

Titanic Dataset EDA Report

1. Overview

This report presents an Exploratory Data Analysis (EDA) of the Titanic dataset based on the training data.

2. Survival Count

About 38% of the passengers survived the Titanic disaster. Survival is indicated by 1, and non-survival by 0.

3. Survival by Gender

Female passengers had a significantly higher survival rate (~74%) compared to male passengers (~18%).

4. Survival by Passenger Class

1st class passengers had the highest survival rate, followed by 2nd, then 3rd class.

5. Age Distribution

Most passengers were between 20 and 40 years old. Some passengers were children or elderly.

6. Embarked Port Distribution

Majority of passengers boarded from Southampton (S), followed by Cherbourg (C) and Queenstown (Q).

7. Correlation Summary

Positive correlation observed between Fare and Survival. Pclass and Survival are negatively correlated.

8. Conclusion

Key factors influencing survival: Gender (female), Passenger class (1st), and Fare amount. This aligns with the historical accounts that women and children were prioritized during rescue.