Introduced By — *John Hunter*
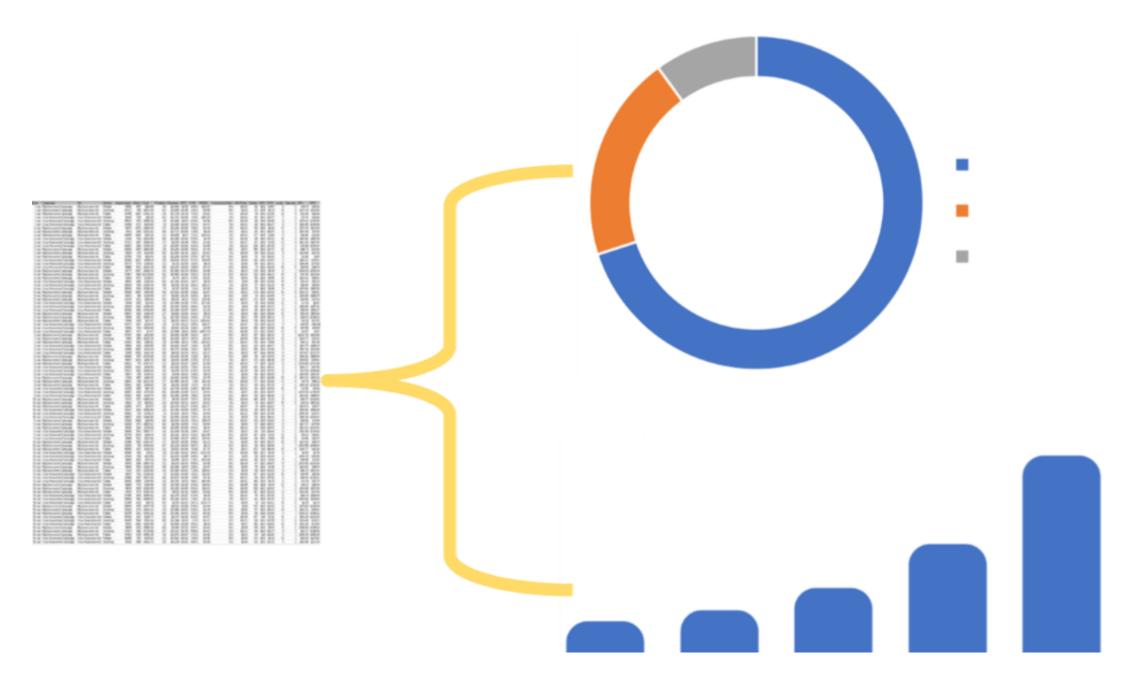
# MATPLOTLIB

**1** Why Data Visualization

**2** What is Matplotlib

**3** Types of plots

**4** Demo on Plots

**5** Matplotlib vs Seaborn

CONTENT
FOR
MATPLOTLIB

**Data Visualization is the method to introduce the information in pictorial or graphical format**

# Advantages of
# Data Visualization

We convey a great number of
information in an effective manner

**Simplifies statistics**

Top 10 Sale across the
globe, Top performing
products.

**Easy to understand**

We can grab more insights
compared to tabluar format.

*Hidden Patterns*

Use of data visualization,
patterns, and trends that
have been hiding in the
shadows of untapped data,
can be brought to the light
of day and acted upon

**Makes an impact**

Data visualization can
materially impact the
speed of decision-making
processes.

# What is Matplotlib?

**MATPLOTLIB IS A PYTHON LIBRARY UTILIZED FOR PLOTTING THE EXCELLENT AND APPEALING GRAPHS**

In Data science visualization is the significant advance. By utilizing visualization we can undoubtedly comprehend than how information is part.

There are loads of libraries to envision the information however among that the matplotlib is exceptionally well known and simple one.

