Final Project Code (Your title here)

5/6/2021

1 Introduction

When COVID-19 first hit the world, no one knew its potential impact. With vaccination underway, there is increasing hope that the world will return back to normal soon. In this project, we hoped to learn more about how the vaccination effort is proceeding in continents around the world, and also compare countries and continents with one another to understand how well they are combating COVID-19. In particular, we examined the global vaccination rate and how the continents and specific countries compared to each other in their vaccination efforts. We obtained our two major data sets, vac_time_series and global_vac_data, from the John Hopkins Coronavirus Resource Center through this link: https://bit.ly/2Q2ey6e (https://bit.ly/2Q2ey6e). The vac_time_series data set records the total number of vaccination and vaccinated people in each country or continent each day over a span of time from December 2020 to May 2021. The global_vac_data shows information on the total amount of administered Covid-19 vaccines and vaccinated people in each country/continent in April. In addition, for section 2.5 we used data from the following links: https://bit.ly/33v1wl6 (https://bit.ly/33v1wl6) (Asian countries) and https://bit.ly/3f5lCaN (https://bit.ly/3f5lCaN) (European countries).

One of the goals of our project was to understand how the continents compared to each other in total number of vaccines administered and whether the average temperature of the continent is correlated with the number of administered vaccines since many vaccines need to be stored at freezing temperatures. We hypothesized that as the continents have increasingly higher temperatures, the number of total vaccine doses administered will be lower since it will be harder to store the vaccines properly. The average daily temperatures in degrees Celsius of the continents are as follows: Europe = 10.7, Asia = 18.3, North America = 15.0, Africa = 22.0, and Australia = 13.9 (World Data, 2020). Considering this, we hypothesized that of the six continents (excluding Antarctica), Africa will have the fewest doses administered, followed by Asia, North America, Australia, with Europe having the highest number.

Another question we investigated was how the U.S. compares to the other two major North American countries, Canada and Mexico, in the total number of doses administered and the proportion of administered vaccines to the respective population of each country. We hypothesized that the three countries will start out the same, but as time passes, the U.S. will have the most total administered vaccines and the highest proportion of total administered vaccines to the population, followed by Canada and Mexico. To answer this question, we created

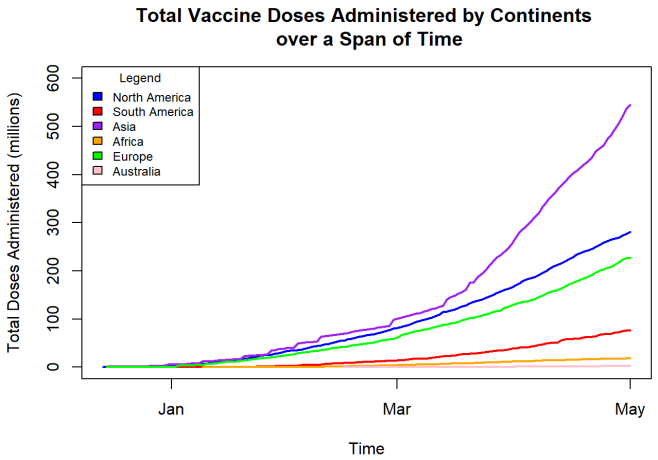
two plots: one with total doses of each country over 1.5 years, and the second plot compared the proportion of total administered doses to the population of each country. The US population is 332,727,601; the Canadian population is 38,039,789; and the Mexican population is 130,169,933 ("List of North American Countries by Population", 2021).

