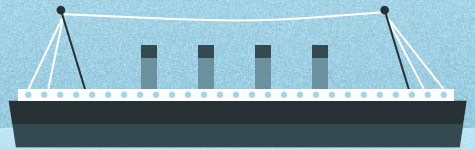
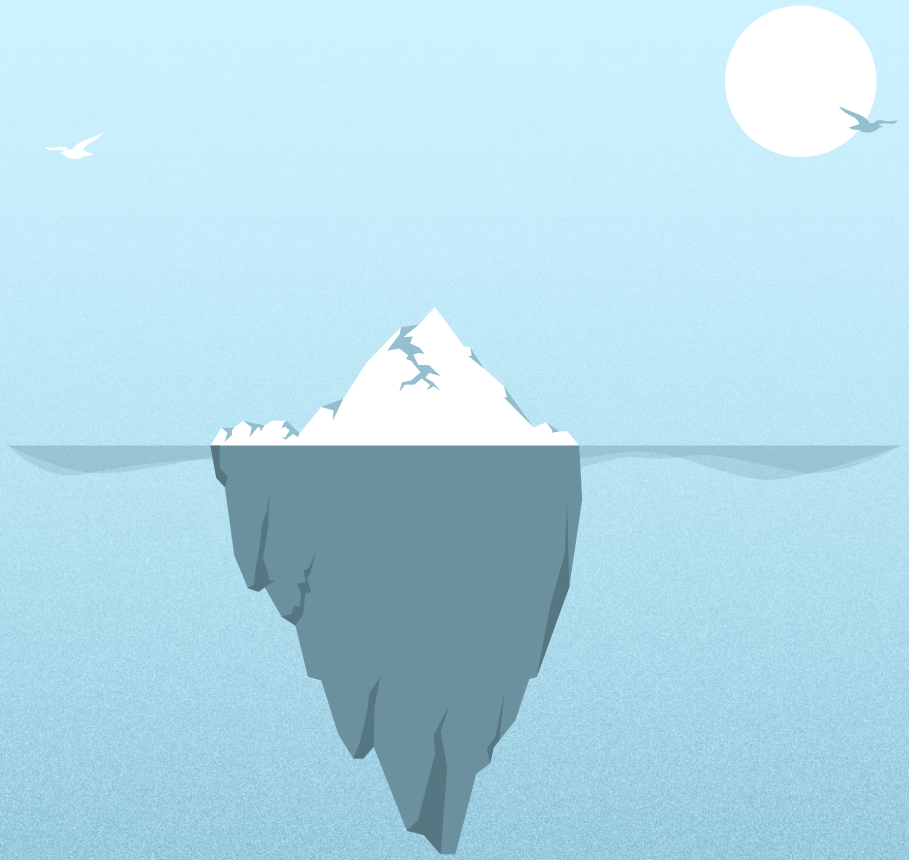
— Data Disaster Case Competition —

UNSW FINTECH SOCIETY



Presented by Priya Nandyal, Dharani Palanisamy & Faiyam Islam



1. Research Questions

Address questions related to our data analysis

2. Data Visualisations

Using Tableau, we explored the dataset and created useful visualisations

3. Conclusion

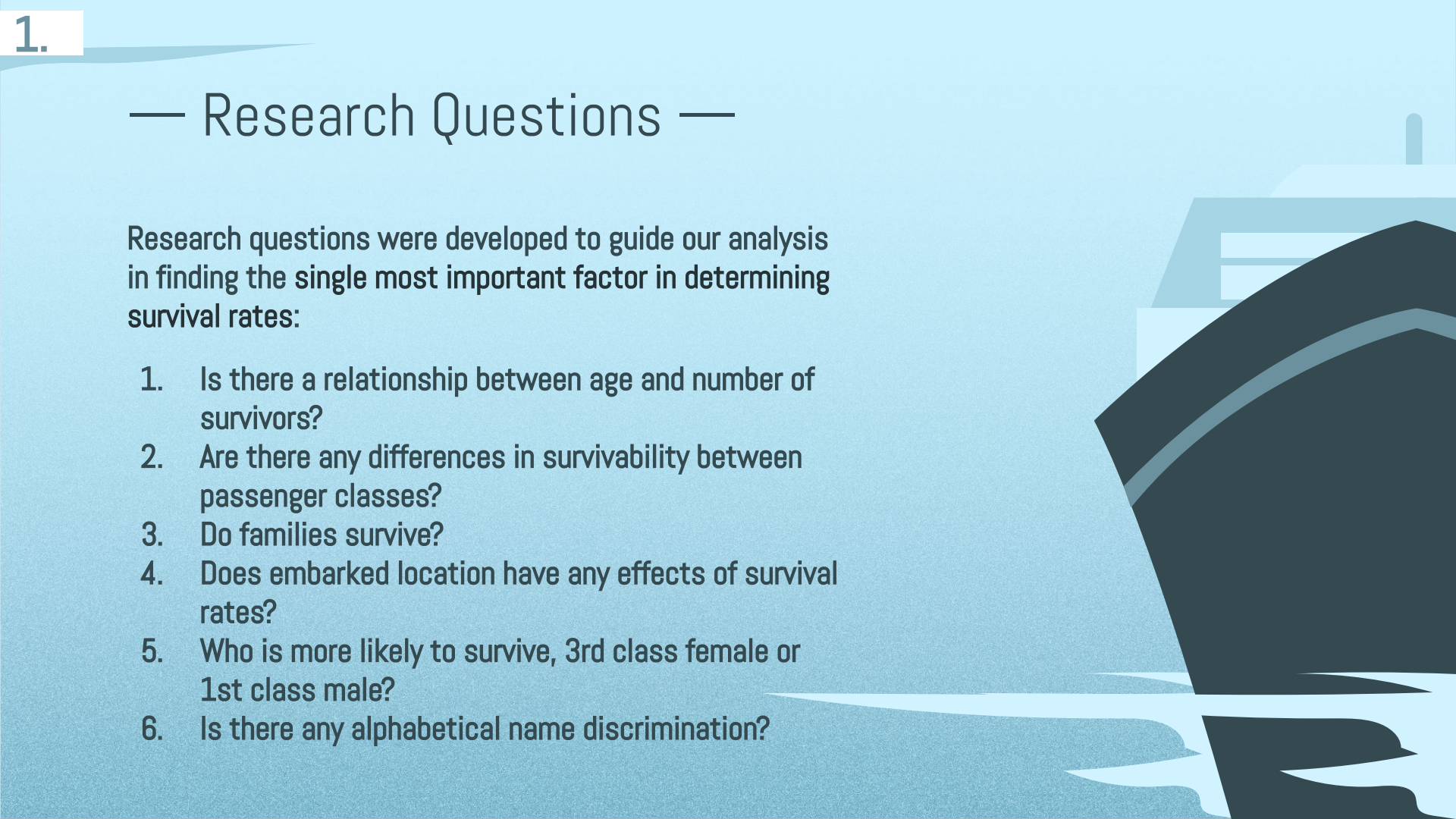
Conclusions obtained from our analysis using Tableau

1.

