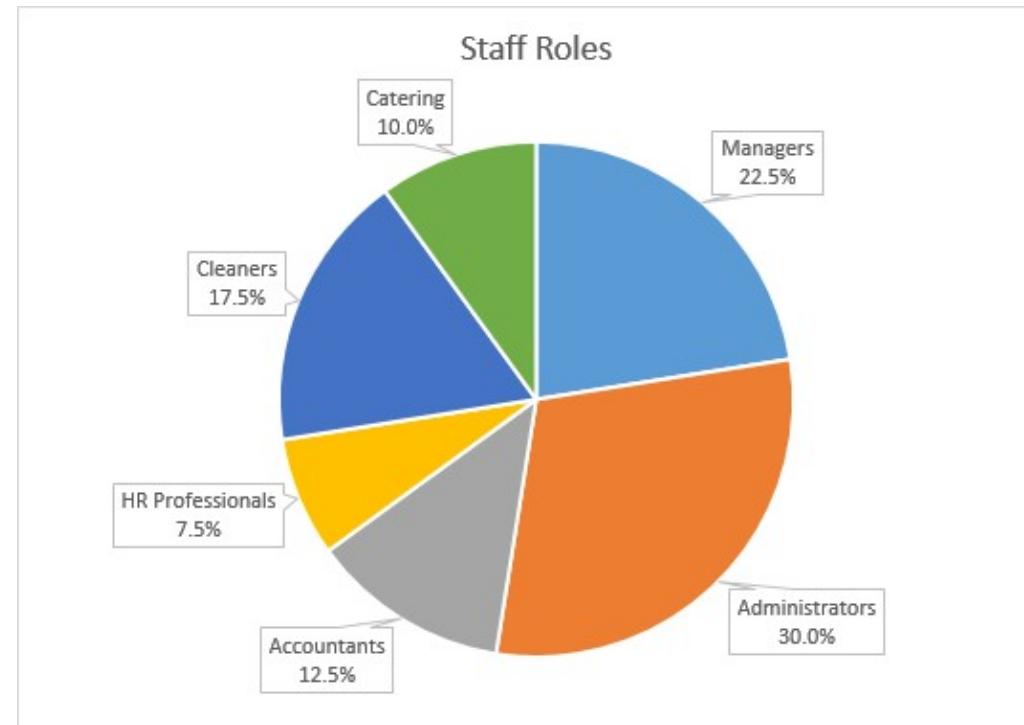
Source:

[https://en.wikipedia.org/wiki/Logarithmic\\_scale](https://en.wikipedia.org/wiki/Logarithmic_scale)

# Scales- Types

## Percent scale

- ▶ It is used to represent part of the whole data, its maximum is 100 percent



Source:  
<https://www.skillsyouneed.com/num/percentages.html>

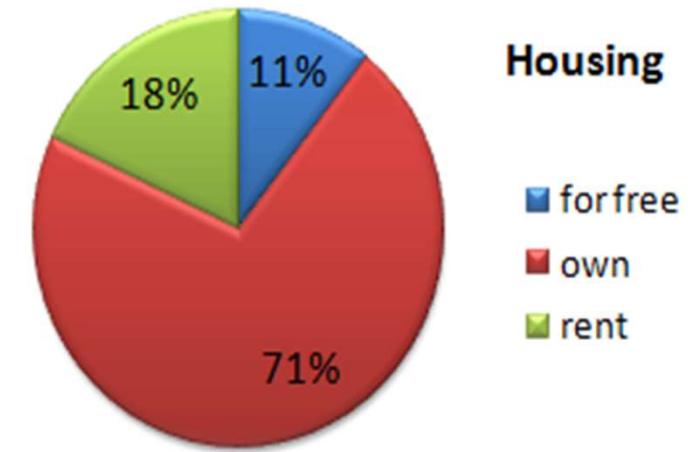
# Scales- Types

## Categorical scale

Example: The housing variable with three categories-

- for free
- own and
- rent

Frequency Table		
Housing	Count	Count%
for free	96	10.67%
own	641	71.22%
rent	163	18.11%



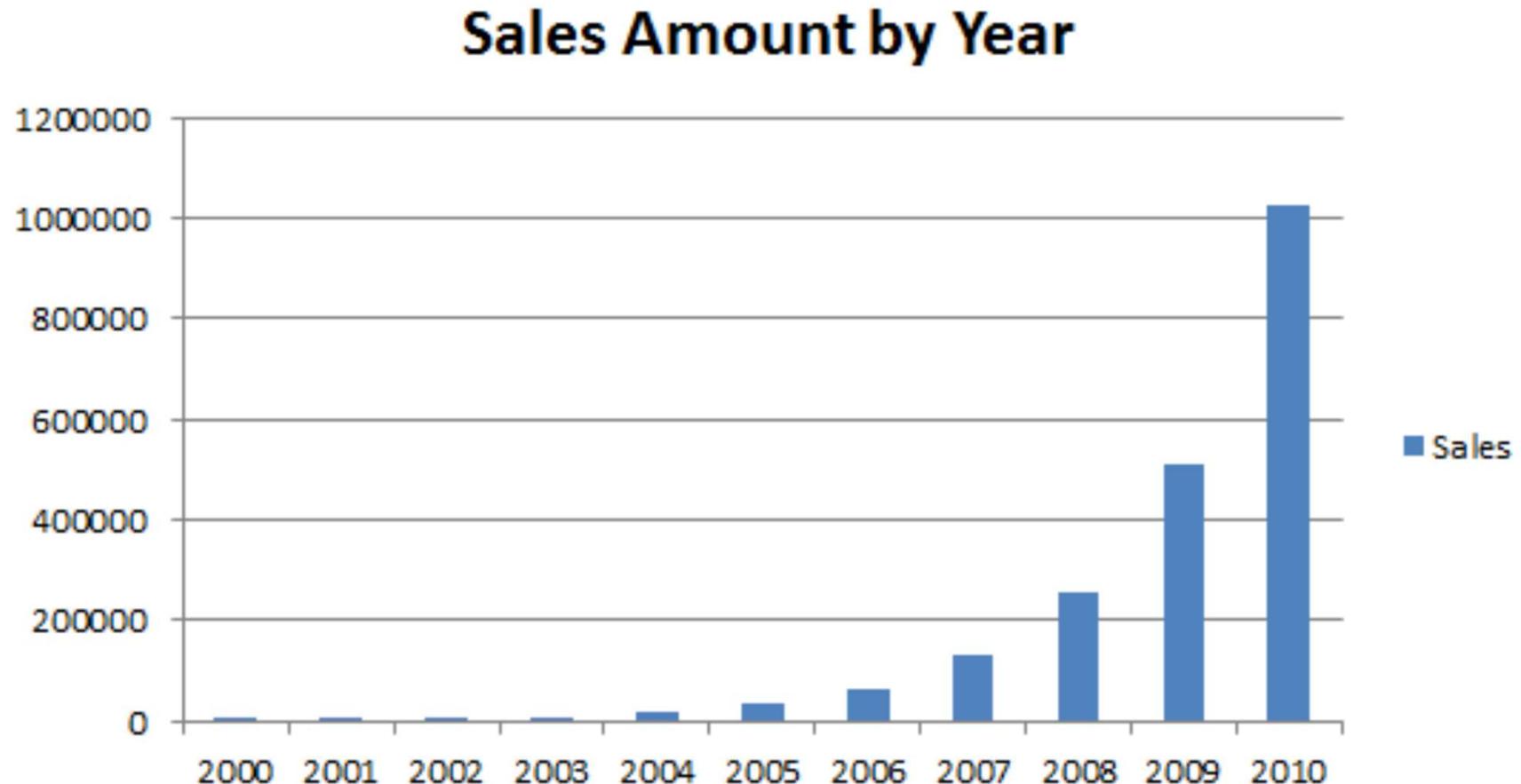
Source:

[https://www.saedsayad.com/categorical\\_variables.htm](https://www.saedsayad.com/categorical_variables.htm)

# Scales- Types

## Time scale

- ▶ Used to plot temporal data on a linear scale, or to divide the temporal data on a categorical scale, such as by year, month or day





# Context

# Context

## The big idea

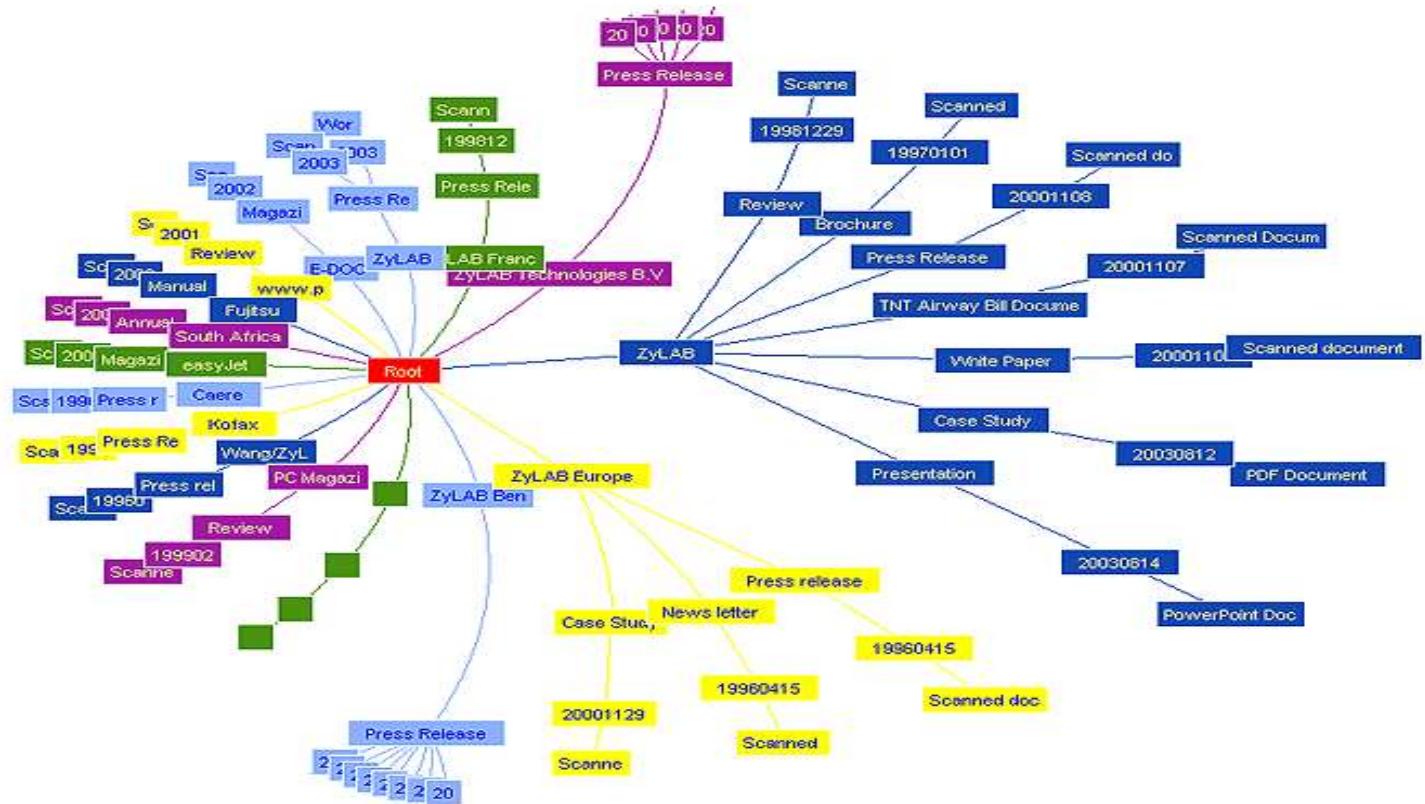
- ▶ *Context* is a data visualisation component that lends to better understanding of who, what, when, where and why of the data
- ▶ *Focus-context problem in data visualization*
  1. The viewer needs both overview and details of the information simultaneously
  2. Information needed in the overview may be different from that needed in the detail
  3. Need to combine both types within a single interactive data visualisation

