Exploratory Data Analysis (EDA) – Titanic Dataset

Objective
The objective of this analysis is to explore the Titanic dataset using statistical and visual techniques to identify patterns, trends, and anomalies that influenced passenger survival.

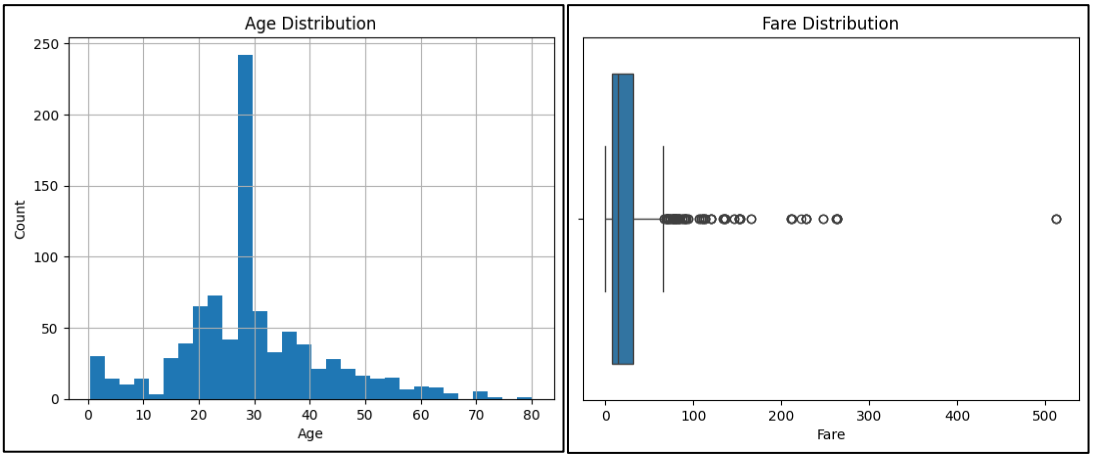
Dataset Overview
The Titanic dataset contains passenger details such as age, gender, passenger class, fare, and survival status. It includes both numerical and categorical variables.

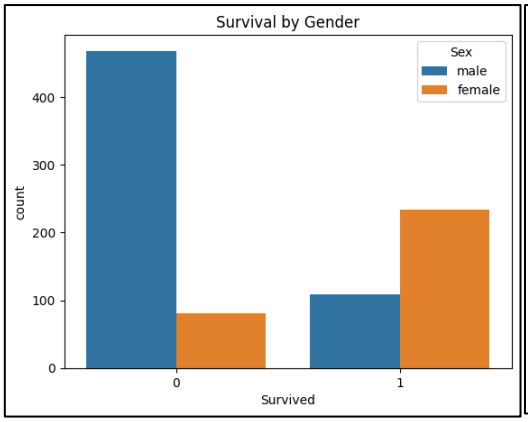
Statistical Analysis
Descriptive statistics show that the average passenger age is around 30 years. The survival rate is below 50 percent, indicating that most passengers did not survive. Fare values vary widely, indicating the presence of outliers.

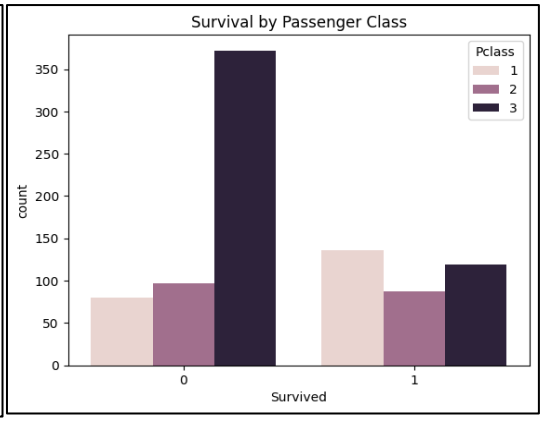
Visual Analysis
Histograms show that most passengers were between 20 and 40 years old. Boxplots reveal that fare distribution is highly skewed. Count plots indicate higher survival among females and first-class passengers. Correlation analysis shows a positive relationship between fare and survival and a negative relationship between passenger class and survival.

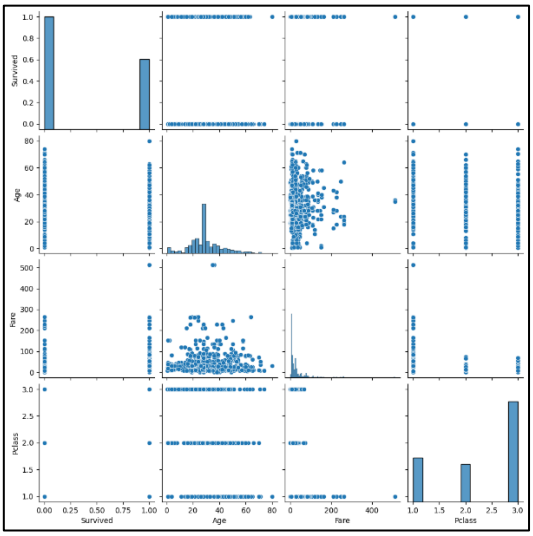
Key Findings
• Female passengers had a higher survival rate than males.
• First-class passengers were more likely to survive than third-class passengers.
• Higher fares were associated with better survival chances.
• Age alone was not a strong factor in survival.
• Fare contains significant outliers.

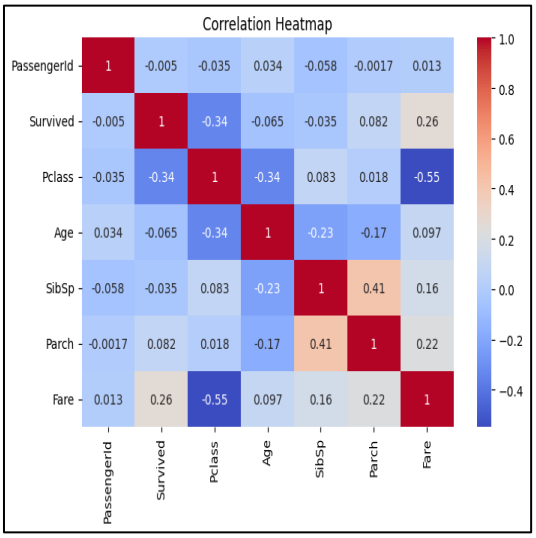
Output:

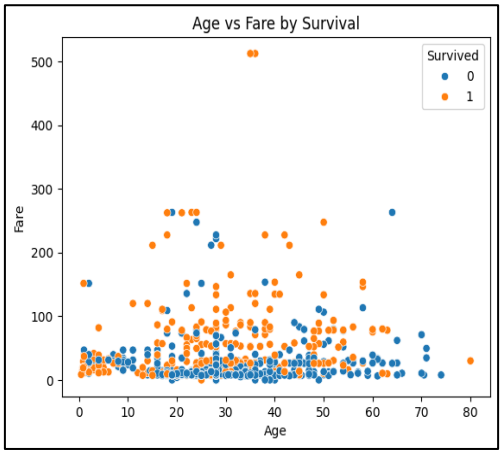












Conclusion

The EDA process helped identify important factors affecting survival outcomes and demonstrated how visual and statistical exploration can provide meaningful insights from data

By: C T Thrisha