

TASK -05 : Exploratory Data Analysis (EDA)

Report of Findings: Exploratory Data Analysis on Titanic Dataset

Introduction: This report presents the findings from an exploratory data analysis (EDA) performed on the Titanic passenger dataset. The objective was to investigate patterns, relationships, and trends that might explain survival outcomes during the Titanic disaster.

Methodology

The analysis was conducted using Python with the following libraries:

Pandas : for data manipulation

Seaborn and Matplotlib : for data visualization

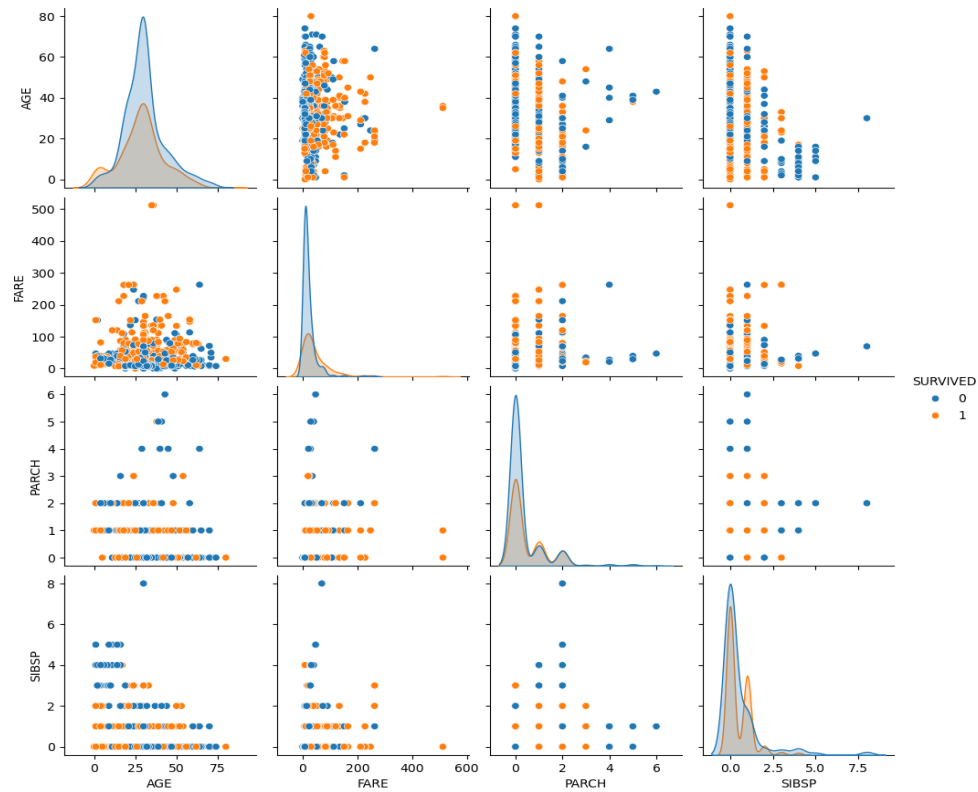
Key techniques applied:

- Descriptive statistics (`.describe()`, `.info()`, `.value_counts()`)
- Pairplots and heatmaps to identify correlations and relationships
- Histograms, boxplots, scatterplots, and countplots for data visualization
- Summary of observations for each visual output

