Best Practices - Use the right Charts

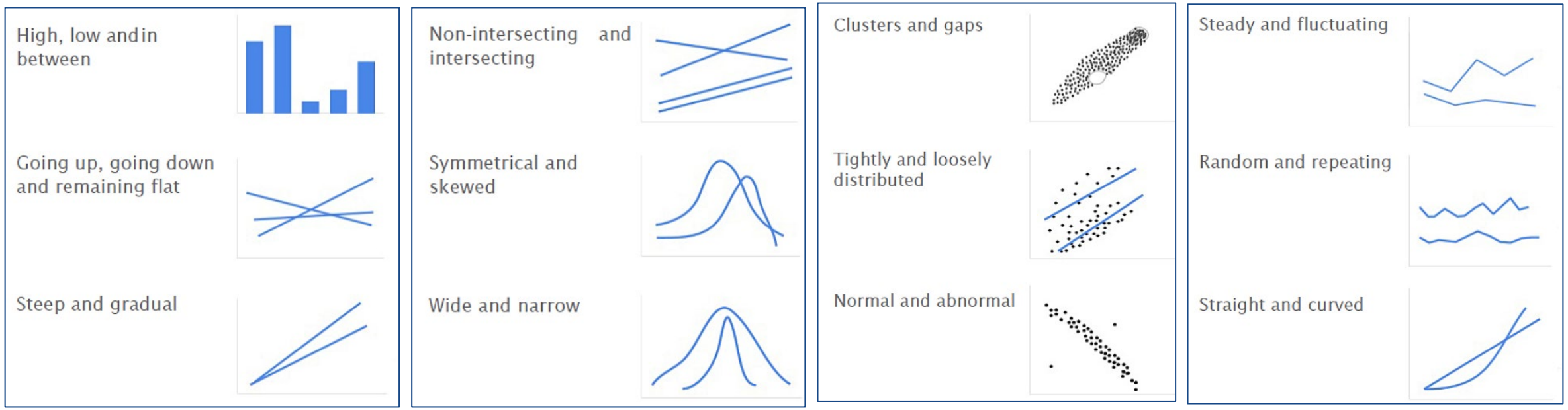
Contents

- Underlying Concepts
- Classification of Charts
- Hands on Workshop

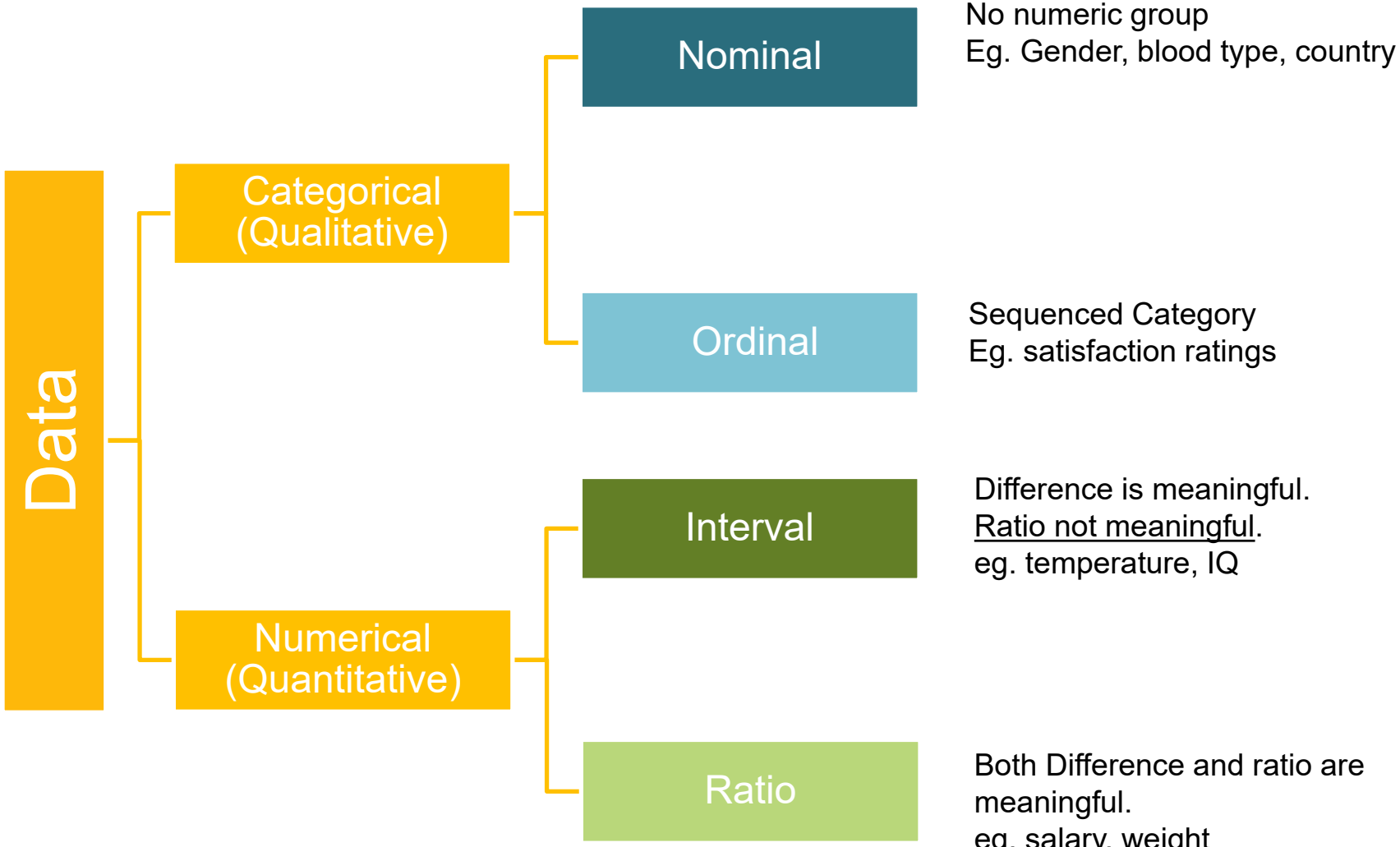
Analytical Patterns

We can use ‘Pre-Attentive Attributes’ to form ‘Analytical Patterns’…

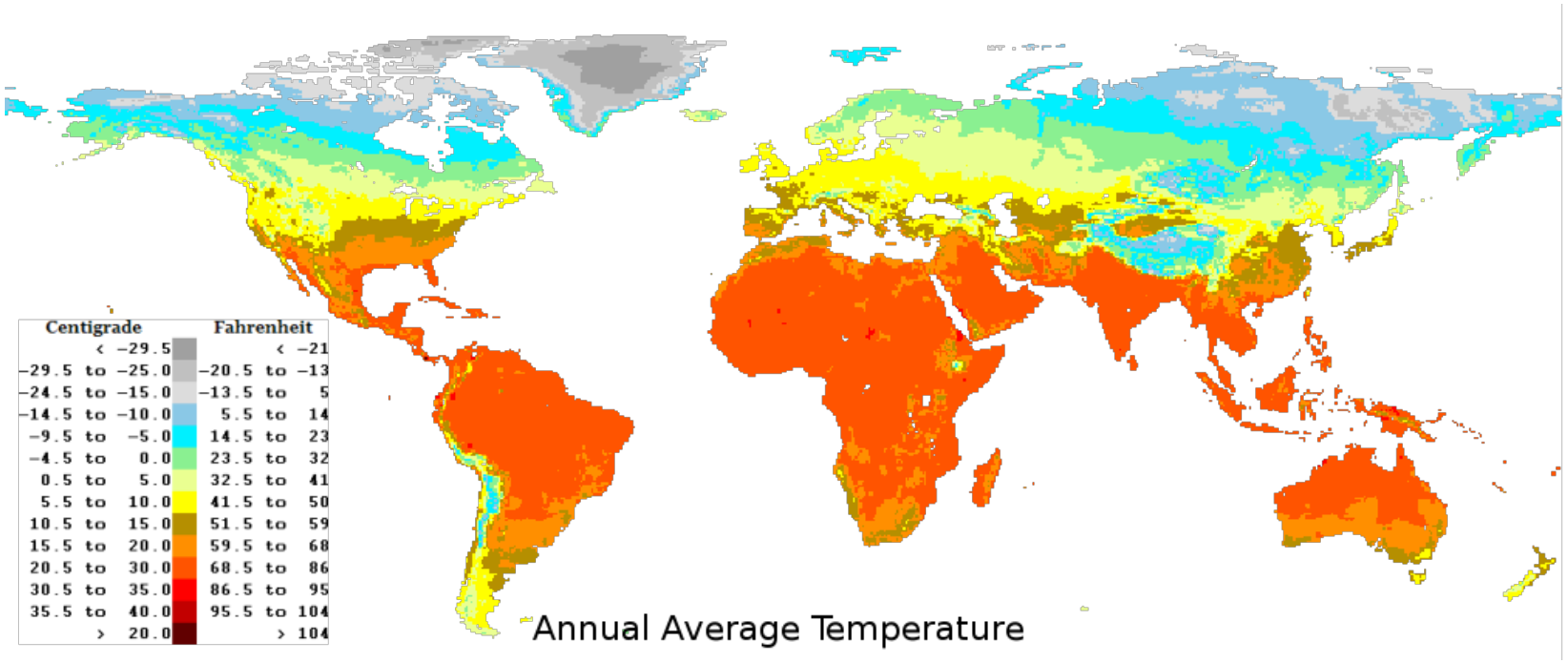
The ‘Analytical Patterns’ provide us with a visual language that allows us to communicate information visually.



Data comes in different shapes

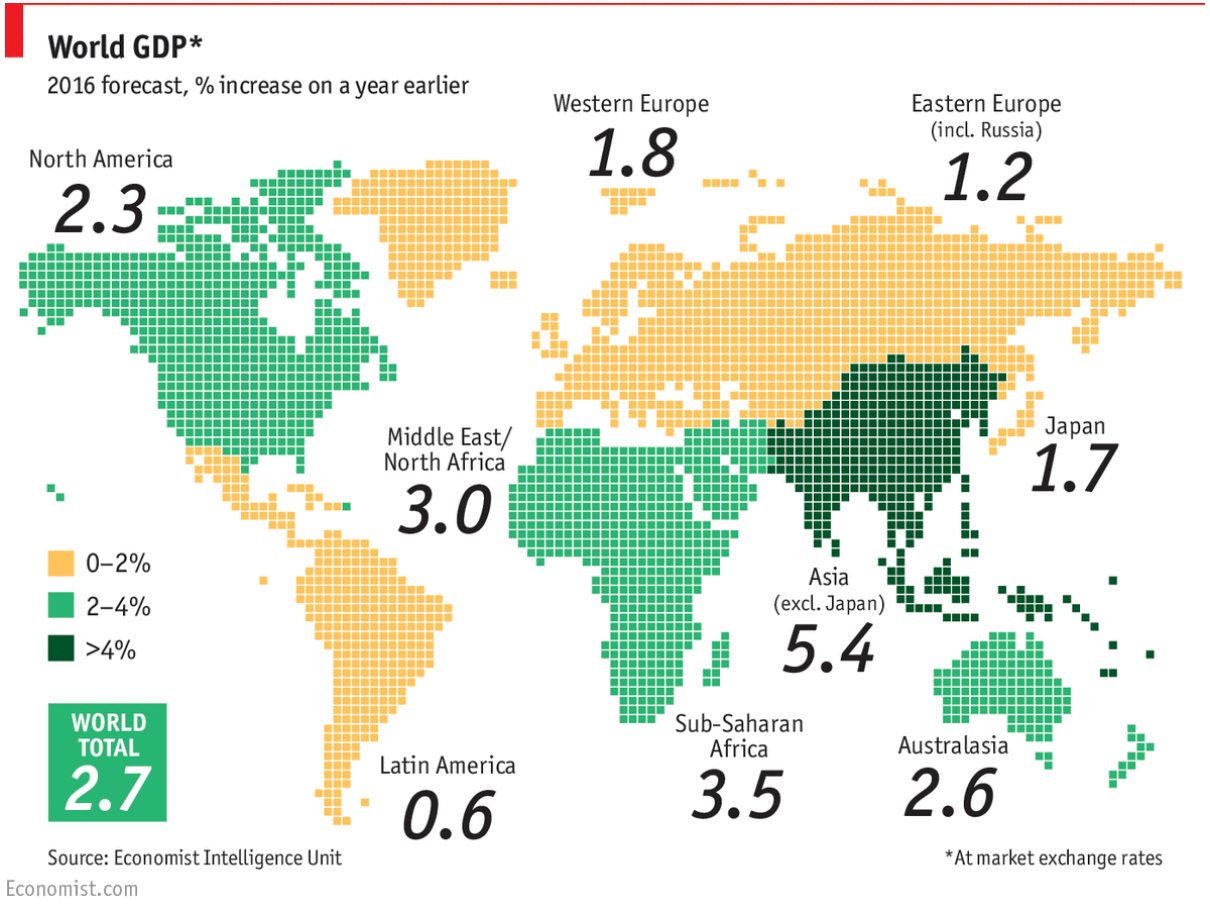


Guess the data type (1/4)



Source: http://www.climate-charts.com/World-Climate-Maps.html#temperature

Guess the data type (2/4)

