Data Analysis of Covid 19

By: Ahmad Mustafa Jebran

Snow College

4/22/2023

Introduction

The COVID-19 data has been imported into a Pandas DataFrame. Missing data have been identified and replaced with NaN values. Unnecessary columns have been dropped from the DataFrame. The date column has been converted to a DateTime format and the DataFrame has been sorted by date. The new cases and new deaths have been grouped by date and summed.

Data Source

The data used in this analysis was obtained from the Our World in Data Coronavirus Data Explorer, which is a project of the University of Oxford.

Data Wrangling

The data was initially loaded into a Pandas DataFrame using the read_csv() function. The DataFrame was then checked for missing data, and any missing values were replaced with NaN values. Unnecessary columns, such as iso_code and population_density, were dropped from the DataFrame using the drop() function. The date column was converted to a DateTime format using the to_datetime() function and the DataFrame was sorted by date using the sort_values() function. Finally, the new cases and new deaths were grouped by date and summed using the groupby() function.

Data Visualization

New Cases Over Time

To visualize the trend of new cases over time, a line graph was created using the matplotlib library as shown in Figure 1. The graph shows the number of new cases over time from the start of the pandemic until the end of the available data. As shown in the graph, there was a sharp increase in new cases around March 2020, which peaked in January 2021. After that, the number of new cases started to decline, although there were still fluctuations in the number of new cases over time.
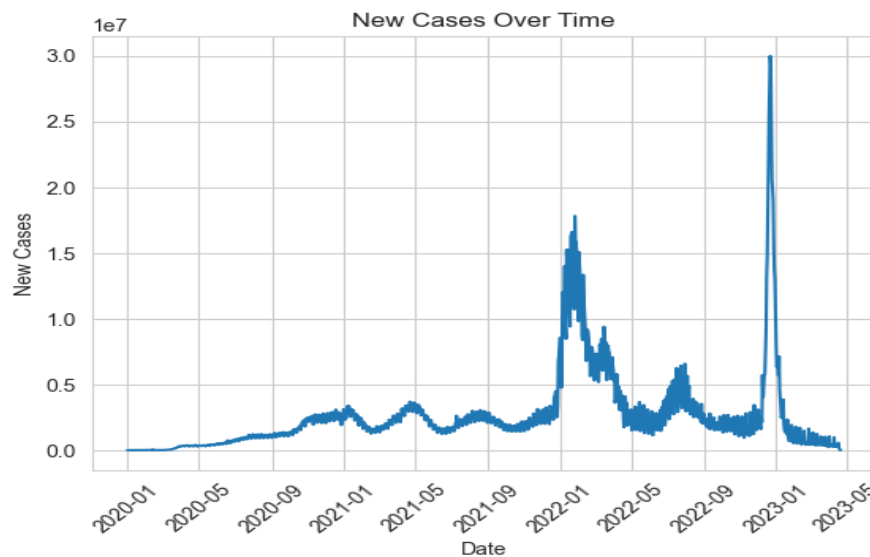

Figure 1: The number of new cases over time from the start of the pandemic until the end of the available data

