THE BILL+SALMAN GROUP EXAMINES THE SINKING OF RMS *TITANIC*

October 10, 2016

WHITE STAR LINE

BOARDING PASS

PERMISSION GRANTED TO COME ABOARD
WHITE STAR LINE'S
R.M.S.
TITANIC

ISMAY, IMRIE & CO.,
34, LEADENHALL STREET, LONDON,
AND
10, WATER STREET, LIVERPOOL
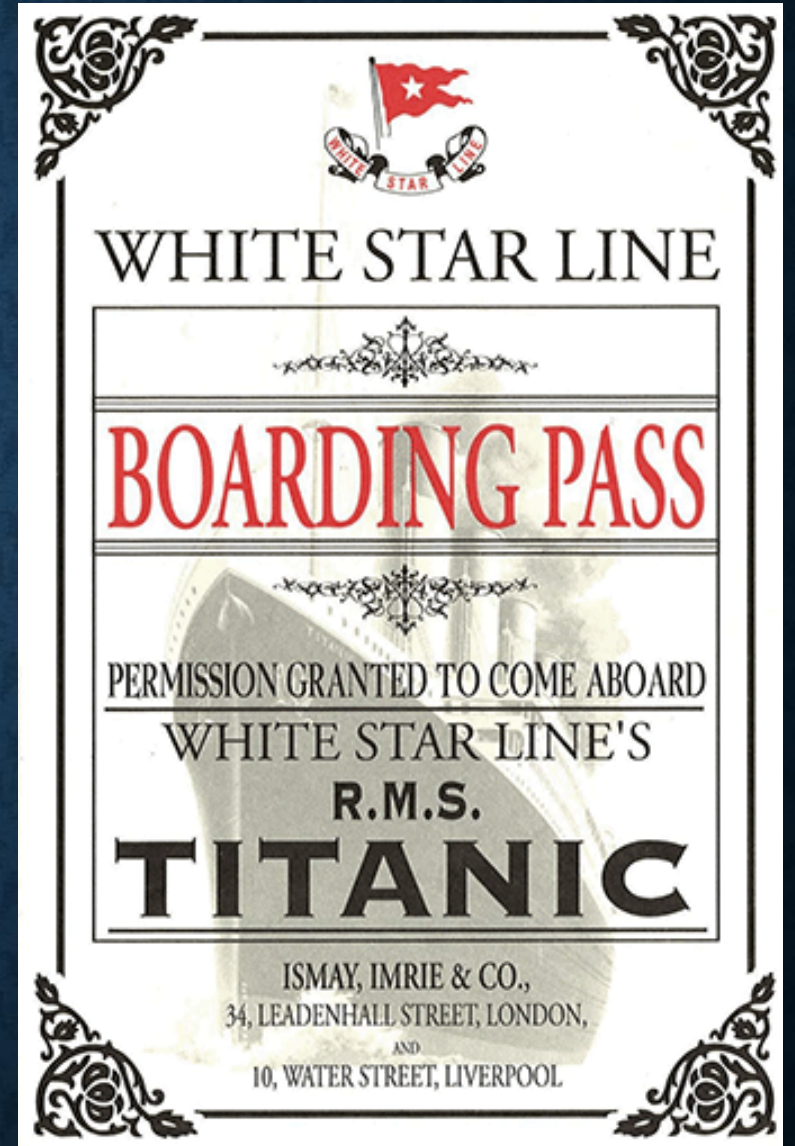
BACKGROUND

On April 14, 1912 at 11:40pm on her maiden voyage from Southampton to New York, the RMS *Titanic* hit the iceberg that would eventually lead to the sinking of the ship less than 3 hours later.