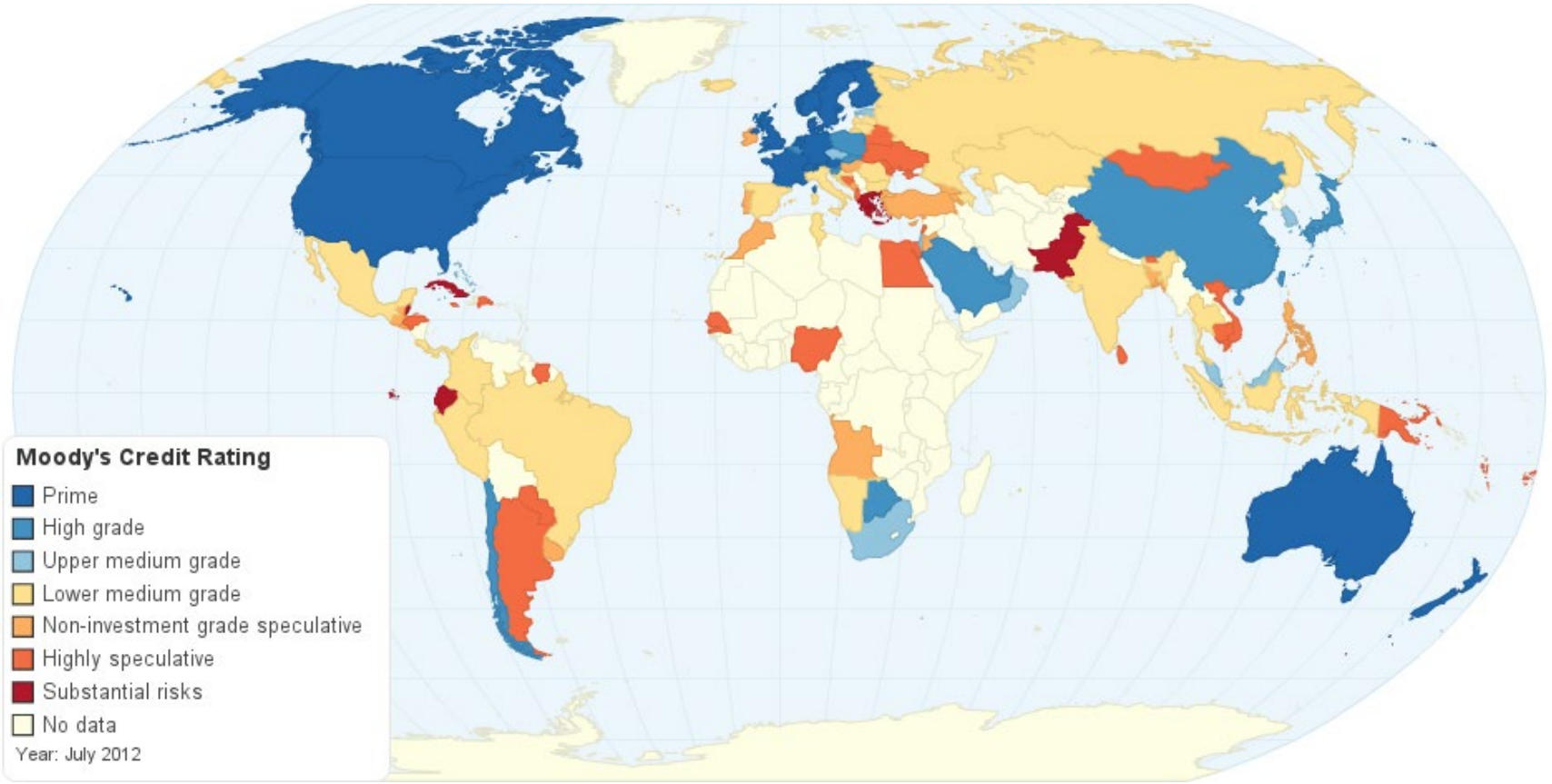


Source: https://www.economist.com/blogs/graphicdetail/2015/12/daily-chart

Guess the data type (3/4)

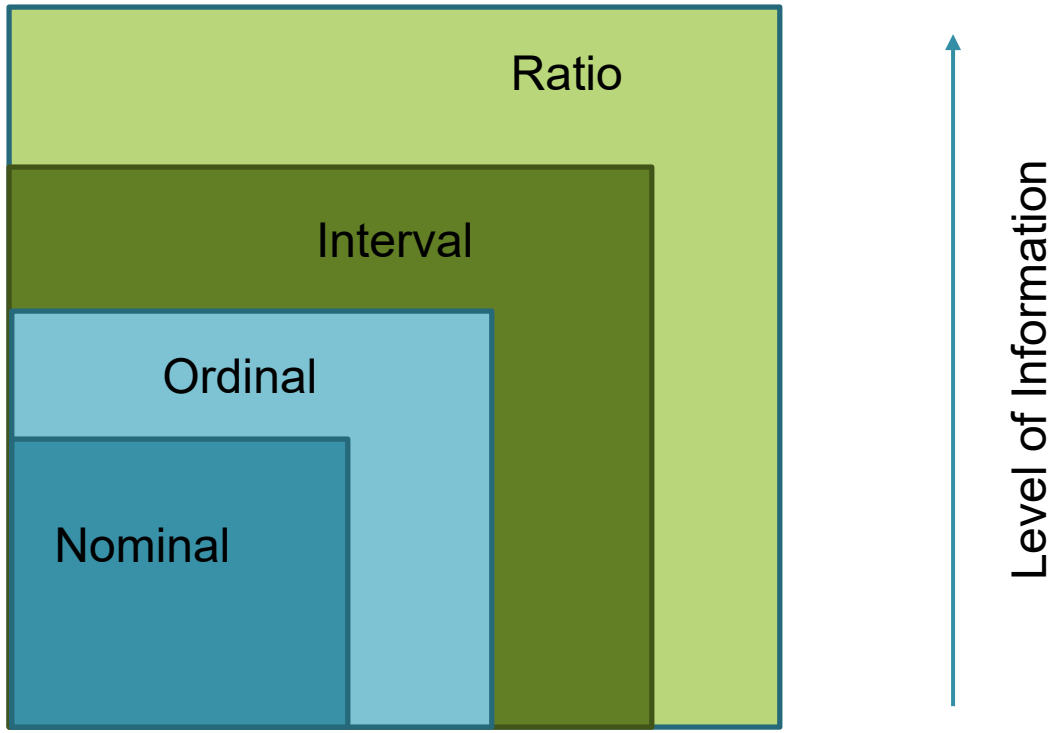
Source: http://www.dailymail.co.uk/news/article-3994504/How-close-live-NUKE-Fascinating-map-reveals-exact-locations-nuclear-bombs-stored.html

Guess the data type (4/4)



Source: http://chartsbin.com/view/1175

Ratio data has the maximum information



Class discussion: How many dimensions can you show on a chart?

Classification of Charts

Summarization

Comparing categories

Part-to-whole relationships

Showing changes over time

Connections and relationships

Maps

Adapted from Andy Kirk, Data Visualization: A Successful Design Process, Packt Publishing 2012

Summarisation

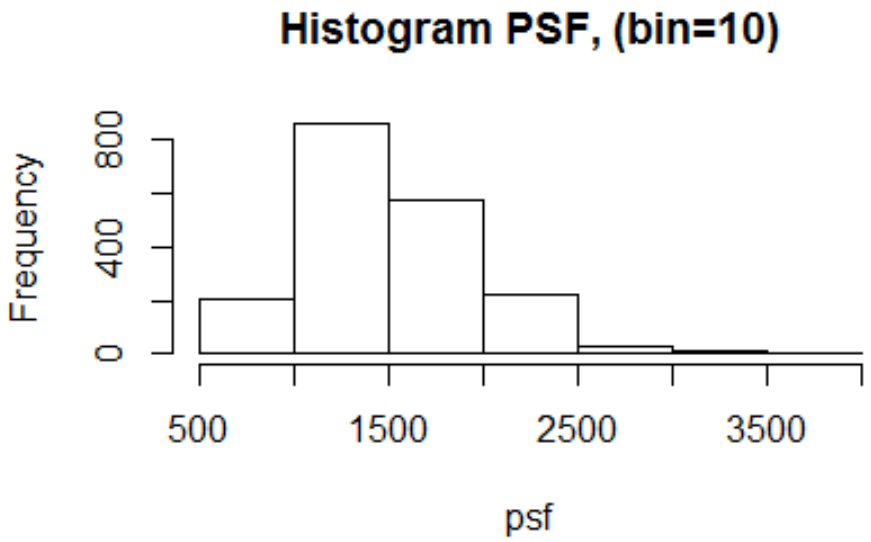
Charts that facilitate the overview of data

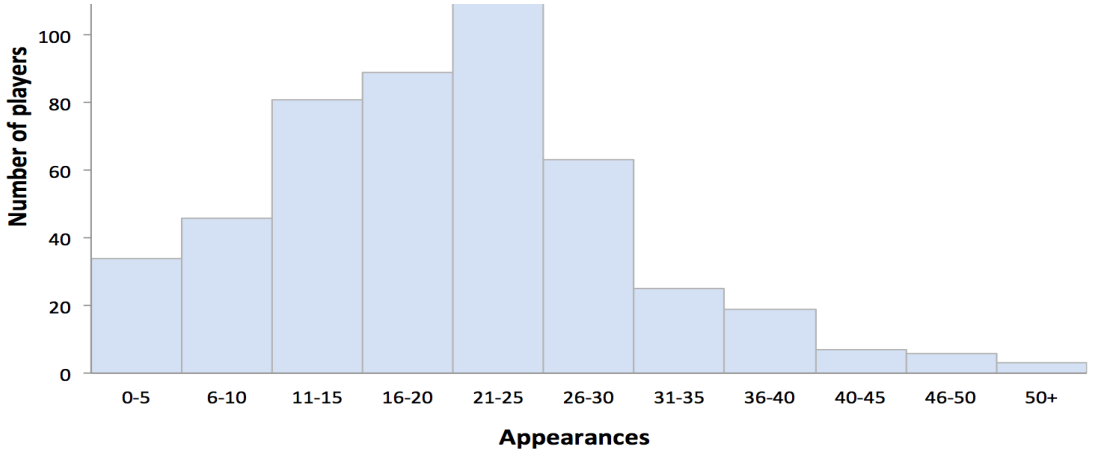
Histogram

- Used for showing **distribution** of data and makes categorizing easier
  - Bin size is important
    - Too small – spurious patterns
    - Too large – might miss important patterns

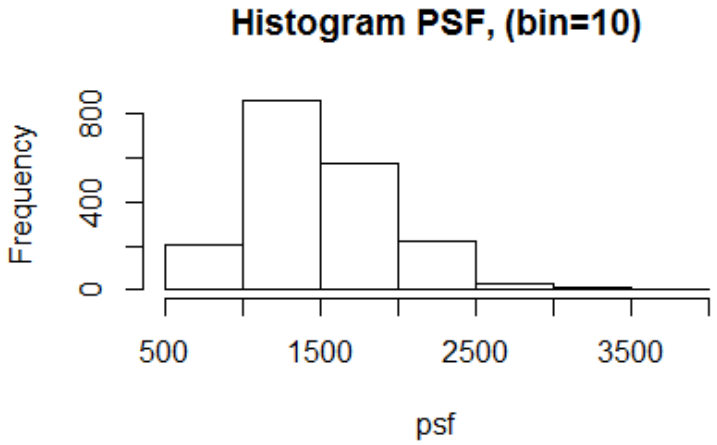
- Histogram vs. Bar Chart
  - Bar Chart is used to display **categorica**l data
  - Histogram is used to display a range of **numeric** values

