

Udacity Data Analyst Nanodegree

Project for the Data Visualization

Author: Hande Keskin Sungur

Summary

This visualization is an animation that shows that the passengers of the Titanic. Specifically, 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, survivorship within each of these categories was explored. Unsurprisingly, women, children, 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/hande.keskin.sungur#!/vizhome/TitanicStory1/TitanicStory>

First page: My initial idea is to move progressively through the data from simple to more complex. The first set of visualizations will simply show the breakdown of these passengers in terms of gender, embarking port and ticket class abroad.

Second page: The survival rate is of our main interest. I created the variable that shows survival rate. Then bar chart is easy to understand and good for comparison. So I used a bar chart which is used to show the differences across survival by the passenger class. I use color cue to visualize data of difference gender and survival situations.

Third page: Stacked bar chart uses color to encode the category of data. I use it to present whether one was survived or perished. Stacked 100% bar chart are used to show the comparison of the chance of survival. I used a grouped stacked bar chart which had the passengers first split by passenger embarking port and each port is further split by sex.

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 and children have a better chance of survival than adult men.

Fifth page: I looked at the chart of survival by looking at the first departure port and the average ticket price here.

Sixth page: My purpose was to explore the possibility that the rich were more likely to survive. Finally. Based on the Data available and the visual analysis, we can conclude that a minor female in pclass 1 from port Cherbourg had higher chances of survival than an adult male in pclass 3 from port Southampton.

Feedback

I posted my story in the Udacity website forum. I gathered feedback from 1 person people and tried to follow Udacity questions guideline and here is the abridged response.

The feedback link can be found below:

<https://discussions.udacity.com/t/titanic-tableau-project-feedback/362332>

Feedback 1 : You prepared very detailed story. At first section you have “passenger by class” 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 graph to bar chart.
- I changed gender color. Because multicolored graphics were not nice

The final story link can be found below:

https://public.tableau.com/profile/hande.keskin.sungur#!/vizhome/TitanicTableauStory_final/TitanicStory