



Clive Guo



Daksh Mukhra



Yvonne Thich

Data Disaster

Greatest Sea Disaster in Modern History



The unsinkable ship



Deaths of two thirds onboard

Who survived the Titanic?

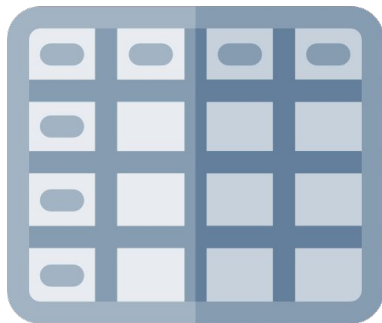
Exploring the impact of sociological constructs

1. How do **socio-economic determinants** affect survival?
2. How do **associative determinants** affect survival?
3. How do **inherent determinants** affect survival?

Data Preprocessing

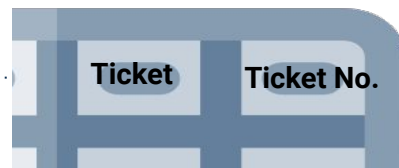
Our Dataset

- 891 passenger observations
- 12 features
- 20% missing values in Age
- Floating value ages
- Inconsistencies in Ticket
- 77% missing values in Cabin



Cleaning Ticket

Optional string prefix before ticket number



Extracted Ticket Number from Ticket column

Cleaning Age

Created new "Title" feature

Interpolated missing ages by finding median ages based on Title and Pclass

Compared floating age values to Title to verify validity of data

Table of Median Ages

Pclass	1	2	3
Title			
Dr	42.5	38.5	-
Mr	40	31	26
Mrs	41.5	32.5	31
Master	4	1	4
Miss	30	24	18

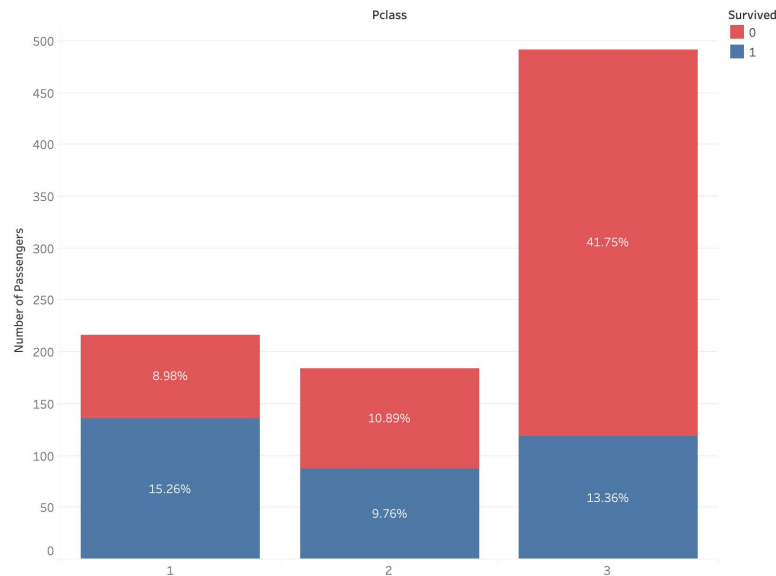
Does higher economic status lead to survival?

Survival Across Passenger Classes

First class 2 times more likely to survive than not

Second class equally likely to survive than not

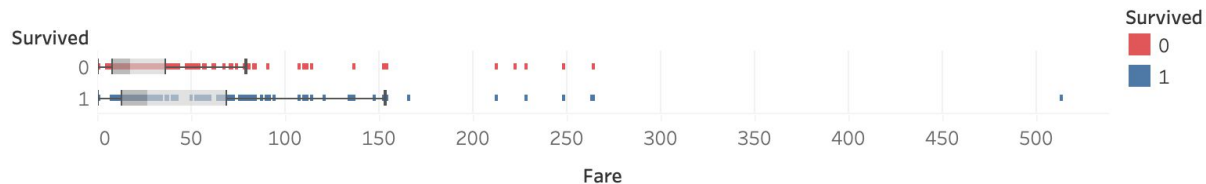
Third class 3 times more likely to **not** survive



Count of Passenger Id for each Pclass. Color shows details about Survived. The marks are labeled by % of Total Count of Passenger Id.

Do people who pay more survive?

