

## Project: Create A Tableau Story

### 1. Dataset Used

The dataset chosen for the project was the Titanic Data available from the Data Set Option link provided. This dataset contains demographics and passenger information from a subset (891 records) of the 2224 passengers and crew on board the Titanic. The link to an overview and description of the data is provided in the Resources section below.

### 2. Summary

The objective of the visualization was to test out two complementary hypotheses about what factors would have enhanced or diminished chances of survival for the passengers and crew when the Titanic struck an iceberg on her maiden voyage and subsequently sank. We postulated that there were two factors that would have determined chances of survival:

- The 'women and children first' policy followed during evacuation to the lifeboats
- Awareness of the danger and the time it would have taken to reach a lifeboat from where one was stationed

The Tableau story includes elements of the background history. Two dashboards were created to illustrate each hypothesis, and a third dashboard showed how the two factors would have interacted and created further nuances to the chances of a person's survival (e.g. a first-class passenger who was male stood a better chance of survival than men in general).

### 3. Design – initial

The link to the initial design of the visualisation is given below:

<https://public.tableau.com/profile/saugata.ghosh#!/vizhome/TitanicSurvivorship-Initial/TheTitanic>

The following elements of the initial design are mentioned below:

- a. Background history – A deliberate decision was made to incorporate the background history to the sinking of the Titanic through interesting visuals of the ship, a map of where it sank and even interesting trivia pertaining to the survivors. This was woven into the narrative of the Tableau story to provide a context especially for those who might be unfamiliar with the marine disaster.
- b. Chart type – Pie charts were used almost exclusively in communicating the findings. Since there were only two categories (Died/Survived) for which percentages had to be shown, it was felt that a pie chart would not be misleading. In order to maintain consistency and not distract the viewer too much from the findings by showing charts of different types it was

- decided to stick with the same type of chart, except for the third interactive dashboard where a bar chart was used to show percentages of passengers travelling in different classes.
- c. Visual encoding – Initially the colours green and red were used to show percentages of those who survived and those who died respectively. This was changed later after receiving feedback (please see below). The proportions of passengers in different classes was shown in greyscale.
  - d. Legends – Explanatory legends for all colour encodings were added to the visuals.

#### 4. Feedback

Feedback for the initial visualisation was sought in the Udacity Discussion Forum. One feedback received from James Cooper is given below.

Thanks for sharing! Here's my comments, hope this is useful to you.

1. What do you notice in the visualization?

Multiple individual pie charts to show survival percentage based on variables. Heavy use of pie charts. Use of Red and Green inadvisable due to colour blindness.

2. What questions do you have about the data?

How was it gathered? What errors might be present?

3. What relationships do you notice?

Massive drop in survival rates for third class women and children despite supposed "Women and Children First" policy.

4. What do you think is the main takeaway from this visualization?

That survival rates differed for different subgroups in the Titanic.

5. **Is there something you don't understand in the** graphic?

I did not understand what Cabin Group meant as there wasn't a map showing how the cabins were split up around the ship.

When you say that Cherbourg passengers "may have been more advantageously placed", could this not be due to one of the other variables, for example small families / first class passengers/ more women and children?

I enjoyed reading the story however I found myself wanting to delve deeper into the interrelationships within the data. For example if when I moused over the piechart for Cabin Group A, it could tell me what percentage of these passengers were 3rd class.

I wasn't sure what your conclusion was / what you felt was the most interesting part.

## 5. Design Changes – Final

The link to the final design of the visualisation is given below:

<https://public.tableau.com/profile/saugata.ghosh#!/vizhome/TitanicSurvivorship-Final/TheTitanic>

- a. The background story elements were retained as it received a positive feedback
- b. A story point giving my conclusion regarding the hypotheses was included as desired.
- c. Consistency in using pie charts was retained for reason given in previous Design section. The added versatility of pie charts in indicating relative size of different groups (apart from percentages of died/survived within each group) was taken advantage of.
- d. The colour coding for 'Survived' was retained as green. The colour coding for 'Died' was changed to a darker shade of blue to take into account the feedback about colour-blind viewers, while still allowing the colours to signify some meaning on their own.
- e. Taking feedback into account the dashboard for 'Proximity to a lifeboat' was changed. An attempt was made to relate the class of travel of passengers to where they embarked at and the kind of cabin they stayed in. Cabin types were related to the corresponding ship decks as given in the deck plan of the Titanic (please see Resources section).

## 6. Resources

1. Dataset Description: <https://www.kaggle.com/c/titanic>
2. Information about the Titanic:
  - a. <http://www.titanicandco.com/inside.html>
  - b. <https://www.encyclopedia-titanica.org/titanic-deckplans/>
  - c. [https://en.wikipedia.org/wiki/RMS\\_Titanic](https://en.wikipedia.org/wiki/RMS_Titanic)
  - d. [https://en.wikipedia.org/wiki/Margaret\\_Brown](https://en.wikipedia.org/wiki/Margaret_Brown)
3. Feedback from Udacity forums:  
<https://discussions.udacity.com/t/feedback-requested-titanic-dataset/309457>

