Shark Attacks are the New Wave: Do Sea Levels Affect the Frequency of Shark Attacks?

For our project, we decided to look at shark attack and sea level datasets. We were most excited to find that there was a correlation between shark attacks and sea level which we will discuss more later.

We obtained two datasets for this project: one containing information regarding global shark attacks, and one concerning sea levels. The shark attack dataset contains interesting information such as the demographics of the person attacked, the country they were attacked in, and whether the attack was fatal, among others, dating as far back as the year 500. The sea level dataset tracks monthly sea levels for thirteen regions in North America beginning in January 1961.

We cleaned the sea level and shark attack data so that they would be easier to combine, or merge, which we then did. We also dropped any columns that we did not use in our analysis to make it easier for us to look at our data. We found the median sea level for each month and added this information to the shark dataset. Considering that the sea level data is only applicable to the United States and Canada, we filtered our merged data to include only shark attacks from these two countries. We then retrieved the counts for the number of shark attacks in each month and added these values to the data for the US and Canada.

We noticed that there were many more shark attacks reported recently compared to before the 1990s, seeing that between 1961 and 1990, there were 456 shark attacks reported compared to 1408 reported between 1991 and 2020. For this reason, we believed that our results could have been skewed due to there being more shark attacks reported in more recent years due to their being more people who want to report them and better technology to keep track of them. We wanted to make sure that our data was as constant as possible and affected the least by changes in technology or by more people wanting to report their shark attacks. As a result, we took a sample of 456 shark attack cases from each of the time periods we mentioned earlier and compared these to each other through a scatterplot and a correlation coefficient heatmap. We found that there was very little association shown in the scatterplot and close to no correlation found with a correlation coefficient of about –0.00079 between both samples of shark attack data. Since both samples did not seem to be correlated at all, in short, the median sea levels are not changing much over time, regardless of the time period.

Then we worked on finding whether there was a correlation between sea level and the amount of shark attacks reported per month from 1961 to 2020. We first created a scatterplot (see Exhibit 1) using the median sea level for each month in comparison and the number of shark attacks reported per month to see if there was any visible correlation. We did see a faint positive correlation between these two variables. To ensure this correlation, we created a heatmap (see Exhibit 2) which shows the correlation coefficient for our variables which ended up being 0.35, a positive correlation which is not too strong, just as we predicted. This was an exciting finding for us because it showed that sea level seems to indeed affect the amount of shark attacks per month, although it may be somewhat slightly. We even further ensured this by performing a t test based on high sea levels versus low ones (in the fourth quartile versus the first quartile of median sea levels for each month) and the amount of shark attacks. Again, we found that the p-value was very low compared to the significance level of 0.05, showing that we should reject the null hypothesis, meaning that there is a significant difference in the number of shark attacks between low and high sea levels.

Exhibit 1: Scatterplot of Number of Shark Attacks vs. Monthly Median Sea Level

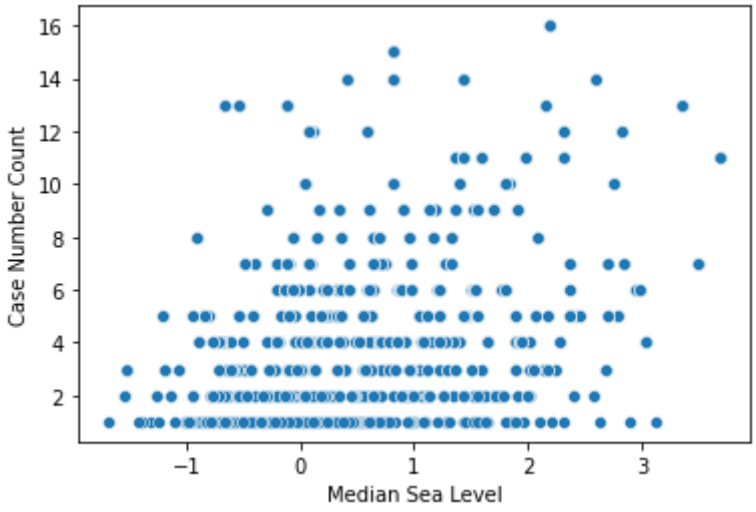
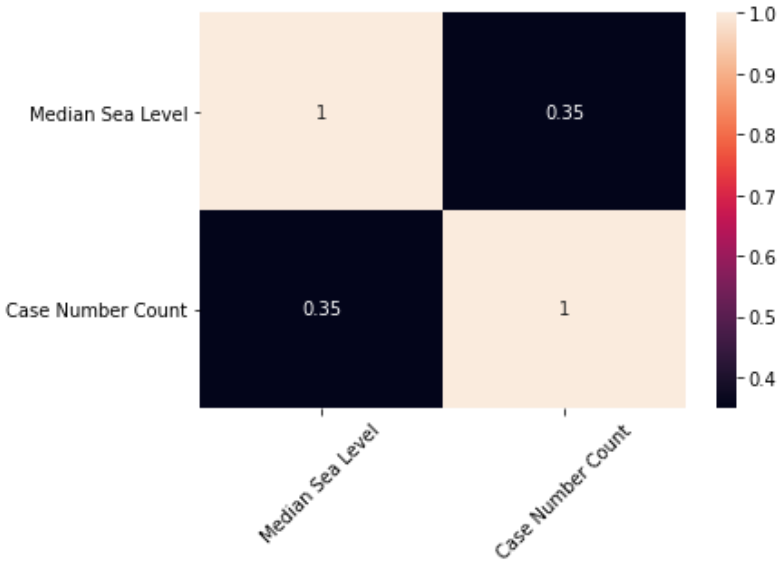


Exhibit 2: Heatmap of Correlation Coefficients



From our findings, we can conclude that sea levels do in fact affect the number of shark attacks due to the positive (though small) correlation we were able to see in a scatterplot we created, the t-test we performed, and the correlation coefficient we found. Further, we believe that due to climate change, sea levels will rise as glaciers melt and oceans expand. As a result, we believe that there might be more shark attacks which occur in the future due to our findings.

To find our GitHub repository which contains the code we wrote in order to analyze our datasets and draw conclusions and more, go to https://github.com/tayloregan1/final_project. To view our dashboard, go to https://taylorandanusha.herokuapp.com/.