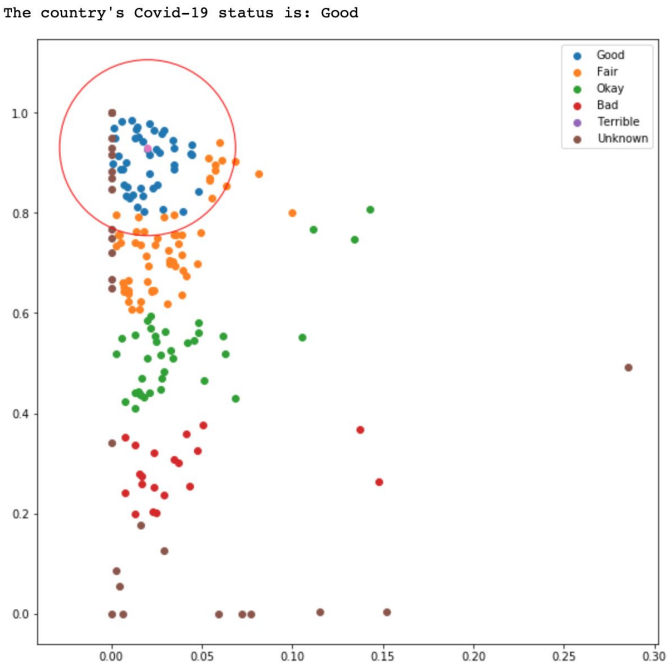
COSC 311 Project 2

Members: Declan Sheehan, Jack Stoetzel

With our COVID-19 dataset from kaggle we have implemented a kNN algorithm to determine the current condition a country is in. Every country in the data set is categorized as either Good, Fair, Okay, Bad, Terrible, or Unknown. These conditions are based on the country's current death to recovery ratio. If the number of people recovered from COVID-19 is greater than the number of people who have died from COVID-19, then the country is in a “Good” condition. The graph below shows that there are groups of countries that share the same condition in close proximity to each other. So, if we were to plot a random point on this graph, we could use our kNN algorithm to determine what condition it should be classified as.



To further visualize this data, we are considering plotting this data on a graph and making predictions based on geographical location as well as recovery to death ratios.

So far, the project is going well and we believe we are working in the right direction. We are planning to implement a linear regression algorithm as our second machine learning algorithm. If we do not believe we are getting conclusive results from it, we will look into another algorithm to implement for our data.