

# **Task 2:- Perform data cleaning and exploratory data analysis (EDA) on a dataset of your choice, such as the Titanic dataset from Kaggle. Explore the relationships between variables and identify patterns and trends in the data.**

## **Data Cleaning in Power BI**

1. Data cleaning was performed directly in Power Query Editor
2. Removed duplicates and unnecessary columns
3. Handled missing values (replaced with average/most frequent values)
4. Renamed columns for better readability
5. Changed data types (numeric, categorical, date)

## **Exploratory Data Analysis (EDA)**

Explored patterns, relationships, and trends in the data.

### Univariate Analysis

- Distribution of Age, Fare and survival.
- Count plot of Gender, Class, and Embarked.

### Bivariate Analysis

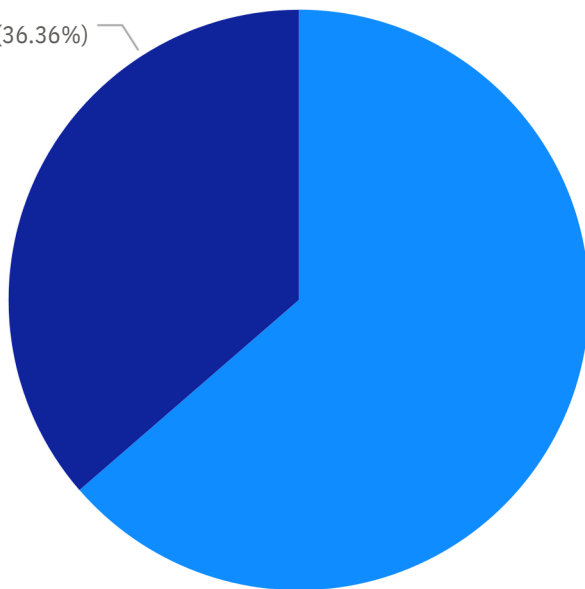
- Relationship between Sex and Survival.
- Relationship between Pclass and Survival Rate.
- Relationship between Age and Fare.

418

Count of PassengerId

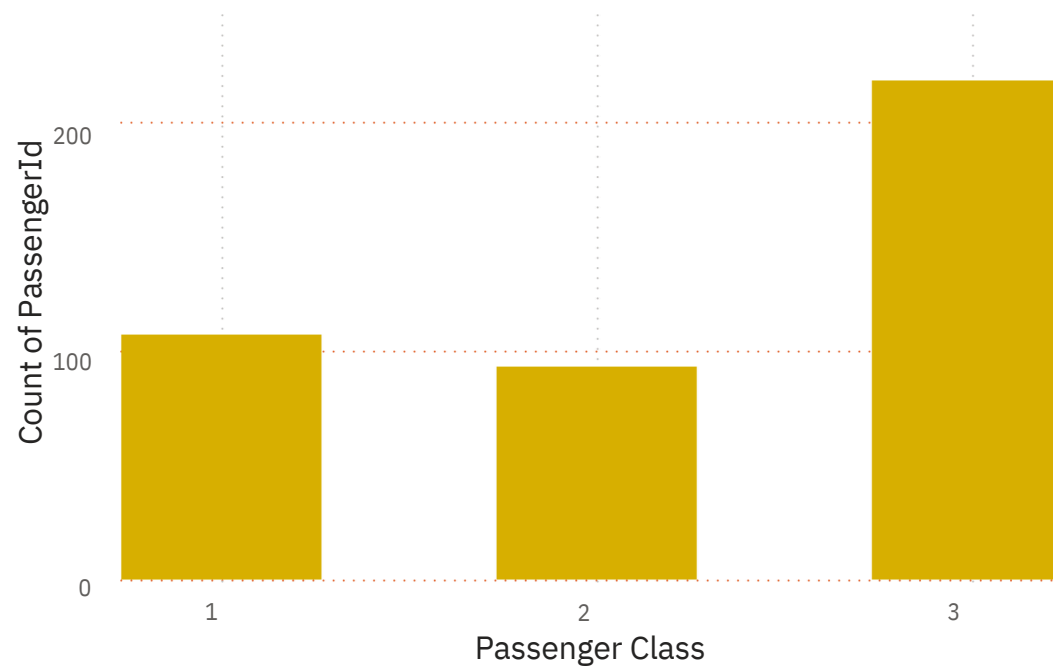
Gender Distribution

152 (36.36%)

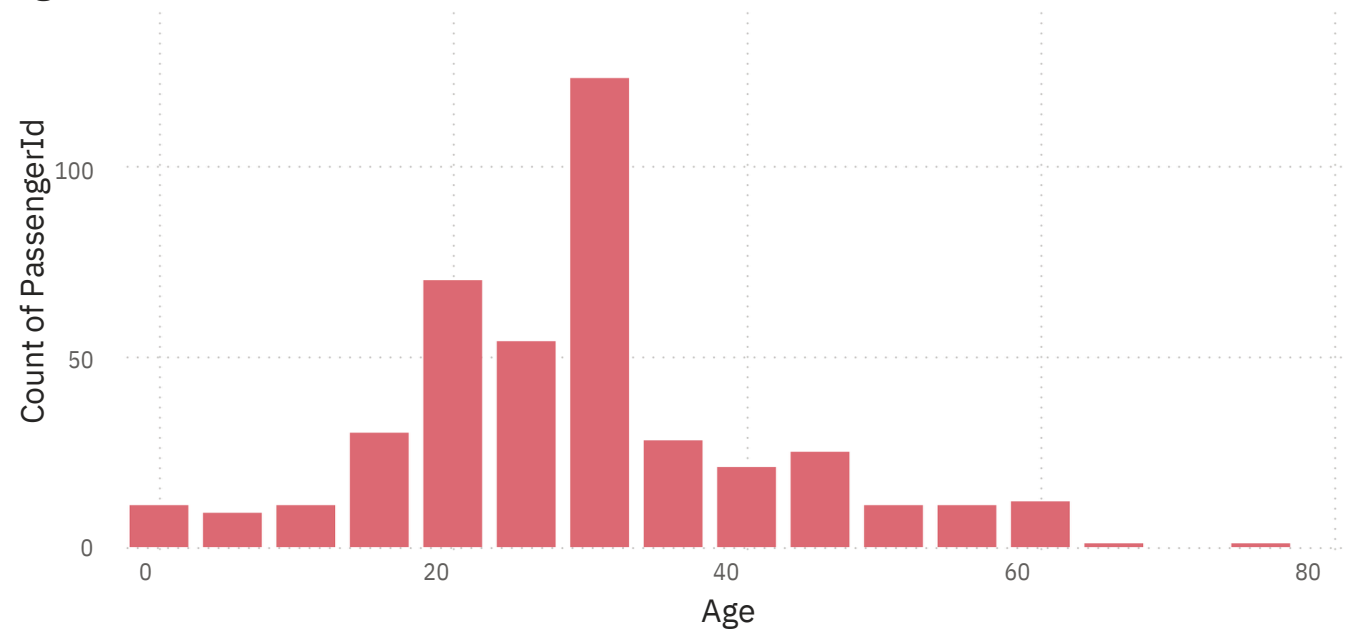


266 (63.64%)

Class Distribution



Age Distribution



Scatterplot of age and fare

