Assignment 3 – Data Visualization

Name: Rabia Abdul Sattar
Roll No: 2225165022
Course: Applied Data Science with AI
Week #: 3
Project Title: Customer Churn Prediction

---

1. Reading Summary

Reading Material:

- Storytelling with Data resources
- Matplotlib and Seaborn tutorials

Key Learnings:

- Different plots (bar, histogram, boxplot, heatmap) explain customer behavior.

Reflection:

I gained an understanding of how visualizations can highlight churn patterns, particularly the impact of contract types and monthly charges.

2. Classroom Task Documentation

Task Performed:

- Created bar charts, histograms, boxplots, and heatmaps using sample datasets in class.

3. Weekly Assignment Submission

Assignment Title: Data Visualization and Exploratory Data Analysis

Steps Taken:

Steps Taken:

Step 1: Import Dataset

1. Loaded the Titanic dataset (train.csv) into a Pandas DataFrame using pd.read_csv().

2. Checked the dataset shape, missing values, and column info.

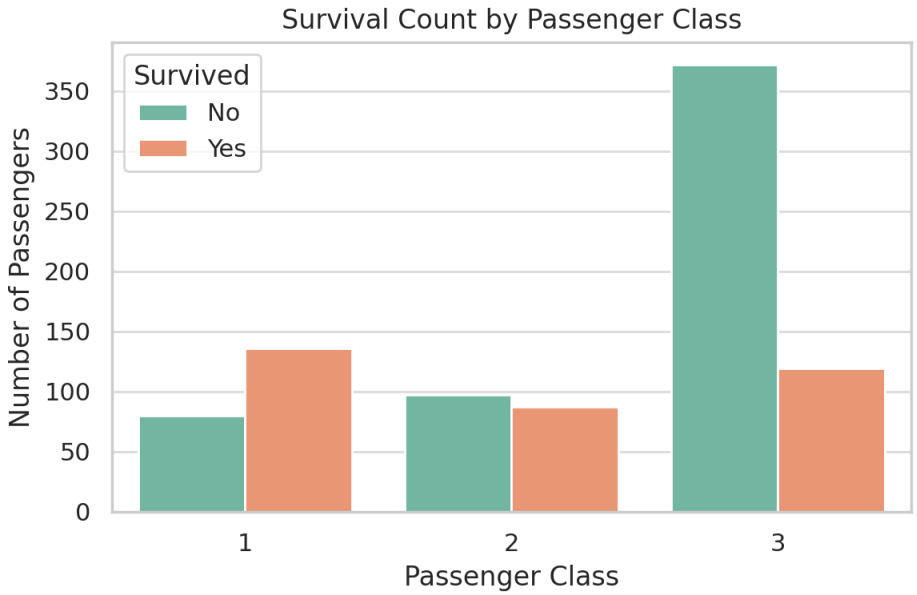
Step 2: Data Cleaning

1. Identified missing values (especially in the Age column).

2. Dropped NaN values for plotting age-related graphs (like histograms & scatter plots).

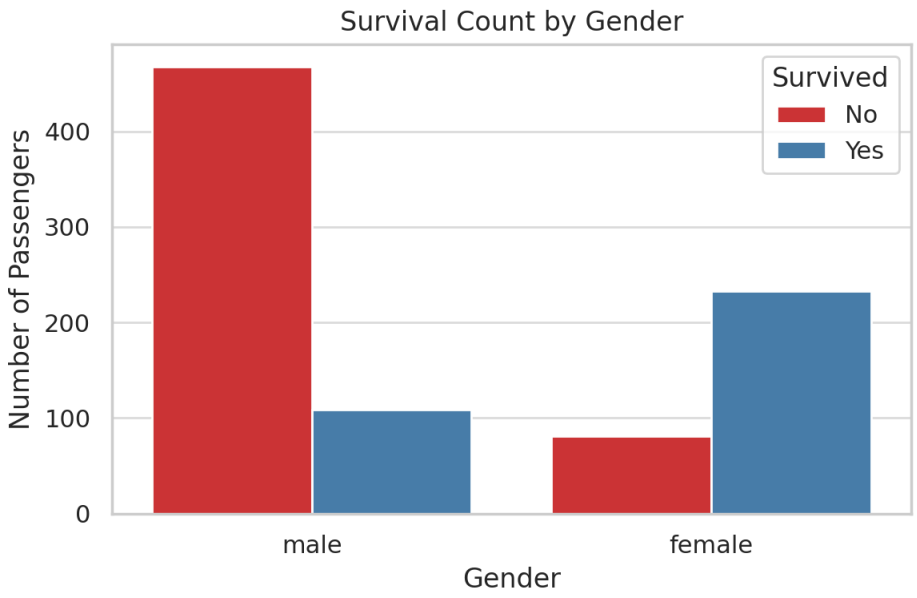
Step 3: Plot 1 – Bar Chart (Survival by Passenger Class)

Insight: 1st-class passengers had higher survival chances than 3rd-class.



