

Data Visualization by Burcu Kurtaran

Summary

This visualization is an animation that shows that the passengers of the Titanic. Specifically, socio-economic status, gender, class, port of embarkment, age, and average fare aboard were explored. There were more men than women, more third class passengers than first class, and the most common age range was between 21 - 30. Next, survival ratio within each of these categories was explored. Unsurprisingly, women, and first class passengers were more likely to survive.

Design

Dataset information

This dataset is taken from Kaggle website. Before I worked with the data I need to clean data and wrangle. Because non-uniform data can corrupt my analysis. I can see that the Age, Embarking port columns are missing entries. In order to fill the missing ages I will use the mean age. In order to fill the missing embarked ports I will use maximum count embarked ports.

The first story link can be found below:

https://public.tableau.com/profile/burcu.kurtaran#!/vizhome/Titanic_410/SurvivedStoryofTitanic

First page: My initial idea is to move progressively through the data from simple to more complex. The first visualization will simply show the survival ratio. I created the variable that show survival with description like survived or died.

Second page: In this page, I explored the socio-economic status of passengers and its relations between average fare, gender and title - is created from passengers' name. Most of passengers have lower socio-economic status and obviously most of them are men. People who are in upper socio-economic status have different titles like Major, Countess, Master etc and their average of fare is the most. Others are generally had title for their gender – Mr or Mrs – and their average of fare is decreasing according to their status.

Third page: Ticket class, port of embarkment and cabin group are likely related to ship. Their relations with survived have been explored in third page. The passengers who bought 1st class ticket survived the most and it could be seen from the graph despite of number of passengers used in columns in bar chart. Cabin groups vs. Survived bar chart could not give us any idea about survived because cabin group information has missing entries. When we look at the port of embarkment vs. survived bar chart, passengers who board from Cherbourg port were survived most in ratios. On the other hand, passengers who board from Southampton port were survived in number of passengers.

Fourth page : "Were some particular passenger groups more likely to survive than others?", in other words "Can some specific set of features push the survival rate upwards ?" Next step is to find a way to answer this question by highlighting some patterns so that my story conveys the overall message to readers. In other words, I have to create a visualization that is answering this question, by giving information between those passengers who survived and those who died. I have seen women have a better chance of survival than adult men. In the title of survived passengers support this question too. But, number of parents and number of siblings could not give us an information about survival.

Feedback

I shared my story with my colleagues in person. I gathered the same feedback from them and tried to follow Udacity questions guideline and here is the abridged response.

General feedback is :

At the second page, you have “Socio-Economic Statu and Title of Passengers” which is bubble graph. But this graph is not precise and it is not easy to understand the difference between bubbles. So you can show the numbers inside this graph or you can change it to the bar graph.

Post-feedback Design

Following the feedback from the 1 feedback, I implemented the following changes:

- I flipped the bubble gragh to horizontal bar chart.

The final story link can be found below:

<https://public.tableau.com/profile/burcu.kurtaran#!/vizhome/Titanic-Final/SurvivedStoryofTitanicFinal>