A BRIEF ANALYSIS OF THE TITANIC TRAGEDY

The data was imported from the internet in the form of a spreadsheet document, which has been attached along with the code and this analysis. Note that this is my first project on data analysis, hence the power of Machine Learning has not been used, since I'm still on its learning stage; however my upcoming project uploads may involve a lot of statistics and machine learning.

In this project I use the power of:

1. Python
2. Numpy, pandas , and seaborn libraries of python
3. Analytics and graphs.

Thinking back to the disaster and observing the given data, following questions come to mind:

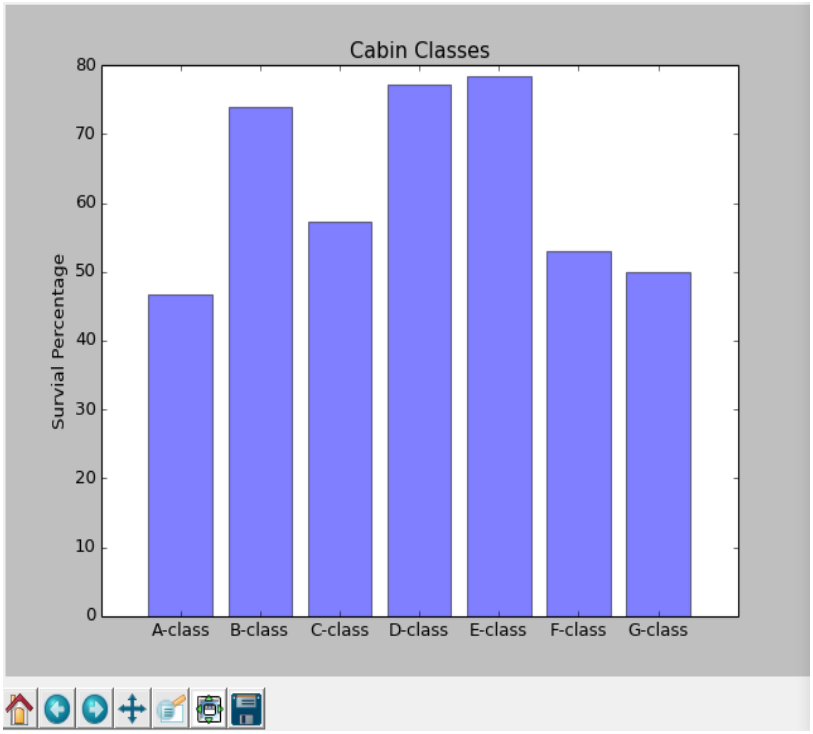
1. What percentage of the population was economically well off, and how likely were they to die as compared to lower wage earners?
2. Which gender was more likely to survive?
3. Which cabin numbers had higher percentage of survival, i.e., which floors were safer compared to others?
4. Which letter cabins were occupied by the higher class, which ones by the middle class, and which ones by the lower class.
5. Looking at the movie adaptation, how likely were women with children/spouse to survive compared to women without children/spouse?

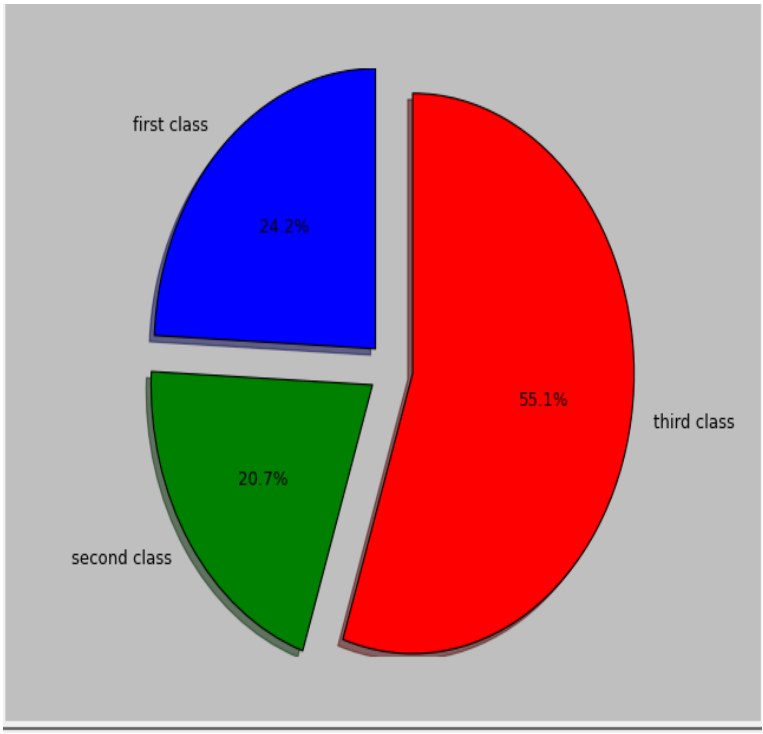
6. How many children were present(<18 years), had nannies/parents with them, and what percentage survived, also which age group of children survived with the highest percentage?
7. Percentage of survival of adults vs Percentage of survival of children(<18)