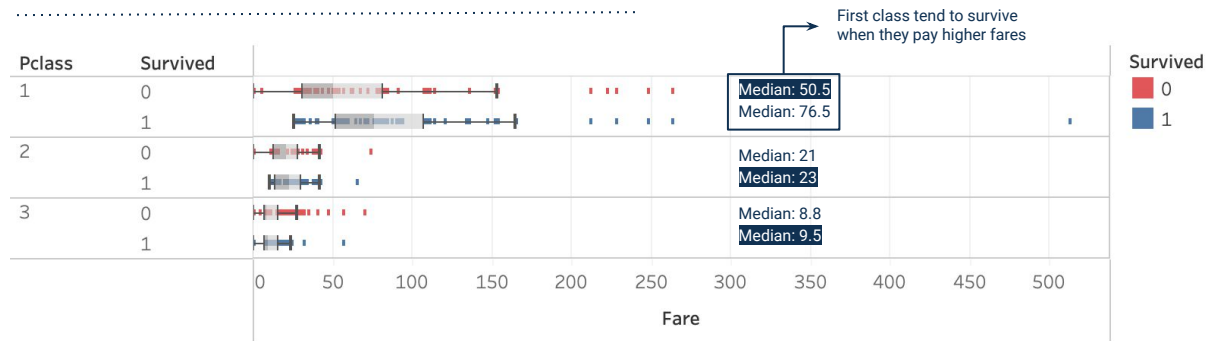
Fare Price vs Survival



Higher passenger classes tend to pay more for their fares

Fare for each Survived. Color shows details about Survived.

Fare Price and Passenger Class vs Survival



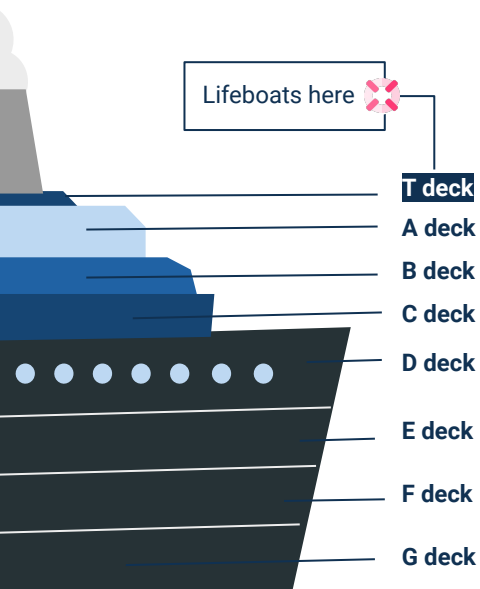
Presence of survivors who paid less than some first class passengers

Paying higher fares to survive does not necessarily hold true

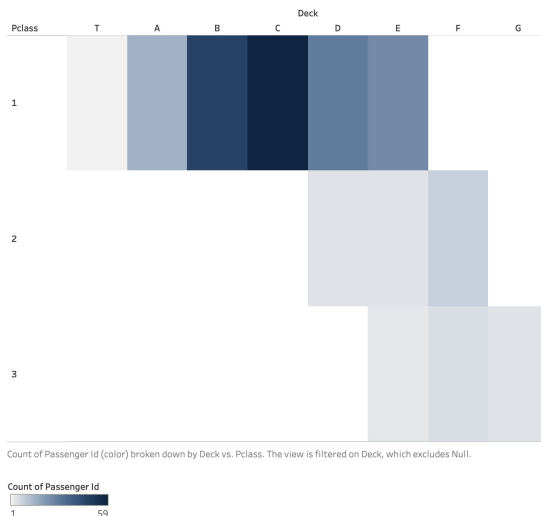
Fare for each Survived broken down by Pclass. Color shows details about Survived.

Does cabin location influence survival?

