Covid 19 World Vaccination Progress

Project Report

By Bria James

Due: December 3, 2024

<u>Objective</u>:

The objective of my project is to take a look and examine global Covid 19 vaccination data to discover trends, patterns, and insights about vaccination efforts all over the world. By looking at factors such as different countries, total vaccinations, and daily vaccinations, I will be able to find out what countries are leading in vaccination data, what areas might need to be looked at for improvement, and see the progress that vaccinations have made throughout time. The data obtained from this project can be used by many roles including policymakers, health organizations, government officials, and even simply students like myself who are interested in learning about where they stand in comparison to the rest of the world. All of this is to help us understand the impact and reach of vaccination efforts and campaigns all over the world.

Introduction:

The Covid 19 pandemic truly did change the world and it was never the same. It affected households across the world on every level, from education with schools having to move online, to professional with many people's careers and livelihoods being halted and completely changed, to socially as well, with people everywhere living with the hesitation or fear that they might catch the virus or spread it to their loved ones. With that, something else that became more prevalent in people's lives around the world was vaccines. The dataset used in this analysis contains information on COVID-19 vaccinations from various countries, including metrics such as total doses administered and daily vaccination rates. This dataset provides a unique opportunity to see how nations have tackled one of the biggest public health challenges in modern history and also assess the progress made in combatting the pandemic.

<u>Method</u>:

Before I can do any kind of analysis, I need to make sure the data is clean and ready for further examination. This means no missing values or duplicates, just data that is accurate, consistent, and meaningful. As I analyzed further, I realized that there actually were a lot of values that I needed to handle. For example, I used the forward and backward fill method for time series columns to fill in what was missing. Then, I decided to impute numeric columns with the median. This includes columns such as total vaccinations per hundred and daily vaccinations per million which had their missing values replaced with the median value of the respective column to minimize bias. Columns with more than 60% missing data were dropped, as they needed more information for meaningful analysis. Once missing data was taken care of, I could then go on to handle outliers and validations. Thankfully, there were no duplicates in this dataset and there were also no invalid dates. Outliers were identified using the Interquartile Range method. For my dataset, an outlier was defined as any value that is significantly lower or higher than the typical range for metrics such as daily vaccinations, total vaccinations, or people fully vaccinated. Rows with extreme values were filtered out to ensure that anomalies did not skew the analysis. Now that the outliers have been removed and the data has been validated, I check back and see that there are no missing values left in this data set. Lastly, I do a little bit of additional cleaning, making sure all of the states' names are capitalized and in uniform. I have now completed the data wrangling and cleaning so now the data set is ready for analysis.

Storytelling: Data Visualization & Interpretation

Now I can examine some graphs that will help us better understand the trends and patterns in the vaccination data. These visualizations provide a clear and concise way to interpret datasets that are otherwise too complicated and large to understand. So first here is a graph of Global Total Vaccinations Over Time:
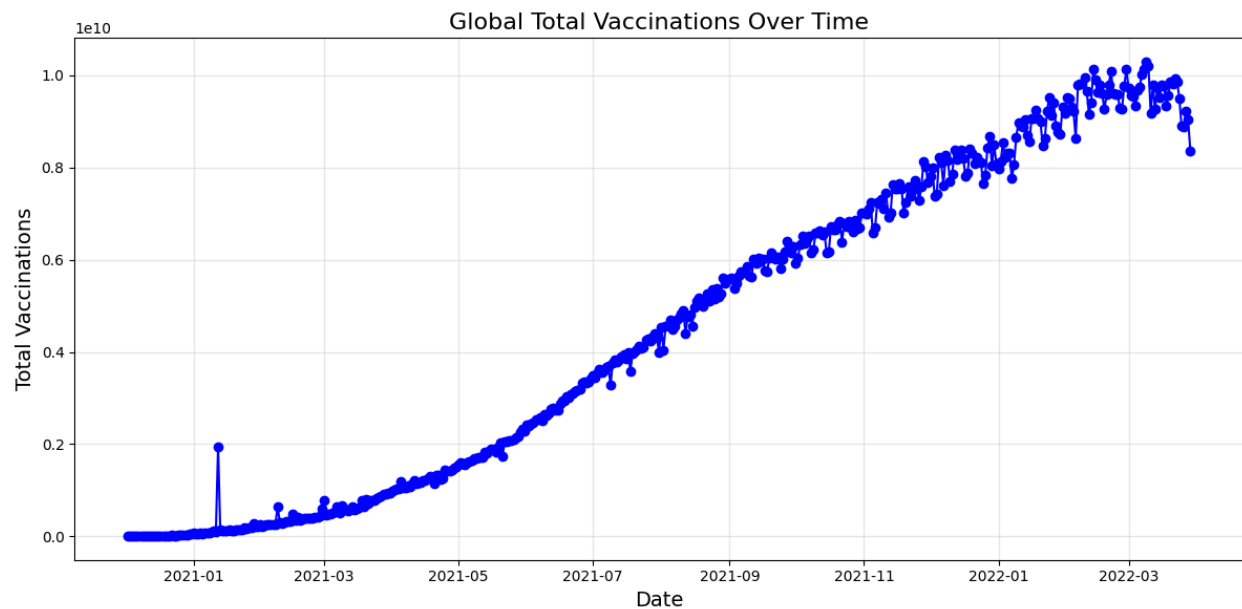

