COVID 19 ANALYSIS

**Dataframe = day\_wise**

**Date**: The specific date on which the data was recorded.

**Confirmed**: The total number of confirmed COVID-19 cases up to that date.

**Deaths**: The total number of deaths caused by COVID-19 up to that date.

**Recovered**: The total number of recovered COVID-19 patients up to that date.

**Active**: The total number of active COVID-19 cases on that date (calculated as Confirmed - Deaths - Recovered).

**New cases**: The number of new confirmed COVID-19 cases reported on that date.

**New deaths**: The number of new deaths caused by COVID-19 reported on that date.

**New recovered**: The number of new recoveries from COVID-19 reported on that date.

**Deaths / 100 Cases**: The ratio of deaths per 100 confirmed cases.

**Recovered / 100 Cases**: The ratio of recoveries per 100 confirmed cases.

**Deaths / 100 Recovered**: The ratio of deaths per 100 recoveries.

**No. of countries**: The number of countries reporting COVID-19 cases on that date. ( Basically how many countries was affected by that time)

Analysis

**Confirmed Cases**

Initially, the number of confirmed COVID-19 cases was very low. However, over time, the number of confirmed cases increased significantly, showing a clear upward trend as the pandemic spread across more dates.

**New Cases**

The number of new cases fluctuated over time. Sometimes it remained constant, while at other times it spiked or dropped. This inconsistency in new cases contributed to the overall increase in confirmed cases since many new cases were not immediately recovering.

**Active Cases**

The number of active cases represents the currently infected individuals at a given time. This number increased linearly over time, reflecting the fact that new cases were not recovering quickly enough to offset the number of new infections.

**New Deaths**

Initially, the number of new deaths was very high, with a peak around mid-April 2020 when daily deaths reached 864. After this peak, the number of new deaths fluctuated, with periods of increase and decrease, likely influenced by the availability and effectiveness of treatments in different regions.

**Total Deaths**

The total number of deaths was very low around March 2020 but increased steadily over time. By June 2020, the total number of deaths had reached a significant peak, reflecting the cumulative impact of the pandemic.

**Recovered Cases**

The number of recovered cases showed a positive trend, increasing over time. Initially, recoveries were very few, but as treatments improved and more people received care, the number of recoveries grew exponentially.

**Deaths per 100 Cases**

This metric shows the proportion of deaths among every 100 confirmed cases. Initially, the death rate was high, peaking around May 2020. Over time, as the number of new cases stabilized and recoveries increased, the death rate per 100 cases decreased.

**Recovered per 100 Cases**

This metric represents the proportion of recoveries among every 100 confirmed cases. Initially low, it peaked in March 2020 when about 55 out of every 100 cases were recovering. After a brief decline, the recovery rate improved again post-April, stabilizing around 47 recoveries per 100 cases by June.

**Deaths per 100 Recovered**

This analysis shows the proportion of deaths among every 100 recoveries. In the early stages (February 2020), this rate was very high, with 134 deaths per 100 recoveries, indicating severe outcomes for many patients. Over time, this ratio decreased significantly, reflecting better treatment outcomes and increased survival rates.

**New Recovered**

The number of new recoveries started low but increased steadily over time, peaking around June 2020. This trend shows that more people were recovering as treatments became more effective and widely available.

**Number of Countries Affected**

This column shows the number of countries reporting COVID-19 cases. Initially, less than 15 countries were affected. By mid-2020, the virus had spread globally, with 187 countries reporting cases, indicating the widespread reach of the pandemic.

**Country\_wise**

1. **Country/Region**: The specific country or region where the data was recorded.
2. **Confirmed**: The total number of confirmed COVID-19 cases in the country/region.
3. **Deaths**: The total number of deaths caused by COVID-19 in the country/region.
4. **Recovered**: The total number of recovered COVID-19 patients in the country/region.
5. **Active**: The total number of active COVID-19 cases in the country/region (calculated as Confirmed - Deaths - Recovered).
6. **New cases**: The number of new confirmed COVID-19 cases reported in the country/region on the latest date.
7. **New deaths**: The number of new deaths caused by COVID-19 reported in the country/region on the latest date.
8. **New recovered**: The number of new recoveries from COVID-19 reported in the country/region on the latest date.
9. **Deaths / 100 Cases**: The ratio of deaths per 100 confirmed cases in the country/region, expressed as a percentage.
10. **Recovered / 100 Cases**: The ratio of recoveries per 100 confirmed cases in the country/region, expressed as a percentage.
11. **Deaths / 100 Recovered**: The ratio of deaths per 100 recoveries in the country/region, expressed as a percentage.
12. **Confirmed last week**: The total number of confirmed COVID-19 cases in the country/region reported in the previous week.
13. **1 week change**: The change in the number of confirmed cases in the country/region over the past week.
14. **1 week % increase**: The percentage increase in the number of confirmed cases in the country/region over the past week.
15. **WHO Region**: The World Health Organization region to which the country/region belongs.

