

# Exploratory Data Analysis (EDA) - Titanic Dataset

**Objective:** Extract insights using visual and statistical exploration.

**Tools Used:** Python (Pandas, Matplotlib, Seaborn)

## Dataset Overview:

The Titanic dataset contains passenger details such as age, sex, class, fare, and survival status.

## Basic Exploration:

1. Used `df.info()` to check datatypes and missing values.
2. Used `df.describe()` to view statistical summary.
3. Used `value_counts()` to analyze categorical distributions.

## Univariate Analysis:

- Survival count distribution.
- Passenger class distribution.
- Gender distribution.
- Age and Fare distributions.

## Bivariate Analysis:

- Survival by Gender shows females had higher survival.
- Survival by Passenger Class shows 1st class passengers had better survival chances.
- Boxplots of Fare indicate higher fares linked to survival.

## Multivariate Analysis:

- Correlation Heatmap revealed Fare and Pclass correlate with survival.
- Pairplots show separation trends between classes and survival.

## Summary of Findings:

1. Majority of passengers were in 3rd class.
2. Women and children had higher survival rates.
3. 1st class and higher fares were linked to higher survival chances.
4. Missing Age values may require imputation for modeling.

**Outcome:** This EDA improved skills in identifying patterns, trends, and anomalies in real-world data.