

Null Hypothesis (H0): The survival rate is not associated with the class of the passenger.

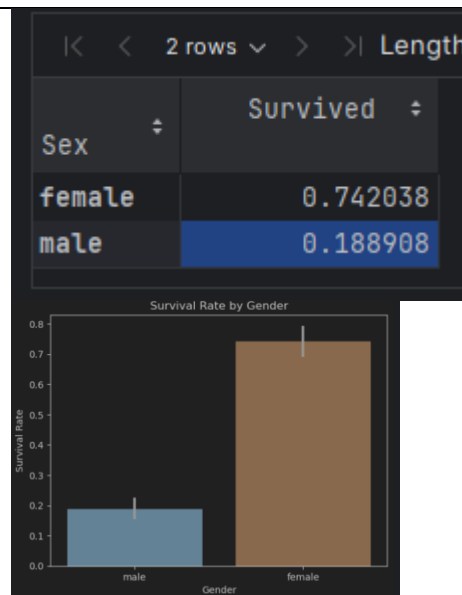
Alternative Hypothesis (H1): The survival rate is associated with the class of the passenger.

Findings:

Class 1 had the highest survival rate at 46.8%, followed by Class 2 at 35.2%, and Class 3 at 18%. The significant difference in survival rates between passenger classes is evident, with Class 1 having more than double the survival rate of Class 3.

These findings strongly indicate a direct association between passenger class and survival rates.

Conclusion: We reject the Null Hypothesis (H0) in favor of the Alternative Hypothesis (H1) because the data clearly demonstrates a correlation between passenger class and survival rates. Specifically, first-class passengers had a significantly higher survival rate compared to third-class passengers.



Determine if the survival rate is associated to the class of passenger:

Null Hypothesis (H0): The survival rate is not associated with the class of the passenger.

Alternative Hypothesis (H1): The survival rate is associated with the class of the passenger.

After conducting EDA on the titanic dataset, we have observed the following insights between PClass and Survival rates:

Findings:

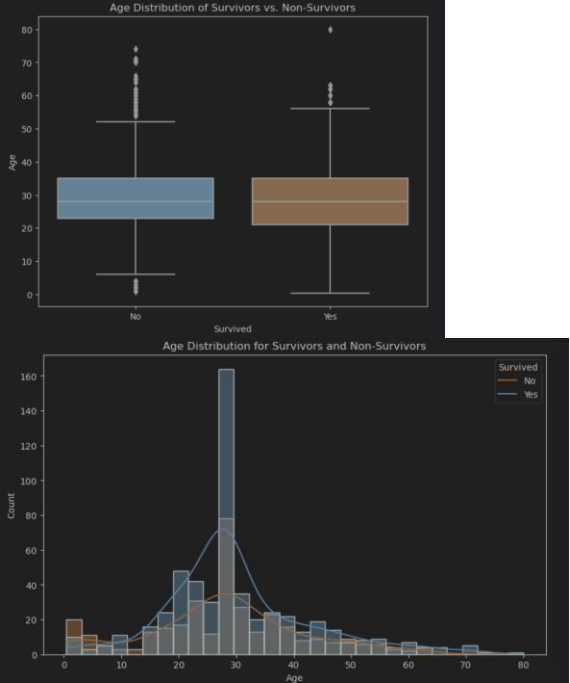
Class 1 had the highest survival rate of 46.8%, followed by class 2, being 35.2% and lastly Class 1 being 18%

The difference in the survival rate of classes is noticeable as Class 1 has more than double the survival rate of class 3

This suggests a strong association between class and survival rates

Conclusion:

We reject the Null Hypothesis (H0), accepting the alternative hypothesis because the data clearly

	<p>shows that passenger class is directly correlated to survival rates.</p> <p>First Class passengers had higher survival rate than 3rd class passengers.</p>
 <p>The figure consists of two charts. The top chart is a boxplot titled 'Age Distribution of Survivors vs. Non-Survivors'. The y-axis is labeled 'Age' and ranges from 0 to 80. The x-axis is labeled 'Survived' with categories 'No' and 'Yes'. The 'No' box is blue and the 'Yes' box is brown. Both boxes show a median around 28, with whiskers extending from approximately 10 to 55. The bottom chart is a histogram titled 'Age Distribution for Survivors and Non-Survivors'. The x-axis is labeled 'Age' and ranges from 0 to 80. The y-axis is labeled 'Count' and ranges from 0 to 160. The legend indicates 'Survived' with 'No' in blue and 'Yes' in brown. The histogram shows two overlapping distributions, both peaking around age 28-30. The 'Yes' distribution (brown) has a slightly higher peak count than the 'No' distribution (blue).</p>	<p>Hypothesis 3: Determine if the survival rate is associated with age</p> <p>Null Hypothesis (H0): The age of passengers is not associated with their survival rate, and there is no significant difference in survival rates among different age groups.</p> <p>Alternative Hypothesis (H1): The age of passengers is associated with their survival rate, and there is a significant difference in survival rates among different age groups.</p> <p>Findings:</p> <p>As the mean and median age of survivors(28.9, 28) and non-survivors(30, 28) is similar, we cannot see any direct correlation between the two variables.</p> <p>The little difference in the mean may be caused by outliers as mean does not take those into account. That is why the more accurate values to consider are the median, which are both 28.</p> <p>The histogram and boxplot also shows a similar shaped graph with both being symetrically shaped and not skewed to any side.</p> <p>Conclusion:</p> <p>We accept Null Hypothesis, rejecting the alternative hypothesis as the median age of survivors and non-survivors is the same, so is the shape of their histogram and boxplot, indicating that there is no correlation between the age and survival rate.</p>