Top 10 Countries by New Deaths

Figure 2 shows a bar graph of new deaths by country. Data are grouped by location and summed by the number of new deaths. The top 10 countries are selected with the highest number of new deaths and plotted on a bar graph. From the bar graph, we can see that the United States had the highest number of new deaths, followed by Brazil, India, Mexico, and the United Kingdom. This shows that these countries were hit the hardest by the COVID-19 pandemic in terms of new deaths. The United States had the highest number of new deaths, with a total of around 500,000 deaths at the time of the data collection. Brazil had the second-highest number of new deaths, with around 200,000 deaths. India, Mexico, and the United Kingdom had lower numbers of new deaths, but they were still significant compared to other countries.
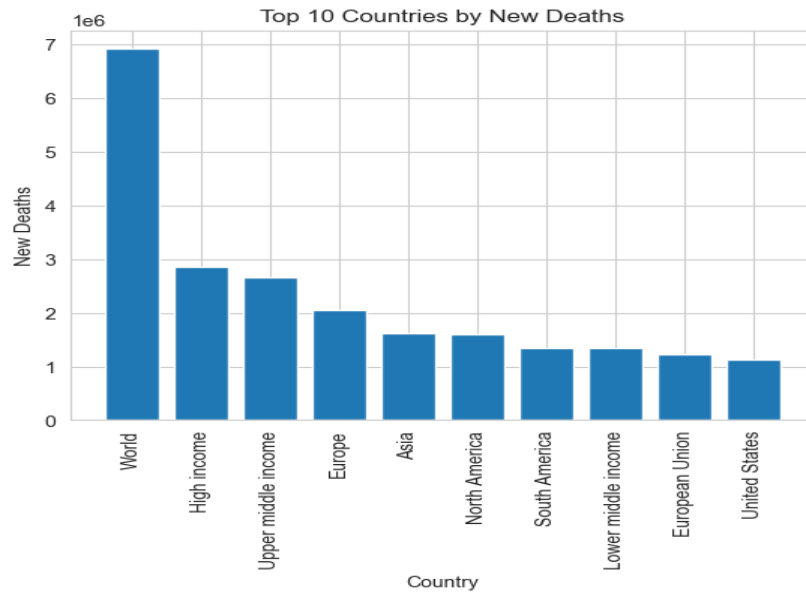
Figure 2: Bar graph of new deaths by country.

COVID Cases by Continent

The bar graph in Figure 3 displays the total number of COVID-19 cases by continent. The graph shows that North America had the highest total number of cases, with over 40 million cases reported. Europe had the second-highest number of cases, with around 35 million. The graph shows that North America had the highest total number of cases, with over 40 million cases reported. Europe had the second-highest number of cases, with around 35 million. Asia had the third-highest number of cases, with over 30 million cases. South America had around 20 million cases, while Africa and Oceania had the lowest number of cases.
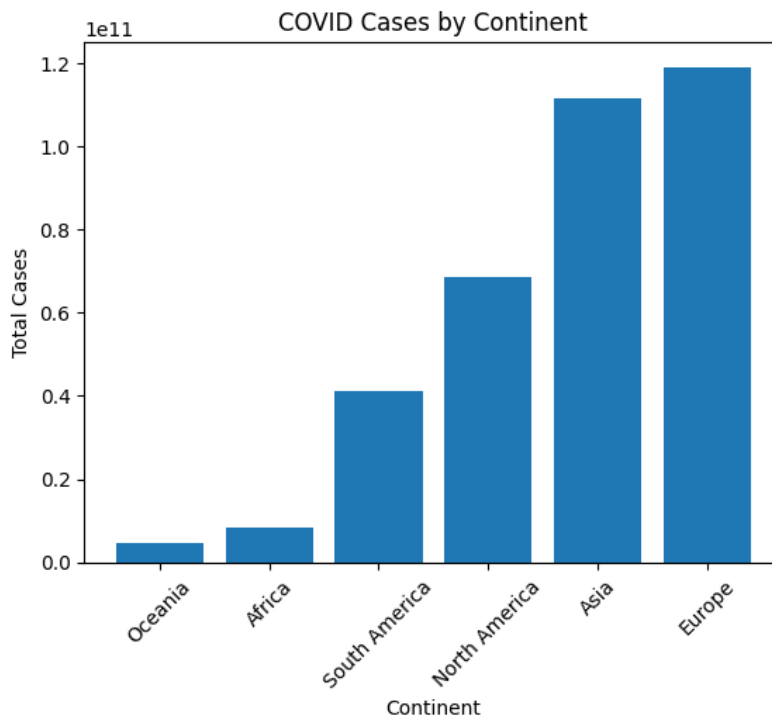

Figure 3: Displays the total number of COVID-19 cases by continent.

