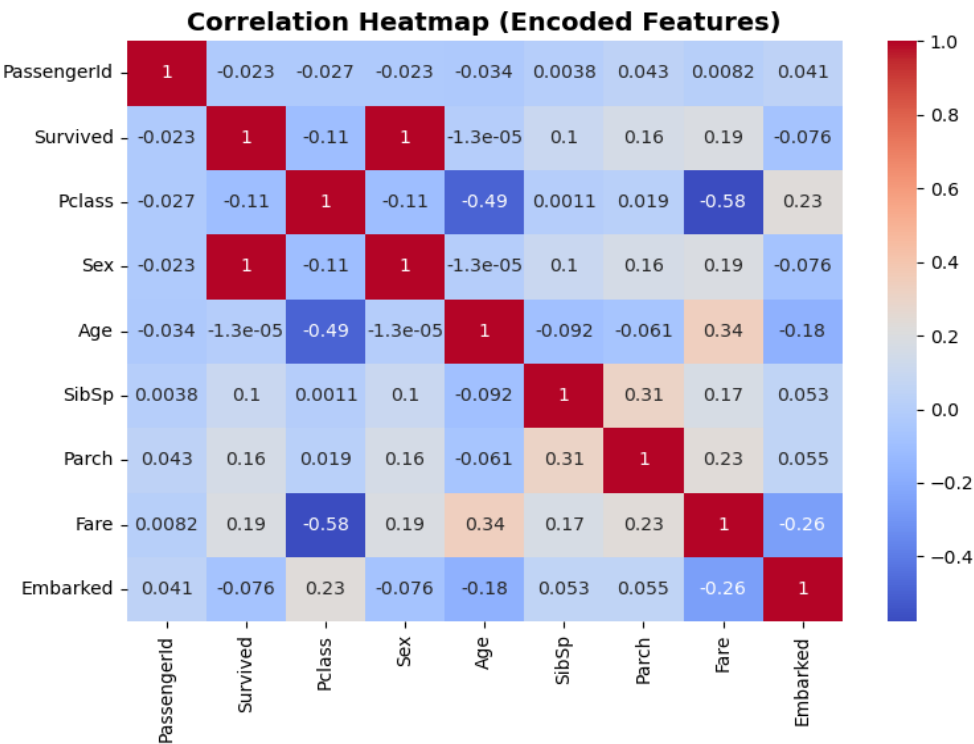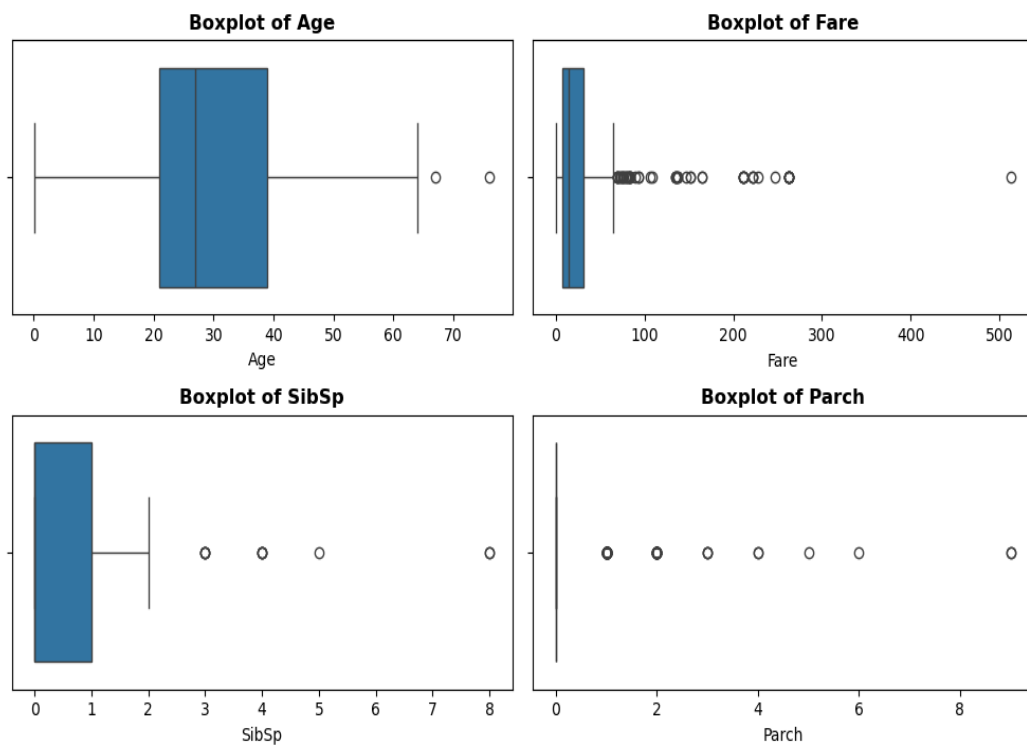# Titanic Dataset Visual Analysis

This report provides a detailed visual analysis of the Titanic dataset. It includes various plots and charts to explore the relationships between different features and the survival outcomes of passengers. Each visual is followed by an explanation of the insights it provides.

