The goal of this analysis is to provide a comprehensive understanding of the Titanic dataset. The analysis includes data cleaning, exploration, and visualization.

Data Loading and Cleaning The dataset was loaded from a CSV file. Initial inspection showed missing values in the Age, Cabin, and Embarked columns. To simplify the analysis, rows with any missing values were removed.

Summary of Findings

- The dataset contains 891 entries with 12 columns.
- After cleaning, 204 entries remained.
- The columns are of the following data types:
    - Numerical: PassengerId, Survived, Pclass, Age, SibSp, Parch, Fare
    - Categorical: Name, Sex, Ticket, Cabin, Embarked

Visualizations

1. **Gender Survival Disparity** (Images 2 & 3):
    - Females had significantly higher survival rates than males (≈75-80% vs ≈20%)
    - This confirms the "women and children first" protocol was followed
2. **Age Distribution** (Image 1):
    - Most passengers were young adults (20-40 years old)
    - Few children (0-10) and elderly (60+)

1. Class Impact (Image 5):
    - Higher classes (1st class) had better survival rates than lower classes
    - Pclass shows moderate negative correlation with survival (-0.25)
2. Fare Correlation (Image 4 & Correlation Heatmap):
    - Higher fares (correlated with higher class) had better survival (0.13 correlation)
    - Wealthier passengers had better access to lifeboats

- Age: Negative correlation (-0.25) - younger passengers had better survival chances
- The strongest predictors of survival were:
    - Being female (highest impact)
    - Traveling in higher class (1st class)
    - Being younger
    - Paying higher fare

Conclusion The analysis provides a basic overview of the Titanic dataset.