

CPSC Assignment 3 – Part 2

Background

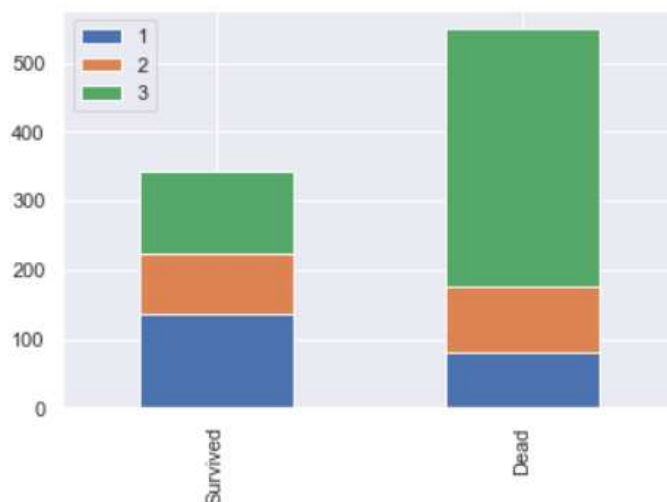
The sinking of the Titanic is one of the most infamous shipwrecks in history.

On 15 April 1912, during her maiden voyage, the widely considered “unsinkable” RMS Titanic sank after colliding with an iceberg. Unfortunately, there weren't enough lifeboats for everyone onboard, resulting in the death of 1502 out of 2224 passengers and crew.

While there was some element of luck involved in surviving, it seems some groups of people were more likely to survive than others. (Anon., n.d.)

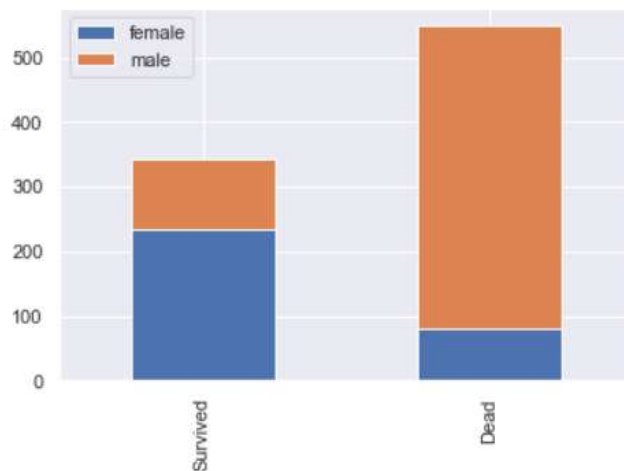
Here are three possible hypotheses that can define for this dataset:

Determine if the survival rate is associated to the class of passenger



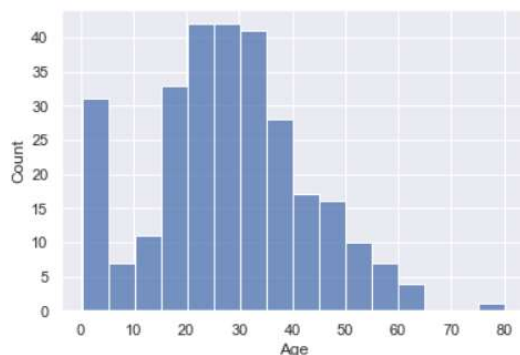
By using Chi-square, we found that if Degree of freedom=2, So decision point of chi square is 5.99, now is 102.88898875696056 which is extremely bigger and far from the Decision point. Chi square in this part is relatively high. The P-value for Pclass is negative (very small $4.549251711298793e-23$) and insignificant too. Hence, reject the null hypothesis. We can say with a 95% Survival rate is dependent on Passenger Class and some leaning of survivability towards better passenger class.

Determine if the survival rate is associated to the gender

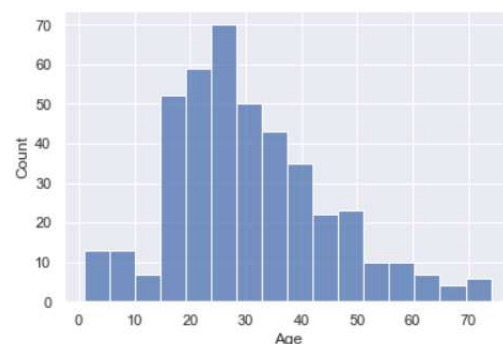


By using Chi-square, we found that if Degree of freedom=1, So decision point of chi square is 3.84, now is 260.71702016732104 which is extremely bigger and far from the Decision point. Chi square in this part is also relatively high. Besides, The P-value for gender is negligible ($1.1973570627755645e-58$), less than 0.05 (5%) and H_0 can be rejected, which means it's correlated. We can say 95% of survival rate for female is greater than Males.

Determine the survival rate is associated to the age



Correlation for Survived = 1 vs Age



Correlation for Survived = 0 vs Age

The histogram for Age shows most of the passengers survived = 1 and survived = 0 are in the age range of 20 – 40. It defines considerably correlation, Although the correlation is not very high.

References

Anon., n.d. *Kaggle*. [Online]

Available at: <https://www.kaggle.com/c/titanic/overview/description>

[Accessed 9 Mar 2022].