Correlation Heatmap

Correlation Heatmap (Encoded Features)

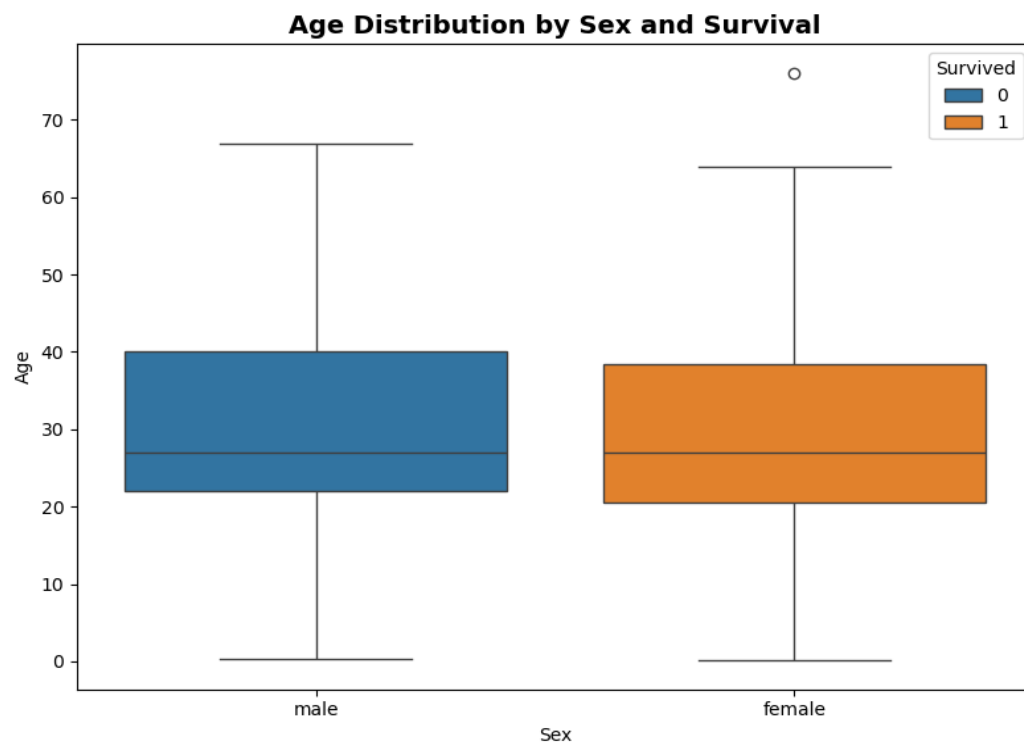|	PassengerId	Survived	Pclass	Sex	Age	SibSp	Parch	Fare	Embarked
PassengerId	1	-0.023	-0.027	-0.023	-0.034	0.0038	0.043	0.0082	0.041
Survived	-0.023	1	-0.11	1	-1.3e-05	0.1	0.16	0.19	-0.076
Pclass	-0.027	-0.11	1	-0.11	-0.49	0.0011	0.019	-0.58	0.23
Sex	-0.023	1	-0.11	1	-1.3e-05	0.1	0.16	0.19	-0.076
Age	-0.034	-1.3e-05	-0.49	-1.3e-05	1	-0.092	-0.061	0.34	-0.18
SibSp	0.0038	0.1	0.0011	0.1	-0.092	1	0.31	0.17	0.053
Parch	0.043	0.16	0.019	0.16	-0.061	0.31	1	0.23	0.055
Fare	0.0082	0.19	-0.58	0.19	0.34	0.17	0.23	1	-0.26
Embarked	0.041	-0.076	0.23	-0.076	-0.18	0.053	0.055	-0.26	1

