When we decided on doing the Titanic for our project we looked into what we thought were the main variables of survival. The datasets we found were already broken up into training and testing for machine learning. These graphs were created using the combined data of both data sets with matplotlib and seaborn.

Embarking Location:

A majority of the passengers boarded in South Hampton, so I filled in the two missing values as such. Most of the class 3 passengers boarded in Queenstown.

This pie graph shows the percentage of survivors from each embarking location: Cherbourg, Queenstown and South Hampton.

The bar graph shows the number of passengers that survived or perished from each embarking location. (Survived is 1 and perished is 0)

Chart, histogram

Description automatically generatedChart, histogram

Description automatically generated

These histograms show survival and perishing counts by age.

Chart

Description automatically generated

This heatmap helps determine if the higher-class passengers (Class 1) had a higher rate of survival than the lower classes (Class 2 or 3). It implies that passenger class contributed to a passenger’s survival. (Survived is 1 and perished is 0)

Chart, bar chart

Description automatically generated

Survival rate of males is around 20%, whereas women is around 75%. Passenger gender played an important role in determining if one was going to survive (1) or perish (0).