TITANIC PRESENTATION

: DATA INSIGHTS

INTRODUCTION

OBJECTIVE:

To gain insights about the data regarding the proportion of passengers that survived aboard Titanic based on their age, class, and sex.

Some questions to answer:

Proportion of passengers that survived based on age

Proportion of passengers that survived based on class

Proportion of passengers that survived based on sex

Proportion of passengers that survived based on the combination of two of the attributes

Proportion of passengers that survived based on the combination of the three attributes

DATASET

THE DATASET CONTAINS DATA OF 887 REAL PASSENGERS ABOARD THE TITANIC. IT ALSO CONTAINS 8 ATTRIBUTES.

THE DATASET IS AVAILABLE AT: https://web.stanford.edu/class/archive/cs/cs109 /cs109.1166/problem12.html

- ATTRIBUTES:
  - Survived: 0 = Died, 1 = Survived
  - Pclass: 1= class 1, 2 = class 2, 3 = class 3
  - Name: complete name of the passenger
  - Sex: gender of the passenger (male, female)
  - Age: age of the passenger
  - Siblings.Spouses.Aboard: number of companions that the passenger has that is considered any one of these
  - Parents.Children.Aboard: number of children and parents accompanying the passenger
  - Fare: how much the passenger paid to board Titanic.

LIBRARIES:

THE LIBRARIES THAT WAS USED IN THE ANALYSIS

- library(ggplot2): for data visualization
- library(dplyr): for data manipulation
- library(DT): provides an R interface to the JavaScript library

DATA SUMMARY

```
    Survived          Pclass                                           Name
Min.    :0.0000   Min.    :1.000   Capt. Edward Gifford Crosby      :  1
1st Qu.:0.0000    1st Qu.:2.000    Col. John Weir                   :  1
Median :0.0000    Median :3.000    Col. Oberst Alfons Simonius-Blumer:  1
Mean    :0.3856   Mean    :2.306   Don. Manuel E Uruchurtu          :  1
3rd Qu.:1.0000    3rd Qu.:3.000    Dr. Alfred Pain                  :  1
Max.    :1.0000   Max.    :3.000   Dr. Alice (Farnham) Leader       :  1
                                   (Other)                          :881
    Sex              Age          Siblings.Spouses.Aboard Parents.Children.Aboard
female:314    Min.    : 0.42    Min.    :0.0000           Min.    :0.0000
male  :573    1st Qu.:20.25     1st Qu.:0.0000            1st Qu.:0.0000
              Median :28.00     Median :0.0000            Median :0.0000
              Mean    :29.47    Mean    :0.5254           Mean    :0.3833
              3rd Qu.:38.00     3rd Qu.:1.0000            3rd Qu.:0.0000
              Max.    :80.00    Max.    :8.0000           Max.    :6.0000


    Fare
Min.    :  0.000
1st Qu.:  7.925
Median : 14.454
Mean    : 32.305
3rd Qu.: 31.137
Max.    :512.329
```
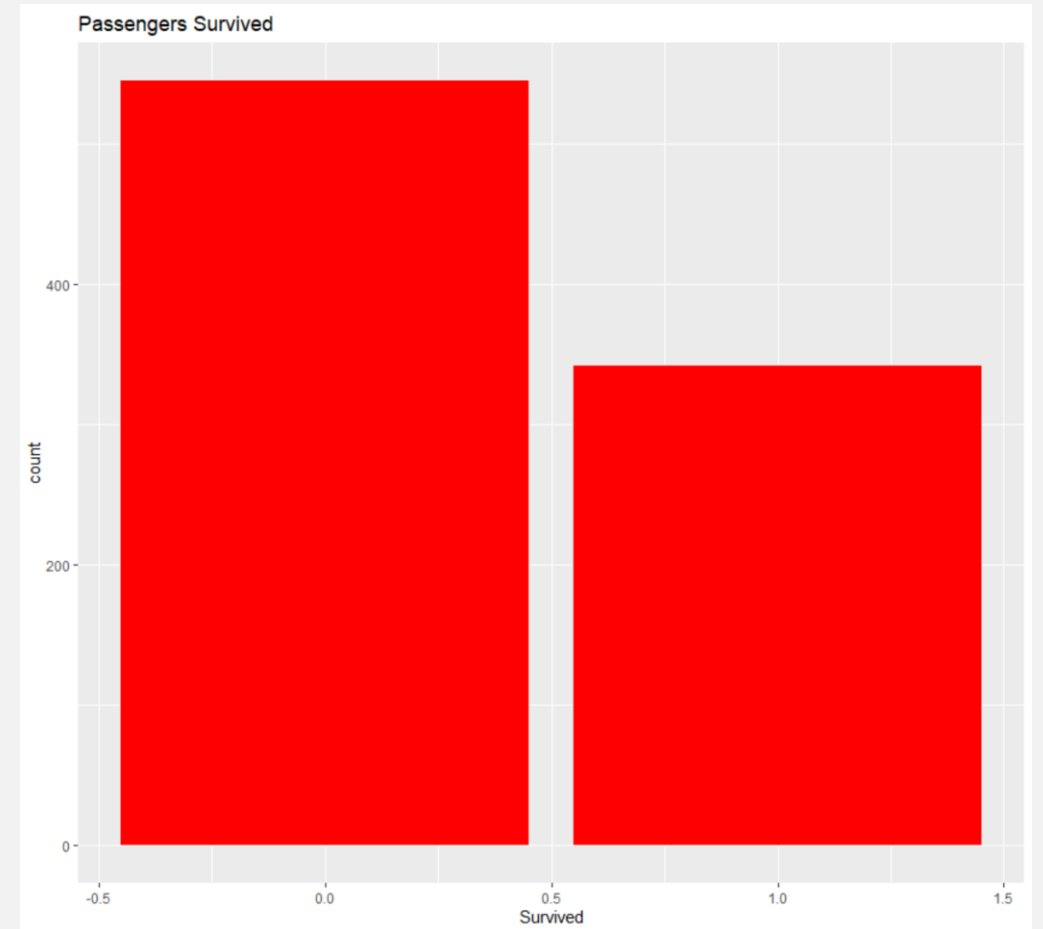
BAR PLOT OF "SURVIVED" ATTRIBUTE

