

Can wealth buy survival?

NATHANIEL SIMON, Marquette University

ACM Reference Format: Nathan Simon. 2018. Can wealth buy survival?. 1, 1 (December 2018), 6 pages. <https://doi.org/10.1145/1122445.1122456>

1. ABSTRACT

The present text explores the titanic environment and how many different factors affected. I was able to see from the three pivotal sections (Age, class, and town) that the dataset had many different implied relations.

It is hoped this in-depth look at this Titanic dataset can inform readers about how age, class, and town embarked factor into survival rates on the crash of the titanic. Another large takeaway that the higher one's wealth is the higher of a chance they have to survive.

2. INTRODUCTION

The goal of this project is to analyze the titanic dataset and derive patterns of activity from within them. I chose this set for its many different components in the dataset to compare as well as the fact that I wanted to me more informed about the titanic statistically. The goal of this project is to use dataset to find relations and possibly make predictions about the titanic based on our data. The outcome of this will be several visual graphs and findings that will help answer the research questions. The other part of this project will be performing a lit review on our sources to inform our discussion and to provide the human context for the results.

We have several research questions that form the guiding framework for our project. The first goal is to find a relation on survival based on the cost of an ticket and class. This includes scrapping the dataset and creating many graphs to help visualize the relation. My second question is finding a relationship between town embarked from, wealth, and survival. This means deriving if there a relation between town and wealth and determining what town were wealthier. After that look at the survival rates from those said towns. Finding these will allow for contextualization of the data set and indications of which topics were most important to the group.

3. Lit review

In 1908, a ship named the *Titanic* began construction in a shipyard owned by Harland & Wolff. The ship was about 882 feet long and 92 feet wide and weighed over 46,000 tons. (1) It was said to be largest and most luxurious ocean liner built up to 1912. With sixteen watertight compartments and the most up-to-date technology for the time, the *Titanic* was advertised as the ship that was unsinkable. On April 12, 1912, the day of its first voyage from Southampton to New York, the ship held 2,206 passengers and was equipped with only twenty lifeboats that could hold a maximum of 1,178 passengers. (2) The passengers came from all different social class es and from very different upbringings, however, all would meet the same fate.

Can wealth buy survival?

On April 14, 1912, an iceberg was spotted by the ship's captain around 11:20 at night. The *Titanic* had received multiple warnings of icebergs from other passing ships at sea

. Despite these warnings, it was too late, and the iceberg dragged across the side of the ship, flooding six compartments and two boiler rooms.(3) Upon impact, the lower-class passengers on the lower level felt something like a feeling like an earthquake and panicked, while the first-class travelers on the upper deck barely felt anything and were not disturbed.

The captain then declared that the ship was going to sink, so the crew members began loading the lifeboats in an attempt to remain calm and organized. First-class members were asked to board the lifeboats first, which caused a friendly by other passengers in different classes. (4) Crew members began sending off the first few life boats only half full due to all the confusion and chaos caused by the lower-class passengers. The captain ordered the crew to leave the ship at midnight, and the final lifeboat left with 1,507 passengers still on board. The ship split in half due to the pressure of the water and submerged completely around 2 in the morning. Overall the crashing of the *titanic* resulted in 1,517 deaths. By the next morning, the *Carpathia (name of the rescue crew)* was able to make the journey to the sunken ship and rescue around 705 survivors with lifeboats.

Over the past century, there have been many studies discussing the *Titanic*. Scholars studying it write from a wide variety of different angles and expertise. For example, there was a study done that researchers did that said the crash could have been from a fire and not due to the iceberg. (5) There are a lot of researchers that say the day of the sinking had many different holes and a lot of untold truth. For example one researchers discussed how the *titanic* could broke in "half" due to the way it hit the iceberg.(6) We can see that there are many things hidden or unknown about the event but we know that there were man people who unfortunately passed away. While reading this article we are going to discover the rates of survival based on many different factors.

