

Visual Analysis and Observations

1. Age Distribution and Boxplot by Passenger Class

The histogram shows that the age distribution is right-skewed, with a majority of passengers being in the 20–40 age range.

The boxplot indicates that:

- Passengers in 1st class tend to be older on average.
- There is a wider spread of ages in 2nd and 3rd classes, with more young passengers in 3rd class.
- Outliers are present across all classes.

2. Correlation Heatmap of Numeric Features

- Fare has a moderate positive correlation with Pclass (likely due to higher fare for higher class).
- Survived is moderately negatively correlated with Pclass, implying that passengers in lower classes had a lower survival rate.
- Correlation values are generally low, suggesting that survival isn't strongly dependent on any single numeric feature.

3. Scatter Plot: Age vs Fare Colored by Survival

- Survivors tend to cluster in specific areas:
 - Higher fares and middle-aged passengers seem to have higher survival.
 - Passengers paying lower fares, particularly in younger or very old age ranges, appear less likely to survive.
- There's a wide range of fares even among passengers of similar ages, indicating fare wasn't age-dependent.

Summary of Findings

- Age and Pclass play important roles in survival, with older and higher-class passengers more likely to survive.
- Fare correlates with both survival and class, suggesting it could serve as a proxy indicator for socioeconomic status.
- There is no single dominant predictor of survival among the numeric features, but combined variables (like Age + Pclass + Fare) may help in predictive modeling.
- Visual analysis supports the well-known trend in Titanic data: passenger class and fare paid influenced survival chances.