THE TRAGEDY

- *Titanic* struck an iceberg on the forward starboard side

- 1,500 lives lost; about two-thirds of the people on board

- 20 lifeboats were carried; there was room for 64

- Seawater temperature in the area was about 28.4 degrees Fahrenheit

- Passengers thought the ship was unsinkable, so they stayed on board. Life boats left with room on them

RMS Titanic   Person   Car   Bus   Airbus A380   Queen Mary 2
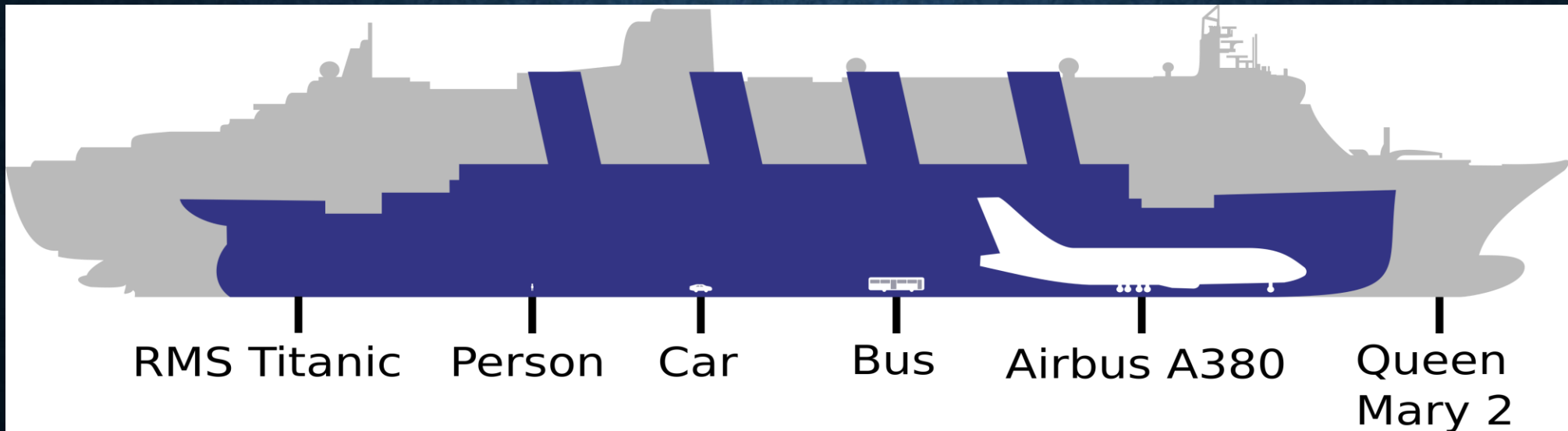
FORMULATION OF A RESEARCH QUESTION

INITIAL QUESTION -- MOTIVATION

Is what we think we know about the *Titanic* really true and can we use data to investigate?

- *Women and children first!*

- *The captain goes down with the ship*

And from popular movies like *Titanic*:

- *The upper class saved themselves at the expense of the lower classes*

Our **starting** question, then was "Did who lived and who died in the *Titanic* disaster follow as these preconceptions would suggest?".

This would not be our final question.

WAS OUR QUESTION S.M.A.R.T.?

- Specific

- Measurable

- Answerable

- Relevant

- Time bound

Criterion	Analysis	Action
Specific	The meaning of "going first" and the upper class saving themselves are too vague	Focus the question on specific variables and outcomes
Measurable	Somewhat; adult vs. child and male vs. female, and survived vs. perished are reasonably measured, but it is difficult to measure someone's class in society	Refine question to include type of ticket as a proxy for class
Answerable	It is difficult to ascertain motives with the available data; we can however see if the data is consistent with various theories	State the question as something that can be determined using the data
Relevant	Due to the continued public interest in the Titanic, the public concern of emergency preparedness during travel, and the social question of the just treatment of all citizens, the Titanic sinking remains a relevant question. However, whether the captain went down with the ship is trivial. Considering the about the crew in general would be more interesting	Consider the crew as a whole rather than just the captain
Time bound	Assumed, but we should explicitly state the limits of the question	Focus question to the immediate disaster

REFINED QUESTION

What are the relationships between the sex, adulthood status, and ticket class or crew status of the people on board the *Titanic* and whether they survived when it sank?

THE REFINED QUESTION VS. S.M.A.R.T.

