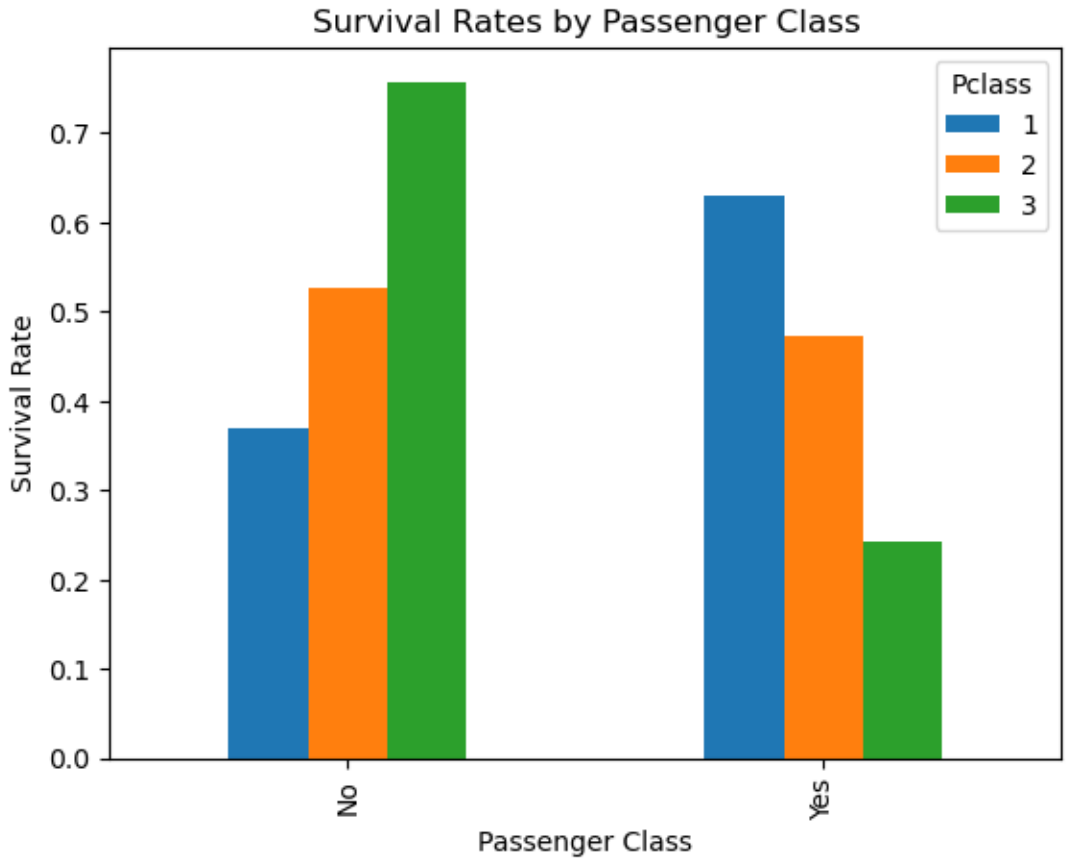
1. Determine if the survival rate is associated to the class of passenger.

The function provides various values including the chi-squared statistic, p-value, degrees of freedom, and expected frequencies under the null hypothesis. The results are then printed out. If the p-value is below 0.05, the null hypothesis can be rejected, and it can be concluded that there is a significant relationship between survival rate and passenger class. In this case, the chi-squared value is large at 102.89, suggesting a significant difference between the expected and observed counts, and supporting the presence of a statistically significant association between survival rate and passenger class.

The p-value is extremely small at 4.55e-23, much less than the usual significance level of 0.05, which strongly contradicts the null hypothesis of independence. This implies that the probability of getting such a high chi-squared value or even greater by chance alone is very low. Thus, the null hypothesis can be rejected, and it can be concluded that there is a significant association between survival rate and passenger class.



