Titanic EDA Feature-level Inferences

Objective

Feature-level inferences derived from the Titanic dataset.

Age

- Most passengers were young adults (~20–30 years old).
- Inference: The majority of travellers were younger, suggesting Titanic was popular among working-age individuals and families.

Fare

- Fare distribution was highly skewed with long tail values.
- Inference: Most paid low fares, indicating economic diversity, while a small group of wealthy passengers paid very high fares for first-class tickets.

Pclass (Passenger Class)

- Higher class (Pclass = 1) had significantly higher survival rates.
- Inference: Wealth and class status impacted access to lifeboats and survival odds.

Sex

- Females had a much higher survival rate than males.
- Inference: The "women and children first" protocol was largely adhered to.

SibSp / Parch (Family Relations)

- Small families (1–2 relatives) had slightly better survival odds.
- Larger families faced lower survival rates.
- Inference: People traveling in smaller family units were more mobile and likely to access lifeboats.

Correlation Between Features

- Fare and Pclass had a strong negative correlation.
- Inference: Higher fares correlated with higher class status, impacting survival odds.

Final Takeaways

- Survival was influenced by age, sex, class, fare, and family connections.
- The Titanic disaster outcomes were shaped by economic status and demographics.