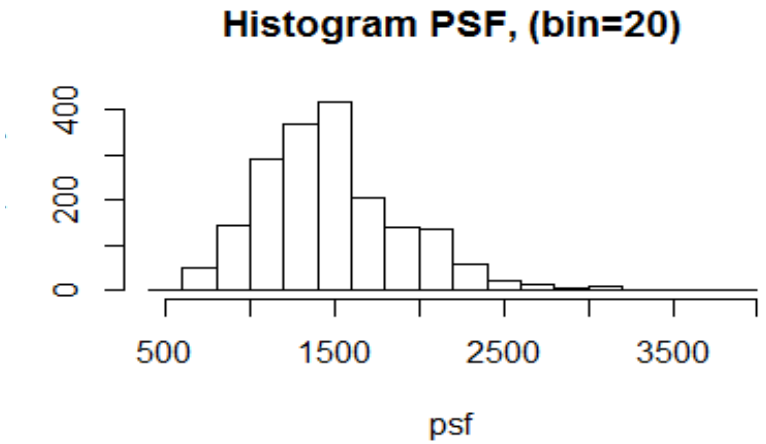
Histograms & Bar Charts

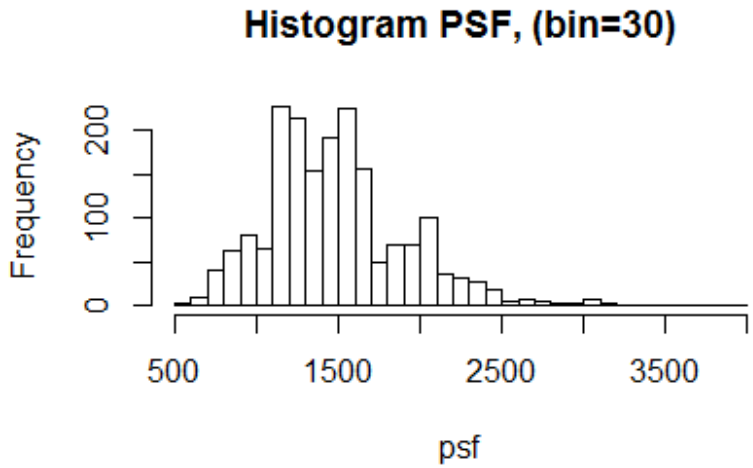


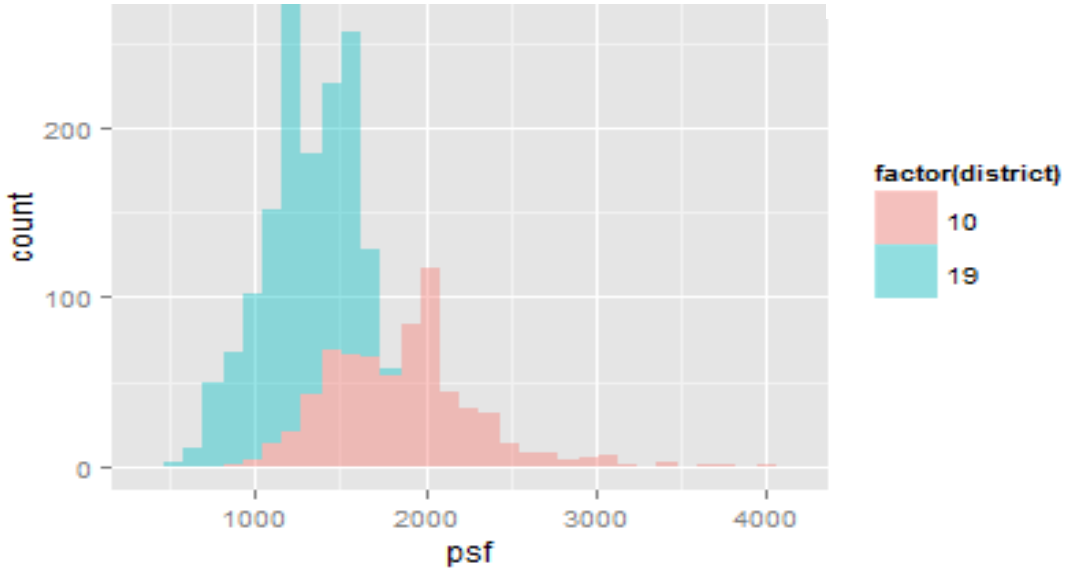


Histograms with bin variation

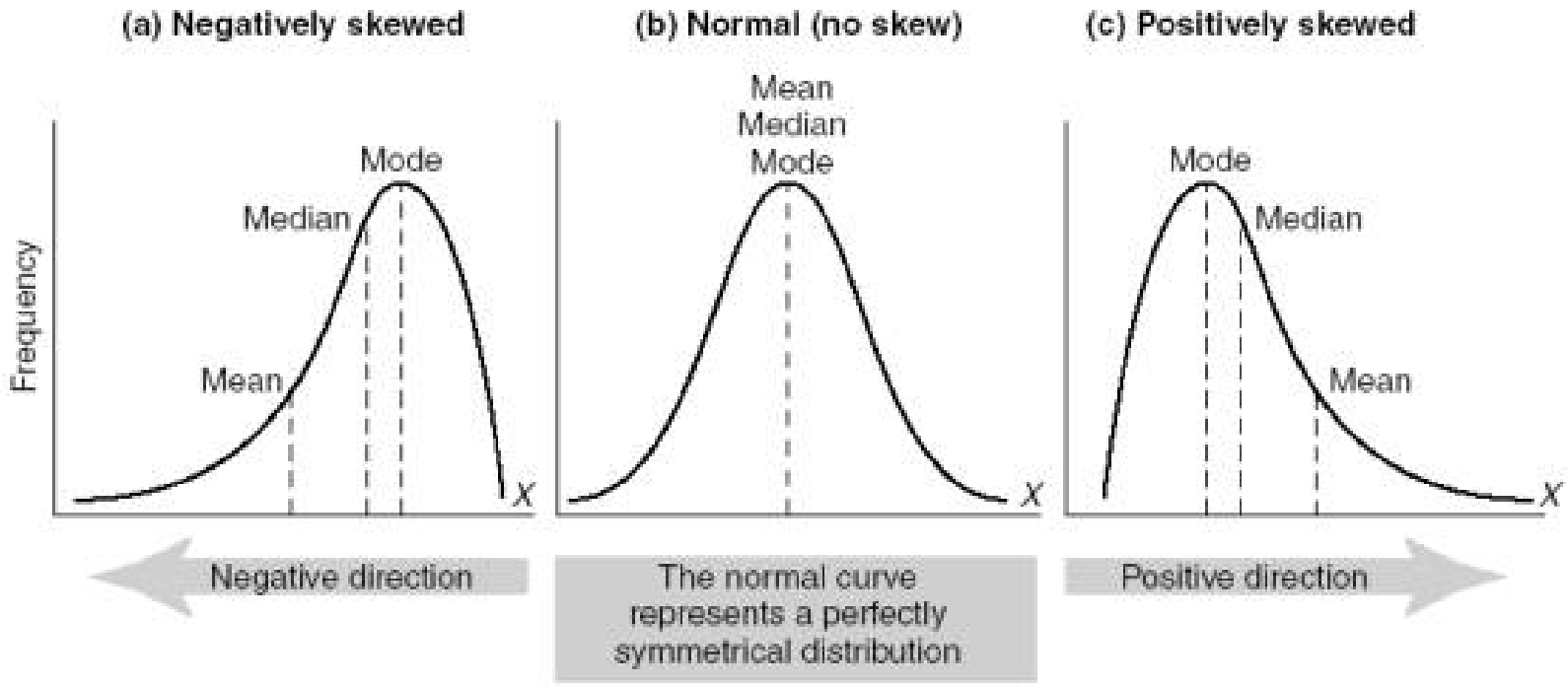






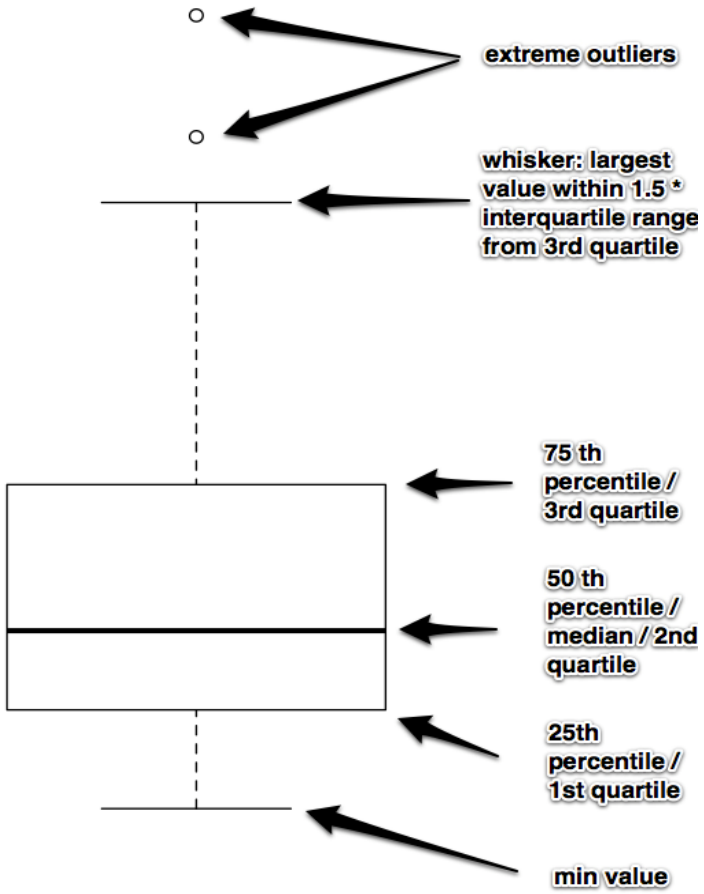


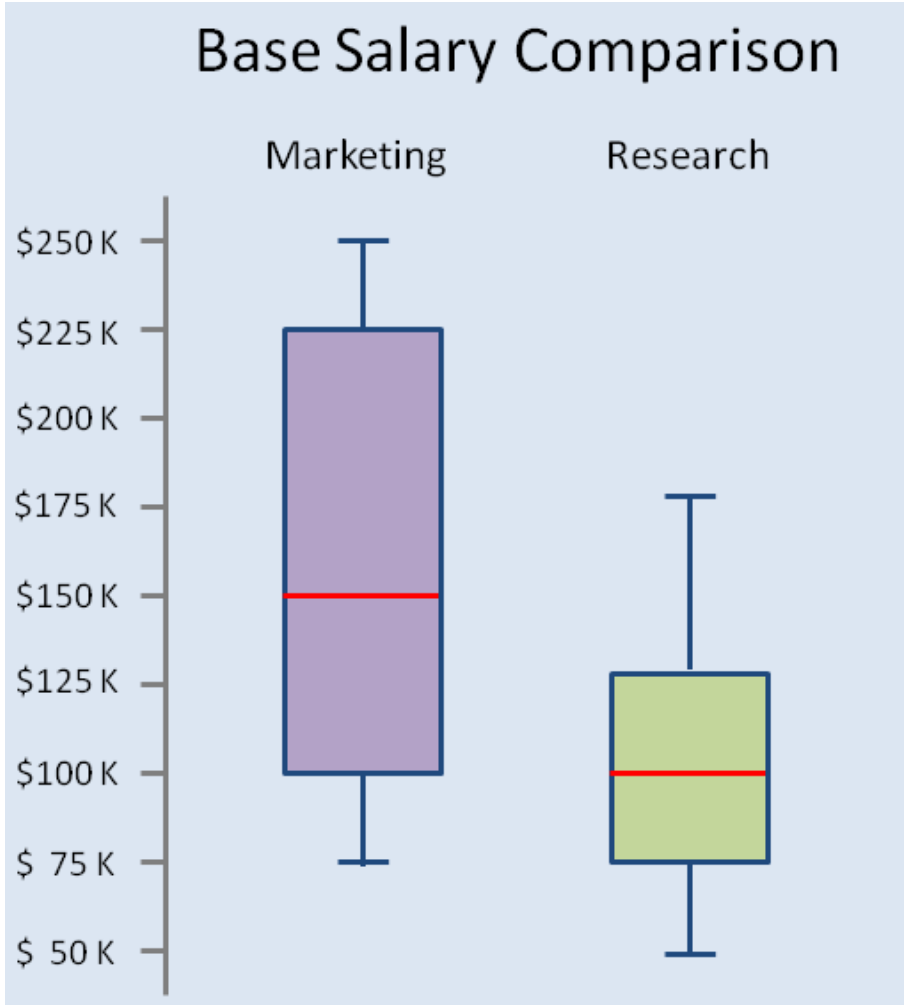
Mean, Median, and Mode



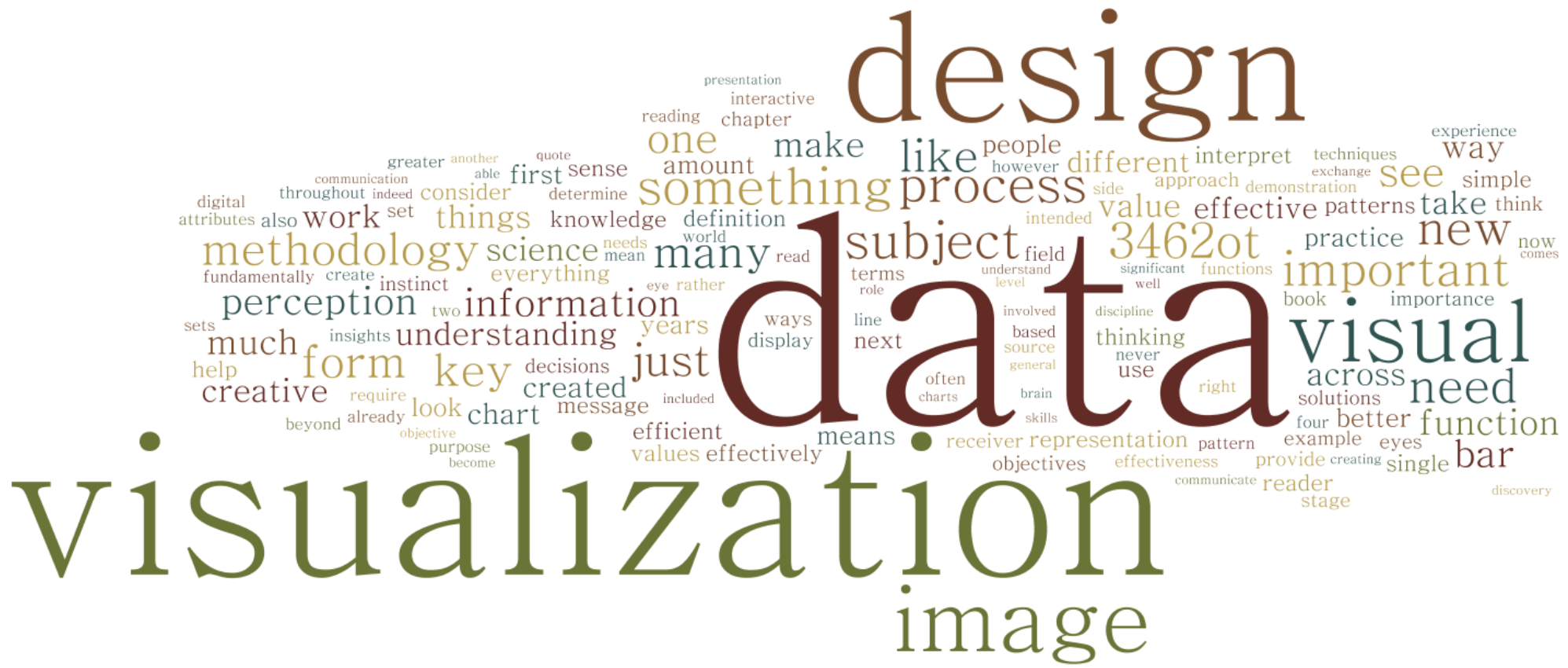
Box Plot

The five-number summary can be represented graphically using a boxplot





Word Cloud



https://www.wordclouds.com/

More Summarising options

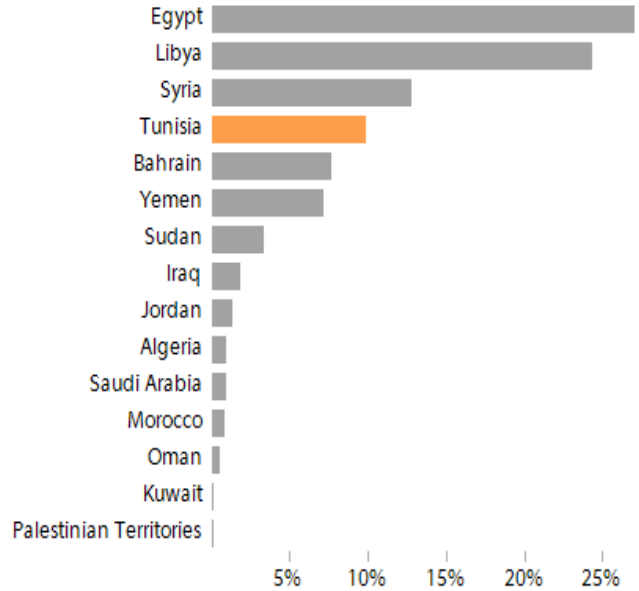
- Conditional Formatting
- Faceting

Comparing Categories

Charts that facilitate comparison across categories

Bar Chart

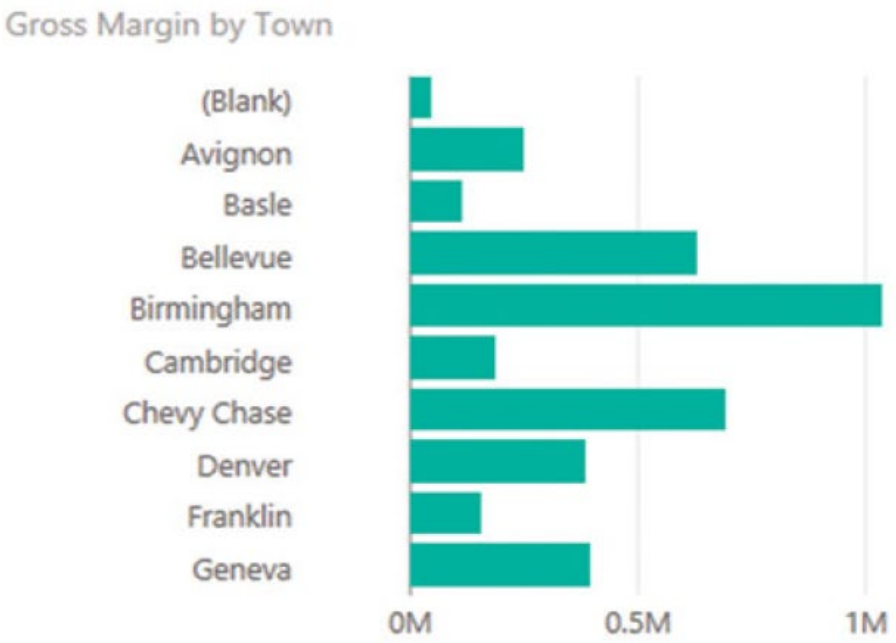
- Used for **comparing categories** or examining data across time periods
  - How sales are distributed across products?
    - Which product produce the highest sales
  - How has sales have changed over time?
    - Which months were good for sales

