PREDICTION OF TITANIC SURVIVAL ANALYSIS

Name, age, gender, survival (0,1), ticket class (1st, 2nd, 3rd), number of children, parents, spouses or siblings aboard Titanic, ticket number, cabin number, passenger fare, and embarkation port (Southampton, Cherbourg, Queenstown,) are all included in this analysis.
A thorough investigation was carried out, with the most fascinating findings being discussed below:

VARIABLES

- Prediction variable, i.e., whether or not they survived the sinking
- Numeric variables such as passenger ids, ages, and so on.
- Categorical variables, such as the ticket's class, and string variables, such as the name, are examples of categorical variables.

pclass: A socioeconomic status identifier (SES),

- 3rd grade = lower
- 2nd grade = middle
- 1st grade = upper

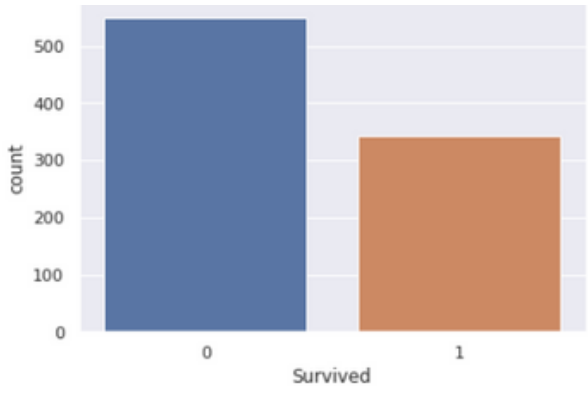
age: If you're under the age of one, your age is fractional.
Family relationships are defined in this way in the dataset

- Brother, sister, stepbrother, and stepsister are all examples of sibling.
- Mistresses and fiancés were not counted as spouses.
- Mother and father are two examples of parents.
- Daughter, son, stepdaughter, and stepson are all examples of children.
- Because some children only travelled with a nanny, parch=0 was assigned to them.

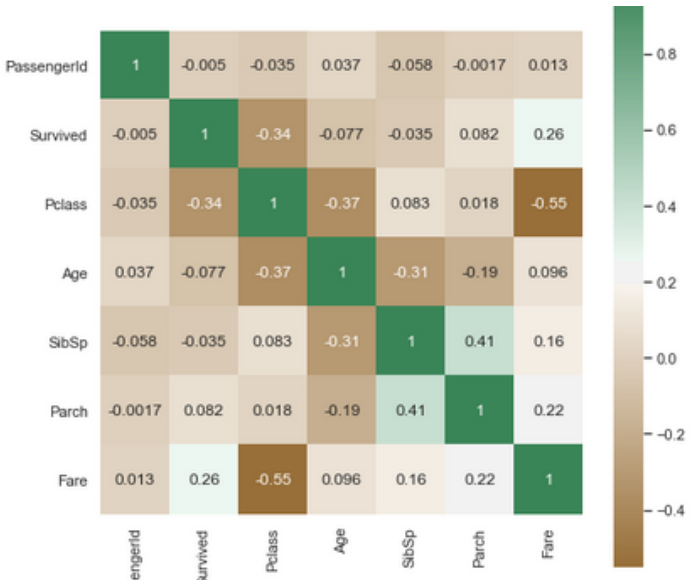
NUMBER OF PEOPLE WHO SURVIVED AND THOSE WHO DID NOT

- Survived = 549
- Not survived = 342
-

The number of passengers who died in the Titanic disaster is more than the number of individuals who survived.



CORRELATION BETWEEN CHARACTERISTICS



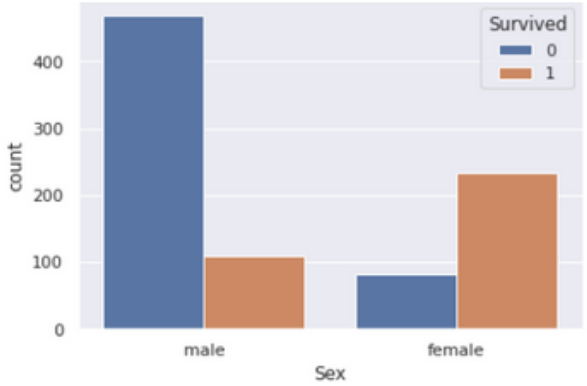
The Correlation Between Features Provides Insights:

- The Pearson's correlation coefficient for fare and pclass is -0.55; obviously, higher fares suggest better ticket classes (lower the class number) and vice versa.
- The target feature and pclass are somewhat negatively linked (-0.34), indicating that the higher the ticket class, the better the chance of survival.
- The features parch, sibsp, and age all follow a similar trend. Both measures, parch and sibsp, have a very minor connection with all features except Age, and both have a negative correlation with Age (-0.19 and -0.31, respectively). The passenger's age determines the number of family members who accompany him or her.
- The features sibsp and parch are positively associated, indicating the number of family members accompanying the traveller.
- The ticket class and pclass are inversely associated (-0.37), meaning that the older you are, the better the ticket class.

