

Create a Tableau Story

Project By: Urmay Shah

Dataset: Titanic Dataset

Links to the first version of your story (e.g. story_v1) and final version of your story (e.g. story_final)

- Link to First Version of the story: [1st Version of the story](#)
- Link to Second Version of the story: [2nd Version of the story](#)

Summary: brief description of the visualization and the main story or findings conveyed.

Dataset Parameters

- After having a brief look at the data frame, I have listed all the parameters with their type
 - Survival: Survival (0 = Not Survived ,1 = Survived)
 - Pclass: Passenger class (1 = Upper class ,2 = Middle class , 3 = Lower class)
 - Name: Name of the passenger
 - Sex: Male/Female
 - Age: Age of the Passenger
 - Sibsp : No of siblings /spouses aboard the Titanic
 - sibling = brother,sister,stepbrother, stepsister
 - spouses = husband, wife
 - Parch: No of parents and children aboard the Titanic
 - Parent = mother and father
 - Child = daughter, son
 - Some children travelled alone or with nanny so they have parch = 0 for them
 - Ticket: Ticket no of passenger
 - Fare: Fare of the ticket
 - Cabin: Contains the cabin number
 - Embarked: Contains port of Embarkation. Contains three value.(C = Cherbourg, Q = Queenstown, S = Southampton)

Summary

1. While exploring the data only by gender, it is found that females were given higher priority as female survival rate(68.13%) is quite higher than male survival rate(31.87%).This comes out to be a ratio of almost 1 : 2.
2. Survival rate per Pclass gave us the insights that Survival rates of various Passenger Class. Here we see that Passenger Class 1 had the most % of survivors (62.96%), whereas Passenger Class 3 had the most number of casualties. Survival rate for class 3 is only (24.24%). The order of survival rate per Passenger Class is Passenger Class 1 > Passenger Class 2 > Passenger Class 3.

3. Although females from class 3 were also given higher priority but not as compared to class 1 & 2. The survival rate of females from higher class is almost (90 to 96%) but for class 3 it is only 50% concluded from Passenger class exploration. It is also noticed that males of class 1 were given the highest priority amongst all men of all classes.
4. While exploring from port of embarkation vs survival, we observed that Port C saw the most survivals while Port S and Q saw the least survivals for both the sexes.
5. An interesting observation found that all children whose age is less than 10 and from class 2 were able to survive. It is also observed that child from class 3 has most % of casualties.
6. On an interesting note it is found that the people who spend more money on ticket (5 Highest Fare) mostly able to survive. And most of aged male (Age>60) were not able to survive.

Design: explain any design choices you made including changes to the visualization after collecting feedback

Design and visualization choices:

- Before starting my exploration and visualization process, I had a brief look at the data. Then I decided the flow of the story.
- I had to group Age, Passenger class and Embarked port data. For Age, I choose grouping over binning because I wanted to divide my data into non-equal bins. Also for embarked port, I also created a set so that null data can be removed from the analysis.
- I have tried to incorporate the best visualization practices as per my knowledge, like proper choice of color and color set, form and special positioning of the data points and tried to maintain a high data-to-ink ratio in my story. I have given different color variation for each analysis so that reviewer can easily connect.
- I have tried to keep each and every visualization simple and clean so that its easily interpretable.
- **For the suggestions and issues pointed out by the reviewer:** I have considered the reviewer's suggestions and I overcame the issues by adding the text to story toolbar.

Summary

Feedback: feedback received from others from the first sketch to the final visualization.

Reviewer Name: Anukriti Arya

Q 1 What do you notice in the visualization?

A. Data analysis on titanic data is done via data visualization techniques.

Q 2: What questions do you have about the data?

- A. 1. Who amongst the two-gender had highest chance of survival
2. What factors affected passenger's survival?

Q 3: What relationships do you notice?

A. Relationship amongst embarked port and survival is depicted. Also, survival per age is show. Survival per fare category is also depicted.

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

A. According to the visualization, it was the females who had a higher chance of survival despite of age, embarked port and fare category.

Q 5: Is there something you don't understand in the graphic?

A. Depiction is well predicted and i understood everything. I have a suggestion though. Please try to encorporate the text that you have entered in the sheet to the story line text where you have written the title because i was not able to read the graphs properly and they seem to overlap with the graphs. Plus, please avoid red and green. as they are not colorblind friendly Colors (though look vibrant in the graph: D)

Rest all is great! Good work.

Resources: list of Web sites, books, forums, blog posts, GitHub repositories etc that you referred to or used in this submission (Add N/A if you did not use such resources).

I have referred to only tableau resources, which are:

- [Data Visualization in tableau, Udacity course](#)
- [Create a story](#)
- [Publishing Data sources and workbooks](#)
- [Hide or Unhide fields](#)
- [Multiple Graphs in one tableau Worksheet](#)
- [Formatting at the worksheet level](#)