Summary of Assignment 3 Part 2

Exploratory data analysis using Titanic Dataset from Kaggle

Introduction:

For the analysis on Titanic Dataset, I have taken 3 hypothesis and by the knowledge of statistics I have analyzed. Hereby, I am submitting my findings below:

Steps:

Initially, the required libraries were imported in Jupyter notebook such as NumPy, Pandas and Matplotlib for statistics and plots. Then, titanic dataset was loaded, moving on the ".head()" was used to observe the first few rows of the dataset.

Summary to each hypothesis:

1. Passengers in higher class cabins had a better chance of survival, as they were given priority access.

   Conclusion: The above hypothesis is **TRUE** that the passengers in higher class cabins had a better chance of survival, as they were given priority access. Both by code and visualization it was proved. In the statistical part, it was observed that "Passenger class 1 had survival rate of 0.629630 which is comparatively higher to that of passenger class 3 which had survival rate of 0.242363." Similarly, in bar graph below, the bar of passenger class 1 was the highest, followed by passenger class 2 and class 3.

2. Women and children had a higher chance of survival compared to men, that was observed during the evacuation of the ship.

   Conclusion: The above hypothesis is **TRUE** that Women and children had a higher chance of survival compared to men, that was observed during the evacuation of the ship. In the statistical part, it was observed that "the survival rate of the women and children is 0.69 which is way higher than that of men." Similarly, from the pie chart it was observed that 68.8% of women and children passengers survived, on the other hand only 31.2% of men passengers survived. There is a negligible difference of 0.2% in code and plot, which can be ignored because both code and visualization parts justify that women and children had survived more than that of men.

3. Children had a higher survival rate than adults.

   Conclusion: The above hypothesis is **TRUE** children had a higher survival rate than adults. In the statistical part, it was observed that "the survival rate of the children was 0.5398230088495575 and the survival.

rate of adults was 0.36118251928020567". By these statistics it can clearly be observed that children had higher survival rate than adults. Similarly, from the bar graph visualization children survived more as compared to adults.