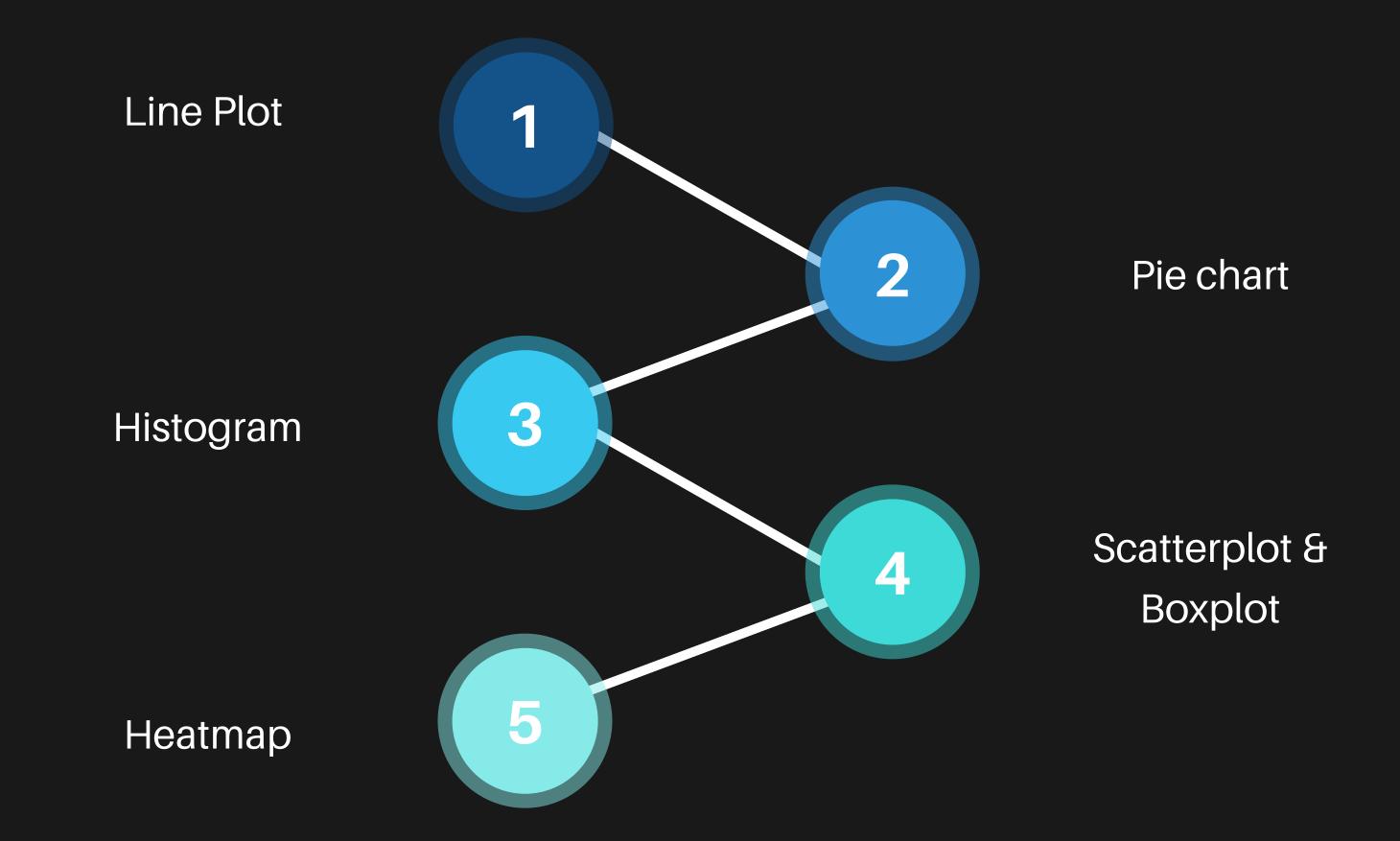# MOST USED PLOTS FOR BASIC ANALYSIS
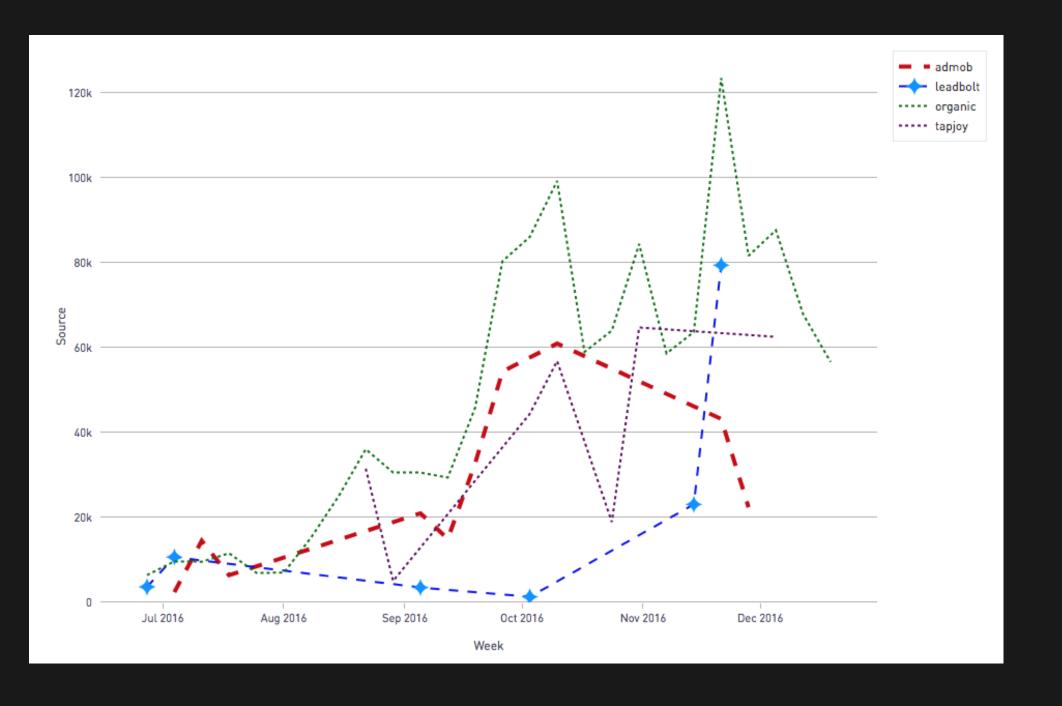
## UNI- VARIATE