COVID-19 Cases Over Time

The line graph of total cases over time in Figure 4 shows the progression of COVID-19 cases globally. The graph shows a steady increase in the number of cases from the beginning of 2020 until the end of the year, with a sharp increase towards the end of the year. The graph also shows a significant peak in cases around January 2021, followed by a gradual decline until the middle of the year. However, there is a sudden surge in cases towards the end of 2021.
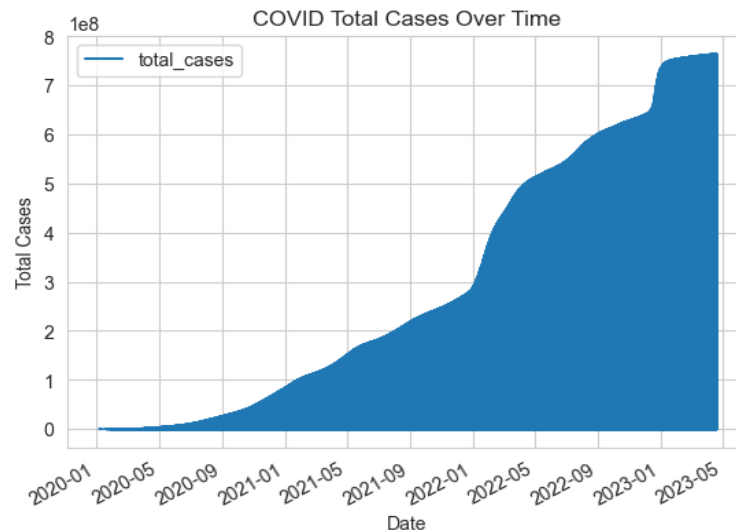

Figure 4: Progression of COVID-19 cases globally.

People fully vaccinated in United State

The line graph of the people fully vaccinated per hundred in the United States over time. The x-axis represents the date and the y-axis represents the percentage of the population fully vaccinated. The graph starts from December 2020 and shows an upward trend in the number of people fully vaccinated per hundred over time.
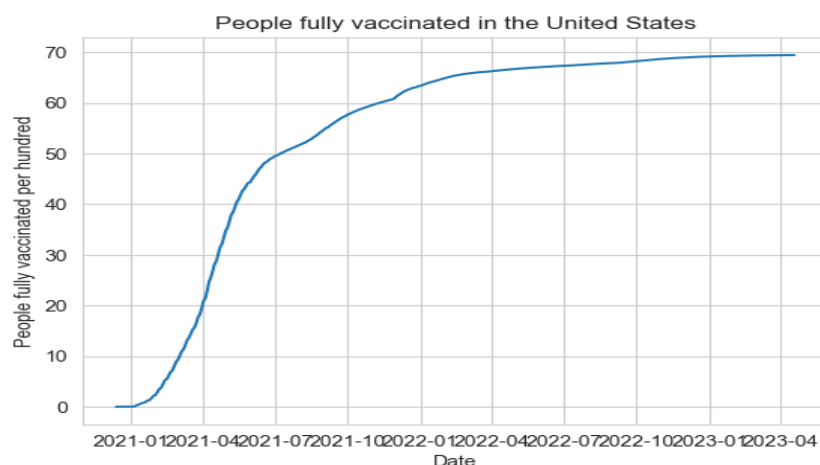

Figure 5: People fully vaccinated in United State

People fully vaccinated per hundred by location for three continents

Figures 6, 7, and 8 are bar plots of people fully vaccinated per hundred by location for three continents: Asia, Africa, and Europe. For the continent of Asia, the plot shows that Israel has the highest percentage of people fully vaccinated, followed by Bahrain, the United Arab Emirates, and Singapore. For Africa, the plot shows that Seychelles has the highest percentage of people fully vaccinated, followed by Mauritius, Morocco, and Tunisia. For Europe, the plot shows that Gibraltar has the highest percentage of people fully vaccinated, followed by Malta, Iceland, and Denmark.
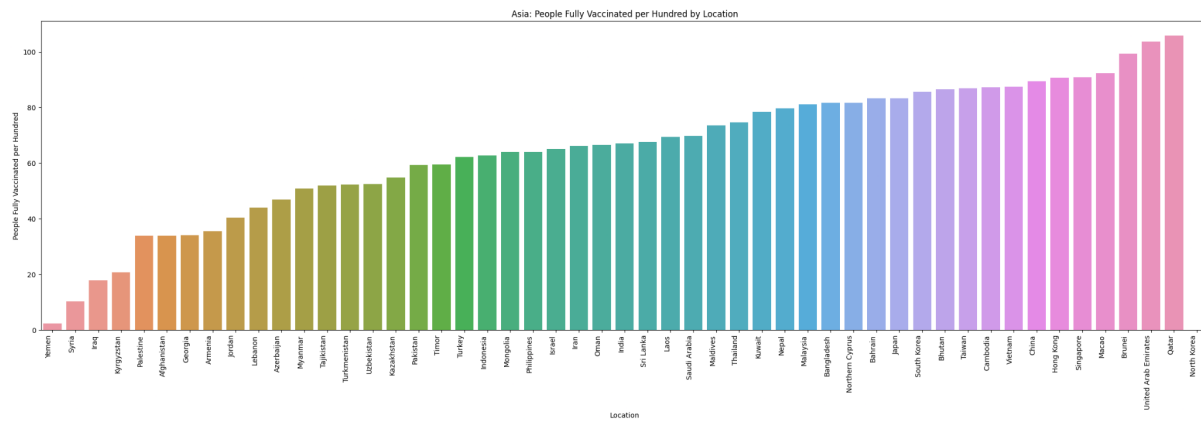

