**Report of findings**

**Summary of Exploratory Data Analysis (EDA)**

Based on the statistical and visual exploration of the Titanic dataset:

**Key Predictors of Survival (Relationships & Trends):**

1. **Gender (Sex):** This is the strongest predictor. Females had a survival rate of ~74%, compared to ~19% for males.

2. **Passenger Class (Pclass):** Survival rate decreased significantly with lower class (1st > 2nd > 3rd), confirming that socio-economic status played a major role.

3. **Fare:** Directly correlated with Pclass, passengers who paid a higher fare were generally more likely to survive.

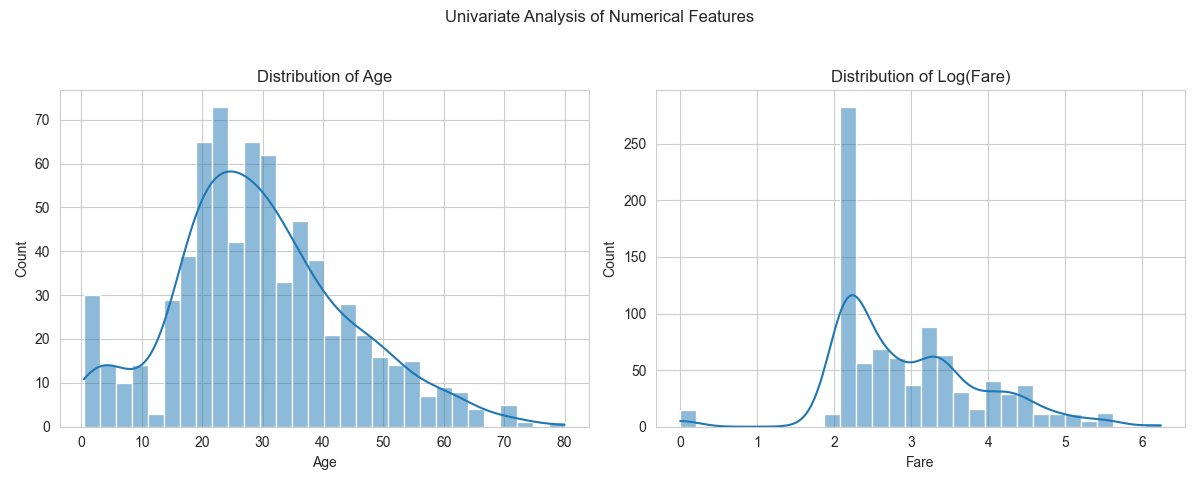
**Data Quality and Anomalies**:

1. **Missing Data**:

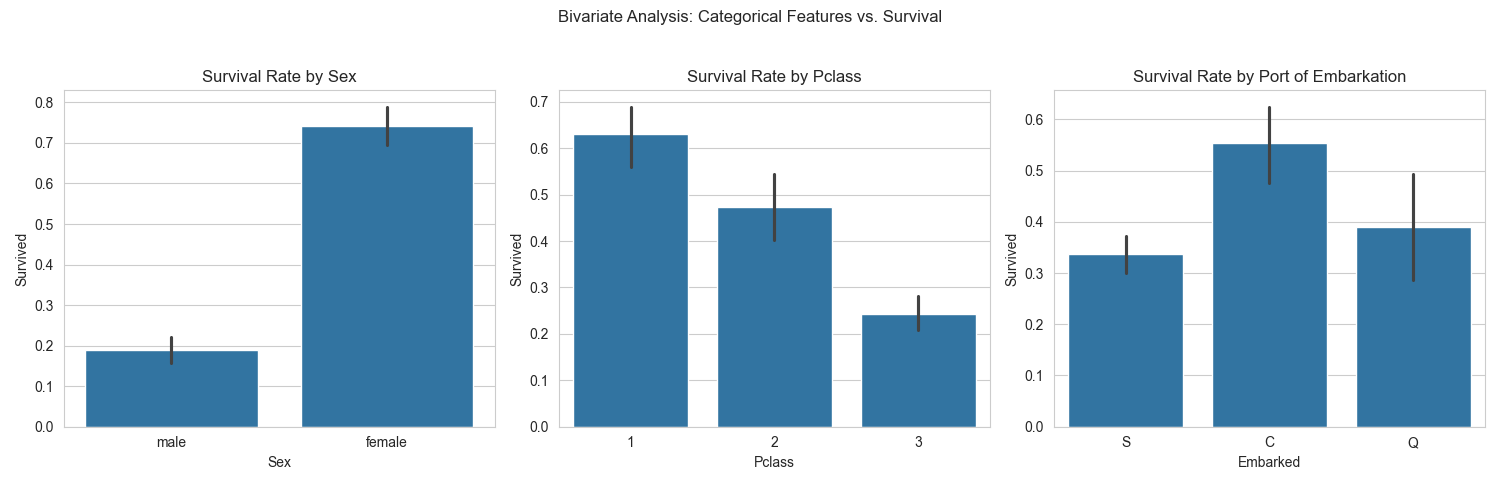
**Cabin** is severely missing data (over 77%) and is likely not useful without extensive feature engineering.

