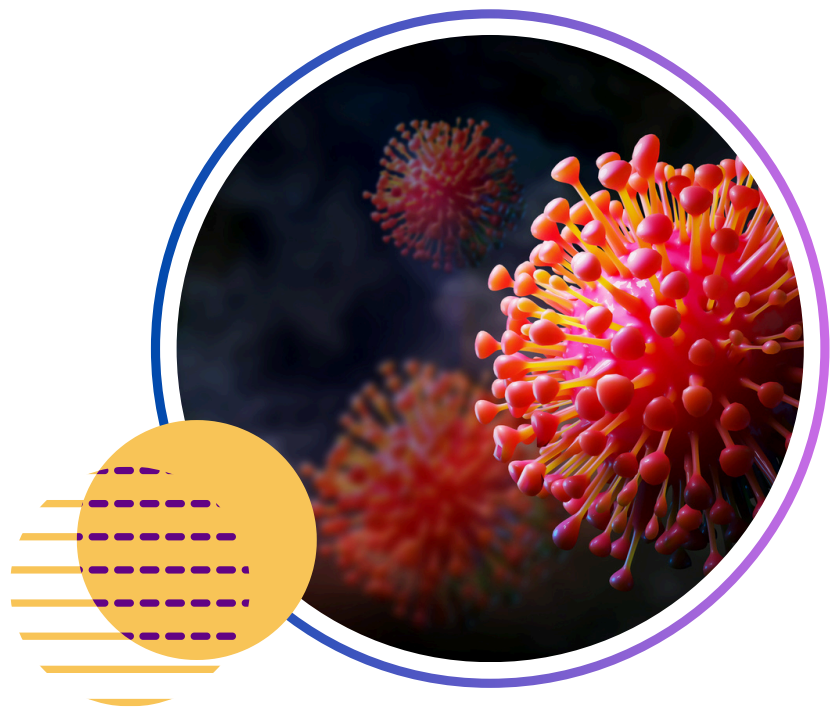


22 June, 2025

Project VII

COVID INTERNATIONAL TRENDS ANALYSIS

Prepared by :
Ananya Jha



Tools Used: Microsoft Excel & Power BI
Dataset Source: WHO COVID-19 Global
Data from kaggle.com



OVERVIEW

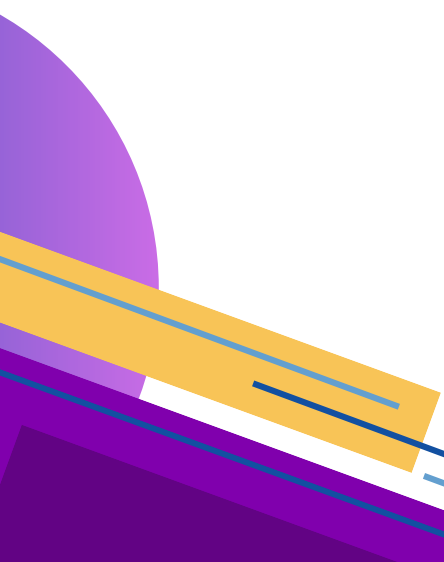
This project explores global trends in COVID-19 cases and deaths using a dataset sourced from Kaggle, derived from the official WHO COVID-19 Global Data. Covering the timeframe from 2020 to 2024, the dataset includes over 323,000 rows across 237 countries, providing a rich foundation for international analysis. To capture longer-term trends and reduce daily volatility, a new column was added to calculate the 7-day moving average of new cases, labeled as the “Moving Average.”

Data preparation and initial exploration were conducted in Excel, including the creation of a second sheet containing multiple pivot tables for comparative and regional insights. Final visualizations and interactive dashboards were developed in Microsoft Power BI, allowing for clear presentation of global and country-specific patterns in both cases and fatalities.

OBJECTIVE

“To identify major global patterns in COVID-19 cases, deaths, and regional differences using data visualization and trend analysis.”

