DATA SCIENCE TASK

Analysis and Insights From Titanic Dataset

By Abhinav Singh

Index

Introduction

The first task was to analyze the dataset of passengers on the titanic with the following details given in the data set:

- Passenger ID
- Whether passenger survived or not
- Ticket class
- Name of the passenger
- Gender of passenger
- Passenger age
- Number of siblings/spouses aboard
- Number of parents/children aboard
- Ticket number
- amount respective passenger paid for ticket
- Cabin number of the passenger
- Port name from which Passenger boarded the Titanic

  Which were Cherbourg, Queenstown, Southampton

Ordered and Color-coded raw data

Before diving into the graphical representations, the raw data should be converted into a much more understandable format segregated into **Male and Female, Adult and Children, Ticket class from 1st to 3rd And the boarding ports going along Cherbourg, Queenstown and Lastly Southampton.**

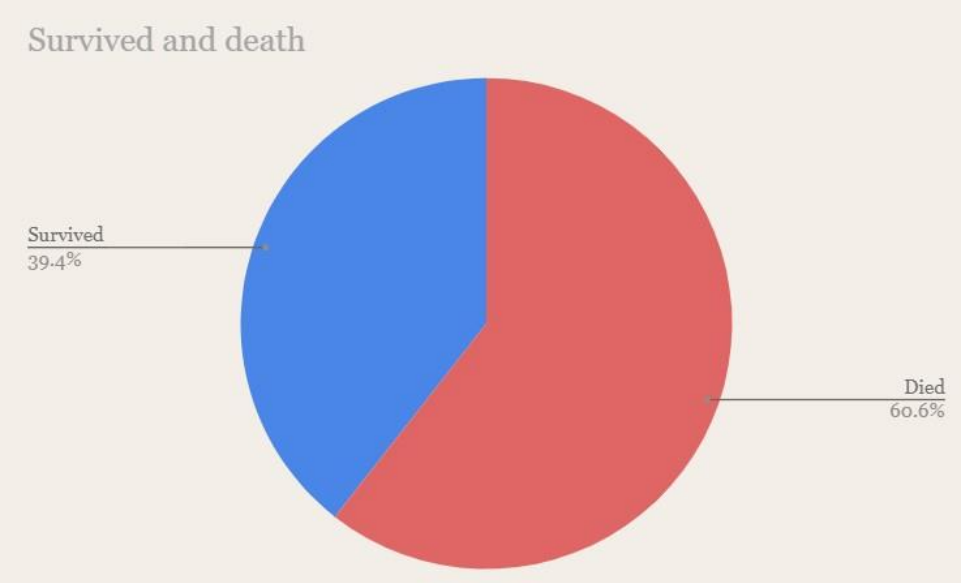
Every data is color coded to make it easier to understand the data without having to read every single value.

ALL OF THESE COLORS WILL BE USED FOR PIE CHARTS AND GRAPHS						
C-Dark yellow 1	Q - Orange	S - Blue				
Non Adult- Pink	Adult - lighter gray 2					
Survived - Cornflower blue	Dead - Light Red 1					
Class 1 - Purple	Class 2 - Light Green	Class 3 - Light magenta 2				
Fare 1 to 20.999 = Light Cyan 1	Fare 21 to 40.999 = Light Orange	Fare 41 to 60.999 = Light Magenta 1	Fare above 60 = Light Cyan 3			
sibling/spouses0 - light redberry 3	sibling/spouses 1 - dark orange 1	sibling/spouses 2 - light red 2	sibling/spouses 3 - light orange 3	sibling/spouses 4 - light red berry 2		
parch 0- dark gray 2	parch 1 - dark gray 1	parch 2 - gray	parch 3 - light green 1	parch 4 - light green 2	parch 5-dark green 1	
Male - light cornflower blue 1	Female - light yellow 1					

