In this project, we analyze 5 data sets, each having crucial data to examine and talk about. The data was already formatted for us. The first data set that I analyzed was the worldwide data. First thing, I wanted to create a map with the index of deaths, just to see what country had the most deaths. I was amazed that some countries had really high death rates and some had really low rates. Then I decided to do a scatterplot maybe I would be able to see the data better. Most of the countries had a similar death rate but there were a few outliers, then I decided to only graph those outliers, and I was amazed that those countries were mostly on different continents, I would assume that they were in the same geographical location. Now we had a daily dataset that consisted of 175 days, and I compared it to our y-axis which is the number of deaths. I did see a pattern that after around day 60 the deaths skyrocketed, which we can see that maybe it took around 2 months for COVID-19 to fully develop. Now I wanted to compare every variable of the dataset, and I concluded that we actually got better at saving lives as time went through. The ratio between new deaths and new recovered is way higher than deaths and recovered implying that the chance of dying of COVID-19 was decreasing over time went through.

Now we have analyzed the USA data set, we can see that we have some places that are not states where we removed them. Let's group them by state and add the confirmed deaths. Now I did a bar graph that compares the death ratio maybe we can see some outliers., there was a strong outlier which was South Dakota, with around a 15% death rate. With the worldwide data, I decided to group them by continent and add all the variables grouped by their continent. I was mostly concerned with Total Cases, Total Deaths, and Total Recovered. Decided to create a bar graph, and the results were pretty similar all continents had similar ratios, but I felt my Y-axis was messing up my conclusion. So I decided to graph them and scale Y axis by a million. As I graphed them we could see a gap between total deaths and total cases, north america had a big gap, and south america and Asia had a similar gap.

After this project, which I really enjoyed, I analyzed 5 datasets and focused on the important information, I learned so much about data visualization, and I understood the power of data visualization. I tried to make the data visitations as simplier as possible for the general public to understand. We can see that COVID-19 has many trends and patterns that are only possible to see with the power of data visualization. I used as many types of data visualization such as barplots, scatterplots, and graphs. Since the data had a big numbers I had to scale it too see it better.