Cross Section of The Titanic

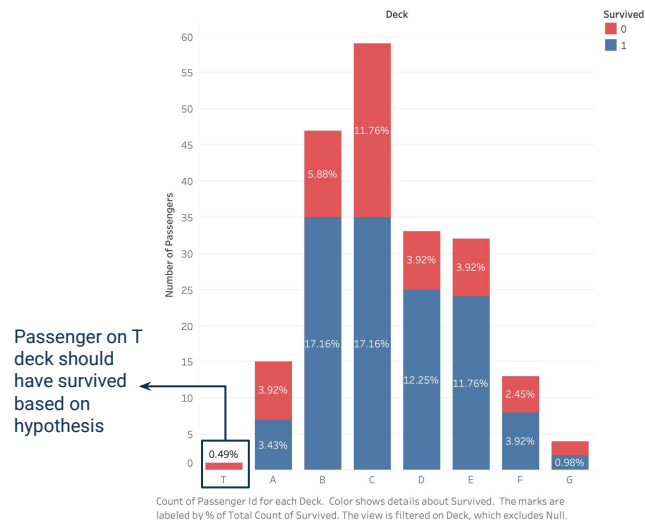


Heatmap of Passenger Class Distribution Across Decks



Survival Across Decks

Proximity to lifeboats does not positively correlate to chance of survival



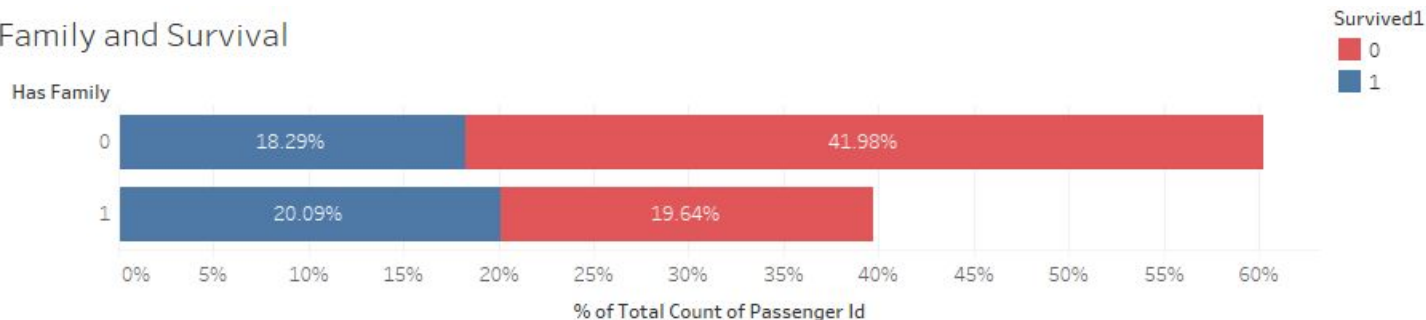
Does having family affect survival?

Family and Survival

60% of passengers did not have a family member onboard

However more than 50% of passengers with family members were third class

Family and Survival



With a family member onboard, there was just over a **50% chance of survival**

VS

Without family members onboard only just over **30% chance of survival**

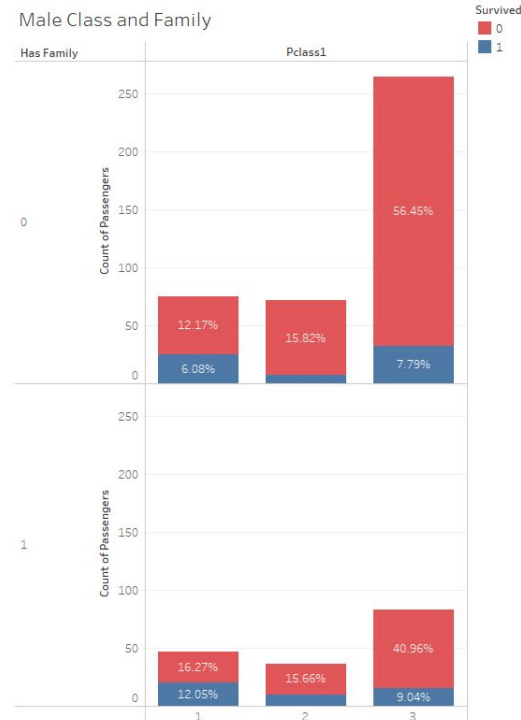
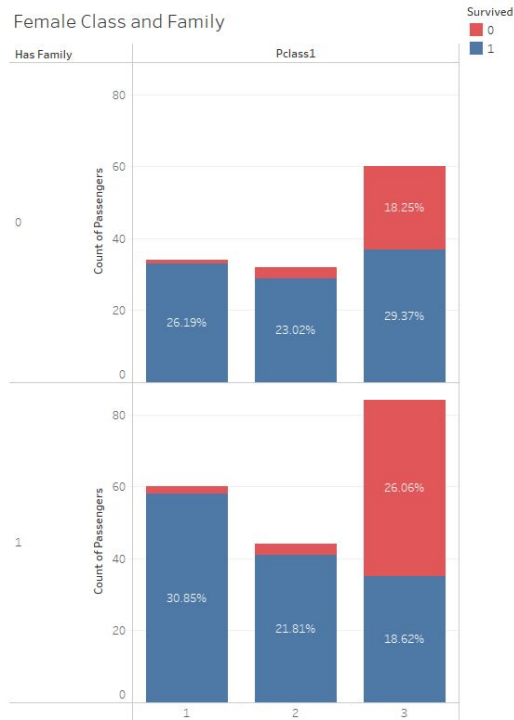
Does class in families affect survival?