2. Determine if the survival rate is associated to the gender.

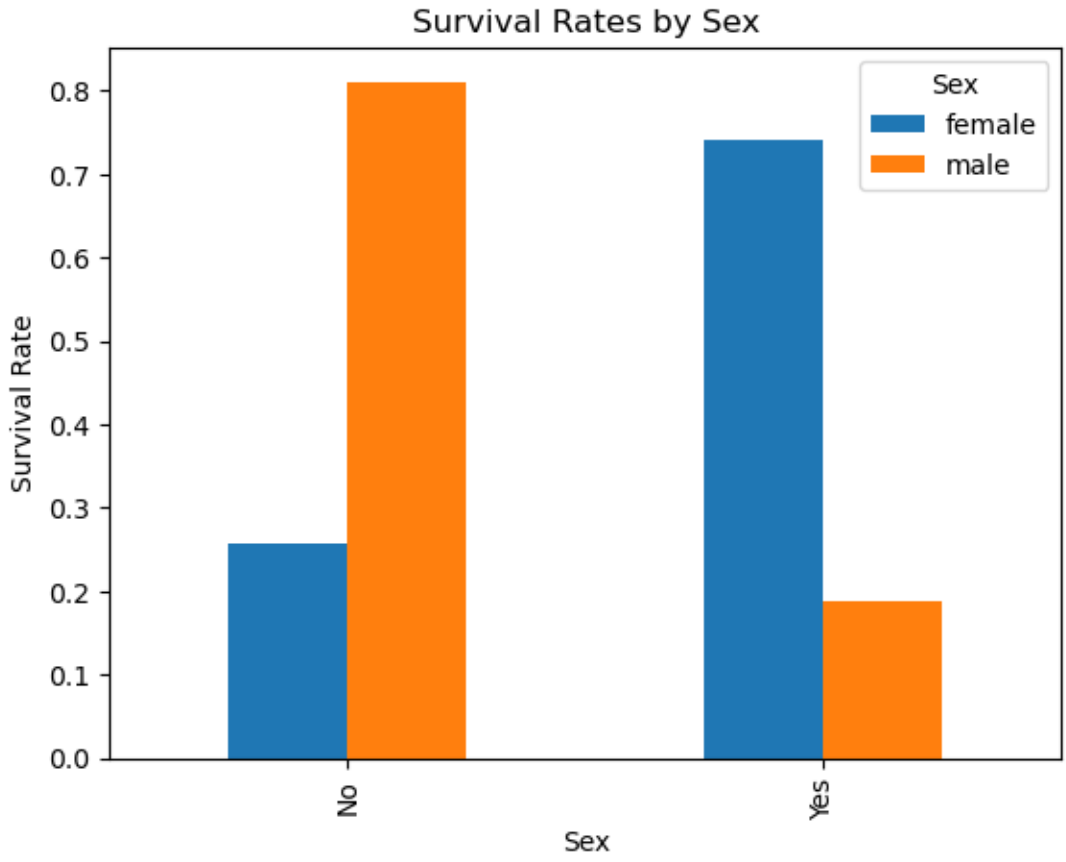
The chi-squared statistic is used to determine the difference between the observed and expected counts in a contingency table, and in this case, it is 260.72. Such a large value indicates a strong association between the variables being studied.

A small p-value of 1.20e-58 shows that the null hypothesis of independence is strongly contradicted. This means that the chances of obtaining such a high chi-squared value or an even greater one by chance are very low, leading to the rejection of the null hypothesis and concluding that a significant association exists between the two variables.

The degrees of freedom are 1, which is calculated by subtracting 1 from the number of rows multiplied by the number of columns, in this case, 2 rows and 2 columns.

The expected counts table shows the expected frequency counts under the null hypothesis of independence, which is calculated using the row and column totals of the contingency table. The observed counts are significantly different from the expected counts, providing further evidence against the null hypothesis.

Overall, these results strongly suggest a significant association between the variables in the contingency table that is unlikely to be due to chance.



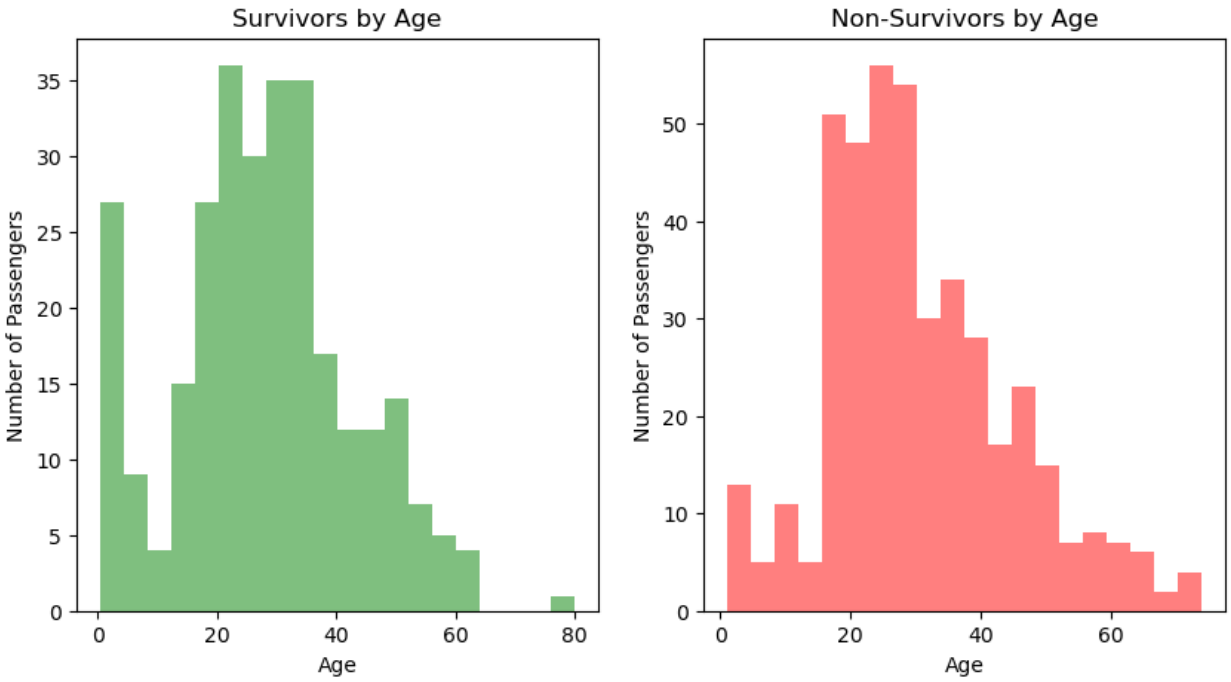
3. Determine the survival rate is associated to the age.

