Data Visualization in Python I

Sat, April 4th. 2023

Data Analytics

Program Zenius Studi Independen Bersertifikat Bersama Kampus Merdeka

1. Introduction to Data Visualization
2. Visualization Pillar and Concept
3. Python Libraries for Data Visualization

Introduction to Data Visualization

Data Visualization

The graphical representation of information and data.

Data Visualization

The graphical representation of information and data.

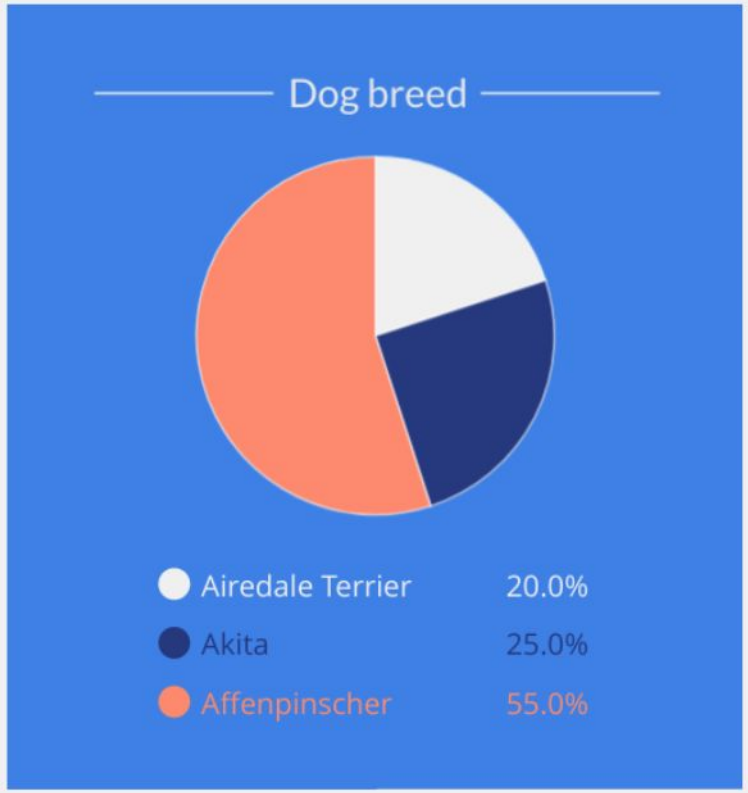
It is a practice of translating information into a visual context, such as a map or graph, to make data easier for the human brain to understand and pull insights from Data

Without Data Viz

Dog breed

Airedale Terrier	Akita
Afghan Hound	Malamute
Kelpie	Barbet
Basset Hound	Basenji
Beagle	Beauceron

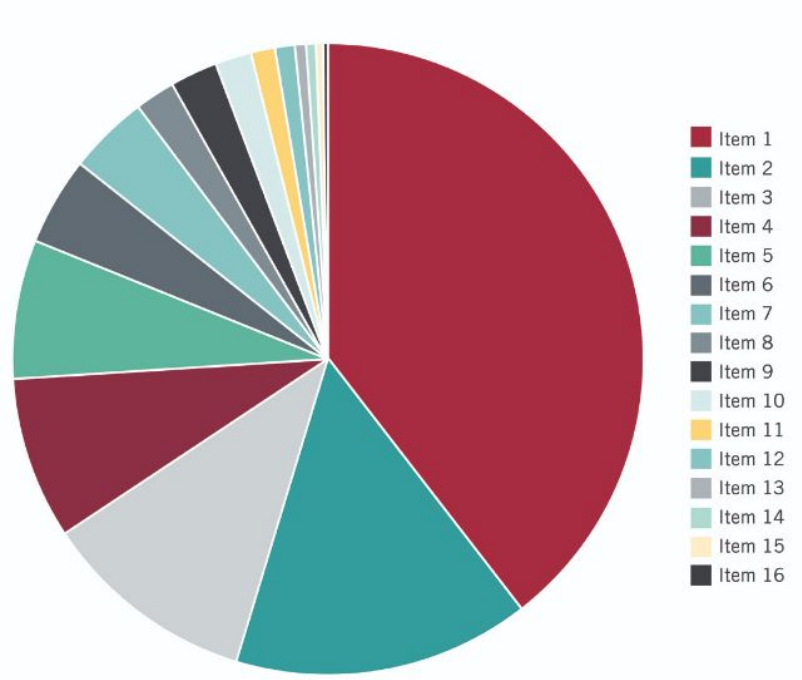
With Data Viz



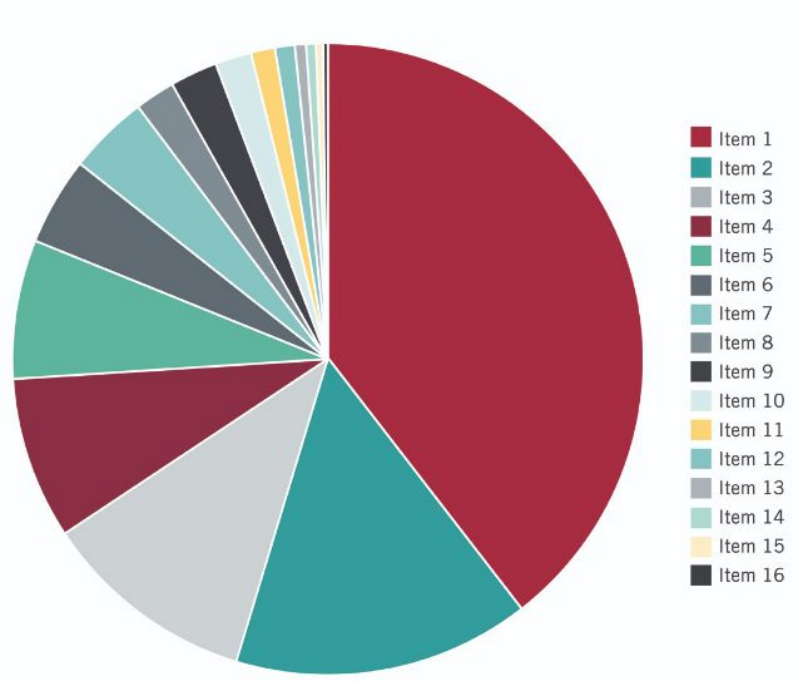
Imagine you are one of the stakeholders.