# Context

# Solving the focus-context problem — distortion

- ▶ Spatially distorts the data presentation to give more room to the designated points of interest, and to decrease the space given to regions away from those points

# Hyperbolic Tree Browser

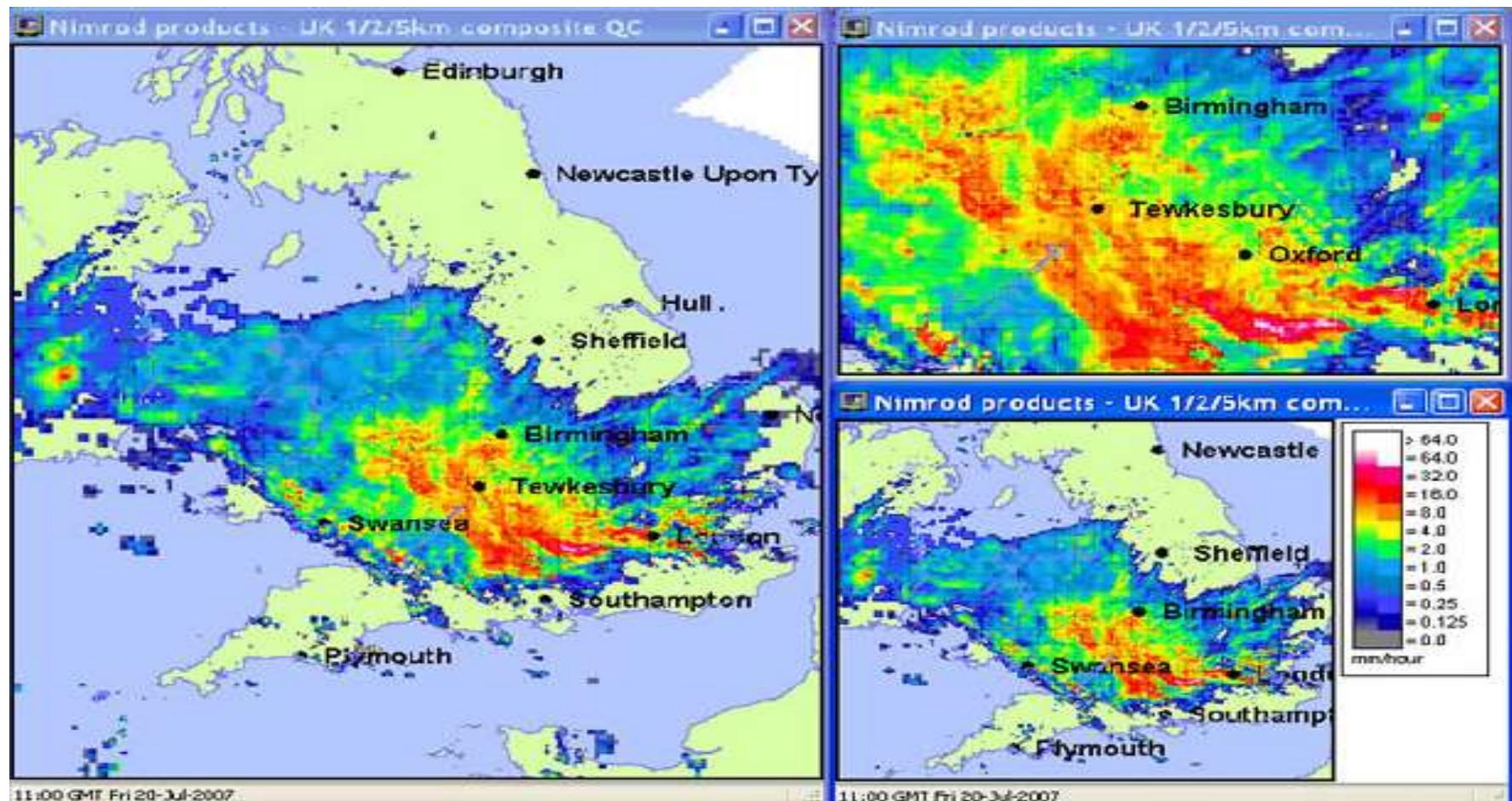


Source: [https://zylab.files.wordpress.com/2010/09/hyperbolic\\_tree.png](https://zylab.files.wordpress.com/2010/09/hyperbolic_tree.png)

# Context

Solving the focus-context problem — rapid zooming

- ▶ Allows viewers to zoom rapidly in and out of points of interest

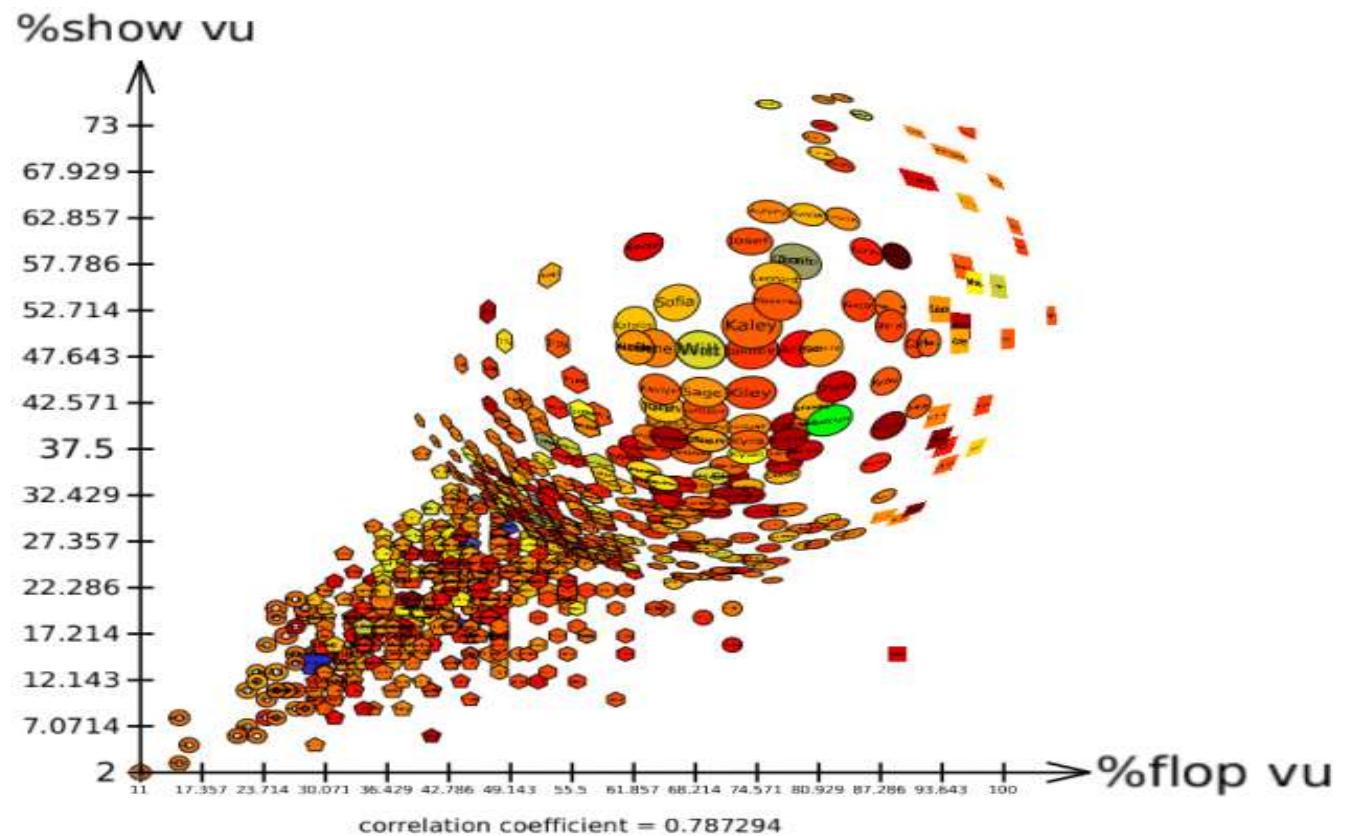


# Context

Solving the focus-context problem — elision

## Elision – Fish Eye Technique

- ▶ Hides parts of a structure from viewers until they are needed



# Context

## Solving the focus-context problem — multiple windows

- ▶ Allows viewers to have one window that shows an overview of the data, and several other windows that show the expanded details