The correlation heatmap shows the linear relationships between encoded features. Negative correlations indicate inverse relationships, while positive correlations show direct relationships. Notably, Fare has a moderate negative correlation with Pclass, and Age is moderately negatively correlated with Pclass.
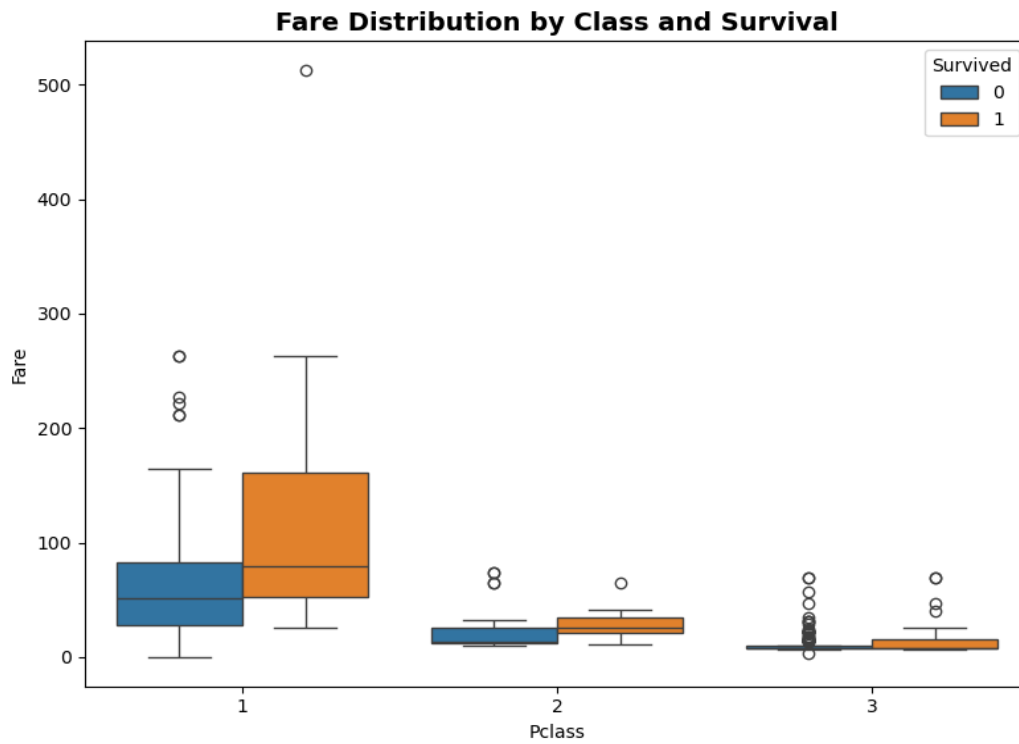
Boxplots

These boxplots display the distribution of Age, Fare, SibSp, and Parch. Outliers are present in Fare, SibSp, and Parch, indicating some passengers paid much higher fares or traveled with unusually large numbers of family members.
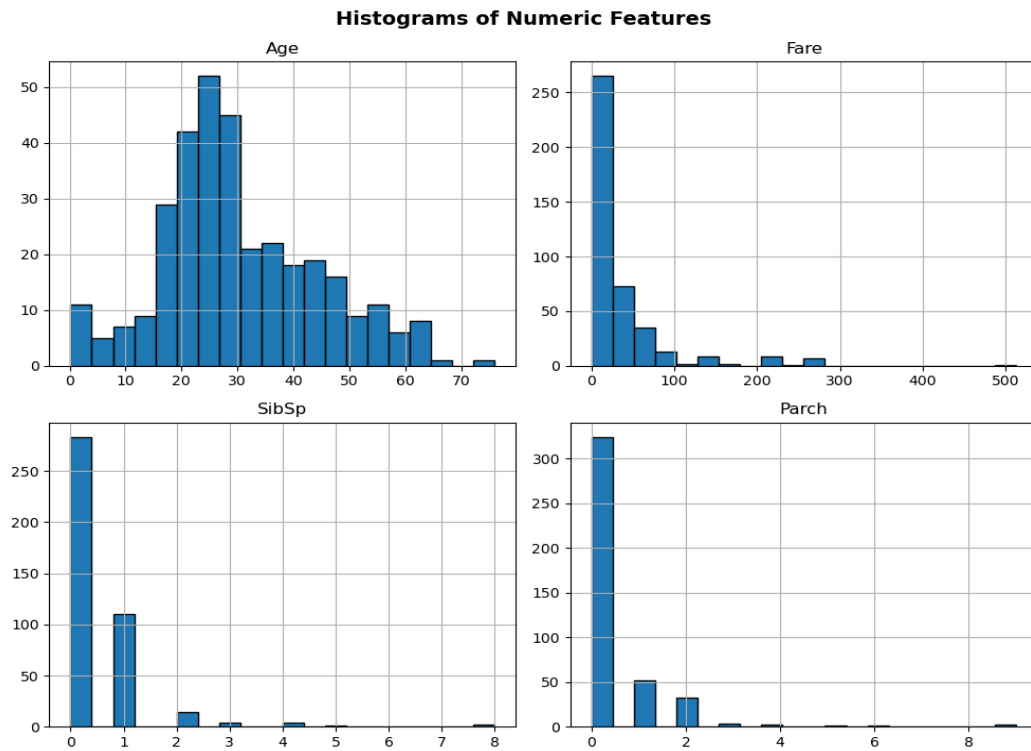
Age Distribution by Sex and Survival

This plot shows the distribution of ages by sex, separated by survival status. Survival rates appear higher among females, with a relatively balanced age distribution.