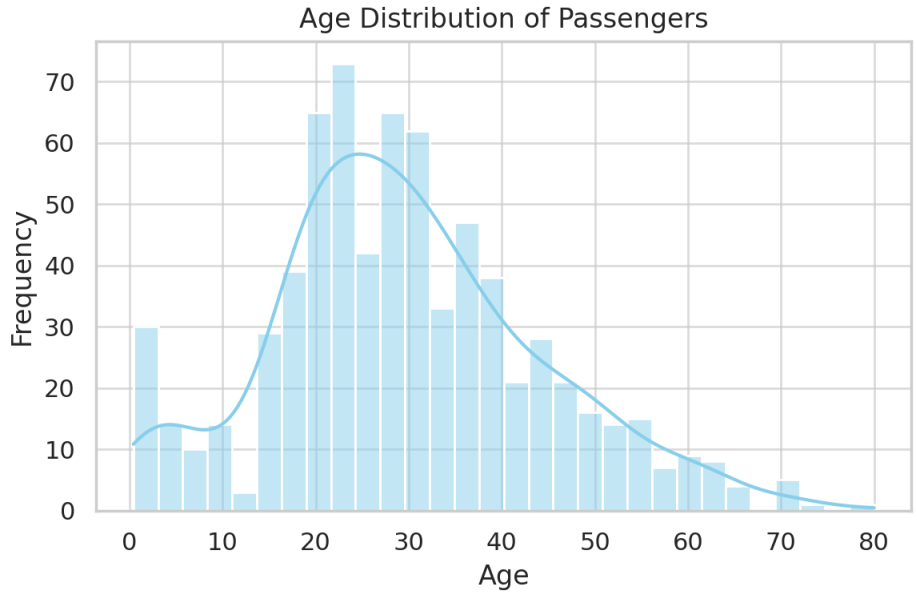
Step 4: Plot 2 – Bar Chart (Survival by Gender)

Insight: Females survived far more than males.



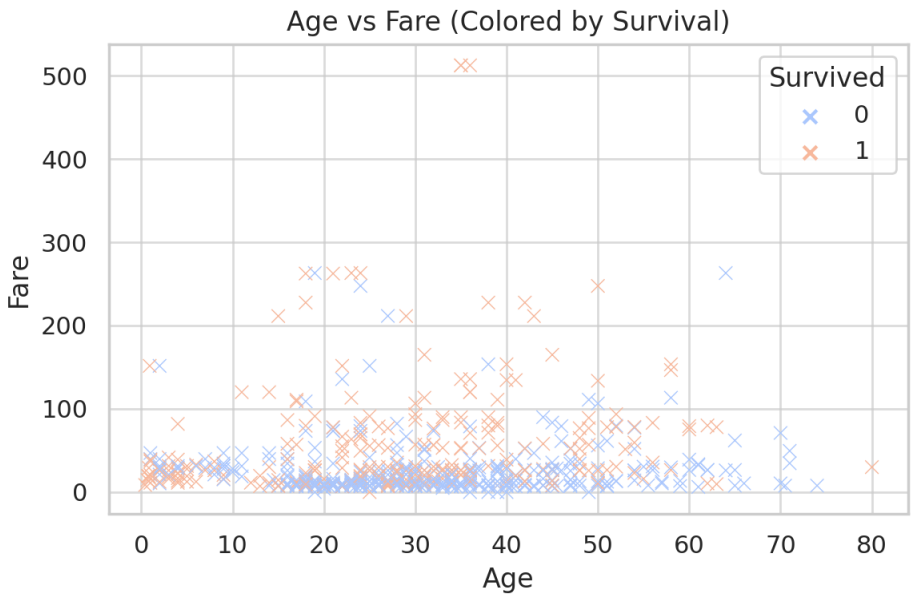
Step 5: Plot 3 – Histogram (Age Distribution)

Insight: Most passengers were between 20–40 years old.

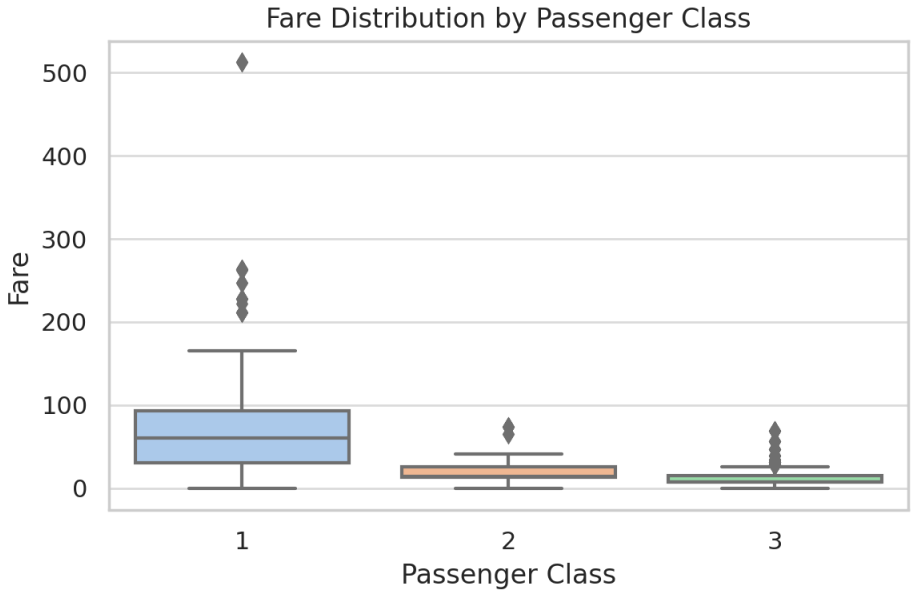


Step 6: Plot 4 – Scatter Plot (Age vs Fare, colored by Survival)

Insight: Higher fare passengers had better survival chances.



Step 7: Plot 5 – Box Plot (Fare by Class)
Insight: 1st-class fares were highest; 3rd-class lowest.



Output:

- Visual patterns showed churn is higher in short-tenure, month to-month, and high-charge customers.

Challenges Faced:

• Handling missing data was a minor issue: the many missing Age values meant I had to decide whether to drop or estimate ages for the histogram and scatter plots.

GitHub Link:

https://github.com/Rabia-Abdul-Sattar/Customer-Churn-Prediction

4. Project Progress Milestone

• Completed first Exploratory Data Analysis (EDA).

• Next week's goal: Perform correlation analysis to identify key features related to churn.

5. Self-Evaluation

☑ I completed all tasks on time.