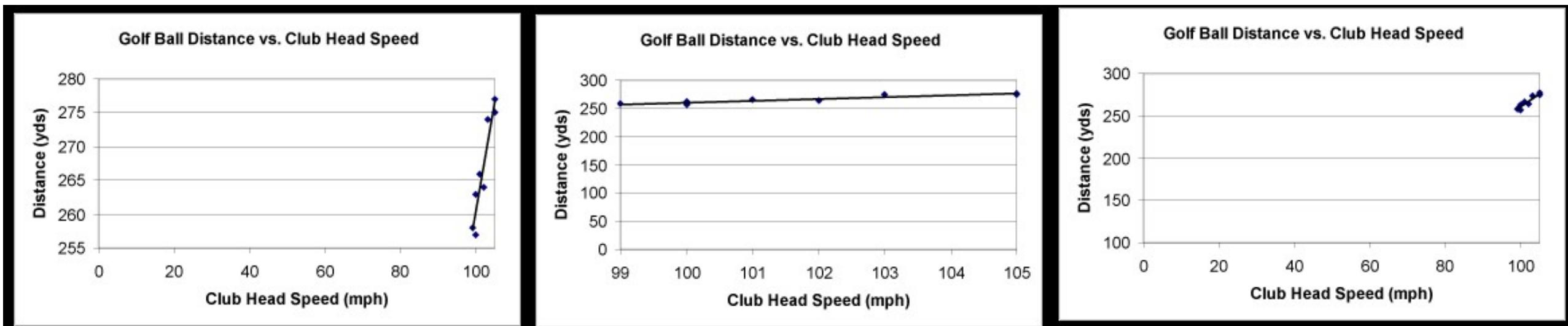




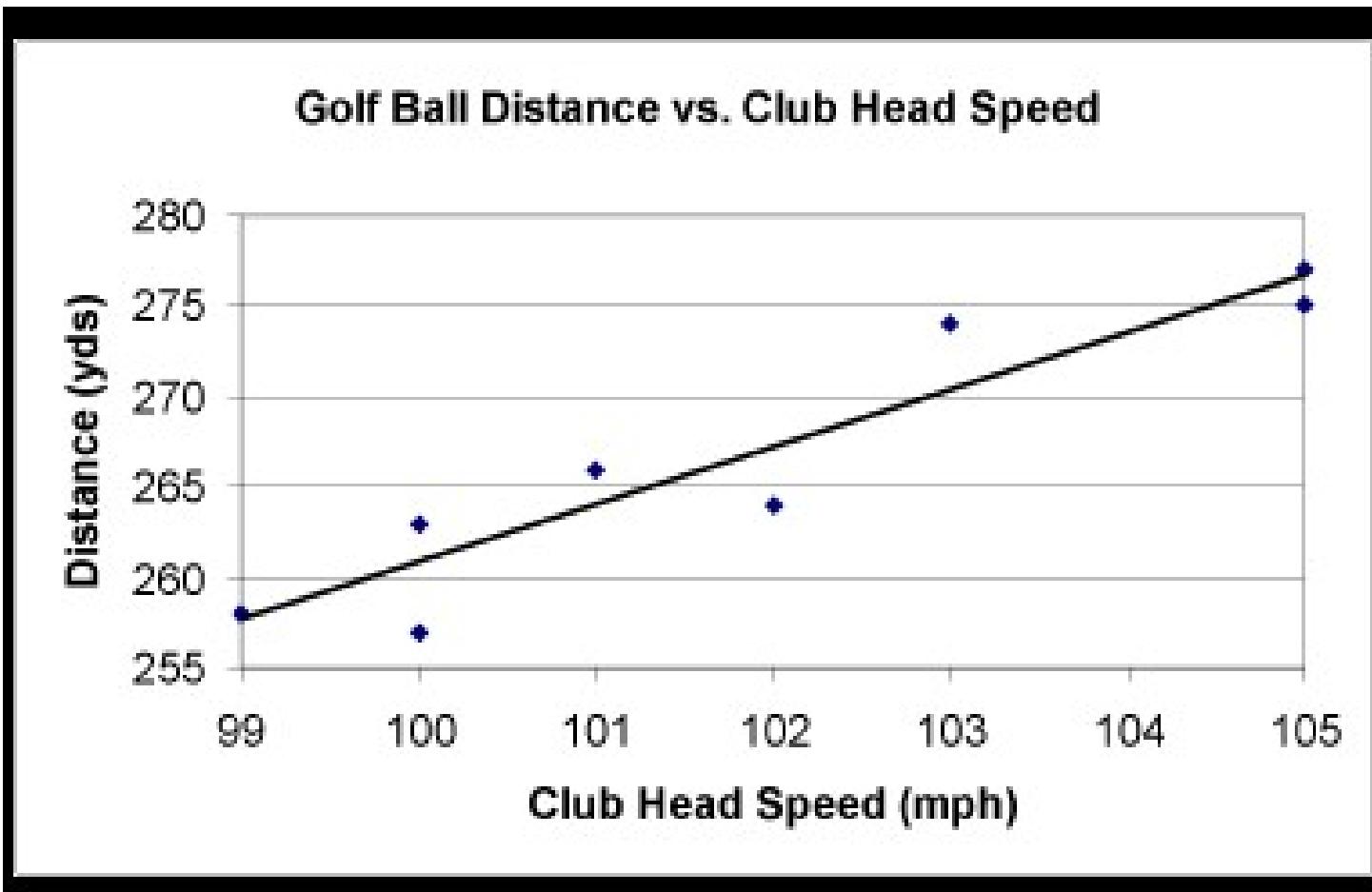
# Class Activity: Data Visualisation Evaluation