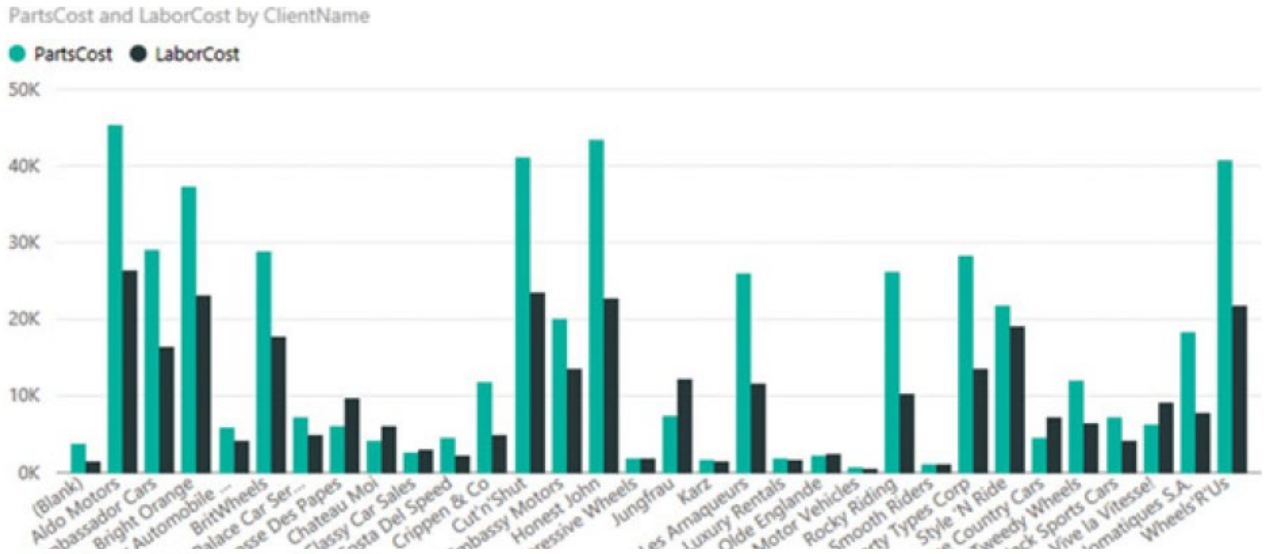


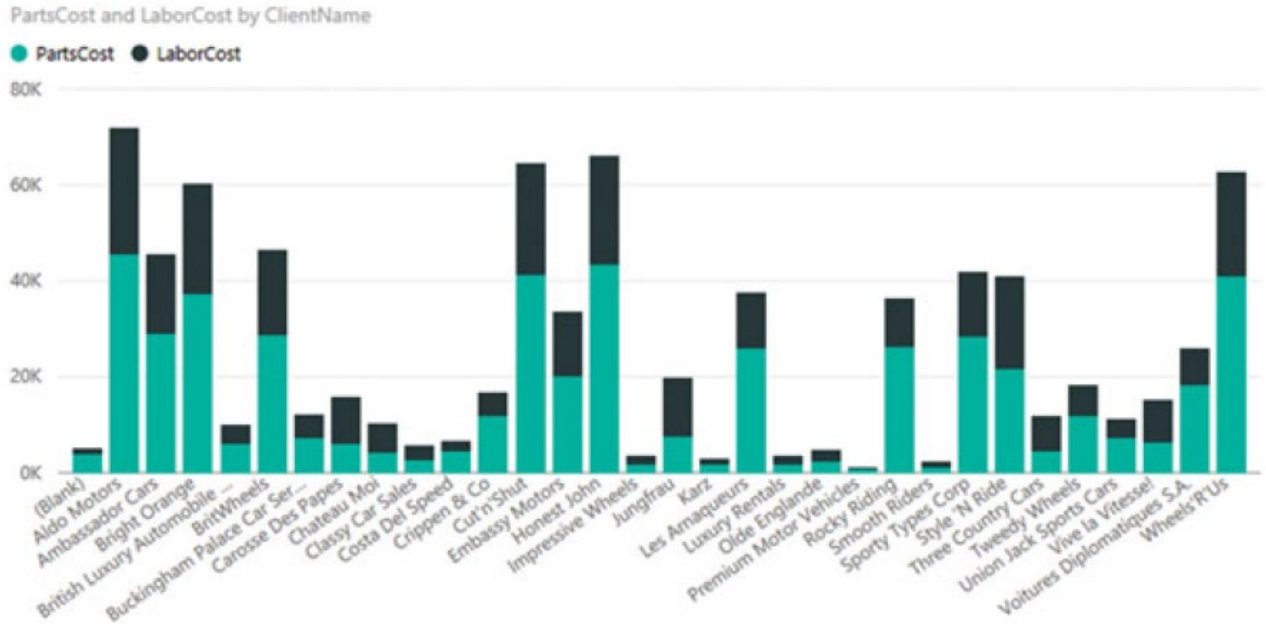
- Probably the most useful and most commonly used visualization of data

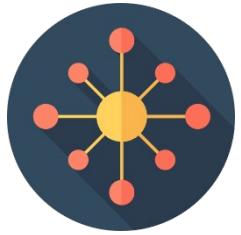
- Able to quickly **draw initial conclusions** about your data and leads to asking more questions

Bar Charts and options







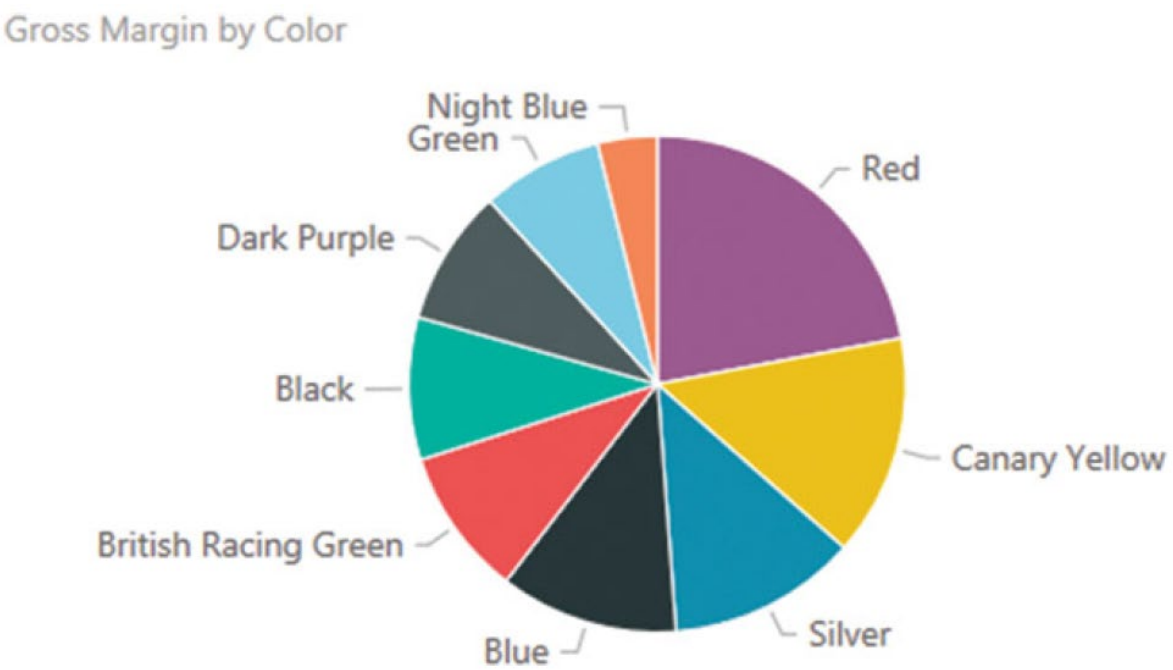


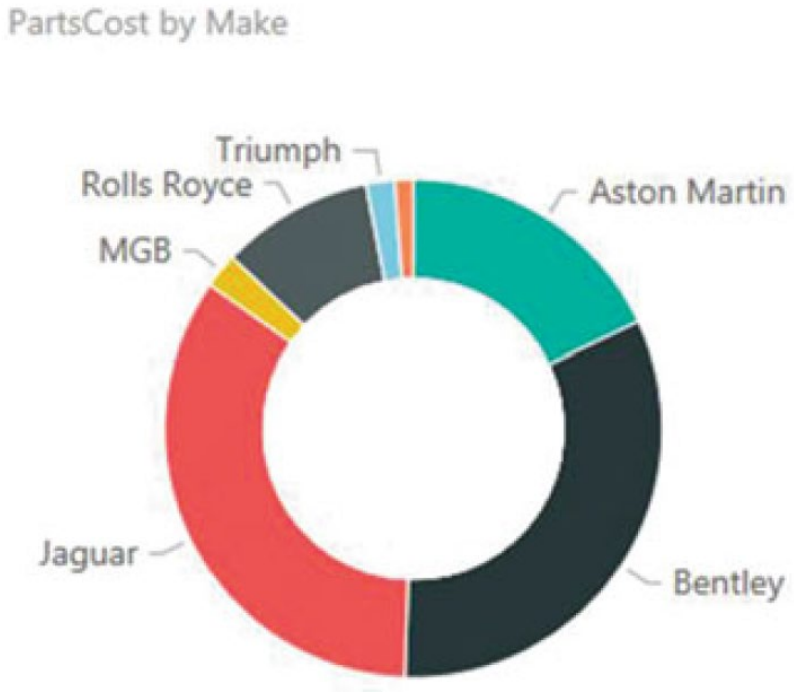
Part-to-Whole Relationships

Charts that help us to assess hierarchical or part-to-whole relationships

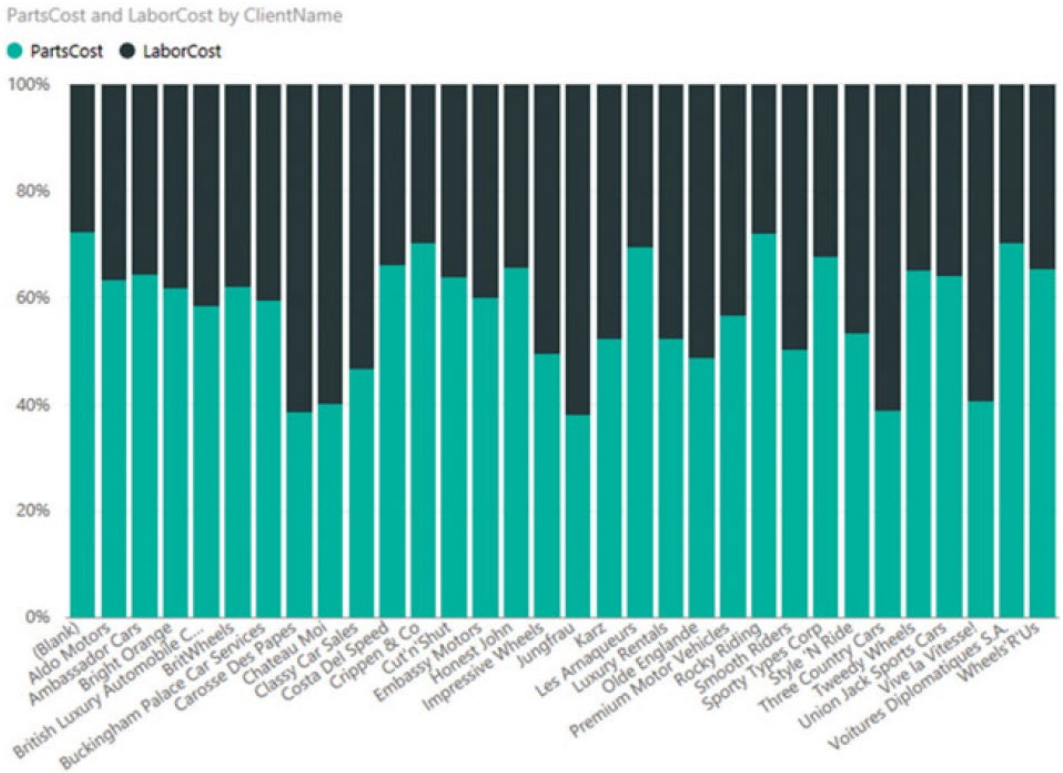
Pie Chart / Donut Chart

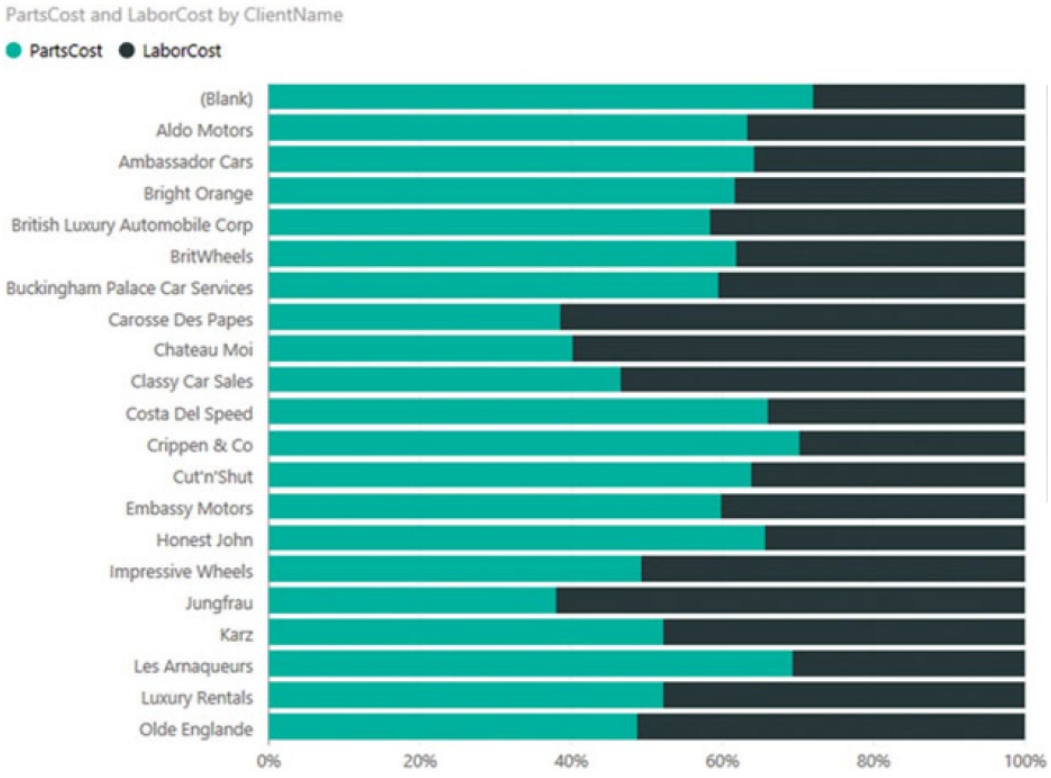
To display a limited set of data for a single series