— Research Questions —

Research questions were developed to guide our analysis in finding the **single most important factor in determining survival rates:**

1. Is there a relationship between age and number of survivors?
2. Are there any differences in survivability between passenger classes?
3. Do families survive?
4. Does embarked location have any effects of survival rates?
5. Who is more likely to survive, 3rd class female or 1st class male?
6. Is there any alphabetical name discrimination?



2.

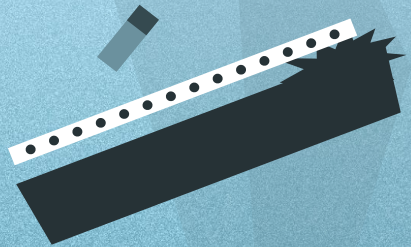
— Initial Data Findings —

Our initial investigation revolves around the question: **what are the most indicative characteristics of a passenger's survival?** We are observe the key variables that are useful in our data visualisations.

Age range	0 - 80
Total passengers	891
Pclass	1-3
Embarked Location	Southampton, Cherbourg, Queenstown
Survived	342

Outliers and null values

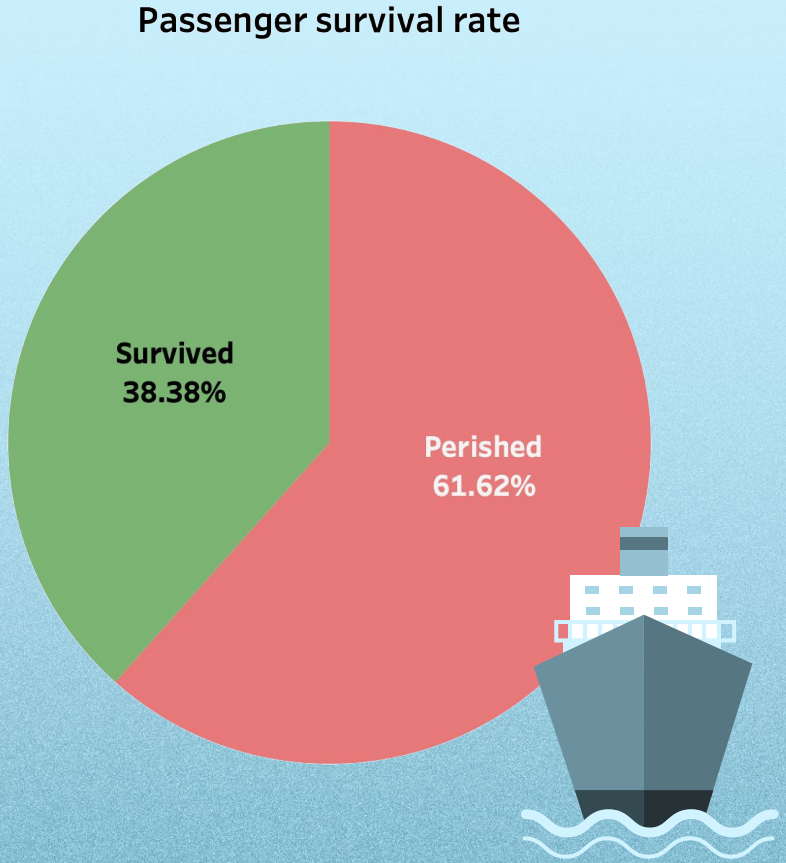
- Age had 177 null values. These rows were removed in the analysis of age and survival as they are are hard to accurately estimate
-
- Embarked had 2 null values. These cells were filled with the most popular embark location: Southampton



2.

— Survived vs. Perished —

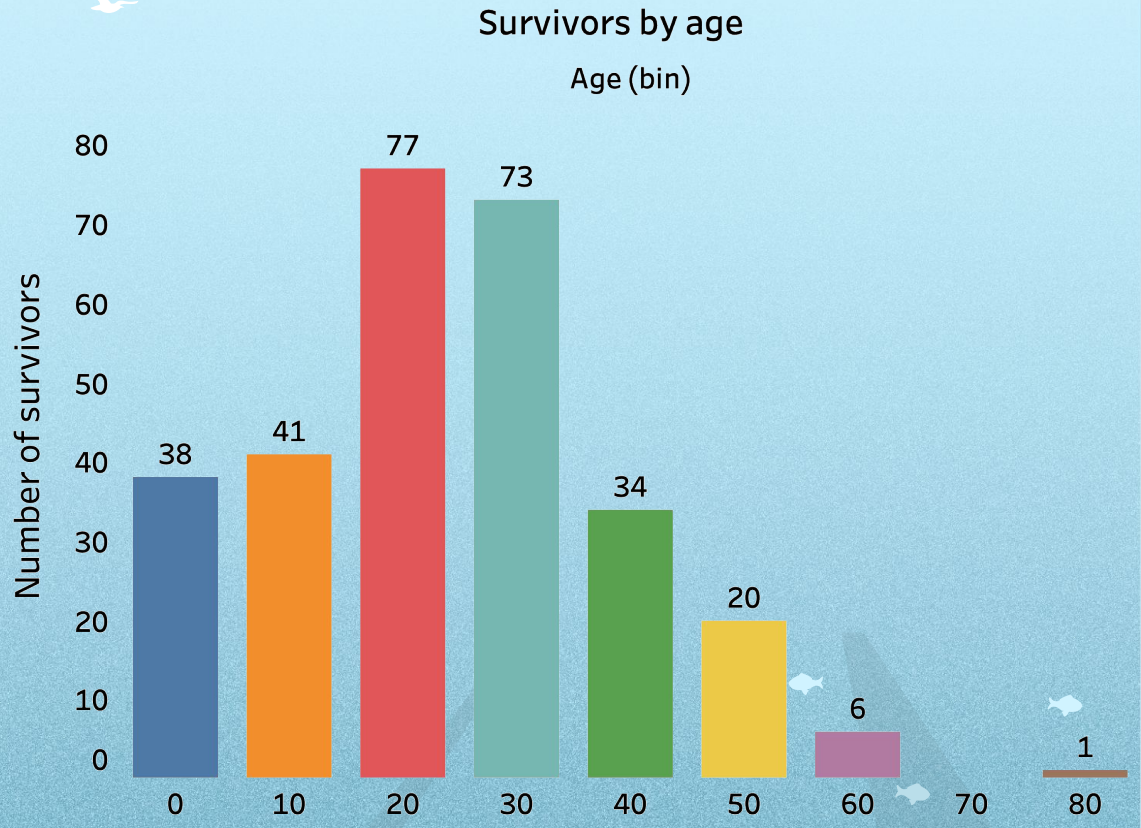
- The sample size of 891 passengers were used
  - From which 342 survived
  - 549 had perished
- Majority of passengers perished, approximately 61.62%
- The remaining 38.38% survived
- The crux of our analysis will be on survival rates