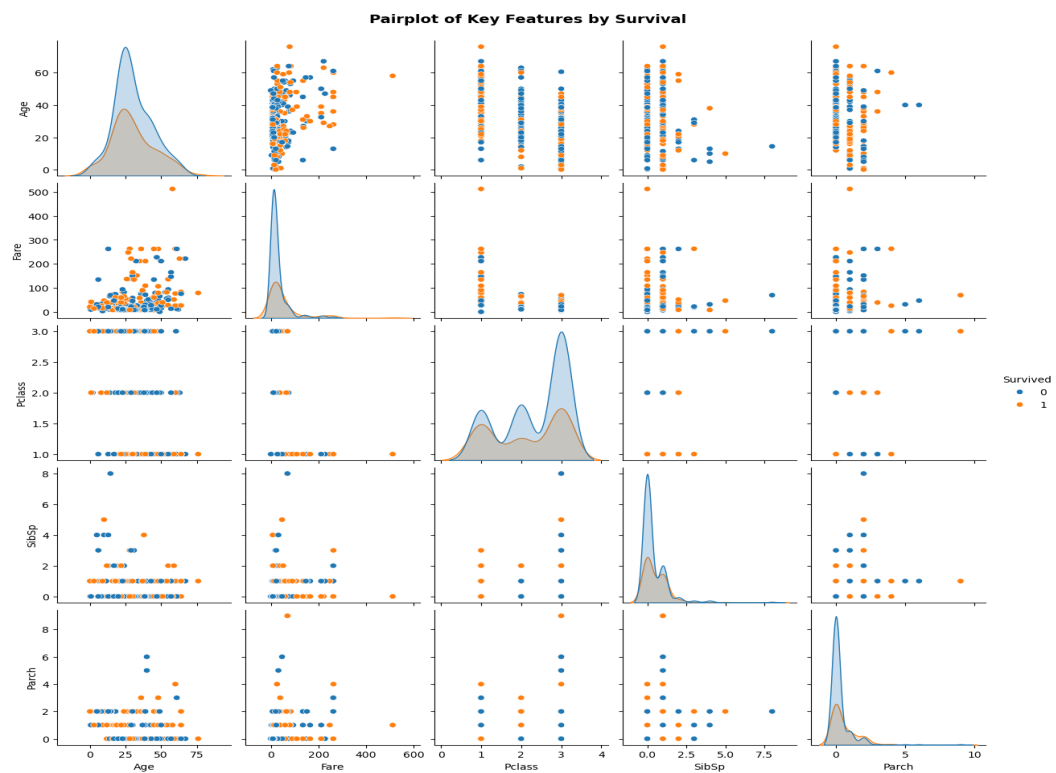# Fare Distribution by Class and Survival

This plot reveals that passengers in higher classes paid significantly higher fares and had higher survival rates, especially in 1st class.

