

Visual Analysis

Covid-19 was a deadly virus that quickly spread in late 2019 and throughout 2020. In this analysis we will look at some data from this time on how different locations were affected through out this period and we will try to see if there was any correlation between the case-fatality rate and the new cases that appeared in different locations in the year 2020. The data for this analysis was obtained by collected data by "Our World in Data". The case-fatality rate was calculated by dividing the total sum of deaths in the year 2020 by the sum of total cases in the year 2020. The new cases in each location were taken by summing up all the new cases that occurred in 2020 in a particular location. Each dot on the graph represents a different location, although this gives us a general understanding of the global scale it does not provide an in-depth view of each countries struggles with Covid-19.

The graphs shows that locations with a lower number of new cases had a lower case- fatality rates and that the generally the greater number of cases the higher the case-fatality rate. The case-fatality rate shows a gradual increase from almost 0.00 at the locations with the least amount new cases up to around 0.10 then decreases as the number of new cases to just under 0.05 at the location with the highest number of new cases in 2020. Most locations were had around a 0.025 case-fatality rate. However, there were two outliers. One location had a generally lower number of new cases but had a higher fatality rate of over 0.25. The other outlier had over $8e7$ cases but only had a case-fatality rate around 0.025. Although the graph did show signs of appositive correlation it was a weak correlation as if a line of best fit were to be drawn it would not go through most of the points.

The two graphs differ by the scale and the colour of the. The second graphs x-axis has been formatted to a log scale. This spreads out the data this makes the data easier to read so that the viewer can more easily interpret the graph compared to the graph without the formatted x-axis. The graph without the formatted x-axis is much harder to read as most of the points are bunched to together, however the outlier stands out more in this graph as the massive difference in new cases in 2020 of the location with the most cases that year can be seen very clearly more so than in the graph with the x-axis formatted to the log scale. The y-axis on the second graph was not formatted so the outlier location with a high case-fatality rate and low number of new cases in 2020 can be easily interpreted from both graphs.