TOOLS USED

1. Excel (Data Preparation)
 2. Power BI (Visualizations)
- 

STAKEHOLDER

1. Public Health Policy Makers
2. International Development Agencies (eg. WHO, UNICEF)
3. Media Researchers
4. Business Leaders Tracking Impact by Countries

DATA PREPARATION

- Cleaned and stored data by country and date
- Created a “Moving Average” column in excel to calculate the 7-Day Moving Average
- Uploaded to Power BI for dashboard creation

KEY VISUALIZATIONS AND INSIGHTS

Total Confirmed Cases

- Tool: Power BI
- Type: Card



Insight:

- Shows the highest total number of reported cases globally.
- Values reached over 100 million.
- Confirms the global scale of pandemic.

Total Confirmed Deaths

- Tool: Power BI
- Type: Card



Insight:

- Shows the highest total number of reported deaths globally.
- Values reached 1 million.
- Confirms the global scale of pandemic.

Highest Daily New Cases

- Tool: Power BI
- Type: Card

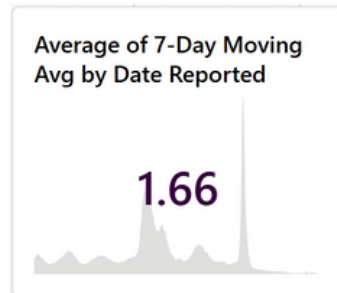


Insight:

- Shows the highest daily new cases reported globally.
- Values reached 7 million.
- Confirms the global scale of pandemic.

7-Day Moving Average

- Tool: Power BI
- Type: KPI

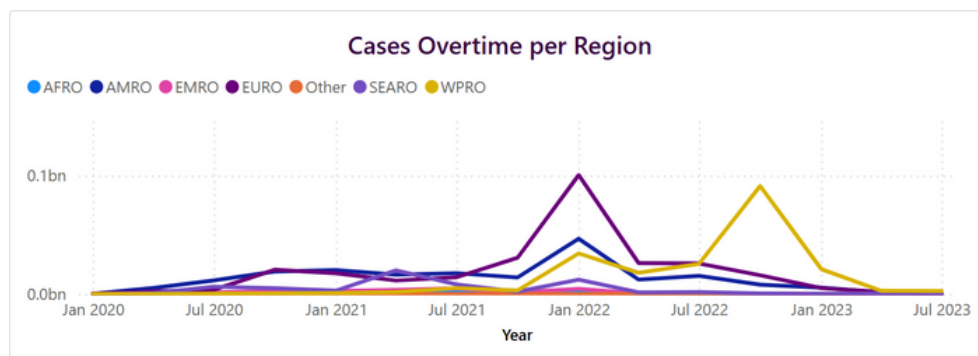


Insight:

- Shows the 7-Day Moving Average of new cases reported globally.
- Average value is higher than 1.
- Average peaked in the final years of the pandemic.

Highest Daily New Cases

- Tool: Power BI
- Type: Line Chart



Insight:

- Shows the Cases Overtime per Region.
- Peak values reached almost 100 million.
- EURO and WPRO were the most affected WHO regions.

New-Cases by WHO Region

- Tool: Power BI
- Type: Visual Table

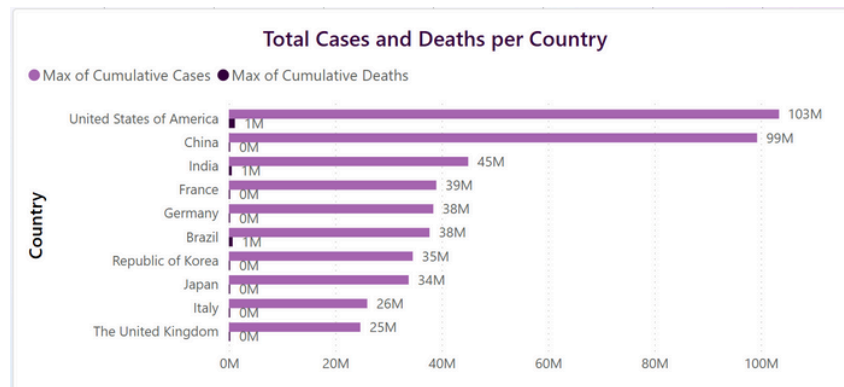
WHO Region	2020	2021	2022	2023	Total
AFRO	1900687	5405787	2144972	118428	9569874
AMRO	35799646	68174960	82286505	7025156	193286267
EMRO	4912291	12254762	6055407	171662	23394122
EURO	26661208	73758771	168522315	7192341	276134635
Other	745	19	0	0	764
SEARO	11973259	33001946	15762226	467606	61205037
WPRO	1087695	10515085	169254955	26426999	207284734
Total	82335531	203111330	444026380	41402192	770875433

Insight:

- Shows the total of new cases across WHO regions reported globally.
- Net total exceeds 75 million.
- EURO, WPRO and AMRO amongst most affected regions.