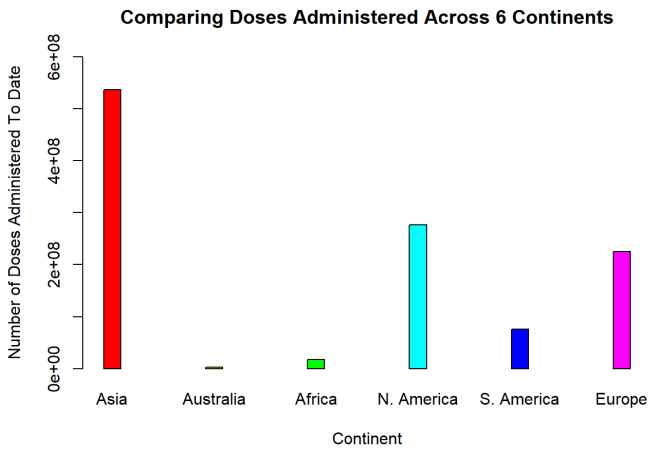
We also investigated how Asian countries performed in their vaccination effort, since Asia is the world's most populated continent. We chose countries including China, Thailand, and Kazakhstan to represent East, Southeast, South, Central, and West Asia; we then compared the number of fully vaccinated people in these five countries. We hypothesized that these countries will start out with roughly the same number of fully vaccinated people. However, as time passes, China would have a significantly higher number of people fully vaccinated than the other four countries and India would have the second-highest number of fully vaccinated individuals. We also researched whether the vaccination rate is correlated with a country's wealth. We hypothesized that countries with higher GDP will be able to administer more vaccines than countries with lower GDP, as they would have more money to buy vaccines. Finally, we also wanted to investigate the relationship between the population size of a country and the percentage of the population that is fully vaccinated. To examine this, we chose the five most and least populated countries from both Asia and Europe and compared their vaccination data. Our hypothesis was that countries with larger populations would have a smaller percentage of their total population vaccinated since a higher population would require more resources, time, and money to vaccinate each person. Thus, we believe smaller countries will have a larger percentage of their population vaccinated.

2 Analysis/Visualizations

To access the data, first download the Excel files that are included in the submission. Then inside the ()of the read_excel(), function enter the downloaded Excel files path on your computer. Where there is a \ in the file, you may need to change it to \\ if you have a Windows machine. I.e. read_excel("path of downloaded excel file")

2.1 Comparing the 6 Continents (Excluding Antarctica)



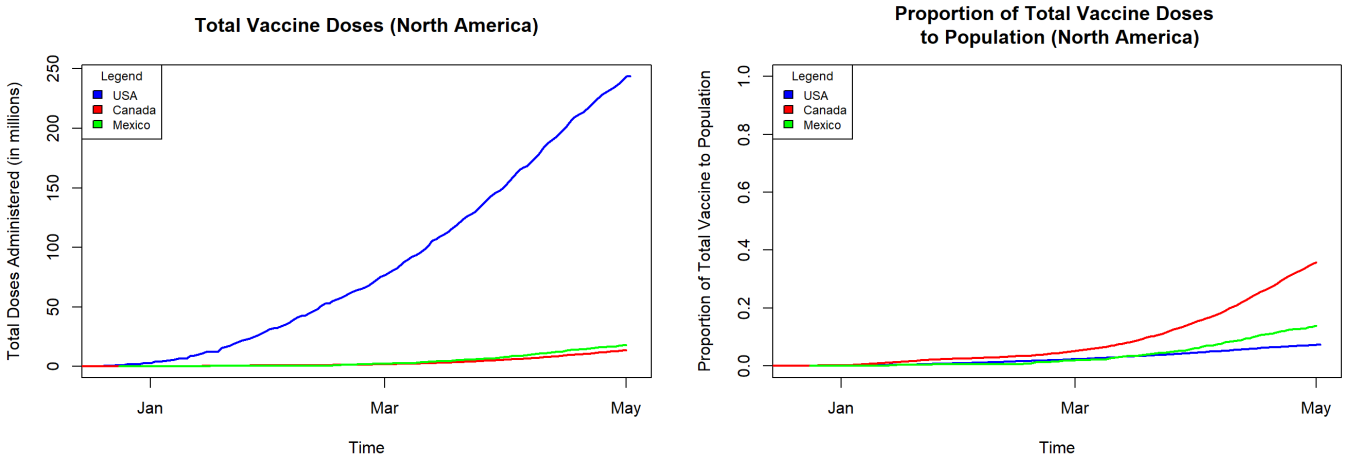


Analysis:

Looking at the line graph, Asia leads in total vaccine dosages administered and is ahead of North America and

Europe starting from late March/early April. It shows that the total amount of vaccine doses administered by North America and Europe is not significantly different. Moreover, it seems like South America administered more doses starting from late March. The total number of doses administered by Australia is not significantly different from the total amount of vaccine doses administered by Africa, and they both have lower numbers of vaccine doses administered than the other five continents. The barplot shows similar results: Asia has the highest number of vaccine doses administered, while Australia has the lowest. The order of number of vaccine doses administered, in order of greatest to least, is the following: Asia, North America, Europe, South America, Africa, and Australia.