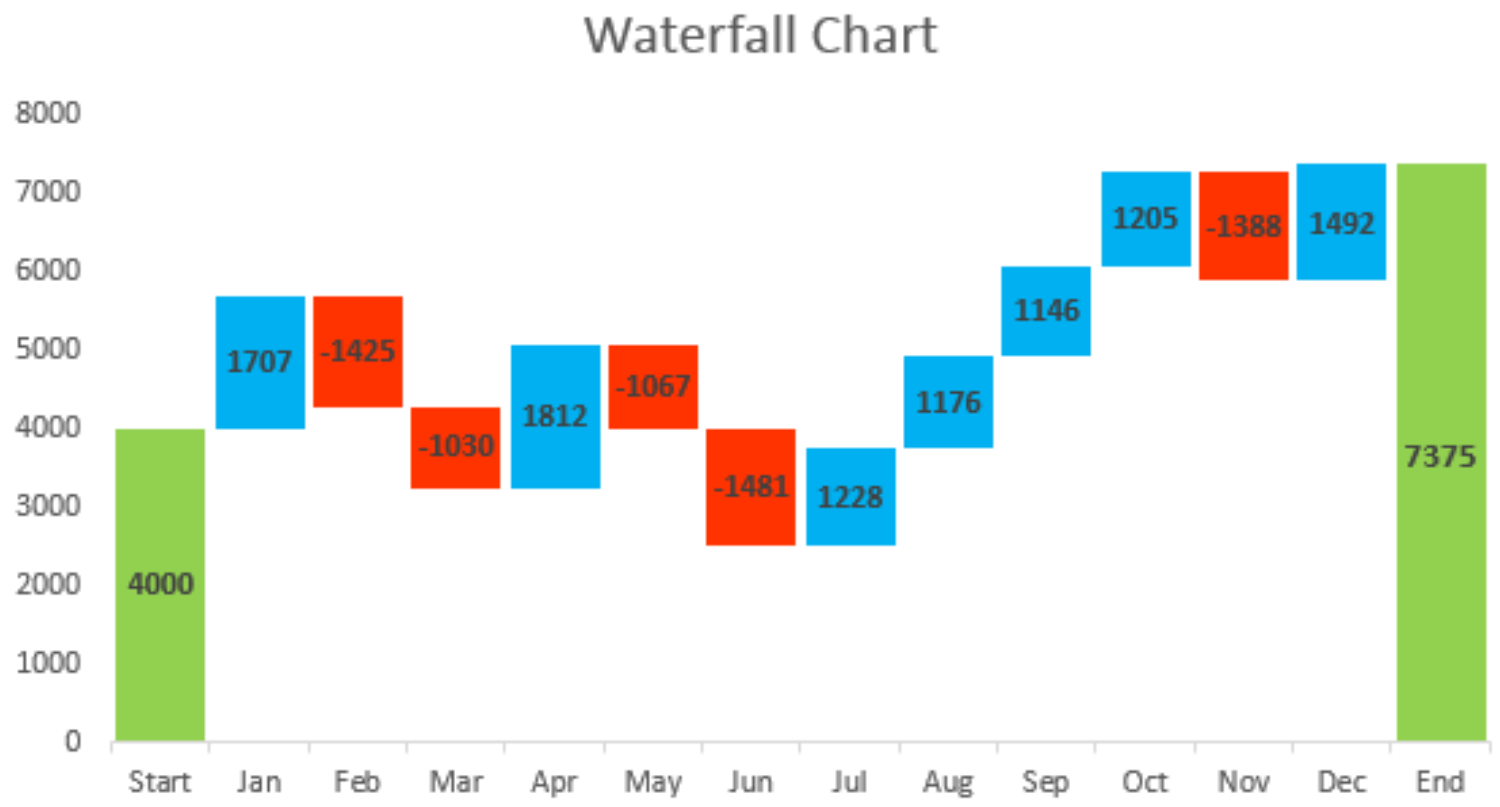
100% Stacked Column & 100% Stacked Column Chart





Compare data from multiple datasets and present each individual data series as a percentage of the total

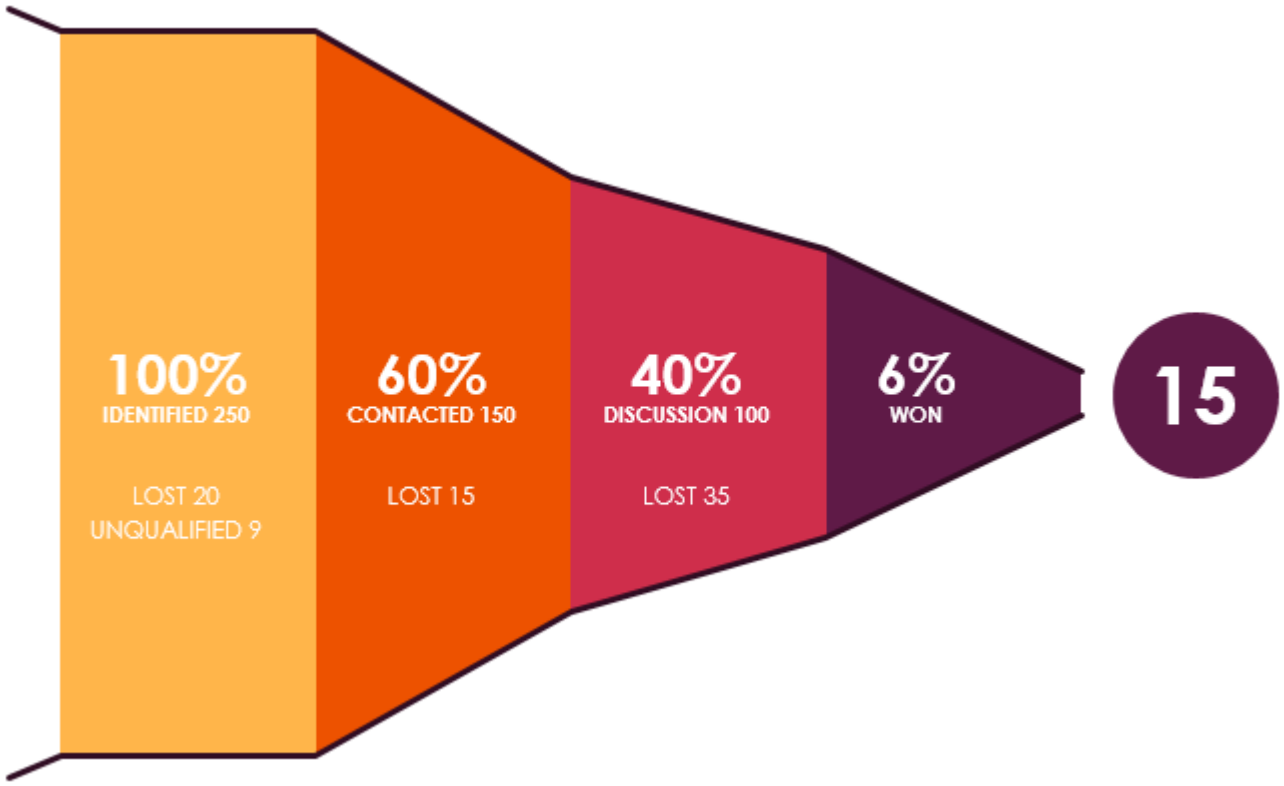
Waterfall Chart



To display changes from an opening state to an end state

Image Source: https://www.ablebits.com/office-addins-blog/2014/07/25/waterfall-chart-in-excel/

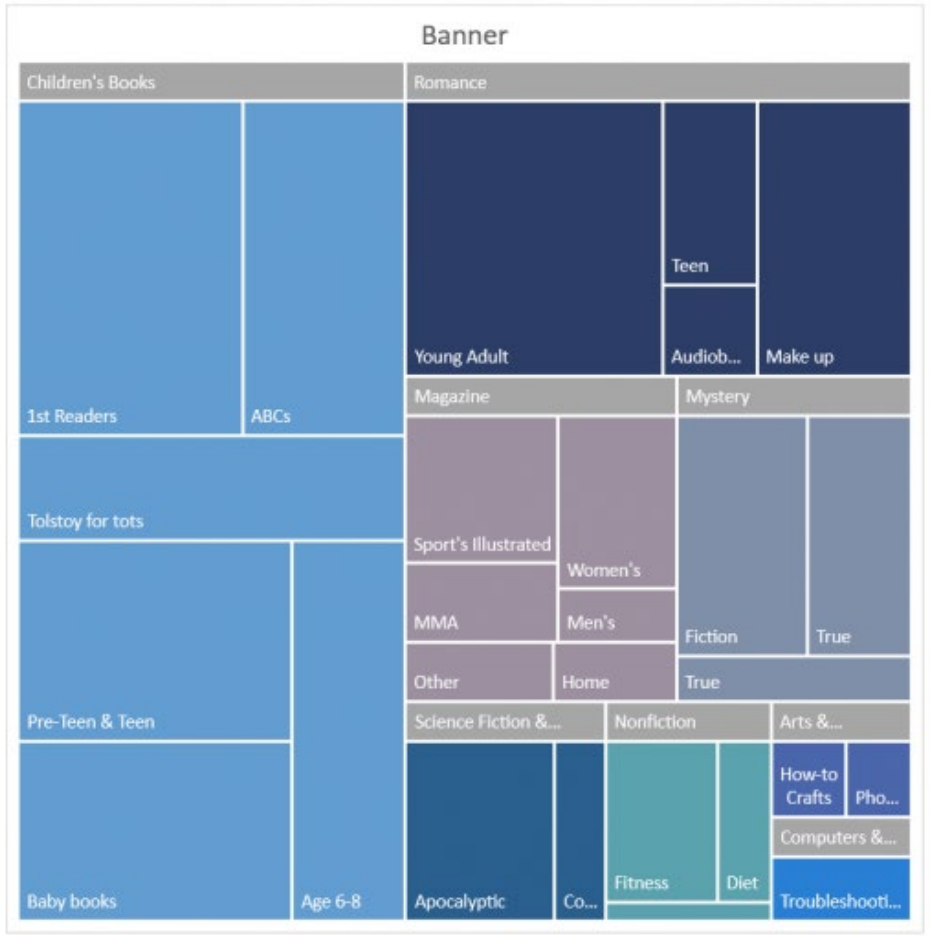
Funnel Chart

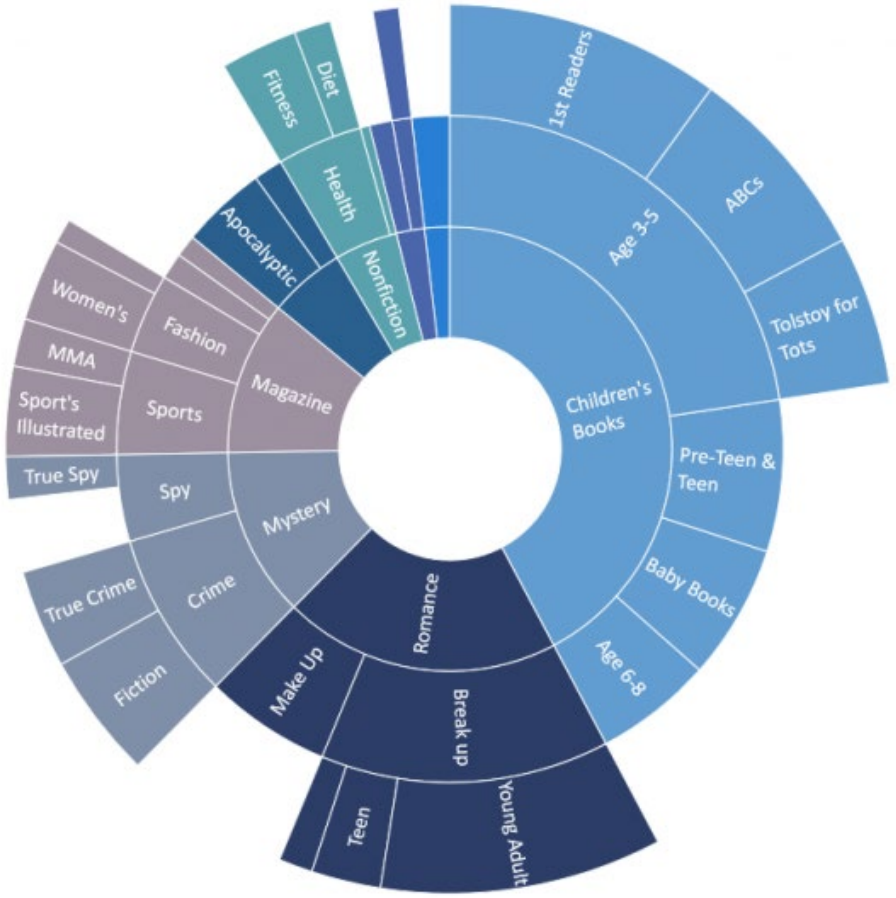


To show how a process filters out numbers

Image source: https://exceldashboardschool.com/sales-funnel-chart/

Tree Maps and Sunbursts





To see the way individual values relate to a total. Sub categories across hierarchies can also be compared. Is equivalent to multiple pie charts!