2.

— Survival by Age —

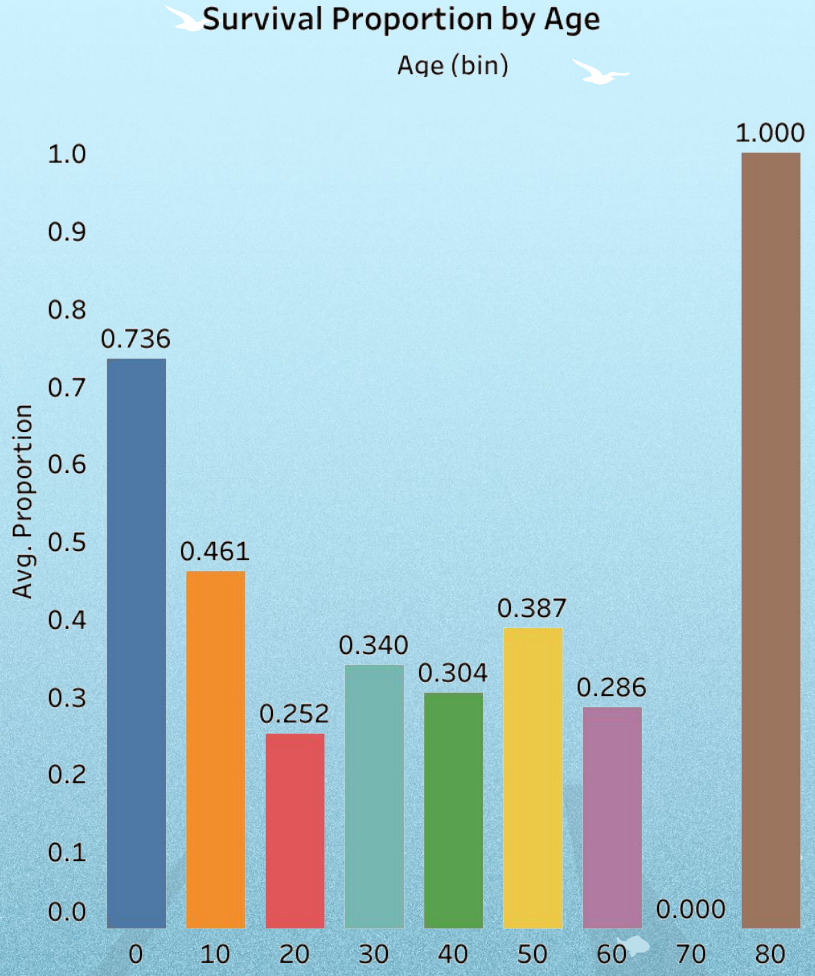
- There are greater number of survivors at the age of 20 compared to passengers of other ages
  - Followed by passengers in their 30s
- More children have survived than elders but not as much as 20 and 30 year olds



2.

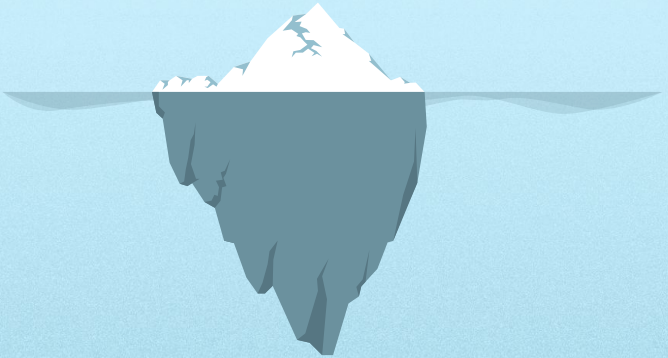
— Survival by Age —

- Children seem to have a greater average proportion in survival rate.
- 80 year olds have the greatest survival rate
  - Due to only 1 passenger being in their 80s, this visualisation can be misleading
- Between ages 20 to 60, there are no notable features, survival rates by age in proportion is relatively consistent