**Histpgram,
Countplot, Pie-chart**

## BI-VARIATE

**Scatter-Plot , Box-plot, Heatmap**

## MULTI-VARIATE

**Line- plot, Bar plot**

# Line plot



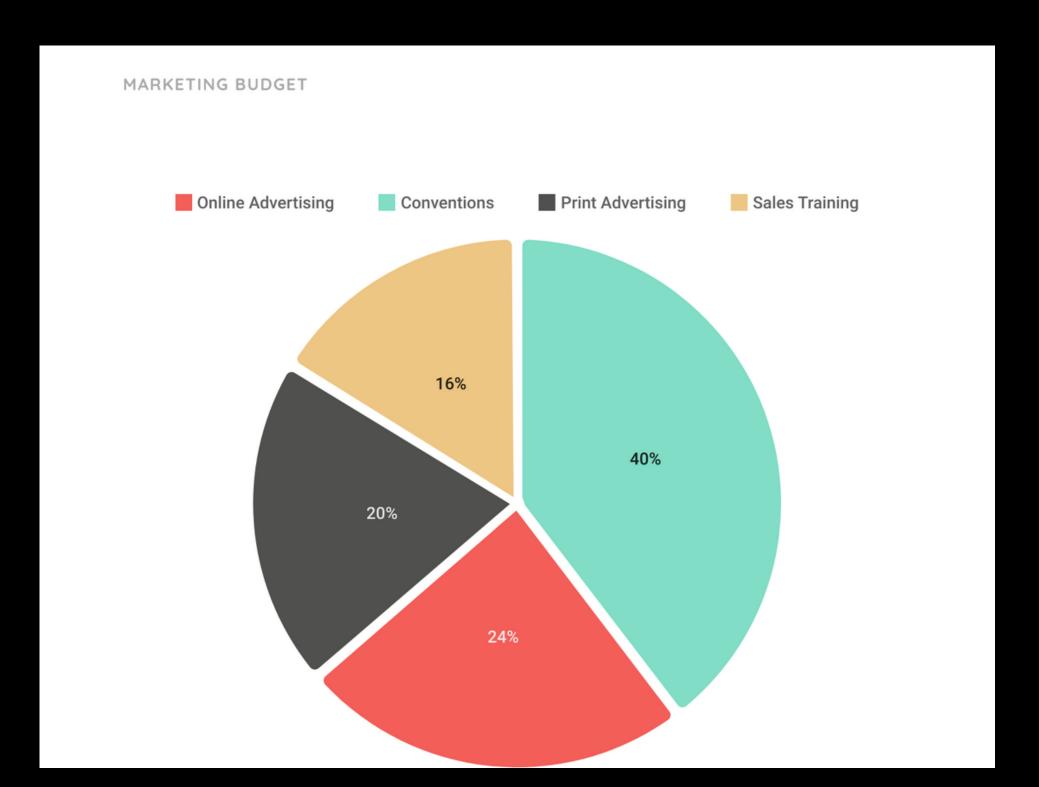**Type of Data Required to plot:**
- **Time Series**

