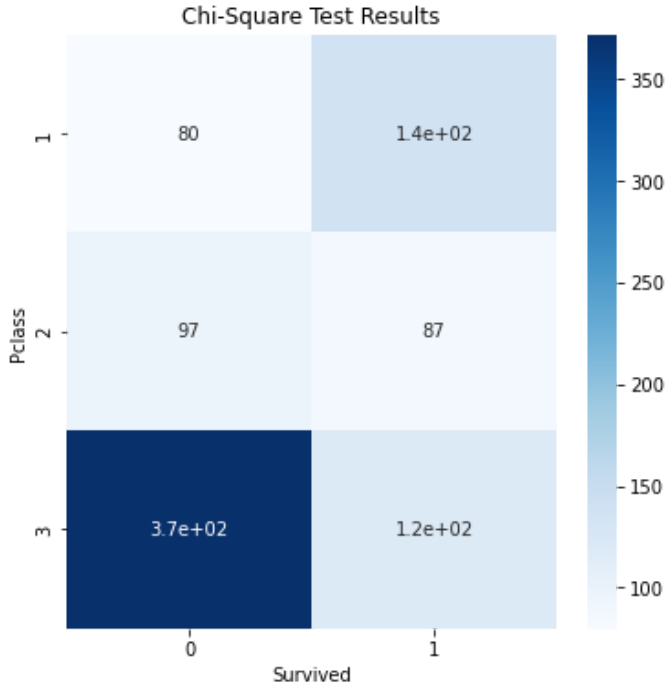
Determine if the survival rate is associated to the class of passenger

As we know Survival rate is represented by variable "**Survived**" and class of Passenger by "**Pclass**" . These both are Categorical Variables so we will be doing **Chi Square Test** on these variables to determine if they are dependent or independent of each other .

Null Hypothesis (H_0) - Survived and Pclass are *independent* of each other .
Alternate Hypothesis(H_A) - Survived and Pclass are *dependent* on each other .

P- value turns out to be less than 0.05 (4.54925e-23) therefore we can say that we reject the Null Hypothesis and both the Variables are Dependent .

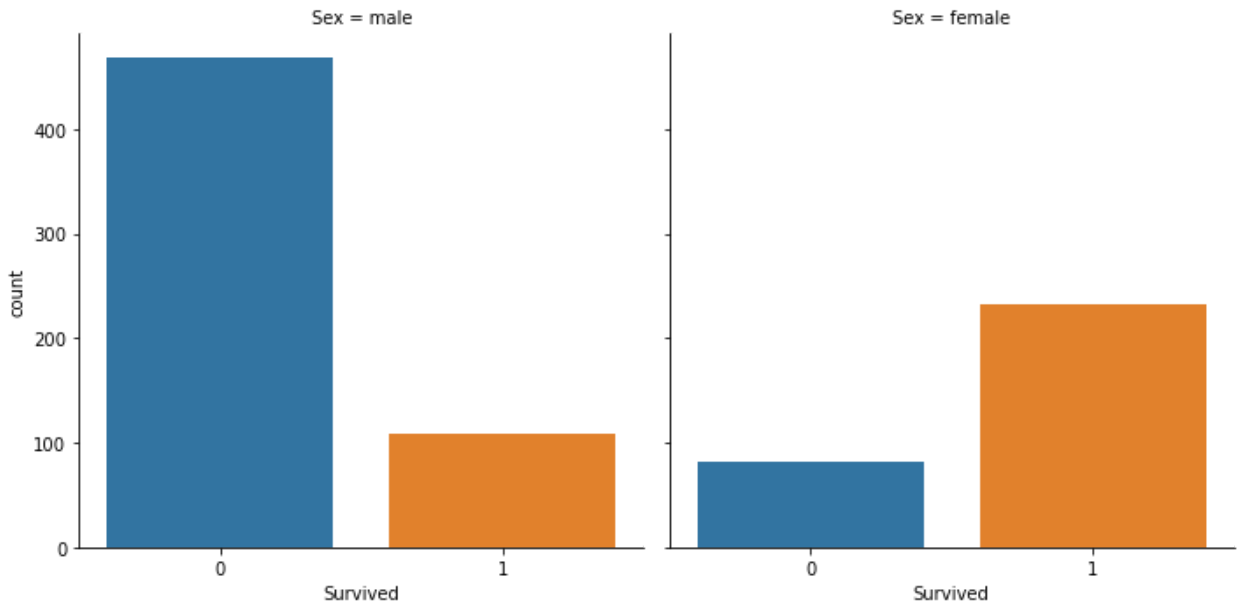
We will be plotting a **Heat Map** to visualize the association between the variables .



As you can clearly see above that there were more survivors belonging to class 1 and as we go down the ladder in Class , we see there were a high number of people who were not able to make it .

Determine if the survival rate is associated to the gender

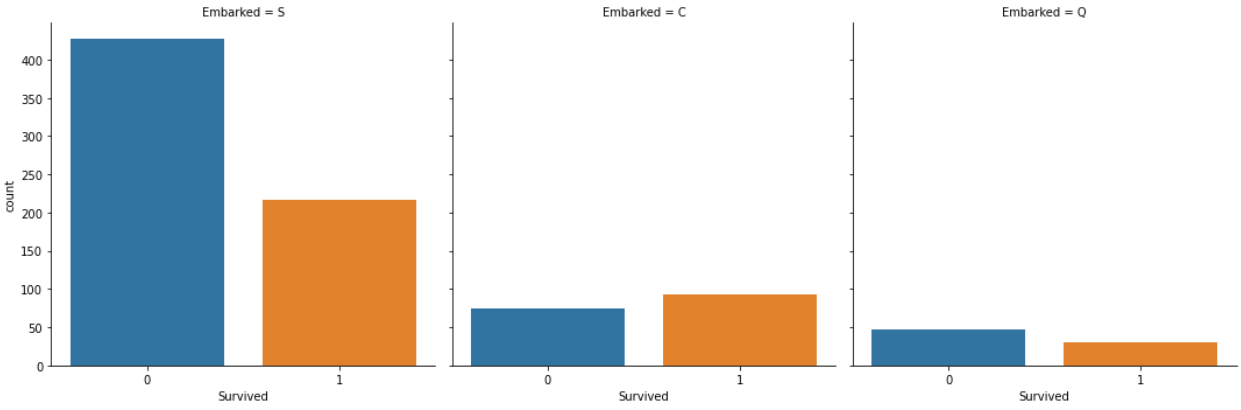
We applied Chi Square Test on "**Survived**" and "**Sex"** variable and came to know that both these variables are dependent on each other . Let us visualize this on a Bar Plot .



As we can clearly see in the plot above , there were a high number of male casualties and when looking at females , we observe that more females survived .

Determine if the survival rate is associated to the Embarked

We applied the Chi Square Test on the "**Survived**" and "**Embarked**" variables and came to know that both these variables are dependent on each other . Let us visualize this on a Bar Plot .



The above plot clearly represents that the majority of the passengers who belonged to Embarked category "S" did not survive and so is the case with category "Q" . Whereas in Category "C" , the number of survivors slightly exceeded those who did not .