Visualizations and Observations

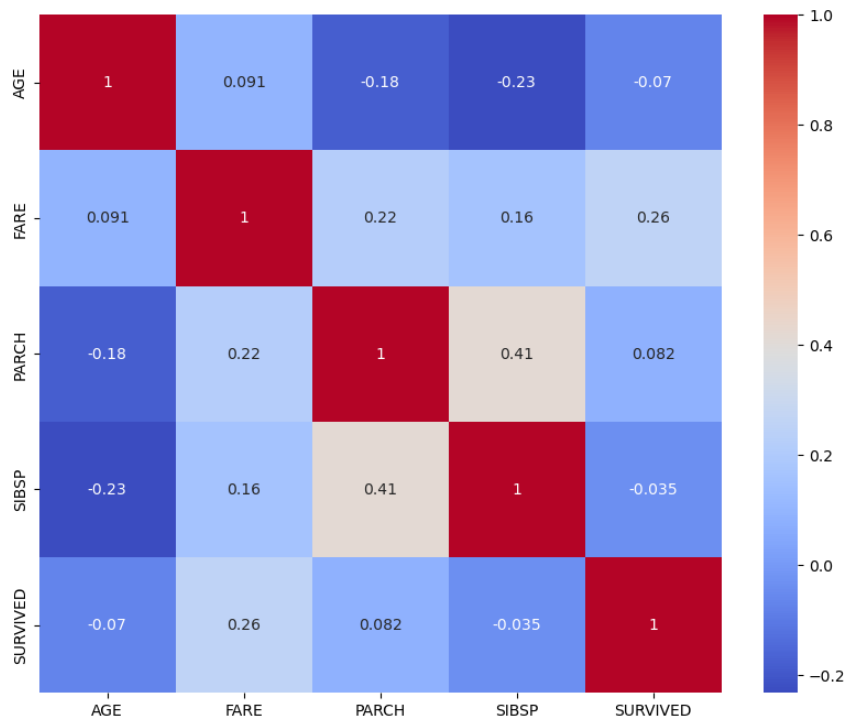
Pairplot:

- Higher fare-paying passengers had higher survival rates.
- Majority of passengers were between 20 and 40 years old.
- Traveling with family (SibSp and Parch) slightly increased survival probability.



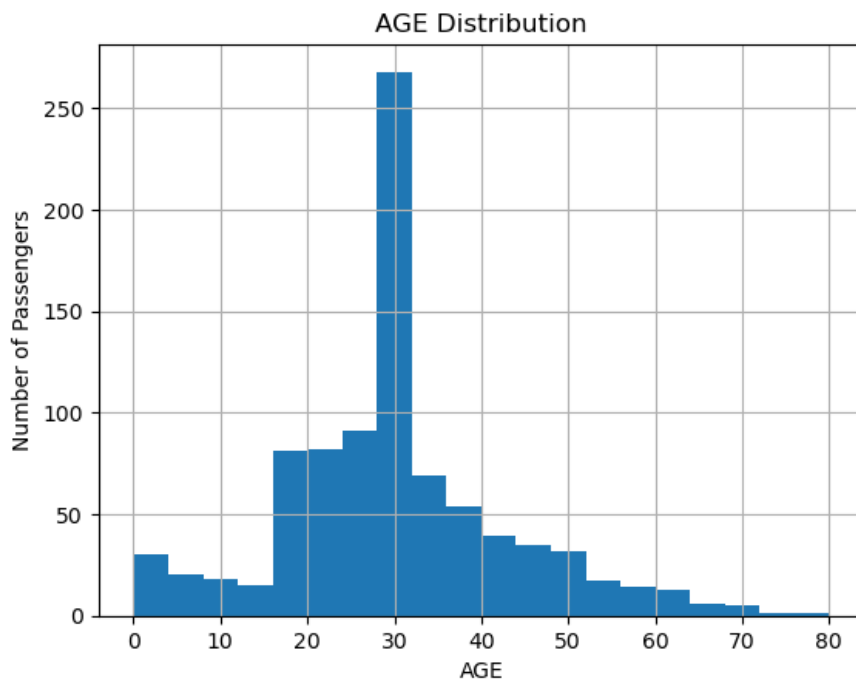
Correlation Heatmap:

- Fare and Survival : Moderately positive correlation (~ 0.26)
- Pclass and Survival : Negative correlation (~ -0.34)
- Age showed a weak relationship with survival.



