Mark Ciganovic
Project 7

https://public.tableau.com/profile/mark.ciganovic#!/
Project7_draft – initial visualization
Project7_final – completed story

Summary

The goal of my visualization is to answer how sex and socioeconomic status affect an individual's probability of survival on the Titanic. The sex of an individual is both very explicit and intuitive in the dataset, since there is a variable titled "sex", and the only two values are either "male" or "female". The socioeconomic status will be reflected in the variable titled "pclass", which is a proxy for socioeconomic status based upon the ticket class, with "1" representing the upper class, "2" representing middle the middle class, and "3" representing the lower class. The two independent variables will be sex and pclass, and the dependent variable will be the probability of survival, and the goal is to create a visualization to present the findings for this question based upon the data.

Design

I chose to use a bar chart to present the findings of my data. The idea was rather straightforward, I wanted to put the independent variables of both sex and pclass in the column section, and the probability of survival in the row section. Calculating the probability of survival was very easy, since there is a variable titled "survived", which has either a value of "0" or "1", in which "0"means the individual did not survive, whereas "1" means the individual did survive. All that I had to do was get the average for survived, and the probability was measured perfectly. I decided to place a label for the probability of survival at the top of the bars, which shows the exact number for the probability of survival at the top of the bars. I also decided to color the bars by the pclass, so that it would be visually more discriminating, and I decided to size the bars by the count of the pclass, so that the width of the bars would show the amount of people in each pclass factored by sex. Finally, if you scroll over a specific bar, it will show the variables of pclass, sex, probability of survival, and count, and there is the option to click on a bar that highlights that specific bar, while dimming the rest. After receiving feedback, I decided to place a count at the bottom of the bars, and to do so, I needed to create a dual axis where I synchronized the axes, and then for one of the axes, I placed the text of the count at the bottom. I also created calculated fields based upon age, so I could factor the bar graphs by age to see how age affects survival. I then made a story that included all three dashboards, which given the data set gives a great deal of information about who survived.

Feedback

The feedback consisted of three points. The first was to change the name of the title to include the word "combination", so that it is specific about sex and pclass influencing probability together, and not separate. The second was to show the exact count of the different groups on

the bars, so people could be exact about the count. The third was a question about what other variables might influence survival besides sex and pclass, and this is why I added age, because with these three, I have presented all the variables in the data set relevant to survival. I would say that I have maximized the information capacity of the data set relevant to probability of survival.

Resources

I did get help on how to create the dual axis, and synchronize the axis, so I could place some of the text at the top, and some at the bottom. Here is the link
https://community.tableau.com/message/673458?et=watches.email.thread#673458