## Summary

I am exploring the Titanic dataset to better understand who was more likely to die. I use 4 common features in the analysis: the person’s gender, age, social class, and whether they were traveling with someone or alone. I have found that lower class males travelling alone perished the most. The visualization I created allows users to explore these 4 common features too.

## Design

Initially, I did not start with a focused question. Although I am familiar with the Titanic dataset, my workbooks and dashboard were scattered. I used the default plots and colors. I felt fairly clumsy in the tool. My lack of experience showed in my first plots. Simple things like scroll bars or aliasing the survived values didn’t cross my mind.

As I developed a question to answer the plots became more focused and creative. I also took time to cleanse the data to the analysis was more meaningful and so it would mean something to a reader.

## Feedback

My initial feedback came from co-workers who use Tableau and are experienced in visual analytics. They explained some simple things like adding a title and maybe presenting a question to the user of what the dashboard is trying to achieve. They suggested to customize the colors I chose to make the visualization special and meaningful. I also received feedback to take a look at Tableau Public to see if I could gain inspiration from others.

Below are the design and feedback iterations with more explanation about initial considerations and then implementing changes based on feedback.

### Design and Feedback Iterations

**Version 1**

<https://public.tableau.com/profile/tania8454#!/vizhome/TitanicSurvialOriginalVersion/Dashboard1>



In the initial dashboard design, I admit I was all over the place. I was getting familiar with the tool and in a way seemed to throw out all analytic learnings.

I chose bar chart for Gender and Age, because I wanted to see the age distribution for each gender and whether they survived. I did not see the histogram option at first and thought this was how distributions were plotted in tableau. It did help me see the large number of null age values and this was something a review also provided feedback to revise.

Next, I was curious to see if the amount someone paid for their ticket indicated whether or not they would survive. For instance, would it be possible that passengers who paid more had better access to life preservers or happened to be dining or relaxing in entertainment halls closer to emergency exits. I also thought maybe age could factor into the ticket price and that the data should be seen from this angle too. These questions led me to produce the Age and Fare graph. It seemed like middle aged men paid more.

The Class and Fare graph was another variation of how much someone paid and their social status. I was looking for outliers to see someone paid more for their ticket than the others. This was not very successful, but the dashboard seemed blank without it. I added it for visual variety.

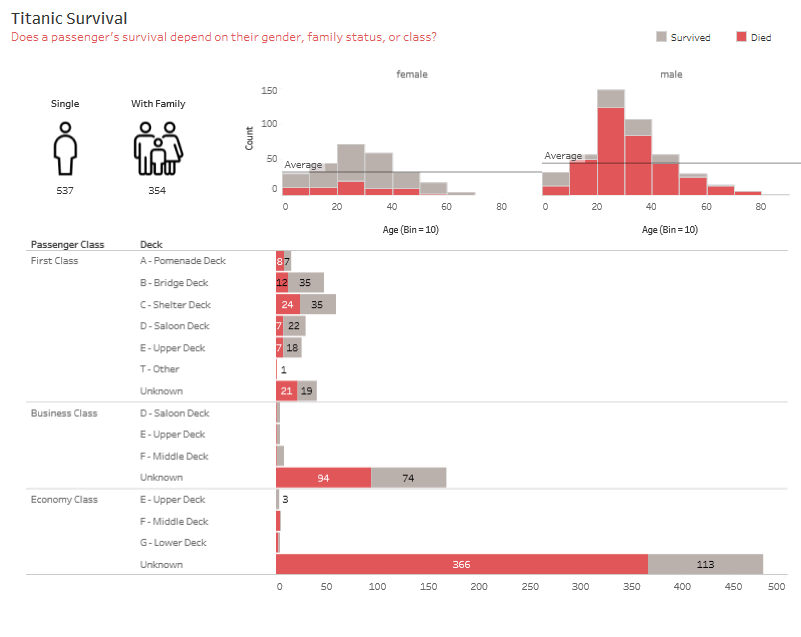
Family status was an included as an exploration of whether people travelled alone or with someone. If more females or men travelled by themselves. Mostly men did this, but I was surprised that a fair number of women did this too.

**Feedback, Version 1**

**I received the following feedback on my initial design:**

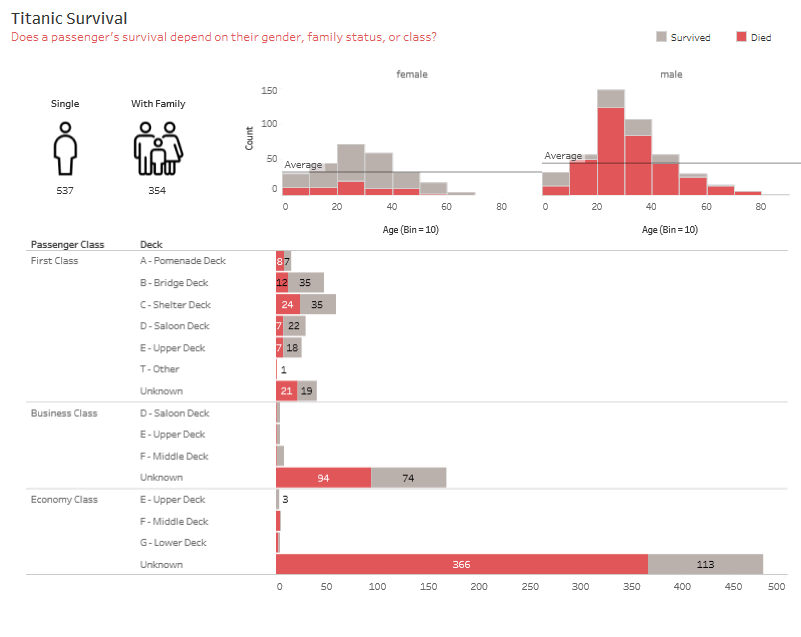
* Enhance colors and make them more meaningful, don’t use the default colors.
* Focus the analysis and include a title and/or the question you are trying to answer.
* Add more graphics or visuals.
* Modify the variables so they mean something to the reader.
* Try using a histogram for the age and reduce scrolling.
* It’s not clear the analysis you are trying to do with the class and fare box plot.
* Take a look at some of the titanic viz’s already posted to tableau public for inspiration.
* Decide how to treat null values.
* Add actions

**Version 2**

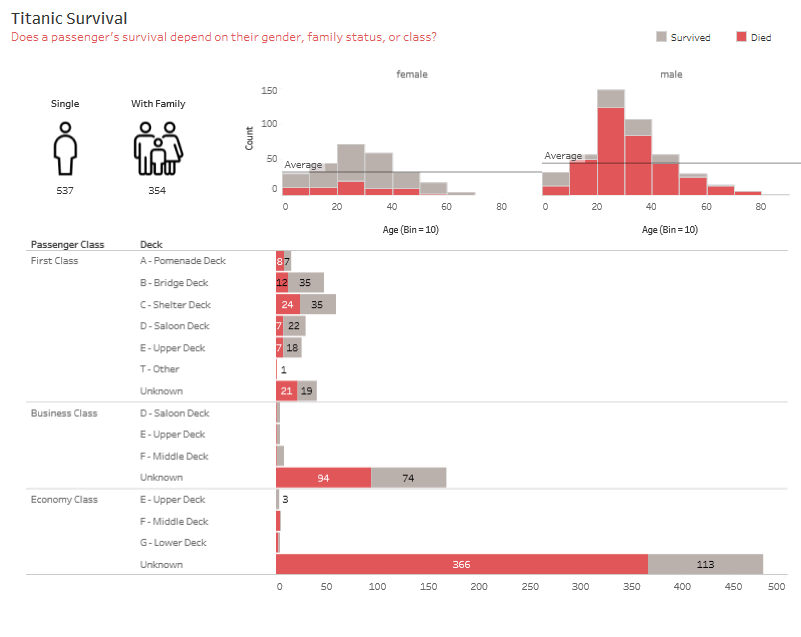


**The initial dashboard needed lots of rework and I took a lot more time and care in version 2. I concentrated on basic design principles like color, names, and plot placement. I removed the chart junk as well. I took time to clean up the data. I overhauled everything about the first version and decided most of the plots were not useful.**