As we can see the total number of vaccinations globally has been on a pretty steady and consistent increase from 2021. To me, this means that all over the world, people have been advocating for the COVID-19 vaccine and the campaigns for these vaccines to go into use have actually been sustained. The rate of increase appears steeper from mid-2021 to late 2021, reflecting the high vaccination periods as vaccines most likely became widely available and distribution networks expanded globally. Toward early 2022, the curve starts to flatten slightly, indicating that vaccination rates may have slowed down. I think this is because many people have already been vaccinated by that point.

Next, I examined global vaccination trends, but on a daily level:
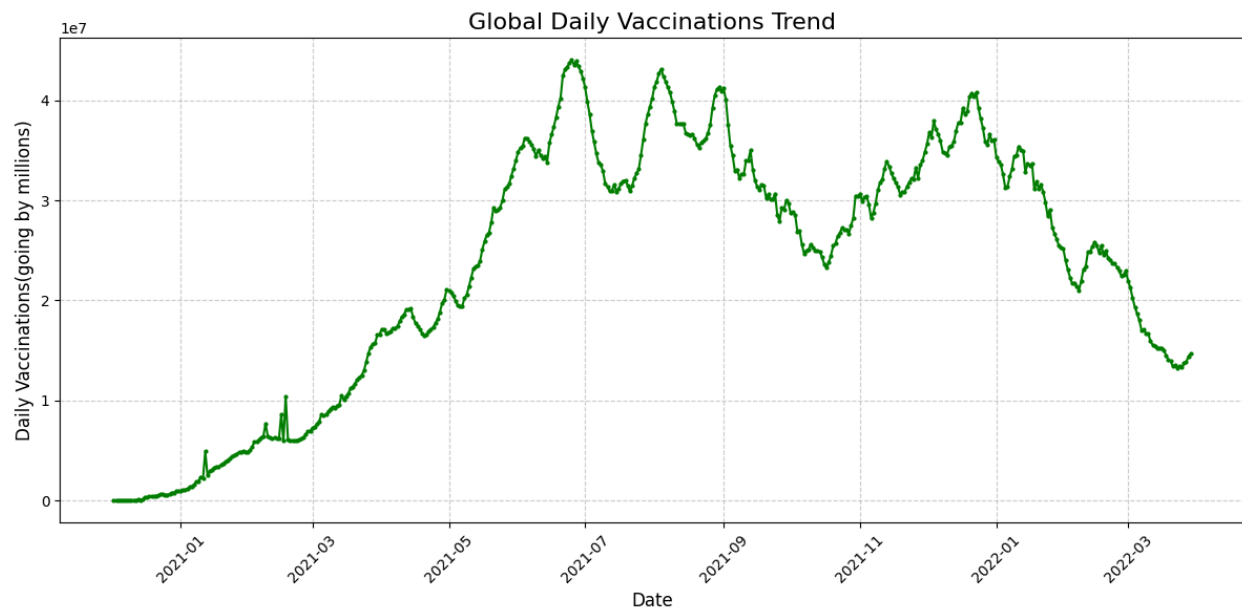
Again, from the beginning to the middle of 2021, we can see a pretty steady increase in the number of daily vaccinations, which proves to me even more that this is when the vaccine became highly available and accessible. Several peaks throughout 2021 suggest moments of maximum vaccination activity, likely driven by policy pushes, vaccine mandates, or even an increased supply. From late 2021 into early 2022, daily vaccination rates began to decline, which to me signifies that campaigns for people to get vaccinated slowed down or there was a reduced need for them as so many people were already receiving them.