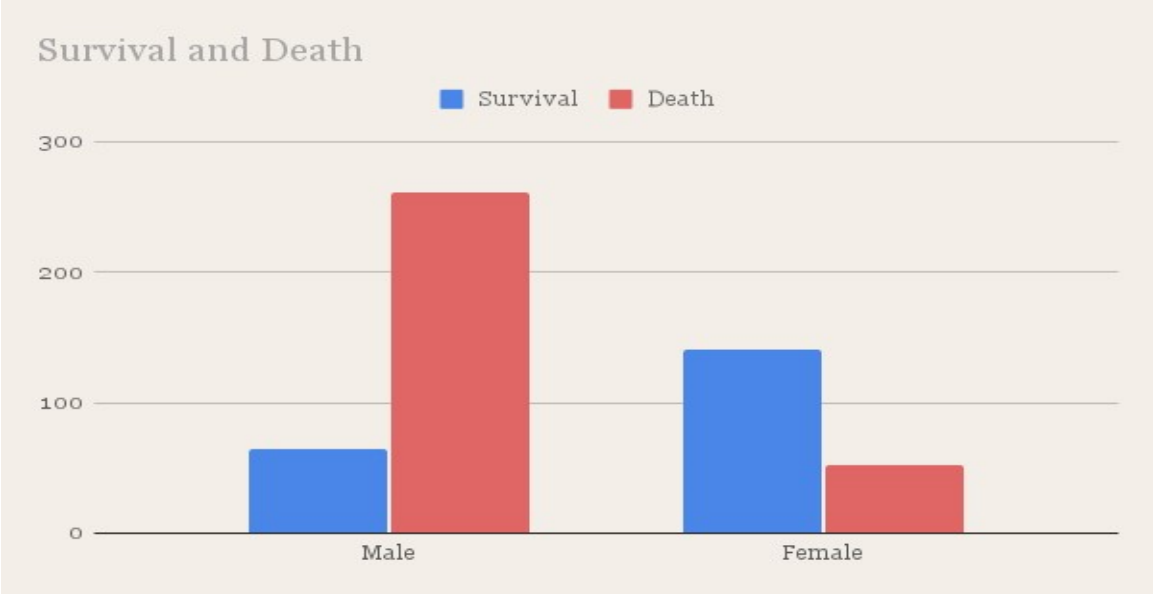
*The google sheets pdf with the ordered and colored data will be in the .zip file.

Data Regarding Survival

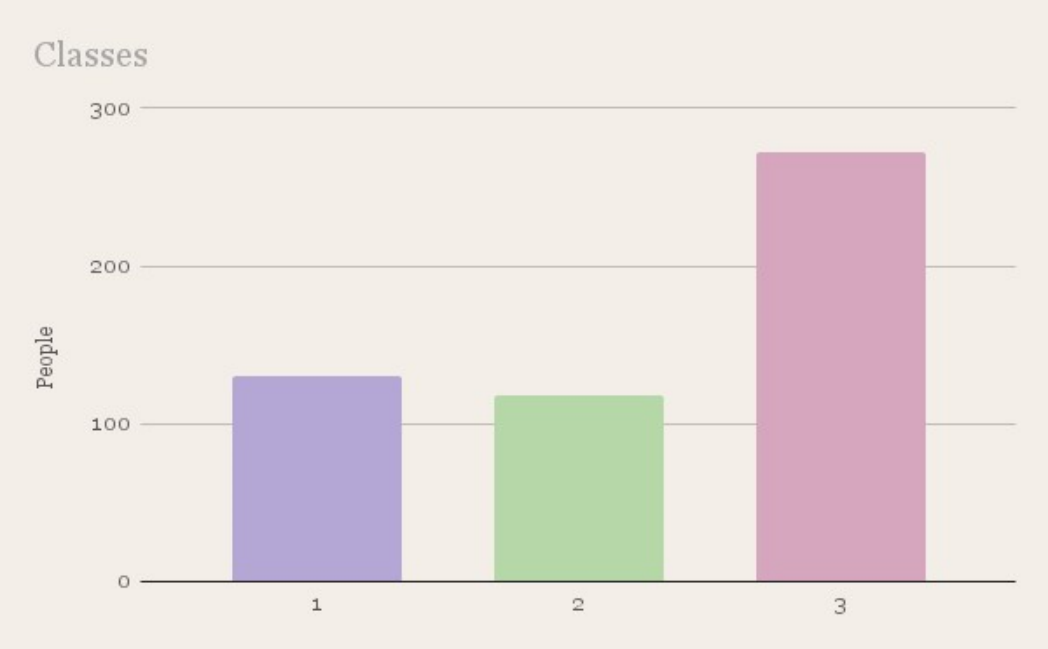
1.Total Survival – Pie chart of the percentage