Analysis

**Confirmed Cases:**

* **Overview**: This represents the total number of confirmed COVID-19 cases in each country and region.
* **Top Countries**:
  + The United States has the highest number of confirmed cases.
  + Brazil ranks second, followed by Russia and India.
  + Pakistan is at the lowest among the top 15 countries.
* **Regional Insights**:
  + The Americas region (primarily the US) leads in confirmed cases.
  + Europe, represented by Russia, is also significantly impacted.
  + India represents the South-East Asia region, ranking fourth.
* **Implication**: The high number of confirmed cases in the US and Brazil indicates a severe spread of the virus, necessitating stringent measures to control the outbreak.

**Active Cases:**

* **Overview**: This is calculated as the number of confirmed cases minus the number of deaths and recoveries, representing the currently active COVID-19 cases.
* **Top Countries**:
  + The United States again tops the list, followed by Brazil, the UK, Russia, and India.
  + South Africa is at the lowest among the top 15 countries.
* **Regional Insights**:
  + The Americas region (primarily the US) has the highest number of active cases.
  + The Africa region has the lowest, with South Africa representing it.
* **Implication**: High active cases in the US and Brazil highlight ongoing transmission, requiring continued public health interventions.

**New Cases:**

* **Overview**: This represents the number of new confirmed cases reported on the latest date.
* **Top Countries**:
  + The United States reports the highest number of new cases.
  + Brazil ranks second, with significant new cases also reported in India.
  + Sweden is at the lowest among the top 15 countries.
* **Regional Insights**:
  + The Americas region, particularly the US and Brazil, continues to report high numbers of new cases.
  + South-East Asia, represented by India, is also experiencing a notable increase.
  + Europe (Sweden) shows lower new cases compared to the top countries.
* **Implication**: The high number of new cases in the US and Brazil suggests ongoing challenges in controlling the spread, while India is also seeing a significant rise.

**Recoveries:**

* **Overview**: This represents the total number of recovered COVID-19 patients.
* **Top Countries**:
  + The United States and Brazil lead in the number of recoveries.
  + Russia ranks third, followed by India.
  + France is at the lowest among the top 15 countries.
* **Regional Insights**:
  + The Americas region (US and Brazil) shows the highest recovery numbers.
  + Europe, represented by Russia and France, also has significant recoveries.
  + India in South-East Asia shows substantial recoveries as well.
* **Implication**: High recovery numbers in the US and Brazil indicate effective recovery efforts despite the high number of cases.

**Deaths:**

* **Overview**: This represents the total number of deaths caused by COVID-19.
* **Top Countries**:
  + The United States and Brazil have the highest death tolls.
  + The UK, Mexico, and France follow.
  + India has a relatively lower death count compared to these countries.
* **Regional Insights**:
  + The Americas region (US and Brazil) again shows the highest death numbers.
  + Europe (UK, France, Spain) has also been significantly impacted.
  + India's lower death count is notable compared to its high number of cases.
* **Implication**: The high death toll in the US and Brazil underscores the severe impact of the pandemic, while India's lower death count relative to its cases suggests differences in healthcare responses or reporting.

**Weekly Changes:**

* **Confirmed Cases**:
  + Brazil and India show notable weekly increases.
  + Argentina is at the lowest among the top 15 countries.
* **Regional Insights**:
  + The Americas and South-East Asia regions, particularly Brazil and India, are seeing substantial weekly increases in confirmed cases.
  + South Africa's position in weekly changes indicates an improving situation compared to earlier data.
* **Implication**: The weekly changes highlight dynamic shifts in the pandemic's impact, with Brazil and India experiencing significant increases and South Africa showing improvements.

**General Observations:**

* **Data Trends**: The data indicates a dynamic and evolving situation, with the US and Brazil frequently topping various metrics, reflecting the pandemic's severe impact in these countries.
* **Regional Variations**: Significant differences are observed between regions, with the Americas and Europe showing high numbers in most metrics, while South-East Asia (India) and Africa (South Africa) show varying impacts.
* **Implications for Policy**: These insights underscore the need for targeted public health interventions, especially in regions with high active cases and new infections, to control the spread and mitigate the impact of COVID-19.

