Assignment 1 – Christina Feiner-Leek – INF6490

While examining the covid 19 global dataset, I initially found myself questioning how to conduct a meaningful analysis with relatively limited metrics and began exploring related sources to answer deeper questions on the long-term impacts of covid across the world. I considered bringing in data on doctors per capita, economic data such as GDP, as well as environmental data surrounding water quality and availability. Ultimately, I took a step back and questioned my initial impression that there were insufficient metrics, and formulated the question; “Did testing reduce the number of serious/critical cases, and/or deaths?”, allowing me to focus on the fields “serious\_or\_critical”, “total\_deaths\_per\_1m\_population”, and “total\_tests\_per\_1m\_population”.

A few considerations needed to be made while using these variables. First, the field serious\_or\_critical was not weighted by one million so could not be adequately used for comparison and had several countries without this data. The result of this was the countries with NULLs were omitted resulting in the dataset size being cut in half. Second, this data is in aggregate so assessing trends is difficult; we can’t say how many tests were used in a period and if there were impacts to covid outcomes during that same timeframe.

For my visualizations, I chose two line charts, and one geospatial chart. The first line chart shows an increase in deaths as testing numbers increase. It’s difficult to determine if this is a significant relationship, and we do not know the timeframe, so I interpret this as inconclusive at this time. The second line chart shows a decrease in serious or critical cases as testing increases. Once again, it is difficult to determine if this is a significant relationship, but I do believe we can assess that testing reduced the number of serious or critical cases in this dataset. The final chart I chose was a geospatial map describing the total tests per continent. What stood out to me was that the greatest number of tests were provided to Africa with Asia was provided the least per capita. None of these visualizations can truly stand on their own, and I look forward to improving my visualization skillset, and feel these individual charts tell important pieces of the covid 19 global dataset story.