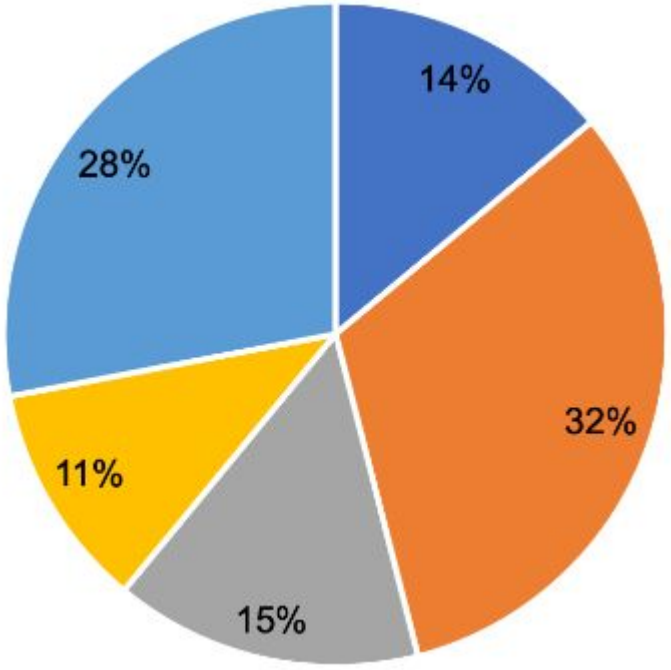
Bad vs Good Visualization

Bad vs Good Visualization

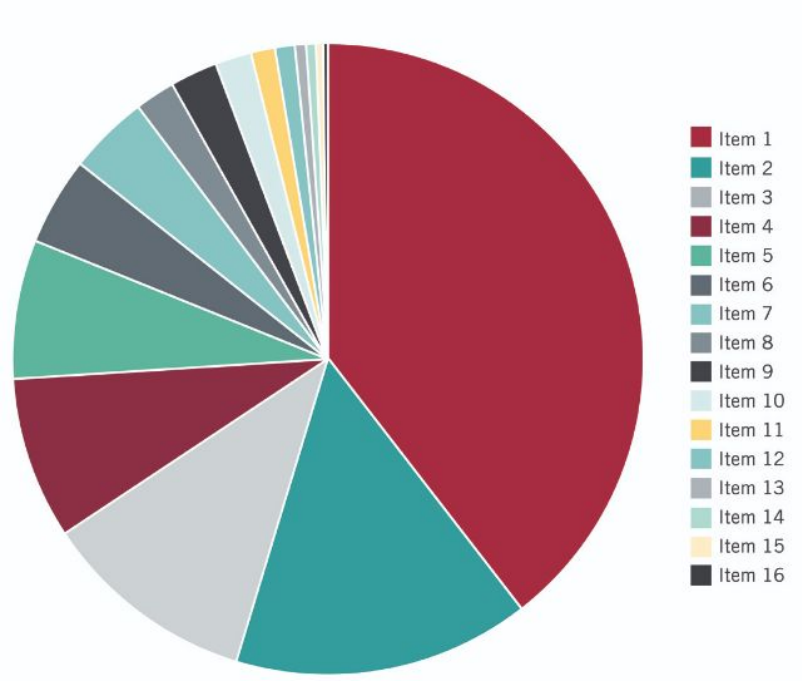


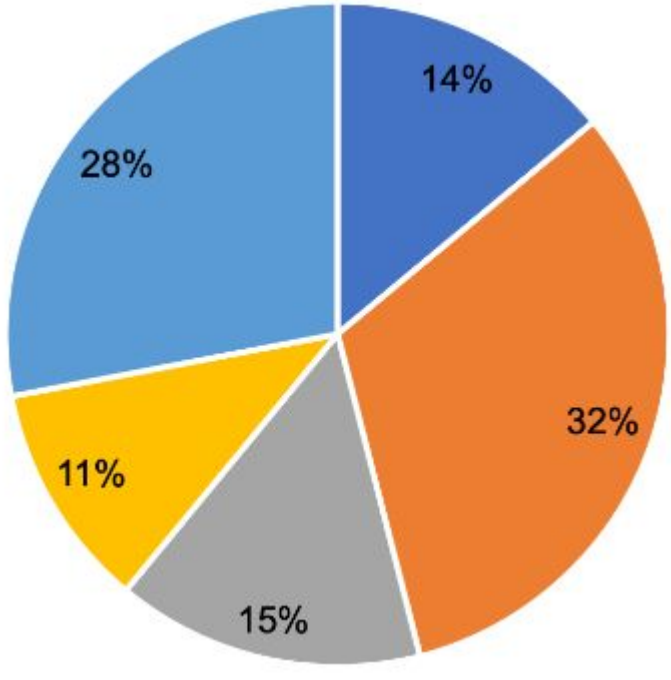
Bad vs Good Visualization



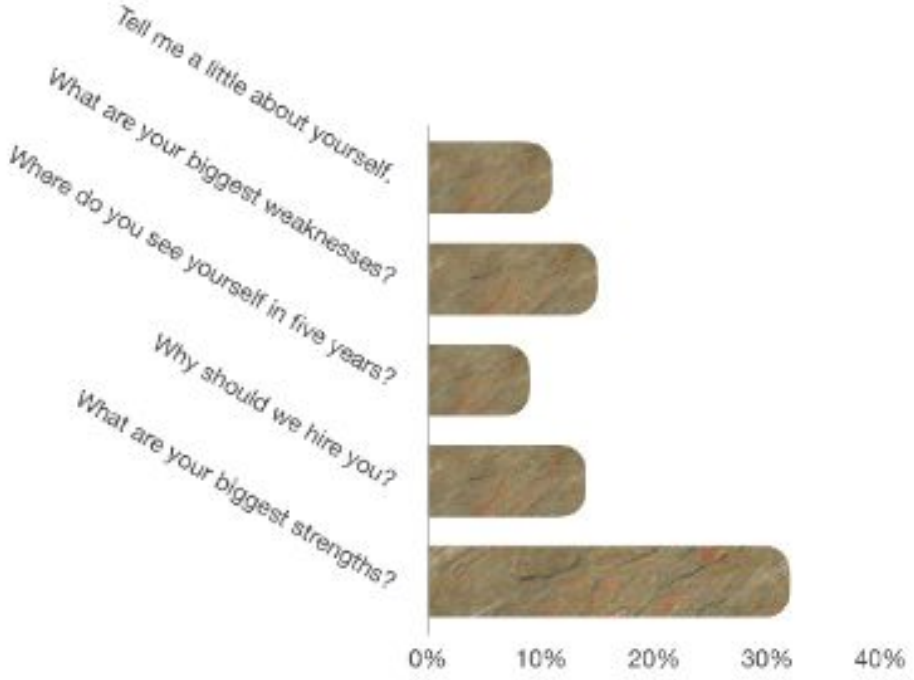


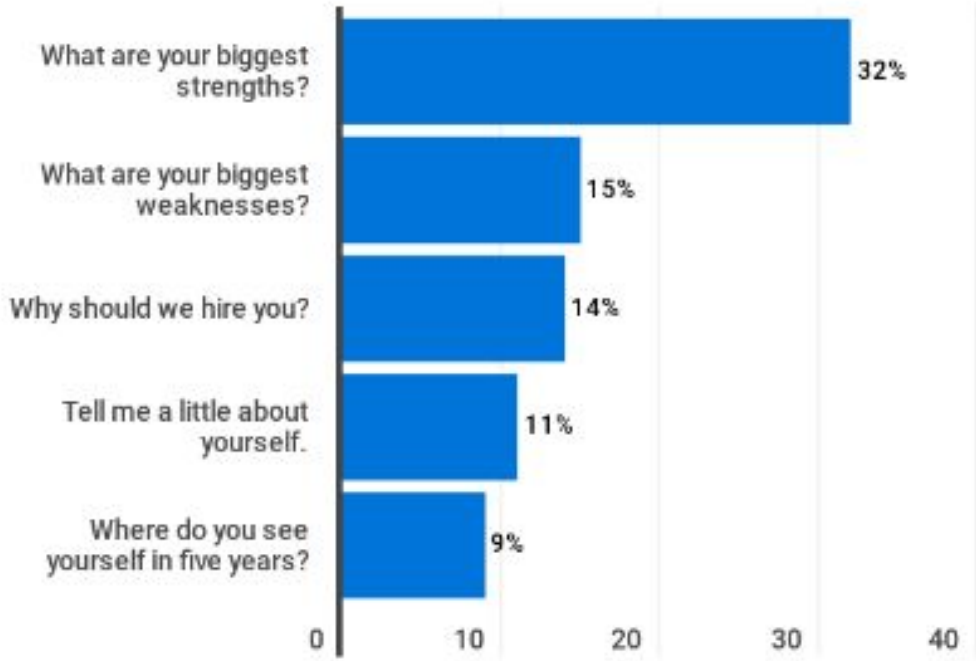
Bad vs Good Visualization



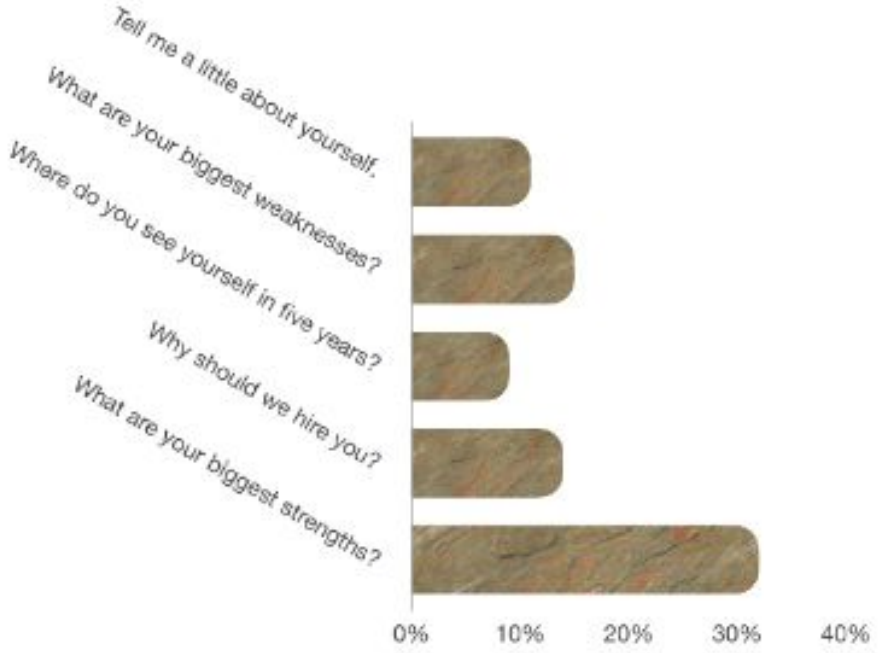


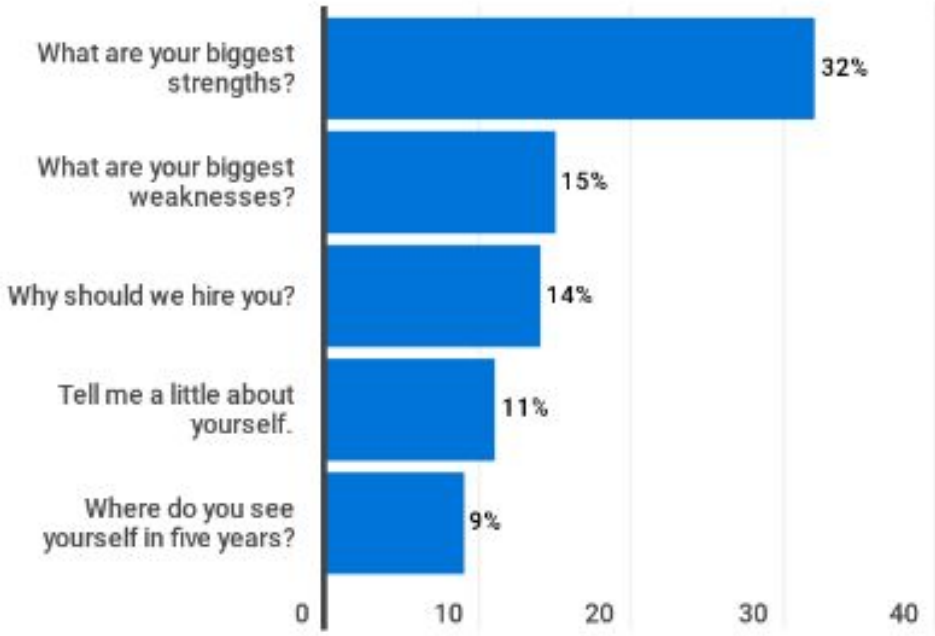
Bad vs Good Visualization