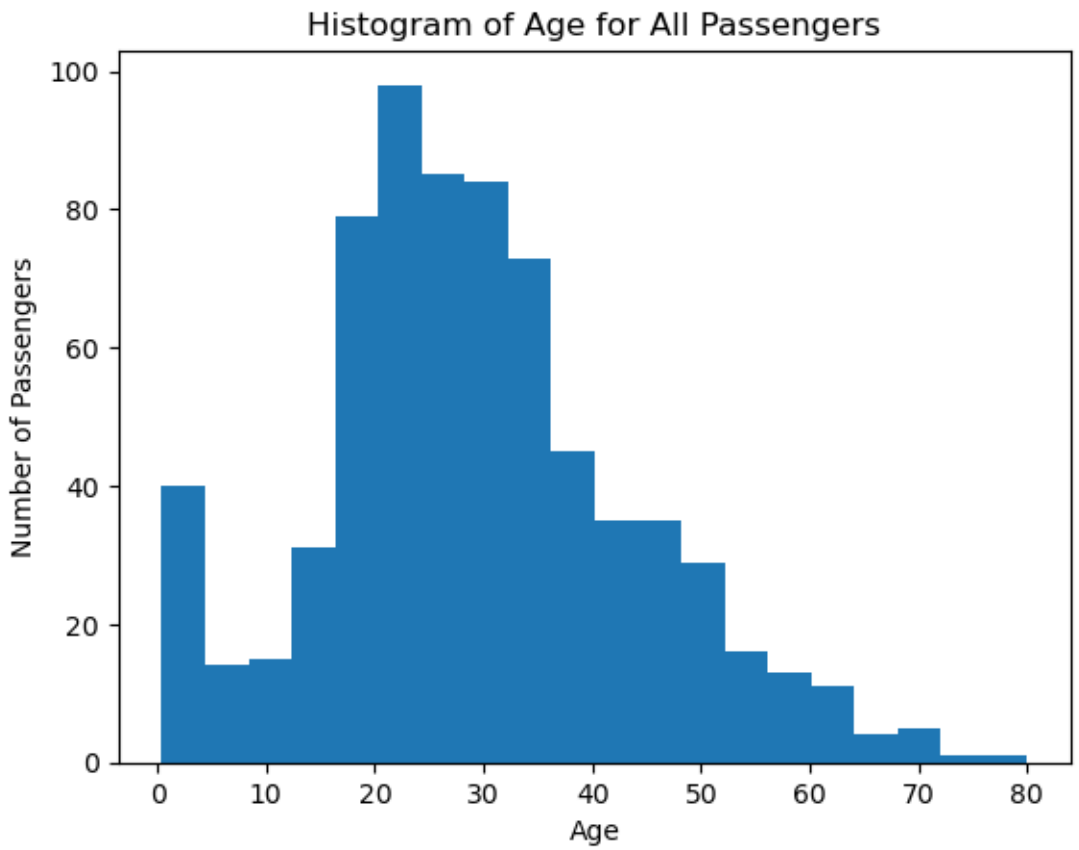
To investigate whether age and survival rate are correlated, we can employ a chi-squared test of independence by creating a contingency table that cross-tabulates age groups with survival status. To achieve this, the first step is to generate age groups by classifying age variable into four groups: children (age < 18), young adults (18 <= age < 35), middle-aged adults (35 <= age < 50), and older adults (age >= 50).

Next, create a contingency table with two rows representing survived and not survived, and four columns indicating each age group. Tabulate the count of passengers in each group who survived and who did not survive, and fill in the corresponding cells of the table.

To test the null hypothesis that survival status and age group are independent, and the alternative hypothesis that they are dependent, perform a chi-squared test of independence. Calculate the chi-squared statistic, the degrees of freedom, and the p-value. Reject the null hypothesis and conclude that there is a significant association between survival status and age group if the p-value is less than a pre-selected significance level (e.g. 0.05).

In case of a significant association, we can investigate the nature of the association further by calculating measures of association such as Cramer's V or odds ratios. Such measures can indicate the strength and direction of the correlation between the two variables.




Histogram of Age for All Passengers
Number of Passengers
Age
0
20
40
60
80
100
0
10
20
30
40
50
60
70
80