

7) Theory

Description of fields in the dataset.

- 1) Survival - Survival (0 = No & 1 = Yes)
- 2) Class - Passenger Class; 1st, 2nd, 3rd class.
- 3) name - Name of passenger
- 4) Sex - Gender of passenger.
- 5) age - Age of passenger.
- 6) sibsp - Number of sibling / spouse aboard.
- 7) parch - Number of parent / children aboard.
- 8) ticket - Ticket Number.
- 9) fare - Passenger fare.
- 10) Cabin - Cabin.
- 11) Embarked - Port of Embarkation (C = Cherbourg; Q = Queenstown; S = Southampton)
- 12) boat = lifeboat (if survived)
- 13) body = Body number (if did not survive and body was recovered).

Observations:

Imported pandas for processing the dataset after reading it into dataframe.

Also, Imported seaborn for data visualization and better data readability.

1) Using boxplot

(In descriptive statistics, a box plot or boxplot is a method for graphically demonstrating the locality, spread and skewness groups of numerical data through their quartiles. In addition to the box on a boxplot, there can be lines (which are called whiskers) extending from the box indicating variability outside the upper & lower quartiles, thus the plot is also termed as the box-and-whisker plot & the box-and-whisker diagram)

We can see that, Among the total passenger base the male who survived were from age group 20 to 40 with few out of the range and above 65. Some of the male who didn't survive were from the age group between 15 to 40.

On the other hand, among the total female passenger who survived were from the age group of 15 to 35, whereas the female passenger who didn't survive were from the age group between 20 to 40.

2) Using listplot. we can see that among total passenger the male who survived were from age group 10 to 65 and who didn't were ranging from 65 to 75.

On the other hand, female who survived were ranging from 10 to 65 among a few who didn't survive were having age around 10.

Conclusion :

studied the titanic dataset and identified relevant patterns.