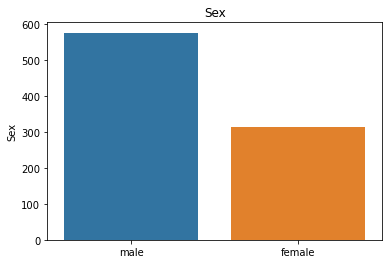
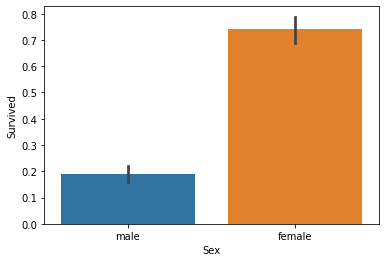
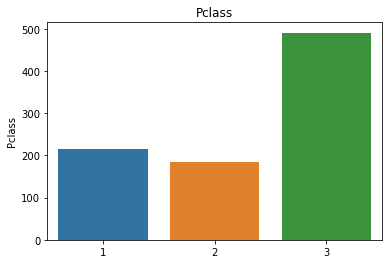
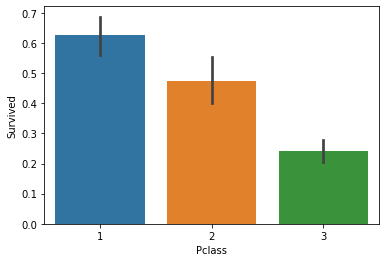
**Intro:** This PDF will go over my 3 findings from the titanic dataset and include associated graphs.

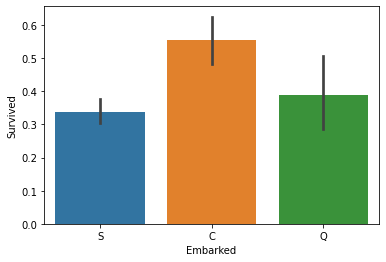
**First hypothesis:**

We began by checking the survival rate with the associated sex as shown in the bar graph below. There are 312 females in the data and 577 males. 231 out of 312 females survived thus giving us a 74% survival rate and 109 out of 577 males survived giving a 18% survival rate. From these proportions alone we can see that there may be significance with gender playing a roll in survival. The bar graph below shows a side-by-side comparison of these survival rates. Emphasizing the ‘women’ in the popular saying “women and children first”.

We can test its statistical significance by doing a Chi Squared test. Below you can see our contingency table for the data. We will assign the array to ‘sexsurv’ to compute the Chi Square analysis. To test our hypothesis, we must create our null and alternative hypothesis. The null hypothesis states that there is no correlation between sex and survival rate. Our alternative hypothesis states that there is a correlation between sex and survival rate. Using out Chi2 Contingency algorithm we found that our data is statistically significant as our statistic is larger than our Degree of freedom. This means we reject the null hypothesis and state there is a correlation between the sex and survival rate.

**Second Hypothesis:** There are 3 classes of passengers in this dataset. Class 1,2, and 3. In our python tables we found that 55% of passengers were classified as class 3 (the lowest class) but had a 24% chance of survival. On the contrary Class 1 was 24% represented but had a 62% chance of survival. The table visualizes the survival rate for each class clearly as the lower class you got the less likely you were to survive. We can again test this statistically with a Chi square analysis. The null hypothesis states that there is no correlation between class and survival rate. Our alternative hypothesis states that there is a correlation between class and survival rate. After our analysis we found that we can reject the null hypothesis and state there is a statistical significance between class and survival.

**Third Hypothesis:**

Chart, bar chart

Description automatically generated We want to see if where a passenger embarked from has any impact on survival rate. We must remove 2 data points as they are null values that we do not want to skew the data. The ports are classified as C, Q, and S which stand for Cherbourg, Queenstown, and Southampton. 72% of passengers boarded from Southampton yet only 33% survived. After another Chi Square analysis, we can see that there is statistical evidence. This means where you embarked affected the survival rate.

Overall, our analysis shows that there is significant correlation with a passenger’s survival rate and their sex, class, and embarked port. These meaning that if you were a first class, women, or embarked from Cherbourg, your survival rate was significantly higher than someone who was a male, 3rd class or embarked from Southampton.