### Data Visualization

**Project Name: statistical machine learning Team ID : PNT2022TMID04163**

**approaches to liver disease prediction**

**Data Visualization:**

* Data visualization is where a given data set is presented in a graphical format. It helps the detection of patterns, trends and correlations that might go undetected in text-based data.
* Understanding your data and the relationship present within it is just as important as any algorithm used to train your machine learning model. In fact, even the most sophisticated machine learning models will perform poorly on data that wasn’t visualized and understood properly.
* To visualize the dataset we need libraries called Matplotlib and Seaborn.
* The Matplotlib library is a Python 2D plotting library which allows you to generate plots, scatter plots, histograms, bar charts etc.

All these libraries come with different features and can support various types of graphs. In this tutorial, we will be discussing four such libraries.

* Matplotlib
* Seaborn
* Bokeh
* Plotly

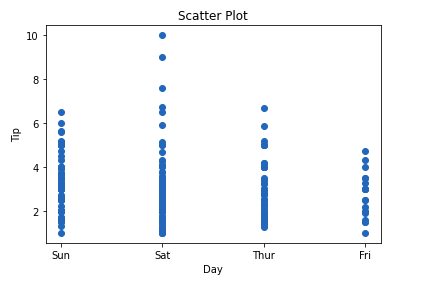
## Matplotlib:

Matplotlib is an easy-to-use, low-level data visualization library that is built on NumPy arrays. It consists of various plots like scatter plot, line plot, histogram, etc. Matplotlib provides a lot of flexibility.

**Program:**

|  |
| --- |
| import pandas as pd  import matplotlib.pyplot as plt      # reading the database  data = pd.read\_csv("tips.csv")    # Scatter plot with day against tip  plt.scatter(data['day'], data['tip'])    # Adding Title to the Plot  plt.title("Scatter Plot")    # Setting the X and Y labels  plt.xlabel('Day')  plt.ylabel('Tip')    plt.show() |

**Output:**



**Seaborn**

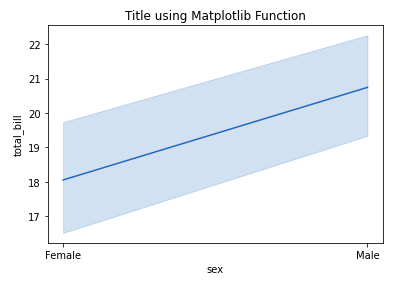
**Seaborn**is a high-level interface built on top of the Matplotlib. It provides beautiful design styles and color palettes to make more attractive graphs.

To install seaborn type the below command in the terminal.

**Example:**

|  |
| --- |
| # importing packages  import seaborn as sns  import matplotlib.pyplot as plt  import pandas as pd      # reading the database  data = pd.read\_csv("tips.csv")    # draw lineplot  sns.lineplot(x="sex", y="total\_bill", data=data)    # setting the title using Matplotlib  plt.title('Title using Matplotlib Function')    plt.show() |

**Output:**



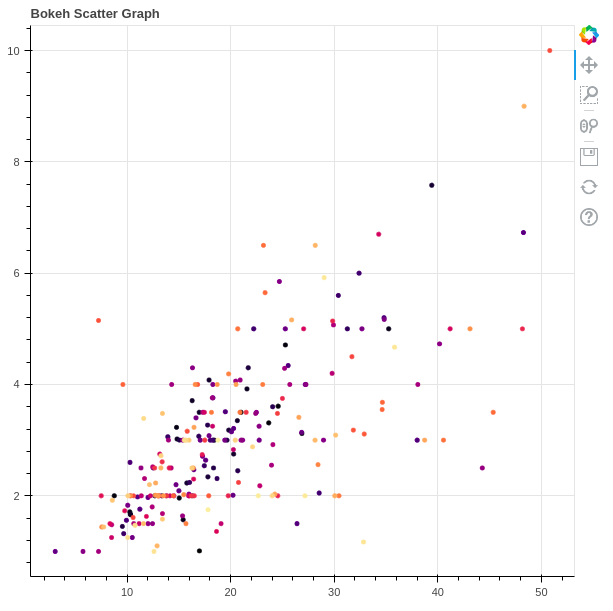
**Bokeh:**

Let’s move on to the third library of our list. Bokeh is mainly famous for its interactive charts visualization. Bokeh renders its plots using HTML and JavaScript that uses modern web browsers for presenting elegant, concise construction of novel graphics with high-level interactivity.

**Example:**

|  |
| --- |
| # importing the modules  from bokeh.plotting import figure, output\_file, show  from bokeh.palettes import magma  import pandas as pd      # instantiating the figure object  graph = figure(title = "Bokeh Scatter Graph")    # reading the database  data = pd.read\_csv("tips.csv")    color = magma(256)    # plotting the graph  graph.scatter(data['total\_bill'], data['tip'], color=color)    # displaying the model  show(graph) |

**Output:**



**Plotly**

* Plotly has hover tool capabilities that allow us to detect any outliers or anomalies in numerous data points.
* It allows more customization.
* It makes the graph visually more attractive.

**Example:**

|  |
| --- |
| import plotly.express as px  import pandas as pd    # reading the database  data = pd.read\_csv("tips.csv")    # plotting the scatter chart  fig = px.scatter(data, x="day", y="tip", color='sex')    # showing the plot  fig.show() |

**Output:**

