

Titanic Dataset - Exploratory Data Analysis (EDA)

Author: Mohd Shahnewaz

Date: 28th April 2025

1. Introduction

This report presents an Exploratory Data Analysis (EDA) of the Titanic dataset. Our objective is to explore patterns, relationships, and trends affecting passenger survival. We used statistical summaries, visualizations, and observations to derive key insights.

2. Dataset Overview

Basic Information:

None

Statistical Description:

	PassengerId	Survived	Pclass	Age	SibSp	Parch	Fare
count	891.000000	891.000000	891.000000	714.000000	891.000000	891.000000	891.000000
mean	446.000000	0.383838	2.308642	29.699118	0.523008	0.381594	32.204208
std	257.353842	0.486592	0.836071	14.526497	1.102743	0.806057	49.693429
min	1.000000	0.000000	1.000000	0.420000	0.000000	0.000000	0.000000
25%	223.500000	0.000000	2.000000	20.125000	0.000000	0.000000	7.910400
50%	446.000000	0.000000	3.000000	28.000000	0.000000	0.000000	14.454200
75%	668.500000	1.000000	3.000000	38.000000	1.000000	0.000000	31.000000
max	891.000000	1.000000	3.000000	80.000000	8.000000	6.000000	512.329200

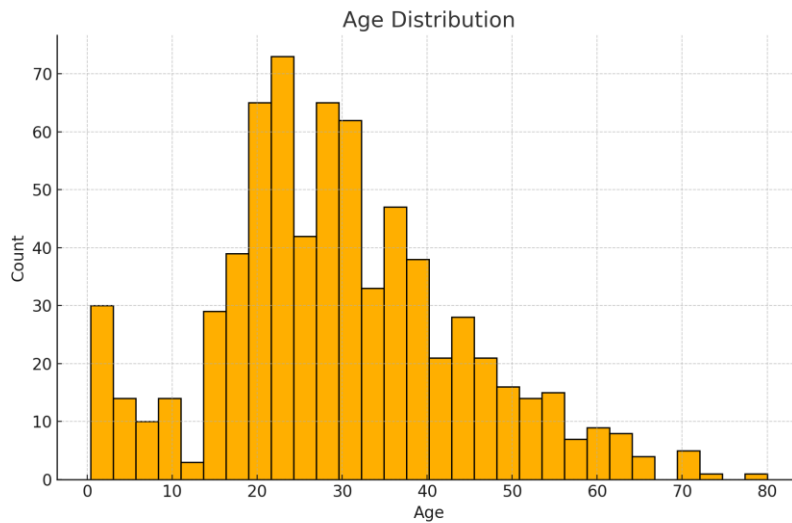
Missing Values:

PassengerId	0
Survived	0
Pclass	0
Name	0
Sex	0
Age	177
SibSp	0
Parch	0
Ticket	0

Fare 0
Cabin 687
Embarked 2

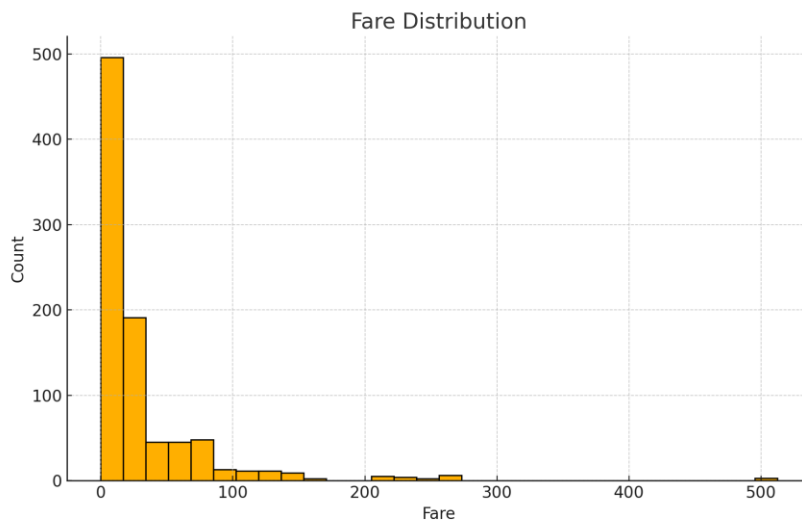
3. Univariate Analysis

3.1 Age Distribution



Observation: Most passengers were between 20–40 years old.

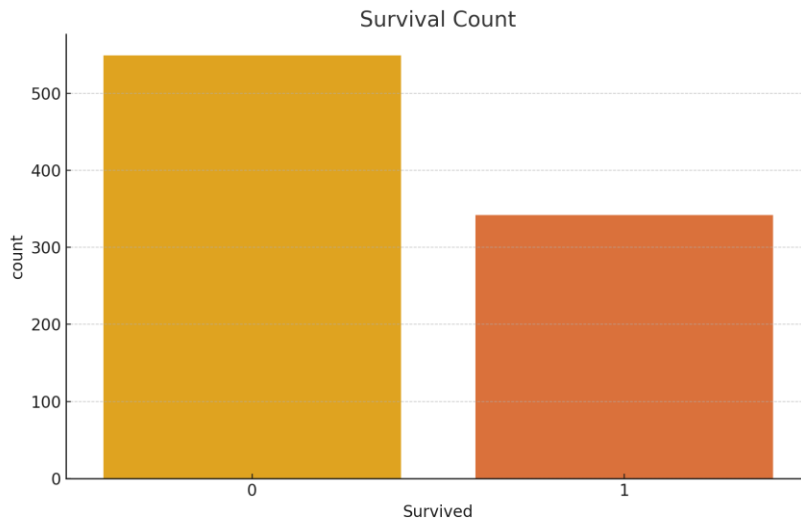
3.2 Fare Distribution



Observation: Majority of fares were lower values; few outliers with very high fares.