Criterion	Analysis
Specific	The factors we wish to consider (sex, age, ticket type or crew member) and the outcomes being measured (survive or perish) are clearly stated
Measurable	The variables have a precise definition
Answerable	Relationships can be ascertained through visualizations and by using statistical techniques
Relevant	The question has been made more interesting by considering the entire crew rather than just the captain
Time bound	The historical event of the sinking of the Titanic bounds the scope of the question

OBTAINING AND PREPARING THE DATA

- Tools
  - R
    - plyr
    - ggplot2
    - Rstudio
  - Python
    - pandas
    - ggplot
    - Jupyter Notebook
- Data
  - R built-in Titanic data, from an online repository maintained by Vincent Arel-Bundock

FORMAT OF THE DATA

32 rows with 5 variables per row

```
# Show the first 10 rows of the original data
head(theData.original)

##     Class    Sex   Age Survived Freq
## 1    1st    Male Child       No    0
## 2    2nd    Male Child       No    0
## 3    3rd    Male Child       No   35
```

WE NEED TO TRANFORM THIS DATA

- Used R / Python code to tidy the data
  - One row per observation
  - → One row per passenger, indicating their Class, Sex, and Age and whether they survived

- From the tidy data, created summary data by (Class, Sex, Age) groups.
  - 14 groups (no Crew children)
  - Each row in the new data frame has number in the group that were rescued, perished, the total, and the percentage that survived in each group