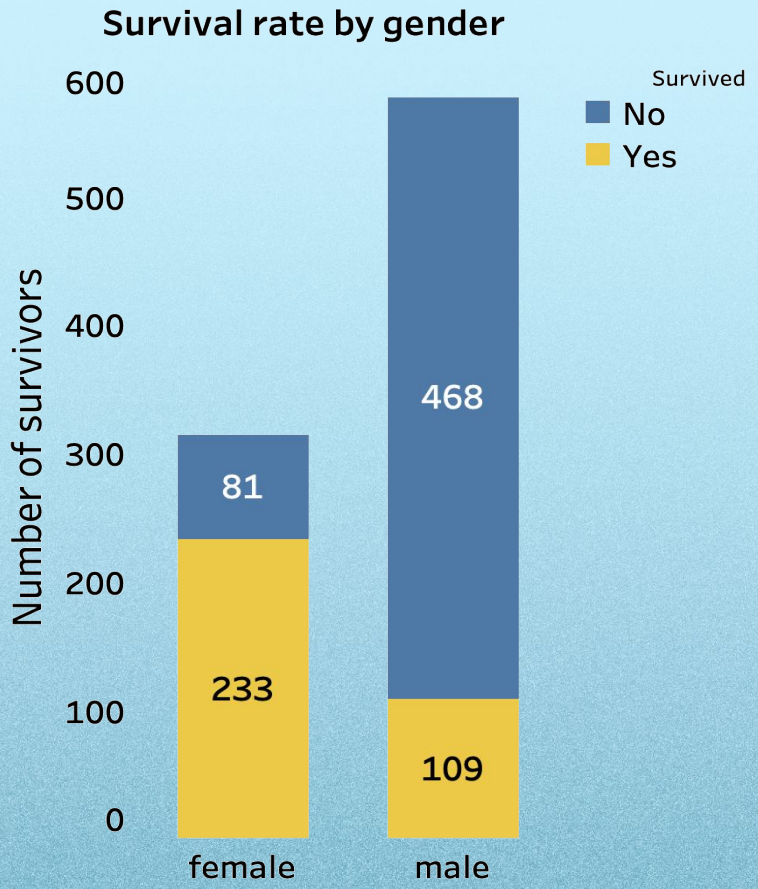


2.

— Survival by Gender—

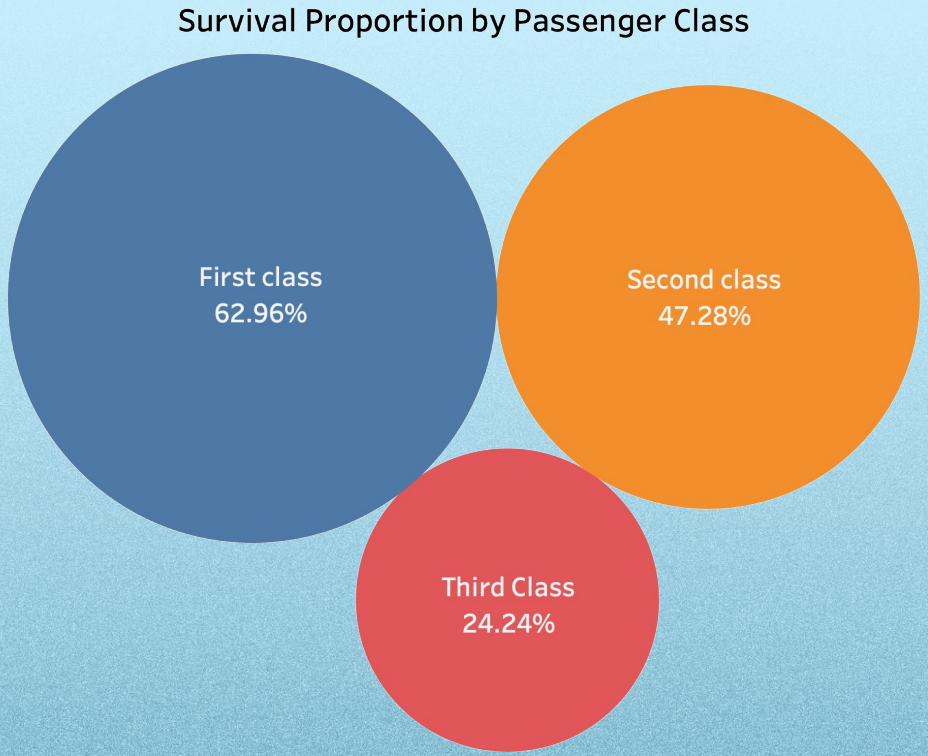


- Approximately double the number of male passengers than female passengers on board
- More females survived
- The proportion of female survivors is higher than the proportion of male survivors



2.

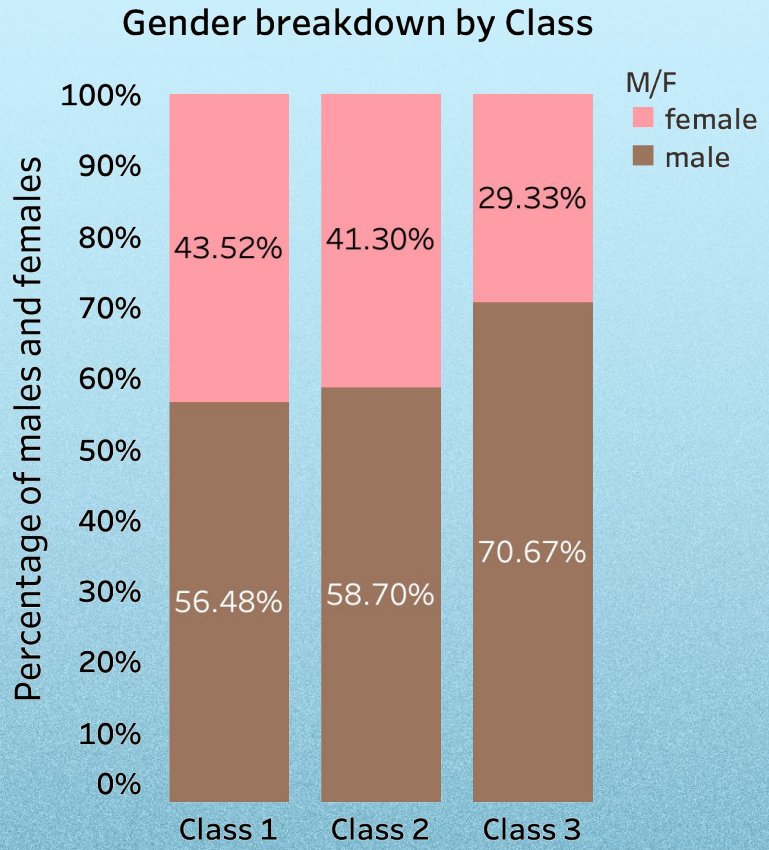
— Survival by Passenger Class —

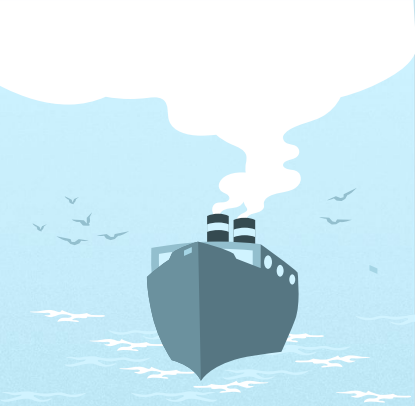


- 1st class has highest survival rate, followed by 2nd class
  - Almost three times greater than survival rate of 3rd class
- The gap between 2nd and 3rd class is greater than gap between 1st and 2nd
- Indicates a strong relationship between passenger class and their chance of survival

2.