**Worldometer\_data**

1. Country/Region: The specific country or region where the data was recorded.
2. Continent: The continent where the country/region is located.
3. Population: The total population of the country/region.
4. TotalCases: The total number of confirmed COVID-19 cases in the country/region.
5. NewCases: The number of new confirmed COVID-19 cases reported in the country/region on the latest date.
6. TotalDeaths: The total number of deaths caused by COVID-19 in the country/region.
7. NewDeaths: The number of new deaths caused by COVID-19 reported in the country/region on the latest date.
8. TotalRecovered: The total number of recovered COVID-19 patients in the country/region.
9. NewRecovered: The number of new recoveries from COVID-19 reported in the country/region on the latest date.
10. ActiveCases: The total number of active COVID-19 cases in the country/region (calculated as TotalCases - TotalDeaths - TotalRecovered).
11. Serious,Critical: The number of COVID-19 cases in the country/region classified as serious or critical.
12. Tot Cases/1M pop: The total number of COVID-19 cases per one million people in the country/region.
13. Deaths/1M pop: The total number of deaths caused by COVID-19 per one million people in the country/region.
14. TotalTests: The total number of COVID-19 tests conducted in the country/region.
15. Tests/1M pop: The total number of COVID-19 tests conducted per one million people in the country/region.
16. WHO Region: The World Health Organization region to which the country/region belongs.

**Insights from Worldometer Data**

**Total Tests:**

* **Overview**: This metric represents the total number of COVID-19 tests conducted by each country.
* **Top Performers**:
  + The United States leads with the highest number of total tests conducted.
  + Russia ranks second, followed by the United Kingdom and India.
* **Regional Insights**:
  + The Americas region, particularly the US, dominates in testing numbers.
  + Europe follows closely with substantial testing in countries like the UK and Russia.
  + India represents the Eastern region, ranking fourth overall.
  + Iran, among the top 15, has the lowest total tests.
* **Implication**: High testing numbers in the US, Russia, and the UK indicate robust testing infrastructure and efforts to identify and control the spread of the virus.

**Deaths per Million Population (Deaths/1M pop):**

* **Overview**: This metric indicates the number of COVID-19 deaths per million people in each country.
* **Top Countries**:
  + San Marino, Belgium, Andorra, the UK, and Spain are among the highest, all located in Europe.
* **Regional Insights**:
  + Europe has the highest death rates per million population, reflecting significant impact in this region.
  + The Americas, specifically the US, also show high death rates per million.
  + India is notably absent from the top 15 list, indicating a relatively lower death rate compared to these countries.
* **Implication**: Europe's high death rates per million population highlight the severe impact of COVID-19 in this region. The absence of India suggests it had a lower relative death rate during the same period.

**Tests per Million Population (Tests/1M pop):**

* **Overview**: This metric shows the number of COVID-19 tests conducted per million people in each country.
* **Top Countries**:
  + Monaco leads in tests per million population, followed by other European countries.
  + Mauritius appears at the bottom of the top 15 list.
* **Regional Insights**:
  + European countries dominate the top spots, indicating extensive testing efforts.
  + India is again absent from the top 15, suggesting a lower testing rate per million population compared to these countries.
  + Notably, China does not appear in this list, indicating a lack of transparent data reporting from China beyond the initial outbreak period.
* **Implication**: The high testing rates per million population in European countries reflect their rigorous testing strategies. The absence of China and India from the top 15 suggests differences in testing transparency and capacity.

**General Observations:**

* **Data Transparency**: China's absence in the tests per million population metric points to issues with data transparency and sharing, as initial data from China was not consistently followed up.
* **Regional Variations**: The data highlights significant regional variations, with Europe showing high death and testing rates, the Americas leading in total tests, and India having relatively lower death and testing rates per million population.

These insights provide a comprehensive view of the global COVID-19 response, highlighting the differences in testing, mortality, and regional impacts.

4o

Full\_grouped

About Data

1. **Date**: The date on which the data was recorded.
2. **Country/Region**: The specific country or region where the data was recorded.
3. **Confirmed**: The total number of confirmed COVID-19 cases in the country/region up to that date.
4. **Deaths**: The total number of deaths caused by COVID-19 in the country/region up to that date.
5. **Recovered**: The total number of recovered COVID-19 patients in the country/region up to that date.
6. **Active**: The total number of active COVID-19 cases in the country/region on that date (calculated as Confirmed - Deaths - Recovered).
7. **New cases**: The number of new confirmed COVID-19 cases reported in the country/region on that date.
8. **New deaths**: The number of new deaths caused by COVID-19 reported in the country/region on that date.
9. **New recovered**: The number of new recoveries from COVID-19 reported in the country/region on that date.
10. **WHO Region**: The World Health Organization region to which the country/region belongs.

