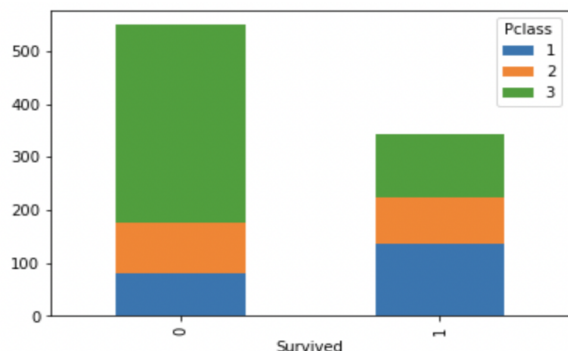


Titanic Dataset Analysis

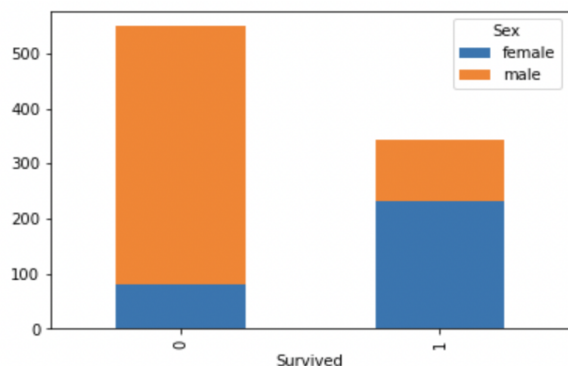
This report's objective is to determine whether there exists a relationship between the survival rate and the class of the passenger, their gender and their age. To test for association between categorical variables, contingency tables were created for each relationship and chi-square tests were performed on those contingency tables. Then, stacked bar plots were built to confirm the results of the chi-square tests.

I. Survival Rate and Class of the Passenger



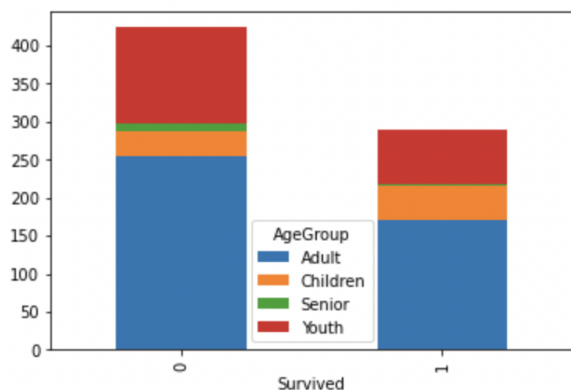
When testing for the association between survival rate and class of the passenger, the chi-square test resulted in a p-value of $4.549251711298793e-23$. This is an extremely small number, smaller than 0.05, which means that there is a statistically significant relationship between these variables. To further prove this point, the stacked bar plot to the left shows that the majority of people that died (represented with 0) were from the lower class (represented with 3).

II. Survival rate and Gender



The association between survival rate and gender was also found to be an statistically significant relationship, as the p-value for the chi-square test was $1.1973570627755645e-58$, an also significantly smaller number than 0.05. The graph to the left also shows this relationship. The majority of people that died in the Titanic were men and the majority of people that survived were female.

III. Survival rate and Age



The 'Age' required some transformations. Since the variable is numerical and no association can be made between numerical and categorical variables, the Age was converted to a categorical variable with four categories from Children, Youth, Adult and Senior. Furthermore, there were some missing values that were not considered for the analysis. In this case, the chi-square test resulted in a p-value of 0.0010709216002096906 , confirming that there is indeed a statistically significant relationship between age and survival rate. The graph to the left tells the

same story, with the majority of people that died and also that survived being adults. This may be because the majority of the people that were in the Titanic were adults.