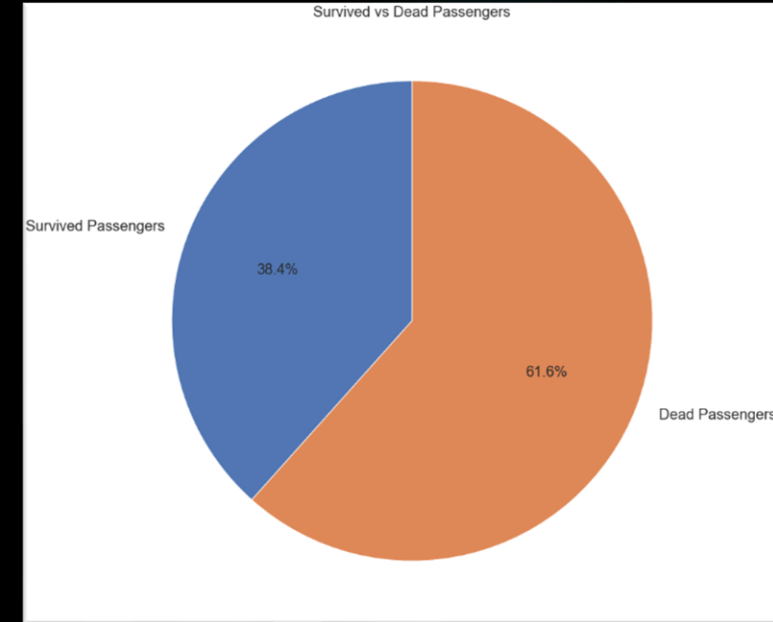


# Situation

Many lives were lost in the crash of the large ship called the Titanic in 1912. Our goal for analyzing the data would be to find any patterns which caused the crash and then we may be able to determine some new safety measures to prevent the repeat of a crash of this magnitude.



Data	Value
Survived Passengers	342
Dead Passengers	549

# Problem

We want to find the variables which may or may not have contributed to the survival of children in the crash. Here we will decide that anyone aged 13 or below is considered to be a child.

# Solution

We will determine which features create a proper correlation with the survival rate of children during the crash.

We can see that this is a normal distribution where we see that children aged between 5 and 6 are the ones that died the most. In addition to that, we also see that young babies (aged 0-1), and early teenagers (aged 12-13) have a 100% survival rate.

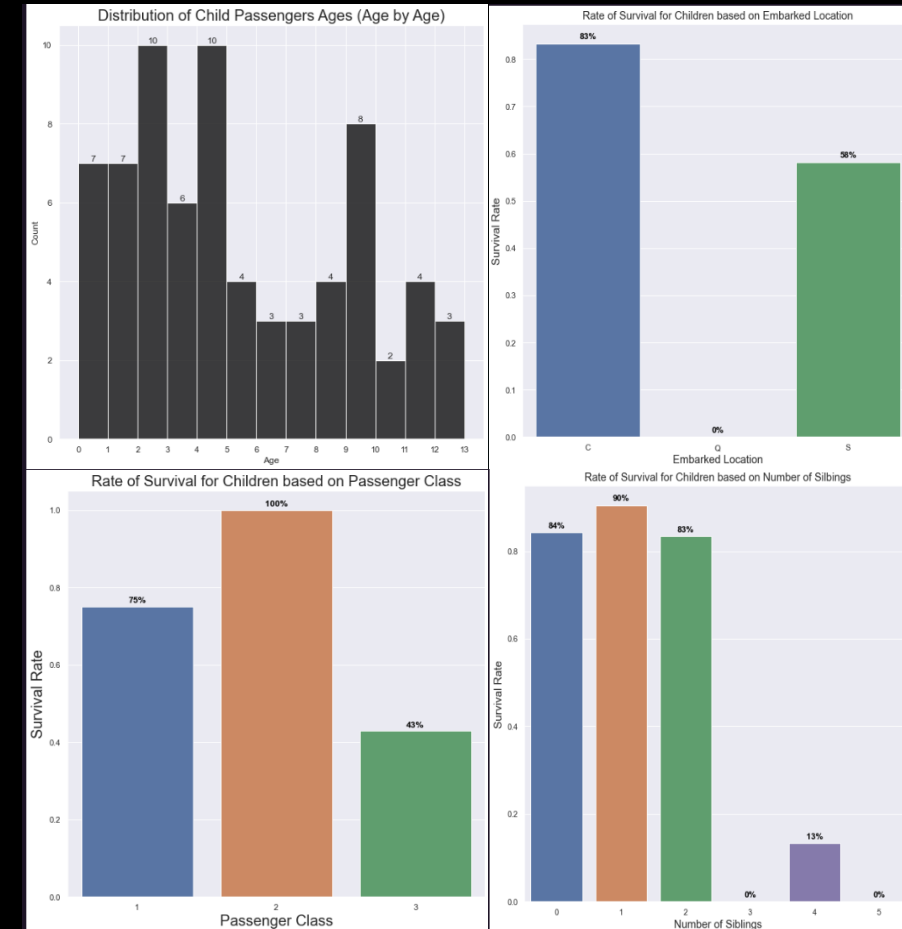
To further continue the analysis, we could check which other factors can be leaded to this. Initially we can check for the passenger class, as there is some differences with the children there.

We have some unexpected results. We initially see that the most number of children are in the 3rd passenger class, however we are getting the lowest amount of survival percentage for that class. It is the same for passenger class 1, where we see 3 out of the 4 children survive. Hence, we can not draw any correlation from this, and would need to continue with another variable of the analysis. Let's try to see if there is any correlation between the port embarked and the survival rate of the children.

Here we can make a clear correlation with the higher the number of siblings a child has, the lower their chance of survival is. We can also simplify this further by saying that if you have 3 or more siblings, your the max chance of survival you have is 13% which is abysmally low.

This leads to two correlations who are intertwined:

- The more siblings you have, the lower your chance of survival will be.
- If all the siblings are from the same family, the lower the chance of survival will be since families will tend to stick together.



# Next Step

Based on our analysis, we can draw 3 different conclusions:

1. Children who are at the extreme spectrum (ages 0-1 and 12-13) have a higher chance of survival, and children in the middle of the spectrum have a lower chance of survival.
2. Children who embark the same port as a family will more than likely perish together.
3. The higher number of siblings you have, the lower your chance of survival get.

While we would need further testing and analysis to fully prove these three observations, we can still apply some steps as to reduce their effects in case another Titanic sails the high seas.

For example, in the future, we can limit the number of siblings a children can have to 2, so that we can prevent such tragedies. We can also apply the same safety and security procedures applied to the children of class 2 to the other class so that this does not happen again.