**Title: Data Visualization Using Python**

**Creating Bar Graphs, Line Graphs, Scatter Plots, and Histograms from a Dataset**

1. **Introduction**

Data visualization is one of the key factors in data analysis. It helps in picturing patterns, trends, and relationships within data through graphical representations. The project illustrates how to plot different plots using Python in Google Colab to understand data from a given dataset.

In this report, we will use Python libraries like **pandas**, **matplotlib**, and **seaborn** to create:

* Bar Graphs
* Line Graphs
* Scatter Plots
* Histograms

**2. Dataset Description**

We'll be working on a dataset contained in a CSV file called tested.csv of the titanic dataset. There are many columns in the said dataset with various values, a number of different categories, numeric data, and perhaps time series data. From these, appropriate columns will be used to produce the following:

**3. Theory of Plots**

**3.1 Bar Graph**

A bar graph is used to represent categorical data with rectangular bars. The height or length of each bar is proportional to the value it represents.  
**Use Case:** Comparing different categories or groups.

**3.2 Line Graph**

A line graph displays information as a series of data points connected by a straight line. It is used to visualize data points over time.  
**Use Case:** Observing trends and changes over time.

**3.3 Scatter Plot**

A scatter plot shows the relationship between two continuous variables using dots on a graph. It helps identify correlations or patterns.  
**Use Case:** Analyzing relationships between two variables.

**3.4 Histogram**

A histogram is a graphical representation that organizes a group of data points into specified ranges (bins). It shows the frequency distribution of a variable.  
**Use Case:** Understanding data distribution and detecting outliers.

**4. Implementation in Google Colab**

**4.1 Setting Up Google Colab**

1. Upload the dataset (tested.csv) to Google Colab.
2. Use the following code to load and preview the dataset.

**5. Codes**

from google.colab import files

uploaded = files.upload()  
 import pandas as pd

df = pd.read\_csv('tested.csv')

print(df)  
#This helps to read the CSV file from the device.

**For Bar Graph**

import matplotlib.pyplot as plt

import seaborn as sns

plt.figure(figsize=(10, 6))

sns.barplot(x=df.columns[0], y=df.columns[1], data=df)  # Adjust columns as needed

plt.title("Bar Graph")

plt.show()

**For Line Graph**

import matplotlib.pyplot as plt

import seaborn as sns

plt.figure(figsize=(10, 6))

sns.barplot(x=df.columns[0], y=df.columns[1], data=df)  # Adjust columns as needed

plt.title("Bar Graph")

plt.show()

**For Scatter Plot**

plt.figure(figsize=(20, 20))

plt.scatter(df[df.columns[1]], df[df.columns[2]], color='r')

plt.title("Scatter Plot")

plt.xlabel(df.columns[1])

plt.ylabel(df.columns[2])

plt.show()

**For Histogram**

plt.figure(figsize=(10,8))

