

# What is Matplotlib?

Matplotlib is an amazing visualization library in Python for 2D plots of arrays.

Matplotlib is a multi-platform data visualization library built on NumPy arrays and designed to work with the broader SciPy stack. It was introduced by John Hunter in the year 2002.

One of the greatest benefits of visualization is that it allows us visual access to huge amounts of data in easily digestible visuals.

Matplotlib consists of several plots like line, bar, scatter, histogram etc.

## **Getting started with Matplotlib**

### **Installation :**

Windows, Linux and macOS distributions have matplotlib and most of its dependencies as wheel packages.

Run the following command to install `matplotlib` package :

```
python -mpip install -U matplotlib
```

### **Importing matplotlib :**

```
from matplotlib import pyplot as plt
```

*or*

```
import matplotlib.pyplot as plt
```

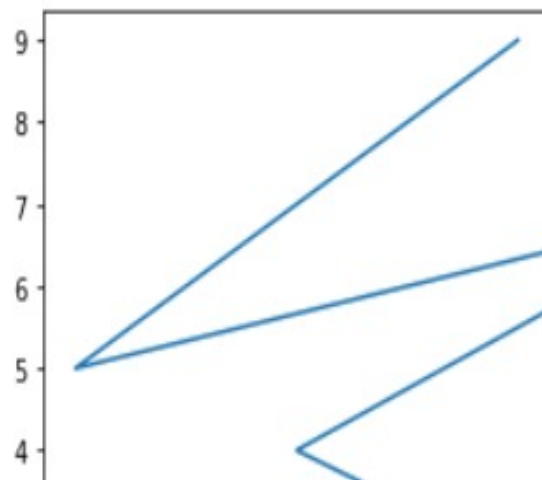
## **Basic plots in Matplotlib :**

Matplotlib comes with a wide variety of plots. Plots helps to understand trends, patterns, and to make correlations.

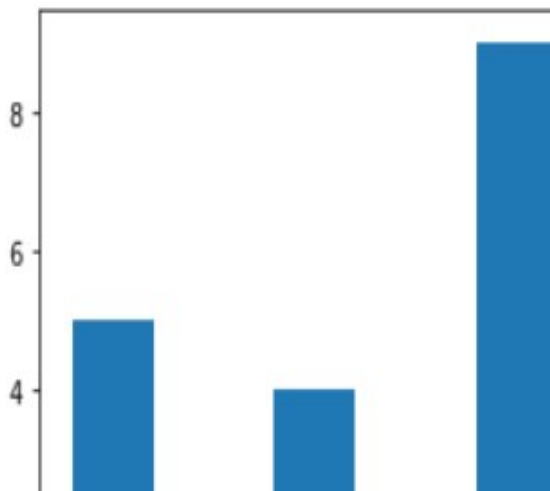
They're typically instruments for reasoning about quantitative information.

Some of the sample plots are covered here.