Image source: https://blogs.office.com/en-us/2015/08/11/breaking-down-hierarchical-data-with-treemap-and-sunburst-charts/?eu=true

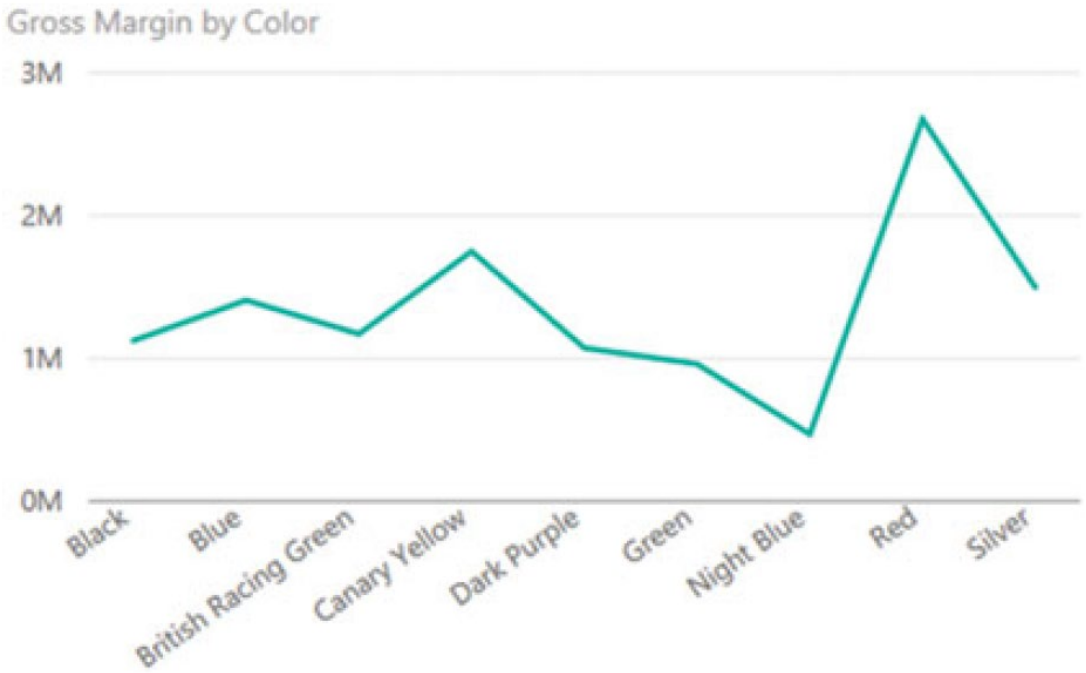
Showing Changes Over Time

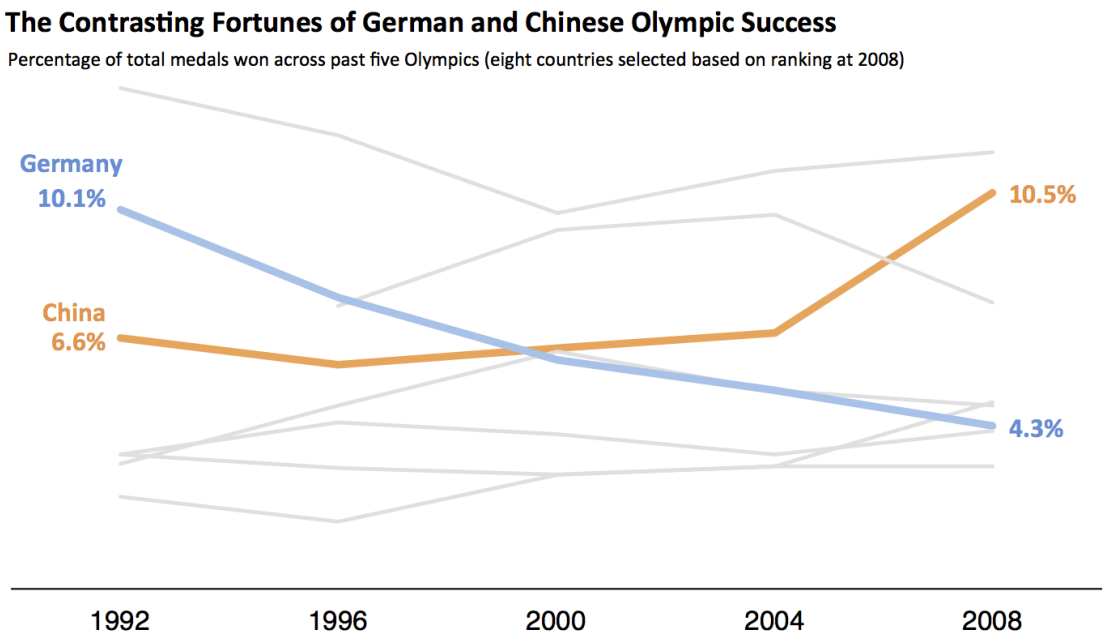
Charts that help show changes over time

Line Chart

- displays the data as a set of points joined by a line
- Used to show the direct **relationship between 2 variables**, and how one variable **changes over time**.
  - show data over a series of numeric values
  - Eg: How sales has changed over the last x years/quarters/month
- Line Plots are useful for **examining trends** over time periods
  - Are sales moving up or down or stagnant?
  - What is happening?
  - Is this seasonal?

Line Chart





displays the data as a set of points joined by a line

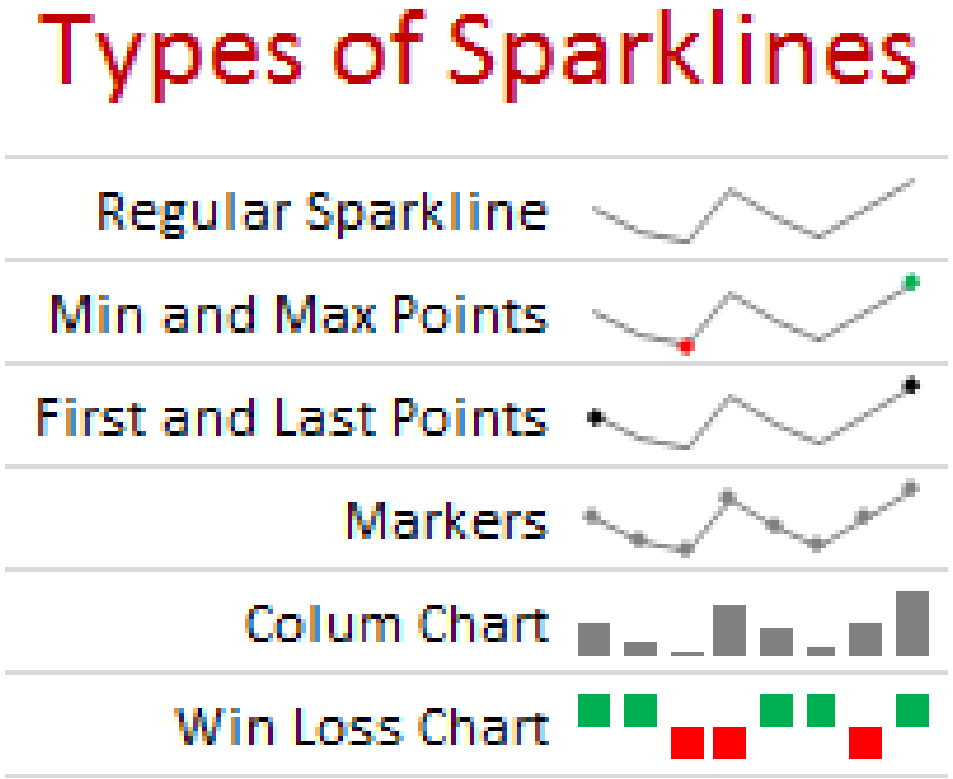
Candlestick Charts

https://www.investingnote.com/stocks/SGX:C6L#/?stock=C6L.SI&tab=all

To show additional information within the same space

- Candlesticks show opening and closing stocks

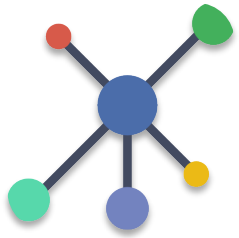
Sparklines



- Tiny charts shown in one cell of a table
- Can at a glance show trends

Gauges

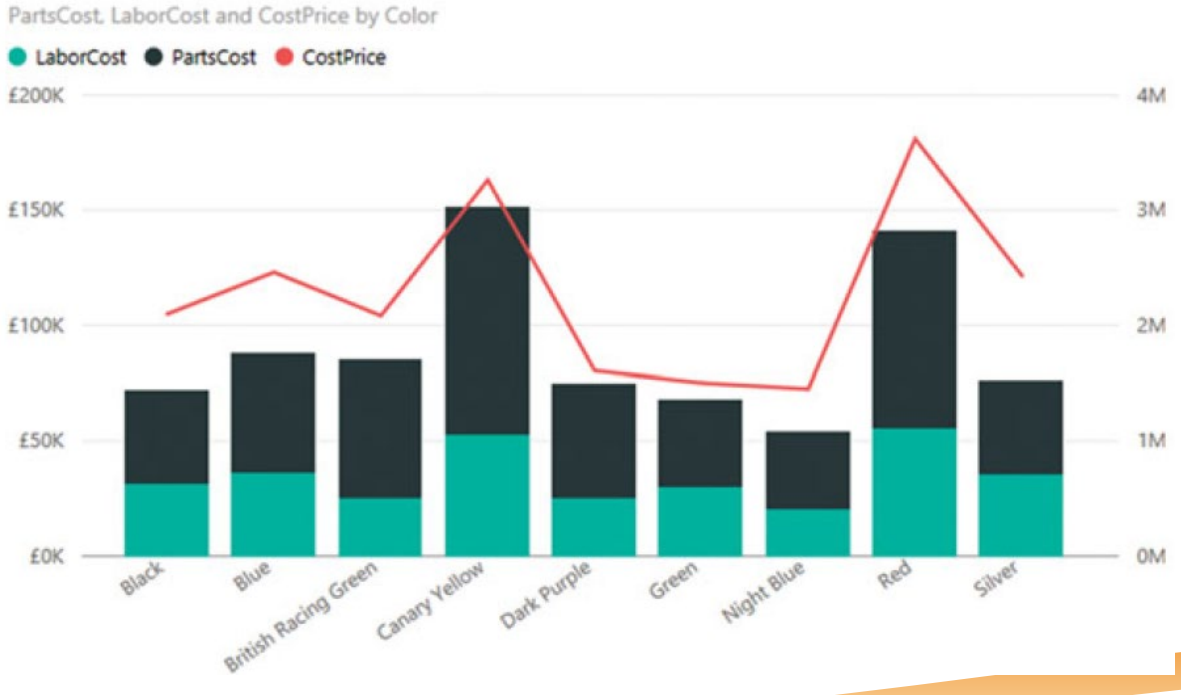
Compare actual to target or see how a metric compares to a key performance indicator (KPI). This could be used to show changes across time if animated



Connections and Relationships

Visualization techniques used to plot connections and relationships

Combo charts

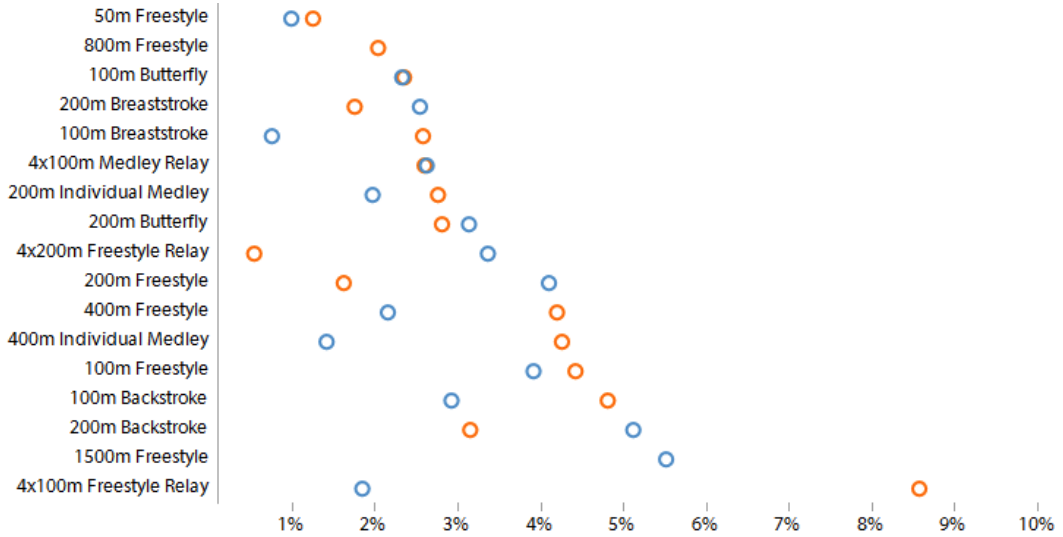


to isolate any interesting correlations

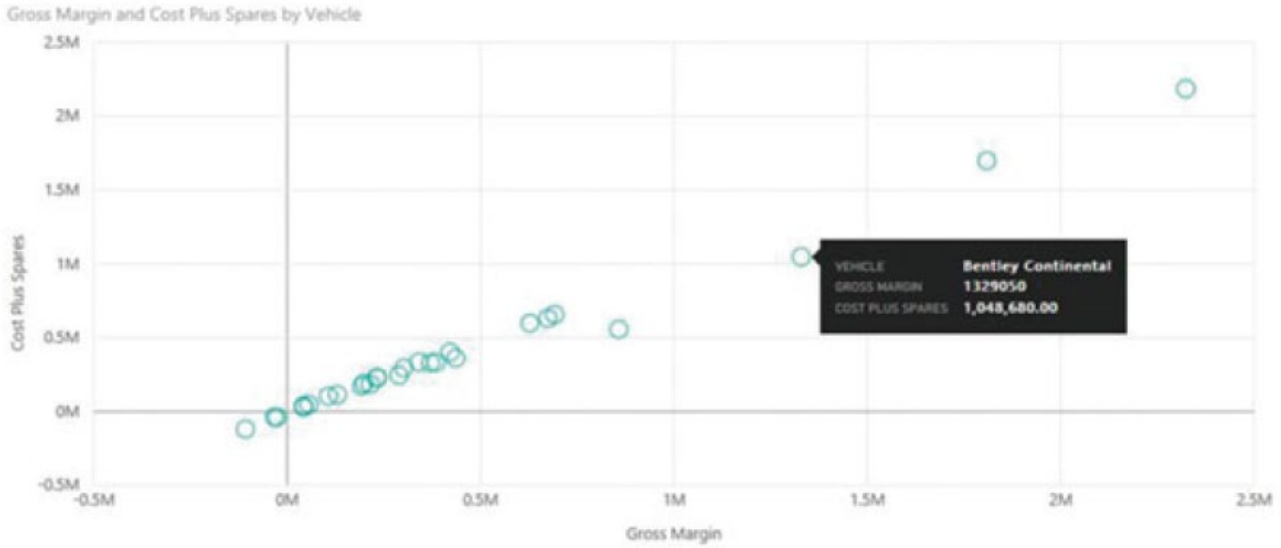
- key advantages of dual-axis charts is that the two X axes can have vastly different scales