— Survival by Passenger Class and Gender —

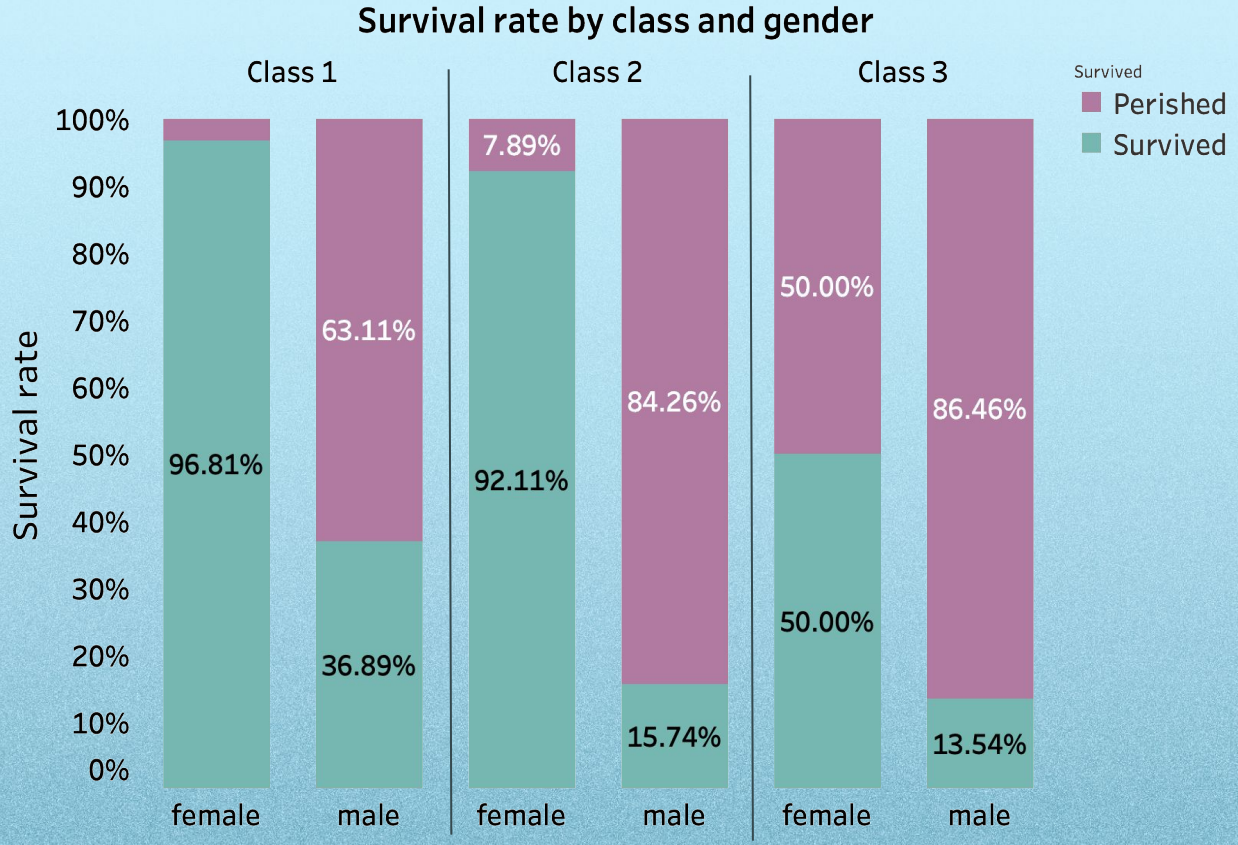




- Higher proportion of men in third class passengers in comparison to first class passengers
- Since females have a higher survival rate,, we can expect a higher survival rate for first class passengers than third class passengers
- Effect of Pclass on survival rate can be partially attributed to gender; Pclass is a weaker variable than gender in determining survival rate

2.

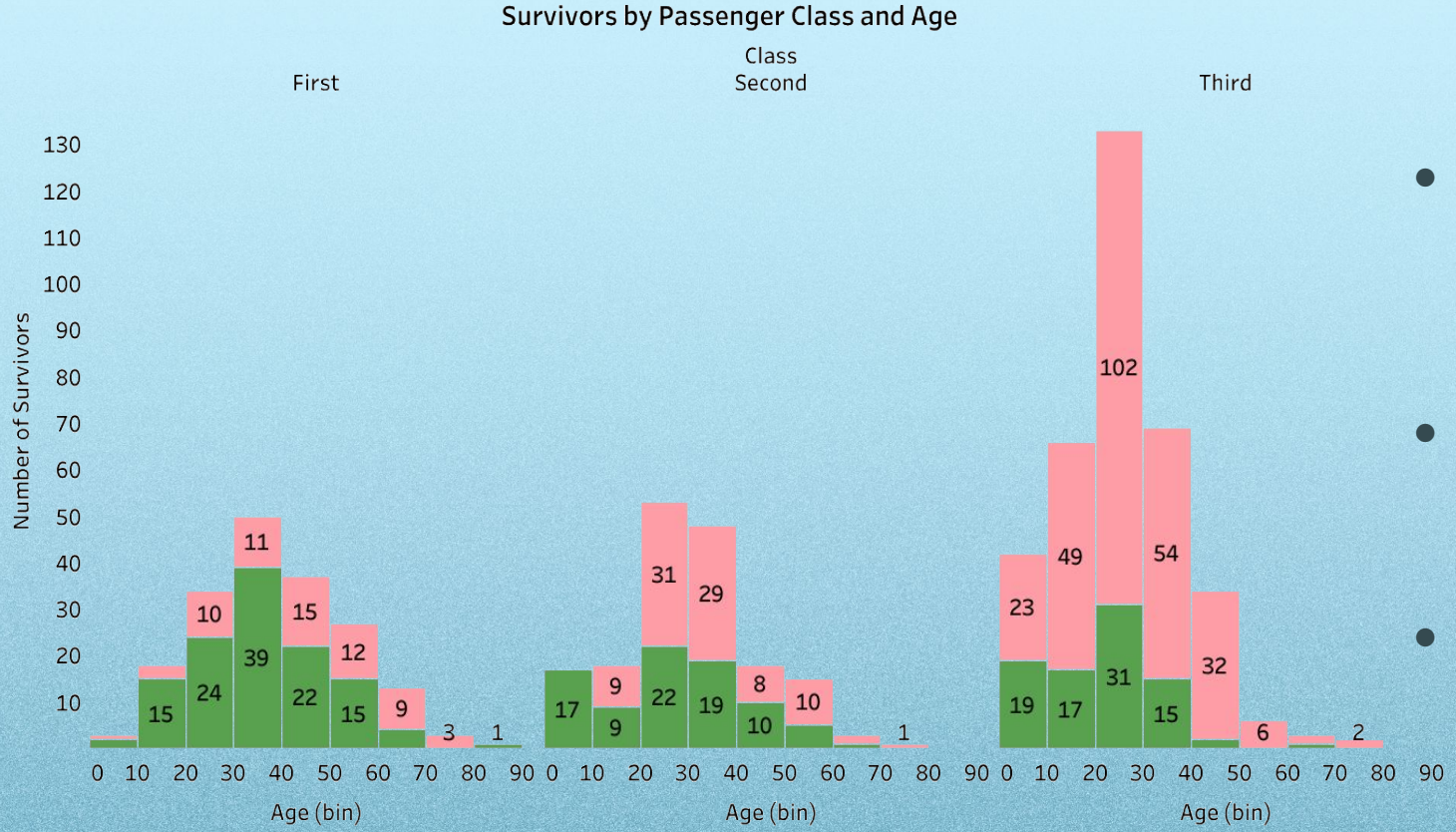
— Survival by Passenger Class and Gender —