4. Methods

The dataset I chose was one that I was able to get from kaggle which showed multiple set of information obtained from the *titanic*. This dataset had 10 different variables (Columns) with many different values associated to them. This includes survival, pclass, sex, age, sibsp, parch, ticket, fare, cabin, embarked. The variable survival refers to the survival of the particular passenger. It was recorder either 1 meaning they survived, or 0 meaning they died. The next variable pclass refer to the ticket class of the particular customer. It is identified as 1(1st class), 2(2nd class), or 3(3rd class). The next variable is sex which is the identified gender of the passenger. It is either male or female. The next variable is age which refers to the numerical age of the passenger. The next variable is sibsp which means the number of siblings/ spouses aboard the *Titanic*. The dataset defines siblings being either brother, sister, stepbrother, stepsister. The dataset defined Spouse as husband, wife (mistresses and fiancés were ignored). The next variable was the parch which refers to the number of parents/ children aboard the *Titanic*. The dataset defined children as daughter, son, stepdaughter, stepson. The dataset defines parents as Mother or Father. In some cases some children traveled with a nanny , therefore their parch value would be 0. The next variable was the Ticket which refers to the ticket number the customers received. The next variable was the fare which reled to the total cost for the customers specified ticket. The next variable was the

Can wealth buy survival?

cabin which refers to the particular cabin number the passengers stayed in. The final variable is embarked which refers to the specific ports the travelers embarked from. It is recorded as either C= Cherbourg, Q=Queenstown, or S=Southampton.

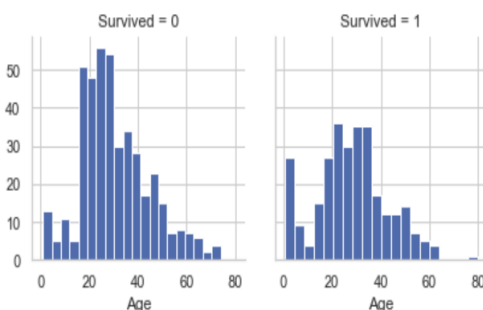
I chose this data set which discusses many different aspects of the Titanic to explore the survival rates based on many different factors. While looking through the dataset I was trying to search for interesting relations. I recalled watching from the film Titanic that the rich were the first ones onto the safety boats and their chances of survival were higher. I wanted to check if that was true. So, the larger Research question I delved into was does money buy you survival. One of the sub questions is there major differences between wealth and the towns the passengers came from.

The first thing I had to do was clean the data in order to see the results more clearly. To start I had to use Jupyter Notebook inside of Anaconda installer to run my python code. To start the cleaning process, I had to load in the correct libraries that I had to use for the project then read the csv file which contained the dataset. I first ran the describe function on the dataset to see some of the ranges and different statistics of each variable in the dataset. I then created a table of gender which helped change the variable of sex into two different variables male and female with values 0 and 1 to determine if they were that specific sex. I also did the same thing with embarked which I split up into 3 different variables which were Cherbourg, Queenstown, and Southampton with corresponding 0 or 1 values if they embarked from that specific city. Next, I deleted the rows cabin and ticket from the dataset. I deleted cabin due to the fact that many of the rows were filled in with the value NaN which indicated that they were not recorded and also the cabin numbers were not a useful variable in order to solve the research question. I deleted ticket since it stated a bunch of random ticket numbers, but I didn't have any relation to the other variables plus it was not useful in order to solve the research question.

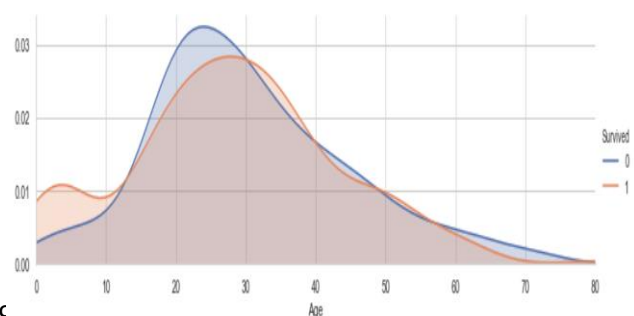
5. RESULTS

4.1 AGE

I first wanted to see the relation that age had on the survival on the passengers. To do so I had to compare the two variables of survival and age. I first had to create a bar chart and to do so I had to make my x axis as the age of the passengers and the y axis as the two variables survive or died. This was able to make two separate bar charts showing survival rates of ages based on if they survived or died. We can see that between the age of 20-40 a large number of passengers died. One key thing this graph shows that ages closer to zero had a high chance of survival. (SEE Fig1) Next, I used the same data from the last graph to create an overlaid shaded graph. This was able to show survival rates on top of each other. This also further shows the ages 20-40 died more than they survived. (SEE Fig2)



019



Can wealth buy survival?

