# [Namya Kathuria] - Research Proposal

One of the most pervasive and widespread issues today is the global pandemic. Despite it having been nearly two years since the initial outbreak, we are still vastly uninformed on many aspects of the virus, its spread, and its cure. A large amount of research is being done to prevent it spreading further, but cases are still incrementing in the majority of countries worldwide. While biological and chemical research can help us in targeting the virus itself and strengthening people against it, we can also take protective and preventative measures nationally to reduce its transmission.

In the past, some countries have been very effective at doing so. New Zealand, for example, in 2020, was able to eliminate the virus completely – albeit temporarily – averaging zero cases per day. There are certain measures the country and government took, such as national lockdown, which contributed to this. By using such success stories as an example, it is possible that others can also reduce, if not eliminate, the spread of COVID-19.

It may be possible for us to rank countries by the effectiveness of their combative measures against the virus. The research would, naturally, take into account the population, as well as the population density, alongside important factors contributing to the capability of the country in targeting the virus, such as GDP or GNP, the amount of funding towards healthcare, the population below the poverty line (who may not be able to afford healthcare), ICU capacity, and potentially other factors not named here, which would also have an influence on how effectively a country is able to tackle the coronavirus. Then, factors regarding the country’s action should be evaluated; possibilities are lockdowns – including the length, and whether it was partial or full – as well as testing rates, vaccination rates, and mask mandates. Lastly, of course, we must take into account the spread of the pandemic by looking at the number of cases, separating serious/fatal cases, and deaths daily.

As aforementioned, this data could allow us to use top-ranked countries as role models for others. By emulating their actions taken in containing the pandemic, it may be possible to achieve similar effects worldwide.

There has been research on this in the past; one notable source is the ranking by the Lowy Institute. However, the factors considered in this ranking vary, including political distributions, and not including economic factors like funding towards healthcare, only considering the economic state overall. The ranking aims more closely to consider a country’s current state, rather than its action taking, when forming their ranking, not providing the insight required to actually work on solving the issue at hand. Furthermore, the ranking is not currently up to date, so the rankings may have changed, and finally, the ranking method used in this research is unstated and may produce different results to the ones we obtain, which may provide an interesting basis for us to compare the two rankings and evaluate the reasons behind these differences in future. An inconvenience of the pandemic is that information is not often widely available, or changes day after day. So, this type of research has not been very effective since the virus has newly emerged and data has been incredibly limited until now, or is constantly changing. While some research has been done, again, stating that by the Lowy Institute as an example, it is often focused on producing statistics rather than analytics on the factors behind the rankings produced.

There is data available regarding lockdown statuses and initial case dates on Kaggle, a popular data science website, which can be utilized alongside other datasets for this research. Data on many of the factors mentioned above, such as GDP, population, and poverty, is available – and reliable – on the official UN Data website. By scraping and combining these datasets, we can form a larger dataset with a collection of information in all the areas we may need.

A program that would be fit for this would be TrueSkill, which I could use to rank countries based on the factors previously identified. Other options may be Elo or the Colley method, in the event TrueSkill encounters problems or proves to be vastly inaccurate.

Since there is a limited timeframe for this research, it is more realistic to only rank countries within a specific region: in this case, countries in Asia. While some countries have responded efficiently and limited the spread of the pandemic, others – such as India – are still struggling, so this ranking may provide insightful results. This may be especially so as Asia is a very diverse continent with countries varying in population density, GDP per capita, and other factors contributing to the effect of their COVID-19 response.

As this research concentration is particularly focused on coding the ranking algorithm itself, the timeline will be unconventional in that the first few weeks may be more code-oriented and less focused on writing, barring the literature review. While there may be modifications, the rough timeline will be as such: by Tuesday 17 August, the literature review should be completed, and an outline of the ranking algorithm defined in pseudocode, if not TrueSkill itself. In the week following, the aim will be to work on the code and complete a preliminary ranking of a small sample of countries to identify any errors and observe and comment on the ranking results produced. As well as this, a first draft of the paper should be complete, using the initial results produced; naturally, a conclusion cannot be drawn. By three weeks from now, the goal will be to have completed the ranking algorithm in TrueSkill and have begun analysis on the results, which will be done, hopefully, by coding some form of dependency algorithm where the quantitative impact of certain factors (on the ranking) can be calculated. Alongside this, a complete draft of the paper should be written, including the literature review, a revised description of the methodology used (in case changes have been made after the initial draft), and a full analysis of the results, which will be updated as analysis progresses. An initial research question will also be identified after the first week, which may be revised to be more precise as the research is worked on.

The takeaway from the completed research, then, is not the ranking itself, but the factors which contributed to producing the ranking: post-research, we must analyze patterns and common factors in top-ranked countries, and the same in bottom-ranked countries to present a hypothesis as to why these countries are in the ranking positions they are, and suggestions on action they can take to then be more efficient in tackling the spread of the pandemic. If, for example, multiple countries within the top 20 expose a trend of cases declining and the implementation of full lockdowns for more than two months, we can consider implementing these lockdowns in other countries and potentially expect a similar result.

During a time where scientific information is not immediately available, we can use other methods – and the implementation of rankings and mathematics – to address problems in ways that may be, in fact, more useful and more plausible by governments and the general public.