Cases and Deaths per Country

- Tool: Power BI
- Type: Bar Chart



Insight:

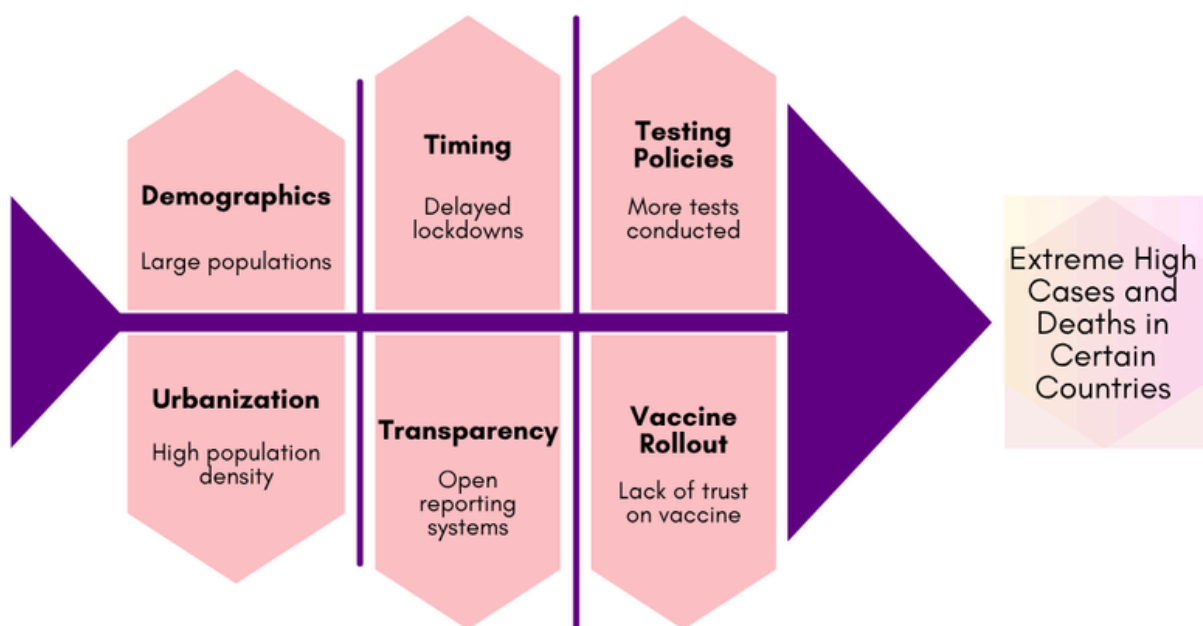
- Shows the Top 10 countries with most Cases and Deaths.
- Peak values reached almost 100 million.
- U.S.A. and China were the most affected countries.

ROOT CAUSE ANALYSIS

Insight 1: Extremely High Total Cases and Deaths in Certain Countries

- Tool: Canva

Countries like USA and China accounted for a disproportionate share of global COVID-19 cases and deaths.



ROOT CAUSE ANALYSIS

Insight 2: High Single-day Spikes in New Cases

- Tool: Canva

Some countries reported unusually high new COVID-19 cases on a single day, far above their usual daily average.

5 WHYS ROOT CAUSE CHAIN

1. Why were there extremely high case spikes on a single day?
 - Because a large number of cases were reported all at once.
2. Why were so many cases reported on that specific day?
 - Because there was a delay or backlog in testing data or result entry.
3. Why was there a delay or backlog in reporting test results?
 - Because the health system was overwhelmed or data was consolidated from multiple regions at once.
4. Why was the health system overwhelmed or data aggregated like that?
 - Because of a sudden surge in testing demand, holidays/weekends, or manual reporting systems in some areas.
5. Why did reporting rely on manual systems or experience surges?
 - Because many regions lacked real-time data infrastructure and had to batch-process results, especially in low-resource or high-population areas.

