📋 Titanic Dataset — Exploratory Data Analysis (EDA) Report

📂 Dataset Information

- **Total entries**: 891 passengers
- **Total features**: 15 columns
- **Feature types**:
  - 4 Integer columns
  - 2 Float columns
  - 5 Object (string) columns
  - 2 Boolean columns
  - 2 Category columns
- **Missing data**:
  - 'Age' → missing in several rows
  - 'Embarked' → a few missing values
  - 'Deck' → majority missing

📈 Value Counts for Categorical Features

- **Sex**:
  - Male: 577
  - Female: 314
- **Embarked**:
  - Southampton (S): 644
  - Cherbourg (C): 168
  - Queenstown (Q): 77
- **Who**:
  - Man: 537
  - Woman: 271
  - Child: 83

📋 Titanic Dataset — Exploratory Data Analysis (EDA) Report

- **Embark Town**:
    - Southampton: 644
    - Cherbourg: 168
    - Queenstown: 77

- **Alive**:
    - No: 549
    - Yes: 342

🔍 Observations from Visual Analysis

1. **Missing Values**:
    - 'Age', 'Cabin' (Deck), and 'Embarked' columns have missing data.
2. **Pairplot Insights**:
    - Survived passengers were more likely to be from higher classes (lower pclass number) and had higher fares.
3. **Correlation Heatmap**:
    - 'Fare' and 'Pclass' show strong correlations with survival chances.
4. **Age Distribution**:
    - Most passengers are between 20 to 40 years old.
5. **Fare vs Survival**:
    - Survivors generally paid higher fares.
6. **Passenger Class Distribution**:
    - Most passengers belonged to the 3rd class (lowest ticket class).
7. **Survival Rate by Gender**:
    - Females had a significantly higher survival rate compared to males.

📋 Titanic Dataset — Exploratory Data Analysis (EDA) Report

📝 Summary of Findings

- Younger passengers and passengers who paid higher fares were more likely to survive.
- Gender was a strong predictor of survival; **females** had much higher survival chances than males.
- First-class passengers had better survival outcomes than second- and third-class passengers.
- Columns like **'Age'** and **'Cabin'** require careful treatment for missing values before proceeding to predictive modeling.

🏁 Conclusion

This EDA highlights key variables influencing survival on the Titanic.
Handling missing data properly and focusing on critical features like **gender**, **fare**, and **passenger class** will be crucial in any future machine learning modeling.