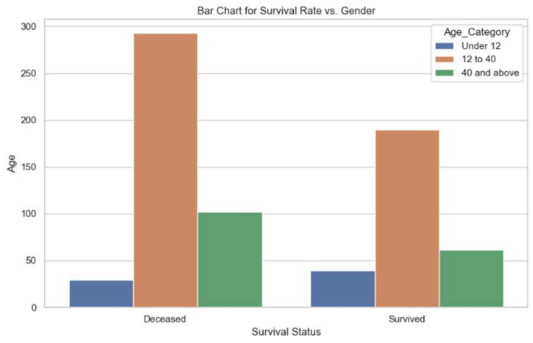
Titanic Dataset

Below are our findings for the Titanic dataset for the 3 given hypothesis.

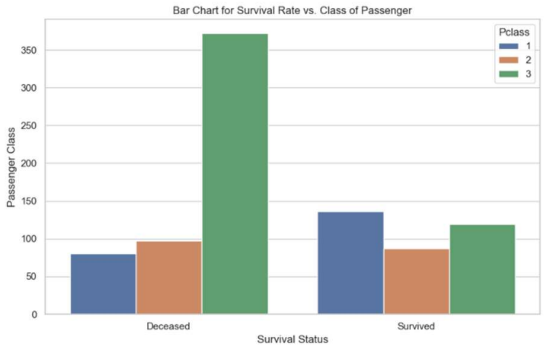
Relationship between Survival Rate and Age



To test the relationship, we grouped age variable into three different groups: children under 12, 12 to 40 years old, and above 40 years old. The distribution between each group does not look similar, indicating that age is dependent of survival rate. Interesting observation: proportion of children under 12 who survived is greater than those who did not survive.

The chi-square test of independence results in p-value of 0.01, confirming our conclusion that age and survival rate are dependent on each other.

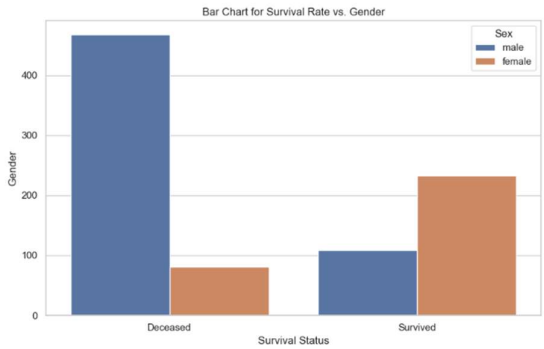
Relationship between Survival Rate and Passenger Class



The distribution of the bar chart does not look similar, indicating that survivability on Titanic is dependent on which class you are in. Passengers in 1st class has greater chance of surviving compared to passengers in the 3rd class. Additionally, proportion of passengers in 1st class who survived is greater than the proportion of those who did not survive.

The chi-square test of independence results in p-value of nearly 0, confirming our conclusion that passenger class and survival rate are dependent on each other.

Relationship between Survival Rate and Gender



Since the distribution of the bar chart does not look similar there is an indication that survival rate is dependent on gender. It is more likely for a female to survive compared to male.

The chi-square test of independence results in p-value of nearly 0, confirming our conclusion that gender and survival rate are dependent on each other.