

Dataset Overview

This dataset contains details of passengers on the Titanic, such as:

- PassengerId
- Survival status (0 = No, 1 = Yes)
- Passenger class (1st, 2nd, 3rd)
- Name, Sex, Age
- Number of siblings/spouses aboard (SibSp)
- Number of parents/children aboard (Parch)
- Ticket number, Fare
- Cabin, Embarkation point

Total entries: 891 rows, 12 columns

🙀 Data Summary (.describe())

Feature	Count	Mean	Std	Min	25%	50%	75%	Max
Age	714	29.70	14.52	0.42	20.12	28.00	38.00	80.00
SibSp	891	0.52	1.10	0	0	0	1	8
Parch	891	0.38	0.81	0	0	0	0	6
Fare	891	32.20	49.69	0.00	7.91	14.45	31.00	512.33

★ Value Counts (Key Columns)

Survived:

- 0 (Did not survive): 549
- 1 (Survived): 342

Sex:

Male: 577Female: 314

Pclass:

3rd class: 491 1st class: 216 2nd class: 184

Visual Insights

- Age distribution is highest between **20–40 years**.
- Most fares lie below **100** but there are few extreme outliers (up to 512).

⊘ Boxplot Observations:

- Survival is higher for females.
- **1st class** passengers have the highest survival rate.
- Fare and Age show outliers in both survived and non-survived.

☑ Heatmap (Correlation):

- Strongest positive correlation: **Fare and Pclass** (more fare = higher class).
- Survival is positively correlated with Fare and Pclass (1st class had better chances).
- Slight negative correlation between **Age** and **Survival**.

Q Key Observations

- Most passengers were young adults (20–40 years).
- **Females** had a significantly higher survival rate.
- 1st class passengers were more likely to survive.
- **Higher fare** generally indicated better survival odds.
- Some **older passengers with high fare** also survived likely wealthy travelers.

Final Summary

This analysis clearly shows that **social status, gender, and age** played a major role in determining survival. The data reflects real historical bias — women and children were given preference, and those who could afford first class had better safety measures.