**I thought of the questions I was trying to answer - who ended up dying and why; was it their gender or age; did it depend on whether they were with family; were first class passengers taken to the lifeboats faster? Eventually, I landed on a more concise question. I targeted the visuals to these questions.**



**I aliased survived values to died or survived as this seemed more useful than 0 and 1. I also made a calculation to determine if someone was single or with family using Parch and Spsp; also more useful than numbers and the abbreviations.**



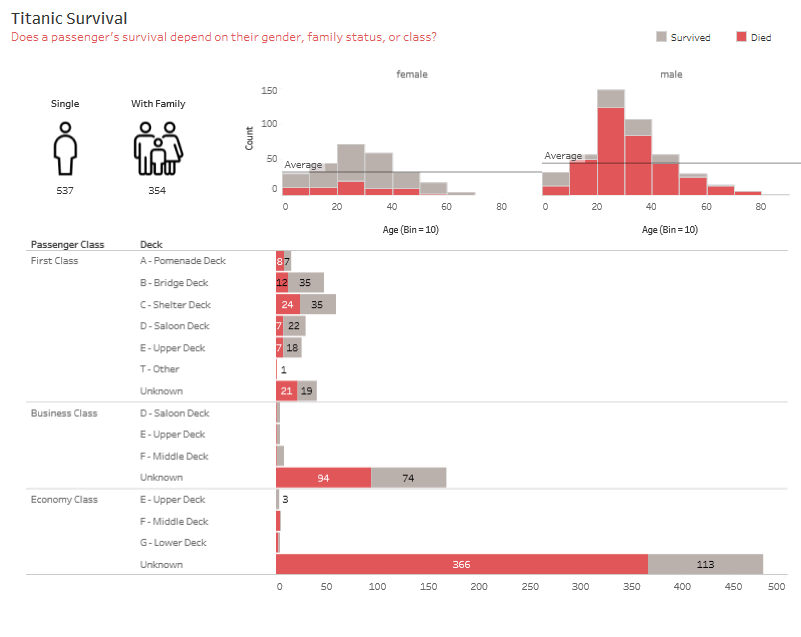
**I started with a lot of bar charts and scatter plots, but ended up (finding) using the histogram to visualize gender and age. In this plot, I treated null ages as 0’s. I liked the 10 bin size because I thought it would be easier for the users to count in the absence of a number for every bar.**

**I tried a lot of different colors, like tones of blue, but decided I liked the contrast of red for those that didn’t survive. Since I wanted to understand who dies, this seemed like a great way to add focus. I used grey for the other color so it was muted and didn’t visually compete.**

**I thought the average line was a nice thing to add to see the average age.**

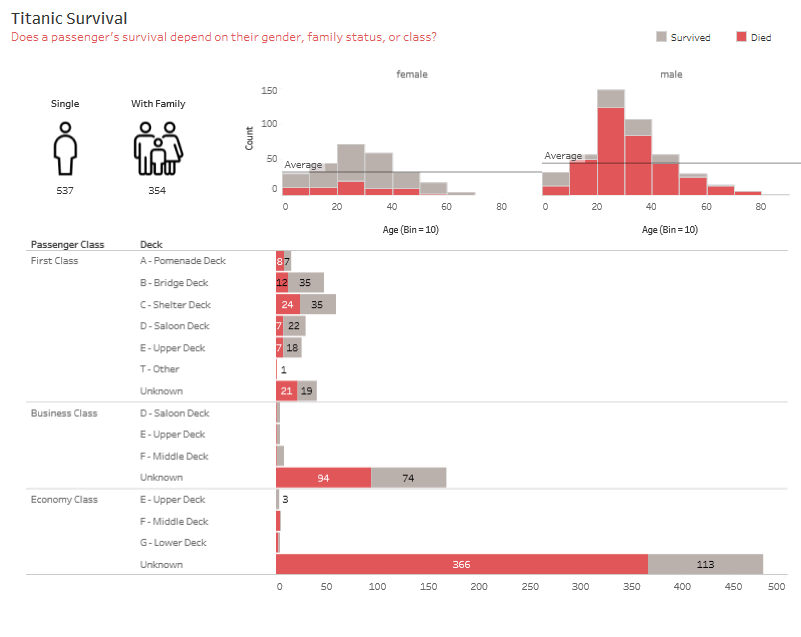
# 

**The idea of having an overall KPI seemed like a good one, so I chose whether someone was travelling alone or with family since I had a visual showing gender, age, and survival. Initially, I did not include pictures, and it seemed really bare. I added these for some visual texture.**



**I liked the idea of continuing to look at class or where someone boarded as another factor for survival. I grouped the tickets by their letters thinking this would indicate where they were located. I found on Wikipedia the translation between the letters and decks. In the event someone wanted to look at the details, I provided a hierarchy using deck and ticket which users can expand/collapse.**

**I played around with the orientation and liked this being horizontal. I added numbers to readers could see the details if they wanted to.**

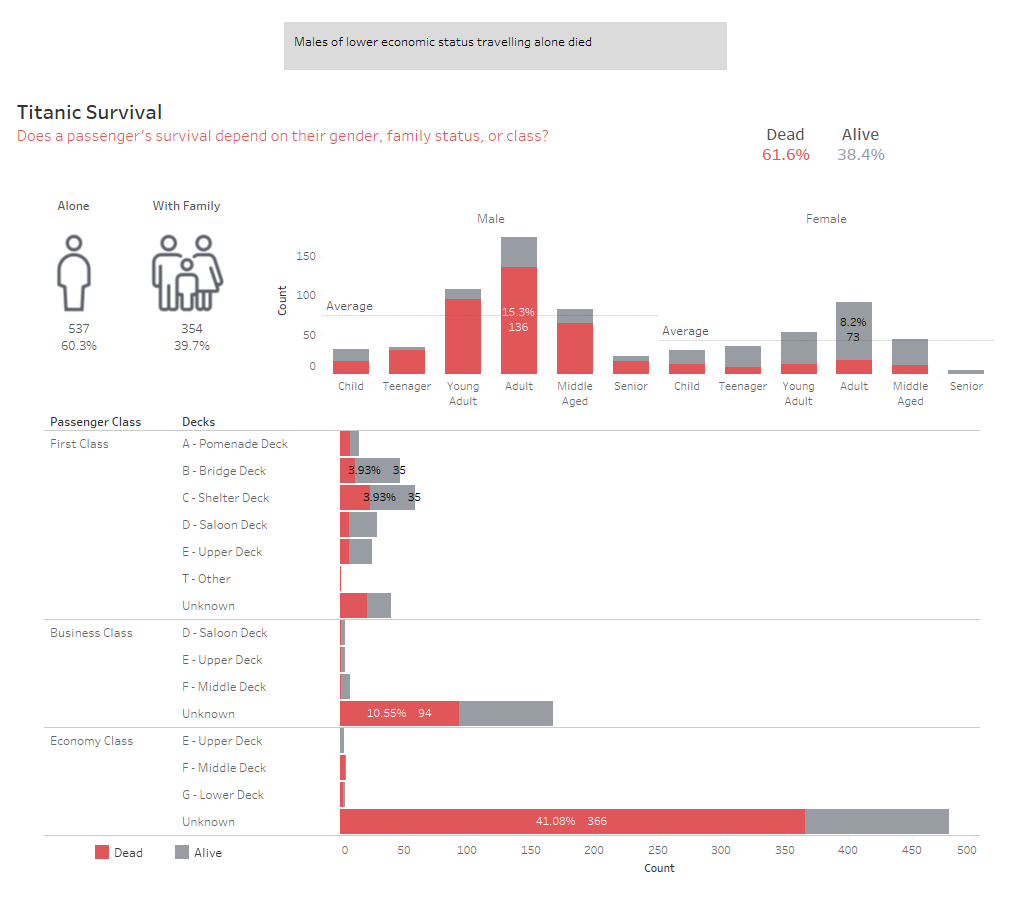


**Feedback, Version 2**

**I received the following feedback on the second iteration:**

* Add percentages to see overall impact instead of the raw numbers
* Include survival rates
* Remove Deck drilling, doesn’t seem useful
* Try age groups rather than bins
* Improved, missing the answer to the question
* Add actions

**Version 3**



I mainly decided to concentrate on readability and answering the main question. I added an overall KPI for survival percentages to the chart and kept same colors also helped keep the dashboard cohesive.