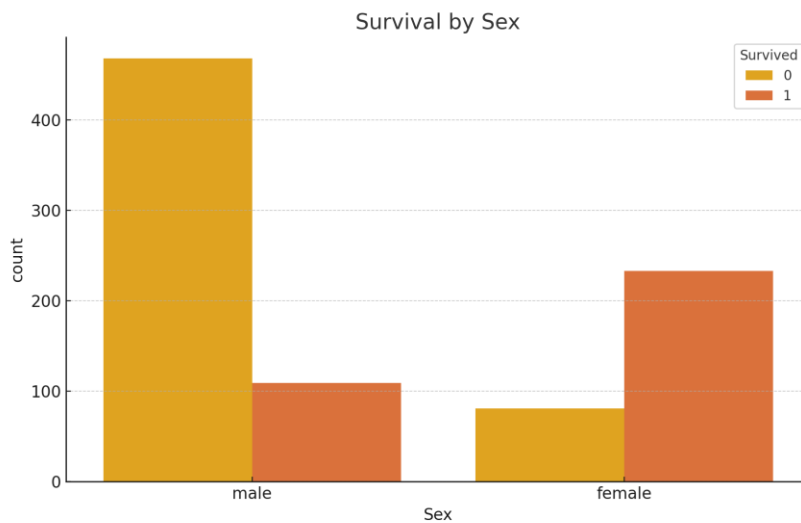
4. Categorical Analysis

4.1 Survival Counts



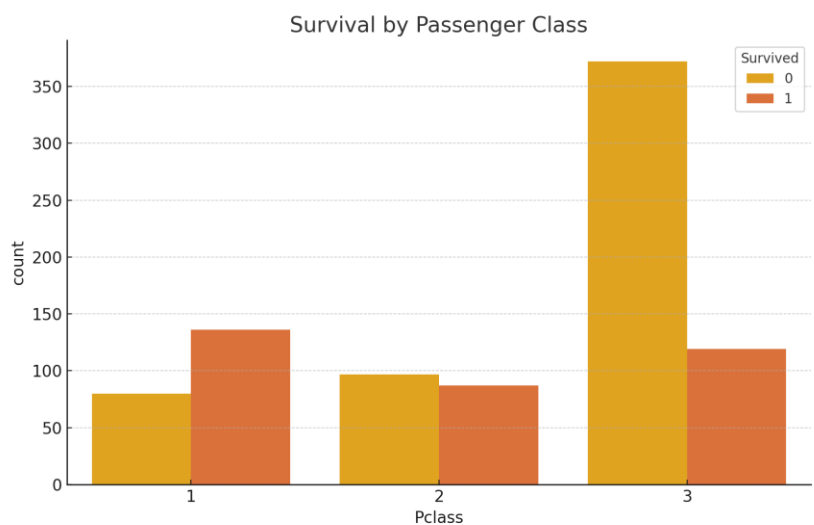
Observation: More passengers died than survived.

4.2 Survival by Sex



Observation: Females had much higher survival rate compared to males.

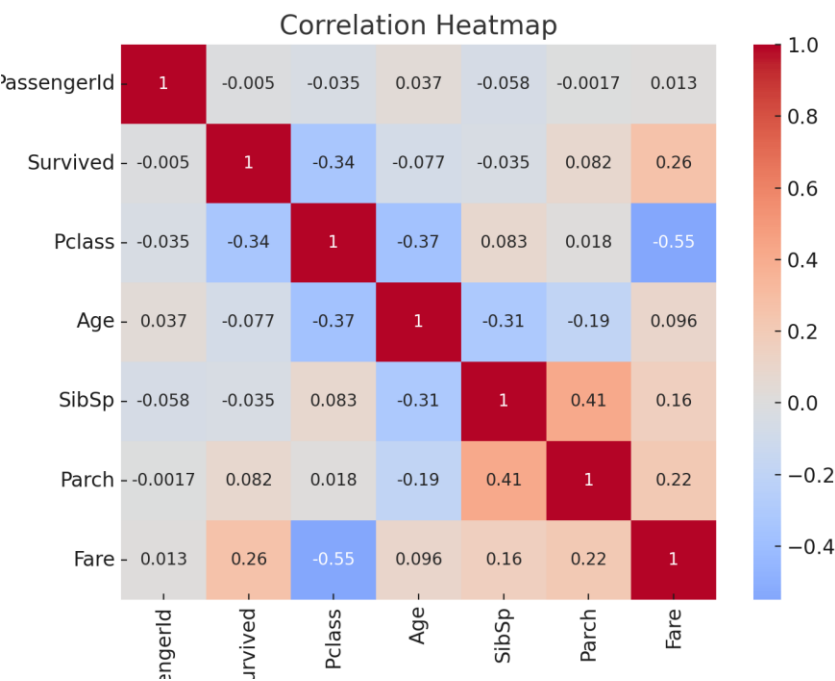
4.3 Survival by Passenger Class



Observation: 1st class passengers had higher survival rates.

5. Multivariate Analysis

5.1 Correlation Heatmap



Observation: Fare and Pclass are negatively correlated. SibSp and Parch are positively correlated.

6. Summary of Findings

- Females had significantly higher survival rates than males.
- Higher class (1st class) passengers had better survival chances.
- Higher fares correlated with better survival.
- Children and young adults had slightly better survival chances.
- Age, Cabin, and Embarked columns have missing values.