This next graph represents the top 5 countries with the highest total vaccinations and their progress over time:

Top 5 Countries by Total Vaccinations Over Time
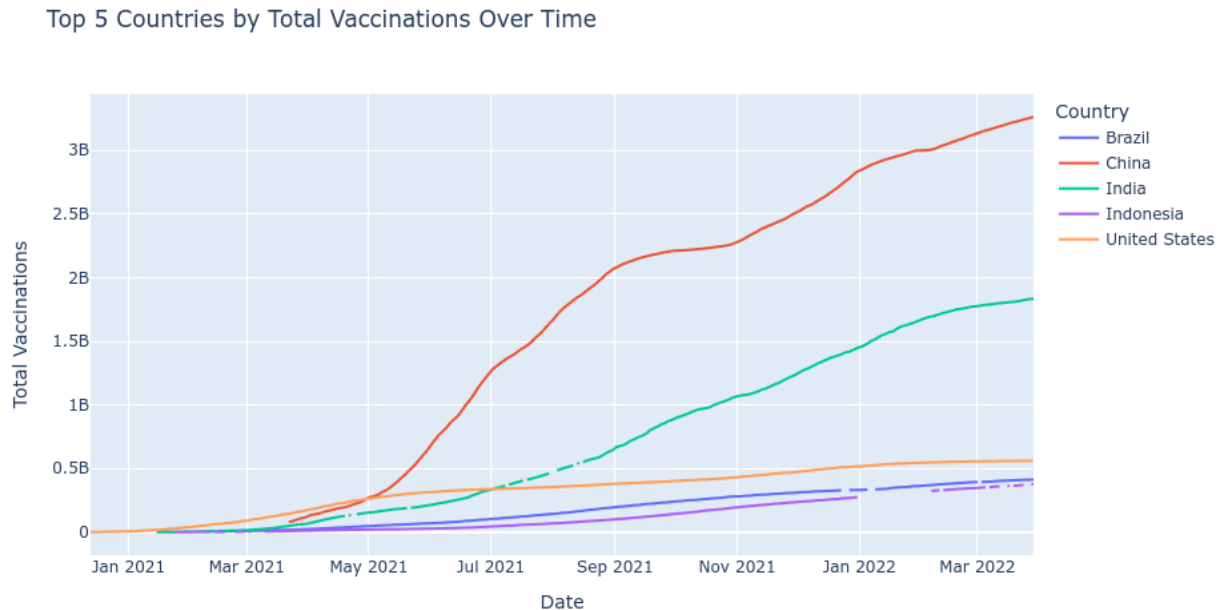
When it comes to total vaccinations, it is clear that China is in the lead. China shows a steep and consistent upward trajectory, far surpassing other countries with over 3 billion total vaccinations by 2022. This reflects China's large population and aggressive vaccination strategy. India follows China with a steady increase, reaching over 1.5 billion total vaccinations. The upward curve highlights the country's mass vaccination campaigns to address its large population as well. The United States, Brazil, and Indonesia show slower but steady progress compared to China and India. These countries reflect efforts to vaccinate large populations at different paces due to varying vaccine availability, population sizes, and public health strategies. The breaks in lines are due to me removing the outliers from the data set.