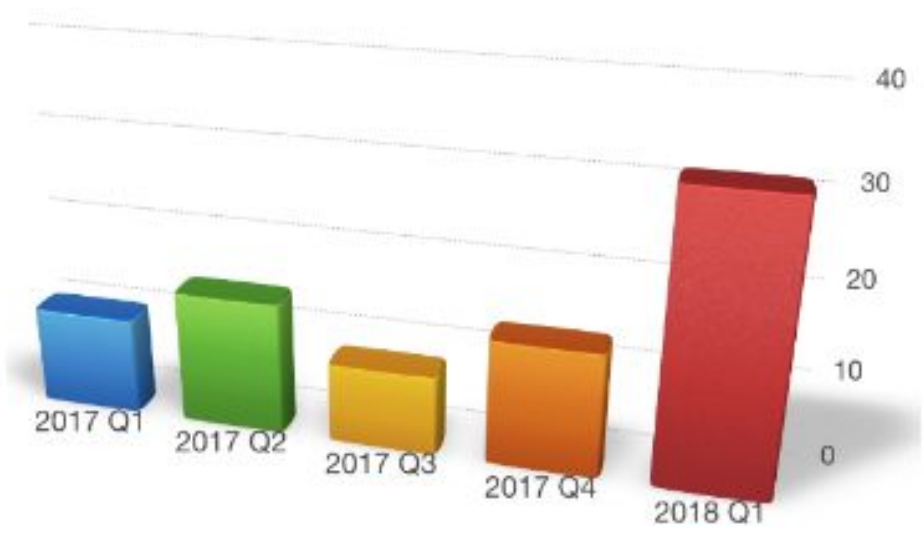


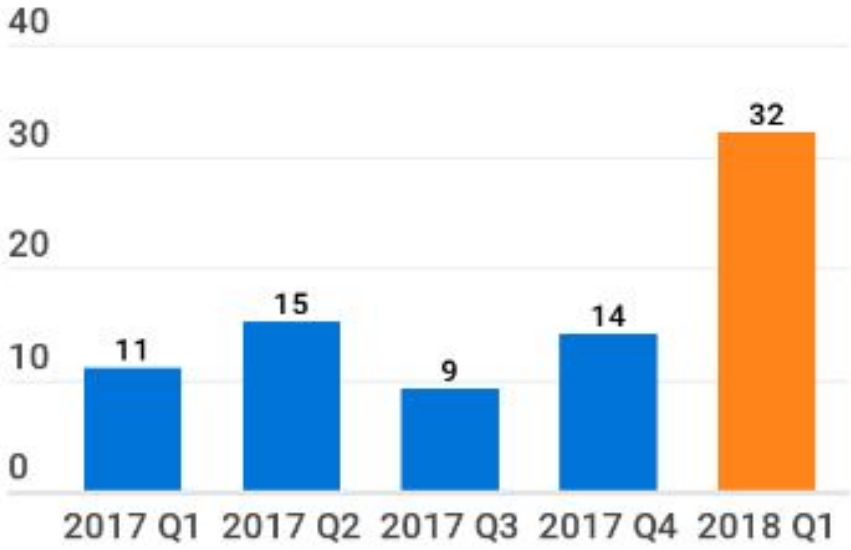
Bad vs Good Visualization



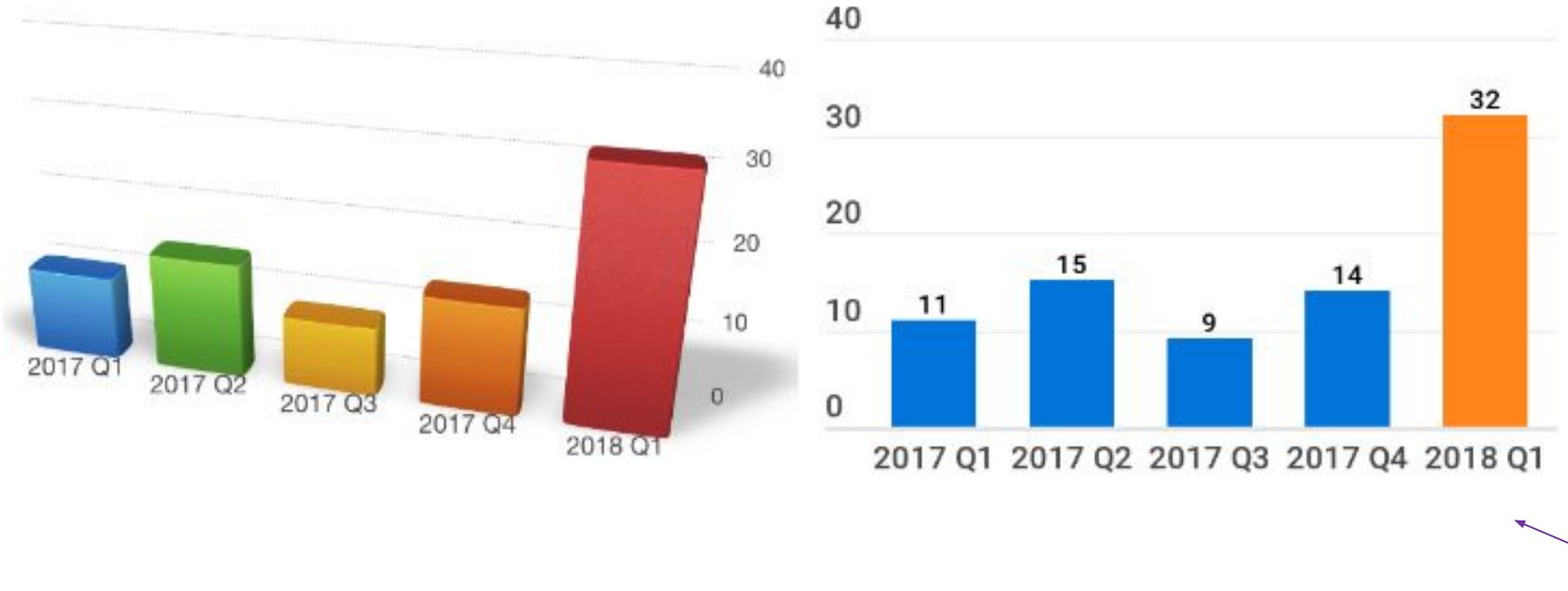


Bad vs Good Visualization

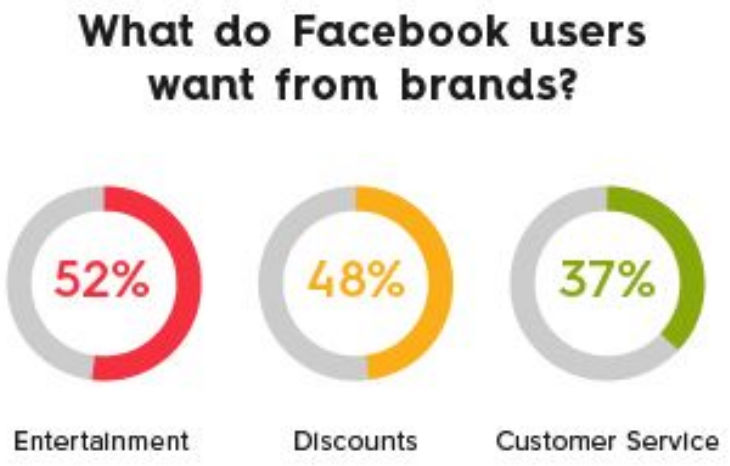


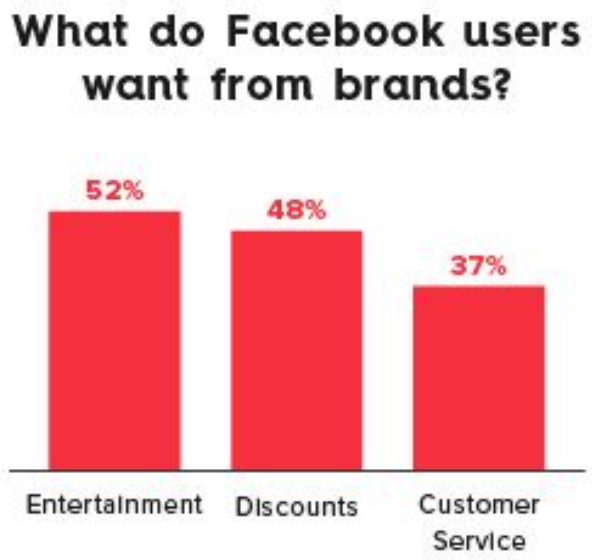


Bad vs Good Visualization



Bad vs Good Visualization

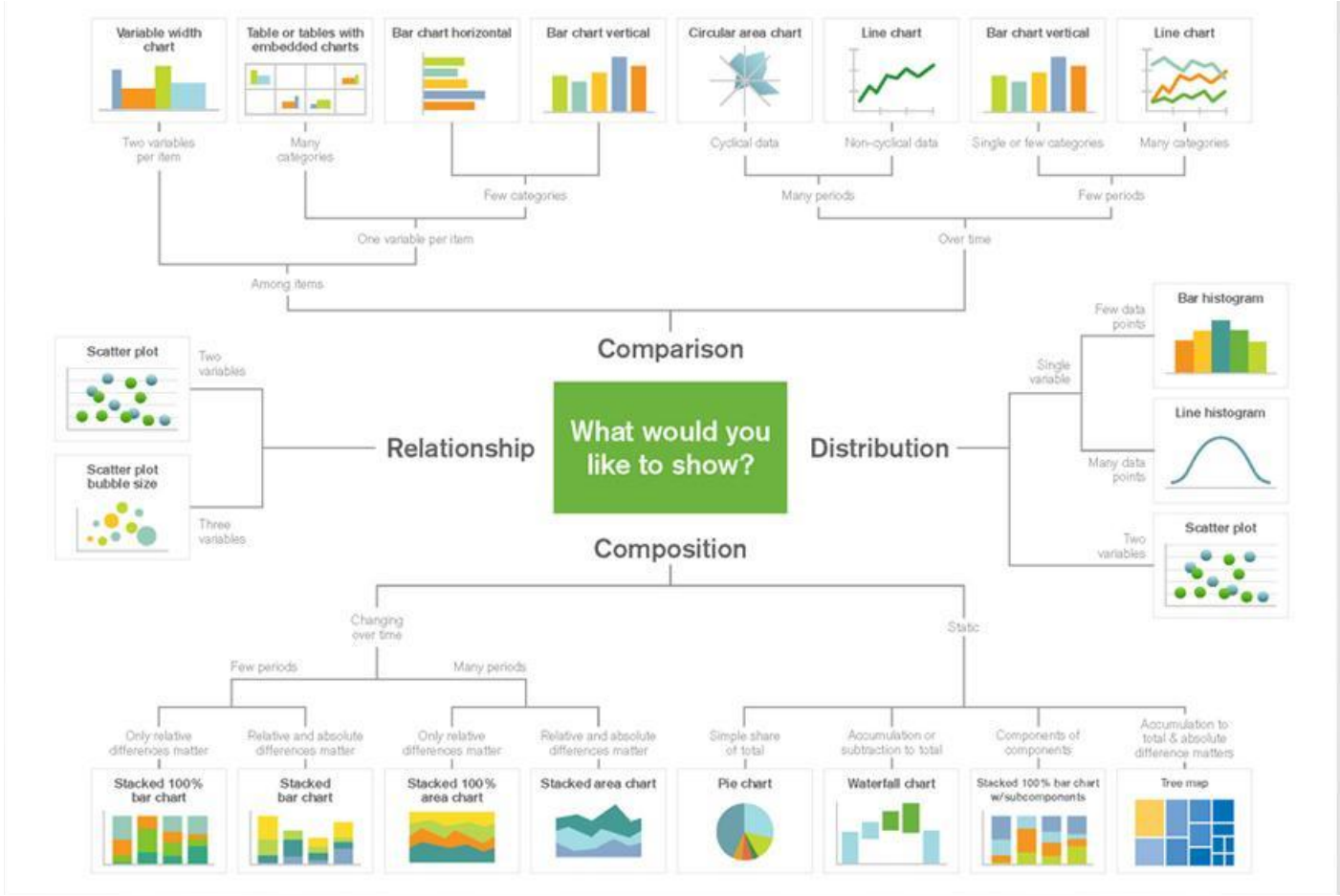




Data Visualization

There are 4 pillars of visualization:

1. **Comparison**
2. **Distribution**
3. **Relationship**
4. **Composition**



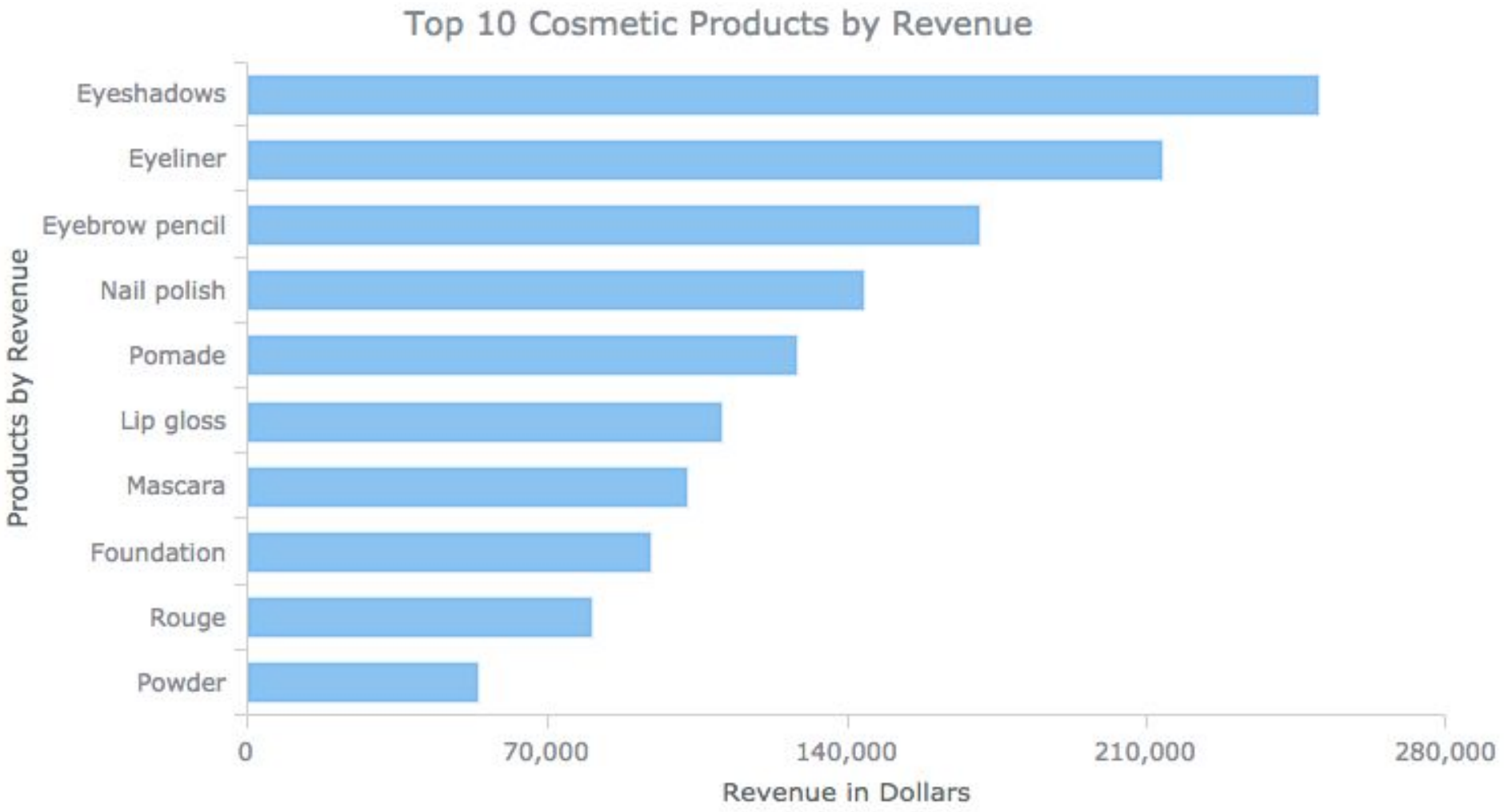
1. Comparison

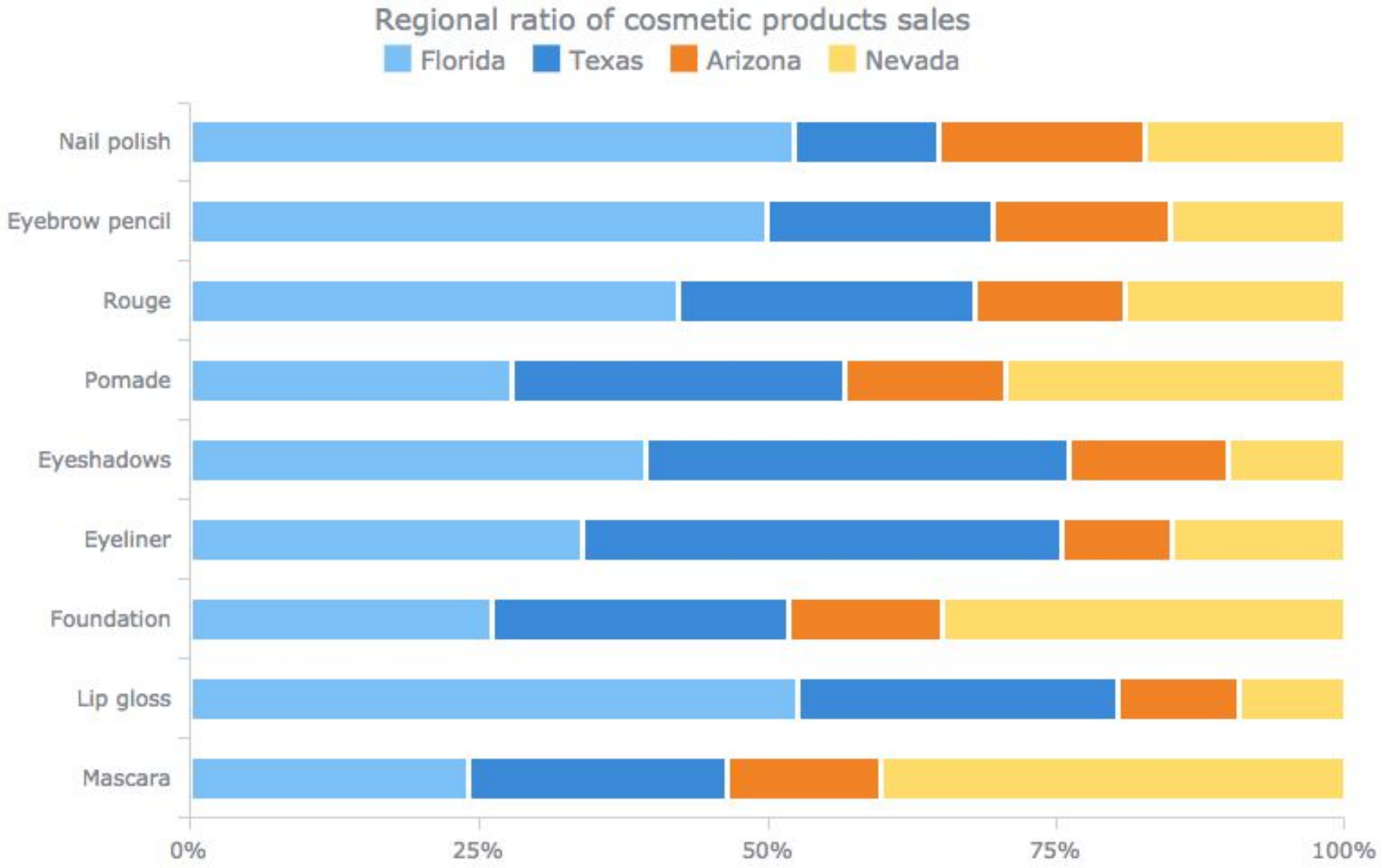
Shows the differences or similarities between values.

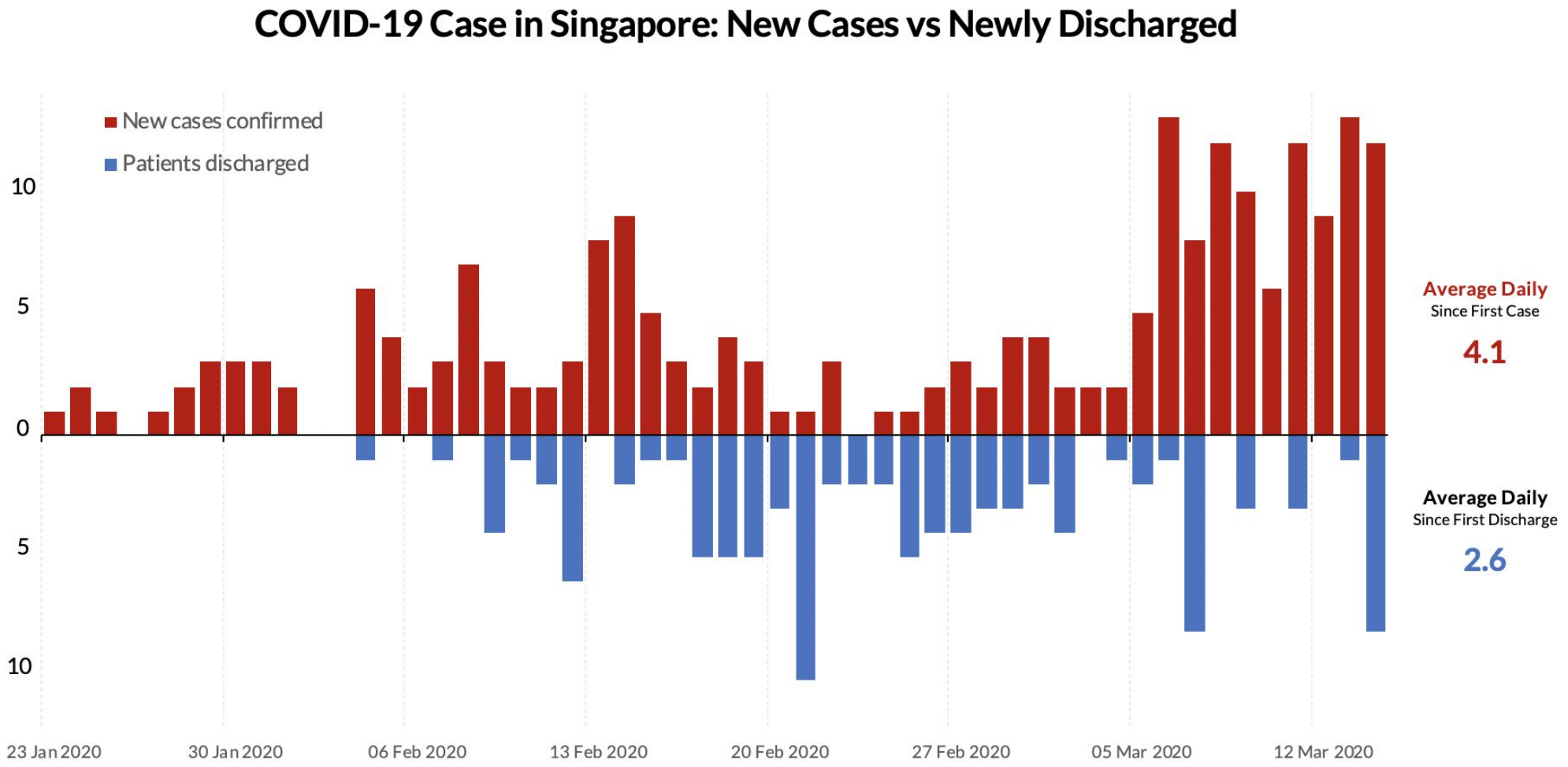
Use Cases:

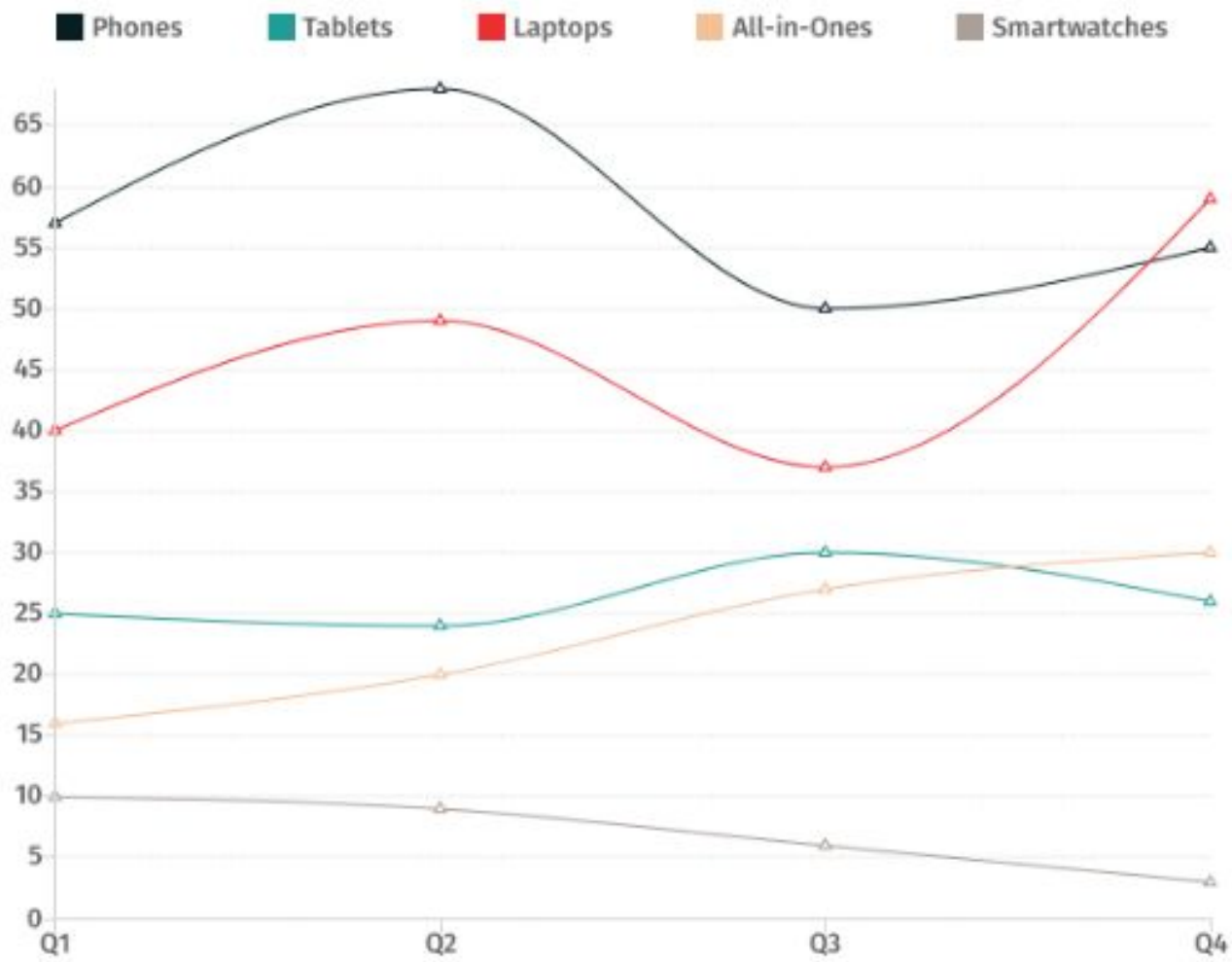
- **Total product sold in different retail branch**
- **Monthly total sales in a company**
- **Salary comparison between different type of role & expertise**

