

Task5

About the data file

Number of data: 891 rows × 12 columns

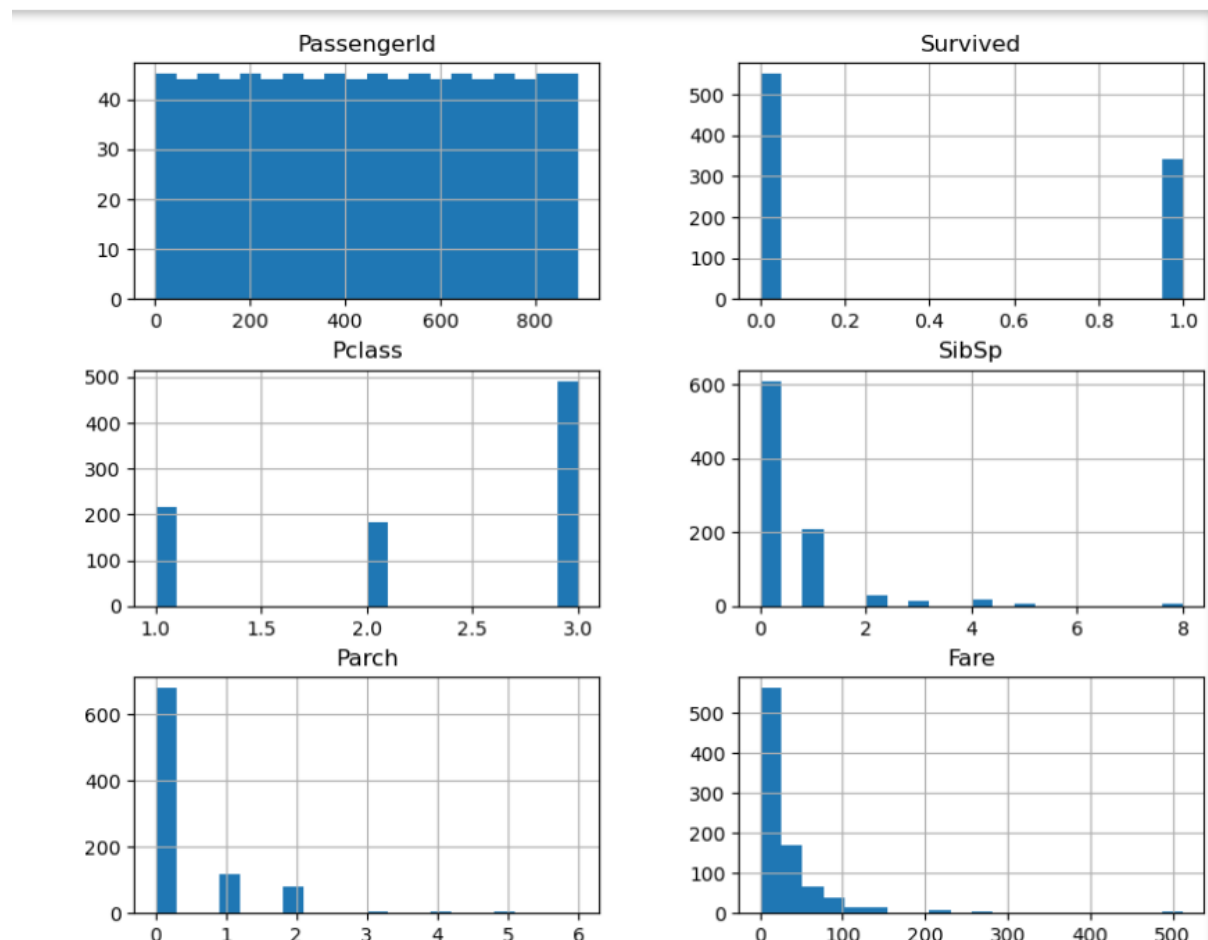
Columns: PassengerId, Survived, Pclass, Name, Sex, Age, SibSp, Parch, Ticket, Fare, Cabin, Embarked

Data types: 6 object (categorical), 5 integer, 1 float

Duplicates: 0 rows

Findings for Visualization

1.



Majority did not survive.