- After every step along the the way, we checked that the transformed data was reasonable using the total number of passengers and summaries of the dataframes
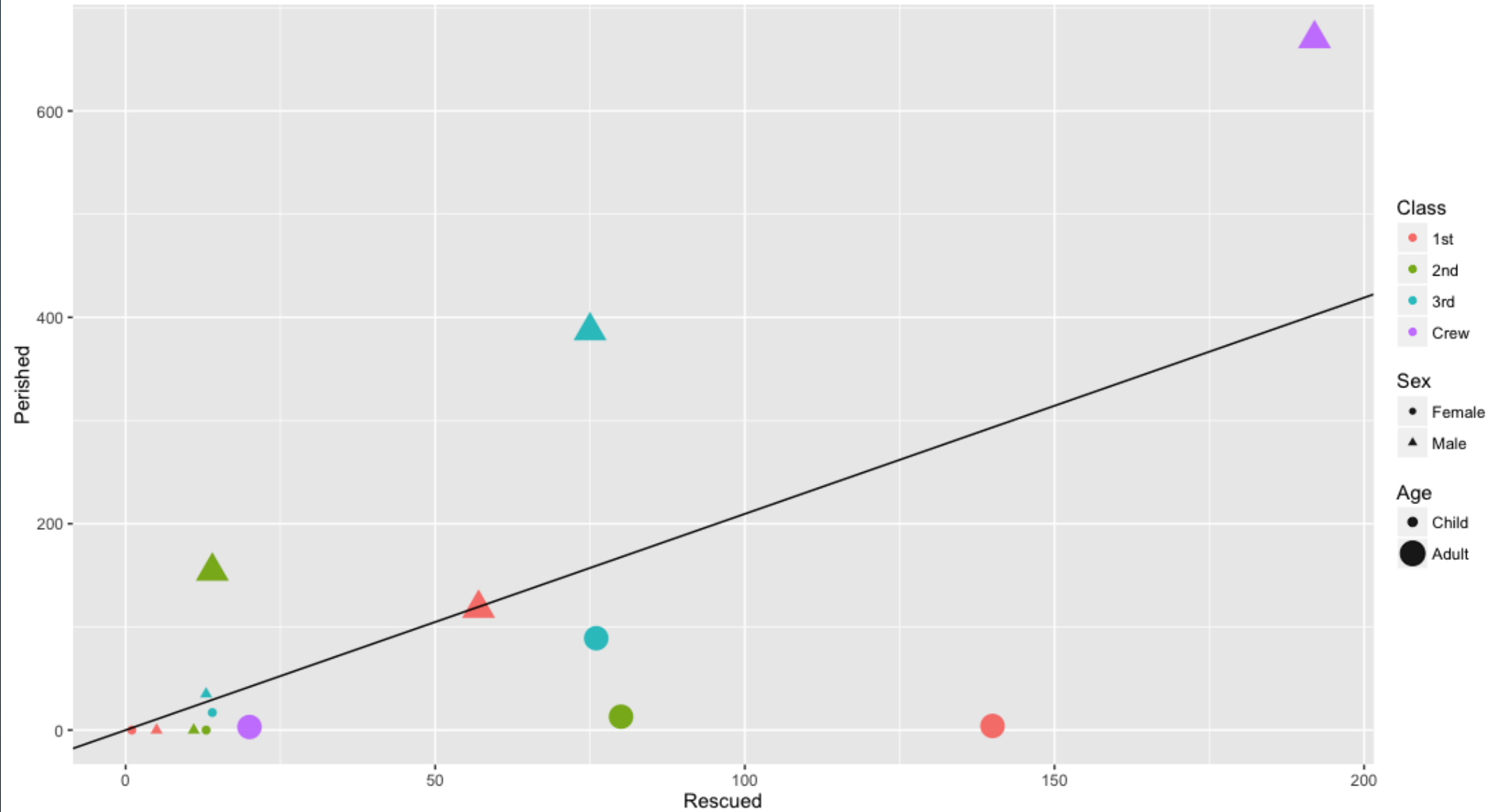
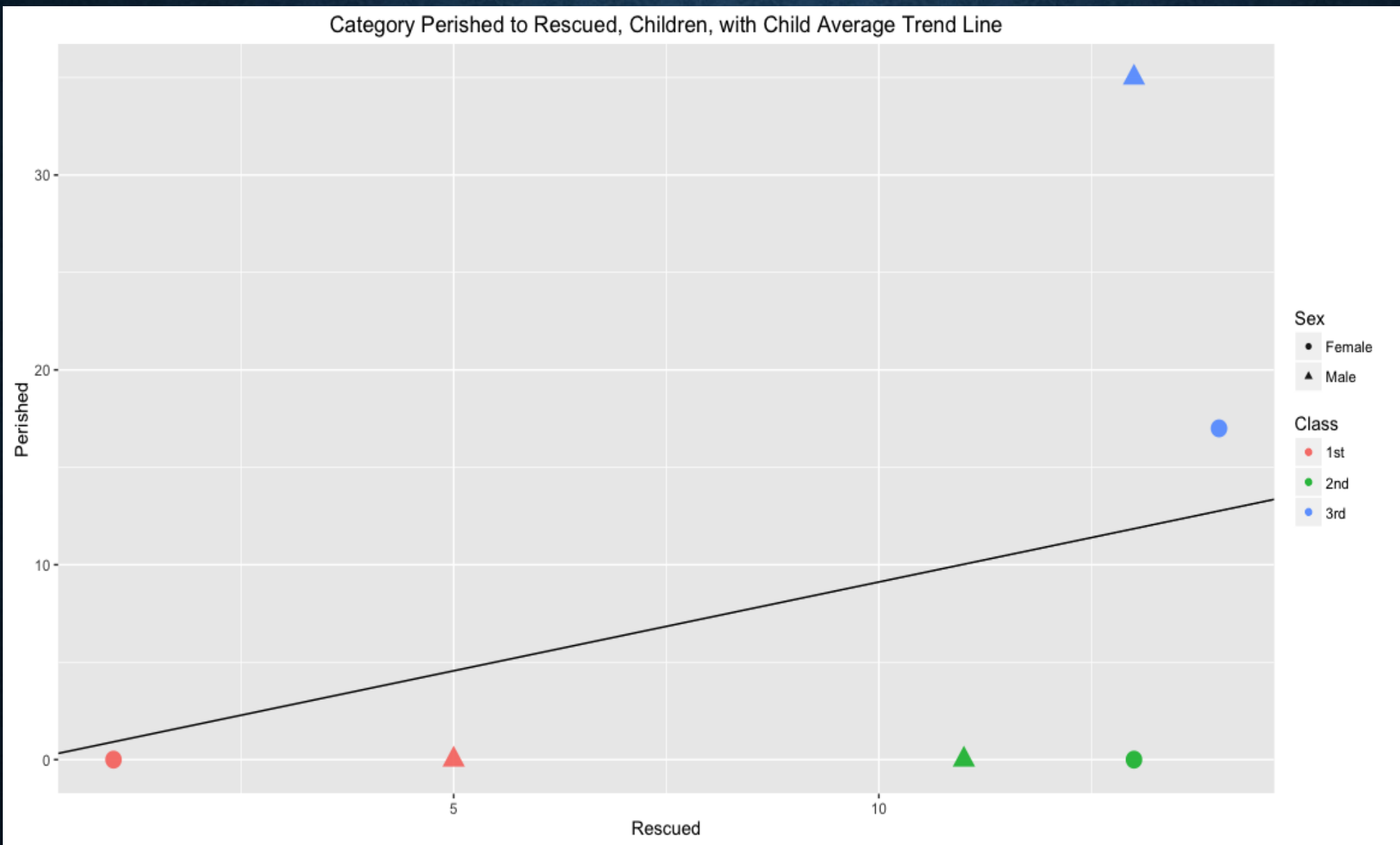*See RStudio html file for details*

