# Seaborn Tutorial

**Prerequisites:**

You should have a basic understanding of computer programming terminologies. A basic understanding of Python and any of the programming languages is a plus. Seaborn library is built on top of Matplotlib. Having basic idea of Matplotlib will help you understand this tutorial in a better way.

In the world of Analytics, the best way to get insights is by visualizing the data. Data can be visualized by representing it as plots which is easy to understand, explore and grasp. Such data helps in drawing the attention of key elements.

To analyse a set of data using Python, we make use of Matplotlib, a widely implemented 2D plotting library. Likewise, Seaborn is a visualization library in Python. It is built on top of Matplotlib.

**Seaborn Vs Matplotlib**

It is summarized that if Matplotlib “tries to make easy things easy and hard things possible”, Seaborn tries to make a well-defined set of hard things easy too.”

Seaborn helps resolve the two major problems faced by Matplotlib; the problems are −

* Default Matplotlib parameters
* Working with data frames

As Seaborn compliments and extends Matplotlib, the learning curve is quite gradual. If you know Matplotlib, you are already half way through Seaborn.

**Important Features of Seaborn**

Seaborn is built on top of Python’s core visualization library Matplotlib. It is meant to serve as a complement, and not a replacement. However, Seaborn comes with some very important features. Let us see a few of them here. The features help in −

* Built in themes for styling matplotlib graphics
* Visualizing univariate and bivariate data
* Fitting in and visualizing linear regression models
* Plotting statistical time series data
* Seaborn works well with NumPy and Pandas data structures
* It comes with built in themes for styling Matplotlib graphics

In most cases, you will still use Matplotlib for simple plotting. The knowledge of Matplotlib is recommended to tweak Seaborn’s default plots.

**Installing Seaborn and getting started**

In this section, we will understand the steps involved in the installation of Seaborn.

### Using Pip Installer

To install the latest release of Seaborn, you can use pip −

pip install seaborn

### For Windows, Linux & Mac using Anaconda

Anaconda (from <https://www.anaconda.com/> is a free Python distribution for SciPy stack. It is also available for Linux and Mac.

It is also possible to install the released version using conda −

conda install seaborn

### To install the development version of Seaborn directly from github

[https://github.com/mwaskom/seaborn"](https://github.com/mwaskom/seaborn)

**Dependencies**

Consider the following dependencies of Seaborn −

* Python 2.7 or 3.4+
* numpy
* scipy
* pandas
* matplotlib

## Importing Libraries

Let us start by importing Pandas, which is a great library for managing relational (table-format) datasets. Seaborn comes handy when dealing with DataFrames, which is most widely used data structure for data analysis.

The following command will help you import Pandas −

# Pandas for managing datasets

import pandas as pd

Now, let us import the Matplotlib library, which helps us customize our plots.

# Matplotlib for additional customization

from matplotlib import pyplot as plt

We will import the Seaborn library with the following command −

# Seaborn for plotting and styling

import seaborn as sb

## Importing Datasets

We have imported the required libraries. In this section, we will understand how to import the required datasets.

Seaborn comes with a few important datasets in the library. When Seaborn is installed, the datasets download automatically.

You can use any of these datasets for your learning. With the help of the following function you can load the required dataset

load\_dataset()

## Importing Data as Pandas DataFrame

In this section, we will import a dataset. This dataset loads as Pandas DataFrame by default. If there is any function in the Pandas DataFrame, it works on this DataFrame.

The following line of code will help you import the dataset −

# Seaborn for plotting and styling

import seaborn as sb

df = sb.load\_dataset('tips')

print df.head()

The above line of code will generate the following output −

total\_bill tip sex smoker day time size

0 16.99 1.01 Female No Sun Dinner 2

1 10.34 1.66 Male No Sun Dinner 3

2 21.01 3.50 Male No Sun Dinner 3

3 23.68 3.31 Male No Sun Dinner 2

4 24.59 3.61 Female No Sun Dinner 4

To view all the available data sets in the Seaborn library, you can use the following command with the **get\_dataset\_names()** function as shown below −

import seaborn as sb

print sb.get\_dataset\_names()

The above line of code will return the list of datasets available as the following output

[u'anscombe', u'attention', u'brain\_networks', u'car\_crashes', u'dots',

u'exercise', u'flights', u'fmri', u'gammas', u'iris', u'planets', u'tips',

u'titanic']

**DataFrames** store data in the form of rectangular grids by which the data can be over viewed easily. Each row of the rectangular grid contains values of an instance, and each column of the grid is a vector which holds data for a specific variable. This means that rows of a DataFrame do not need to contain, values of same data type, they can be numeric, character, logical, etc. DataFrames for Python come with the Pandas library, and they are defined as two-dimensional labeled data structures with potentially different types of columns.