Line charts are utilized to follow changes over short and extensive stretches of time. At the point when smaller changes exist, line charts are smarter to utilize than Bar Plots. Line charts can likewise be utilized to analyze changes over a similar timeframe for more than one group

**Advantages of Line Graphs**
- Line graphs are easy to read and to plot.
- They are useful for making comparisons between different data sets.
- They are useful for showing changes over periods of time

# Pie Chart

A pie chart is a type of graph in which a circle is divided into sectors that each represents a proportion of the whole. Pie charts are a useful way to organize data in order to see the size of components relative to the whole, and are particularly good at showing percentage or proportional data.
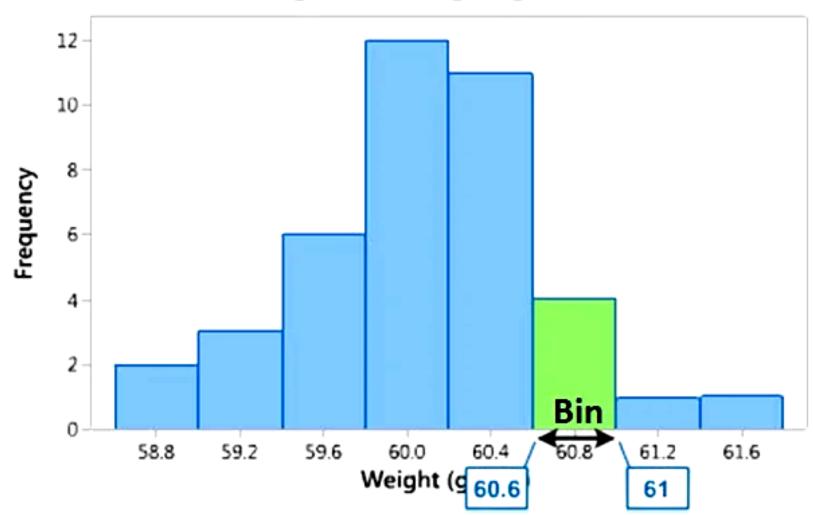
Type of Data required to plot
Discrete + Countinues



MARKETING BUDGET

■ Online Advertising    ■ Conventions    ■ Print Advertising    ■ Sales Training

16%

40%

20%

24%

**Histogram is a graphical representations of the frequency distribution of the data in bar form**

- A "histogram" is used for plotting the occurrences of score frequency in a **"continuous data set"**. This data set is further divided into classes and they are referred as bins
- Summarizes data from a process that has been collected over a period of time
- Provides a quick representation of the spread and centering of the process
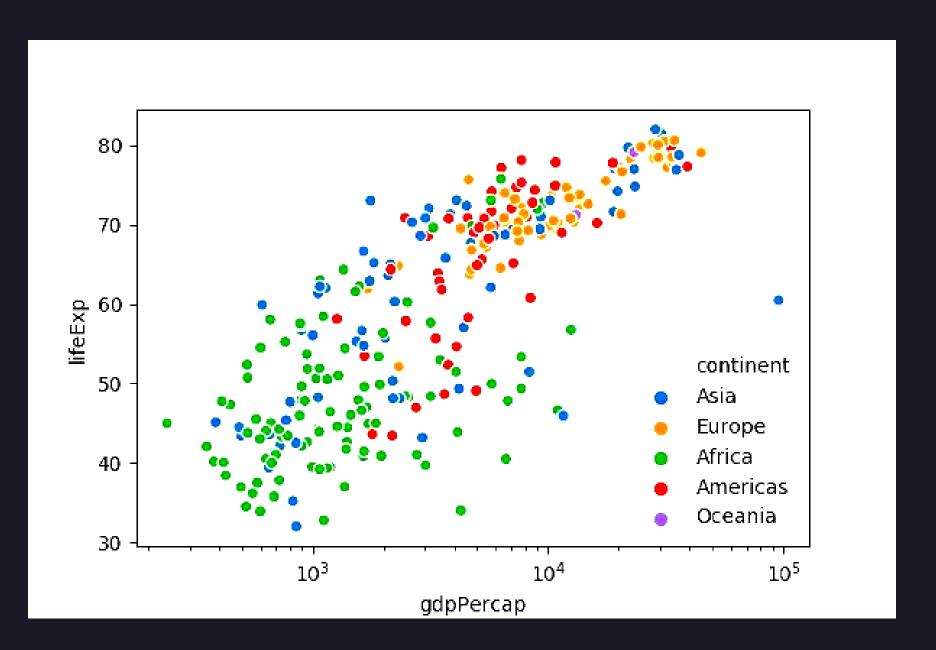- To display large amounts of data values in a relatively simple chart form