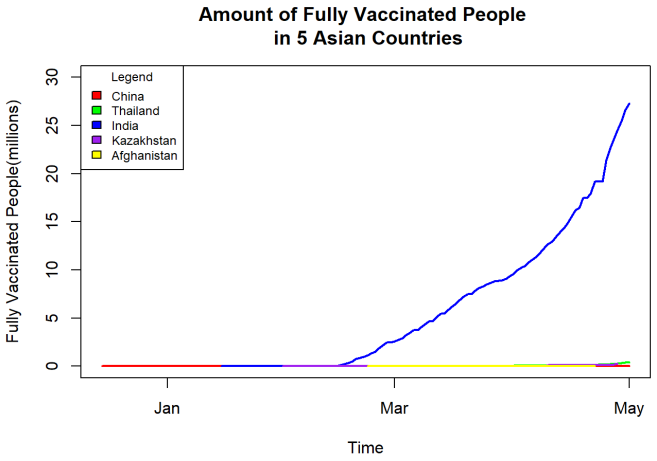
2.2 Comparing the 3 Major North American Countries

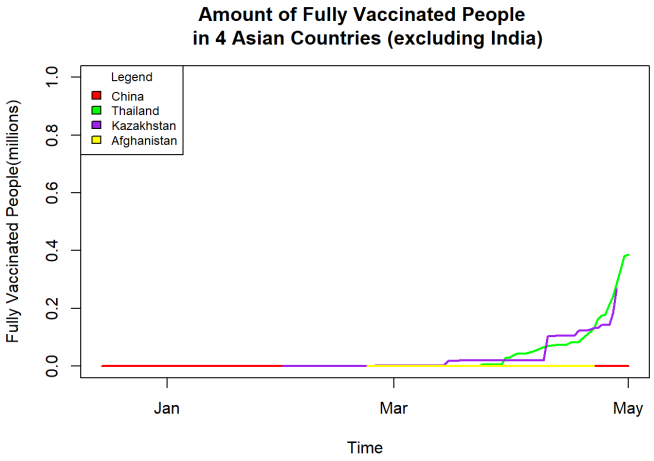


Analysis:

From the first plot, starting from late January/early February, the U.S. lead in total vaccine dosages administered. However, the total amount of vaccine doses administered by Canada is not that different from the total amount of vaccines administered by Mexico, and it appears Mexico has administered more vaccines than Canada in the end. The second plot shows that in the proportion of total vaccines to population, the three countries started at about the same proportion, but Canada currently has the lead, followed by Mexico, with America last.

2.3 Comparing Number of Fully Vaccinated People in 5 Asian Countries

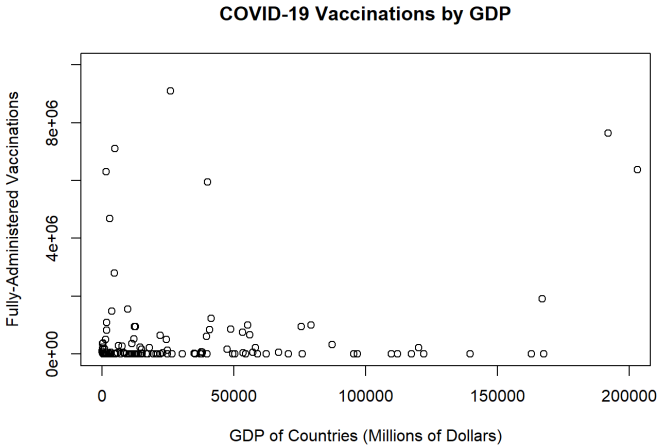




Analysis:

The second graph was included because in the first plot, India's high number of fully vaccinated people is causing the other four countries (China,Thailand, Kazakhstan, and Afghanistan) lines to blend together. Thus, in the second plot, we decided to plot only the trend line for those four countries as to get a clearer picture of how they compare to each other in the number of fully vaccinated people. o The result of the graphs were very surprising, especially the trend line depicting the number of fully vaccinated people. From the resulting graph, starting from February in 2021, India lead in the number of fully vaccinated people as there was almost no fully vaccinated people in China, Thailand, Kazakhstan, and Afghanistan as compared with India. Matthew's graph. Looking at the second plot, it appears that Thailand is second and Kazakhstan is third.