Figure 1: Bar graph of Age vs amount survived and died

Figure 2: Shaded overlay graph of Age vs amount vs. survived and died

4.2 Class

Next, I wanted to take a look at the relation between class and survival. In doing so I had to compare the two variable of the fare (cost of ticket) and survival. I started by taking each separate class and splitting them up. Then I had to take the sum of each class with the those who survived. Then I plotted a bar chart which showed the relation between those who the specific classes and those who survived. This graph was able to show that upper class was the highest in the graph followed by middle then lower for the survival rates. (SEE Fig 3) After I wanted to see if there was any relation of class and gender based of survival. To do so I had to set the x value as the class, the y as those who survived and the hue as the sex. We can see from this graph that upper- and middle-class females had the highest survival rates. We also see that lower-class men had the lowest rate of survival in the chart. (SEE Fig4) Finally one key factor I wanted to see is if there was any relation between class and age on survival. To do this I had to create a chart. I started by making the x axis as the age values. Then the columns of the chart had to be the amount of people who survived or died. Finally, I set the rows of the graphs as the different classes. Overall it printed 6 different graphs. One of the biggest things you see on this graph is that the middle-aged lower class had the lowest rates of survival since it has the highest peaks in the survived=0 graph. (SEE Fig5)

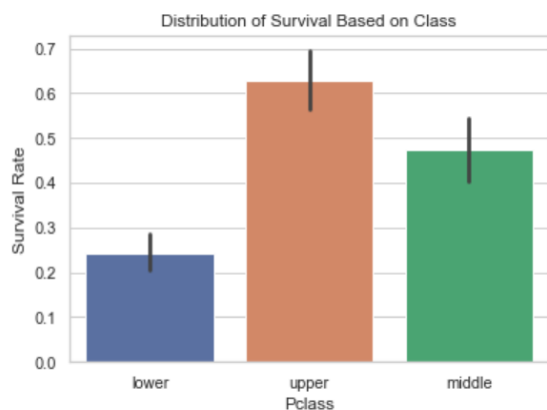


Figure 3: Bar graph of Class vs. Survival rate

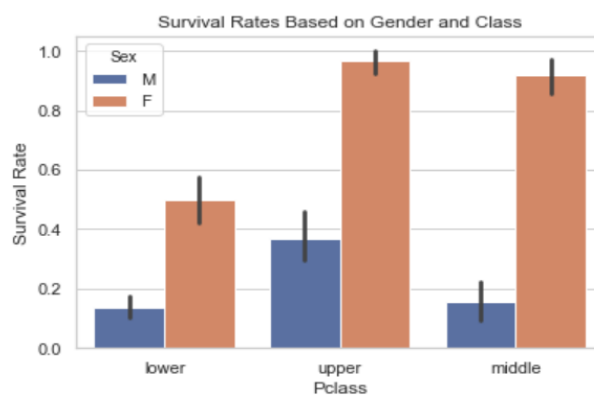


Figure 4: Bar graph of Class vs survival rate vs Gender

Can wealth buy survival?

