CSE317 Software Project V

Project Report

GROUP – 3

Submitted to

Ms. Fariha Jahan

Submitted by

Abruna Afrin (191-15-12303)

SIUM BIN NOOR (191-15-12516)

Tanjil Rahman (191-15-12536)

Abdul Hamid (191-15-12387)

Samiul Islam Abir (191-15-12559)

Title: Passenger Dataset.

Abstract: The dataset contains the number of passengers survived. Our project goal is to show the data visualization on the number of passengers survived. We are using Python to develop our project.

Introduction: Data visualization means the information that we get from a data showing in a graph, chart. We can connect it with image for visualization. In our project, we are using the passenger data. We have taken this dataset from the Kaggle platform. Raw data link: - (https://www.kaggle.com/hesh97/titanicdataset-traincsv). We Pre procced the data by using Python Library (Pandas, Matplotlib, Numpy) and complete the visualization of our project.

Methodology: Our project dataset has 891 rows and 16 columns. After preprocessing the columns are Passengerid, survived, Pclass, name, gender, age, Sibsp, parch, ticket, ticket cost, country.

Bar Graph:
```
sample_data.plot(x="Pclass", y="Ticket Cost", kind="bar",color=["red"])
plt.ylabel('Ticket Cost')
plt.title("Ticket Cost by Pclass(Bar Graph)")
sample_data
```
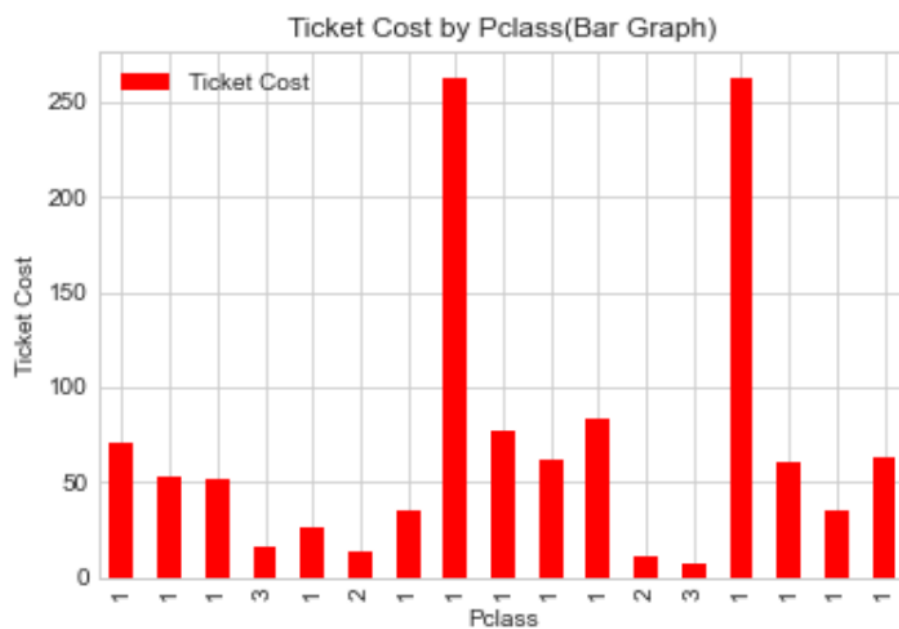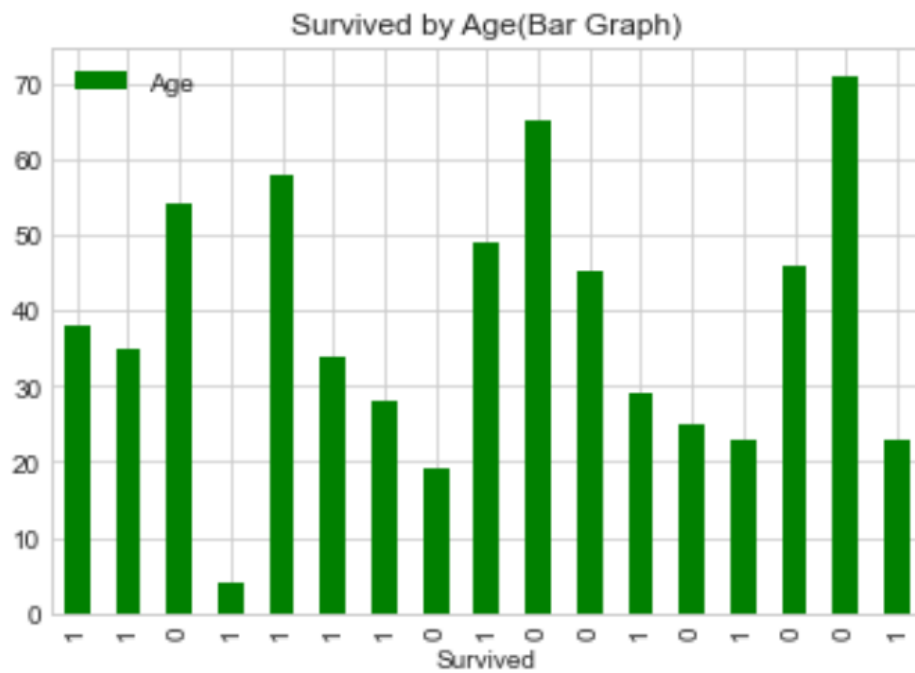