- First class males have a lower survival rate than third class females
- A distinct drop in survival rates with people of lower classes
- Gender seems to have a more prominent impact in survival rates within each class
- Third class passengers have a much lower survival rate.

2.

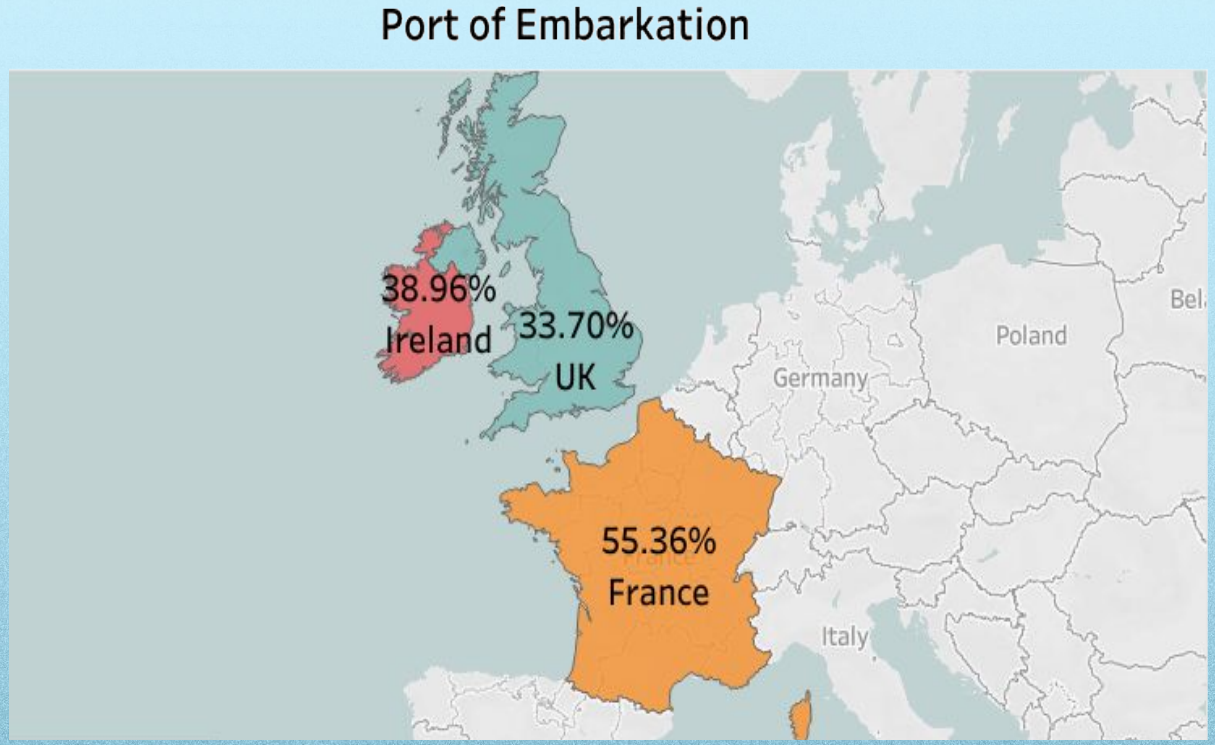
— Passenger Class & Age vs. Survival —



- A significant proportion of first class passengers survived, regardless of age
- Third class passengers had significantly lower survival rates
- Survival rates by age are relatively consistent across passenger classes

2.