Gender, Class and Family

Third class females with family had just **40%** chance of survival compared to all **other groups** with over **50%** chance of survival

Male survival was consistent across all passenger classes regardless of having family or not

Reasoning: Third class females may have felt more inclined to stay with family considering immigration context



Does companionship affect your chances of survival?

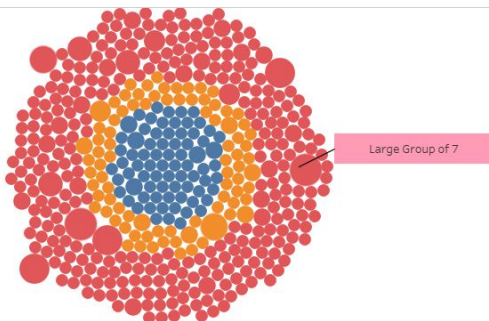
Ticket Grouping & Survival

Survived

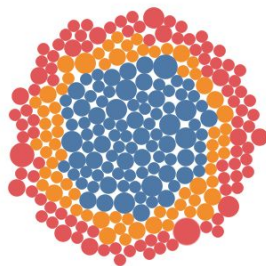
Pclass:



0



1



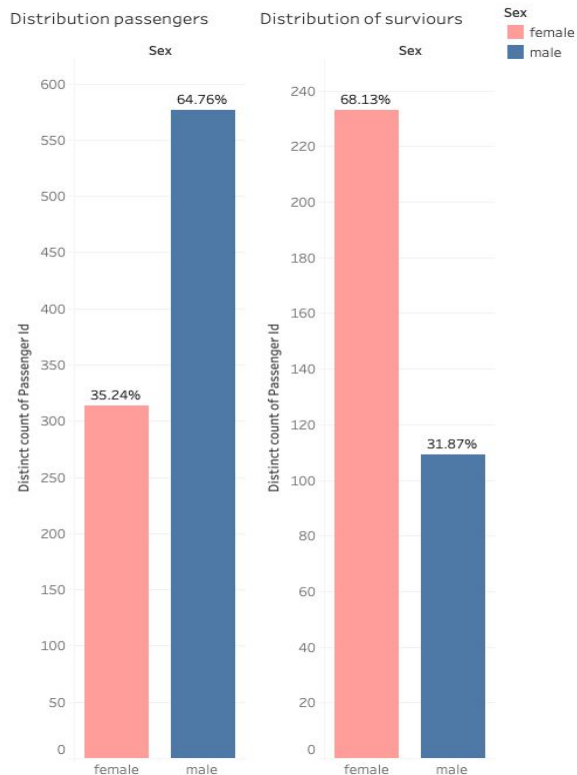
Ticket Grouping

'Groups' could be friends or families that booked under the same ticket

No groups of more than 5 survived - all large groups of 5 or more were from **third class**

Reasoning: **Third class** groups and families may be more reluctant to leave friends and family behind - **idea of immigration to start new lives may become obsolete without companionship and family**

Does sex affect survival?



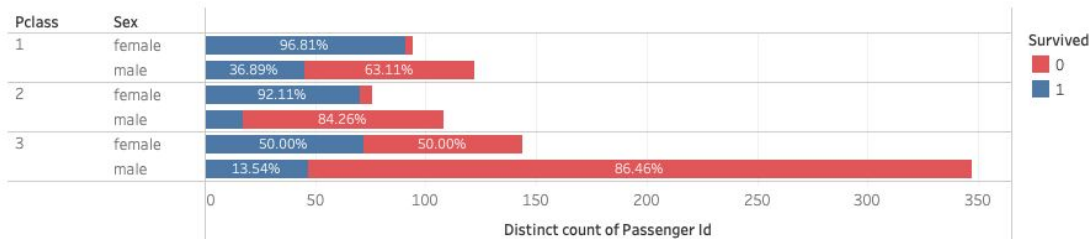
Survivors Across Sex

Females made up a disproportionate number of survivors

Looking at the distribution of sex within the survivors, **68%** of all survivors were female despite only making up **36%** of all passengers.

The Interplay of Class and Sex

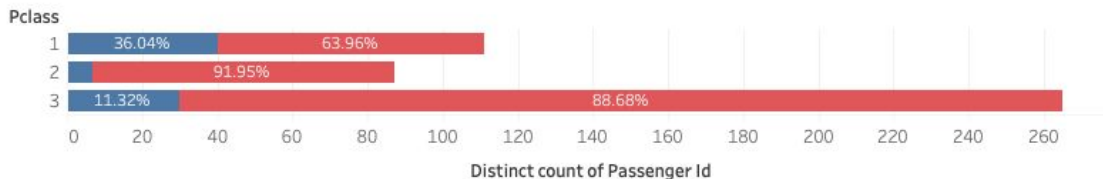
Distribution of Class and Sex



Females overwhelmingly have a greater chance of survival compared to males across all three classes

Approximately only 8% of adult males in class 2 survived

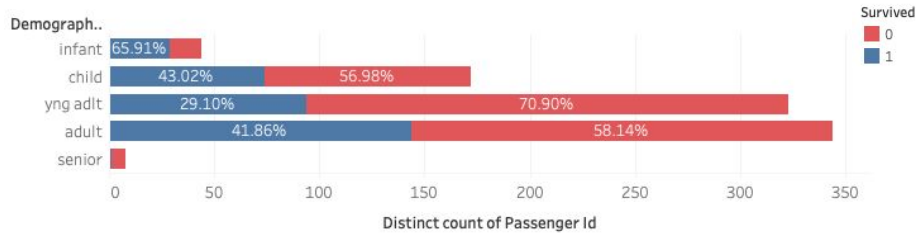
Survival Across Adult Males and Class



This is our first sign of an **altruistic act** that has been committed on the Titanic - men have purposefully given up their spots in the lifeboats for women.

The Interplay of Age and Gender

Distribution of Survivors Across Age Bins



Infants had the greatest chance of survival with a survival rate of **66%**.
However it is **not the case the older you are the lower your survival rate**.
Young adults had a lower survival rate than adults.

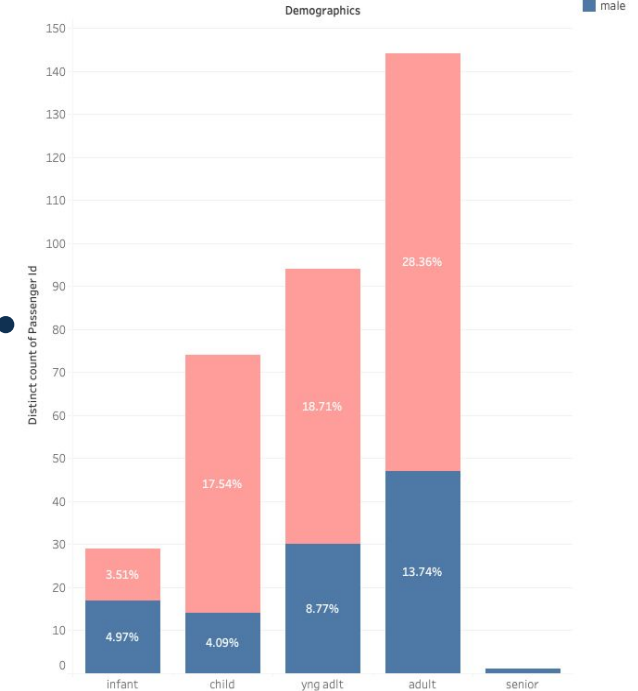
Effect of Sex on Survival Taking Account of Age

Females have a greater survival rate in children, young adults and adults

Only 23% of all survivors were adult males

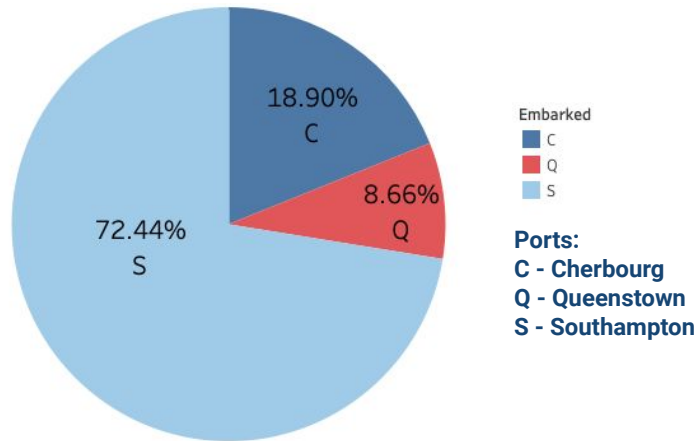
Sex still takes priority in determining survival