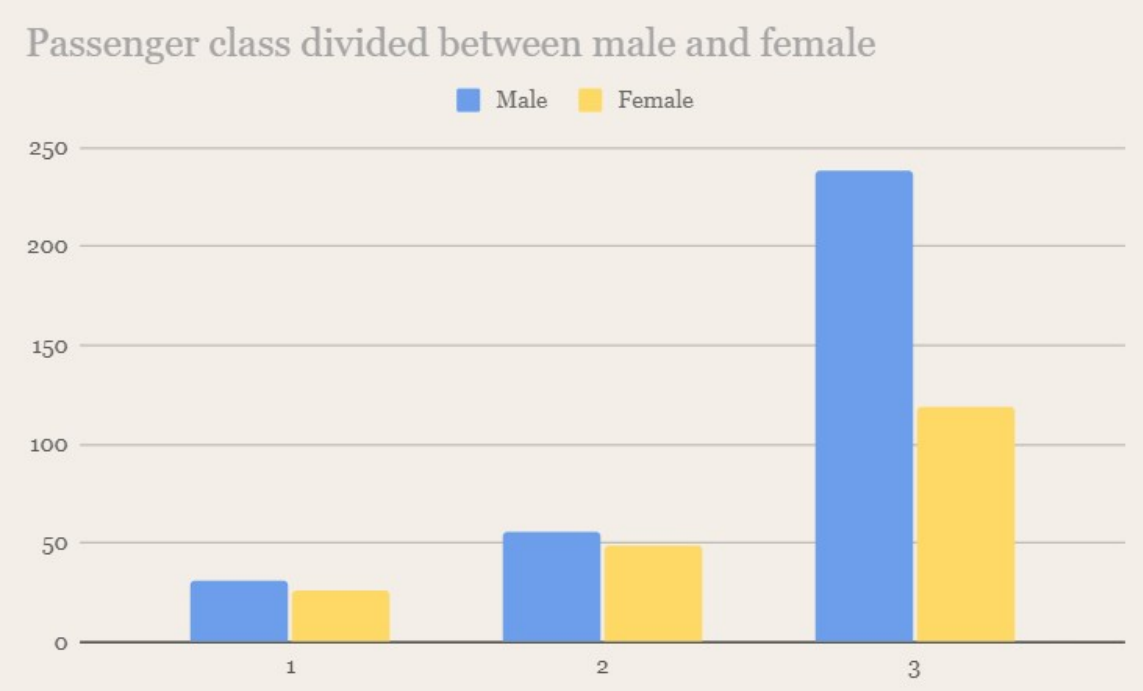
2.Male and Female Survival and Death Percentage in form of bar graph



