

Covid Cases Analysis by Zain Gadagkar

The objective of this project is to analyze and visualize the COVID-19 data to gain insights into the trends, patterns, and key metrics related to the pandemic. The dataset contains various columns such as total cases, total deaths, total vaccinations, and other relevant metrics for different locations over time.

Through data visualization, we aim to:

- Track the daily or weekly count of new COVID-19 cases to monitor the rate of infection.
- Analyze the reproduction rate (R-value) quarterly over time to understand the rate of transmission of the virus.
- Visualize testing metrics such as total tests conducted, positivity rate, and testing capacity quarterly to assess the adequacy of testing efforts.
- Show the progress of COVID-19 vaccination efforts, including total vaccinations administered, the percentage of the population vaccinated, and vaccination rates over time quarterly.

By visualizing these key metrics, we can provide valuable insights into the spread of the virus, the effectiveness of testing and vaccination efforts, and the overall impact of COVID-19

Data Preprocessing:

To handle missing values, I followed these steps:

- Loaded the data into a DataFrame and removed unwanted columns not relevant to the analysis.
- Explored the dataset to identify columns with missing values using methods like `info()` and `isnull().sum()`.
- Carefully observed the nature of missing values and the impact they might have on the analysis.
- Filled missing values with appropriate substitutes, such as filling zeros (0) in place of null values, where applicable.
- For columns with missing dates, if necessary, formatted the dates to ensure consistency and usability in the analysis.

By handling missing values appropriately and calculating daily returns accurately, I ensured the integrity and reliability of the data analysis and visualization process.

Find total countries where total_deaths is greater than 1000000.

The locations in which total deaths have crossed 1M mark are 10:

Asia

Europe

European Union

High income

Lower middle income

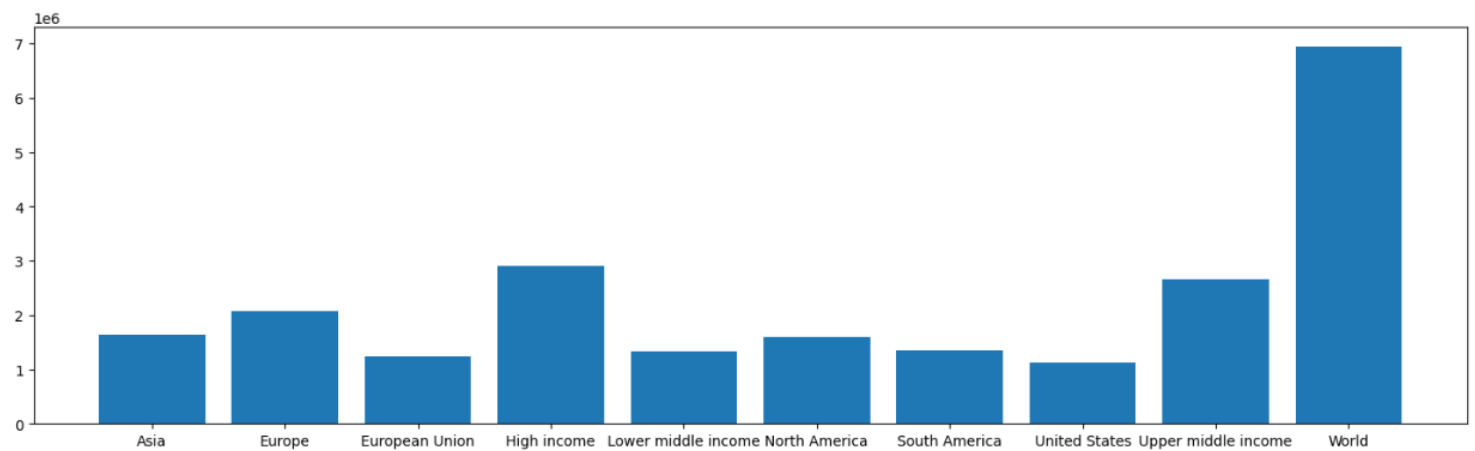
North America

South America

United States

Upper middle income

World



I have performed the below analysis and EDA:

- **Total Cases:** Visualize the total number of COVID-19 cases over time to understand the overall trend of the pandemic.
- **New Cases:** Show the daily or weekly count of new COVID-19 cases to monitor the rate of infection.
- **Total Deaths:** Display the total number of COVID-19 deaths over time to assess the impact of the pandemic on mortality.
- **New Deaths:** Visualize the daily or weekly count of new COVID-19 deaths to track changes in mortality rates.
- **Reproduction Rate:** Show the reproduction rate (R-value) over time to understand the rate of transmission of the virus.
- **Hospitalizations:** Display the number of COVID-19 patients hospitalized over time to assess the burden on healthcare systems.
- **Testing Metrics:** Visualize testing metrics such as total tests conducted, positivity rate, and testing capacity to assess the adequacy of testing efforts.
- **Vaccination Metrics:** Show the progress of COVID-19 vaccination efforts, including total vaccinations administered, percentage of population vaccinated, and vaccination rates over time.

Based on the COVID-19 data analysis, several key findings emerge:

- **Peak Period:** The analysis reveals that the COVID-19 cases peaked between 2020 and 2022, with the highest number of cases recorded on July 18, 2023, reaching 768,237,024. Similarly, the highest number of deaths occurred on the same day, totaling 6,951,664.
- **Reproductive Rate:** The reproductive rate (R-value) was observed to be at its peak during the third quarter of 2021, indicating a high rate of transmission of the virus during that period.
- **Control Measures:** There is evidence to suggest that control measures, such as increased testing, were effective in managing the pandemic. The positive rate of COVID-19 cases peaked during the first quarter of 2022, suggesting a high prevalence of the virus. However, with the rollout of vaccinations starting from late 2020 and reaching maximum rates during early 2022, there was a subsequent decline in the spread of the virus.
- **Impact of Vaccination:** The data indicates that the vaccination efforts played a crucial role in controlling the pandemic, leading to a decrease in the transmission rate and ultimately contributing to the containment of COVID-19.

These findings underscore the importance of implementing control measures, such as increased testing and vaccination campaigns, in mitigating the spread of COVID-19 and reducing its impact on public health. Additionally, the analysis highlights the effectiveness of data-driven approaches in understanding and addressing public health challenges.

Please go through the below charts for more insights:

