COVID-19 Daily Cases Analysis in Pakistan & Asia

1. Introduction

This project focuses on analyzing the spread of COVID-19 in Pakistan and across several Asian countries using real-world datasets. The goal is to visualize daily confirmed cases, detect trends over time, and compare countries' pandemic responses. This type of analysis helps improve awareness and decision-making during pandemics.

2. Dataset Description

- Source: Johns Hopkins University CSSE COVID-19 Data
- Format: CSV
- Link: https://github.com/CSSEGISandData/COVID-19
- Data includes daily confirmed COVID-19 cases for all countries globally.
- Fields used: Date, Country/Region, Cumulative Confirmed Cases

3. Tools & Technologies Used

- Python (v3.x)
- Jupyter Notebook
- Pandas (for data manipulation)
- Matplotlib (for data visualization)
- FPDF (for PDF generation)

4. Data Processing Steps

- Loaded the global time-series dataset using Pandas.
- Filtered data for Pakistan and selected Asian countries.
- Converted dates to datetime format.
- Transposed data to make dates the index.
- Calculated daily new cases using `.diff()`.
- Smoothed results using 7-day rolling averages.

- Created comparison graphs for Asian countries using line plots.

- Generated visualizations to compare wave patterns among countries.

5. Data Visualizations

The following graphs were generated using Matplotlib:

1. Daily New Cases in Pakistan

   - A simple line graph showing daily confirmed cases in Pakistan.

2. 7-Day Rolling Average of Daily Cases

   - Smoothed line plot to reduce daily fluctuation and highlight trends.

3. Cumulative COVID-19 Cases in Pakistan

   - Total number of confirmed cases growing over time.

4. Daily New Cases Comparison across Asian Countries

   - A multi-line plot comparing Pakistan, India, Bangladesh, Nepal, Sri Lanka, China, Iran, Indonesia, Philippines, and Japan.
   - Helps visualize how different countries experienced peaks and waves.

5. Smoothed Trends with 7-Day Rolling Average (for Asian Countries)

   - A smoother version of the multi-country comparison using rolling average to highlight key patterns and remove noise.

6. Observations & Insights

- Pakistan experienced multiple COVID-19 waves between 2020 and 2023.

- India's peaks were among the highest in Asia.

- Rolling averages helped reduce noise and improve trend visibility.

- Countries showed different peak timings, indicating varied containment measures.

COVID-19 Daily Cases Analysis in Pakistan & Asia

- South Asian countries showed sharp increases during Delta variant spread in 2021.

- Japan and China had smoother, more controlled case growth curves.

- The multi-country graph revealed contrasting pandemic trajectories across Asia.

- The use of 7-day rolling averages enhanced readability in complex data trends.

7. Conclusion

This analysis shows the importance of real-time data monitoring and visualization. It helped identify trends, detect peaks, and compare pandemic spread across countries. Such analysis can support public health policy and readiness for future outbreaks. The comparison of Asian countries further enhances understanding of regional pandemic patterns and responses.

8. References

- Johns Hopkins CSSE COVID-19 Dataset: https://github.com/CSSEGISandData/COVID-19

- Pandas Documentation: https://pandas.pydata.org

- Matplotlib Docs: https://matplotlib.org/stable/index.html