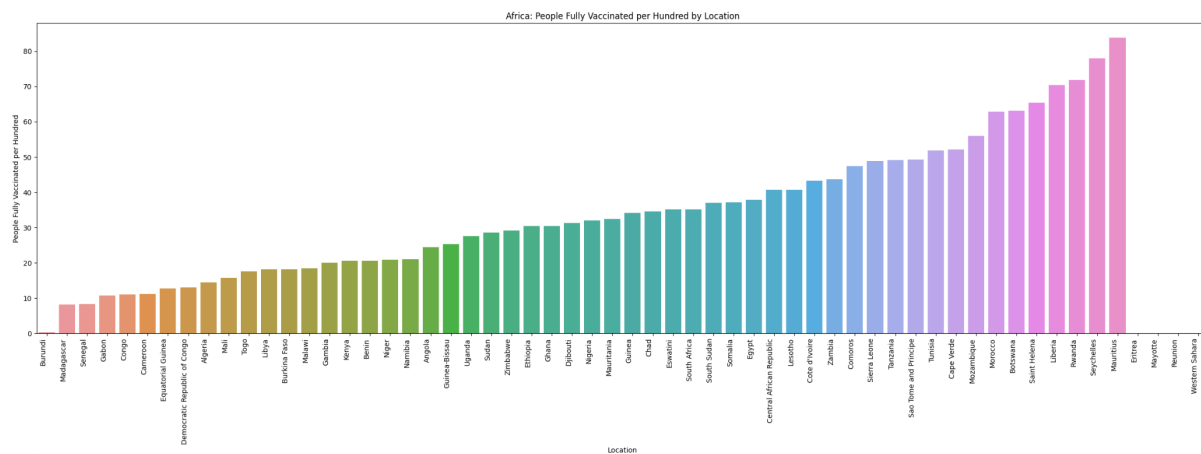Figure 6: People vaccinated in Asia per Hundred location

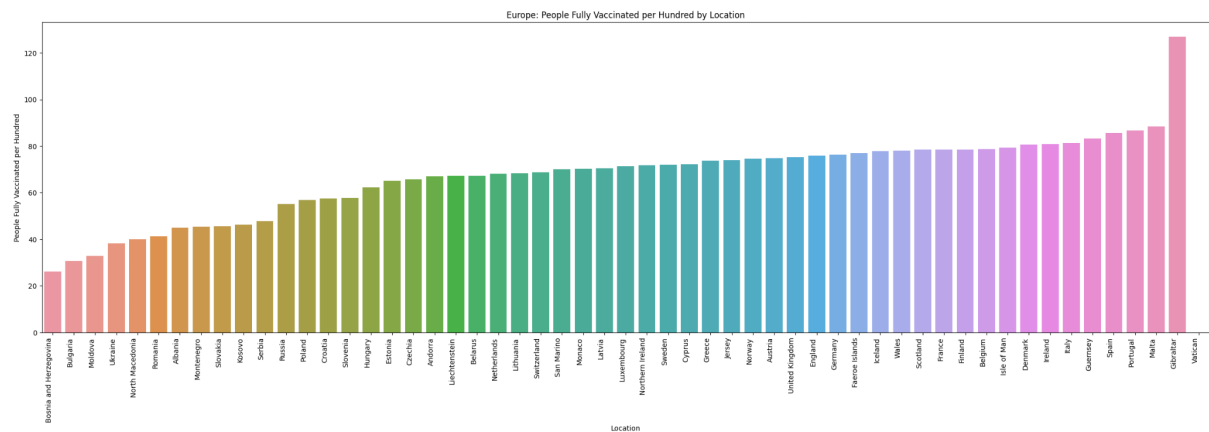Figure 7: People vaccinated in Africa per Hundred location