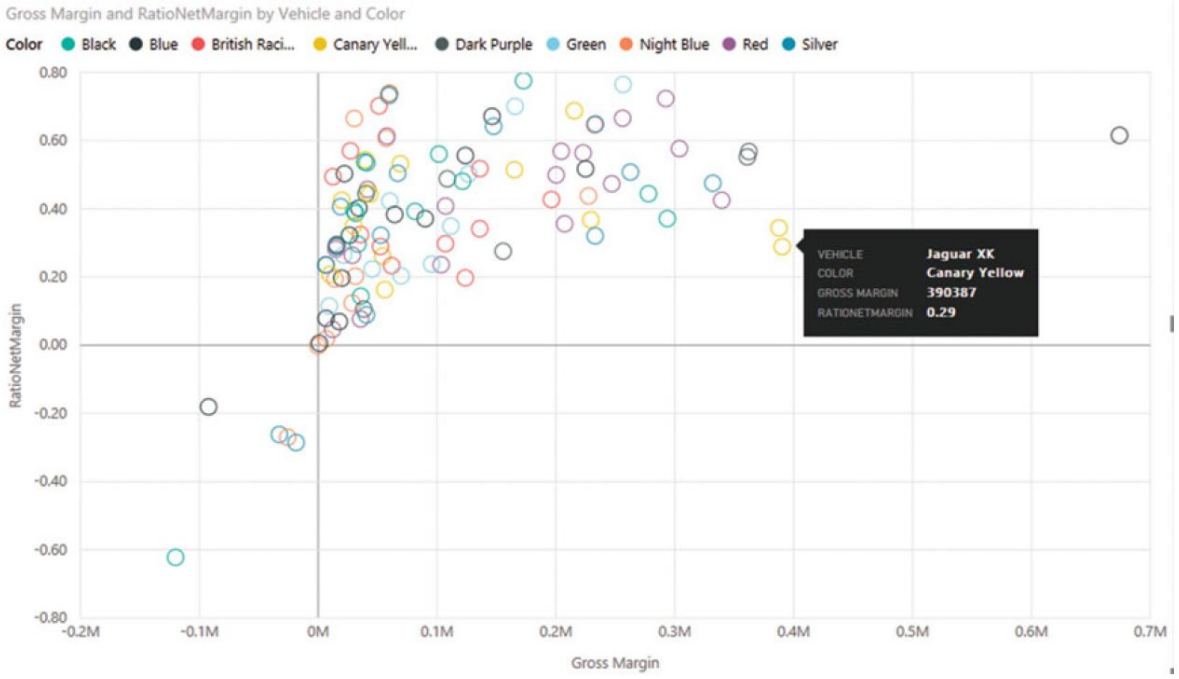
Scatter Plots

- Used for **analyzing relationships** between two numeric variables
  - often used to prove or disprove a relationship between the two variables.
    - Example: see how sales and profit relate
  - Allows a view of skewness of the distribution
    - can easily **detect Outliers**.
- Scatter plots can be more useful if you draw **trend lines**.
- Bubble Plot - is a scatter chart with a third piece of data included

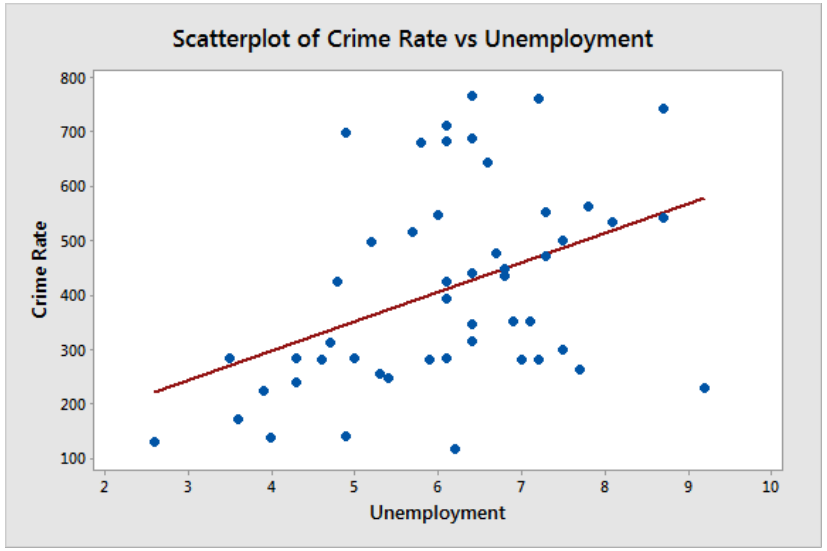


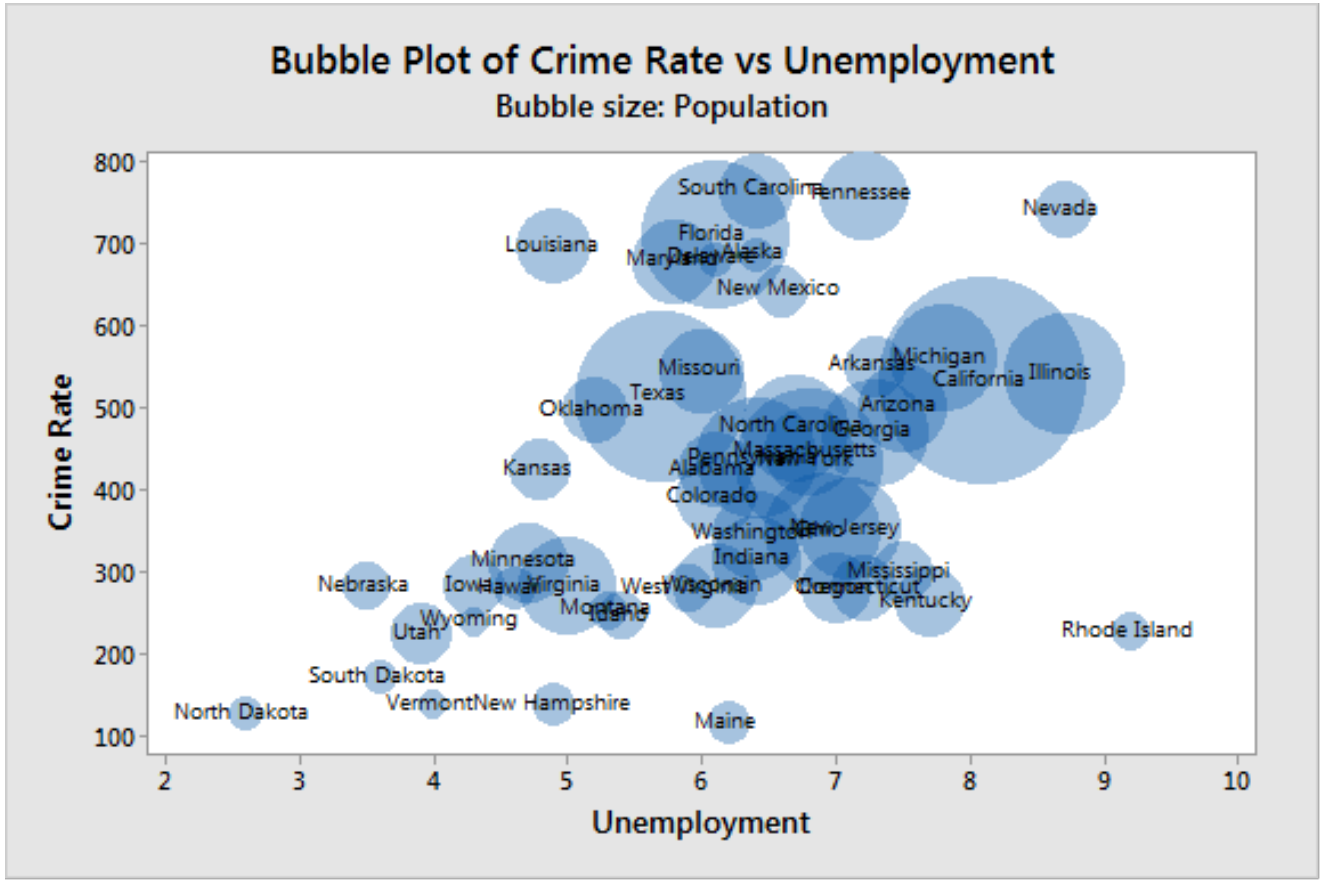
Scatter Plot



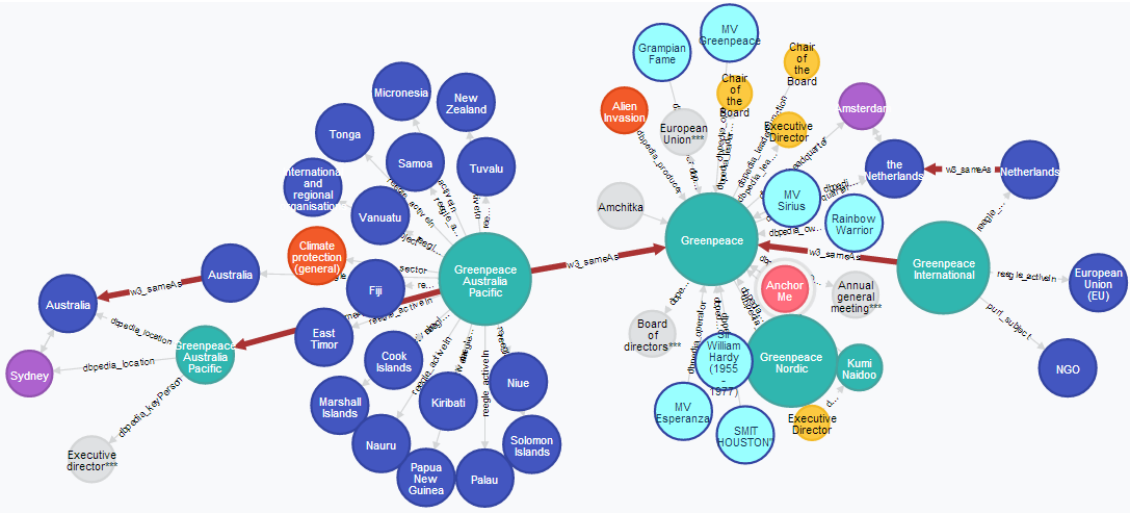


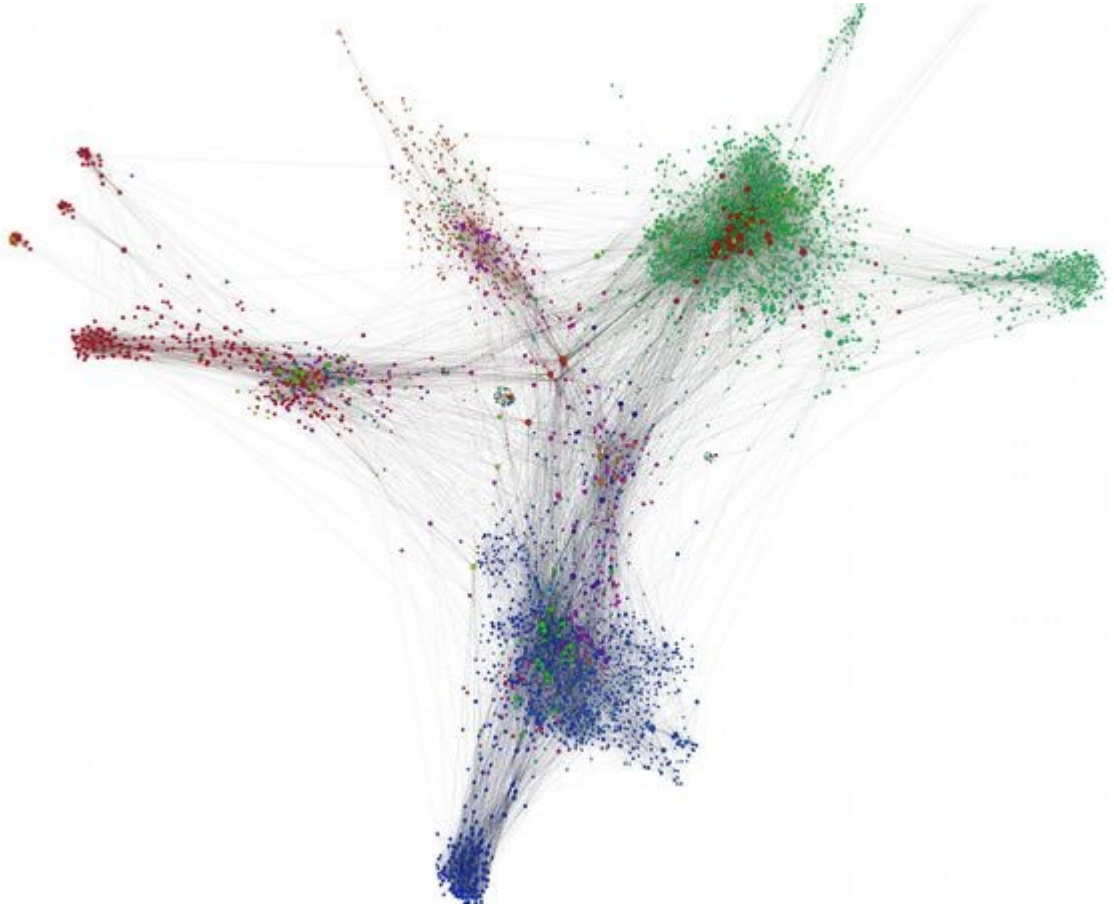
Bubble Plot adds another dimension to a Scatter Plot