Women and Children



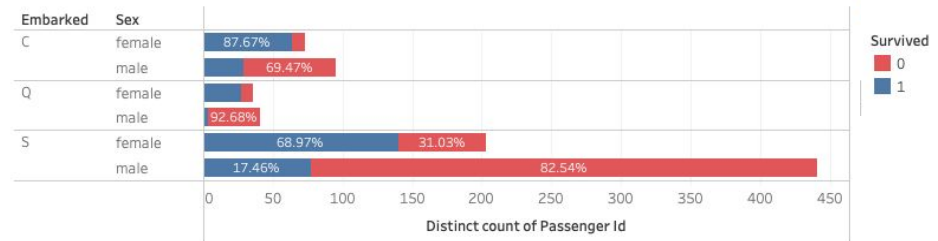
Does location embarked affect survival?

Distribution of Passengers Across Embarkation Points



73% of all passengers embarked from Southampton and consequently **most survivors are from Southampton**

Survival Rate From a Particular Port



Breaking down by sex and class, irrespective of where someone embarked, **females still have a survival rate greater than or equal to males.**

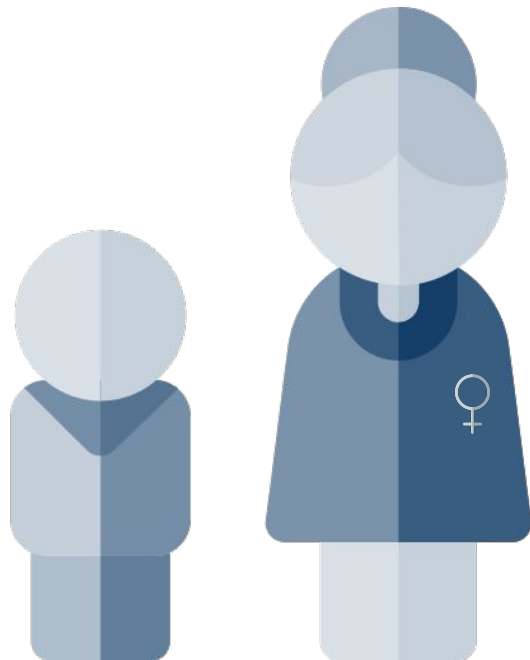
Conclusion

So who survived the Titanic?

Sex was the single most important factor in determining who was going to survive

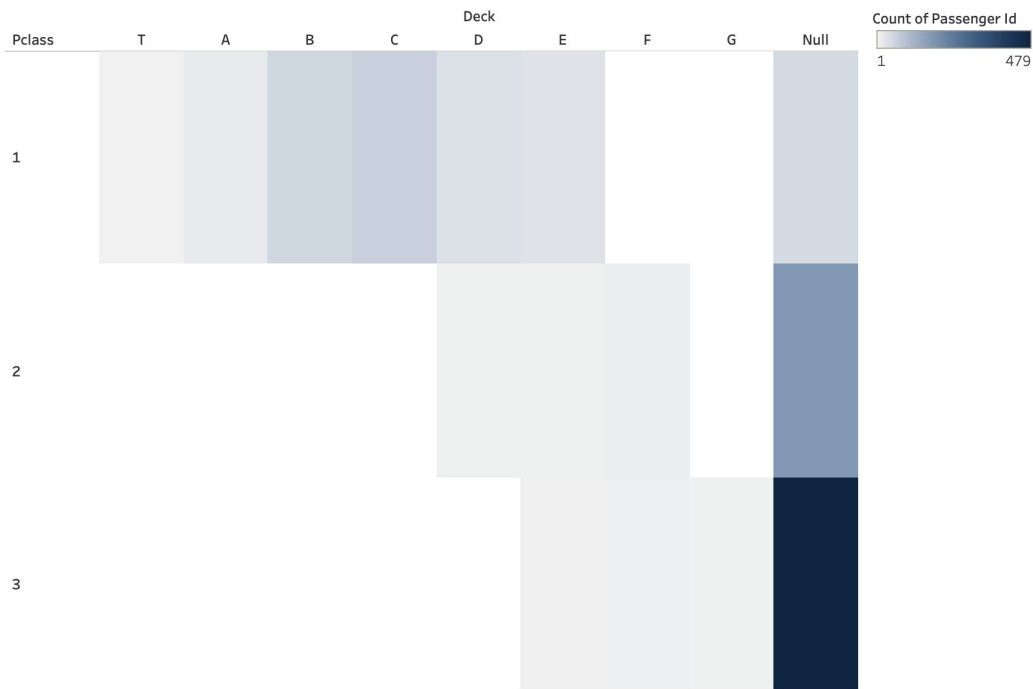
Survival rates will be influenced if you are in a higher class or whether you are younger.