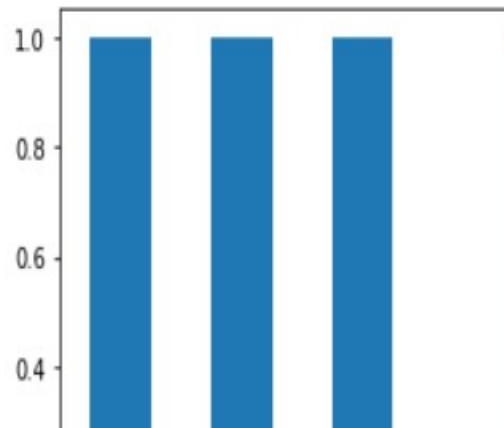
### **Line plot :**



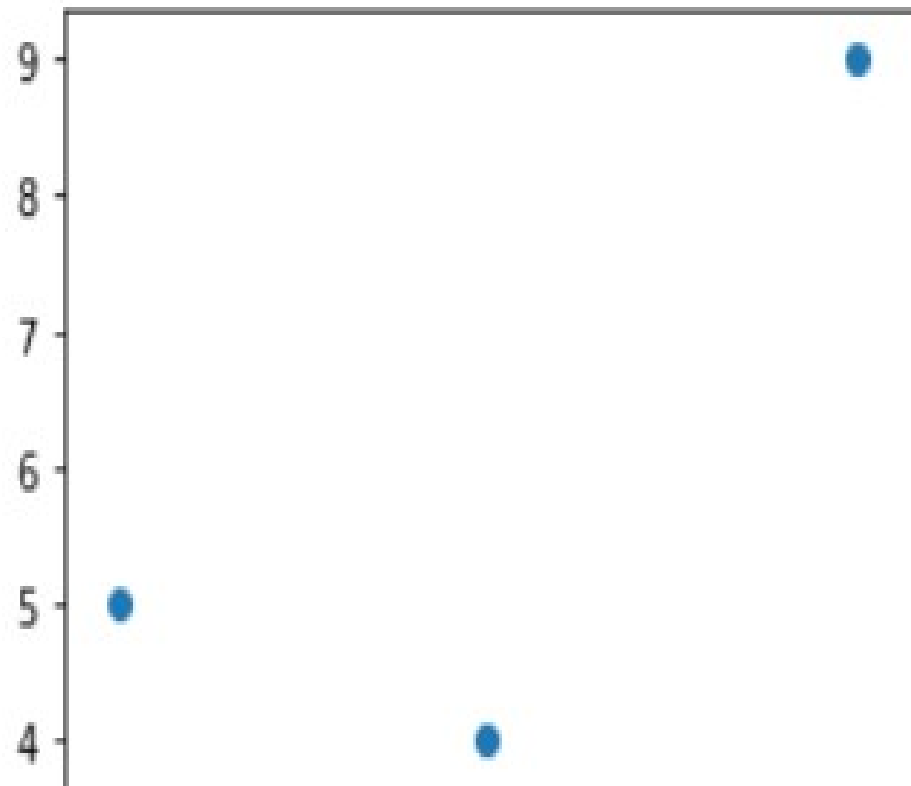
**Bar plot :**



**Histogram :**



## Scatter Plot :



## **What is Seaborn?**

Seaborn is an amazing visualization library for statistical graphics plotting in Python. It provides beautiful default styles and color palettes to make statistical plots more attractive. It is built on the top of matplotlib library and also closely integrated to the data structures from pandas.

Seaborn aims to make visualization the central part of exploring and understanding data. It provides dataset-oriented APIs, so that we can switch between different visual representations for same variables for better understanding of dataset.

## **Different categories of plot in Seaborn**

**Relational plots:** This plot is used to understand the relation between two variables.

**Categorical plots:** This plot deals with categorical variables and how they can be visualized.

**Distribution plots:** This plot is used for examining univariate and bivariate distributions

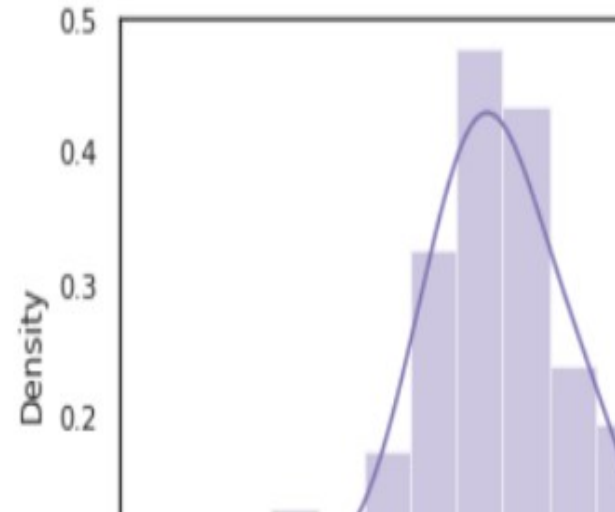
**Regression plots:** The regression plots in seaborn are primarily intended to add a visual guide that helps to emphasize patterns in a dataset during exploratory data analyses.

**Matrix plots:** A matrix plot is an array of scatter plots.

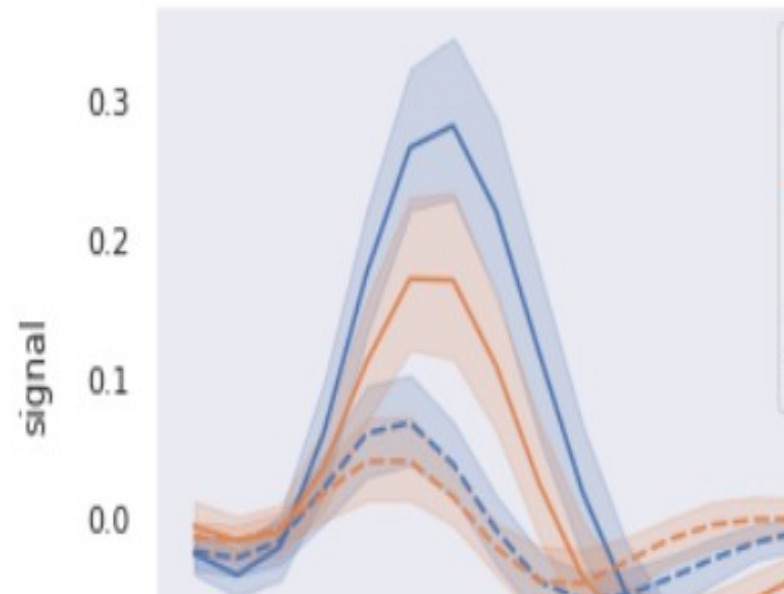
**Multi-plot grids:** It is an useful approach is to draw multiple instances of the same plot on different subsets of the dataset.