Bar Graph 1

Bar Graph:
```
sample_data.plot(x="Survived", y="Age", kind="bar",color=["green"])
```
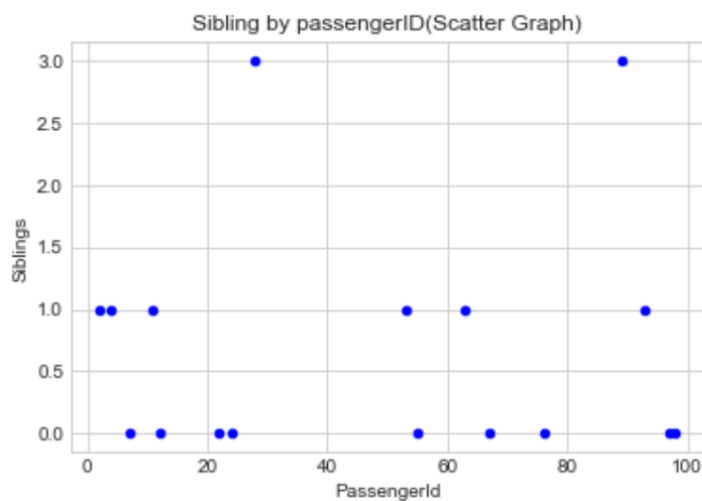
plt.title("Survived by Age(Bar Graph)")
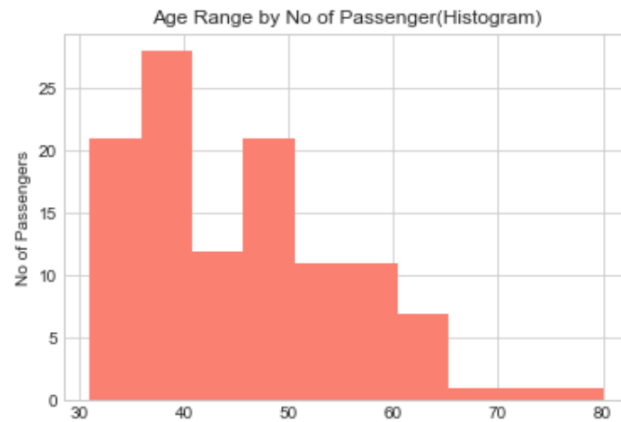sample_data

Bar Graph 1.1

Scatter Plot:
sample_data.plot(x="PassengerId", y="Siblings", kind="scatter",color=["blue"])
plt.title("Sibling by passengerID(Scatter Graph)")
sample_data

Scatter Plot

Histogram:

```
plt.style.use('seaborn-colorblind')
over_30['Age'].plot.hist(color=["salmon"]) #over_30['Age'].hist(color=["salmon"])
plt.ylabel('No of Passengers')
plt.title("Age Range by No of Passenger(Histogram)")
```

Text(0.5, 1.0, 'Age Range by No of Passenger(Histogram)')

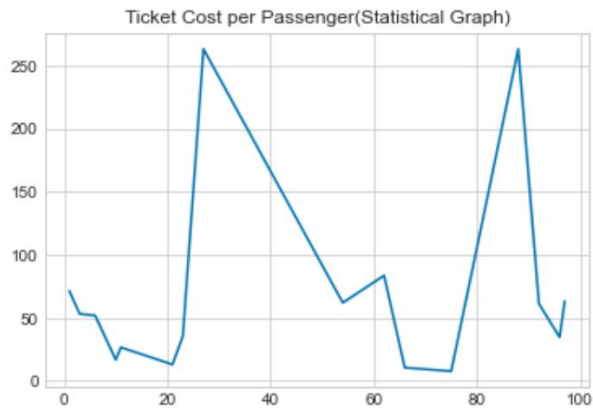**Histogram**

Statistical Graph:

```
plt.style.use('seaborn-whitegrid')
sample_data["Ticket Cost"].plot()
plt.title('Ticket Cost per Passenger(Statistical Graph)')
```

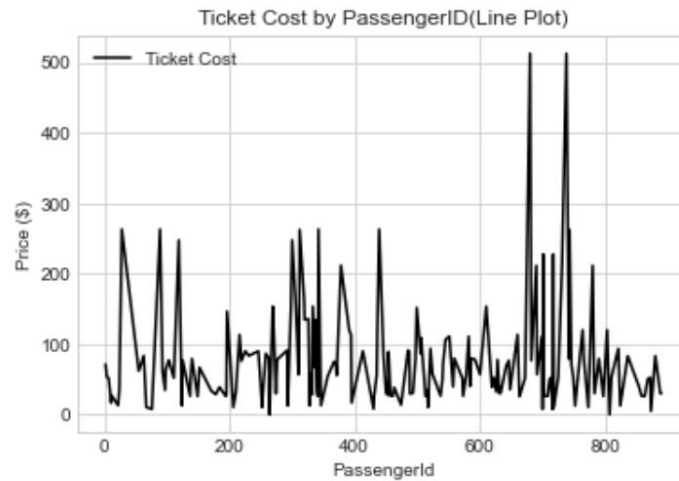Text(0.5, 1.0, 'Ticket Cost per Passenger(Statistical Graph)')

Statistical Graph

Line Plot:

df.plot(x="PassengerId" ,y="Ticket Cost",color=["black"])
plt.title("Ticket Cost by PassengerID(Line Plot)")
plt.ylabel("Price ($)")

Text(0, 0.5, 'Price ($)')

Line Plot

Conclusion: The report contains the visualization of our project on Python Library. We use Python Library(Pandas, Matplotlib, Numpy) and Jupyter Notebook and Google Colab. Finally, it's a well platform for visual dataset. Here is the project linl: tanjil15-12536/SPV_Project_2 (github.com)