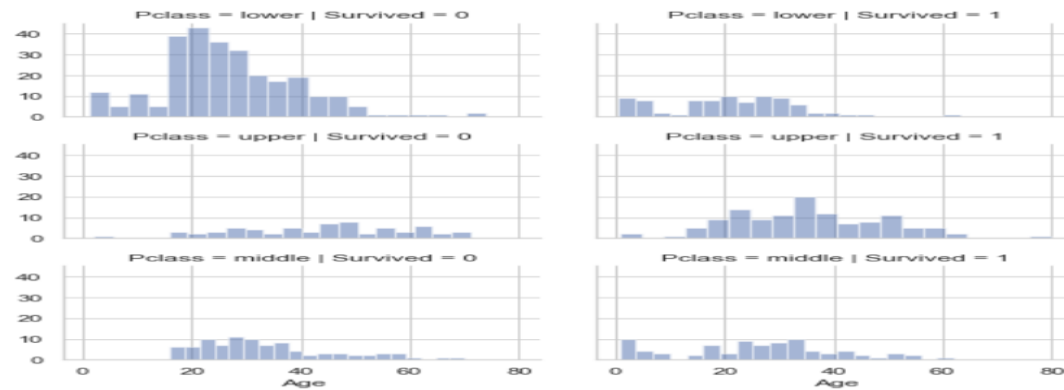


Figure 5: Bar graph of Age vs survival and class

4.3 Towns

Looking at the relation between classes we saw many different survival rates. I wanted to see if there were higher fares from different towns and if there are survival trends between different towns. To do so I first a relation between fares, gender, towns and survival rates. I assigned the x axis as gender and the y axis as the fare of the ticket. Next, I set the column as survival rate and the rows as what town they embarked from. This created 6 unique graphs. We can see that fares and survival was the highest in the town Cherbourg and the lowest ticket cost and survival rate was recorder in Queenstown. (SEE fig6) Next, I wanted to compare class and survival rates based on gender and town. To do this I had to set class as the x values and survival rate as the y. I wanted to differentiate gender, so I assigned gender to different colors. Males associating with blue and females associating with orange. Finally, I set the row sections as the town that the passengers embarked from. This displayed three distinct graphs. At first glance we can see that overall Both Cherbourg and Southampton had larger survival rates compared to Queenstown. (SEE Fig 7)

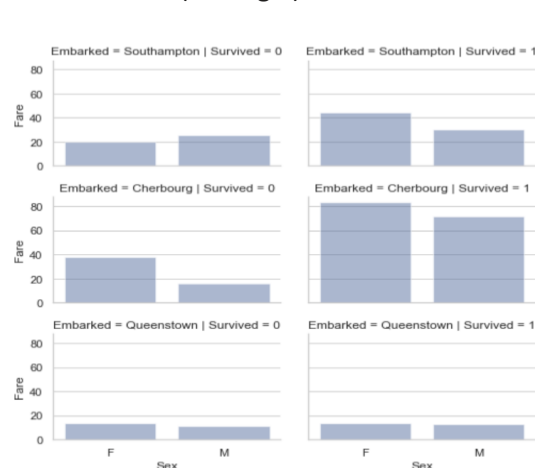


Figure 6: Bar graph of Age vs amount survived and died

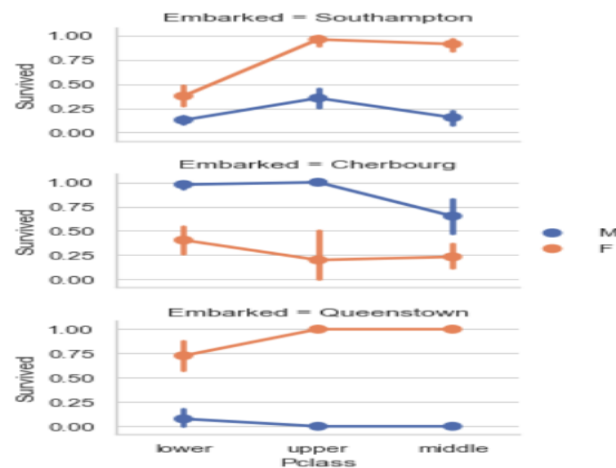


Figure 7: Bar graph of class vs survival rate and gender

4.4 Heat map survival on age and fare

Finally, I wanted to see a heat map relation between survival age and fare. To do this I had to assign the x axis as age and the y axis as the cost of each individuals ticket. Next, I created a color map that

Can wealth buy survival?

correlated with survival rates. So higher survival rates would be more red and lower survival rates would be bluer. This created a scatter plot with every passenger's survival. We can see that a majority of points that were over \$100 were more red. (SEE Fig8)