3.Passenger Classes Bar Chart



4.Passenger Classes Bar Chart between genders



5.Mean, Median of age and youngest to oldest

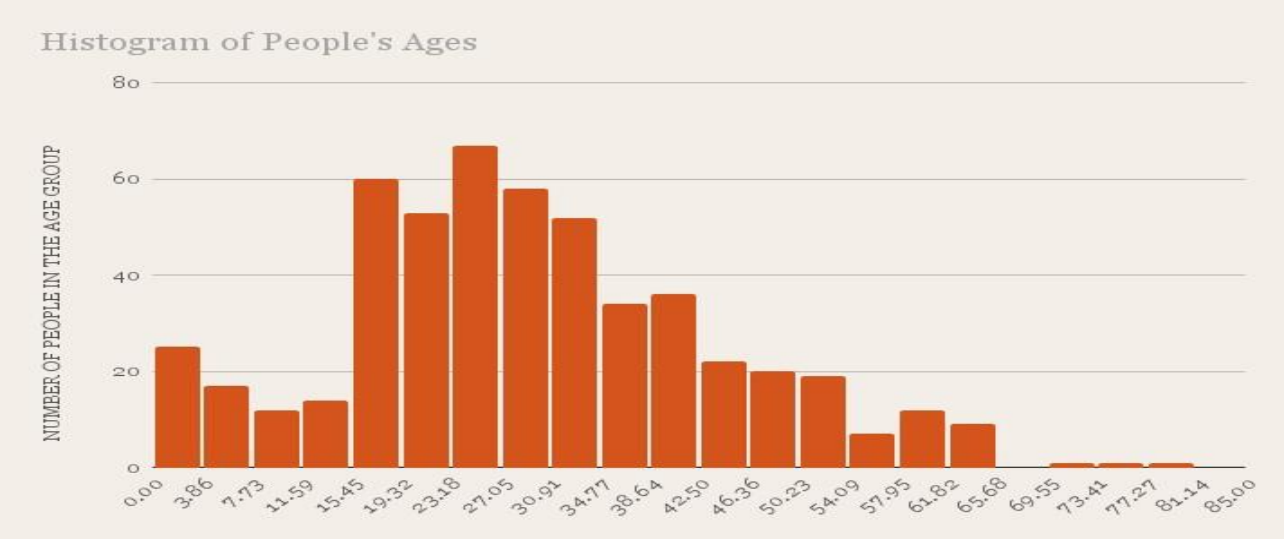
Mean: 29.34

Median: 28

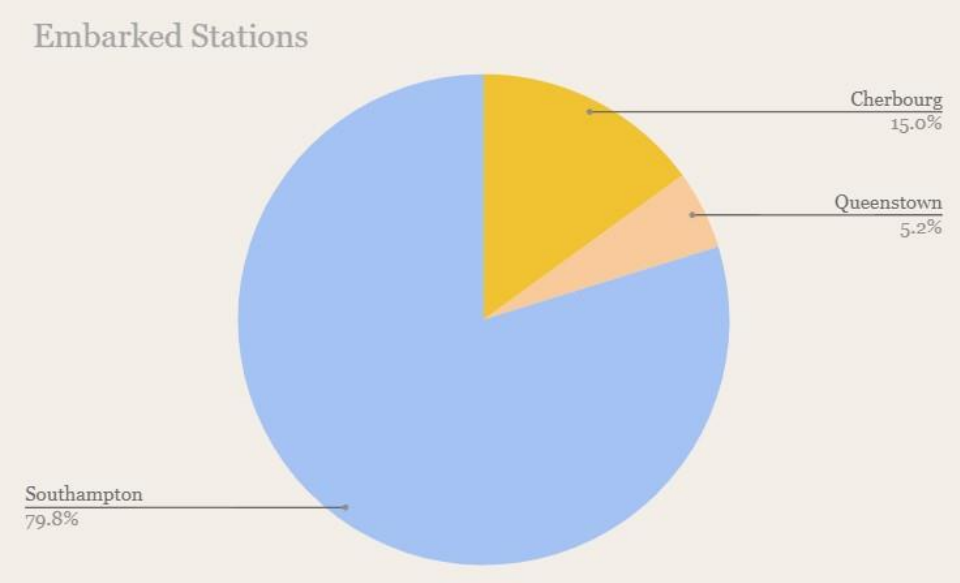
Youngest: 0.42, Youngest surviving passenger: 2 females of 0.75 age