Next is a heatmap showing the correlations between all the numerical columns and a bar graph showing countries that use more than one vaccine and the count of vaccines used per country:
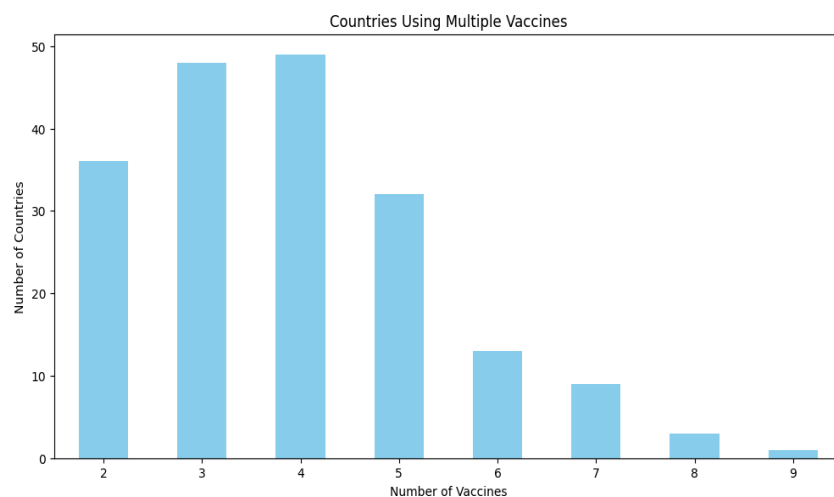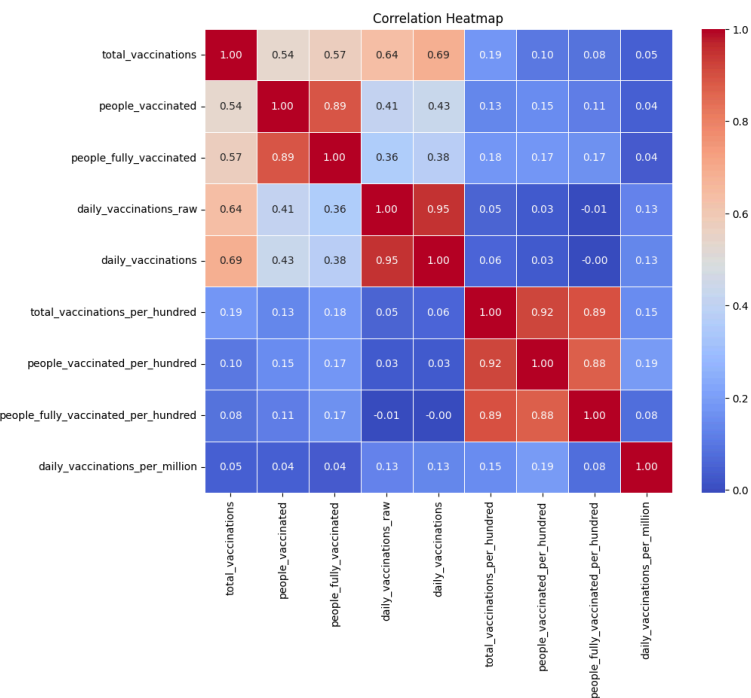
The heatmap shows a strong correlation between total vaccinations with daily vaccinations (0.69), indicating that higher daily vaccination rates lead to higher total vaccinations. People vaccinated and people fully vaccinated also have a very high correlation, which suggests that as more people are partially vaccinated, many also progress to full vaccination. Categories like daily vaccinations per million and people vaccinated per hundred have weaker correlations with other variables, reflecting their relative independence from the total population metrics.

As for the Multiple Vaccines Graph, this shows us that the majority of countries use 3 or 4 different vaccines, with each category having nearly 50 countries. This suggests that countries are relying on a diverse range of vaccines to make sure their populations are being vaccinated.

Next is to examine a graph for Daily Vaccinations per Capita and the comparison of Vaccinated vs Fully Vaccinated:

Daily Vaccinations Per Capita (Top 5 Countries)

Comparison of People Vaccinated vs Fully Vaccinated
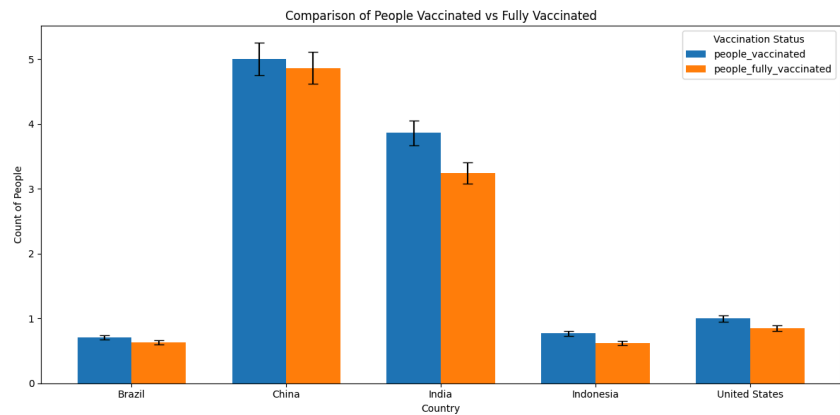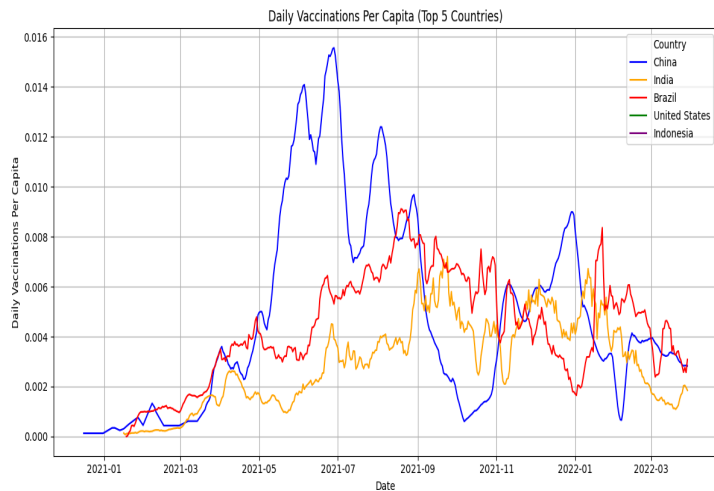
For vaccinations per capita, China shows sharp peaks, especially mid-2021, indicating intense vaccination campaigns. Brazil maintains steady vaccination rates with less dramatic fluctuations. India shows gradual growth in early 2021, followed by a decline after reaching its peak. With the comparison bar plot, you can see that China has the highest number of both partially and fully vaccinated individuals, indicating its large-scale vaccination efforts. India seems to have the largest difference between their vaccinated and fully vaccinated population.