I changed the aliases again to keep the language concise. I’m not sure if it really added much, but I liked the direct words. I also cleaned up some of the aliases, so all labels appeared capitalized.

Per feedback of this iteration, I removed the drilling on Deck.

The numbers seemed busy especially once I added percentages, so labelled the maximum values only.

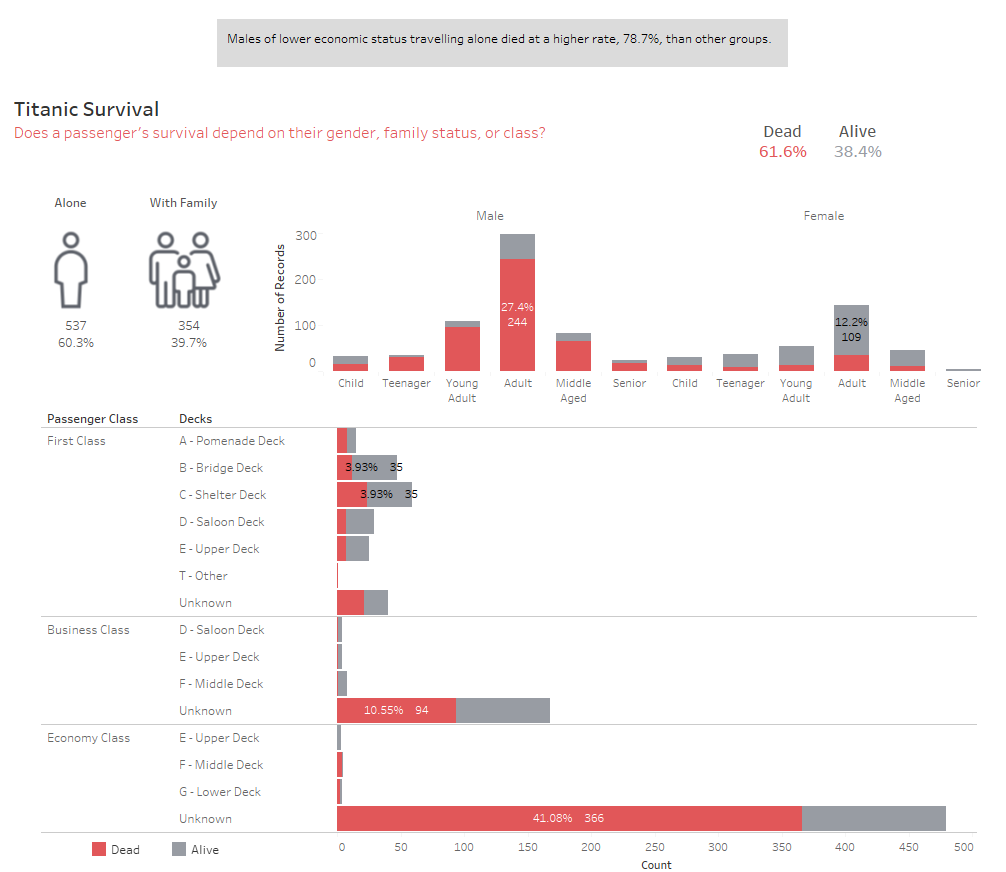
I tried age groups rather than bins, so now we can better see whether someone was a child or a senior. The feedback on the age bins was that it seemed impersonal and it was easier to understand who someone was with an age label than number.

I added actions! Each chart now filters the other charts and it was easier to answer the question about who died more.

**Feedback**

* Clean up tooltips
* Include overall insights to guide the readers.

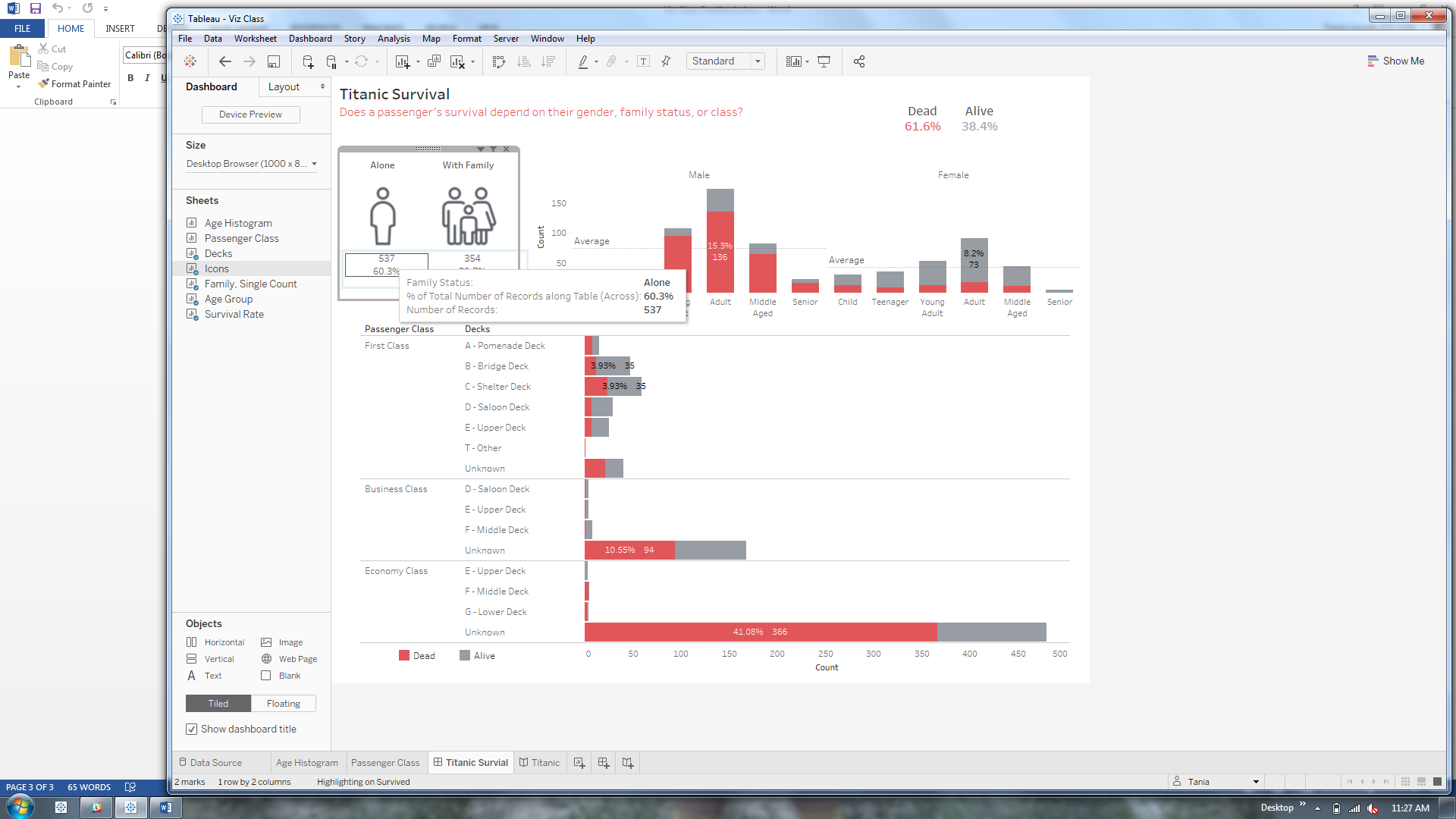
**Version 4**

<https://public.tableau.com/profile/tania8454#!/vizhome/TitanicSurvialVizClass/Titanic>