Figure 8: People vaccinated in Europ per Hundred location

Total Covid cases in Asia, Africa, and Europe

The bar graphs in Figures 9, 10, and 11 show that the number of COVID-19 cases varies significantly across countries located on the same continent. For instance, the total number of COVID-19 cases in European countries ranges from about 10,000 to more than 4 million. Similarly, the total number of COVID-19 cases in African countries ranges from a few hundred to about 7 million. The highest number of COVID-19 cases in Asian countries is about 17 million, while the lowest is about 300 cases. These graphs highlight the need for different strategies to manage and control the spread of COVID-19 in different countries based on their unique situations.
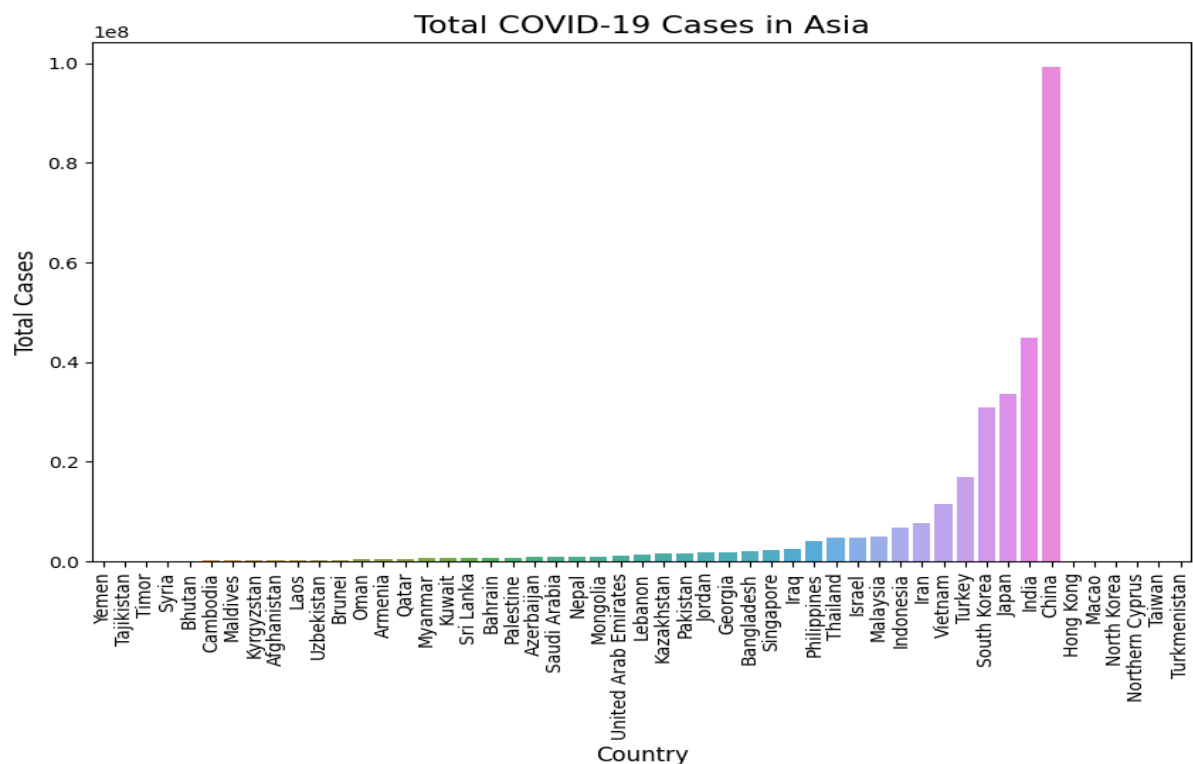