Lastly, I decided to do a bubble chart:

Vaccination Momentum: Daily Vaccinations vs Coverage
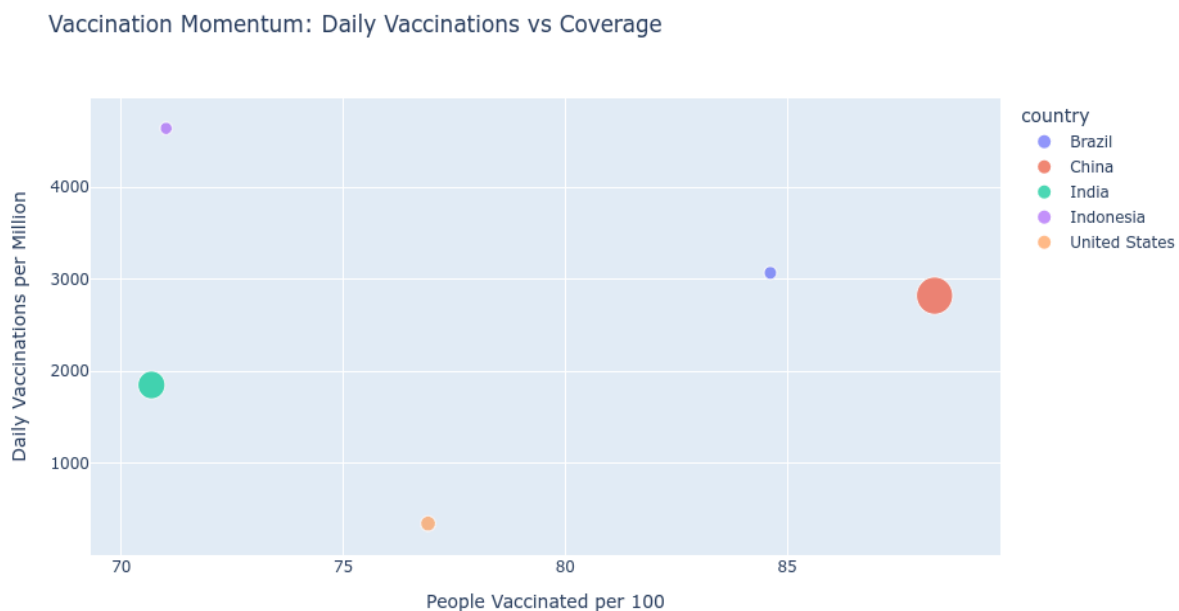
I thought it would be nice to see how countries compare in terms of their vaccination momentum, combining daily vaccination rates, population coverage, and overall progress. The x-axis represents People Vaccinated per 100, the y-axis represents Daily Vaccinations per Million, and the bubble size represents Total Vaccinations. This way I can tell how China has the highest vaccination coverage (around 85 people vaccinated per 100) but a relatively lower daily vaccination rate compared to countries like Indonesia, suggesting they might be more focused on maintaining immunity or administering boosters.

Conclusion:

In conclusion, this project led me to learn many things about the great efforts that people put in globally to ensure various populations were vaccinated. I recall hearing about anti-vaxxers and it seemed like they were such a large population, but it was refreshing to examine data like this and see that the whole world really was trying to combat Covid in the ways we could. By examining metrics like total vaccinations, daily vaccination rates, and population coverage, I gained insights into how different countries have approached their vaccination rollouts. Countries like China and India demonstrated large-scale efforts to vaccinate their populations rapidly (especially since they are so large), while others, like Indonesia, highlighted steady momentum despite lower overall coverage. I would have liked to see more demographic-specific groups, like ages, gender, and race, just to take an even closer look at the data. Overall, the project was very informative.

Reference:

Our World in Data: https://ourworldindata.org/coronavirus (Used for background research)

Acknowledgement:

I would like to sincerely thank my Professor, Dr. Nerolu, for their guidance and support throughout this course. As a non-STEM major, and liberal arts student, learning data analysis and working on this project was new and challenging for me. Your patience and clear explanations made a significant difference in helping me navigate this unfamiliar territory. I truly appreciate the time and effort you put into making this course accessible and engaging for students like me!