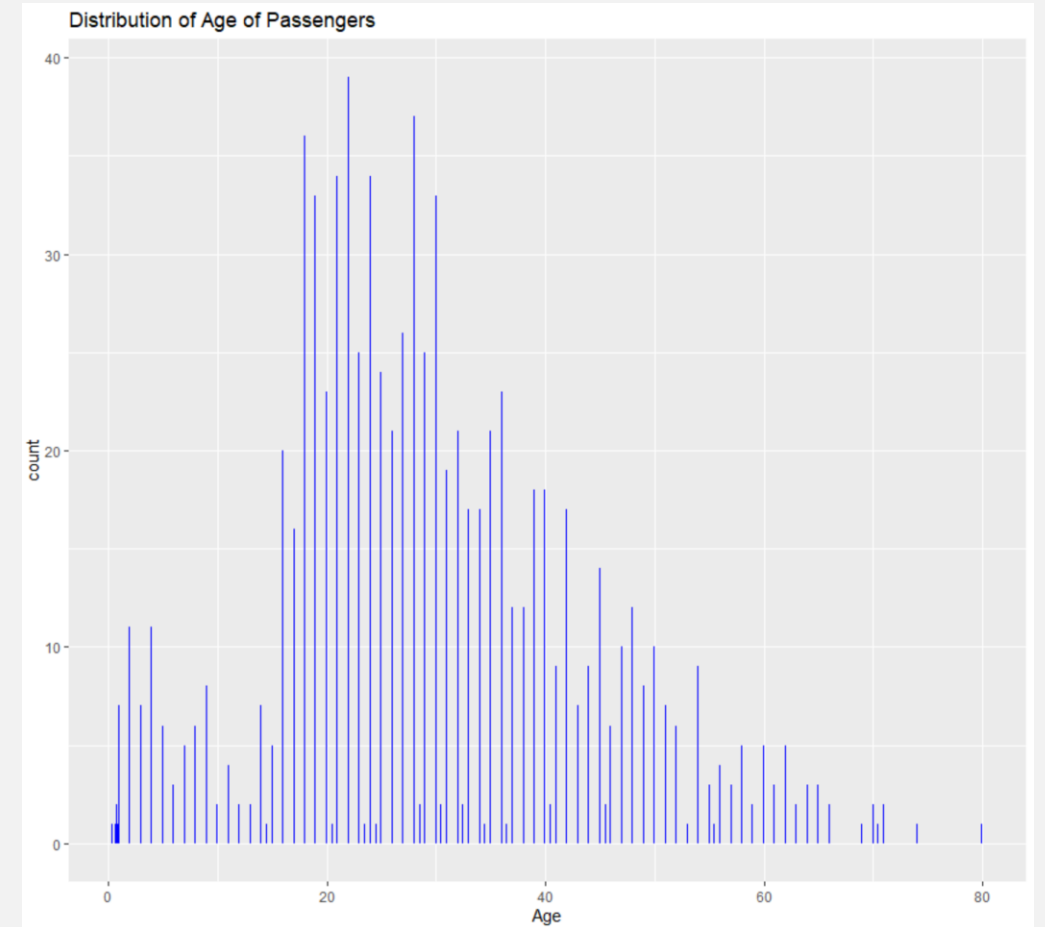FROM THE PLOT, THERE ARE MORE PASSENGERS THAT DID NOT SURVIVE WITH 545 PEOPLE COMPROMISING THIS NUMBER.