- **Example of Charts : Bar Chart, Line Chart**







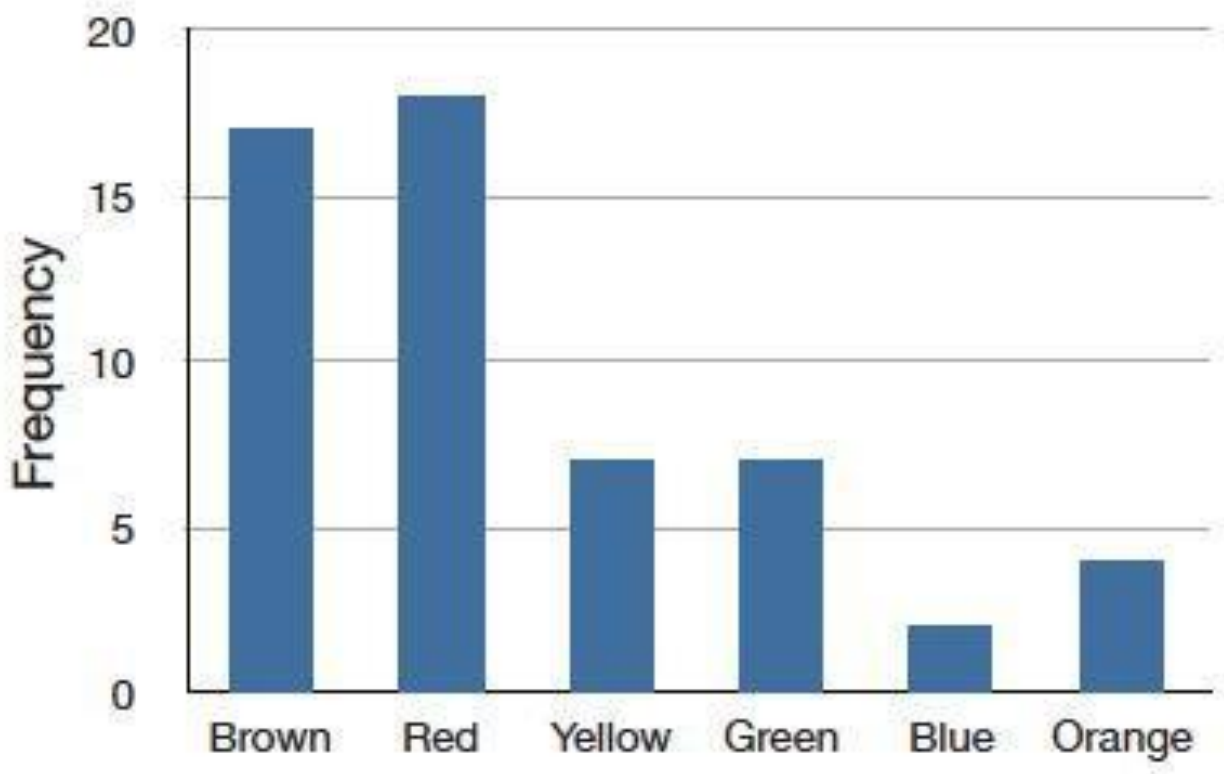


2. Distribution

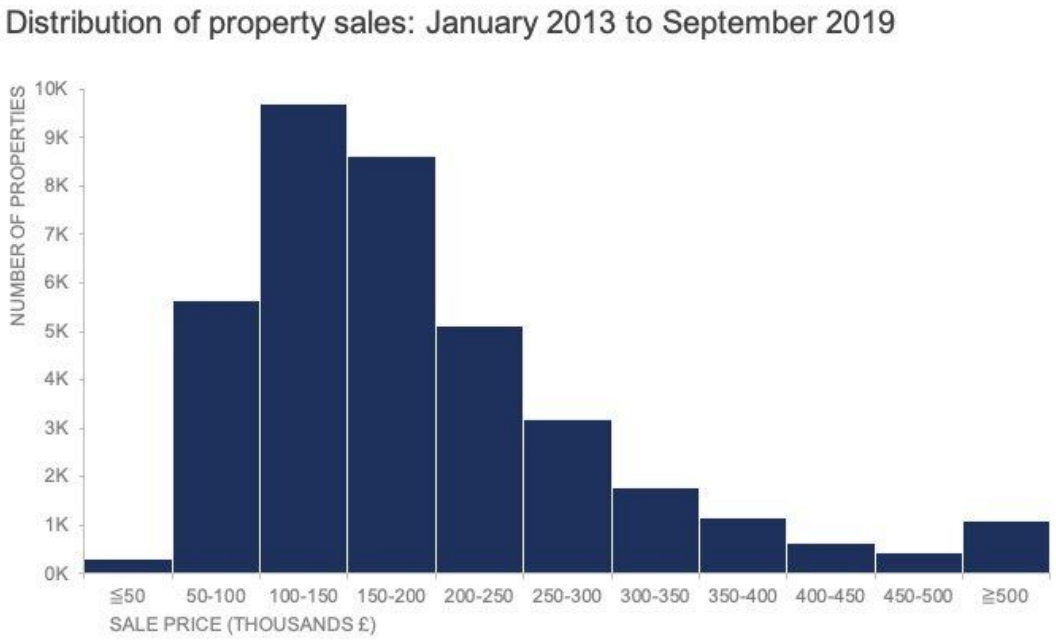
Shows the occurrence of data. Having this visualisation on hand will ease us in figuring information of the centrality, spread and skewness of data as well as pointing out the extreme values, missing or non-typical values such as outlier

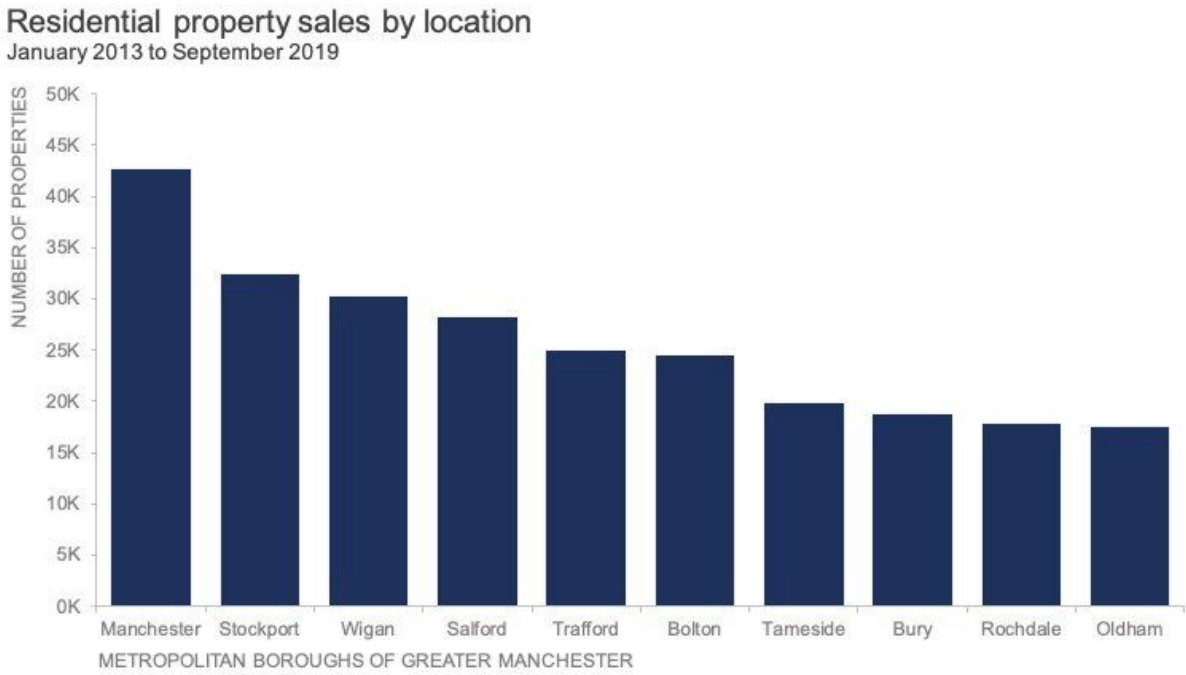
Example of Charts : Histogram, Boxplot, Scatterplot