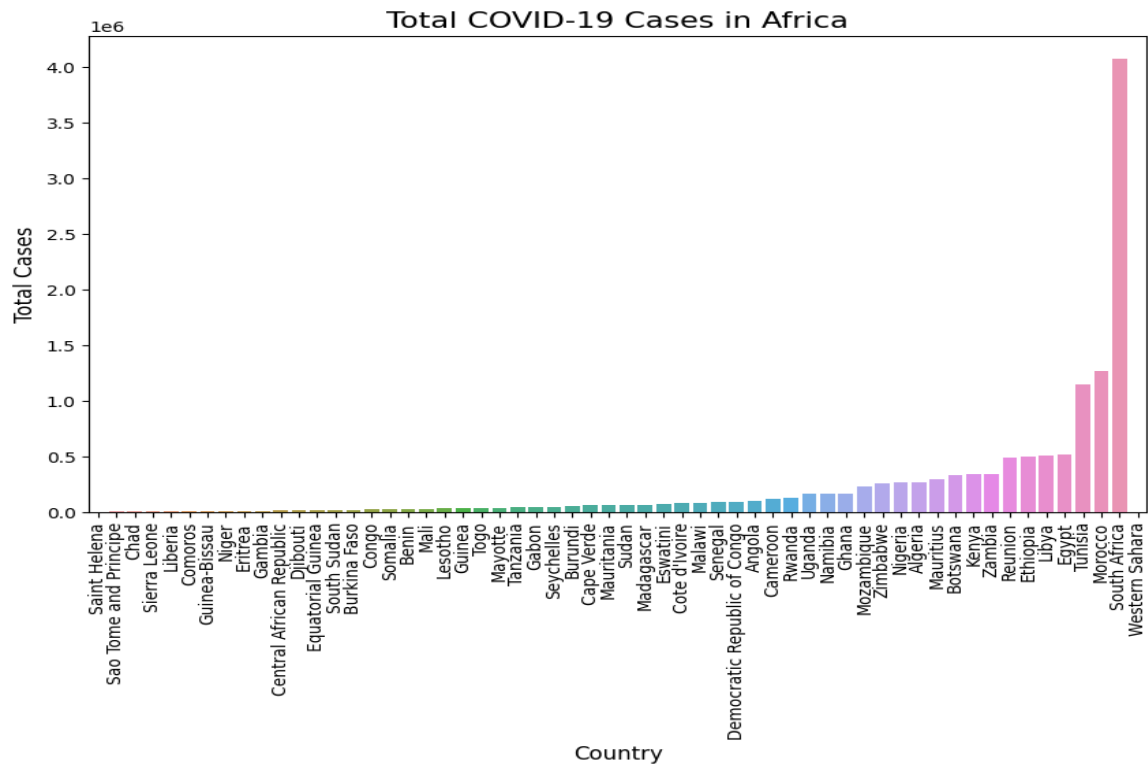Figure 9: Total Covid 19 Cases in ASIA
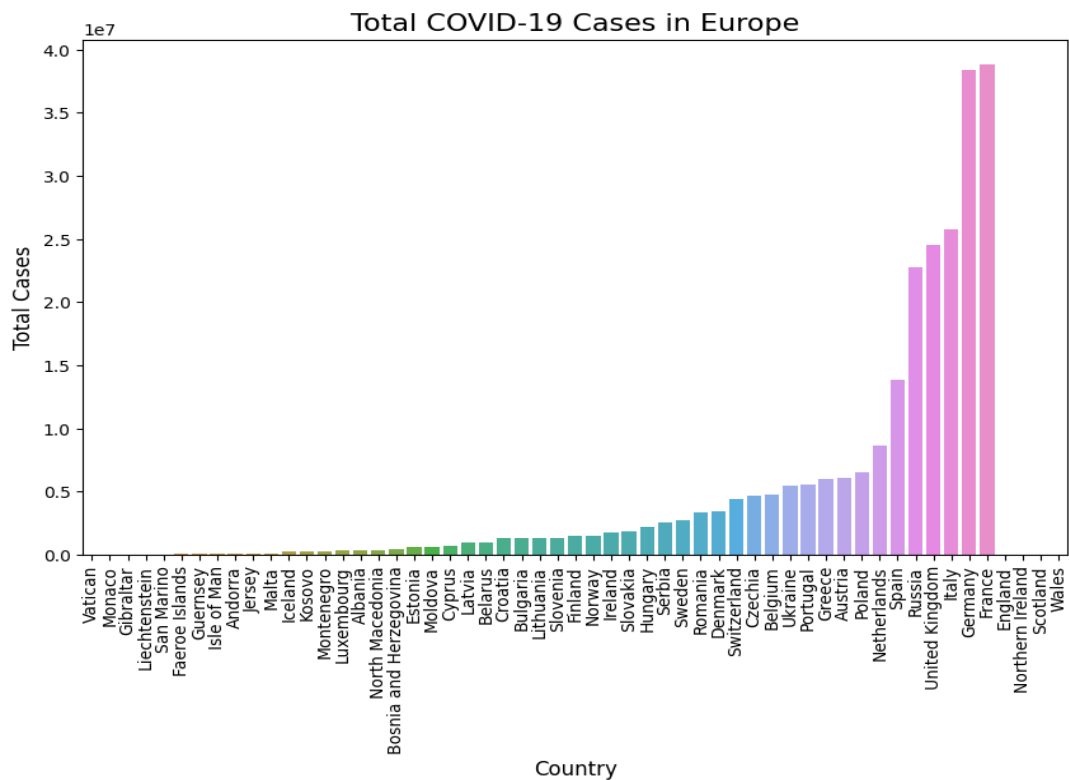
Figure 10: Total Covid 19 Cases in AFRICA

