# Report of Findings - Titanic Dataset EDA

## 1. Introduction

This report presents exploratory data analysis (EDA) performed on the Titanic dataset using Python (Pandas and Matplotlib). The objective was to uncover trends, relationships, and patterns in the data to support further analysis.

## 2. Histogram - Age Distribution

The histogram of passenger ages shows:
- Most passengers were between 20 and 40 years old.
- A smaller number of passengers were children (under 10) and elderly (above 60).
- The distribution is slightly right-skewed, with a few passengers older than 70.

Observation: The Titanic carried mostly young and middle-aged adults, which could influence survival patterns.

## 3. Boxplot - Age by Survival Status

The boxplot comparing age distributions for survivors and non-survivors shows:
- Median age is similar for both survivors and non-survivors.
- Survivors include more younger children, suggesting possible prioritization during rescue.
- Outliers exist in both groups, representing elderly passengers.

Observation: Younger passengers, especially children, appear to have a slightly higher chance of survival.

## 4. Scatterplot - Age vs Fare

The scatterplot of Age vs Fare reveals:
- Most passengers paid fares under 100, but a few paid significantly higher fares.
- Higher fares were generally associated with older passengers and possibly higher-class cabins.
- Survival was more common among those with higher fares.

Observation: Fare amount may be linked to passenger class, which in turn influenced survival rates.

## 5. Heatmap - Correlation Between Numeric Variables

The heatmap of correlations between numeric features shows:
- Pclass is negatively correlated with Fare (lower class -> lower fare).

# Report of Findings - Titanic Dataset EDA

- Fare has a positive correlation with Survival.
- Age shows weak correlation with Survival, indicating other factors played a stronger role.

Observation: Passenger class and fare were stronger indicators of survival than age alone.

## 6. Summary of Key Insights

- The Titanic's passengers were mostly aged 20-40, with fewer children and elderly.
- Higher survival rates were observed among females, children, and higher-fare passengers.
- Passenger class strongly influenced both fare and survival probability.
- Age alone was not a strong survival predictor, but very young passengers had an advantage.

## 7. Conclusion

This EDA highlights the importance of socio-economic factors in survival chances. In the Titanic disaster, passenger class and fare were stronger determinants of survival than age alone.