Exploratory Data Analysis (EDA) Summary - Titanic Dataset

Pairplot (Survived, Pclass, Age, Fare)

- Passengers in 1st class (Pclass=1) had higher survival rates.

- Younger passengers and those who paid higher fares were more likely to survive.

- Strong relationships exist between Pclass, Fare, and Survived.

Histogram of Age

- Most passengers were aged between 20 and 40.

- Distribution is right-skewed with a few older passengers.

- KDE shows young adults were the most common group.

Boxplot of Fare by Pclass

- Fare decreases with increasing class number (1st > 2nd > 3rd).

- Pclass 1 has the highest median and variability in fare.

- Outliers in all classes indicate variability in ticket prices.

Scatterplot: Age vs Fare

- No strong correlation between Age and Fare.

- Most passengers were young and paid low fares.

- High fares were paid by people of all ages, mostly 1st class.