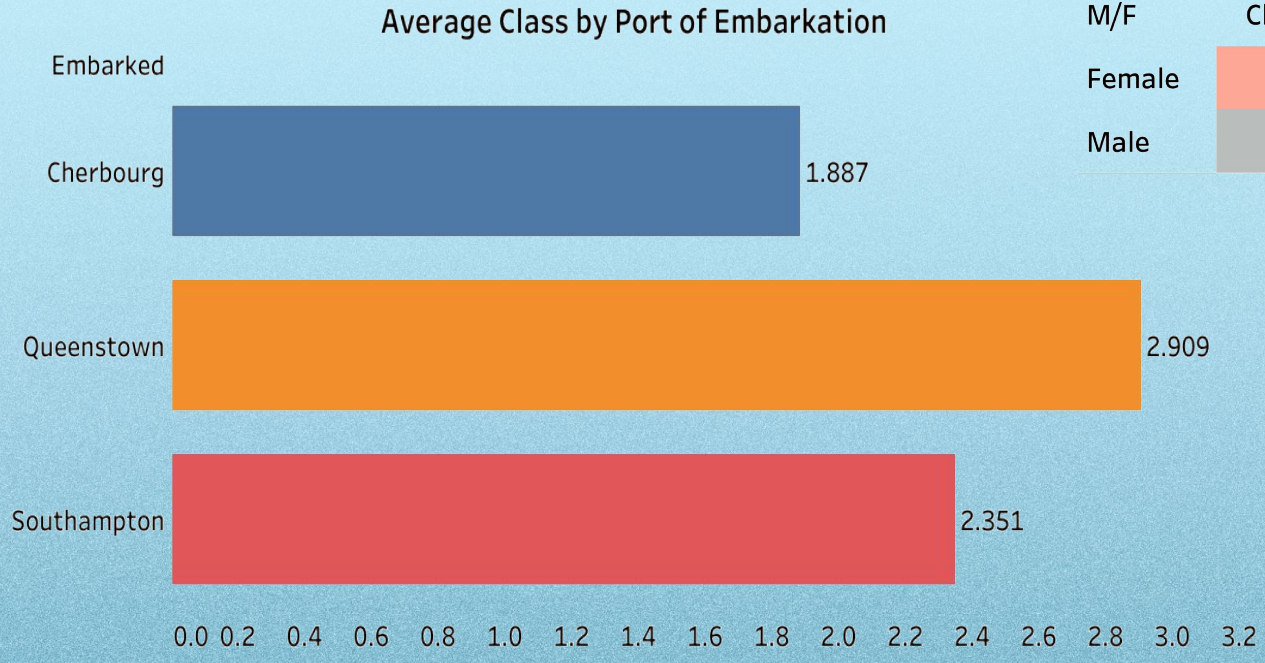
— Survival by Embarkation —



- Cherbourg (France) has the highest survival rates.
- Southampton (UK) has the lowest survival rates

2.

— Passenger Class vs. Embarked Location & Gender —



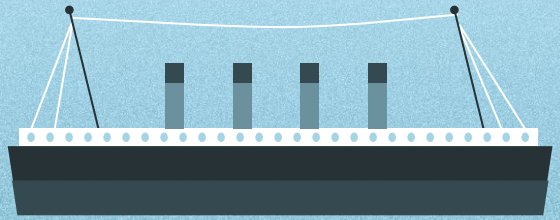
Gender percentage by location

M/F	Embarked: Cherbourg	Queenstown	Southampton
Female	43.45%	46.75%	31.52%
Male	56.55%	53.25%	68.48%

2.

— Passenger Class vs. Embarked Location —

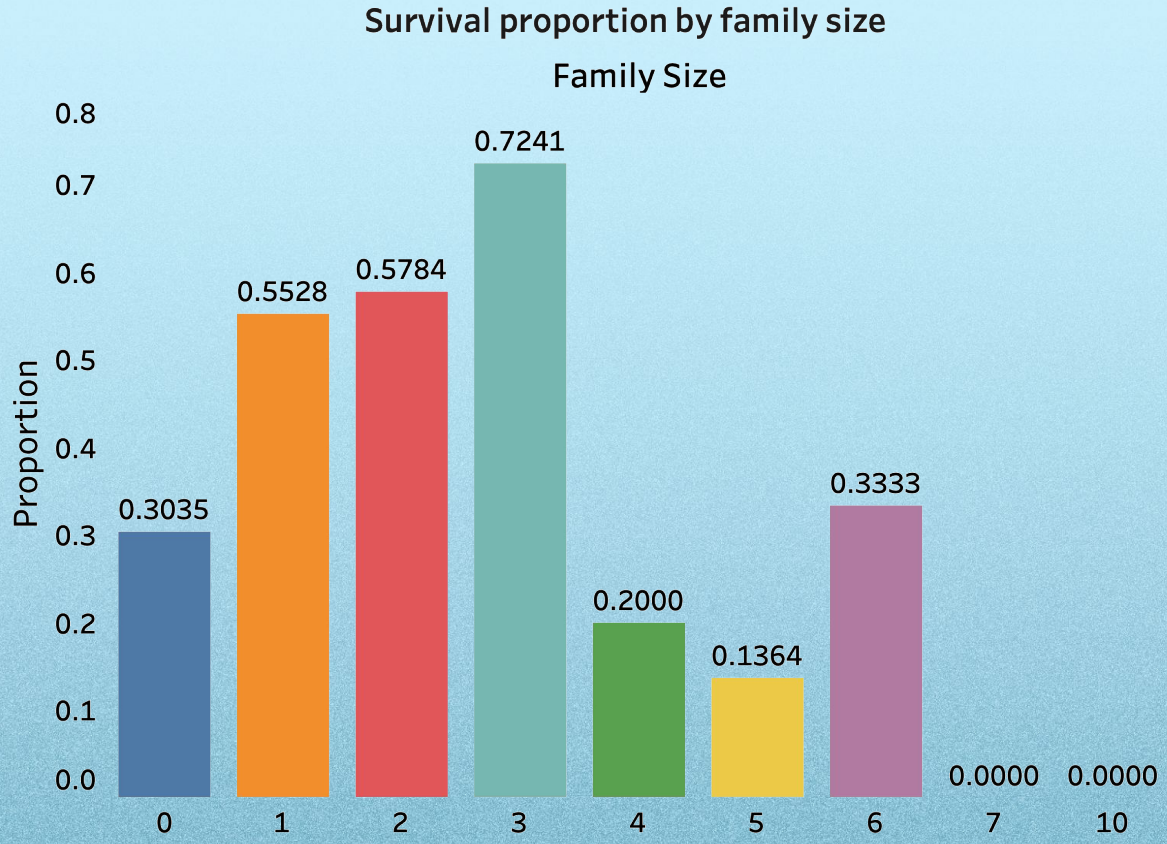
- Cherbourg makes the highest classes of people on average, with the highest survival rates.
- Queenstown predominantly has the lowest classes of people. However, it has the second highest survival rate. This can be attributed to the effect of sex; it has 47% female population in comparison to Southampton with 32% female population.



2.