2.4 GDP and Vaccination

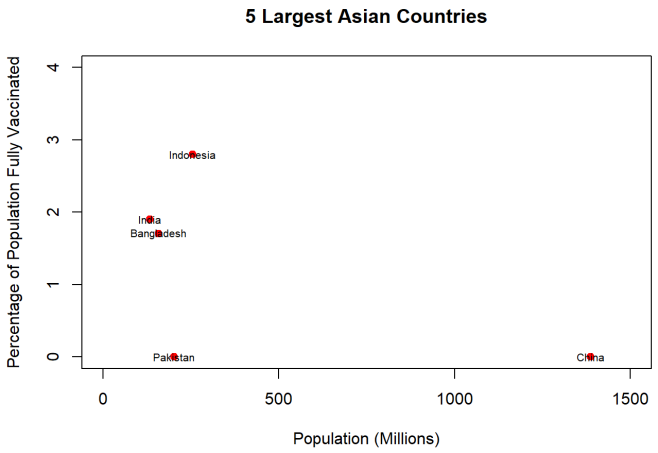


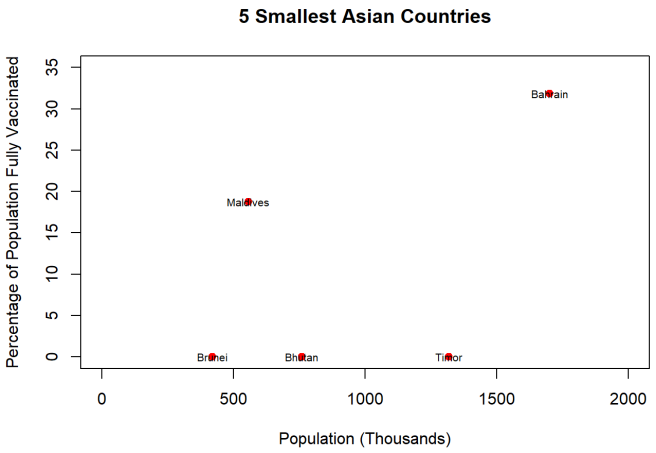
Analysis:

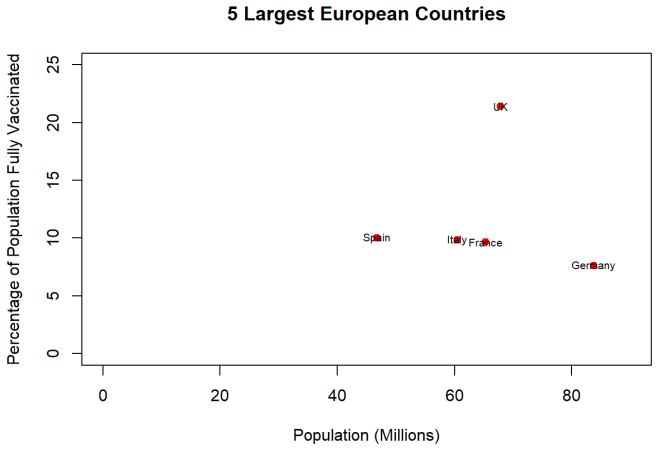
The data is accurate as of April 2021 and considers individuals who are fully vaccinated (having received both doses of the two-dose vaccine). It is scattered across varying GDP's, and a few countries with low GDP's have high vaccination rates. However, countries with 0 vaccinations are especially concentrated in the region of low GDP. We see that, although some countries with higher GDP's have zero vaccinations, the amount of countries with 0 vaccinations is much higher in the low-GDP area than it is with the wealthier countries.

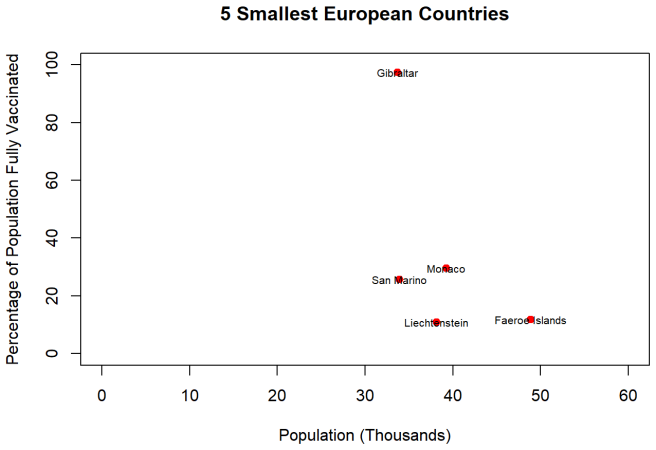
2.5 Population Size and Percentange of Population

Vaccinated









Analysis:

On average, we see that the countries with a smaller population size have a higher percentage of their population fully vaccinated. It is interesting that some of the biggest countries (population wise) such as China and Pakistan had 0% of their population fully vaccinated at the time this data was recorded. It is also interesting that many other small Asian countries had 0% of their population fully vaccinated. All European countries, however, had some percentage of their population fully vaccinated. In some small countries such as Gibraltar, the percentage of the population fully vaccinated was close to 100%. This reinforces the notion that smaller countries have a larger percentage of their population fully vaccinated.