Graph data



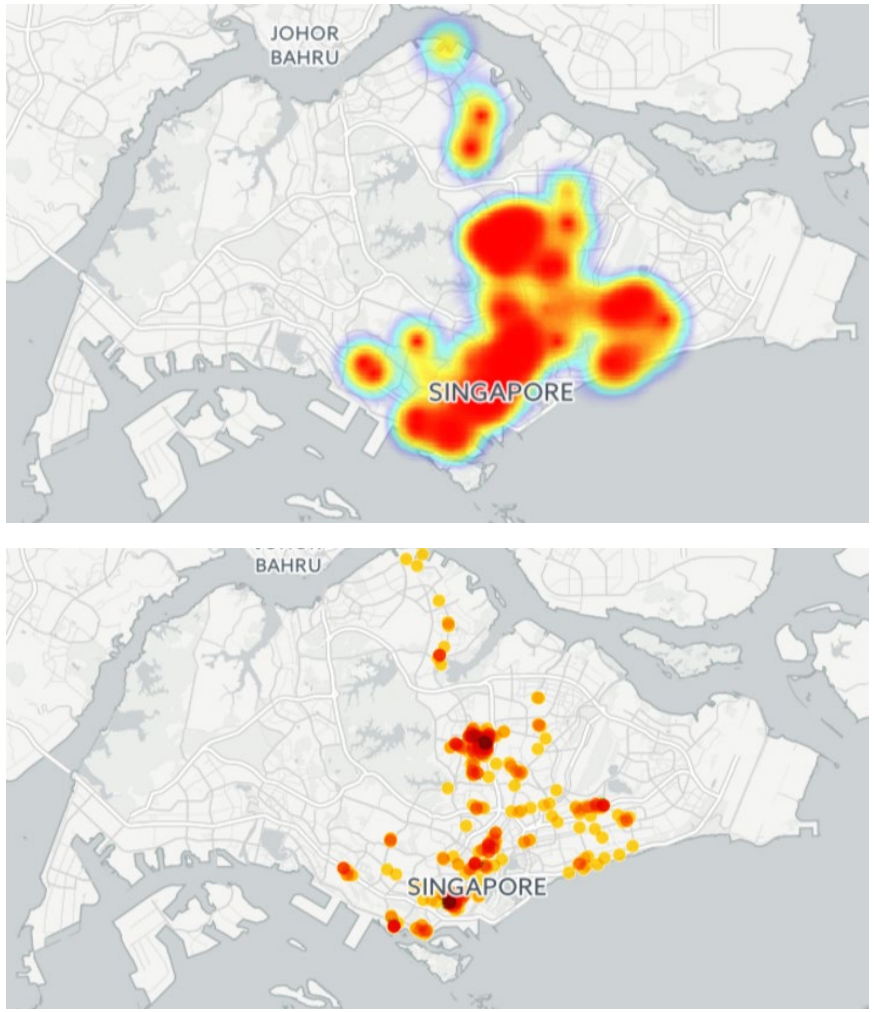


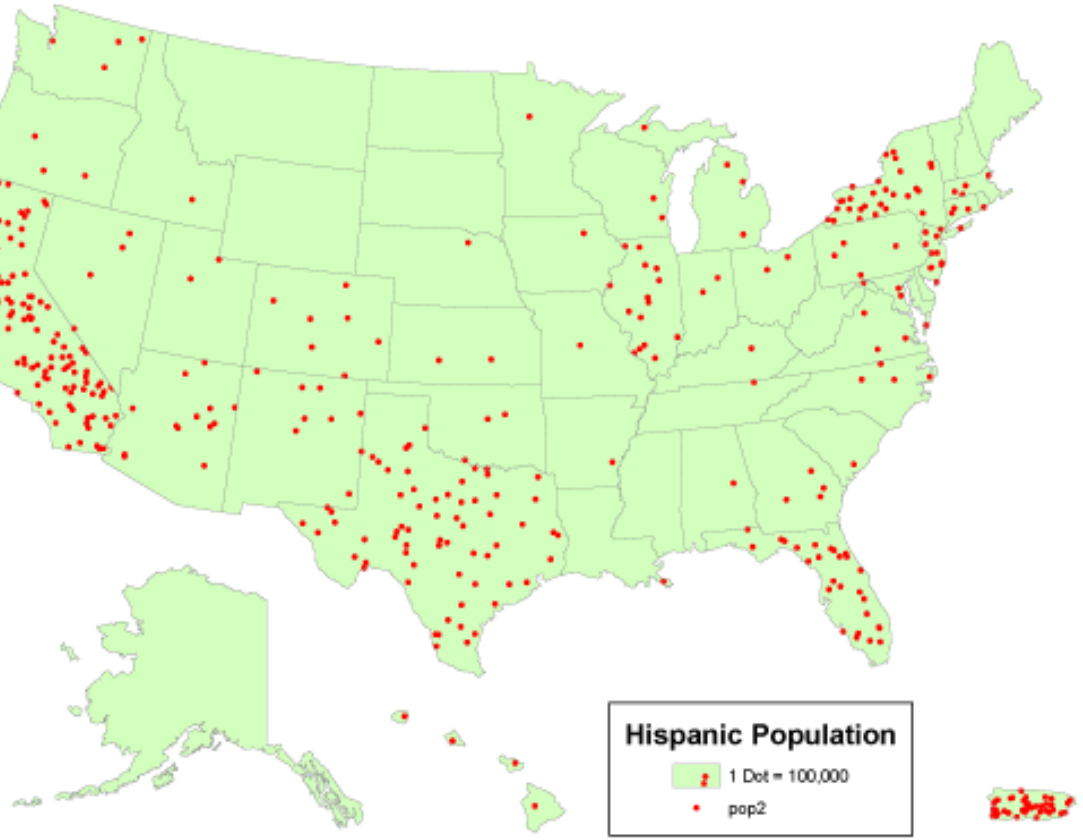
- Two kinds of uses: Meta Data Analysis and Large Graph patterns

Maps

Different ways of mapping data

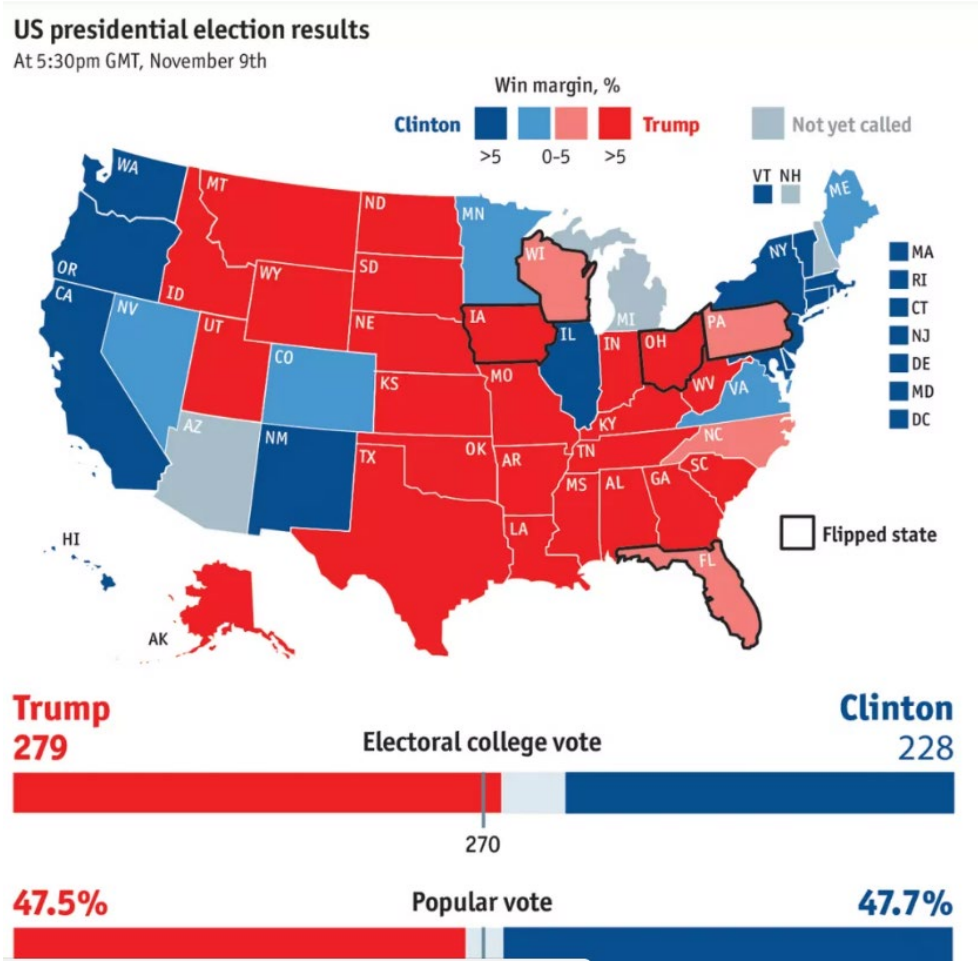
Choropleth Maps (Heatmap) vs Dot Plot Map





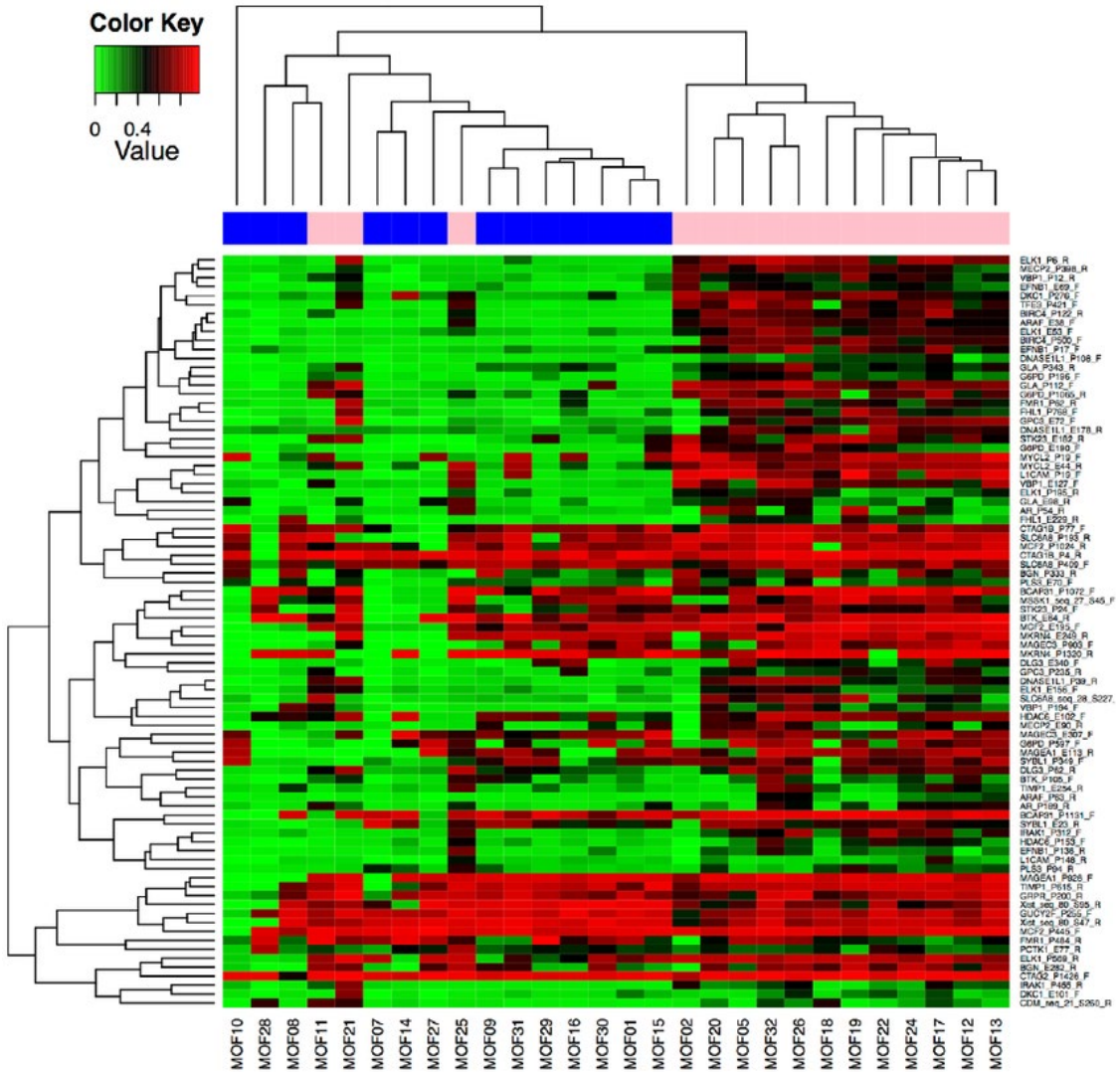
Bubble Plots

Heat Maps using Shape Files



Source: http://geoawesomeness.com/top-23-maps-charts-explain-results-2016-us-presidential-elections/

Heatmaps



- Used for **visualizing trends** in datasets with large categories
- Need not be geographical
- Useful in initial exploratory analysis and in presenting findings

Image source: https://www.researchgate.net/figure/Gender-Heatmap-of-X-linked-CpG-Loci-DNA-methylation-is-involved-in-the-transcriptional_fig1_233888916

Workshop: Pick the visuals

- Pick required charts for the storyboards
- Pick required supporting visuals
- For each chart mock the Analysis Output Data underneath the chart. Use post it notes and pictures if need be
- Place the charts on to a powerpoint presentation and present

PowerPoint tips

- Smart Art
- Alignment
- Wingdings
- XL paste options