Titanic Data Story

Anthony “Ant” Cassetta

Summary:

My visualization comes from a data set of 889 records of passengers on the Titanic. This visualization will demonstrate percentages of survival based on demographic information such as gender, point of origin and class.

First Draft:
https://public.tableau.com/profile/antcassetta#!/vizhome/Ant_Titanic_Project_v1/Story1?publish=yes

Final Draft:
https://public.tableau.com/profile/antcassetta#!/vizhome/Ant_Titanic_Project_final/Final?publish=yes

Design:

My design consists of Maps, Bar charts and Bubble charts, I included legends for Gender (Female/Male) and class (1st, 2nd, 3rd). These legends provide

Initial Design:

My initial story was 5 pages long built on 5 dashboards, the total number of sheets among these dashboards is 8. The first four pages contained a map, corresponding to a specific break down passenger demographic data which was reflected in either a bar or bubble chart. In the case of a bubble chart I encoded size information to visually display the difference between regions or class.

Example: Page one had a map of the three ports of origin The United Kingdom, Ireland and France respectively. Below the map was a bar chart displaying the count of Female and Male passengers of each port. Upon clicking one of these ports, say Ireland, the bar chart would filter exclude other ports displaying only the bars representing Ireland's data. Allowing the user to more easily focus on specifics of the data. Each of the first four pages followed this pattern. The last page contained only summary data about the final findings in bar chart form.

The map was chosen as a clear visual for the three point of origin for each passenger. The size of each country also grants perspective when compared against each quantity of passengers recorded. Example: Ireland is a small country and it had the smallest group of passengers.

Bar charts are a good choice for population counts and survival percentages as they are used to compare values and in understanding the distribution of data. They easily summarise the data sets in a single visual. This is why they are used in this project.

The bubble charts were used in place of bar charts in the case of splitting data by class in order to make use of visually encoding size according to count. I wanted the user to have a very clear visual understanding of which class had the most passengers. I also felt it was necessary to use a different chart type for a different aggregation in order to avoid monotony. If every chart were of the same type for different aggregations it is often harder to pick out the differences in the data.

The colors chosen on the map are meant to differentiate countries a selection may be made on. Each country is a unique color to indicate unique data and the rest of the map is a one flat color indicating no date is associated. The colors chosen for gender are the complimentary colors blue and orange. These colors are double encoded along with axis labels in order to be more user friendly. It's very easy to quickly pick out what bar is female(blue) and what bar is male (orange).

Feedback:

After Asking a friend who is familiar with data science for feedback I received the following:
"Maps are cool in visuals I like them, but in this case they're not adding much. What do the colors of the countries mean, anything?"
"Gender label and color double encoding is more UI friendly, but not adding information about the data. I can see why you include it but you don't have to."
"What's the point of the story, who survived?"

Design Revision:

After receiving feedback I realized a few things, my story was too long and ambiguous.

I whittled my story down to two pages. A page introducing the dataset with total counts of passengers from a specific region and and class and a summary page containing survival statistics in the form of bar charts.

In regard to the critique on the map. I explained that the colors were just meant to indicate and differentiate clickable spots on the map. Reducing the number of pages and therefor maps also keeps the excitement of the map for the user without compromising the story with excess clutter.

In regard to the critique on the double encoding of the gender via label and color distinction. I kept this design choice in place. I think the addition of consistent distinct color makes it easier for a user to quickly observe the trends in the data throughout the different plots.

The critique about the overall point of the story was handled by reducing the number of pages to the intro and the summary. This makes a very brief yet clear story out of all passengers on the Titanic, the one feature that made you more likely to survive was being female.

Resources:

I Choose Hue
http://tools.medialab.sciences-po.fr/iwanthue/

Paletton
http://paletton.com/#uid=33i0u0kllllaFw0g0qFqFg0w0aF