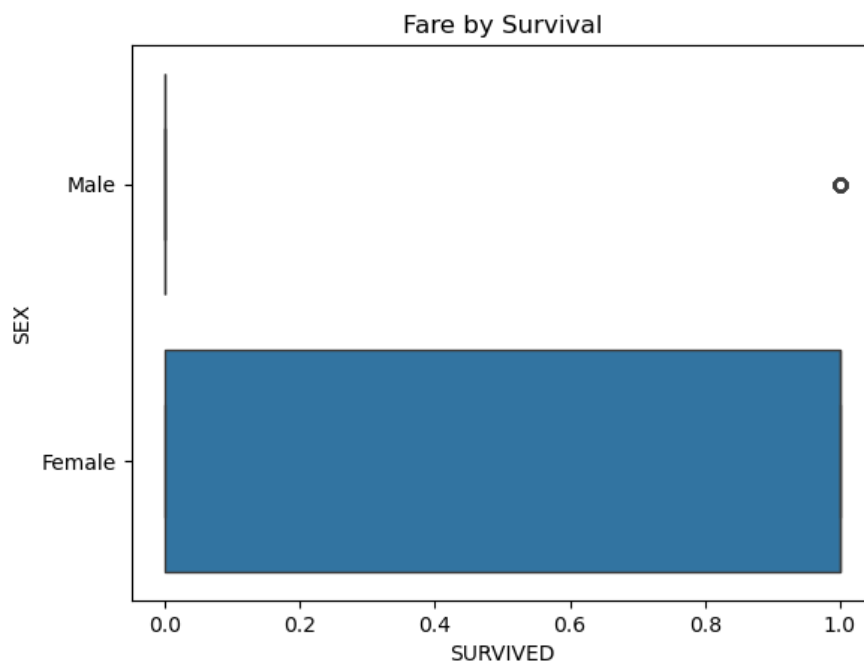
Histograms:

- Most passengers were young adults (20–40 years old).
- Fare distribution was highly right-skewed with a few extreme outliers.



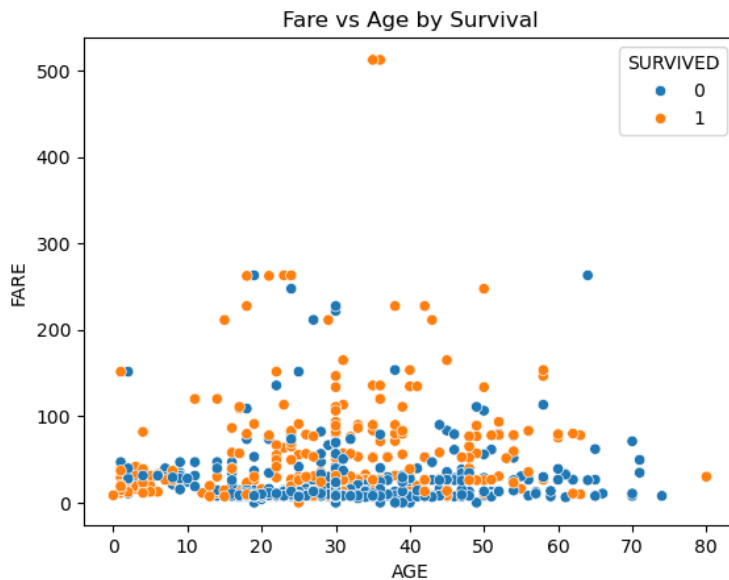
Boxplots:

- Survivors had higher median fares than non-survivors.
- Outliers with extremely high fares mostly survived.



Scatterplots:

- Higher-fare and younger passengers clustered in survival group.
- Some high-paying outliers had a strong association with survival.



Key Findings

- Ticket class and fare amount strongly influenced survival odds.
- Females were prioritized for rescue, resulting in higher survival rates.
- Most passengers were young to middle-aged adults.
- Family presence (SibSp, Parch) had a minor positive influence on survival.
- Wealth and social status (reflected through fare and class) significantly impacted survival chances.

Conclusion

This exploratory analysis of the Titanic dataset reveals clear survival patterns influenced **by gender, ticket class, and fare price**. The findings align with historical records that suggest women, children, and wealthy passengers had higher survival odds.