Figure 11: Total Covid 19 Cases in EUROPE

Total deaths and total new cases

The figure has four subplots, each displaying the top 10 countries in terms of new cases, new deaths, total cases, and total deaths. The bar plots provide a quick and easy way to compare the top 10 countries across the four metrics. As of the data cutoff date in September 2021, the United States had the highest number of total cases and deaths, while India had the highest number of new cases and deaths. Brazil and Russia also ranked high in all four metrics. The bar plots also highlight the large disparities in COVID-19 outcomes between countries, with some countries like New Zealand and Australia having a very low case and death numbers compared to others.
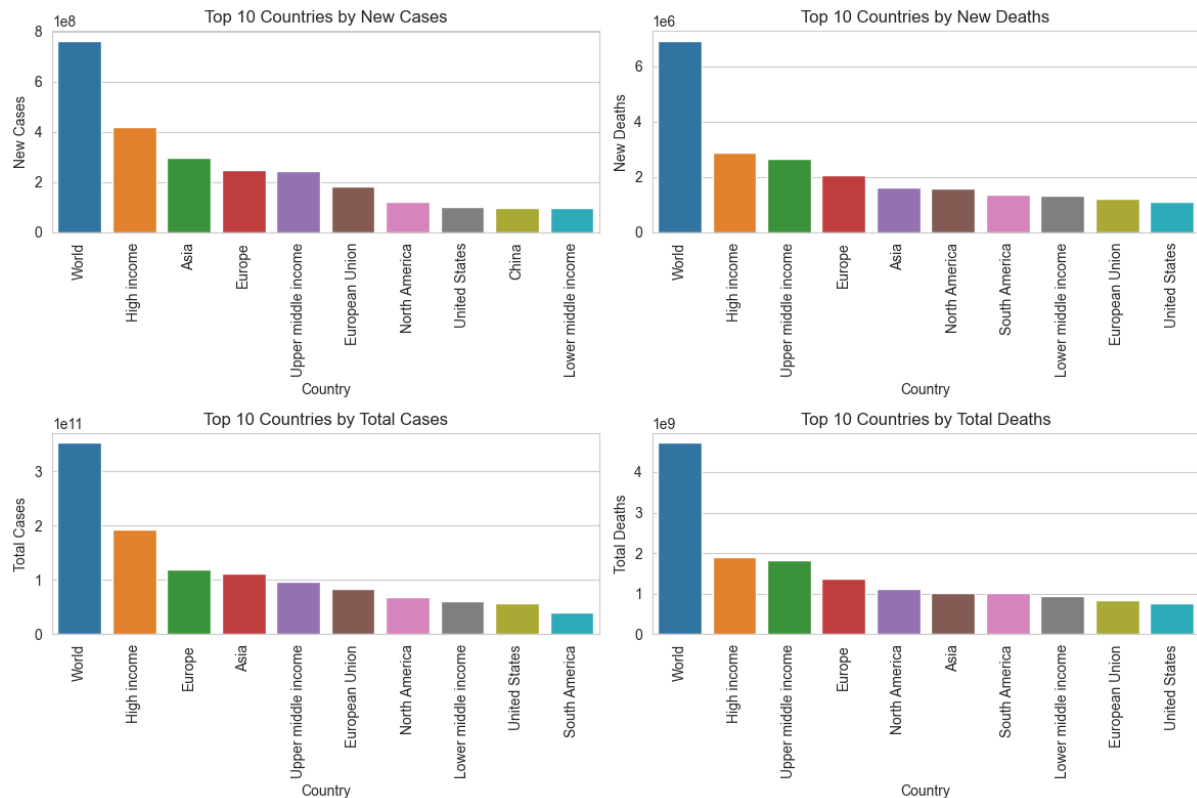