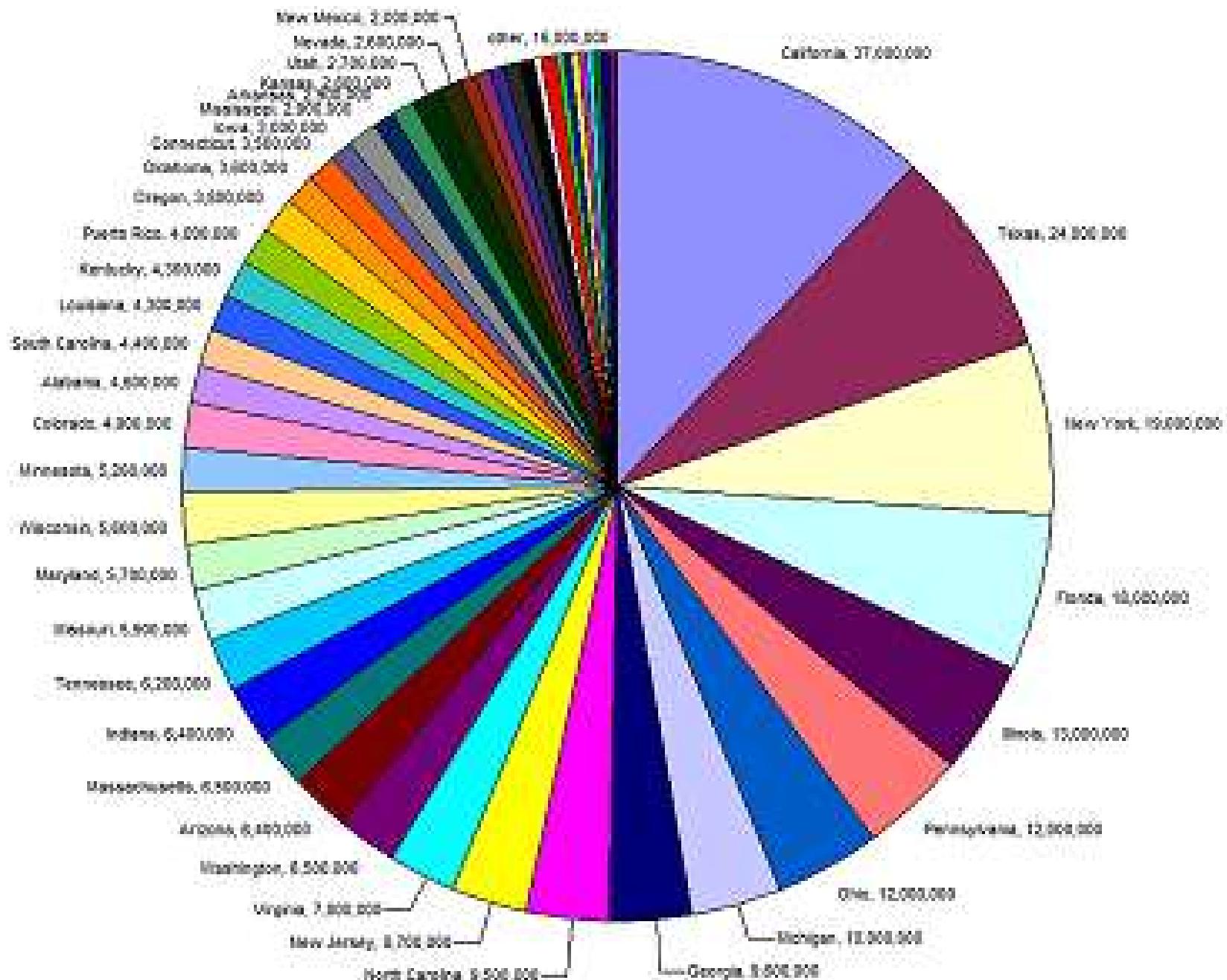
# Class Activity Task

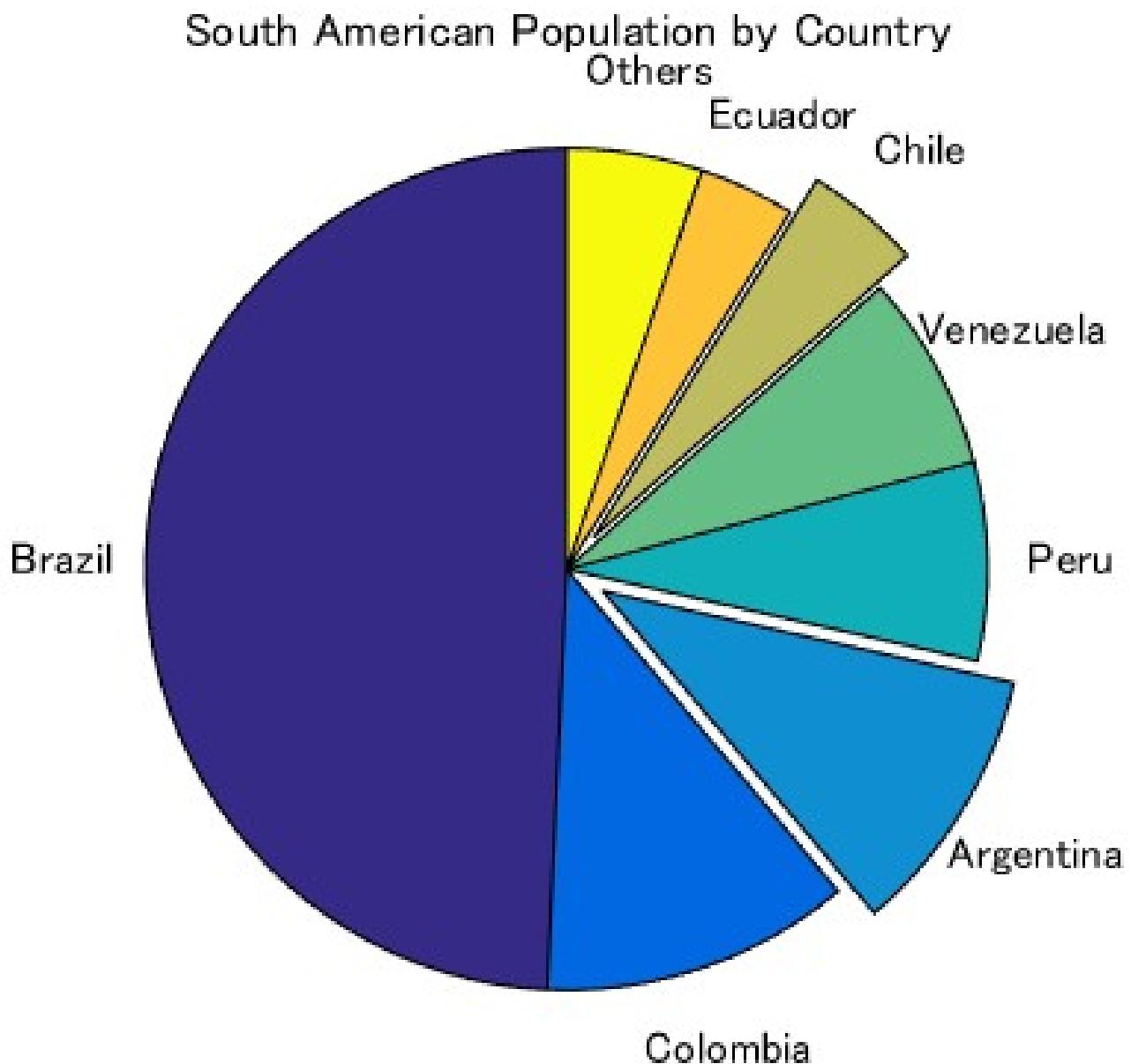
1. Sit with your GBA's team mates (optional)
2. Discuss why the plots/charts are bad visualisation example and suggest ways to improve the plots/charts.



# Class Activity Task







# Tableau (Class Activity)

# Tableau (Class Activity)

- ▶ Sit in your GBA groups (optional)
- ▶ Ensure that you have a working copy of Tableau Desktop installed on your computer
- ▶ 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

# Table Join

The screenshot shows the Power BI desktop interface. On the left, the 'Connections' pane lists a single connection to 'Global Superstore 2016' (Microsoft Excel). Below it, the 'Sheets' pane contains three tables: 'Orders', 'People', and 'Returns'. A 'New Union' section is also present. In the center, a 'Join' dialog box is open, showing two data sources: 'Orders' and 'Returns'. The 'Join' tab is selected, displaying four join types: Inner, Left, Right, and Full Outer. The 'Inner' join type is highlighted. Below the tabs, the 'Data Source' is set to 'Order ID' and the 'Returns' table is set to 'Order ID (Retu...'. A note says 'Add new join cla...'. At the bottom of the dialog, there are 'Sort' and 'OK' buttons. To the right of the dialog, a preview table is shown with columns: Row ID, Order ID, Order Date, Shipping Date, Ship Mode, and City. The preview data includes rows for order IDs 30191, 30190, and 25438.

1. Connect to data: Global Superstore 2016
2. Drag Orders to right side
3. Double click Orders
4. Drag Returns to right side, next to Orders
5. The default join type is inner join. Order ID is set to be the key for join

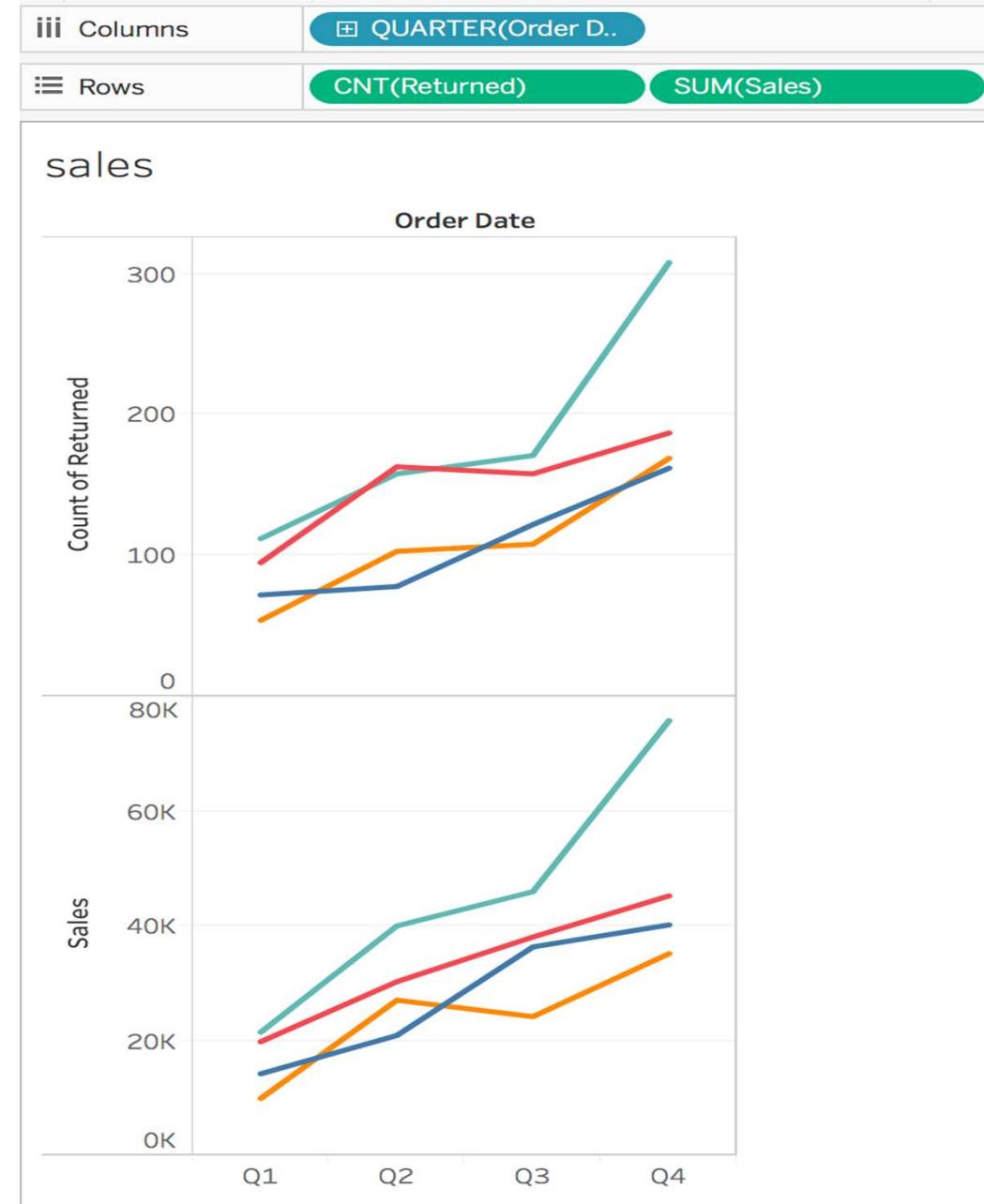
# Table Join

Join Type	Result
Inner 	<p>The resulting table contains values that have matches in both source tables. When a value doesn't match across both source tables, it is dropped entirely.</p>
Left 	<p>The resultant table contains all values from the first source table and corresponding matches from the second source table.</p> <p>When a value in the first source table doesn't have a corresponding match in the second source table, you see a null value in the resulting table.</p>
Right 	<p>The resultant table contains all values from the second source table and corresponding matches from the first source table.</p> <p>When a value in the second source table doesn't have a corresponding match in the first source table, you see a null value in the resulting table.</p>
Full outer 	<p>The resultant table contains all values from both the source tables.</p> <p>When a value from either source table doesn't have a match with the other table, you see a null value in the resultant table.</p>



# Table Join

1. Order date → Columns
2. Click “+” sign before Year, and move Year to color
3. Returned → Rows, Measure → count
4. Sales → Rows