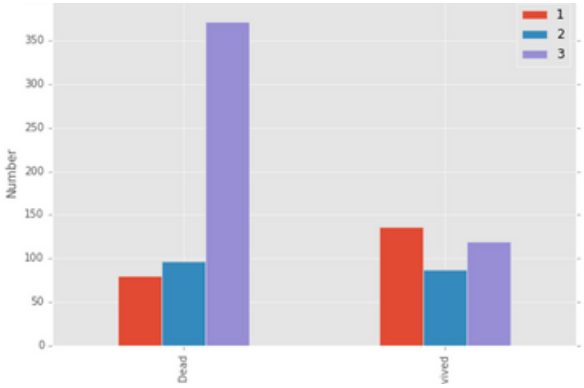
FEMALE OR MALE SURVIVORS COUNT

- Male = 577
- Female = 314

More than female passengers, male passengers died in the Titanic disaster. As a result, during the Titanic disaster, females were given more priority in terms of being saved.

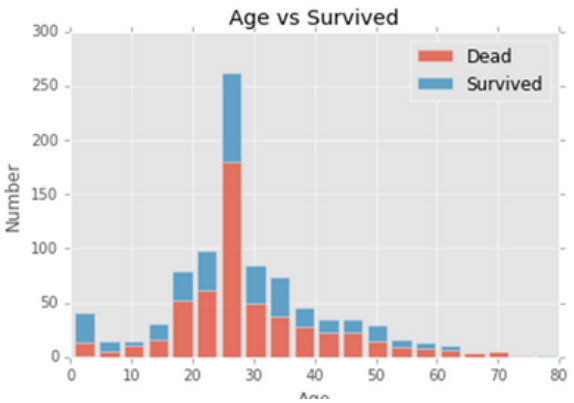


PORT EMBARKED

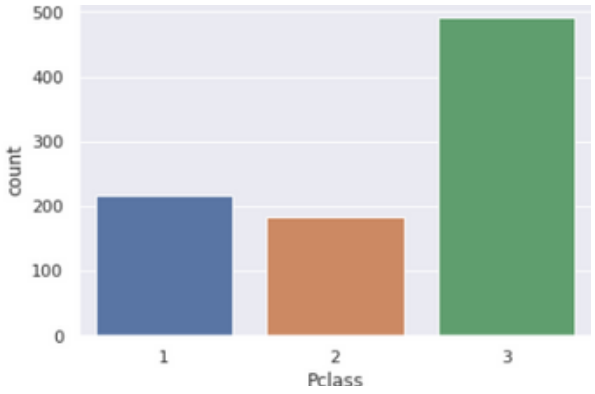


The graph above demonstrates that Port C is more expensive on average than other ports. It also illustrates that individuals who departed from port C were having somewhat better probability of surviving.

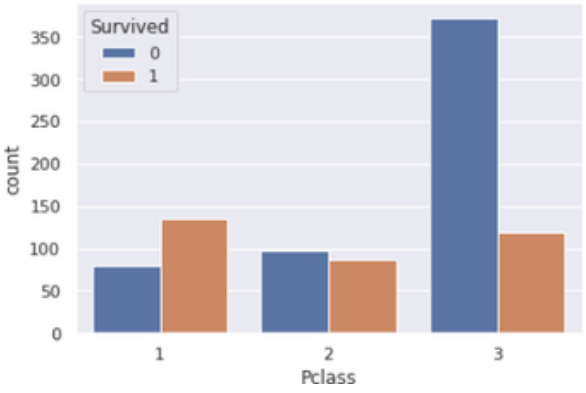
AGE



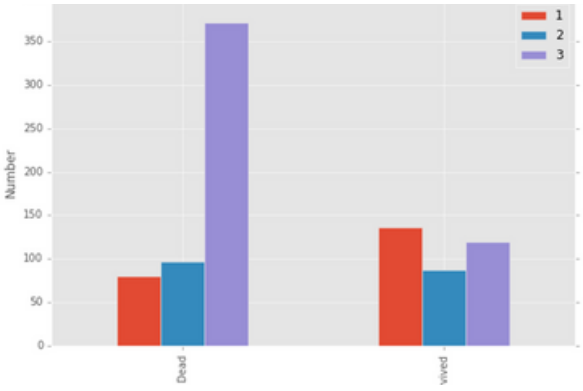
BASED ON PROXY FOR SOCIO-ECONOMIC LEVEL, THE NUMBER OF PEOPLE WHO SURVIVED



BASED ON A PROXY FOR SOCIOECONOMIC CLASS, THE NUMBER OF PEOPLE THAT SURVIVED WAS CALCULATED.



CLASS OF TICKET(PCLASS)



Above graph demonstrates that passengers of pclass 1 tickets are supposable to survive while passengers of pclass 3 tickets are supposable to not survive as pclass 1 tickets are more expensive on average than pclass 3 tickets.