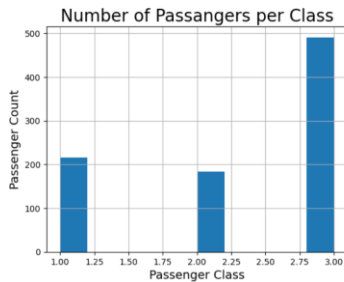
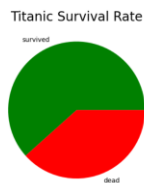


Analysis of Titanic Ship and its Passengers

Bar Graph: My Bar Graph that I made displays the number of passengers per class on the Titanic. This display of data shows how there was a significant number of passengers within the 3rd class, which is the lowest class on the ship. The second class and first-class number of passengers were very close in count, but there were more 1st class passengers than 2nd class passengers



Pie Chart: The pie chart shown is the survival rate of all passengers on the Titanic. The green shows that they survived and the red shows the passengers who didn't survive.

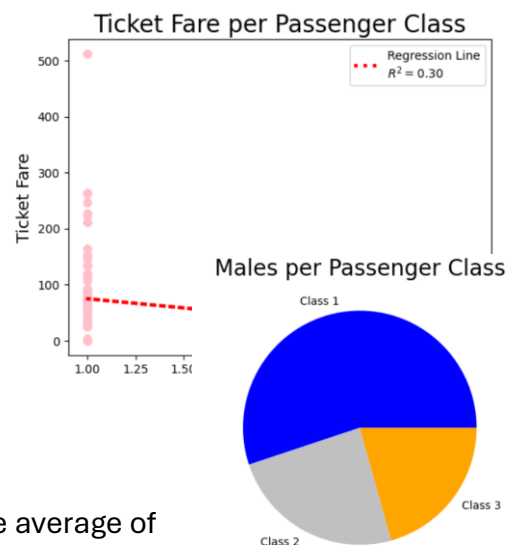


Scatter Plot:

The Scatter plot involves the ticket price average per passenger class. The graph displays a negative r value and a line of best fit that is nothing like we have seen with our other scatter plots and regression lines.

Additional Figure: The additional figure I chose to display is another pie chart displaying the male count per passenger class. Males on the titanic were majority in the first class, and in the second and third class there was an close split.

Contingency Table: My contingency table shows the average of females and males that had survived on the titanic ship. In my chi squared test I used a p value, which came out to be one, to test if there was any difference in male and female survival rates. My null hypothesis was that there was no



difference in survival rates of male and females. My alternative hypothesis was that there was a difference in survival rates. Given that the p value of 1 was greater than 0.05, we fail to reject the null hypothesis.

```
#Chi Squared Test on the sex and survival rate
sex_survived = pd.crosstab(titanic_data["Sex"], titanic_data["Status"], normalize = "columns")
print(sex_survived)

Status    Alive    Dead
Sex
female    0.681287  0.147541
male      0.318713  0.852459

c, p, dof, expected = scipy.stats.chi2_contingency(sex_survived)
print("The P-value is: ", p)

The P-value is:  1.0

print(sex_survived - expected)

Status    Alive    Dead
Sex
female    0.266873 -0.266873
male     -0.266873  0.266873
```

Numerical Data: Two data sets that would make sense to compare to each other would be ticket price and age. In today's world the cost of a ticket to anything, such as sports or an amusement park, the ticket price is dependent on the age of the person. Typically the older the person, the more expensive the ticket. Seeing these two categories can determine if there was an evolution of ticket pricing, or if pricing had stayed the same throughout time.