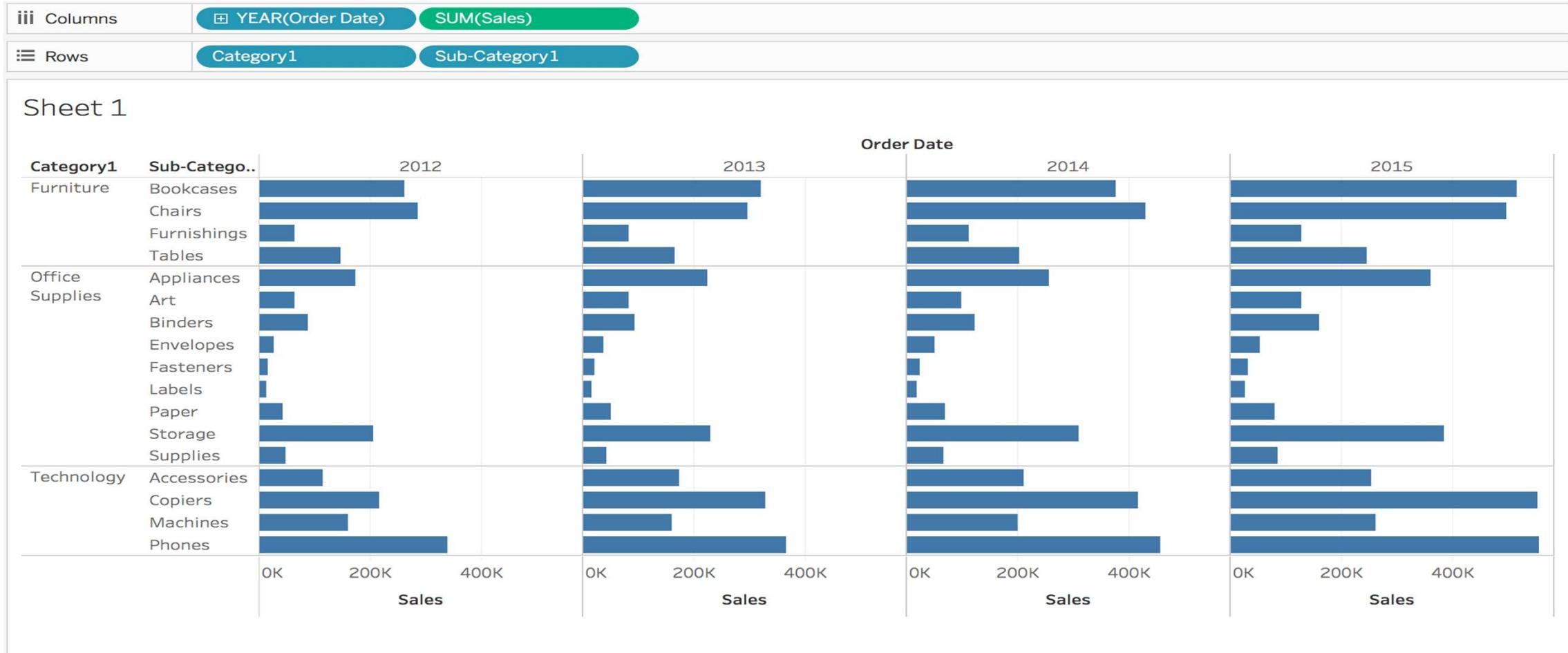
# Cross-database Join

1. Go to home page, connect to Sales 2016
2. Go to Data Source, click “Add” under connections
3. Text Files → Products 2016
4. Double click “Sheet1” on the right
5. Drag “Products 2016.csv” to the right
6. Go to a new worksheet

The screenshot shows the Power BI Data Source interface. On the left, under 'Connections', 'Sales 2016' (Microsoft Excel) is selected. Under 'Sheets', 'Sheet1' is listed. On the right, a 'Sales and Products 2016' connection is shown with two data sources: 'Sales' and 'Products 2016.csv'. A blue circle highlights the join icon between them. Below the connection pane, there's a 'Sort fields' button and a 'Data source order' dropdown. A preview table displays five columns: Category1, Sub-Category1, Product Name, Row ID, and Order ID. The first two rows of the preview table are:

Category1	Sub-Category1	Product Name	Row ID	Order ID
Furniture	Bookcases	Atlantic Metals Mobi...	34320	CA-2015
Furniture	Bookcases	Atlantic Metals Mobi...	33522	CA-2014

# Cross-database Join



# Data Blending

1. Go to home page, connect to Office City
2. Data Source: click the highlighted button
3. Excel → Coffee Chain
4. Go to a new worksheet

The screenshot shows the Tableau desktop application. On the left, the 'Connections' pane lists 'Office City Microsoft Excel'. Below it, the 'Sheets' pane shows 'Purchases' selected. A tooltip is displayed over the 'Office City' connection in the connections list, highlighting the 'New Data Source' option. At the bottom, a preview of the blended data is shown in a grid:

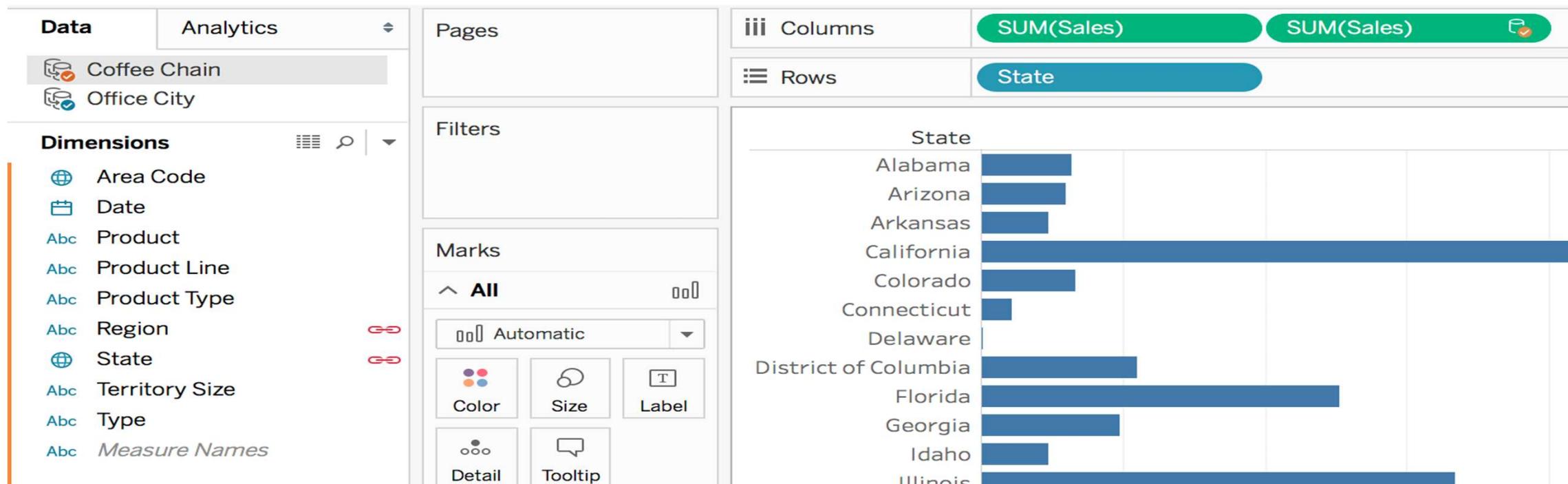
Customer Name	Goods Category	Goods Sub-Categ.
John Doe	Electronics	Smartphones
Jane Smith	Office Supplies	Labels
Mike Johnson	Books	Textbooks
Nancy Lee	Food	Snacks
Peter Williams	Clothing	Shirts
Quinn Brown	Household	Cleaning Products
Ryan Green	Automotive	Tires
Samantha White	Books	Textbooks
Tommy Black	Electronics	Smartphones
Ursula Grey	Office Supplies	Labels
Vincent Red	Books	Textbooks
Wendy Blue	Electronics	Smartphones
Xavier Purple	Office Supplies	Labels
Yvonne Orange	Books	Textbooks
Zachary Green	Electronics	Smartphones

# Data Blending

Discuss to identify:

- Primary and Secondary data sources

- 1.Data: Office City, Sales (Office City) → Columns
- 2.State (Office City) → Rows
- 3.Data: Coffee Chain, Sales (Coffee Chain)→ Columns



# Pivot Data from Columns to Rows

Pivot from wide format 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

To long format, date as rows:

1. Select all the date columns
2. Click “Pivot”
3. Rename the columns

#	#	
Resolved Incidents	Resolved Incidents	
01/09/2015	01/10/2015	
2	0	
2	1	
12	13	
0	1	

A context menu is open over the third row of the table, showing options: Rename, Copy Values, Hide, Create Calculated Field..., Pivot, and Merge Mismatched Fields.

Date	#	Abc
	Pivot	Resolved Incidents
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

# Split

Split “Employee” column:

1. Select “Employee” column, select “Custom Split”
2. Use the separator “-”, and split off “All”
3. Rename the new columns

A screenshot of a data grid interface. A context menu is open over the 'Employee' column header. The menu items are: Rename, Copy Values, Hide, Aliases..., Create Calculated Field..., Create Group..., Split, Custom Split..., Add Data to Pivot, and Describe... . The 'Custom Split...' option is highlighted.

Abc Resolved Inciden...	=Abc Resolved Incidents	=Abc Calculation	=Abc Calculation	=Abc Calculation
Employee	Employee	Location	Employee ID	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		

Custom Split



How should this data be split?

