Titanic Data Analysis

Lalitha Pragada



Requirements for submission:

A PDF or HTML file containing your analysis. This file should include:

* A note specifying which dataset you analyzed
* A statement of the question(s) you posed
* A description of what you did to investigate those questions
* Documentation of any data wrangling you did
* Summary statistics and plots communicating your final results

 If the code you used to perform your analysis is not included in the above, you can attach the code separately in .py file(s).

 A list of Web sites, books, forums, blog posts, github repositories, etc. that you referred to or used in creating your submission (add N/A if you did not use any such resources).

# Data Set Chosen:

* I picked to analyze the **titanic** data to find out some facts about the tragedy.

Questions:

* How many were survived from the given data set?
* How many were male and female?
* Was male population survived more or female?
* What was the average of people travelling and what was the average age of the survivors?
* Were the families (with siblings/spouses or parents/children) or elders or kids tried to be saved more?
* What was the average traveler size in each Class and was there any significance to see if a specific class people were most of the survivors?

## Data Wrangling:

* The total number of passengers from the given set was first collected.
* Then I looked at the total number of passengers who were survived.
* From them, I got the total number of male passengers and there by the total female.
* Also, I tried to get the number of passengers in each Class.
* Then the no. of passengers who had some sort of family travelling with them. I also looked at them per Class.
* I calculated the average survivor’s age and the age of the maximum survivors.

--- All of this analysis is included in the ipython – HTML document.

## Conclusions:

1. I find that more female passengers were saved than male. (more than half of the survivors were female.)

* 233 were female and 109 were male survivors.

1. There was no significance about the Class they were travelling in.

Pclass Survived

1 136

2 87

3 119

1. I also find that single passengers were among the most survivors. Passengers travelling with kids, parents or siblings or spouses were least of the total survivors.
2. Average survivor’s age is 28.3436896552.

It also does not show that elders or kids were among the most survivors.

**References:**

* All the videos from my Udacity – Intro to Data Analysis Course ( Mainly and importantly)
* <https://www.kaggle.com/c/titanic/data>