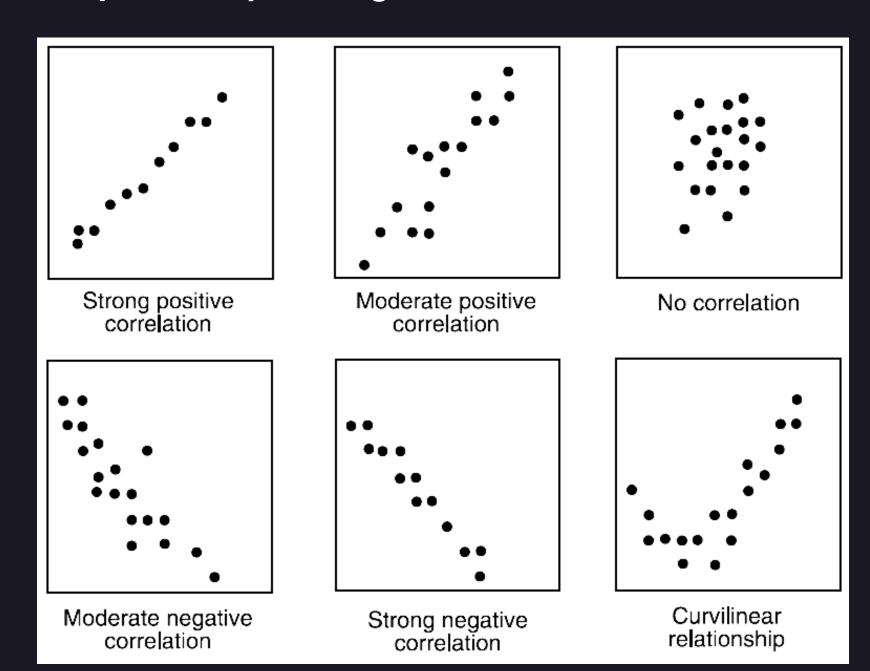**Type of Data required to plot:**
- Continues Data

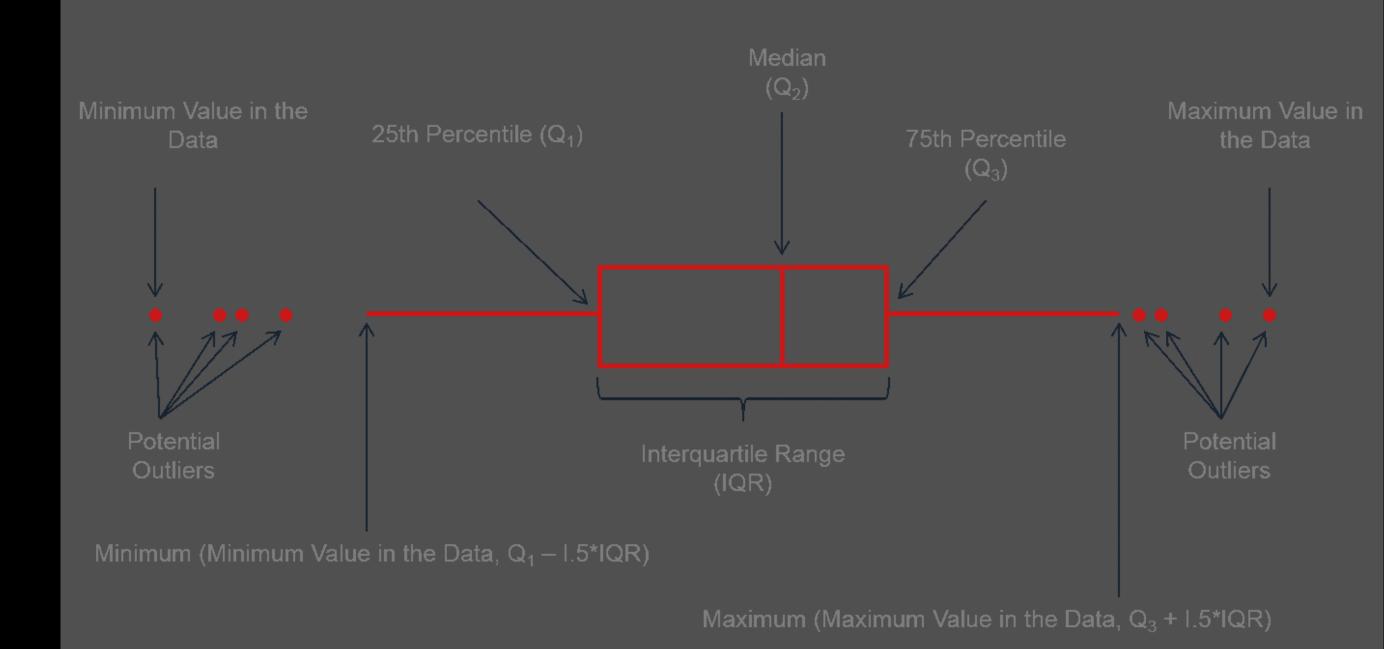# A scatterplot is one of the best plot to find the relationship between two numerical variables