Analysis

**Confirmed Cases:**

* **Trend Overview**: The confirmed cases column represents the total number of COVID-19 cases reported in each country/region up to the specific date.
* **Initial Outbreak in China**: Initially, China had the highest number of confirmed cases.
* **Rising Numbers in US and UK**: Over time, the US and UK saw significant increases in their confirmed case numbers.
* **Global Total**: From the graph, we can see that the total number of confirmed cases across all countries surpassed 8 million.
* **Visualization Insights**: The graph highlights the rapid spread of the virus, with the US and UK emerging as major hotspots after the initial outbreak in China.

**Deaths:**

* **Definition**: This column represents the total number of deaths caused by COVID-19 in each country/region up to that date.
* **Initial Spike**: The visual shows a steep increase in deaths initially.
* **Linear Increase**: After the initial spike, the number of deaths continued to increase but at a slower, more linear rate.
* **Comparison with Confirmed Cases**: While confirmed cases surged rapidly, the death rate, although significant, showed a less steep increase after the initial spike.
* **Trend Analysis**: This suggests that while the virus spread quickly, efforts to manage and treat the disease may have had some impact on the death rate over time.

**Active Cases:**

* **Definition**: Active cases represent the number of currently infected individuals.
* **US Spike**: The US showed a significant spike in active cases, indicating a large number of ongoing infections at certain points.
* **Trend Insights**: The graph helps visualize the burden on healthcare systems, particularly in the US, which had to manage a high number of active cases.

**New Cases:**

* **Definition**: New cases indicate the number of new COVID-19 infections reported on each specific date.
* **Singapore's Low Numbers**: Singapore consistently reported lower numbers of new cases.
* **US and Brazil**: Both countries saw substantial increases in new cases, with Brazil at one point surpassing the US.
* **France's Sudden Spike**: France experienced a sudden spike in new cases around mid-April, doubling its numbers in a short period.
* **Chile's Increase**: Chile also showed a sudden spike, reflecting the rapid spread of the virus in specific regions.
* **Visualization Highlights**: These spikes and trends illustrate how different countries experienced surges at different times, highlighting the uneven spread of the pandemic.

**Observations:**

* **China**: Initially the epicenter, but managed to control the spread over time.
* **US and UK**: Became major hotspots with rapidly increasing confirmed and active cases.
* **France and Brazil**: Experienced significant spikes in new cases, indicating sudden outbreaks.
* **Overall Trend**: The global data shows a clear progression of the pandemic, with various countries facing different challenges at different times.

By analyzing the full\_grouped dataset, we can observe the dynamic nature of the COVID-19 pandemic across different regions and time periods. The visualizations provide a clear picture of how the virus spread, the impact on various countries, and the effectiveness of measures taken to control the outbreak.

**Analysis on Combining Full\_Grouped and Day\_Wise Data**

By combining the full\_grouped and day\_wise datasets, we aimed to analyze the percentage of confirmed cases and deaths by country over time. This merging was done on the date column to track the progression of the pandemic in each country.

**Percentage of Confirmed Cases:**

* **Early Trends**: Initially, a significant number of cases were reported from Malaysia, but China quickly became the country with the highest percentage of confirmed cases.
* **China's Decline**: Over time, the percentage of confirmed cases in China decreased as other countries started reporting more cases.
* **US Surge**: The United States saw a substantial increase in the percentage of confirmed cases, eventually surpassing all other countries.
* **Italy's Brief Rise**: Italy experienced a spike in confirmed cases for a short period before the numbers began to decline.

**Percentage of Deaths:**

* **Initial Dominance by China**: At the beginning of the pandemic, China had the highest percentage of deaths.
* **US Deaths**: Although the US had a high number of confirmed cases, the percentage of deaths was relatively lower.
* **UK's High Death Rate**: The United Kingdom showed a higher percentage of deaths compared to its percentage of confirmed cases.
* **Italy and France**: Both Italy and France had periods where their death percentages were notably high.
* **China's Reduction**: Over time, China's death percentage significantly reduced as the situation improved.
* **Brazil's Consistency**: Brazil's death and confirmed case percentages remained relatively proportional throughout the period analyzed.

**Observations:**

* **China**: Initially had the most significant impact but managed to control the spread and reduce the death rate over time.
* **US**: Became the epicenter of the pandemic with the highest percentage of confirmed cases by the end of the analyzed period.
* **UK, Italy, and France**: Showed high death rates at certain points, indicating severe outbreaks.
* **Brazil**: Maintained a consistent ratio of deaths to confirmed cases, reflecting the ongoing struggle to manage the pandemic effectively.

By merging these datasets and analyzing the combined data, we gained valuable insights into the progression and impact of COVID-19 across different countries and regions. The visualizations helped highlight key trends and shifts in the pandemic's trajectory, allowing for a better understanding of how various nations were affected over time.