3 Conclusions

From our analysis, there were some surprising and expected conclusions. First, in section 2.1, both the bar plot and the line graphs showed that the highest number of doses is administered by Asia and the lowest is by Australia. This refuted our hypothesis. The order from the most to least administered doses was as follows: Asia, North America, Europe, South America, Africa, Australia. Additionally, in section 2.2, our hypothesis was shown to be partially correct. Although the U.S. leads in doses administered, it is last in proportion of vaccines delivered to the population. Moreover, Canada has the highest proportion of vaccine to population although it had less total administered doses than Mexico.

These results can be explained by several factors. One of them is the economical state of the continents. The share of the global GDP by continent is the following: Asia = 36.9%, North America = 28.9%, Europe = 23.9%, South America = 5.1%, and Africa = 3.1%, and Oceania = 2.1% (Oceania includes Australia) (Ghosh, 2020). Interestingly, the order of continents by share of global GDP, highest to lowest, is the same as the order by number of doses administered. Wealthier continents likely have more resources available to collect this data than poorer continents, meaning their numbers are higher. In addition, wealthier continents can spend more money on purchasing vaccines while poorer ones cannot afford to. In conclusion, it seems GDP would be a better indicator to predict trends for the number of vaccine doses administered than the temperatures of continents.

This explanation through GDP matches up with the comparisons in section 2.4.The resulting plot largely supported our hypothesis that countries with higher GDP will have higher total vaccinations than those with lower GDPs. Although there were some countries with high GDP and low administered vaccines and vice versa, the majority of the countries with dosages were in the lower range of GDP as shown by the plot; this reinforced our hypothesis. Another explanation is the difference in population: Asia has the largest percent of the global population at 61.8% (Ghosh, 2020). This could explain why in section 2.1, both the line graph and the bar chart showed that Asia leads consistently over the past few months. Moreover, in section 2.2, comparing the 3 North American countries, this could also explain why the U.S. leads in total administered doses.

However, a higher population may mean that a higher total vaccine administration would only cover a small percentage of the population since more resources and time would be needed to obtain and administer enough vaccine to the entire population. This could explain why in the second plot comparing the proportion of administered vaccines to population, the United States was last. In section 2.5, we explored this relation and found that nations with higher populations generally had less percentage of their population fully vaccinated, while countries with smaller populations had higher percentage of their population vaccinated. Another interesting result from section 2.5 was that many countries in Asia, even those with high population and high GDP, like China, had 0% of their population fully vaccinated. This finding is similar to the results of section 2.3: India leads in vaccinations while China has 0% of its population fully vaccinated. This could be explained by the fact that since China has a developed health care system and stricter regulatory policies, the proportion of COVID-19 cases in China has been greatly reduced, causing people to feel less compelled to fully vaccinate. Compared with China, India is more eager to complete vaccination, especially given its current circumstances.

Our methods have their strengths and limitations. One of the strengths of our methods is that our line graphs allowed us to observe the trend over a period of time of vaccination efforts in countries/continents which would be beneficial to us if we were to predict future vaccination efforts. The strengths of the scatter plots used in the last two subsections are that they can show correlation between two variables, hinting at possible relationships between them. The limitations of our methods though is that our scatterplots show only correlation and cannot be used to show any causational relationship. Other limitations of our methods are that they only allow us to focus on one or two variables in the global vaccination effort when, in reality, global vaccination effort is often influenced by multiple factors at once.

4 Future Work

With more time, tools, and resources at our disposal, we potentially could have analyzed larger quantities of data over a longer period of time. As more of the world's population continues to become vaccinated, more and more research is available every single day. With access to more time and resources, we could have had the opportunity to analyze more of this most recent data. Having access to more recent data and COVID vaccination records would lead to an increasingly more accurate analysis of how the COVID vaccine is being distributed and whether or not it is effective. For example, the percentage of people fully vaccinated in a country

such as China has now increased from 0%, at the time when we first examined this data to now 4% (311 million people). (ourworldindata.org/coronavirus) This is one example of something we as a group could have taken into consideration when we were doing our research if we had access to more time and resources. Other issues we could explore include more of the reasons for disparities in vaccine administration in countries or why some countries performace in vaccine rollout is different from what we had previously hypothesized.