Oldest:80, Oldest surviving passenger: 1 male of 74 age

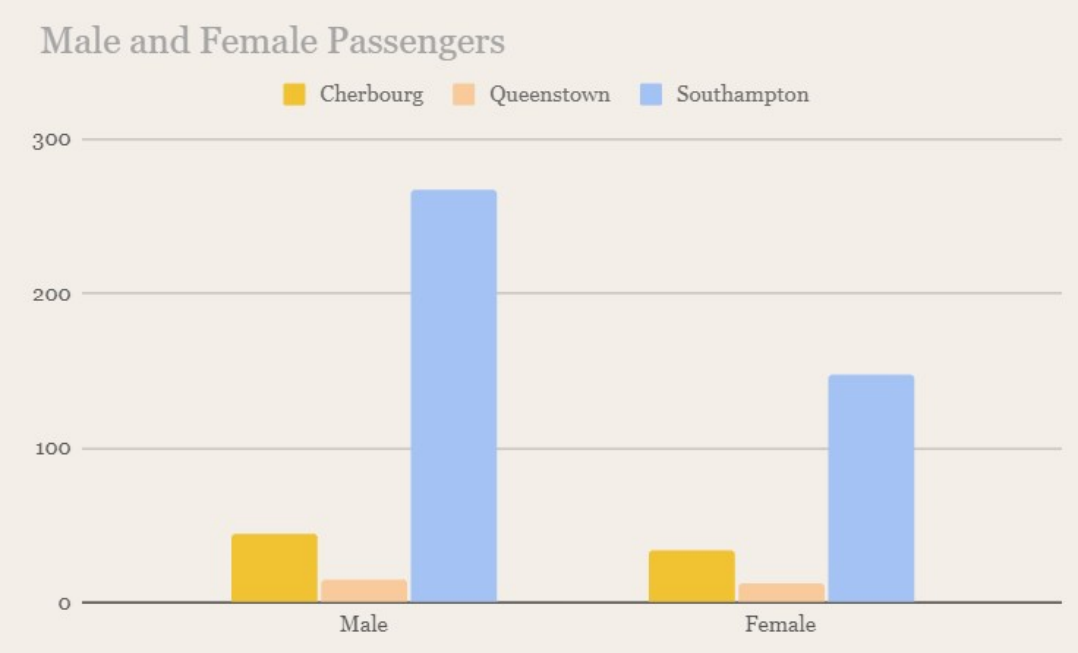
6. Histogram passengers' ages



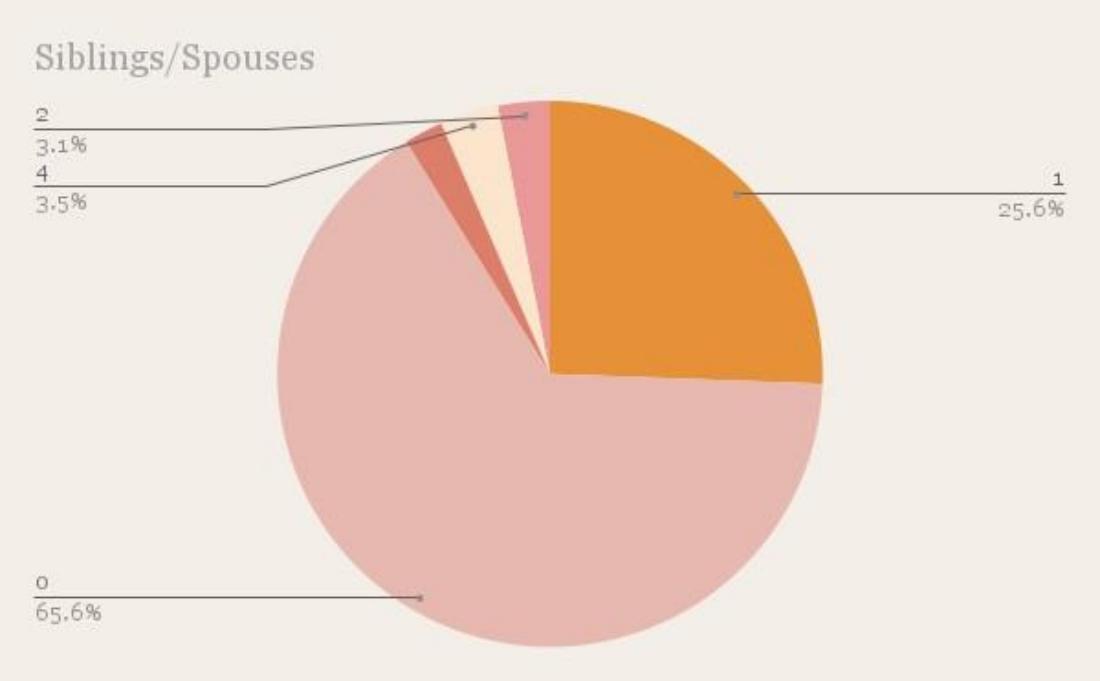
7.Percentage of People boarding from various ports using a pie chart



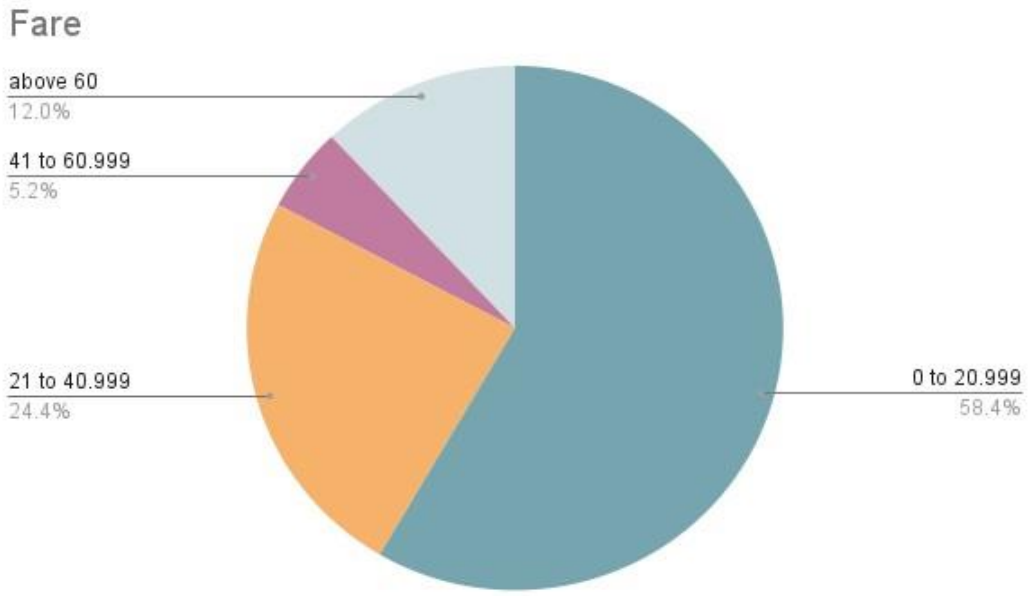
8.Bar graph comparison of male and female passengers from each port



10. Sibling and spouses distribution in form of pie chart



10.Fare division



References

- www.kaggle.com/code/omarmohamedyehia/titanic-databasevisualization
- towardsdatascience.com
- https://my-learning.w3schools.com/

*The MS word file wasn't properly converted into the pdf file, apologies for the weird formatting in the pdf.