

Titanic Dataset - Summary of Findings

1. Survival by Gender

- Females had a much higher survival rate than males.
- Gender was one of the strongest predictors of survival.

2. Survival by Class & Fare

- First-class passengers (higher fares) had better chances of survival.
- Third-class passengers (lower fares) had the lowest survival rate.
- Fare distribution was highly skewed, with some very expensive tickets.

3. Age and Survival

- Children had higher survival rates compared to adults.
- Most passengers were between 20–40 years old.
- Survival generally decreased with age.

4. Family Size (SibSp & Parch)

- Small families (1–2 members) had higher survival chances.
- Very large families had lower survival rates.
- Passengers traveling alone also had poor survival chances.

5. Correlation Insights

- Fare and Pclass were strongly correlated (wealthier passengers traveled in higher classes).
- SibSp and Parch were moderately correlated (families traveled together).
- Survival correlated positively with being female, higher fare, and lower passenger class.

6. Missing Data

- Age (177 missing), Cabin (687 missing), and Embarked (2 missing).
- Cabin is too incomplete to be useful without heavy feature engineering.

✓ Overall Insights

- Gender, passenger class, and fare were the most influential factors in survival.

- Women, children, and wealthy passengers had the best survival outcomes.
- Men, older passengers, and those in third class faced the greatest risk.