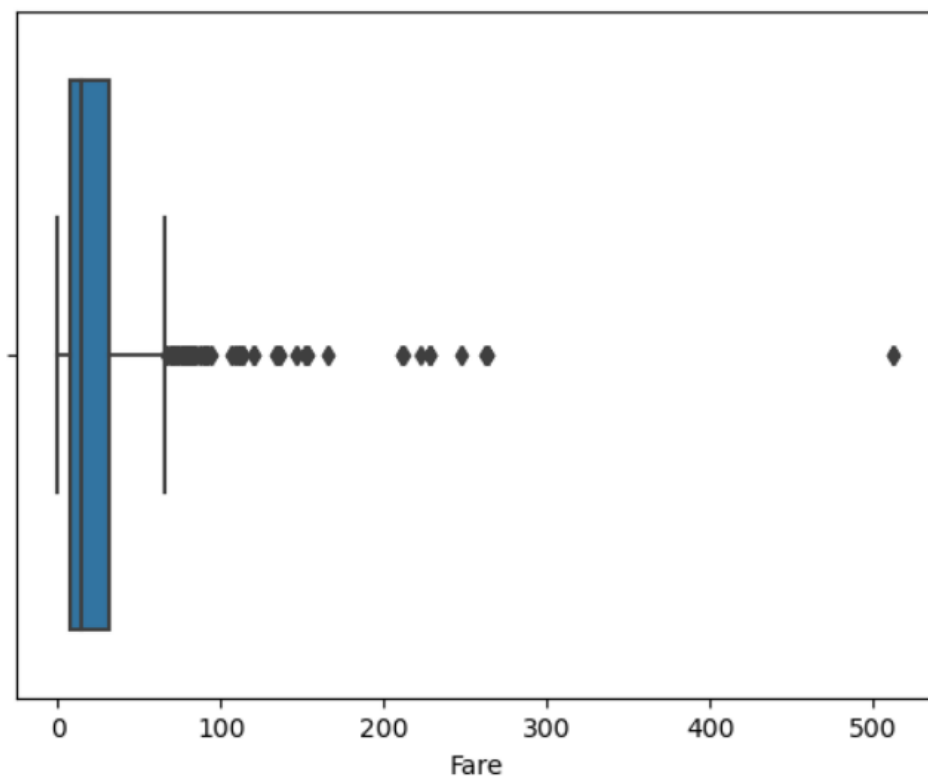
Most passengers were 3rd class.

Most traveled alone (SibSp=0, Parch=0).

Fares mostly low, few high outliers.

Data shows class and fare may affect survival.

2.

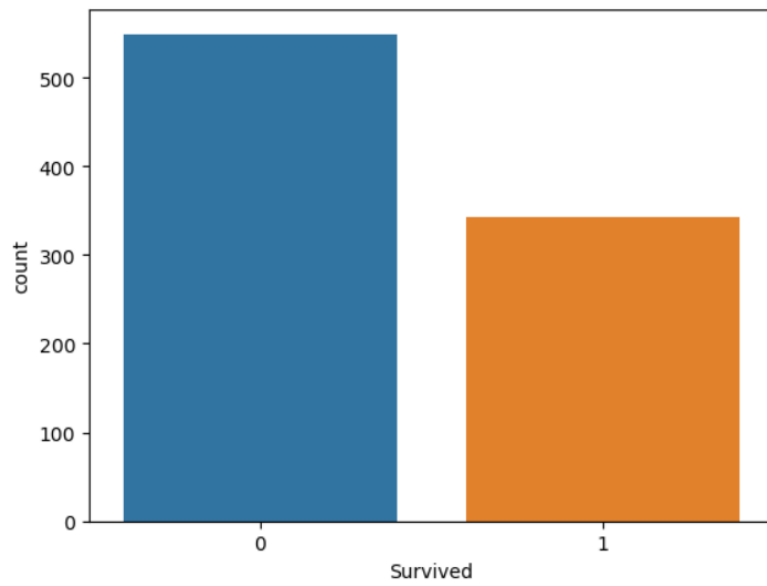


Most fares are concentrated below ₹100.

There are many outliers — passengers who paid much higher fares, some exceeding ₹500.

The median fare is low compared to the maximum.

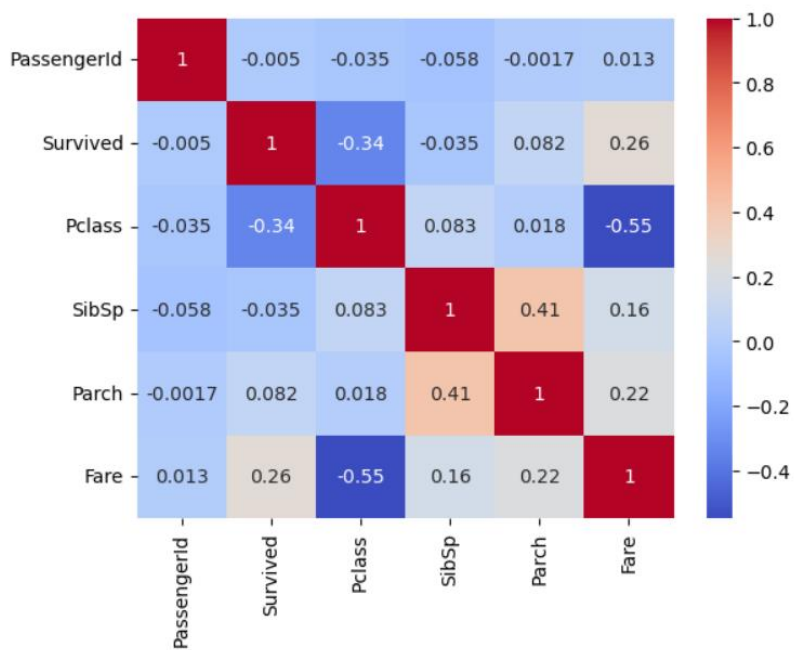
3.



0 (Did not survive): More than 550 passengers.

1 (Survived): Around 340 passengers.

4.



Survived & Pclass: -0.34 → Higher class (lower Pclass value) linked to better survival.

Survived & Fare: 0.26 → Higher fare linked to higher survival chances.

Pclass & Fare: -0.55 → First-class tickets were more expensive.

SibSp & Parch: 0.41 → More siblings/spouses onboard correlates with more parents/children onboard.

Most other correlations are weak or near zero.