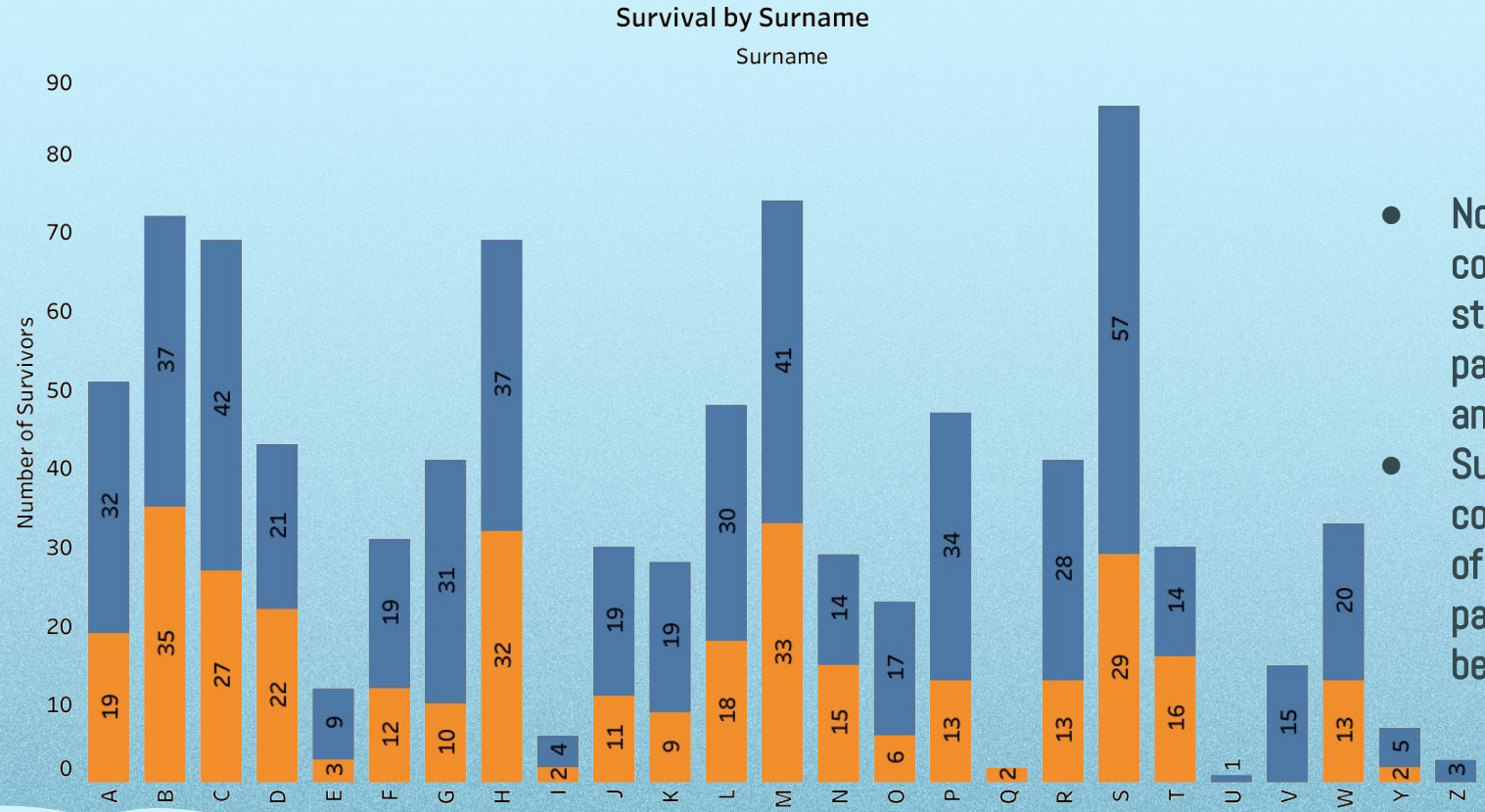
— Survival by Family Size—

- The survival rate of a person increases up to a family size of 3.
- Survival rates for families with 4 or more family sizes decrease dramatically.



2.

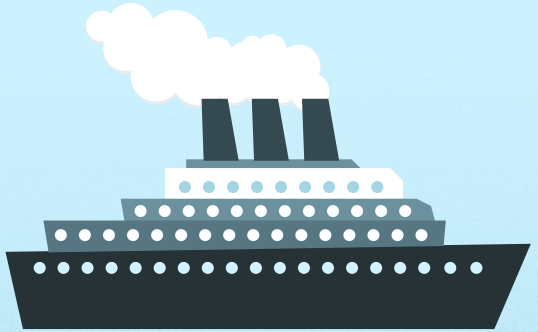
— Survival by Surname —



- No significant correlation between starting letter of passenger's surname and their survival
- Survival rate is consistent regardless of which letter a passenger's surname begins with

3.

— Conclusions —



1. Is there a relationship between age and number of survivors?
- Children have a higher survival rate
- Oldest age group bracket (70 years - 80 years) has 100% survival rate

2. Are there any differences in survivability between passenger classes?
- Notable differences in survivability in passenger classes
- The effect is partially attributable to the predominant gender within these classes

3. Do families survive?
- Survival rates increases for family sizes upto 4, possibly due to increases support
- Survival rates suffer a huge drop for family sizes greater than 4

4. **Does embarked location have any effects of survival rates?**

- The effect of embarked location on survival rates is quite negligible
- However, when combined with passenger class and gender, there seems to be a more noticeable effect
- On average, we have discovered that Cherbourg has the highest classes of passengers on average with the highest survival rate

5. **Who is more likely to survive, 3rd class female or 1st class male?**

- We have concluded that 1st class males have a lower survival rate than third class females

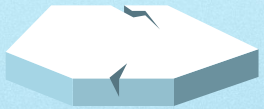
6. **Is there any alphabetical name discrimination?**

- Our analysis indicates no significant correlation between the starting letter of a passenger's surname and their corresponding survival rate
- Results showcase that there is consistency and no notable patterns

3.

What is the single most important factor in determining survival rates?

1.



Gender

Gender was the most influential factor in determining survival rates. Proportion of female survivors is considerably greater than males.

2.

Age

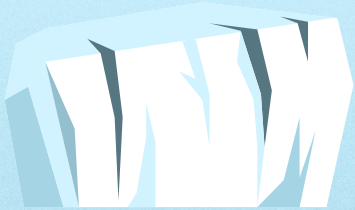
Age also had a significant effect as more younger passengers survived compared to elders.

3.

Pclass

Passenger class showcased a strong relationship with survivability, however its individual effect is minimised due to the effect of age and gender.

4.



Embarked Location

Embarked location indicated the weakest relationship with survival rates.

UNSW FINTECH SOCIETY

— Data Disaster Case Competition: PDF —

—Thank You—