Figure 20: Total deaths and total new cases

Covid-19 Cases Per Million People

This visualization uses Plotly to create a scatter plot of COVID-19 cases per million people. From the plot, we can see that there is a positive correlation between total cases per million people and new cases per million people, as expected. We can also see that the highest number of cases per million people is in Europe and North America, while the lowest number of cases per million people is in Africa and Asia.
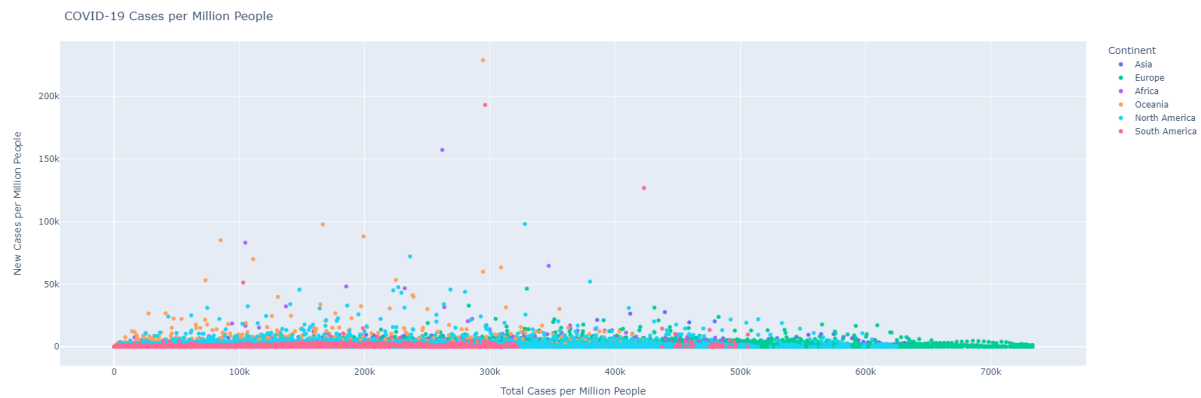

Figure 21: Covid-19 Cases Per Million People

Conclusion

The COVID-19 pandemic has had a significant global impact, with a sharp increase in new cases in early 2020. Although cases have started to decline, continued efforts are needed to control the virus. Future analysis could investigate factors contributing to its spread and develop predictive models.

The bar graph highlights the countries most severely impacted by COVID-19 in terms of new deaths. This information can be used to target interventions and resources.

The global impact of the pandemic is shown in the graph, with certain continents most affected in terms of total cases. This can inform policies and interventions aimed at controlling the virus.

The line graph shows the pandemic continues to pose a significant threat globally, with a spike in cases towards the end of 2021. The analysis could investigate contributing factors.

The United States has successfully vaccinated a large percentage of its population, with a continuously increasing trend. Other countries' vaccination trends could be analyzed to gain a more comprehensive understanding of global progress.

Some countries have higher vaccination rates than others, regardless of continent. Continued monitoring of vaccination rates globally is important to ensure all countries have access to vaccines.

The bar plots compare the top 10 countries across four metrics, highlighting large disparities in COVID-19 outcomes between countries. Some countries have very low case and death numbers compared to others.