ANALYSIS

Distribution of Survival Percentages of Groups
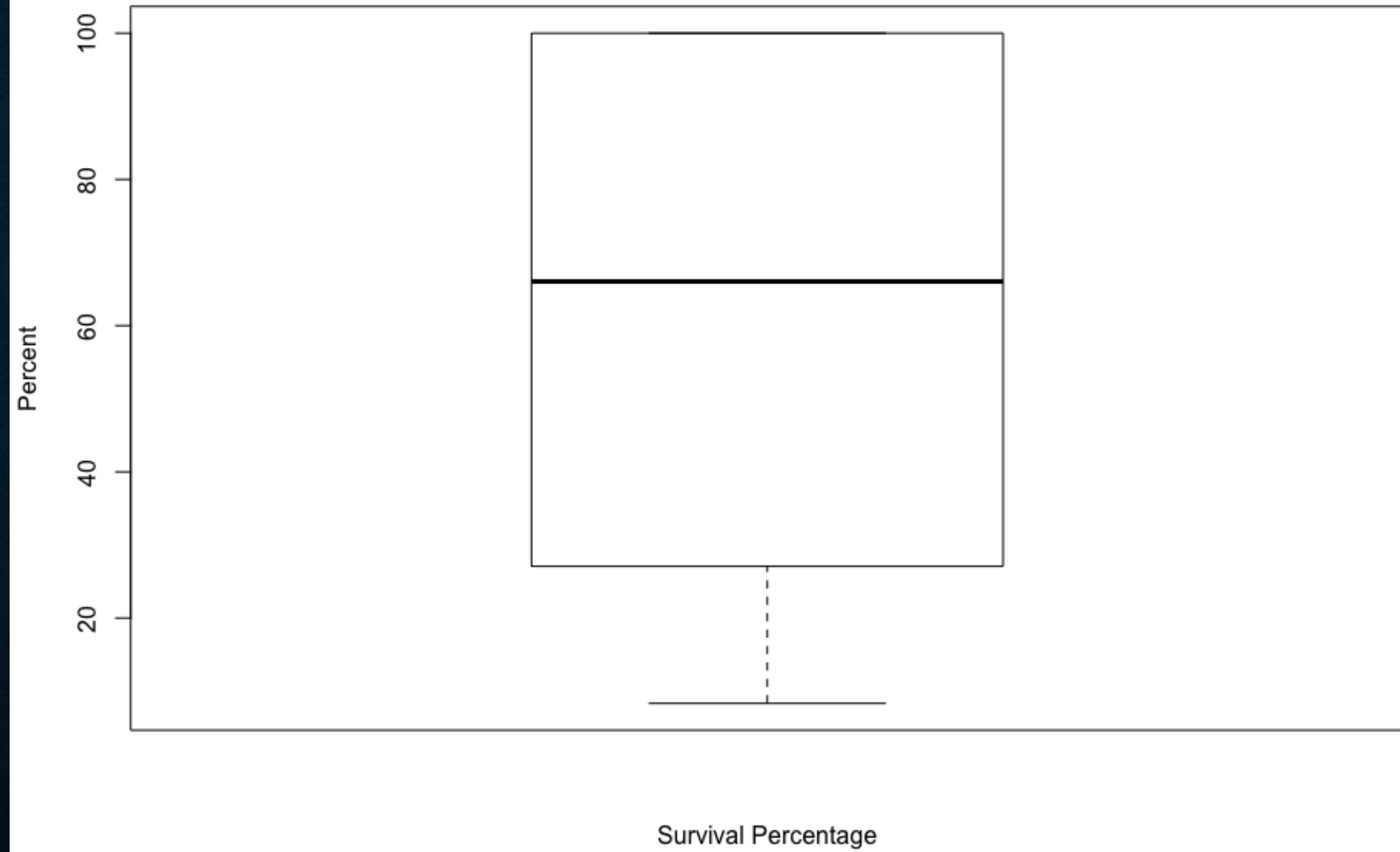
Perished to Rescued, All,
with Ship Average Trend Line