Histogram

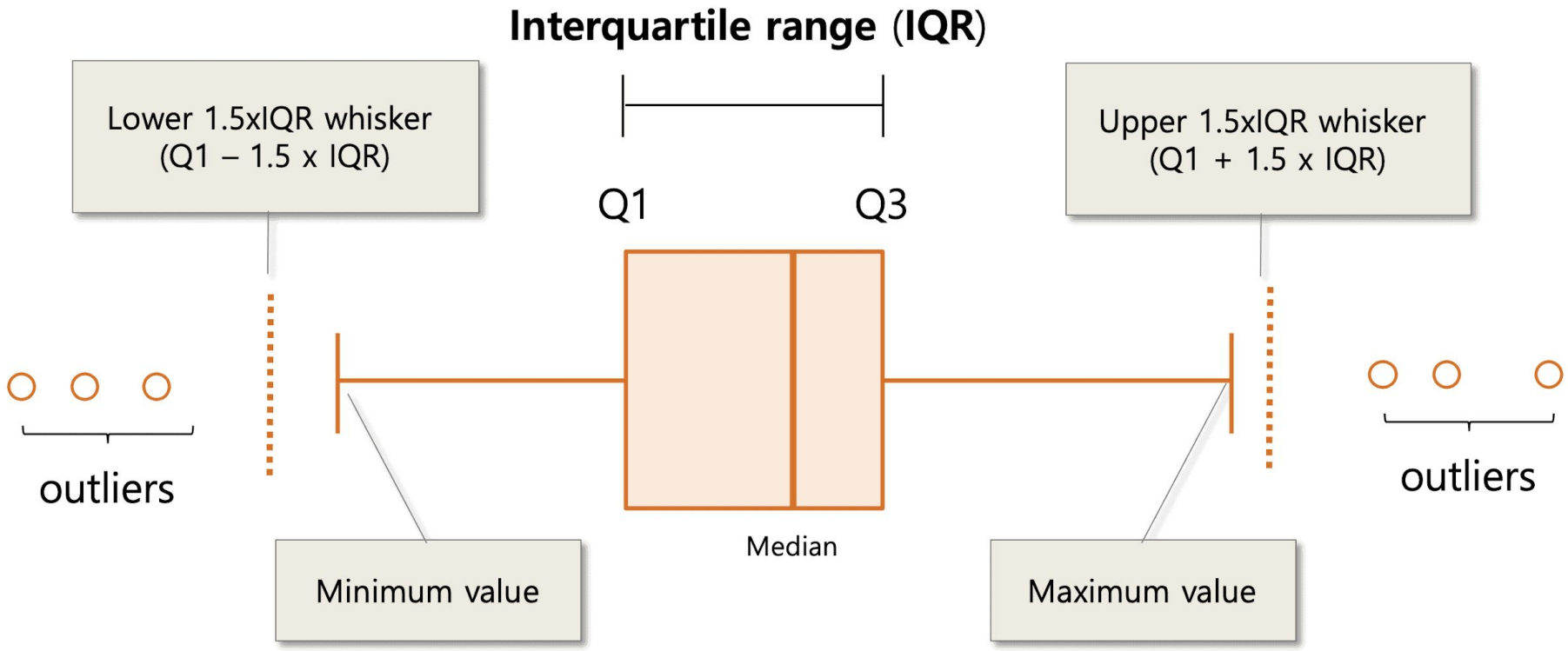


Histogram

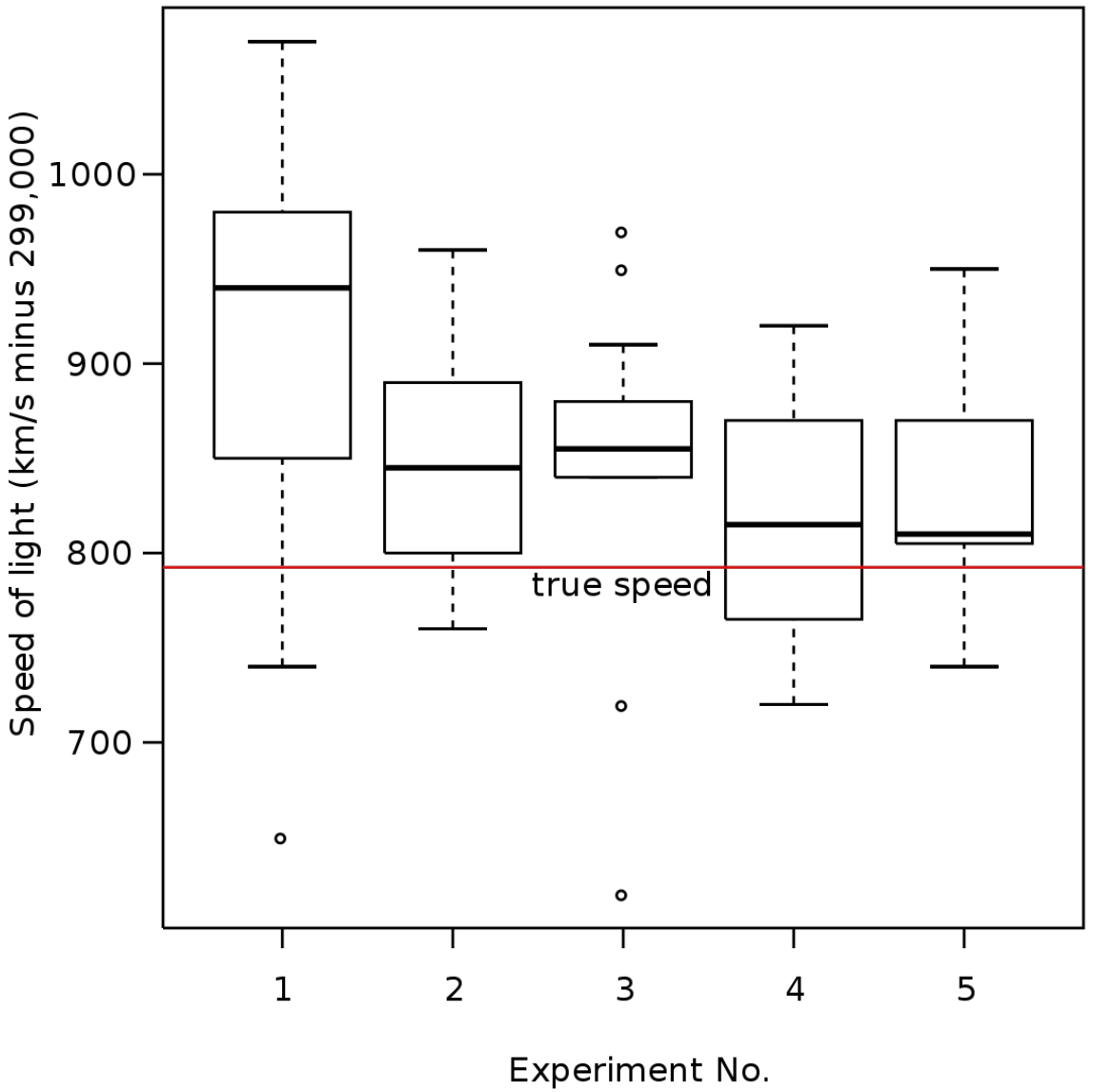




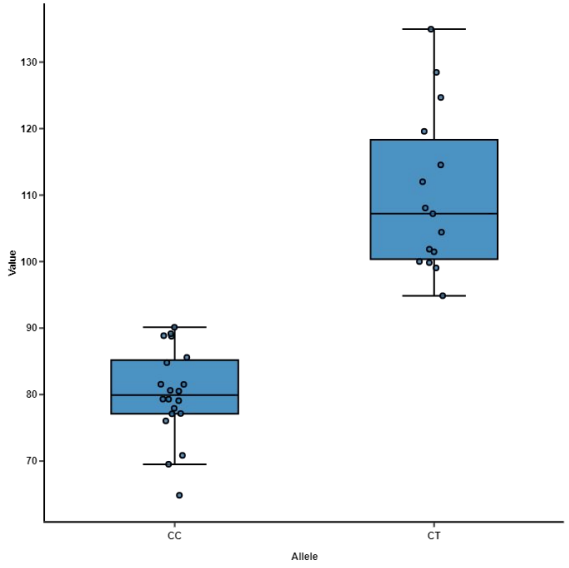
Boxplot



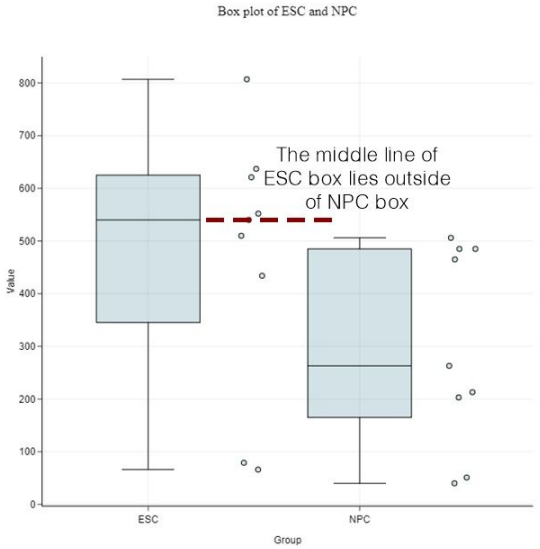
Boxplot



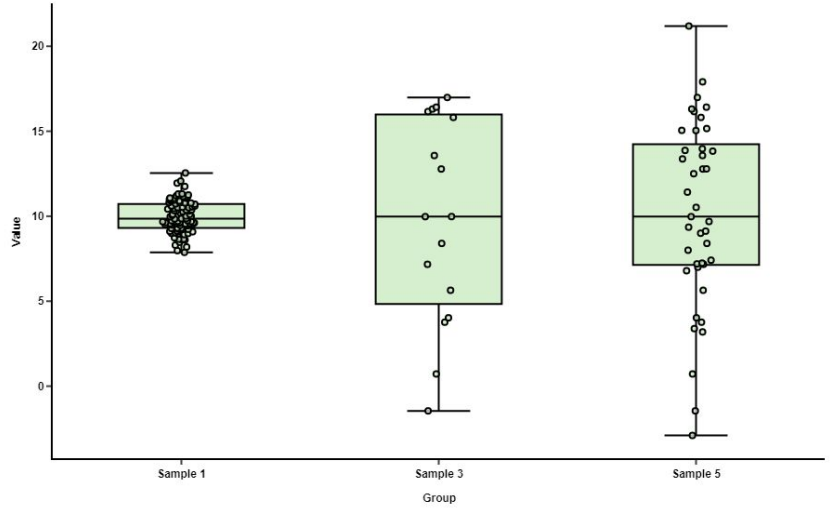
Boxplot



A

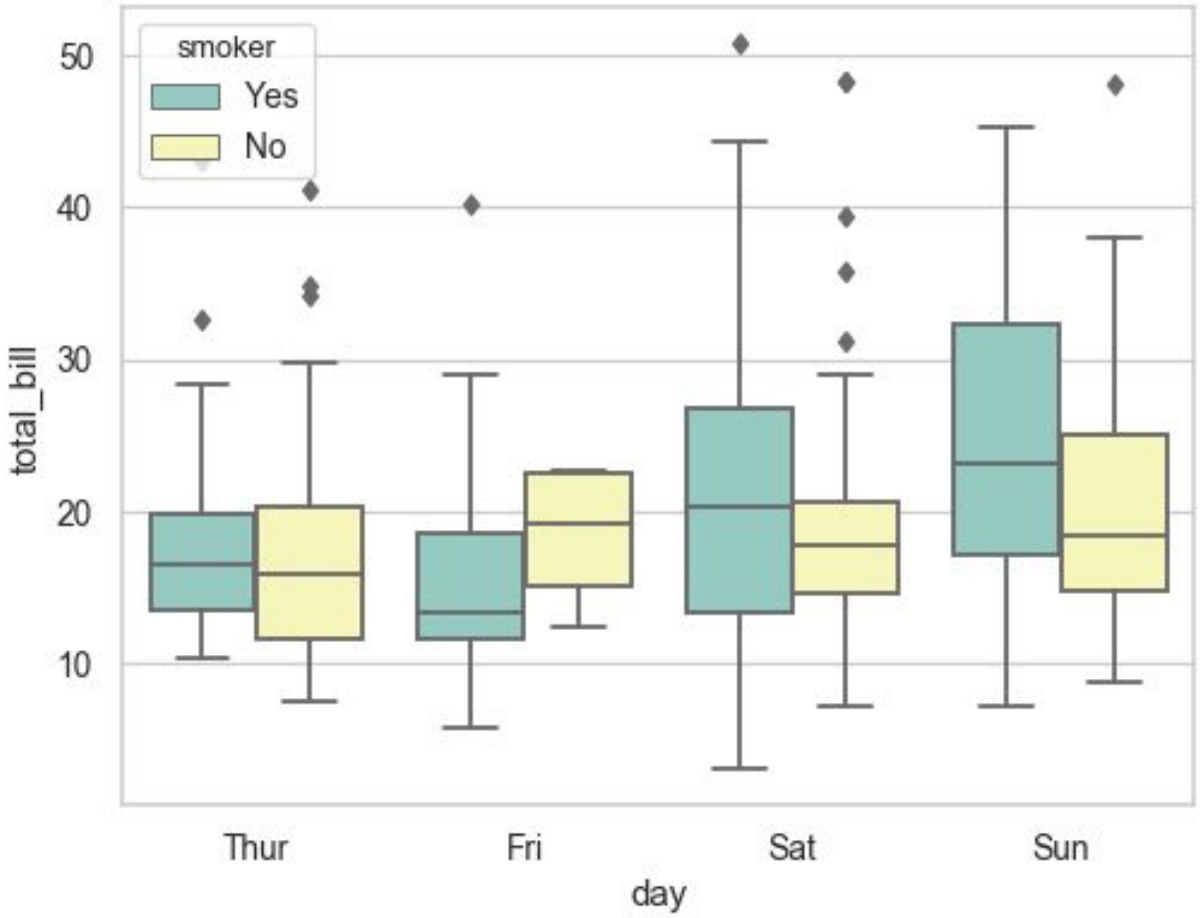


B



C

Boxplot



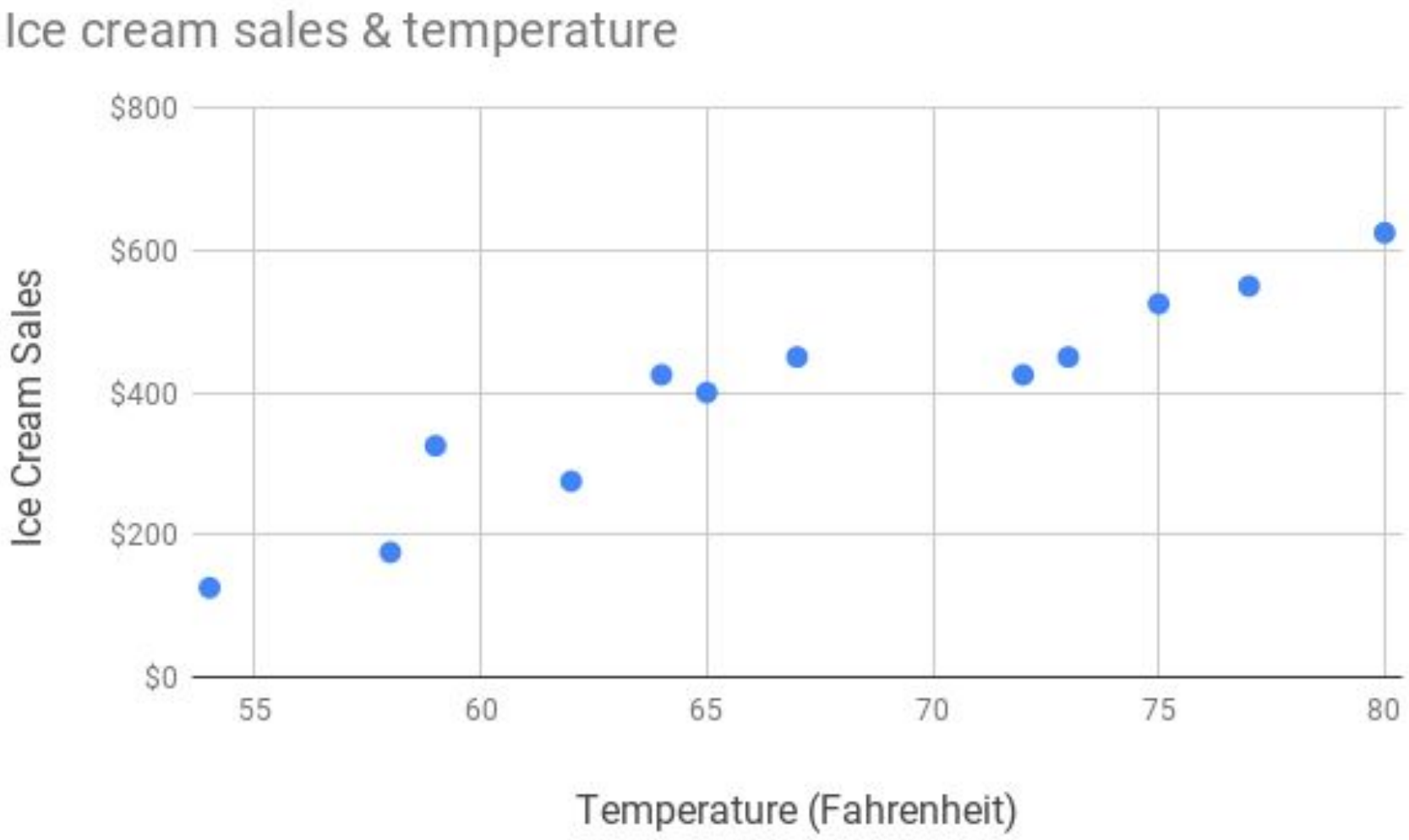
3. Relationship

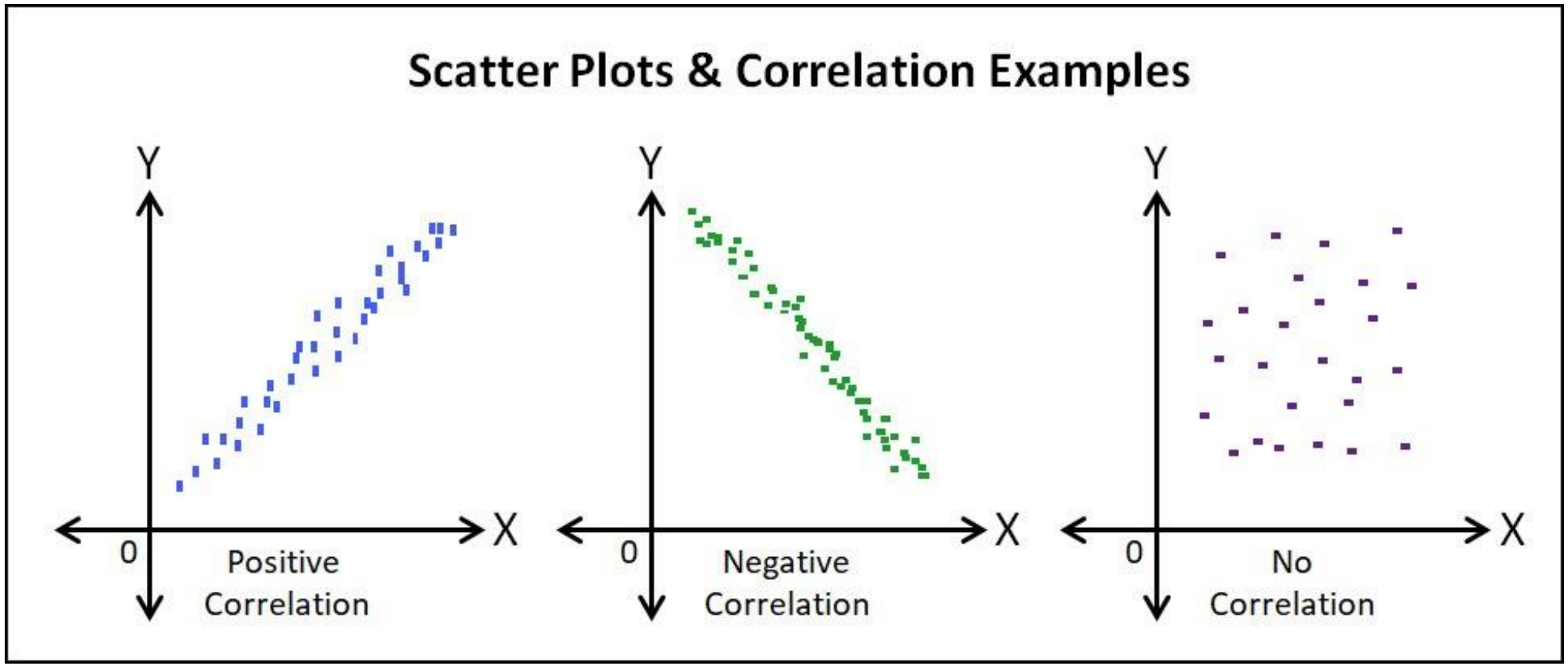
Shows the relationship and correlation between two or more variables in data.

For example :