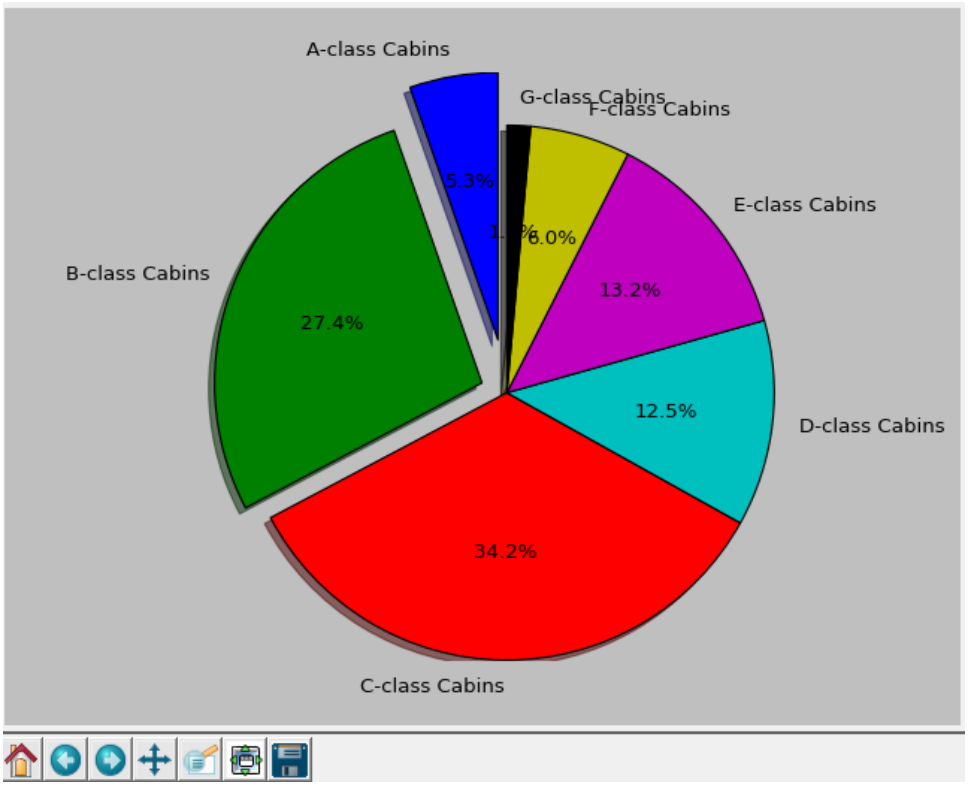
I also tried to draw graphs to represent data wherever possible.

To answer the questions mentioned, my findings were :

➔ 549 people survived out of 891 people
➔ 62.962962963% of rich people survived
➔ 24.2362525458% of poor people survived
➔ 18.8908145581% of all males survived(out of 577 males)
➔ 74.2038216561% of all females survived(out of 314 females)







➔ total women with spouses/siblings/children : 64
➔ total women without spouses/children/siblings : 129
➔ total women with spouses/siblings/children who survived 48, total women without spouses/siblings/children who survived : 103
➔ Survival percentage of women with spouses/siblings/children : 75.0%, Survival percentage of women without spouses/siblings/children : 79.8449612403%

this observation shows that women as a whole were equally likely to survive regardless of whether they had family or not. doing the same observation for men...

➔ total men with spouses/siblings/children : 40
➔ total men without spouses/children/siblings : 342
➔ total men with spouses/siblings/children who survived 6, total men without spouses/siblings/children who survived : 63
➔ Survival percentage of men with spouses/siblings/children : 15.0%, Survival percentage of men without spouses/siblings/children : 18.4210526316%

this observation shows that men as a whole were equally likely to survive regardless of whether they had family or not
And also that women were more likely to survive than men in the given relational states.

> children less than the age of 3: 24
> children less than the age of 6: 20
> children less than the age of 9: 10
> children less than the age of 12: 14
> children less than the age of 15: 10
> children less than the age of 18: 35

> Survived children less than the age of 3: 15
> Survived children less than the age of 6: 16
> Survived children less than the age of 9: 5
> Survived children less than the age of 12: 3
> Survived children less than the age of 15: 6
> Survived children less than the age of 18: 16

➔ total children : 113
➔ total survived children : 61
➔ Percentage of survived children : 53.98%.

➔ total adults : 601
➔ total survived adults : 229
➔ percentage of survived adults : 38.10%

Links :

1. kaggle for titanic data : https://www.kaggle.com/c/titanic/data
2. Stack overflow
3. Udacity for course completion and project :

https://classroom.udacity.com/courses/ud170/lessons/5442503234/concepts/54766487950923