COVID-19 Global Data Tracker — Analysis Report

Project Overview

This project analyzes the global progression of the COVID-19 pandemic using real-world data from **Our World in Data (OWID)**. The analysis focuses on trends in **cases, deaths, and vaccination rollout** across selected countries: **Kenya, India, and the United States**.

🔍 Key Findings

- **The United States** had the highest total number of COVID-19 cases and deaths among the three countries, with several waves visible across the timeline.
- **India** exhibited a massive case spike in early 2021, followed by a rapid and large-scale vaccination campaign that improved its public health metrics by late 2021.
- **Kenya** reported significantly lower case and death numbers, but the data may reflect underreporting or testing limitations. Its vaccination progress was slower but steady.
- **Vaccination Trends**:
    - India surpassed the other two countries in vaccination speed post-July 2021.
    - The U.S. had an early vaccine rollout but plateaued in full coverage.
    - Kenya saw consistent progress with limited spikes.

📈 Visual Highlights

- **Line charts** showing total cases and deaths over time reveal infection waves and national response patterns.
- **Vaccination coverage graphs** illustrate disparities in access and uptake.
- An **optional choropleth map** visualizes global vaccination rates, making it easy to identify leading and lagging regions.

⚠ Data Limitations

- Missing or incomplete data for some countries was filled with zero or interpolated values.
- OWID sources rely on national reporting; underreporting is possible.
- Death rate and vaccination percentages are based on total population, which doesn’t account for age-specific eligibility or testing capacity.

Future Extensions

- Add **per capita metrics** (cases/deaths per million).
- Apply **7-day rolling averages** to smooth out noise in daily case counts.
- Build an interactive **dashboard** using Streamlit or Plotly Dash.
- Include more countries or group regions (e.g., EU, Sub-Saharan Africa).

📚 Tools Used

- **pandas** — for data handling and cleaning
- **matplotlib / seaborn** — for core visualizations
- **plotly express** — for optional interactive map
- **Jupyter Notebook** — for report-style presentation