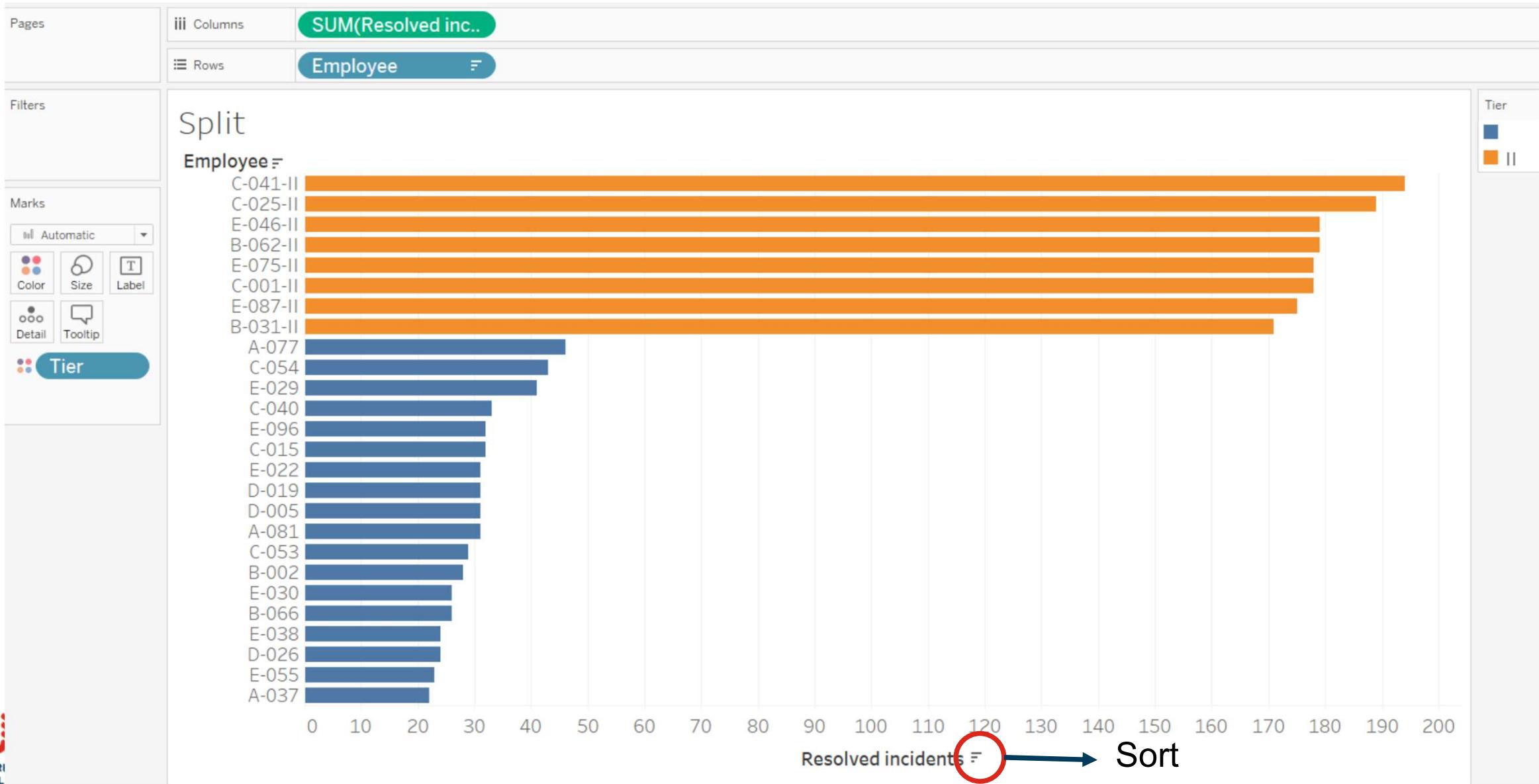
Use the separator -

Split off All 1 columns

OK

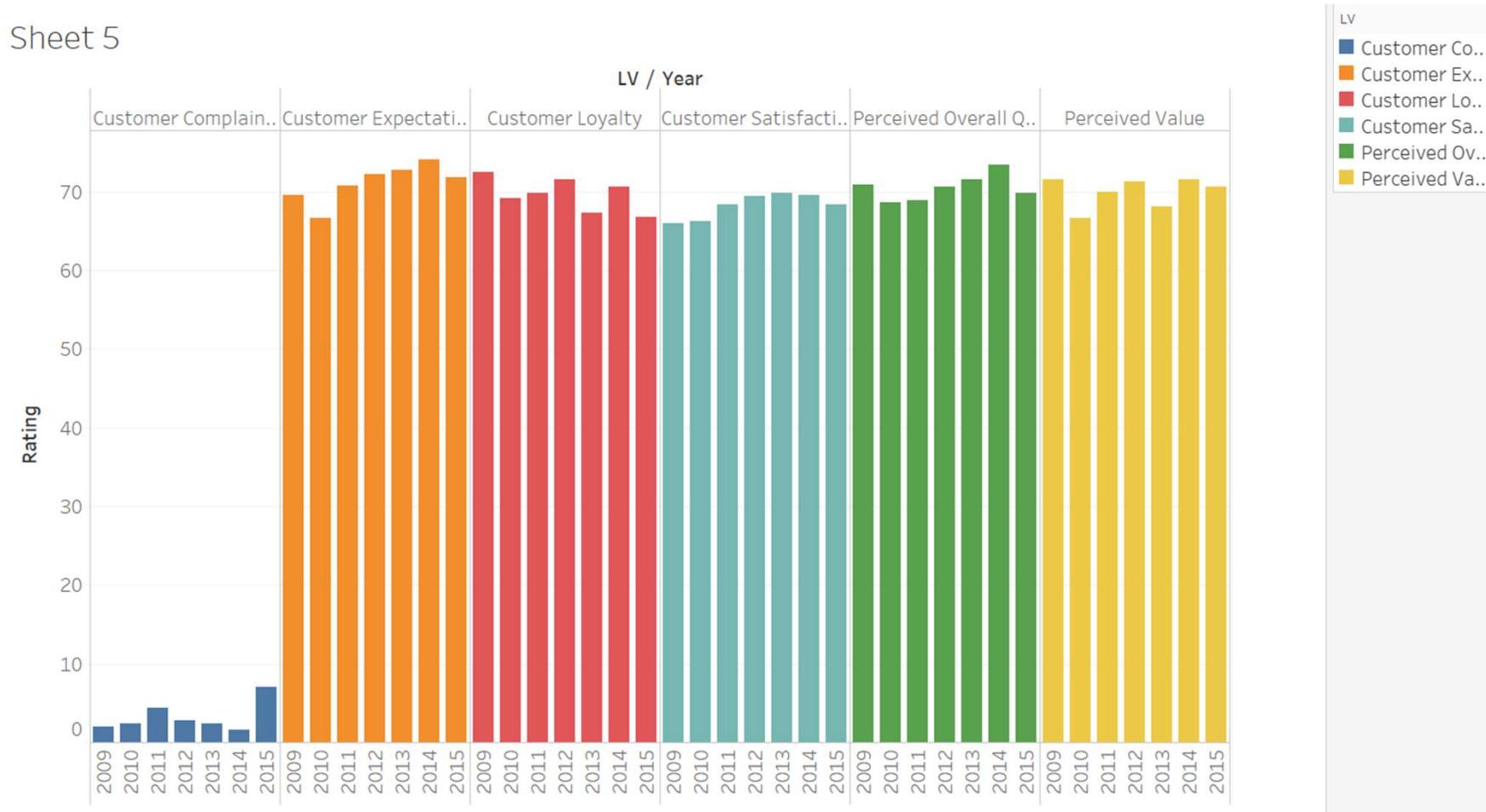
Cancel

# Split



# Exercise

- Use the dataset, `survey_results_wide_format` (`survey_results`)
- Pivot the dataset from wide format to long format
- Create the following chart



# Tableau (Class Activity)

1. Sit with your GBA's team mates (optional)
2. Follow your instructor for the following exercises:
  - Data:
    - global\_superstore\_2016.xlsx (orders)
    - Coffee Chain.xlsx and Office City.xlsx

# Calculation: Aggregate VS Record-Level

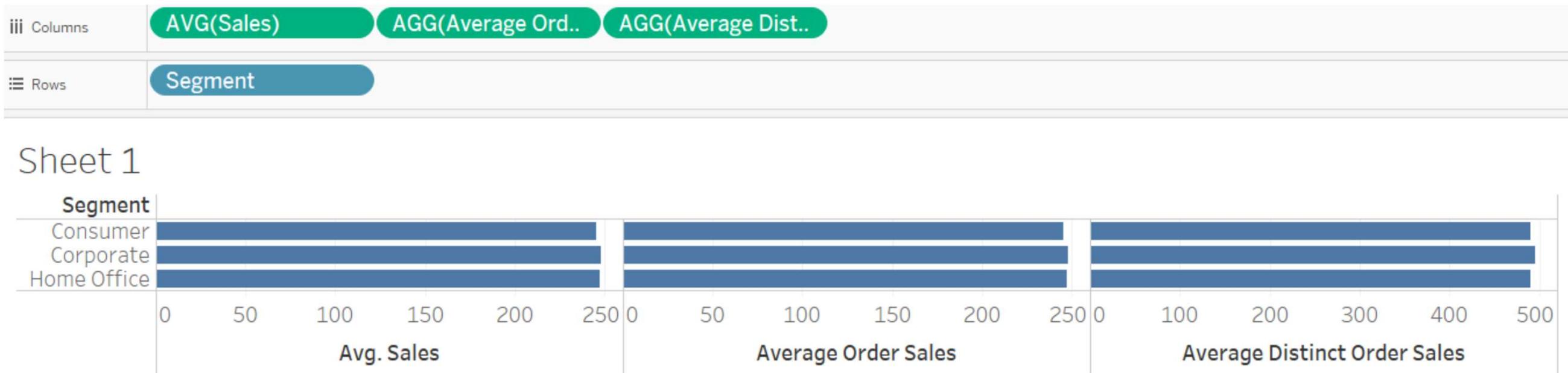
Data: Global\_superstore\_2016 (orders)

Average Order Sales

```
sum([Sales])/count([Order ID])
```

Average Distinct Order Sales

```
sum([Sales])/countd([Order ID])
```



# Create a Calculated Field

The screenshot shows the Tableau Data pane. At the top, there are tabs for "Data" and "Analytics". Below the tabs, a connection named "Orders (global\_superstore\_2016)" is listed. On the right side of the pane, there is a context menu with a red circle highlighting the "Create Calculated Field..." option. To the right of the menu, a modal window is open with the following details:

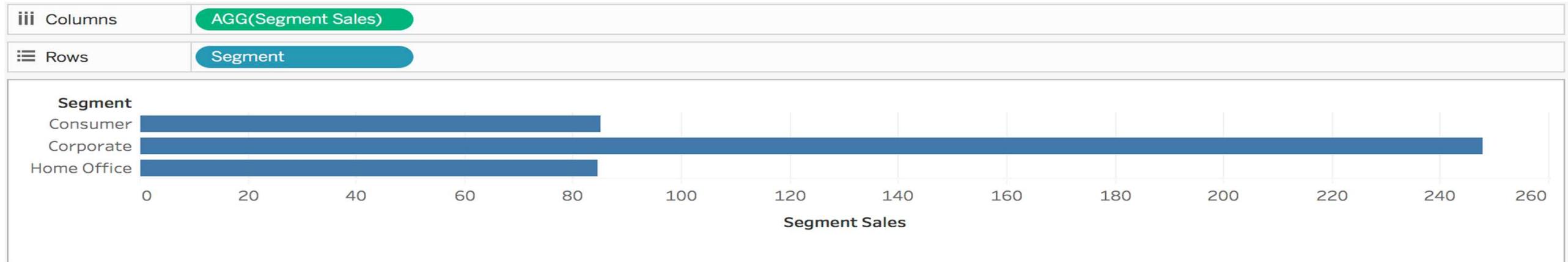
- Title bar: Average Order Sales
- Text input: `sum([Sales])/count([Order ID])`
- Message: "The calculation is valid."
- Status: "1 Dependency" with a dropdown arrow, "Apply" button, and "OK" button.