Below trend line – saved at above-average rate. 2nd,& 3rd class plus crew adult males did worse

Children saved at above-average rates, except 3rd class

LOGISTIC ANALYSIS

- We did a simple analysis to see if there were statistically significant relationships

- All non-first class passengers were at greater risk.

- For Age, being a child had a higher survival chance.

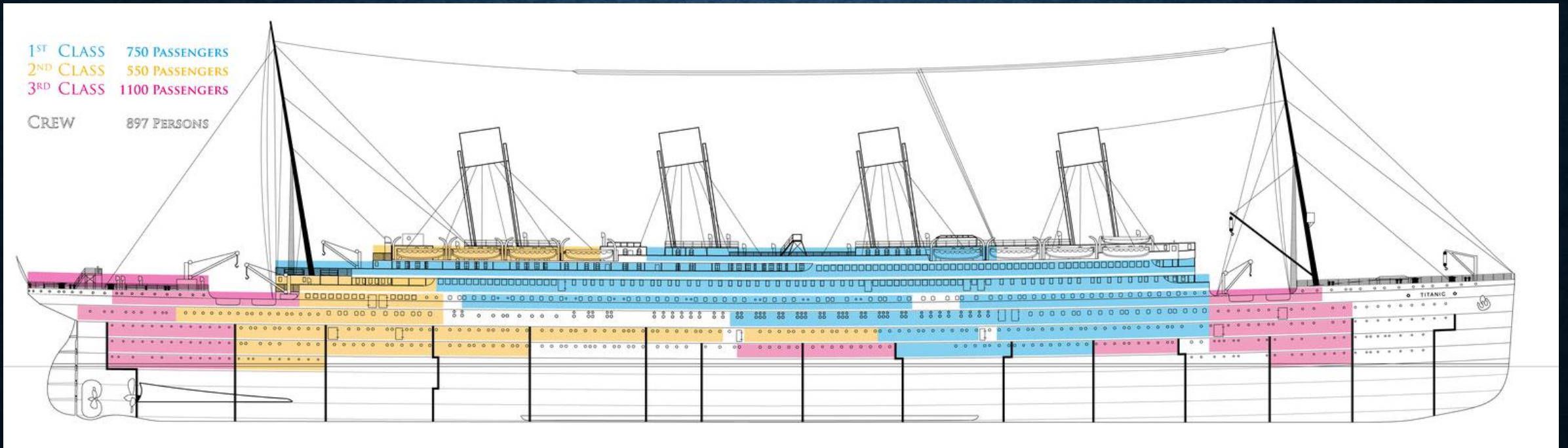- And for Sex, men had a lower rate of survival.
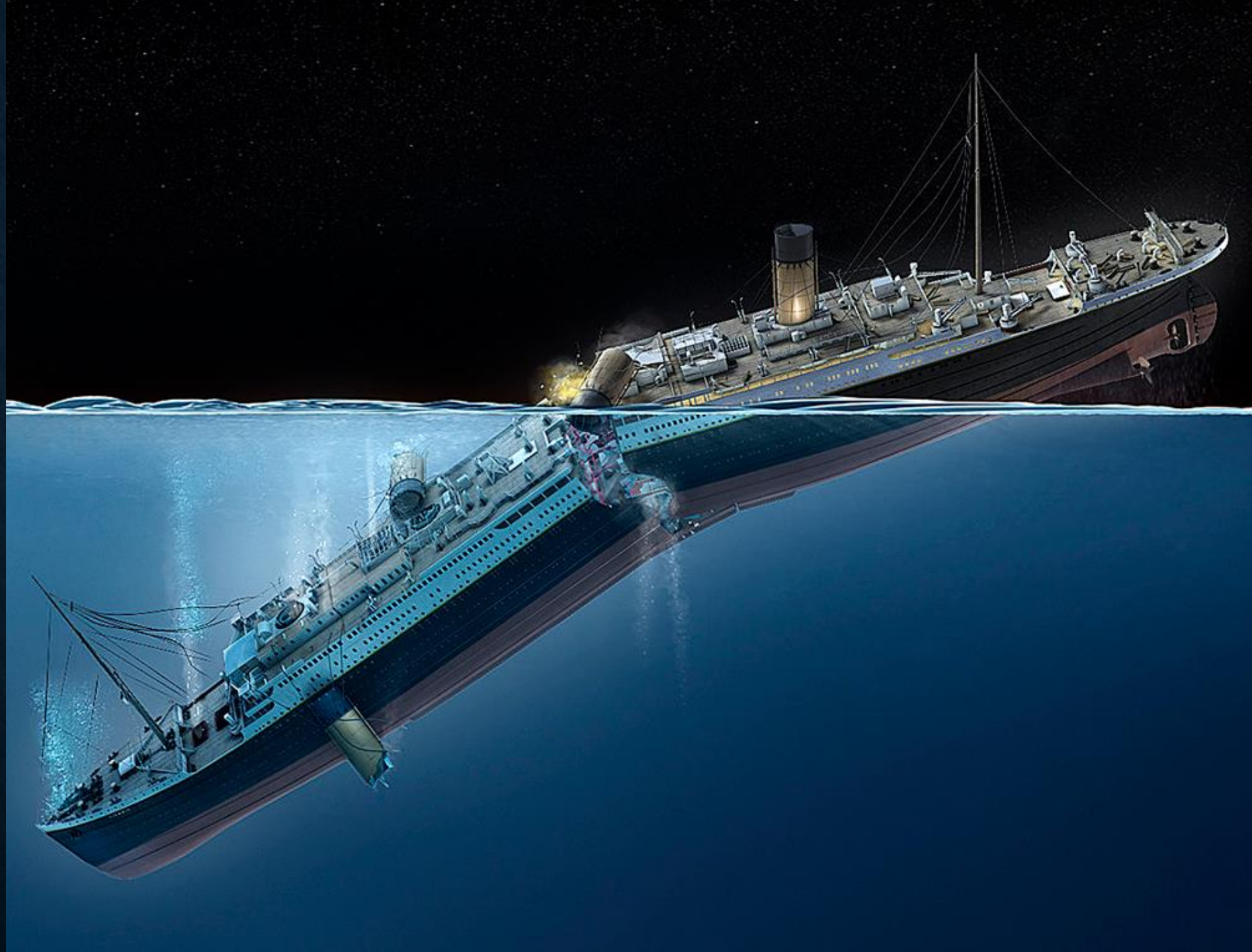
CONCLUSION

- The data tells a compelling story consistent with our preconceptions and consistent with the background research.

- The survival on board *Titanic* was related to ticket class, sex, and age. Women and children did stand a better chance of surviving in first and second class, but the losses in third class were devastating.

- The number one risk factor to not survive we found was to be a man, and this applied across all other categories.

FURTHER QUESTIONS

- Was the location of the third class berths near the ice berg impact site a contributor to the disproportionate rate of loss of the third class passengers? This question could be studied with data giving the berth location of each survivor and of those that died.

- As the disaster unfolded, what were the communication patterns among the crew and did this lead to changes on how the lifeboats were filled? We only have data from the disaster as a whole; it would be interesting to study the event as a time-series of unfolding events, such as the loading of each lifeboat.

- What became of the people who survived? It would be interesting to compare *Titanic* survivors to the general population for signs of Post Traumatic Stress Disorder and other mental conditions.

Berths by Class

1ST CLASS   750 PASSENGERS
2ND CLASS   550 PASSENGERS
3RD CLASS   1100 PASSENGERS

CREW   897 PERSONS

TITANIC

FINAL WORDS

Our final conclusion is that it was a privilege to study this data, and in so doing, connect with the life and death struggle of fellow humans at peril on the sea.

100TH ANNIVERSARY