Histograms

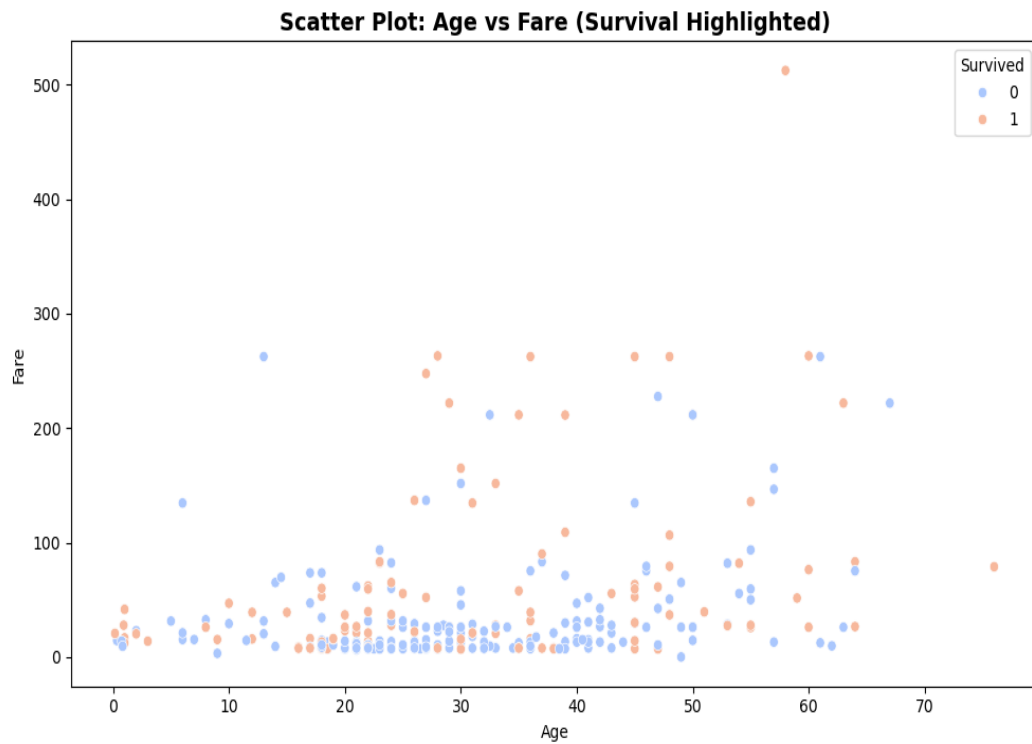
Histograms of Numeric Features

The histograms show the frequency distribution of numerical features. Most passengers were in their 20s–30s, paid relatively low fares, and had few family members aboard.

Pairplot
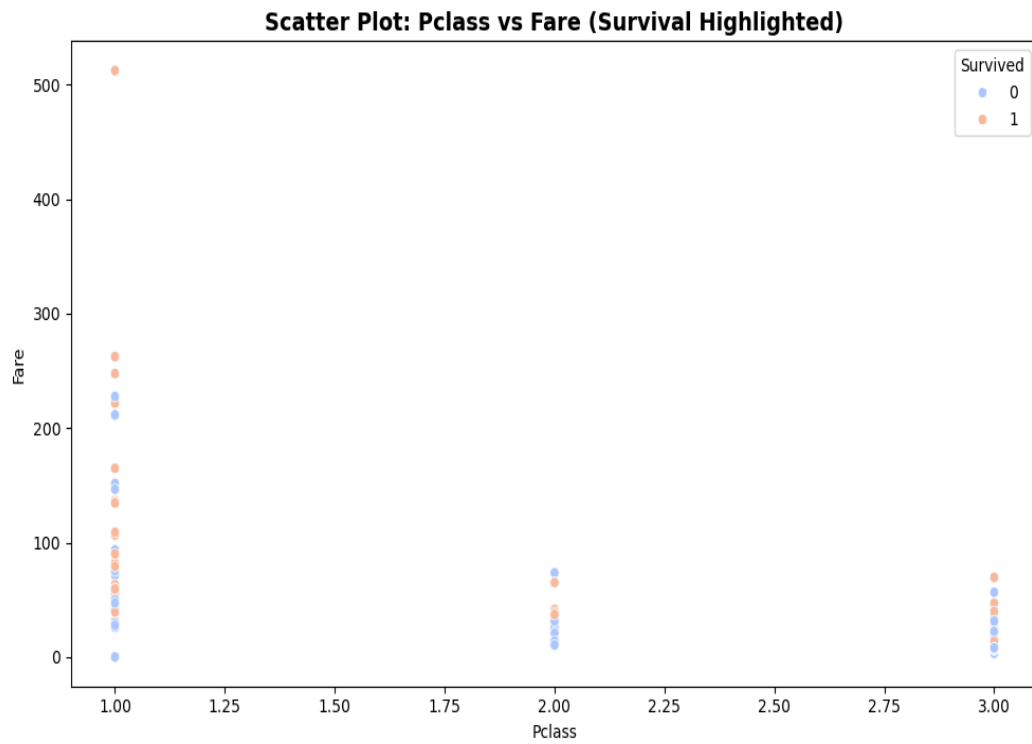

Pairplot of Key Features by Survival

The pairplot visualizes relationships between key features, colored by survival. It shows clear separations in Fare and Pclass between survivors and non-survivors.

Scatter Plot: Age vs Fare

This scatter plot shows the relationship between Age and Fare, with survival highlighted. Survivors are more common among higher-fare passengers, especially in certain age ranges.

Scatter Plot: Pclass vs Fare

This scatter plot demonstrates the relationship between Pclass and Fare, with survival status highlighted. First-class passengers paid higher fares and had better survival outcomes.