- People are taller as they get older
- Price of house depends on the number of rooms
- Sales of Ice cream rocketed due to increase in temperature

Example of Charts : Scatterplot, Heatmap



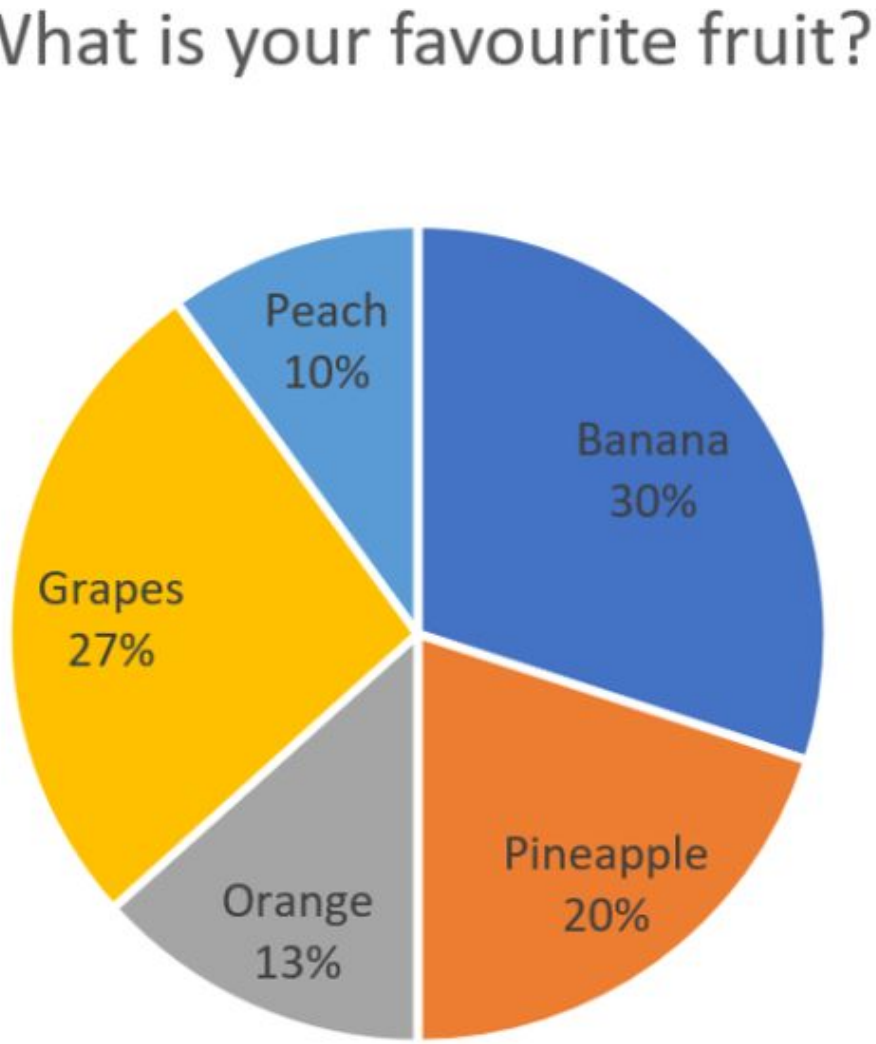


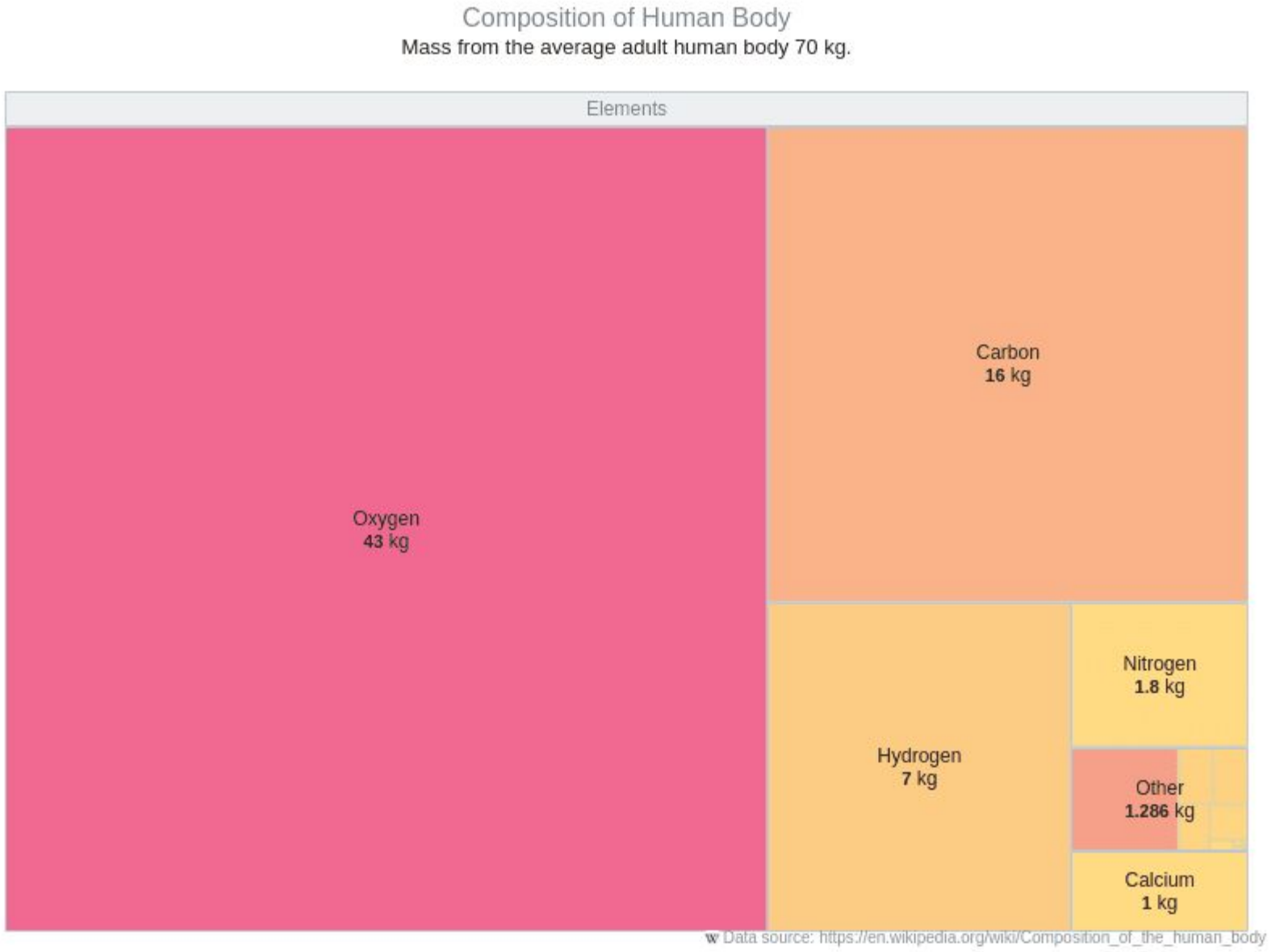
Scale of correlation coefficient	Value
$0 < r \leq 0.19$	Very Low Correlation
$0.2 \leq r \leq 0.39$	Low Correlation
$0.4 \leq r \leq 0.59$	Moderate Correlation
$0.6 \leq r \leq 0.79$	High Correlation
$0.8 \leq r \leq 1.0$	Very High Correlation

