Titanic Story

By Jiffin James

Story_v1=
https://public.tableau.com/profile/jj2272#!/vizhome/TitanicDataStory/Story1?publish=yes

Story_final= https://public.tableau.com/profile/jj2272#!/vizhome/TitanicDataStory2/TitanicStory

Summary

On April 15, 1912, during her maiden voyage, the Titanic sank after colliding with an iceberg, killing 1502 out of 2224 passengers and crew. This sensational tragedy shocked the international community and led to better safety regulations for ships. One of the reasons that the shipwreck led to such loss of life was that there were not enough lifeboats for the passengers and crew. Although there was some element of luck involved in surviving the sinking, some groups of people were more likely to survive than others, such as women, children, and the upper-class.

Design

- The first story point shows information on the various ports of embarkment. I converted the type of variable (Embarked) into a geographical one and plotted those ports on a map using latitude and longitude information from Wikipedia. I used stacked bar plot to show the total number of passengers from each of these ports and the survival statistics for each of these ports. From the visualization it is clear that all the ports exist in Europe (United Kingdom, France, Ireland).Most of the passengers embarked from Southampton but Cherbourg had better survival rate among these ports.
- Reason-Maps can help us visualize location of certain places easily. Example-Most of us might not know in which country or region ports like Southampton, Cherbourg or Queenstown (now Cobh) exist, but using map we can immediately visualize their location. I used stacked bar plot to show the survival statistics i.e. how many of the passengers survived from each of these ports.
- The second story point has information regarding survival rates among male and female passengers along with the age distribution. From the Distribution by Gender plot it is quite clear that most the passengers were male but females had better survival rate. Similarly from the Age Distribution plot it is clear that most of the male and female passengers were in their early 20s.Another point which can be easily seen is that survival rate for female passengers was very high across all the age categories.
- Reason-I used stacked bar plot because it helps others understand the differences between classes easily. Here too, I used Survival status to visually encode information (color).
- The last story point has information regarding socio-economic class and survival rates. First plot shows the average fares of each class(88 .05 for 1st class and 13.23 for 2nd class).The second plot explores the survival rate among these classes and passenger gender. It is quite clear that almost 90% of female passengers from 1st class survived. Finally I included a table to show the exact figures.

Feedback

- It was difficult to estimate the count of various variable because of poor use of labels .
- User had less control over the plot because of poor use of filters.
- Color encoding of Average fare for each class(Story point 3) was confusing because red color was used to represent class 3(socio-economic status) but from the beginning of the story red color was used to represent the passengers who couldn't survive.

Actions

- Included filter control for most of the plots.
- Rectified Ambiguous use of color.
- Included text labels for most of the plots.

Resources

- https://en.wikipedia.org/wiki/RMS_Titanic
- https://www.tableau.com/learn/training
- https://www.kaggle.com/c/titanic