ONLY 342 PASSENGERS SURVIVED THE SINKING.

THIS COMPRISE ONLY 38.6% OF THE TOTAL PASSENGERS IN THE DATASET.

Passengers Survived

BAR PLOT OF "AGE"

MOST OF THE PASSENGERS IN THE DATASET ARE IN THEIR LATE TEENS TO EARLY 30s.


Distribution of Age of Passengers

BAR PLOT OF "PASSENGER CLASS"

OF THE 887 PASSENGERS, MOST OCCUPIED THE THIRD CLASS. IN FACT, THERE ARE 487 PASSENGERS, OR AROUND 55%, OF THE OVERALL NUMBER.

NEXT WOULD BE THE FIRST CLASS. THIS COMPRISE 216 PEOPLE, OR 24% OF THE TOTAL NUMBER.

THE SECOND CLASS COMPRISE OF THE SMALLEST GROUP OF THE THREE. THIS ONLY HAS 184 PASSENGERS, OR 21%.

Distribution of Class
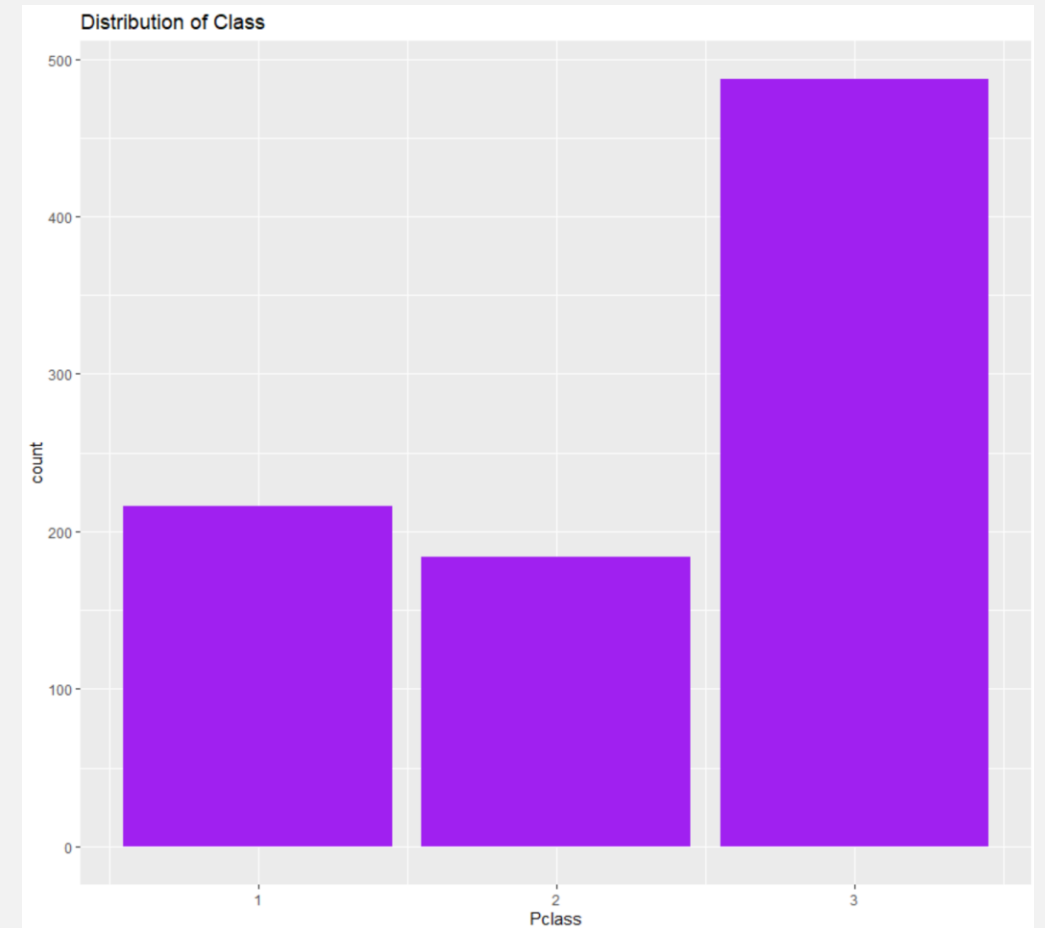
BAR PLOT OF THE "SEX" ATTRIBUTE