# 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
```



[More information:](#)

- [https://help.tableau.com/current/pro/desktop/en-us/calculations\\_aggregation.htm](https://help.tableau.com/current/pro/desktop/en-us/calculations_aggregation.htm)
- [https://help.tableau.com/current/pro/desktop/en-us/calculations\\_calculatedfields\\_aggregate\\_create.htm](https://help.tableau.com/current/pro/desktop/en-us/calculations_calculatedfields_aggregate_create.htm)

# Aggregate Functions

## **Attribute:**

```
IF MIN([dimension]) = MAX([dimension]) THEN MIN([dimension]) ELSE "*" END
```

- It can provide a way to aggregate dimensions when computing table calculations, which require an aggregate expression.
- Using Attribute in a table calculation (optional). Example:  
[https://help.tableau.com/current/pro/desktop/en-us/calculations\\_aggregation.htm#aggregating-dimensions](https://help.tableau.com/current/pro/desktop/en-us/calculations_aggregation.htm#aggregating-dimensions)

Segment Sales

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

# Exercise

1. Use the data Office City and Coffee Chain
2. Data blending to show the sales from Office City (primary data) and Coffee Chain
3. Create a calculated field: for each state, calculating combined sales from office city.xlsx and coffee chain.xlsx



# Exercise

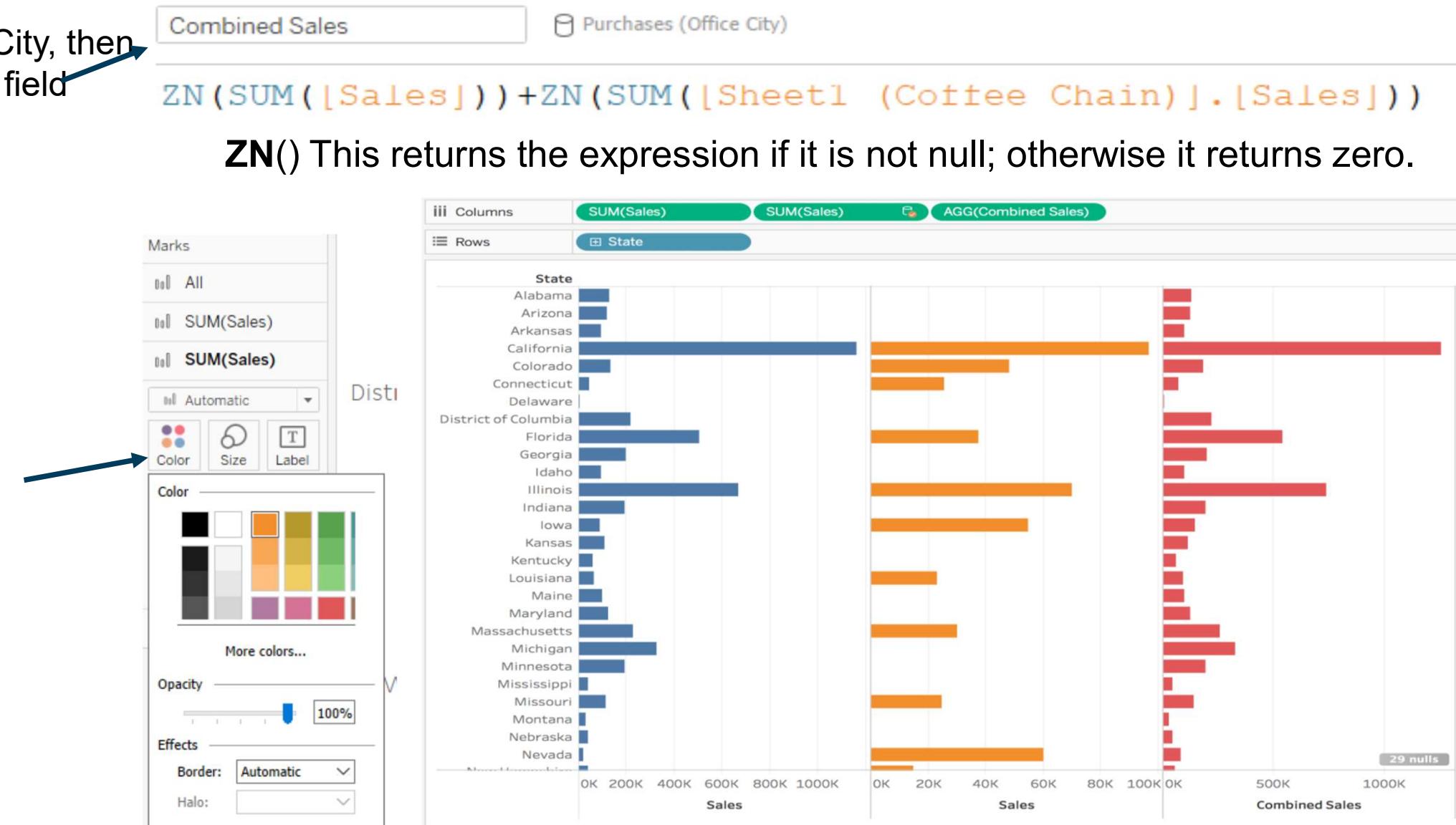
1. Create a calculated field: for each state, calculating combined sales from office city.xlsx and coffee chain.xlsx

Go to Data, Office City, then  
create a calculated field

Combined Sales  
`Purchases (Office City)`  
`ZN(SUM([Sales])) + ZN(SUM([Sheet1 [Coffee Chain]].[Sales]))`

**ZN()** This returns the expression if it is not null; otherwise it returns zero.

Change the  
color in  
“Marks” for  
each of the  
“Sales”  
columns

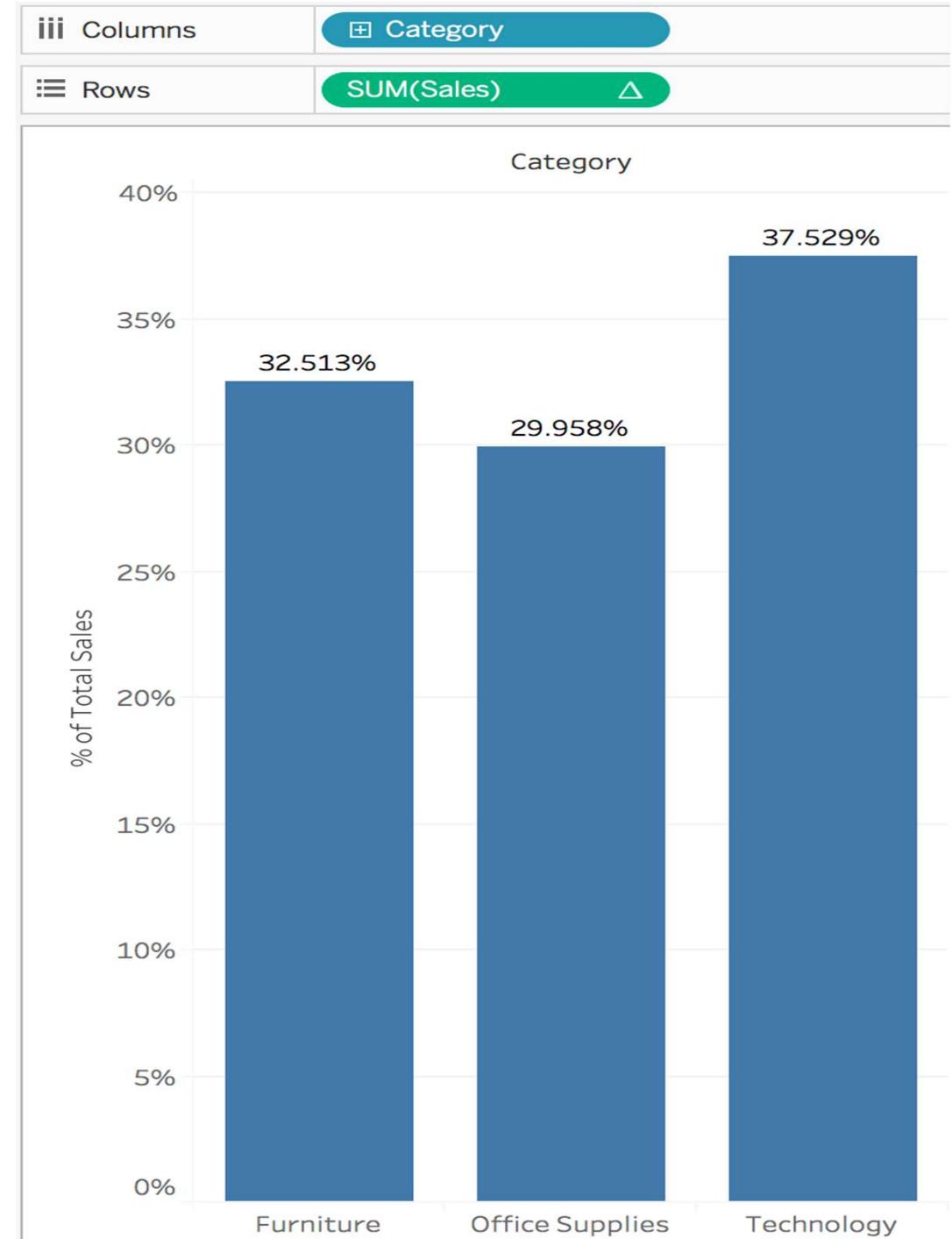


# Quick Table Calculation

## percent of total

Data: Global\_superstore\_2016 (orders)

1. Sales → Rows
2. Category → Columns
3. Go to “Rows” and select drop down menu of “Sum(sales)”.
4. Select “Add Table Calculation” and change Calculation Type to “Percent of total”.
5. Add mark label: press and hold control then drag sum(sales) delta to “Label” in Marks



# Quick Table Calculation

running total: apply when one dimension is time



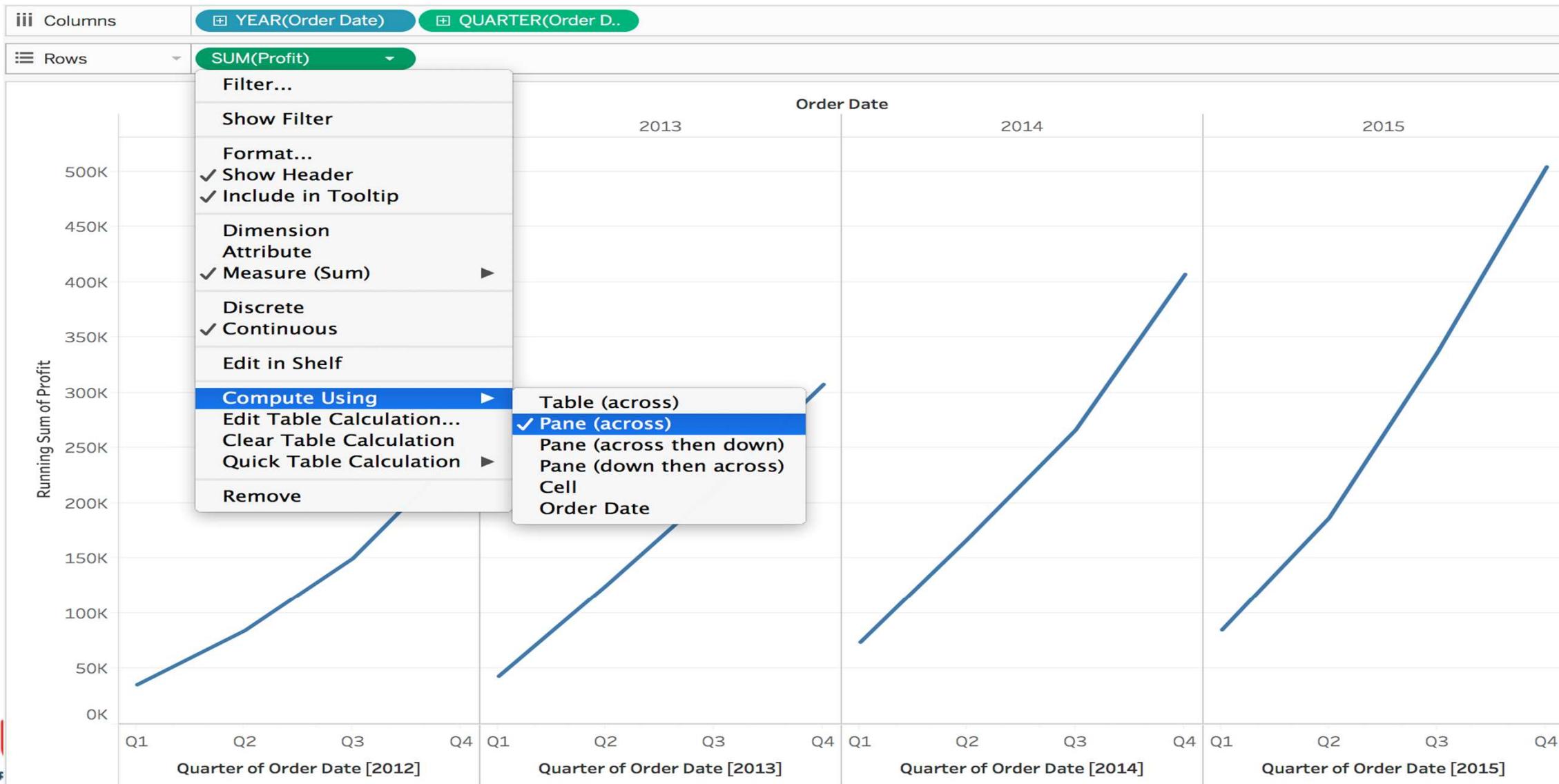
# Quick Table Calculation

running total for each year: Is this chart correct?



# Quick Table Calculation

running total for each year:



# Quick Table Calculation

Compare “Compute Using”:

The screenshot shows the Tableau interface with a data view and a table calculation dialog.

**Data View:**

- Pages:** Shows 'iii Columns' and 'YEAR(Order Dat..)'.
- Filters:** Empty.
- Marks:** Set to 'Automatic'. Includes controls for Color, Size, Text, Detail, and Tooltip. The 'SUM(Profit)' measure is selected.
- Sheet 2:** Displays a table with 'Quarte...' (row) and 'Segment' (column) dimensions. The 'Order Date' column is further subdivided into '2012', '2013', '2014', and '2015'. The data is as follows:

Quarte...	Segment	Order Date			
		2012	2013	2014	2015
Q1	Consumer	15,783	39,239	80,536	127,822
	Corporate	14,220	27,424	45,015	65,094
	Home Office	5,732	12,467	27,585	45,329
Q2	Consumer	25,734	75,010	126,495	180,640
	Corporate	18,423	40,250	69,574	101,722
	Home Office	4,640	15,187	27,814	43,036
Q3	Consumer	34,529	82,576	134,172	201,652
	Corporate	20,048	44,427	75,475	124,363
	Home Office	10,961	25,470	41,619	74,810
Q4	Consumer	41,292	86,312	150,361	239,126
	Corporate	32,056	63,203	110,947	150,029
	Home Office	25,524	44,792	73,697	113,834

**Table Calculation Dialog:**

Table Calculation: Running Sum of Profit

**Calculation Type:** Running Total

**Compute Using:** Table (across)

- Table (across)
- Table (down)
- Table (across then down)
- Table (down then across)
- Pane (down)
- Pane (across then down)
- Pane (down then across)
- Cell

**Specific Dimensions:**

- Year of Order Date
- Quarter of Order Date
- Segment

Restarting every \_\_\_\_\_

Add secondary calculation

Show calculation assistance

# Quick Table Calculation

To have both sales and table calculations:

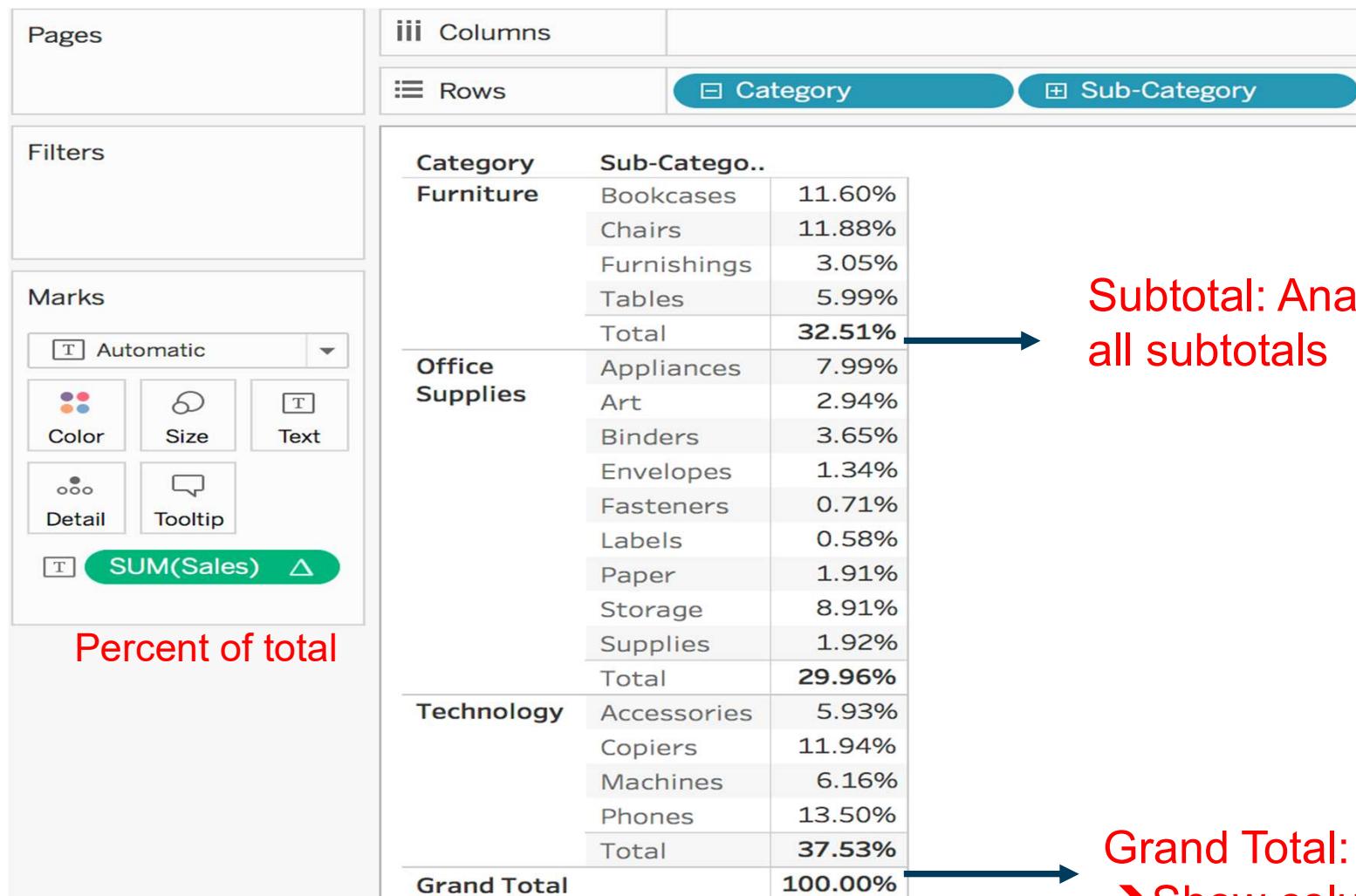
1. Category → Rows
2. Subcategory → Rows
3. Measure Names → Columns
4. Measure Values → Marks (Text)
5. Table calculation for SUM(Sales), percent of total
6. Drag another SUM(Sales) to Measure Values list
7. Remove unwanted measure values

The screenshot shows the Tableau interface with the following components:

- Pages:** Shows "iii Columns" and "Measure Names".
- Filters:** Shows "Measure Names" selected.
- Marks:** Shows "Automatic" selected, with options for Color, Size, Text, Detail, and Tooltip.
- Measure Values:** Shows "Measure Values" selected, with two entries: "SUM(Sales)" and another "SUM(Sales)" below it.
- Sheet 7:** Displays a table with the following data:

Category	Sub-Catego..	Sales	% of Total S..
Furniture	Bookcases	1,466,572	11.60%
	Chairs	1,501,682	11.88%
	Furnishings	385,156	3.05%
	Tables	757,042	5.99%
	Total	4,110,452	32.51%
Office Supplies	Appliances	1,010,536	7.99%
	Art	371,613	2.94%
	Binders	461,869	3.65%
	Envelopes	169,217	1.34%
	Fasteners	89,495	0.71%
	Labels	73,350	0.58%
	Paper	241,788	1.91%
	Storage	1,126,813	8.91%
	Supplies	242,811	1.92%
Total	3,787,493	29.96%	
Technology	Accessories	749,237	5.93%
	Copiers	1,509,436	11.94%
	Machines	779,060	6.16%
	Phones	1,706,824	13.50%
	Total	4,744,557	37.53%
<b>Grand Total</b>	<b>12,642,502</b>	<b>100.00%</b>	

# Exercise



Subtotal: Analysis → Totals → Add all subtotals

Grand Total: Analysis → Totals  
→ Show column grand totals

More Information:

[https://help.tableau.com/current/pro/desktop/en-us/calculations\\_tablecalculations.htm](https://help.tableau.com/current/pro/desktop/en-us/calculations_tablecalculations.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 <u>local data</u> 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 from the source.