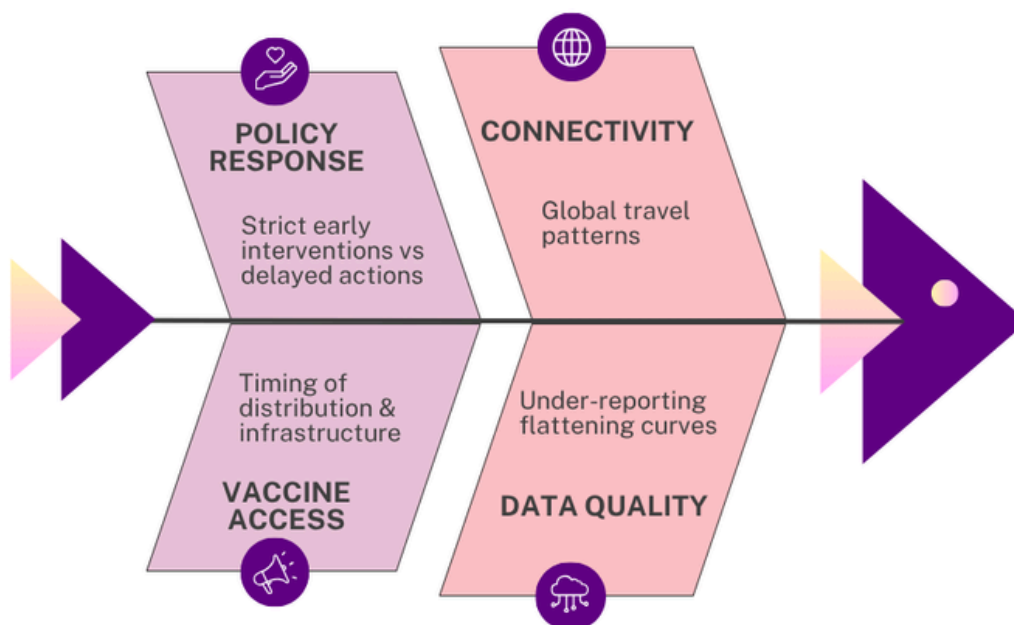
ROOT CAUSE ANALYSIS

Insight 2: Regional Variation in Trends

- Tool: Canva

The line chart revealed different outbreak patterns across WHO regions. For example, the EURO, WPRO and AMRO regions had much higher peaks compared to other regions.

REGIONAL VARIATION IN TRENDS






CONCLUSION

The COVID-19 pandemic has had an uneven and evolving impact across the globe. This project, using WHO data from 2020 to 2024, revealed that a handful of countries such as the United States, and China accounted for a disproportionately high share of global cases and deaths. Daily case numbers were volatile, often affected by reporting delays, which made moving averages more reliable for trend analysis.

The project also uncovered significant regional variation—some WHO regions experienced sharp surges followed by rapid declines, while others maintained steadier, prolonged curves. Through visual dashboards and pivot analysis, key patterns in transmission, data integrity, and policy effects were identified.

CLOSING STATEMENT

This project emphasizes the importance of looking beyond raw daily numbers and adopting structured, comparative analysis to draw meaningful insights from complex datasets. It highlights how data storytelling through Power BI can make global health patterns more accessible and actionable for diverse stakeholders.





SELF-REFLECTION

This was my most complex and data-heavy project so far, with over 300,000 rows and 237 countries. I learned that working with big data is not about doing more, but about doing things smarter. A major learning for me was calculating and analyzing the 7-day moving average. I understood how daily case numbers can be misleading due to delays or inconsistent reporting. Creating this trend metric taught me the importance of smoothing techniques and how they help reveal the real story hidden behind raw data.

On the visualization side, I became much more confident with Power BI. I used KPIs and cards, visual tables, trend chart and top-10 filters to make the dashboard more focused and useful. I also practiced layout design by using Snap to Grid and consistent formatting.

This project pushed me to think deeper — not just what is happening, but why. Using root cause analysis and the 5 Whys framework, I explored the reasons behind major spikes and regional patterns. This step made me feel like a true analyst: going beyond charts, into reasoning and interpretation.

Finally, I improved my communication and storytelling. I learned that visuals alone are not enough — insights must be clearly written and explained. Every chart became a sentence in a larger narrative. That mindset shift was powerful.

