

Titanic Dataset - EDA Summary Report

Key Statistical Insights:

- Overall survival rate: 38%
- Female survival rate: 74%
- Male survival rate: 19%
- First class survival rate: 63%
- Child survival rate: 59%

Key Observations:

1. **Gender** had a significant impact on survival. Females had a much higher survival rate.
2. **Passenger Class** mattered. Higher class passengers had a better chance of survival.
3. **Age** was also a determining factor. Children were prioritized during evacuation.
4. **Fare** paid and **family size** impacted survival. Passengers with medium-sized families and higher fare tickets survived more.
5. **Missing data** in 'age' and 'embarked' was imputed with median and mode values respectively.
6. **Feature engineering** added new insights: family size, age group, and fare group.
7. **Correlation heatmap** showed strongest correlation with survival for: sex (0.54) and pclass (-0.34).

Conclusion:

The Titanic dataset shows clear evidence of demographic and socioeconomic factors influencing survival during the tragedy. Women, children, and higher-class passengers were given priority during rescue efforts.