**Age** has a notable number of missing values (177) and will require imputation (filling) for most modeling efforts.

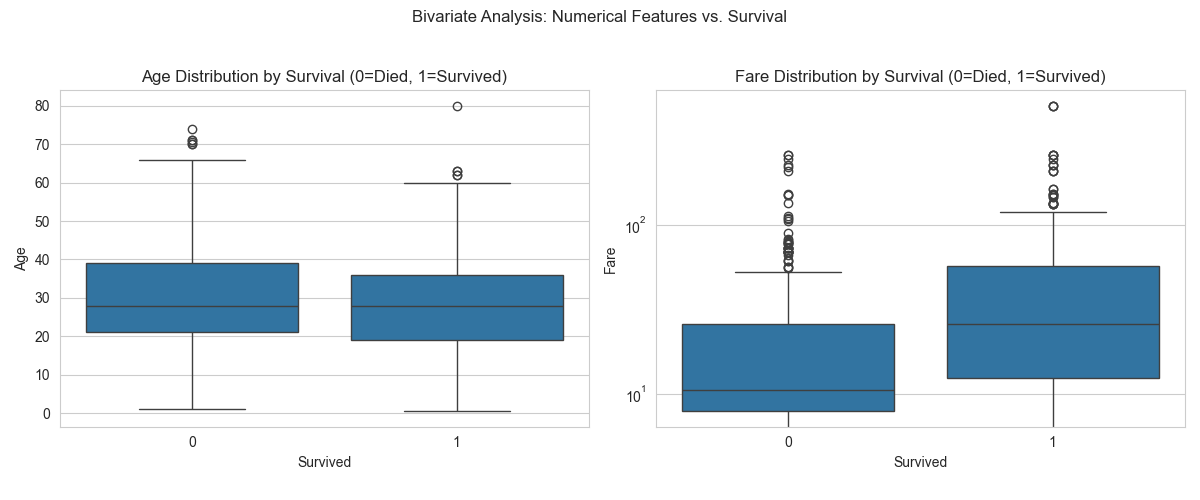
2. **Outliers**: The **Fare** column is highly right-skewed and contains extreme outliers, suggesting a log transformation may be beneficial for predictive modeling.