4. Composition

Shows the composition of one or more variables in absolute numbers or in percentage (portion)

Example of Charts : Pie Chart, Treemap, Stacked Bar Chart





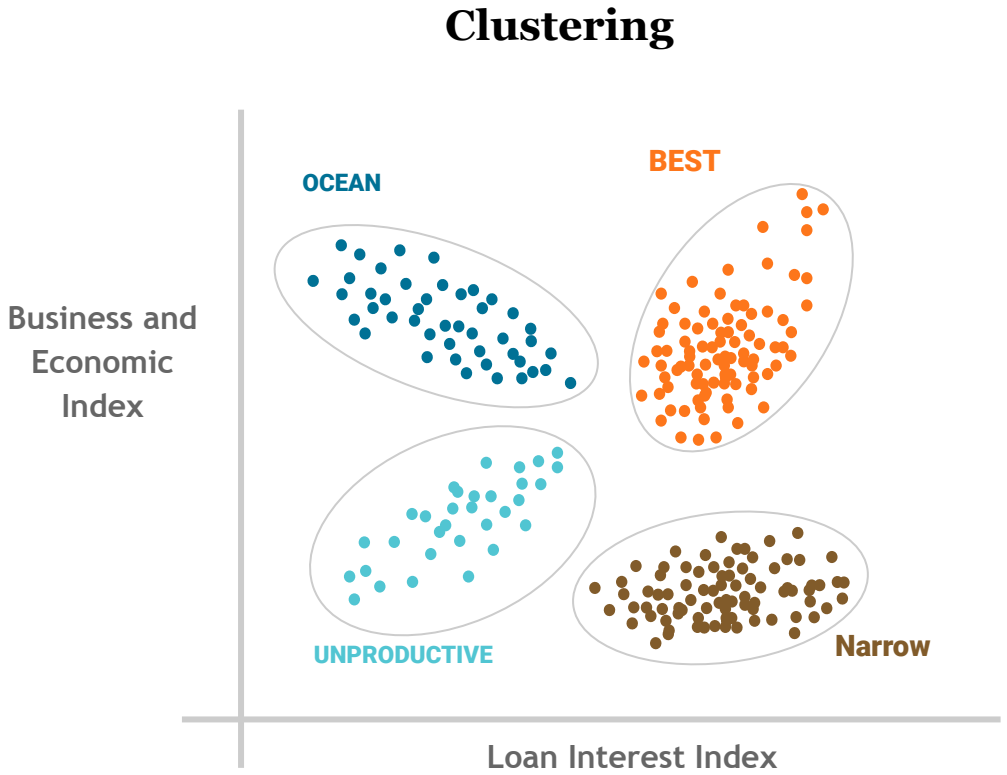
Example

Example

Loan Interest Index

Business and Economics Index

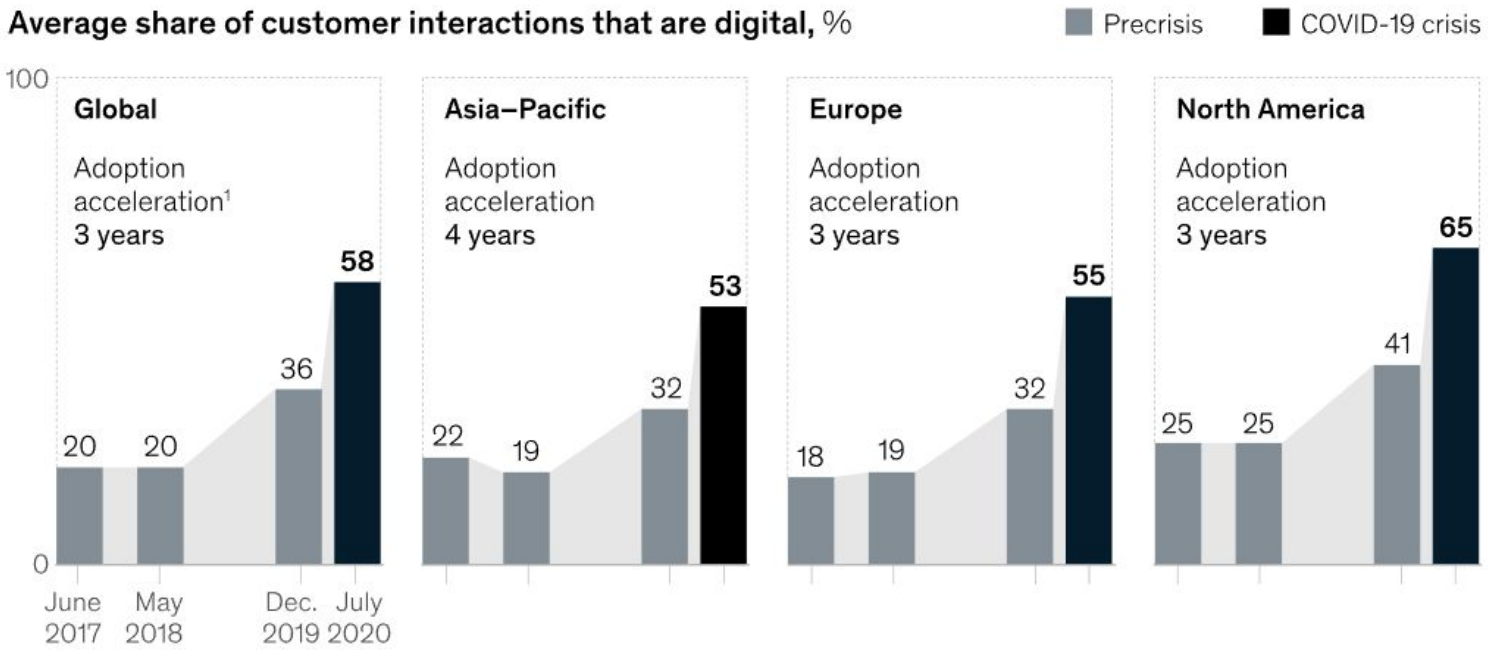
MACHINE LEARNING



Example

Example

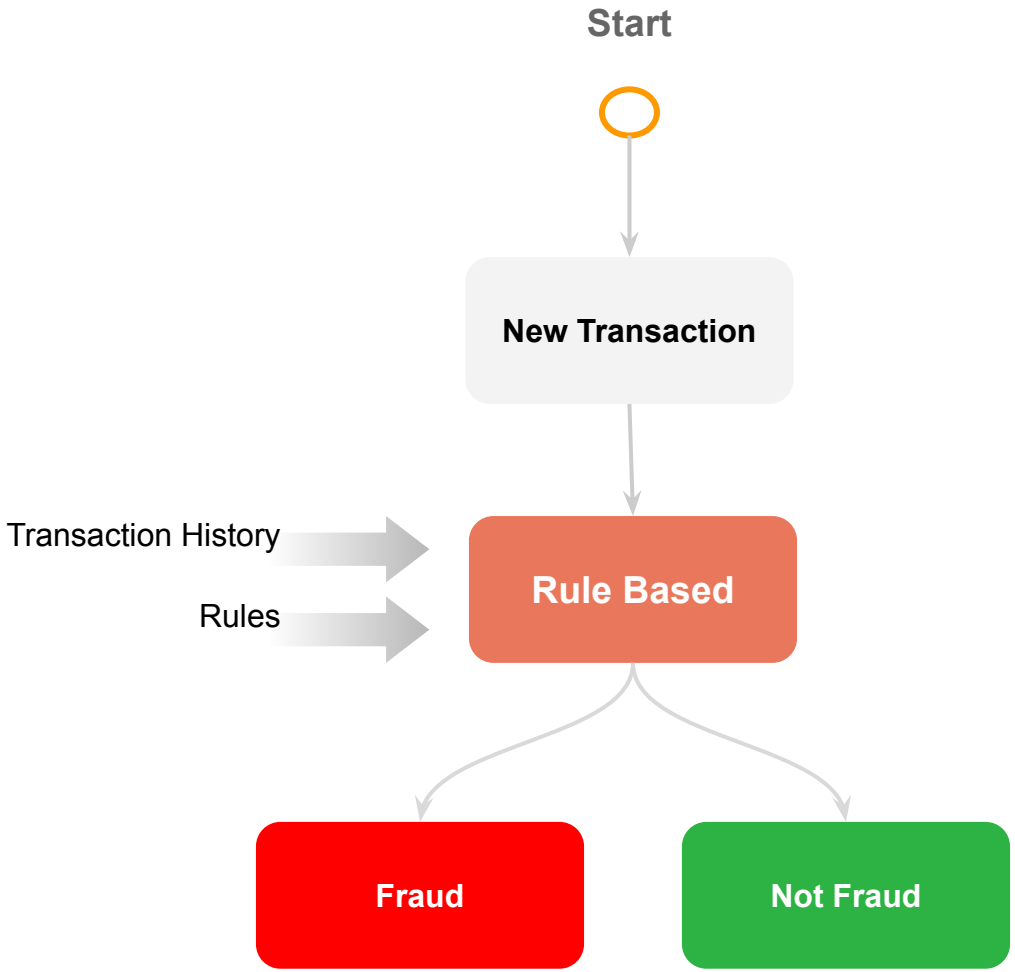
The COVID-19 crisis has accelerated the digitization of customer interactions by several years.



[1]Years ahead of the average rate of adoption from 2017 to 2019.

McKinsey & Company

Other Chart



Python Libraries for Data Visualization

Data Visualization Libraries in Python

These are commonly used Python library for data visualization :

1. Matplotlib
2. Seaborn
3. Ggplot
4. Plotly
5. Bokeh

And so on

Data Visualization Libraries in Python

These are commonly used Python library for data visualization :

1. Matplotlib
2. Seaborn
3. Ggplot
4. Plotly
5. Bokeh

And so on

Let’s Code

Reference

https://act-on.com/blog/data-visualization-101-how-to-make-better-pie-charts-and-bar-graphs/
https://blog.glisser.com/ultimate-guide-to-using-data-visualization-in-your-presentation
https://infogram.com/blog/do-this-not-that-bar-charts/
https://www.anychart.com/blog/2017/04/12/data-comparison-chart-type-visualization/
https://onlinestatbook.com/url?sa=i&url=https%3A%2F%2Fonlinestatbook.com%2F2%2Fintroduction%2Fdistributions.html&psig=AOvVaw0LDaKSlPAxCZvqN0KFgbHo&ust=1649037301165000&source=images&cd=vfe&ved=0CAsQjRxqFwoTCPi46vfk9vYCFQAAAAAdAAAAABBp
https://help.ezbiocloud.net/wp-content/uploads/2020/04/%E1%84%89%E1%85%B3%E1%84%8F%E1%85%B3%E1%84%85%E1%85%B5%E1%86%AB%E1%84%89%E1%85%A3%E1%86%BA-2020-04-19-%E1%84%8B%E1%85%A9%E1%84%8C%E1%85%A5%E1%86%AB-9.41.35.png
https://s3.amazonaws.com/app-production-uploads/2020-04-20/5e66b6bcbc42fed480182615/5e9db7f506adee8c4a958a3d_rich-text-file-2020-04-20T14-55-4-080Z.imagepng
https://static.anychart.com/images/gallery/v8/tree-map-charts-composition-of-the-human-body.png