

1. Data Summary

- **Numerical Features:** Age, Fare, SibSp, Parch
- **Categorical Features:** Sex, Pclass, Embarked, Survived

`.info()` and `.describe()`:

- **Missing Values:**
 - Age: 177 missing
 - Cabin: 687 missing
 - Embarked: 2 missing
- **Data Types:** Mixed types with object and numeric.

2. Univariate Analysis

- **Survival Rate:**
 - 38% survived, 62% did not.
- **Gender:**
 - ~65% male, ~35% female.
- **Class Distribution:**
 - Most passengers were in 3rd class.

Plots:

- **Histogram:** Age is right-skewed with many younger passengers.
- **Bar Plots:** Show clear class and gender imbalance.

3. Bivariate/Multivariate Analysis

- **Survival vs Gender:**
 - Women had a much higher survival rate.
- **Survival vs Class:**
 - 1st class passengers had the highest survival rate.
- **Pairplot:** Correlation of survival with Pclass, Age, and Fare.

Heatmap:

- SibSp and Parch show some correlation with each other.
- Fare shows mild correlation with Pclass.

4. Handling Missing Values

- Dropped irrelevant columns: Cabin, Ticket
- Imputed Age using median.
- Filled Embarked with mode.

5. Outlier Detection

- **Boxplot** of Fare shows significant outliers.

- Age also has mild outliers.

Key Insights

- Gender and class had strong impact on survival.
- Higher fare often related to higher survival.
- Family size affected survival slightly.