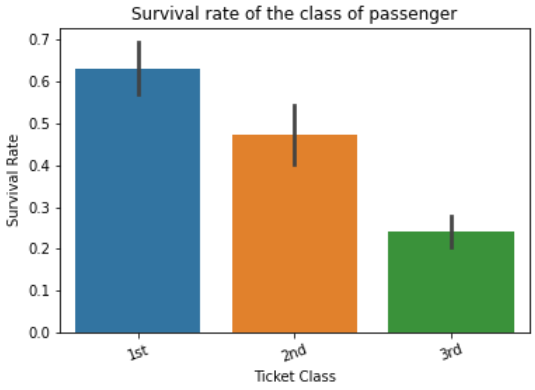
Hirokazu Nakajima
ID: 100379308
CPSC 4800 W01
Due date: 11 Mar, 2022

Assignment 3

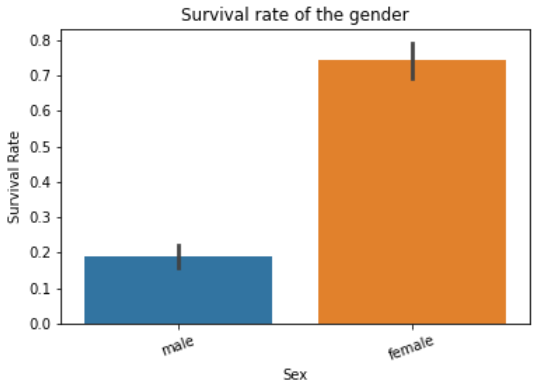
Association of survival rate and class of passenger

We calculated the survival rate from the survived column and made the bar chart with the Pclass column. According to the bar chart, it obviously seems that there is a relationship between the survival rate and the ticket class. To analyze more statistically, we computed chi-squared value and Cramer's coefficient of association, which resulted in 102.90 and 0.543 respectively. From the above result, there is a relatively strong correlation between them, which means people who purchased upper-grade tickets tended to be survived.



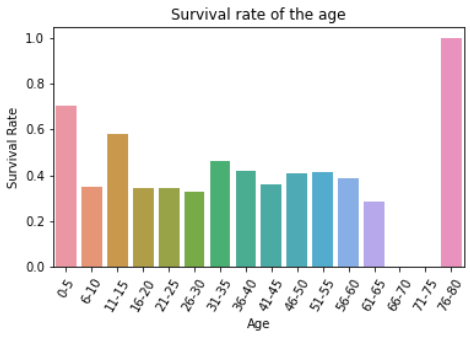
Association of survival rate and gender

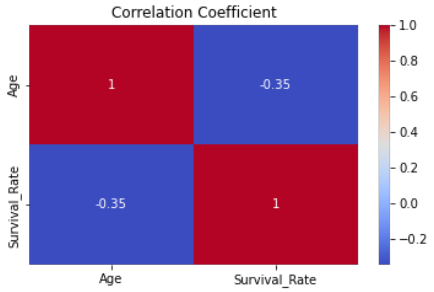
We investigated the association between survival rate and gender in the same manner as the survival rate and the passenger. Consequently, it seems the bar chart has a strong relation between them and the chi-squared value and Cramer's coefficient of the association are 263.05 and 0.54 respectively. The relation between survival rate and gender was more robust than between survival rate and passenger class and the female was more highly likely to survive than the male.



Association of survival rate and age

First of all, we converted the continuous variable on the "Age" column to a categorical variable as the "Age_bin" column, which has bins every 5 years old. According to the bar chart between the survival rate and age, the survival rate is inclined to decline with increasing age. Please note that the value on 76-80 bin was considered as an outlier because the bin contains only one value. For more analysis, we calculated the correlation coefficient, which was – 0.35 and found there is a slightly negative relationship between survival rate and age. Younger people tend to survive moderately compared to older people.





Summary

Relation with survival rate was strong in order of female, ticket class, and age. Young women who bought 1st class ticket tended to be most survived.

Github Repository

CPSC-Assignment-3 https://github.com/hnakaji/CPSC-Assignement-3