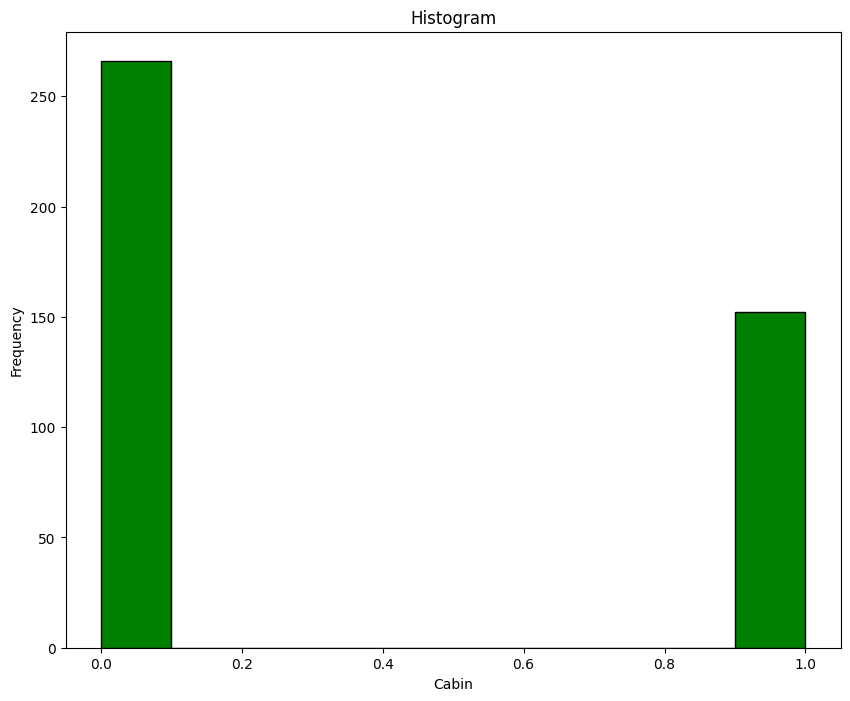
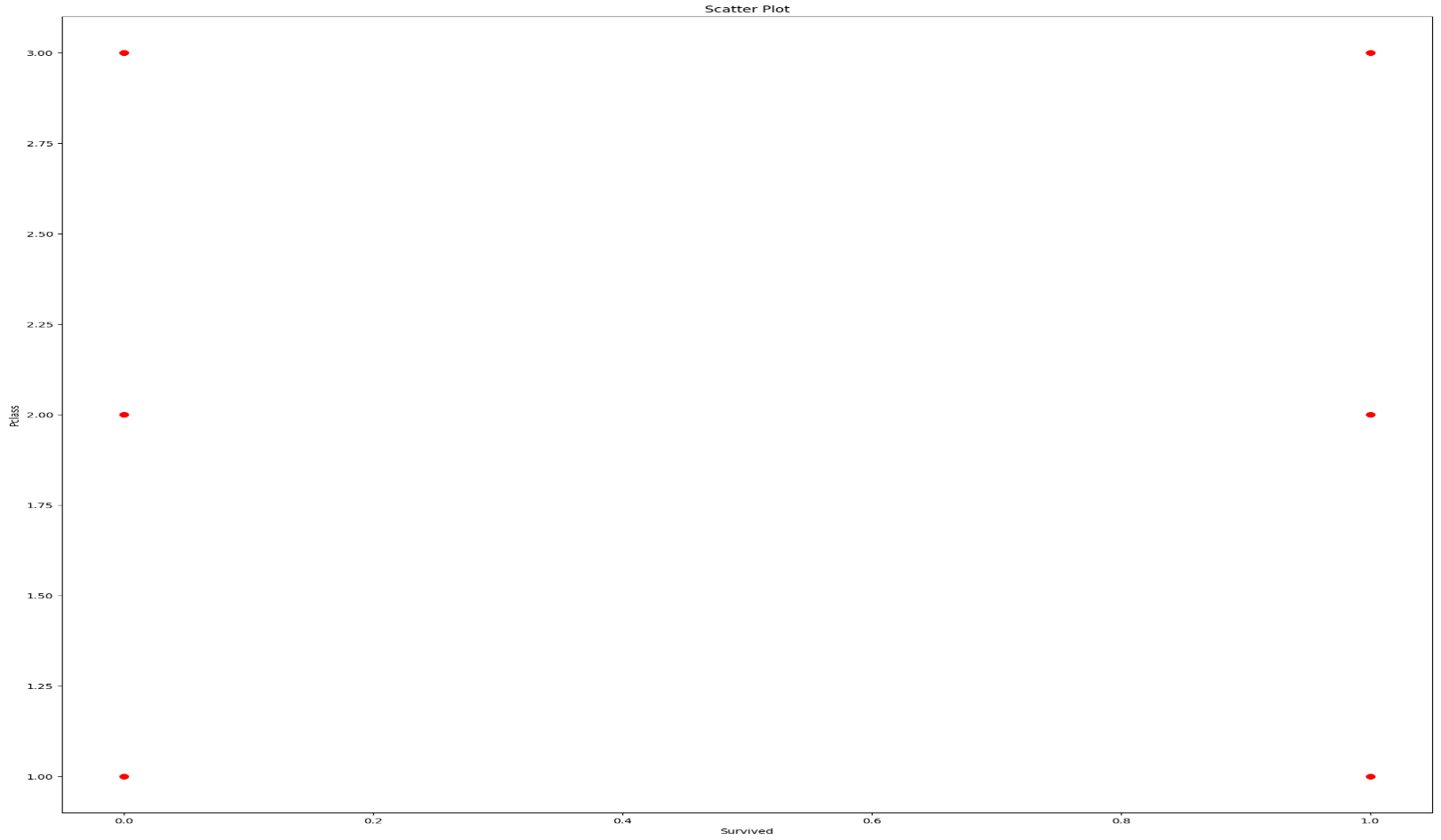
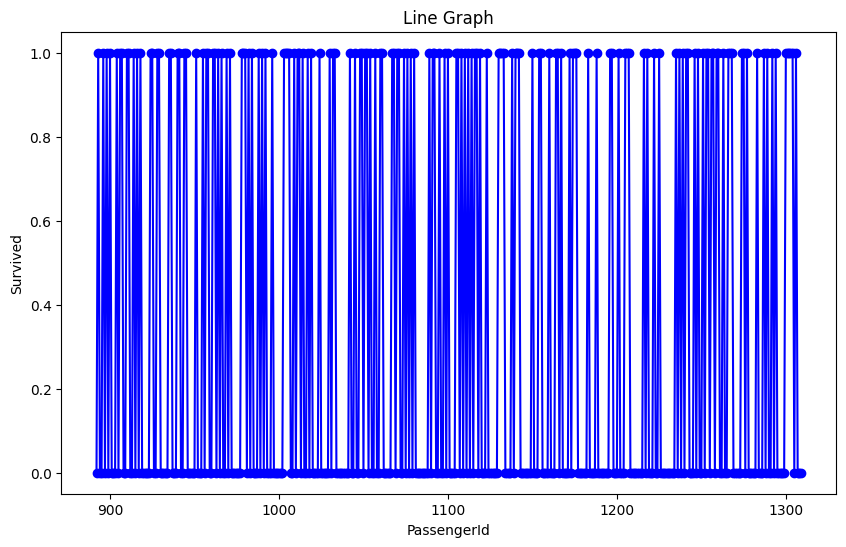
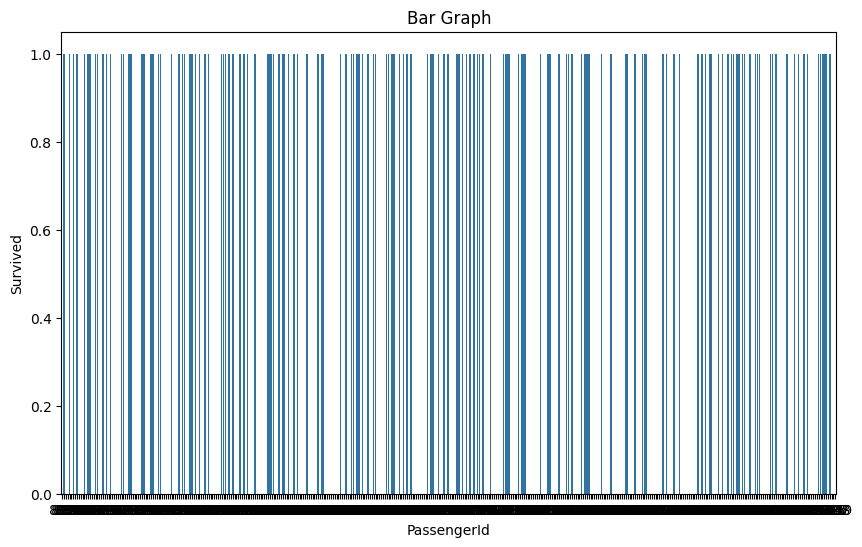
plt.hist(df[df.columns[1]], bins=10, color='g', edgecolor='black')

plt.title("Histogram")

plt.xlabel(df.columns[10])

plt.ylabel("Frequency")

plt.show()

****