Exploratory Data Analysis (EDA) Report — Titanic Dataset

1. Introduction

The goal of this analysis is to explore the Titanic passenger dataset to extract insights related to survival patterns. The dataset contains features including age, sex, passenger class, fare, number of siblings/spouses aboard (sibsp), number of parents/children aboard (parch), embarkation port, and deck.

---

2. Data Overview

- **Number of rows and columns:** See dataset summary.
- **Data types:** Mix of categorical and numerical variables.
- **Missing values:** Notable missingness in age, deck, and embark_town.
- **Unique values:**
  - Sex: male, female
  - Class: First, Second, Third
  - Embarked: Cherbourg, Queenstown, Southampton

---

3. Univariate Analysis

Numerical Features:

- **Age:** Right-skewed distribution; median age around 28; some missing values.
- **Fare:** Skewed distribution with extreme outliers representing higher-paying passengers.
- **SibSp & Parch:** Most passengers had 0–1 family members aboard.

Categorical Features:

- **Sex:** Majority male passengers.
- **Class:** Highest counts in 3rd and 2nd class.
- **Embarked:** Most passengers boarded at Southampton.

Observation: Socioeconomic patterns are evident from class and fare distributions.