THERE ARE 314 FEMALE PASSENGERS ABOARD THE SHIP. THIS IS AROUND 35% OF THE 887 TOTAL.

THE MALE PASSENGERS COMPRISED 65% OF THE TOTAL NUMBER.
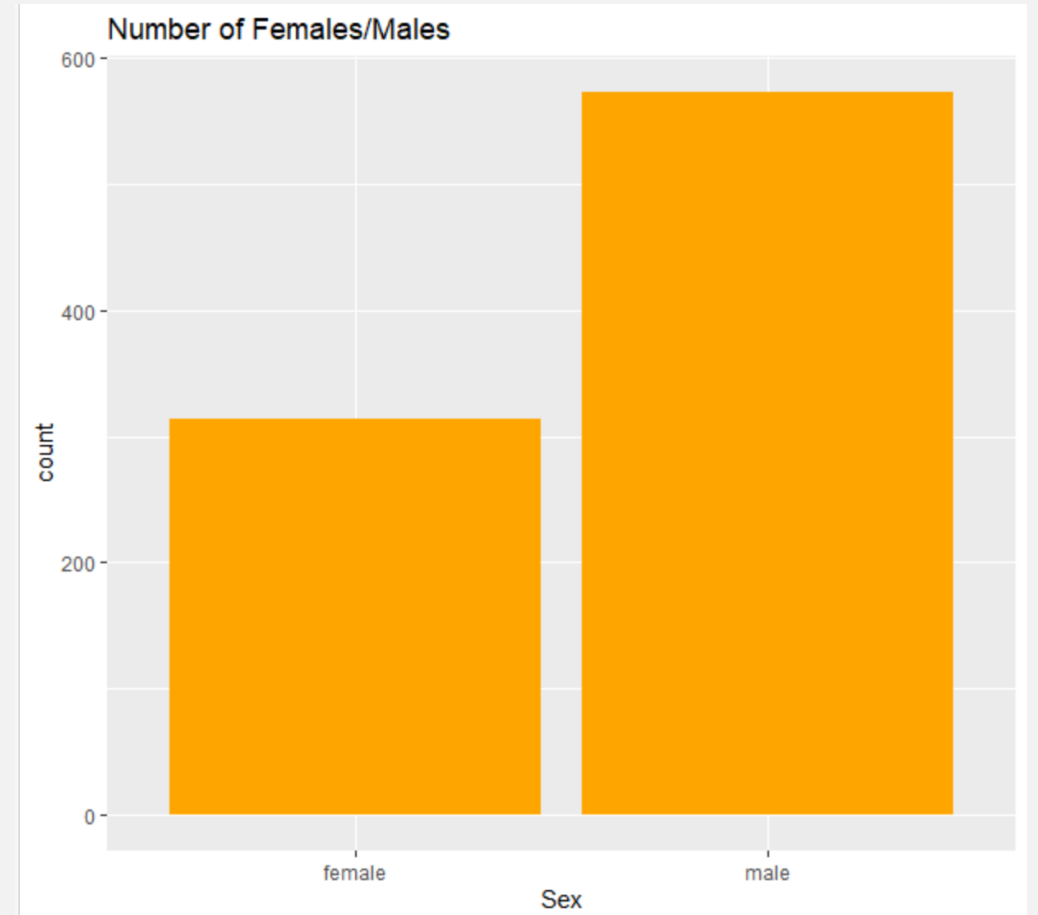
PROPORTION OF SURVIVORS BASED ON CLASS, SEX, AND AGE:

The <u>first graph</u> show the proportion of passengers that survived based on the class tickets. The results showed that 15.3% of the first class, 9.8% of the second class, and 13.4% of the third class passengers survived.

	Pclass	Survived	number	perc
1	1	0	80	9.01916572717024
2	1	1	136	15.3325817361894
3	2	0	97	10.9357384441939
4	2	1	87	9.80834272829763
5	3	0	368	41.4881623449831
6	3	1	119	13.4160090191657

The <u>second graph</u> show the proportion of passengers that survived based on sex. 26.3% of the female and 12.3% of the male passengers survived. Female passengers had the highest survival rate.

	Sex	Survived	number	perc
1	female	0	81	9.13190529875986
2	female	1	233	26.2683201803833
3	male	0	464	52.3111612175874
4	male	1	109	12.2886133032694

The <u>third graph</u> show the proportion of passengers that survived based on age. For this, I changed the variable. Any passenger whose age was greater than or equal to 18 was given 1, and any passengers less than 18 is 0. Only 8.8% of the minors survived while 29.8% of the adults did.

	age	Survived	number	perc
1	0	0	88	9.92108229988726
2	0	1	78	8.79368658399098
3	1	0	457	51.52198421646
4	1	1	264	29.7632468996618

THE PROPORTION OF PASSENGERS SURVIVING FOR EACH AGE/SEX CATEGORY (I.E., FOR ADULT MALES, CHILD MALES, ADULT FEMALES, CHILD FEMALES).

These were the following survival rates for men, women, boys, and girls:

Men = 9.2%

Women = 20.5%

Boys = 3.0%

Girls = 5.7%

Women were more likely to survive, and boys were the least likely. Though the rates for the boys and girls can be attributed to their small numbers in comparison to the adults.

	age	Sex	Survived	number	perc
1	0	female	0	29	3.26944757609921
2	0	female	1	51	5.74971815107103
3	0	male	0	59	6.65163472378805
4	0	male	1	27	3.0439684329195
5	1	female	0	52	5.86245772266065
6	1	female	1	182	20.5186020293123
7	1	male	0	405	45.6595264937993
8	1	male	1	82	9.2446487034949

THE PROPORTION OF PASSENGERS SURVIVING FOR EACH AGE/SEX/CLASS CATEGORY. WHICH GROUP HAD THE HIGHEST MORTALITY IN THIS DISASTER.

The <u>first graph</u> shows the survival rate of the first class passengers:

Men = 18.9%

Women = 37.5%

Boys = 1.9%

Girls = 4.6%

The <u>second graph</u> shows the survival rate of the second class passengers:

Men = 4.3%

Women = 30.4%

Boys = 4.9%

Girls = 7.6%

	age	Sex	Survived	number	perc
1	0	female	0	1	0.462962962962963
2	0	female	1	10	4.62962962962963
3	0	male	0	1	0.462962962962963
4	0	male	1	4	1.85185185185185
5	1	female	0	2	0.925925925925926
6	1	female	1	81	37.5
7	1	male	0	76	35.1851851851852
8	1	male	1	41	18.9814814814815

	age	Sex	Survived	number	perc
1	0	female	1	14	7.60869565217391
2	0	male	0	6	3.26086956521739
3	0	male	1	9	4.89130434782609
4	1	female	0	6	3.26086956521739
5	1	female	1	56	30.434826086957
6	1	male	0	85	46.195652173913
7	1	male	1	8	4.34782608695652

The <u>third graph</u> shows the passenger survival rate of the third class passengers:

Men = 6.8%

Women = 9.2%

Boys = 2.9%

Girls = 5.5%

Looking at the survival rate of each class, boys or men have the highest mortality by gender. In fact of the three classes, the second class has the highest mortality for men. To better understand why second class men have low survival rate, we look at the raw numbers of men and women in each class.

	age	Sex	Survived	number	perc
1	0	female	0	28	5.74948665297741
2	0	female	1	27	5.54414784394251
3	0	male	0	52	10.6776180698152
4	0	male	1	14	2.87474332648871
5	1	female	0	44	9.03490759753593
6	1	female	1	45	9.24024640657084
7	1	male	0	244	50.1026694045174
8	1	male	1	33	6.7761806981 5195

REASON WHY SECOND MALE PASSENGERS HAVE HIGH MORTALITY

The **first graph** shows that second class male passengers have the lowest number of the three classes.

The bar plot illustrates that of the three classes, the second class passengers comprise of the lowest number of passengers.

	Pclass	Sex	number	perc
1	1	female	83	11.5117891816921
2	1	male	117	16.2274618585298
3	2	female	62	8.59916782246879
4	2	male	93	12.8987517337032
5	3	female	89	12.3439667128988
6	3	male	277	38.4188626907073

Adult Female and Male Passengers by Class
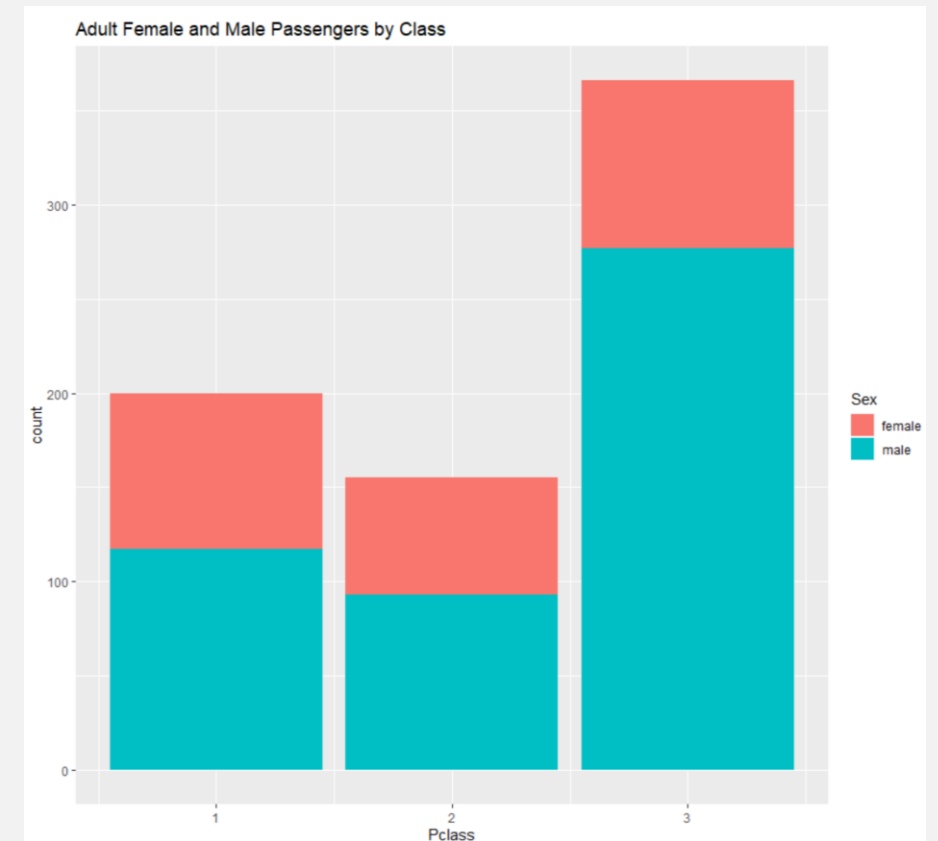
WHY THE PERCENTAGE OF MEN IN THIRD CLASS WHO SURVIVED WAS HIGHER THAN IN THE SECOND CLASS?

Looking at the stacked box, illustrates the comparative survival rates in each class.

From this, it was easier to see that there were the fewest adult male passengers overall in second class. While the men in second class had the lowest survival rate, the raw number of second class adult male passenger deaths was still very small compared to the overall number of men lost in the other classes.