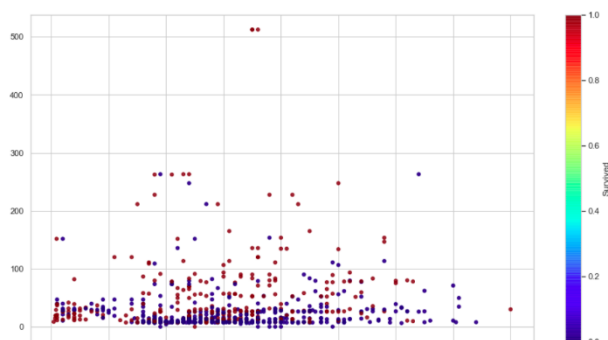


Figure 8: Bar graph of Age vs amount survived and died

6. IMPLICATIONS/ANALYSIS

We can imply multiple conclusion based on the graphs we just visualize. Looking back at the age subsection we can imply from both figure 1 and figure 2 that from ages 20-40 had the lowest rate of survival. We can see that many middle-aged passengers of the titanic died. From the class subsection we can see in graphs 3 and 4 class definitely played a role in survival rate. Both graphs clearly show that highest rates from survival were from the upper class, then the middle, and finally the lower. We can also see that females were more regarded to save due to the fact that in figure 4 we see that females who were associated with the upper class had the highest peak on the graph for survival. This can imply that the safety of the upper class and females of upper and middle class were the highest priority during the crash saving efforts of the Titanic. Looking back in the town subsection we can see from figures 6 and 7 that Cherbourg and Southampton had higher fares and survival compared to Queenstown. This could imply the passengers from both Southampton and Cherbourg were richer and also had a higher chance of surviving the crash of the titanic. We imply that Cherbourg is the richest town of the three due to the fact that it had the highest average cost per ticket. We can also imply that Queenstown is the poorest town of the three since it had the lowest average fare. We also see that Cherbourg which the richest town was had the highest survival rate while Queenstown which is the poorest town had the lowest survival rate. This can imply that if you are from a rich town or have a higher average ticket cost you were likely to survive. In other words, we can see that passengers who were of more wealth had the highest chance of survival. Finally taking a look back at the heatmap section we can see that most of the dots were blue (which identifies with lower survival rates) were towards the bottom. This can imply chances of survival with a low-ticket cost was low. This backs up the point that the rich had the highest chance for survival. With all these implications put together we can make one final implication by saying, money can buy you survival.

Can wealth buy survival?

7. CONCLUSION

Overall, I saw many different connections the dataset displayed. But some of the major takeaways from observing titanic trends was that money can buy you survival during the titanic crash. I was able to visually see this in multiple graphs. First, I started to look at correlation between age and survival rates. I saw that middle-aged passengers had the lowest rate of survival. Next, I looked at class versus the rate of survival. From the multiple graphs I was able to see that the higher the class on is in, the more likely it is for them to survive. We can also see an emphasis on upper- and middle-class females' survival since it was higher than most. We can see that wealth class can buy one their survival. Next, I wanted to take an in-depth look at the towns the passenger embarked from. From the multiple graphs I was able to see that Cherbourg was the richest town due to the fact that it had the highest average cost per ticket and Queenstown was the poorest town since the ticket cost were drastically lower than the other towns. I was also able to see that survival was higher in towns that were wealthier. Finally, in the heat map section we were able to see each individual chance of survival based of their ticket fare. We were able to see that ticket cost over \$100 had the highest chance to live. This shows that money can buy a passenger on the titanic his survival

8. References

- 1.) Tikkanen, Amy. "Titanic." *Encyclopædia Britannica*, Encyclopædia Britannica, Inc., 13 Oct. 2019, www.britannica.com/topic/Titanic.
- 2.) "'Unsinkable' Titanic Sinks." History.com, A&E Television Networks, 13 Nov. 2009, www.history.com/this-day-in-history/unsinkable-titanic-sinks.
- 3.) Frey, Bruno, S., David A. Savage, and Benno Torgler. 2011. "Behavior under Extreme Conditions: The Titanic Disaster." *Journal of Economic Perspectives*, 25 (1): 209-22., <https://www.aeaweb.org/articles?id=10.1257/jep.25.1.209>
- 4.) Shutes, Elizabeth. "The Sinking of the Titanic, 1912." The Sinking of the Titanic, 1912, Eyewitnesstohistory, www.eyewitnesstohistory.com/titanic.htm.
- 5.) Wwww.dailysabah.com. "New Research Claims Titanic Sank Due to Enormous Fire on Board, Not Iceberg." DailySabah, Wwww.dailysabah.com, 2 Jan. 2017, www.dailysabah.com/history/2017/01/02/new-research-claims-titanic-sank-due-to-enormous-fire-on-board-not-iceberg.
- 6.) Bressan, David. "A Geological Study Of The Titanic Shipwreck Site." Forbes, Forbes Magazine, 14 Apr. 2019, www.forbes.com/sites/davidbressan/2019/04/14/a-geological-study-of-the-titanic-shipwreck-site/#3c87d8c6431a.
- 7.) <https://www.kaggle.com/c/titanic/data>