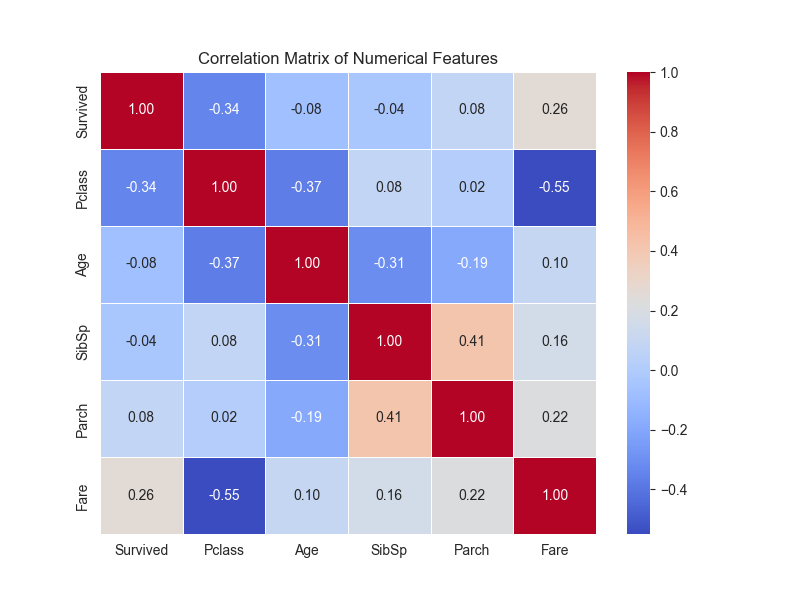


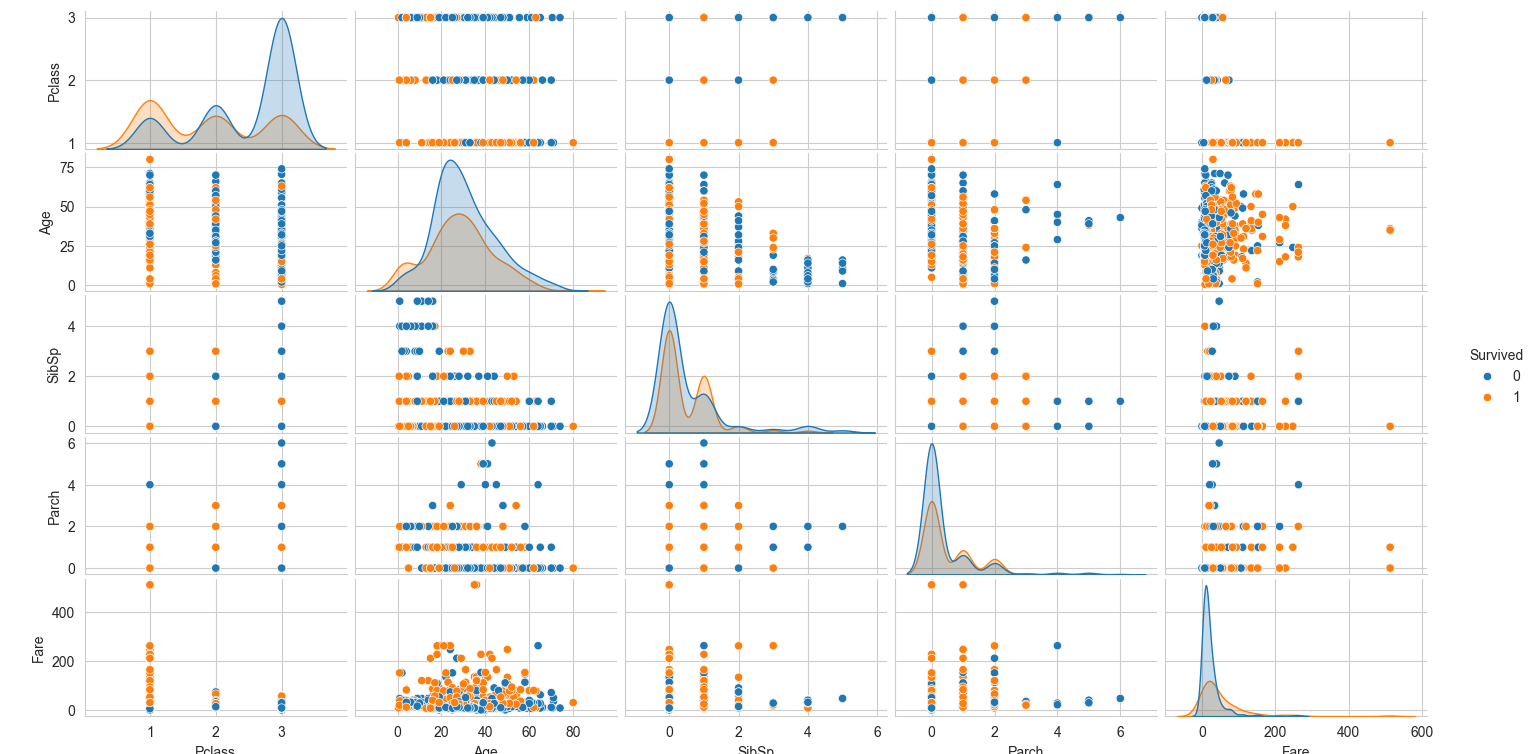
Age is somewhat normal, Fare is highly skewed.



Females and 1st class passengers had the highest survival rates.







Pclass is strongly negatively correlated with Survived.