1. Shows the correlation between the variables .

2. Shows the minimum/maximum and outliers in the dataset
3. Easy to find the clusters.
4. Easy to interpret large datasets



Types of Data Required to plot:
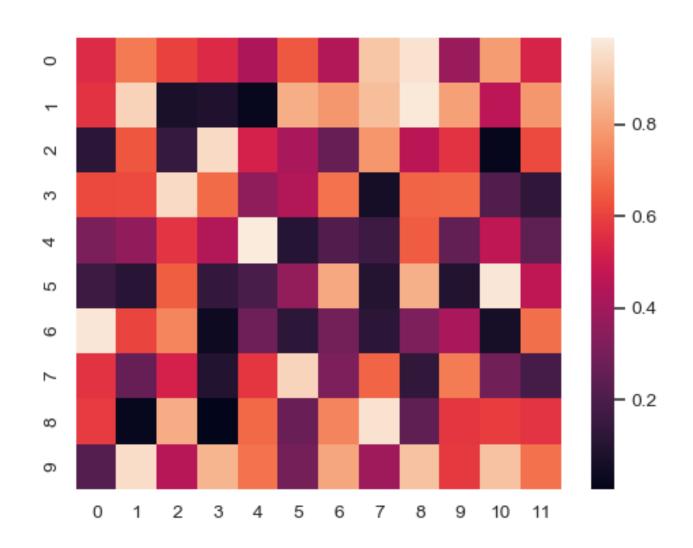- Continues VS Continues

A boxplot is a standardized way of displaying the distribution of data based on a five number summary ("minimum", first quartile (Q1), median, third quartile (Q3), and "maximum"). It can tell you about your outliers and what their values are. It can also tell you if your data is symmetrical, how tightly your data is grouped, and if and how your data is skewed.

**Type of Data required to plot.**
- **Continues vs Continues**
- **Discrete Vs continues**

Minimum Value in the Data

25th Percentile ($Q_1$)

Median ($Q_2$)

75th Percentile ($Q_3$)

Maximum Value in the Data

Potential Outliers

Potential Outliers

Minimum (Minimum Value in the Data, $Q_1 - 1.5*IQR$)

Interquartile Range (IQR)

Maximum (Maximum Value in the Data, $Q_3 + 1.5*IQR$)

# HEAT MAP



A heat map is a two-dimensional representation of information with the help of colors. Heat maps can help the user visualize simple or complex information.

Type of **Data Required** to plot .
- Temp
- Correlation coefficient
- Pressure
- Market Budget

# MATPLOTLIB VS SEABORN

## Matplotlib

1. Matplotlib is mainly used for basic plotting.
2. It is highly customizable and powerful
3. It is multi-platform Datavisualization tool, therefore it is fast and efficient

## Seaborn

1. Seaborn will provide a variety of visualization patterns.
2. It has tools, built-in Statistical functions which helps us to find hidden patterns in the dataset
3. It have builtin themes for better Visualization