# **Summary of Findings**

After conducting an exploratory data analysis (EDA) on the Titanic dataset, the following key insights were observed:

#### 1. Survival Rates

- Overall, a higher proportion of passengers did not survive the Titanic disaster (~62% deaths vs ~38% survivors).
- Females had a significantly higher survival rate (~74%) compared to males (~19%).

### 2. Impact of Passenger Class

- First-class passengers had the highest survival rate (~63%).
- Third-class passengers were the majority but had the lowest survival rates (~24%).
- Socio-economic status (ticket class) strongly influenced chances of survival.

## 3. Age Distribution

- Most passengers were between 20 and 40 years old.
- Children (younger passengers) had relatively higher survival rates compared to adults.
- Median age of survivors was slightly lower than that of non-survivors.

#### 4. Fare Influence

- Passengers who paid higher fares had a higher probability of survival.
- The **fare distribution** was **highly skewed**, with most passengers paying low fares and a few paying very high fares.

## 5. Embarkation Port

- Most passengers boarded from port 'S' (Southampton).
- Survival rates varied slightly based on embarkation point, but class and gender had stronger effects.

# 6. Correlation Insights

- Fare had a positive correlation with Survival.
- Pclass had a negative correlation with Survival (higher class = lower numeric value = better survival chances).
- Other features like Age had very weak correlation with survival.

## **Final Conclusion:**

- Survival during the Titanic disaster was strongly influenced by **gender**, **socio-economic class**, **and fare paid**.
- Women, children, and first-class passengers were prioritized during the evacuation.

The EDA of the Titanic dataset reveals important human and socio-economic factors that played a critical role in survival — highlighting class privilege and gender-based rescue priorities during one of history's most infamous maritime disasters.