# Social Hierarchy and Survival Likelihood on the Titanic

Joshua Tice

Introduction

The objective of this project was to create an explanatory data visualization using Tableau to convey an insight into the Titanic dataset. I wanted to show evidence that social hierarchy and social norms had a strong influence on which passengers survived the sinking of the doomed vessel. In particular, traveling in first class, being a woman, or being a child greatly increased one's chances of survival. Additionally, the social status of the passenger could offset influences from the other two features (*i.e.*, age or sex).

Design

Initially, when I began designing the visualization, I was only going to focus on survival likelihood as a function of Sex and Class. Specifically, I wanted to convey the message that being female or being in first class correlated strongly with greater survivorship. My thought was to display the number of passengers in each segment with bar charts. Then, I was going to apply color to indicate which fraction of each segment survived. My initial sketch is shown below.
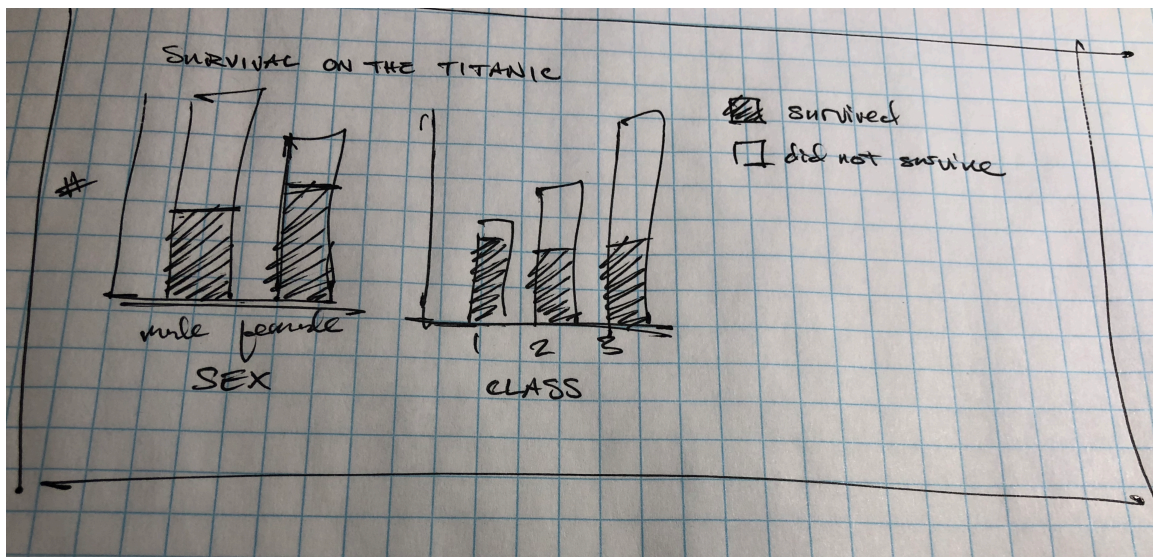
I translated the sketch into a Tableau dashboard, but I thought that the initial concept was too simple. I added several additional features to convey extra information. First, I added a tooltip that would display the exact number of passengers in each segment along with the segment information. Second, I added an action to the dashboard that would filter the data when the tooltip hovered over a specific segment. For example, if I hovered the tooltip over the bar for Class 1, the data in the chart for Sex would be filtered to show only Class 1 data. The user could

use this filter to show that Class could partially offset the disadvantage of being male, for example. Finally, I added an additional chart that showed a histogram of the Fare data, which I thought would reinforce the message conveyed by the Class chart.

Link: Visualization – Dashboard 1.0

Feedback

At this point in the design, I requested feedback from a colleague. Without telling her the exact message I wanted to convey, I asked her to think out loud as she tried to decipher the main take-away of the visualization. Immediately, I saw that the interactive filter was a major distraction rather than an aid. My coworker spent several minutes trying to figure out what the filter did, and I eventually had to explain its purpose. After the clarification, she spent another several minutes trying to pick apart complex relationships between Sex, Class, and Fare rather than focusing on the simple high-level message. The Fare chart was particularly distracting. After finally mentioning the take-away I was looking for, my colleague asked if any correlation could be seen between survival and age. I initially did not include age information because I thought the data would be too messy if it was displayed as a histogram. (Admittedly, I added an unnecessary histogram in its place – the Fare data). After digesting all this feedback, I came to the conclusion that I had designed the visualization to be more of an exploratory data analysis tool for myself rather than a clear communication tool for a second party.

To address the feedback, the first thing I did was to remove the Fare data, as it was a redundant indicator of social status, and it was distracting. In place of the Fare data, I added data about passengers' age. The data suggested a natural division between children and adults at age 15, so I divided the ages into these two categories. Overall, I thought this simple combination of three graphs showed a much clearer picture than the previous dashboard.

Link: Visualization – Dashboard 2.0

I showed the re-worked visualization to my colleague for a second round of feedback. She quickly came to the conclusions that I was trying to communicate. I noticed, however, that she intuitively wanted to talk about the data in terms of "fractions" rather than absolute numbers. Having the fraction of passengers listed rather than the absolute number of passengers seemed to be more useful. Following a helpful post on Tableau's help page [1], I was able to display percentages over the bars for easy visualization. I decided to keep the absolute number of passengers in the tooltip, but I removed all other extraneous information.

Link: Visualization – Dashboard 2.1

After the final round of feedback, my coworker was able to digest the information in the chart very efficiently. She quickly realized that being in a higher class, being a

child, and being female suggested a higher chance of survival. However, the chart did not tease out the nuance related to Class – namely that Class could override predictions made by Age and/or Sex. Initially, I tried to put back in the filter activated by hovering over a particular bar.

Link: Visualization – Dashboard 2.2

However, I still felt that the filter action was not explicit or intuitive enough. Instead of a dashboard, I transitioned to the story format available in Tableau. I summarized the advantages of Class, Age, and Sex independently, and then on the final page of the story, I showed a bar graph with all three segmentations combined. This final bar chart clearly showed that males were very unlikely to survive unless they were in first class. Females and children were highly likely to survive, unless they were in the lowest class.

Link: Visualization – Story 1.0

Conclusion

At the conclusion of the exercise, I had collected three insights that seemed relevant. While the visualizations presented here were fairly simple and straightforward, I still needed to exert significant effort to make the communication as efficient as possible. Secondly, I needed to consider my audience, which made feedback crucial. Finally, I was impressed by the (relative) simplicity of Tableau. Without requiring significant coding knowledge, Tableau made aesthetically pleasing visualizations easily accessible.

References

[1] "Measure values and measure names", Tableau help pages, https://onlinehelp.tableau.com/current/pro/desktop/en-us/datafields_understanddatawindow_meavalues.html, Accessed Feb. 24, 2018.