## Some basic plots using seaborn

**Dist plot** : Seaborn dist plot is used to plot a histogram, with some other variations like kdeplot and rugplot.

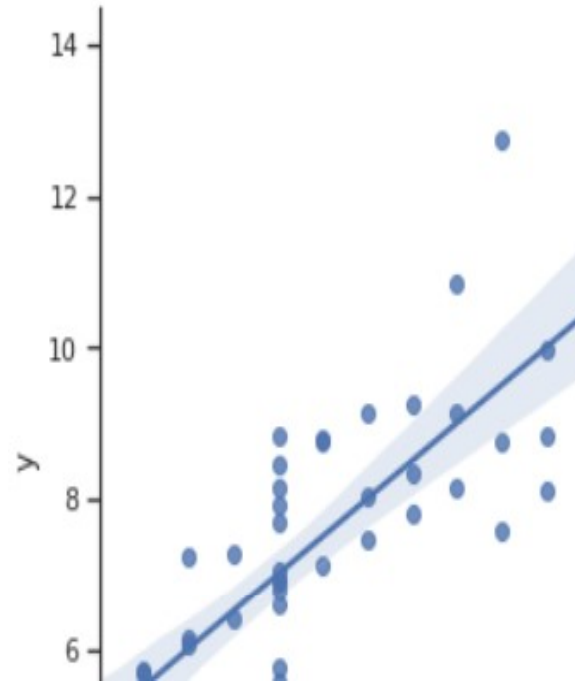


**Line plot** : The line plot is one of the most basic plots in the seaborn library. This plot is mainly used to visualize data in the form of time series, i.e. in a continuous manner.





**Lmplot** : The Lmplot is another most basic plot. It shows a line representing a linear regression model along with data points on the 2D-space and x and y can be set as the horizontal and vertical labels respectively.



## **Matplotlib vs Seaborn**

### **1.Functionality:**

**Matplotlib:** Matplotlib is mainly deployed for basic plotting. Visualization using Matplotlib generally consists of bars, pies, lines, scatter plots and so on.

**Seaborn:** Seaborn, on the other hand, provides a variety of visualization patterns. It uses fewer syntax and has easily interesting default themes. It specializes in statistics visualization and is used if one has to summarize data in visualizations and also show the distribution in the data.

## **2.Handling Multiple Figures:**

**Matplotlib:** Matplotlib has multiple figures can be opened, but need to be closed explicitly. `plt.close()` only closes the current figure. `plt.close('all')` would close em all.

**Seaborn:** Seaborn automates the creation of multiple figures. This sometimes leads to OOM (out of memory) issues.

## **3.Visualization:**

**Matplotlib:** Matplotlib is a graphics package for data visualization in Python. It is well integrated with NumPy and Pandas. The pyplot module mirrors the MATLAB plotting commands closely. Hence, MATLAB users can easily transit to plotting with Python.

**Seaborn:** Seaborn is more integrated for working with Pandas data frames. It extends the Matplotlib library for creating beautiful graphics with Python using a more straightforward set of methods.

## **4.Data frames and Arrays**

**Matplotlib:** Matplotlib works with data frames and arrays. It has different stateful APIs for plotting. The figures and axes are represented by the object and therefore plot() like calls without parameters suffices, without having to manage parameters.

**Seaborn:** Seaborn works with the dataset as a whole and is much more intuitive than Matplotlib. For Seaborn, replot() is the entry API with 'kind' parameter to specify the type of plot which could be line, bar, or many of the other types. Seaborn is not stateful. Hence, plot() would require passing the object.

## **5.Flexibility:**

**Matplotlib:** Matplotlib is highly customizable and powerful.

**Seaborn:** Seaborn avoids a ton of boilerplate by providing default themes which are commonly used.

## **6.Use Cases:**

**Matplotlib:** Pandas uses Matplotlib. It is a neat wrapper around Matplotlib.

**Seaborn:** Seaborn is for more specific use cases. Also, it is Matplotlib under the hood. It is specially meant for statistical plotting.