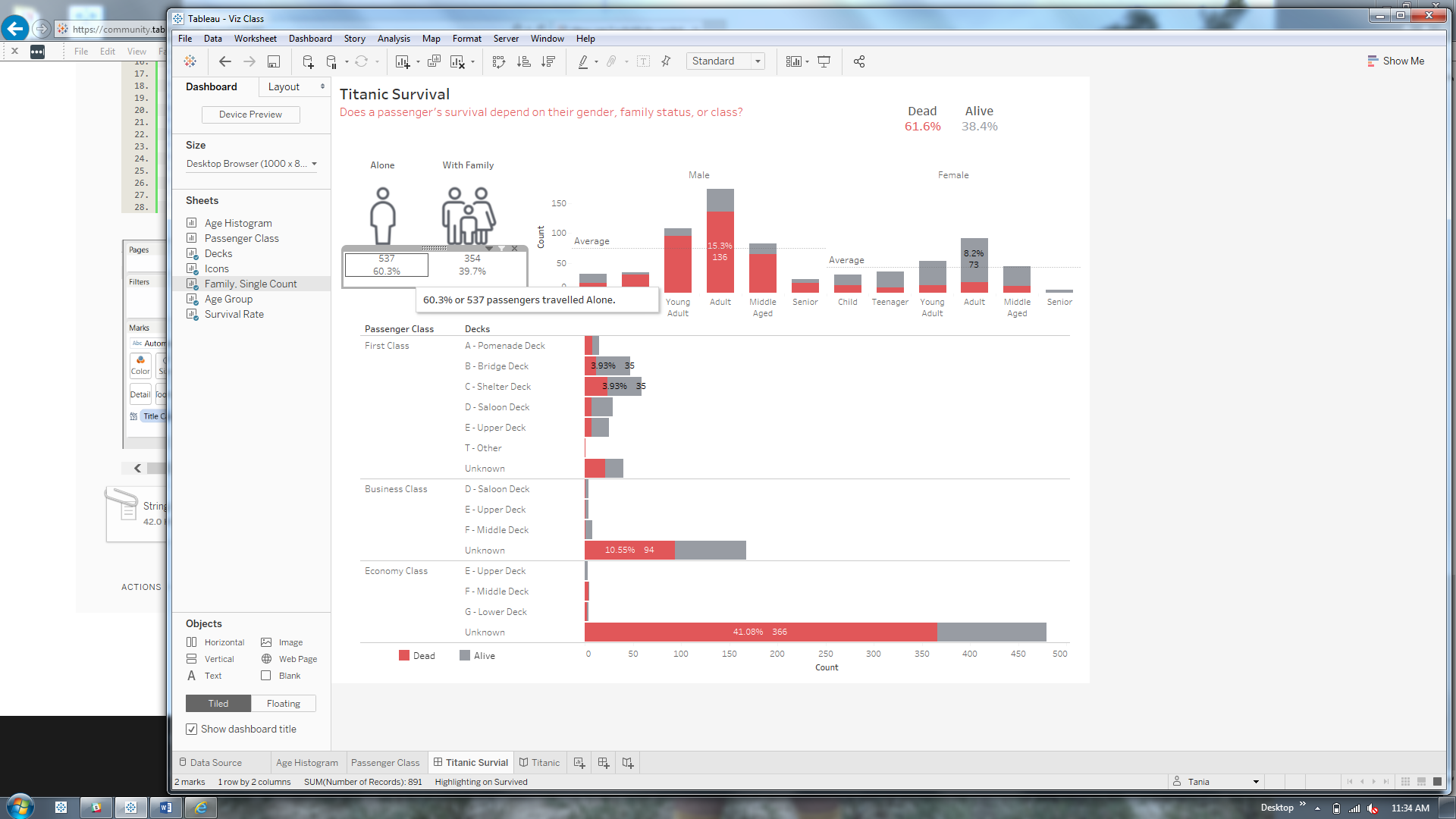
I removed the average line. I realized this was showing the average count, not the average age. I also changed nulls from 0 age, which would get bucketed into the child age group, to adult based on the true average age for the genders.

All of the tooltips were defaults and I hadn’t considered changing these. They seemed to provide the right information. However, some feedback I received was about why we couldn’t make these into sentences that would be easier for a regular person to read. I experimented with that idea and decided on these tooltip revisions.

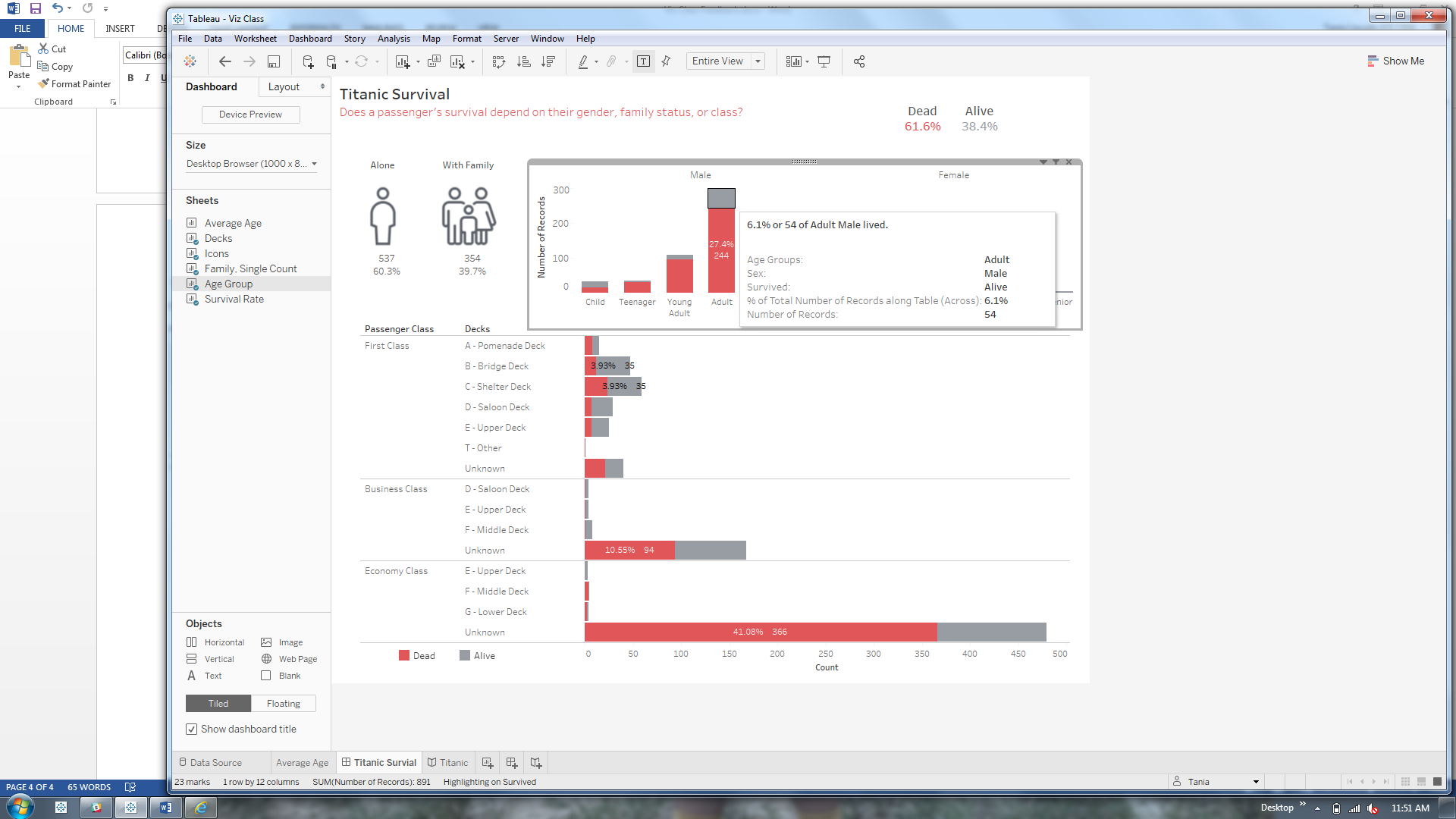
Before:



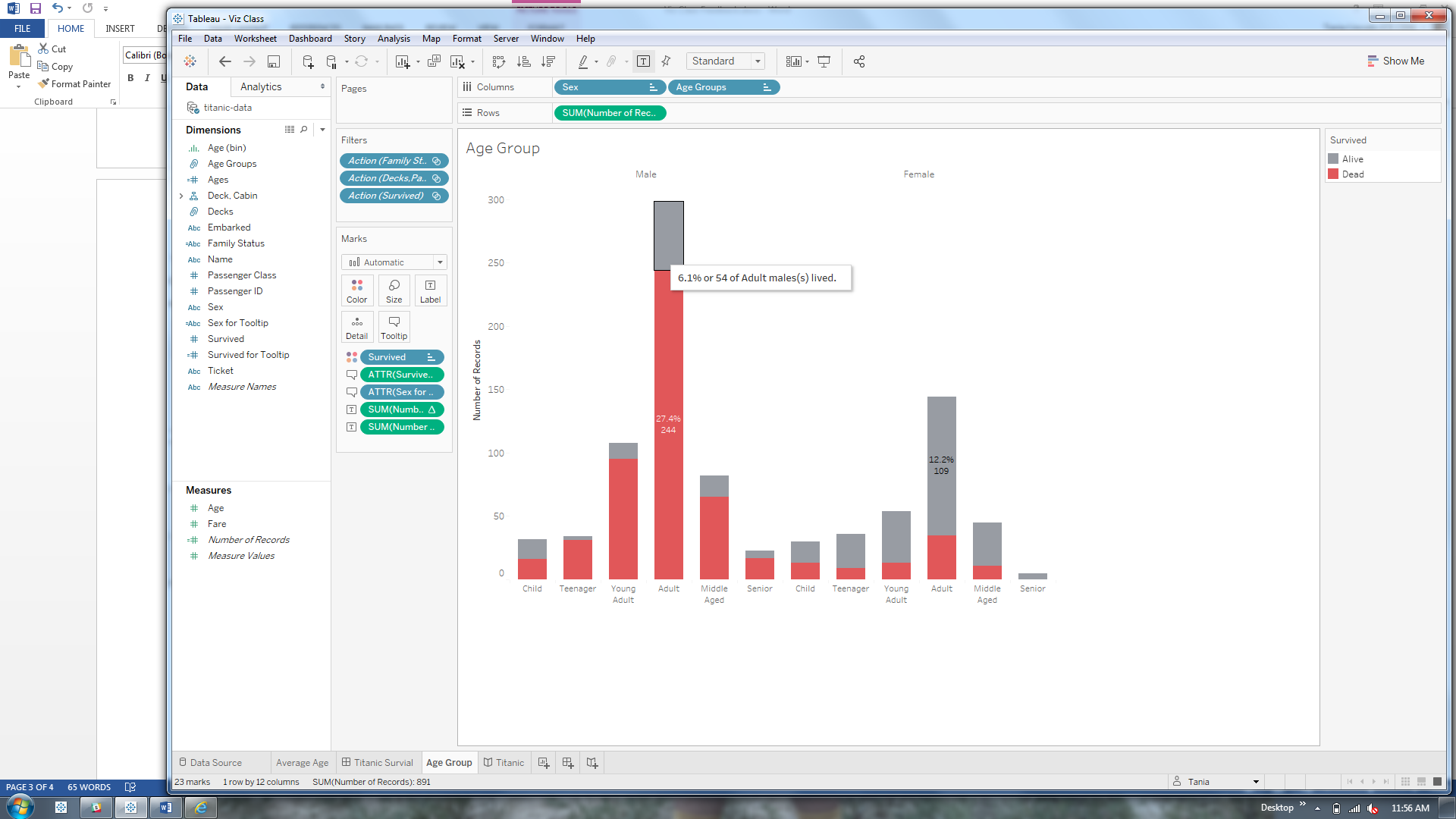
After:



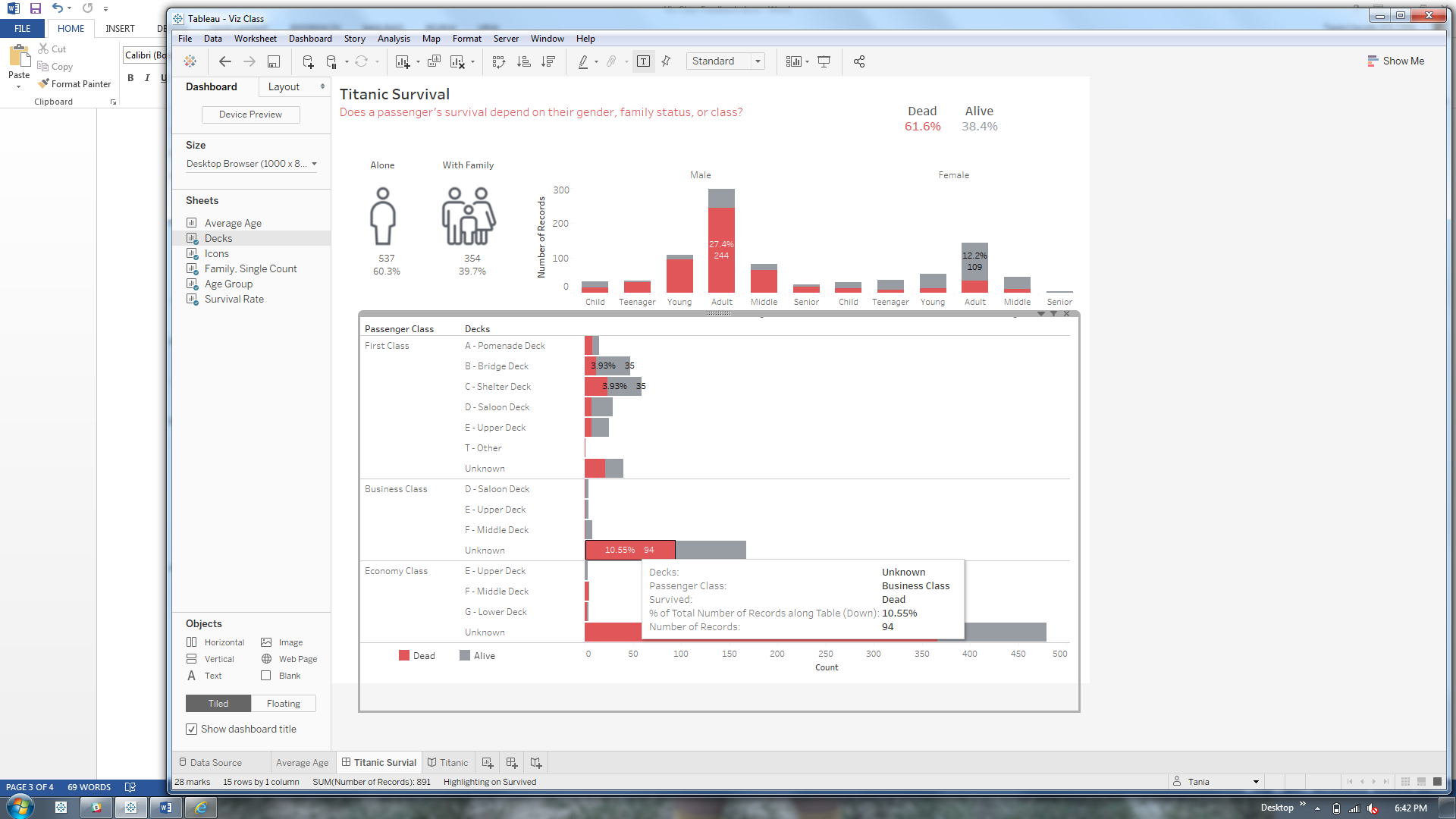
Before:



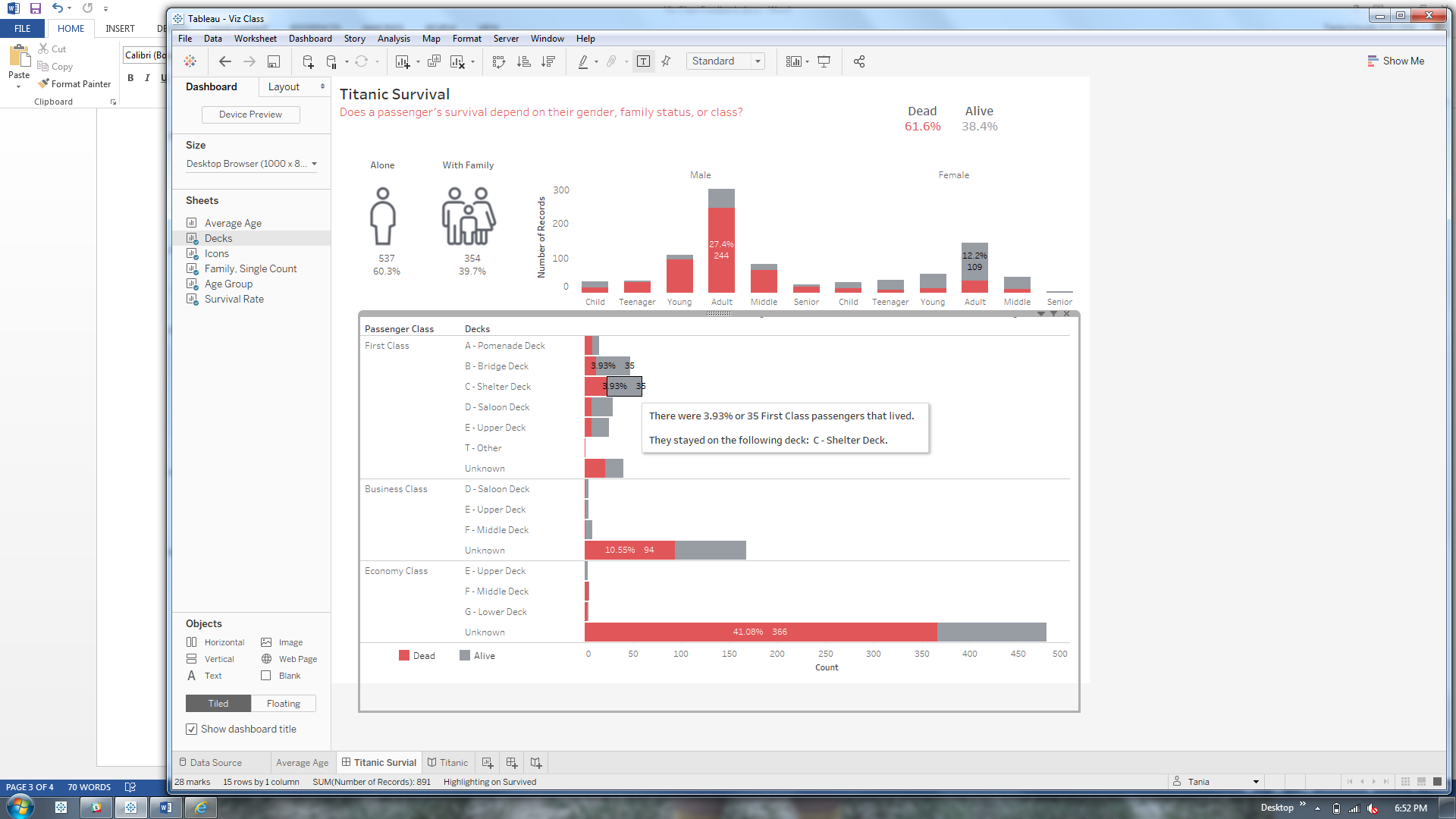
After:



Before:

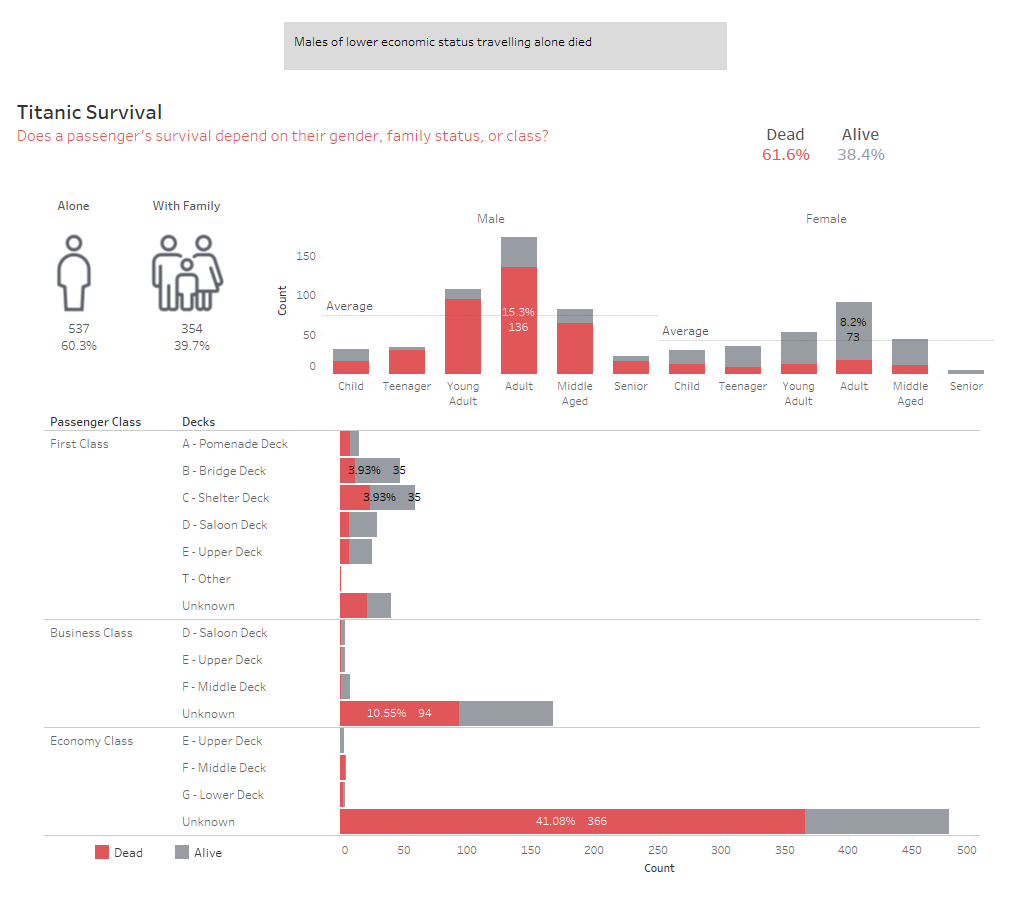


After:

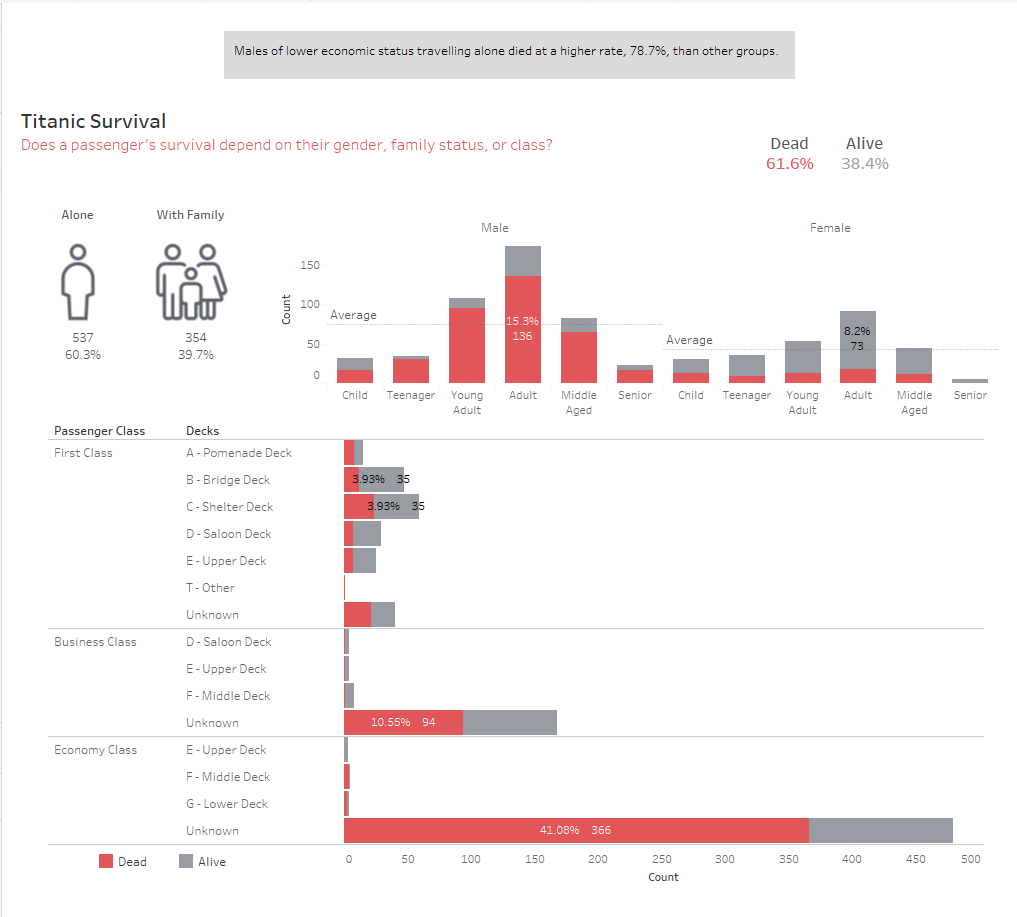


The story headed was also generic. I modified the language to match the dashboard and included the percentage rate.

Before:



After:



**Feedback**

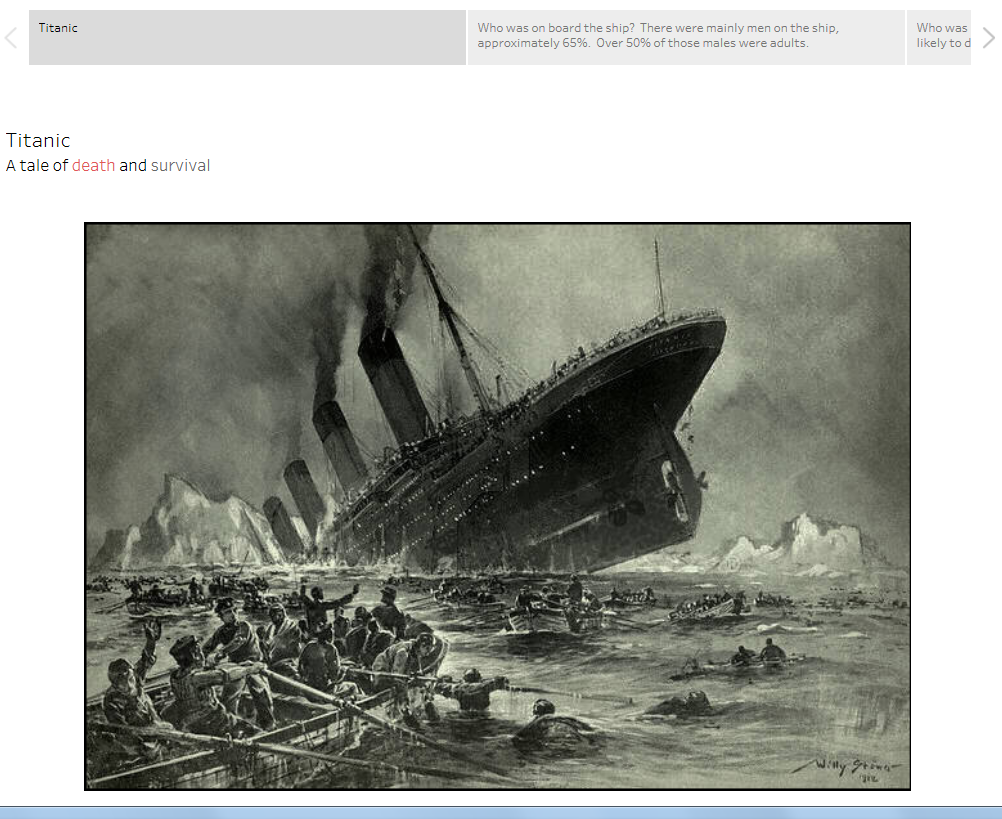
I received the following feedback:

* Use Tableau story to foster communication between reader and viz

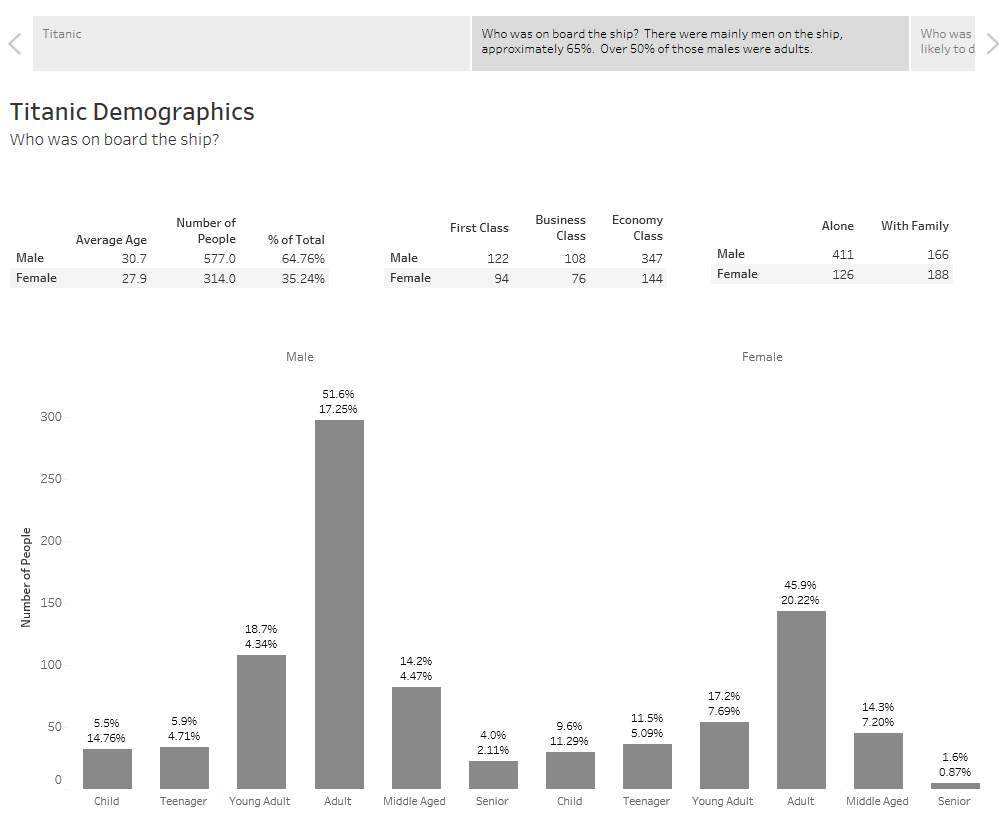
**Version 5 (final)**

<https://public.tableau.com/profile/tania8454#!/vizhome/TitanicStory_5/Titanic>

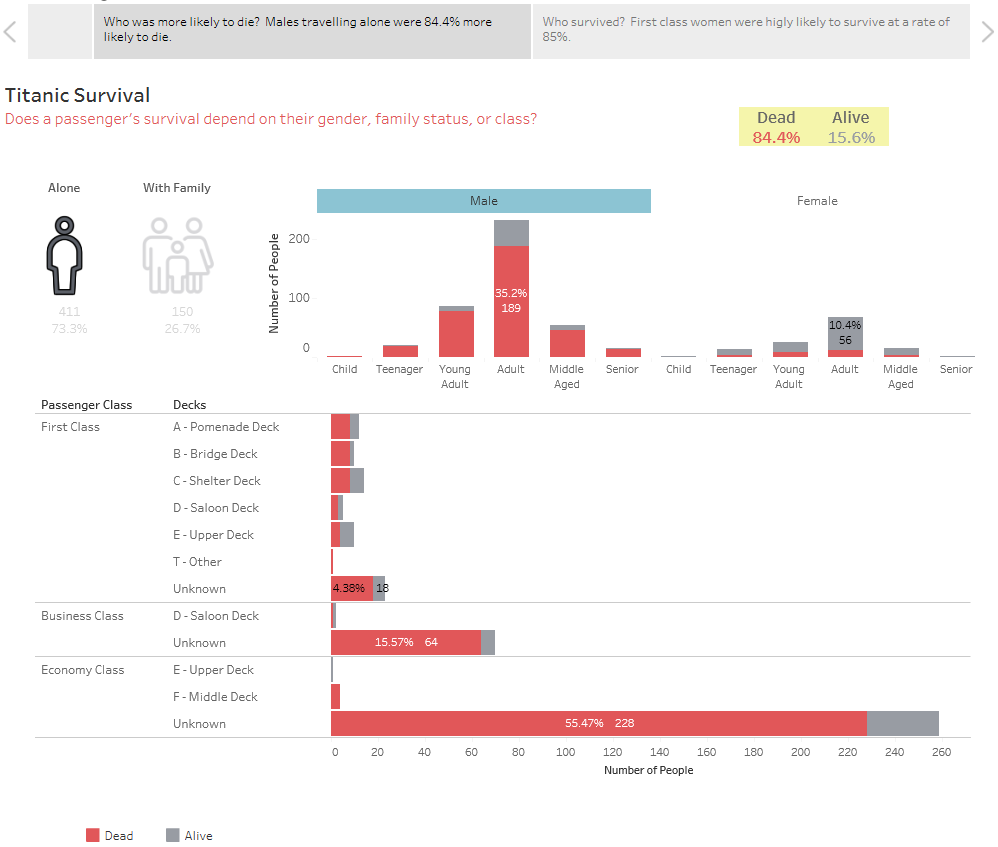
I decided to create 4 slides based on the feedback of needing to create a story to build communication between the visualization and user. I added title page.



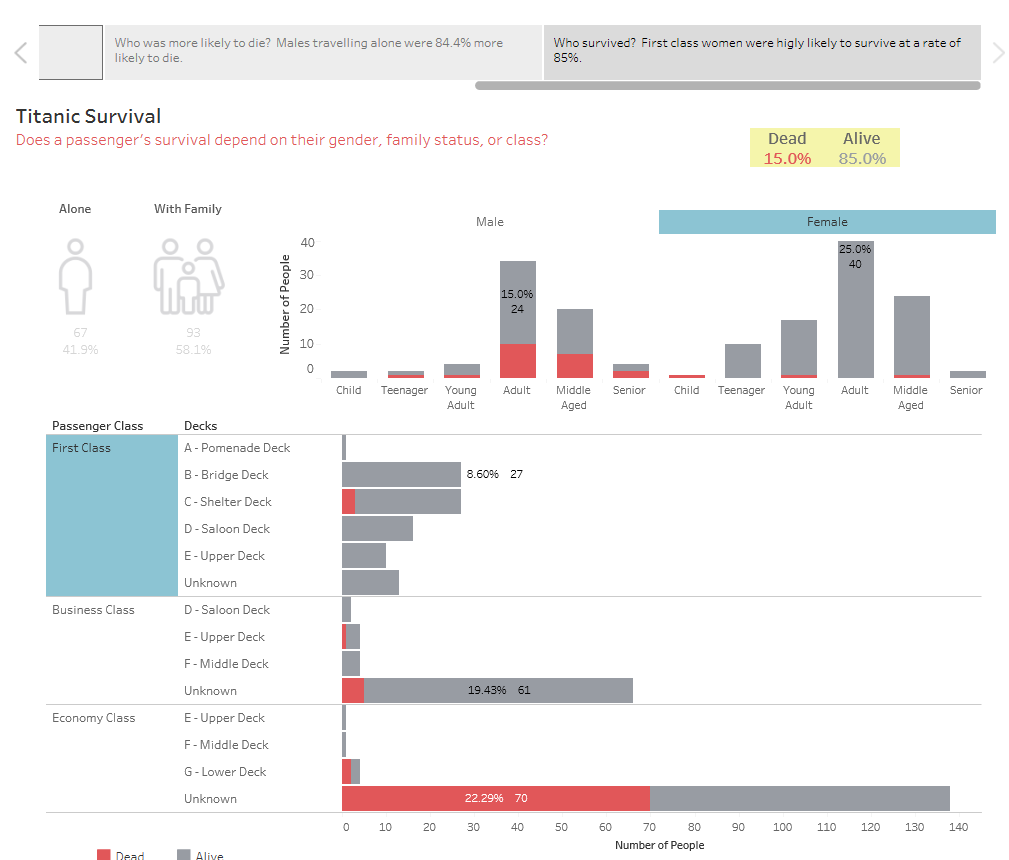
I included a page about who was on board the ship. This was part of my exploration too.



Next, I showed who was most likely to die.



Followed by who was most likely to survive.



## Resources

I used these resources for inspiration on colors, KPIs, and some analysis ideas.

<http://public.tableau.com/views/TitanicDataAnalysis/TitanicAnalysis?:embed=y&:showVizHome=no&:display_count=y&:display_static_image=y&:bootstrapWhenNotified=true>

<https://public.tableau.com/en-us/s/gallery/predicting-survival-titanic>

<https://public.tableau.com/en-us/s/gallery/surviving-titanic>

I also read the Tableau Guide for Designing Efficient Dashboards.