Ultimately if you split any combination of groups by sex, **females had a greater chance of survival.**



Appendix

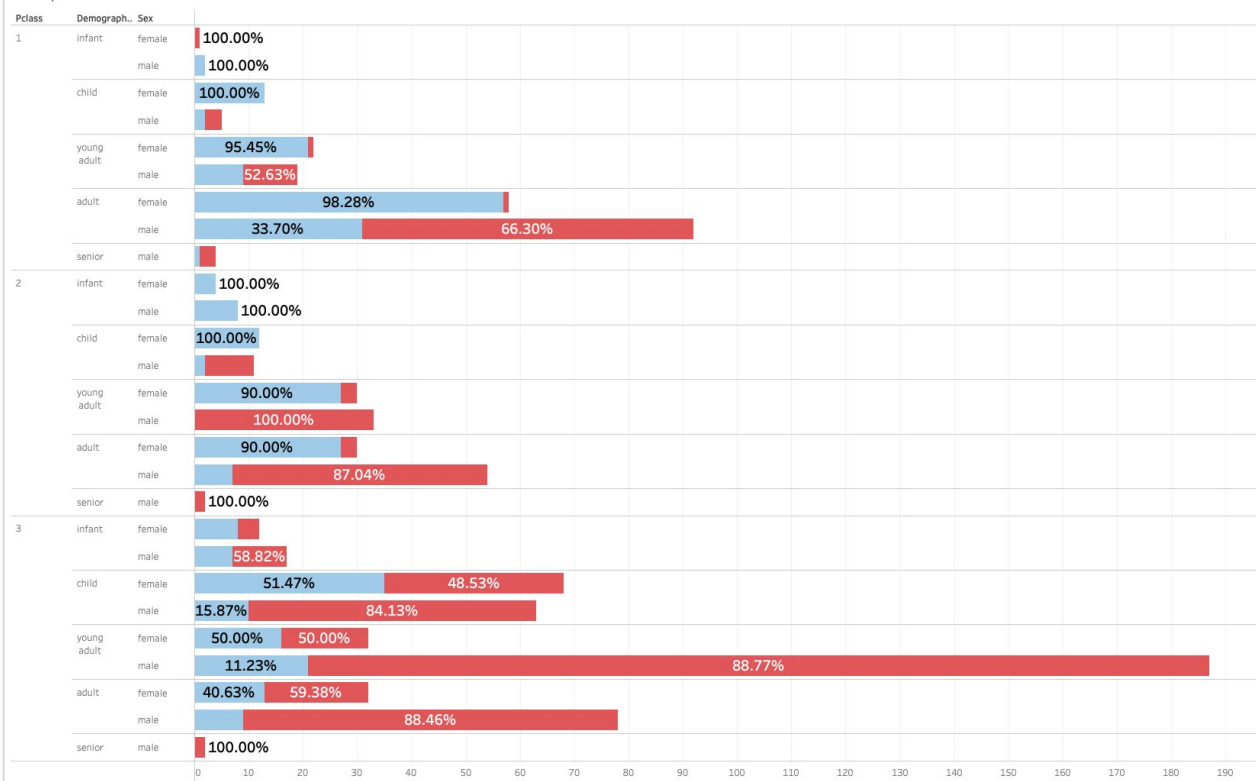
Deck vs Pclass



Count of Passenger Id (color) broken down by Deck vs. Pclass. The view is filtered on Deck, which keeps 9 of 9 members.

Appendix

Class/survival



The Importance of Sex

Sex is the most important factor in determining survival

There is not a single combination of factors in which a male has a higher chance of survival than a female across all ages and classes.