# COVID-19 Data Analysis Project

## Objective

The aim of this project is to analyze global COVID-19 data using Python and excel understand the trends of total cases and deaths across countries, and visualize the pandemic’s impact through data storytelling.

## Tools Used

- Python (Pandas, Matplotlib, Seaborn) – for data cleaning and visualization  
- Excel – for initial inspection and cleaning  
-  
- Google Colab – for coding and analysis

## Data Source

The dataset was collected from Our World in Data:  
https://ourworldindata.org/covid-deaths  
(File used: owid-covid-data.csv)

## Data Cleaning Steps

1. Removed rows with null values in location, date, and total\_cases  
2. Filtered out regional and aggregated entries (like “World”, “Asia”, etc.)  
3. Kept only relevant columns:  
 - location, date, total\_cases, new\_cases, total\_deaths, new\_deaths, population  
4. Converted date column to datetime format for analysis

## Analysis Performed

- Time series analysis of total cases over time  
- Comparative view of top 3 most affected countries  
- Analysis of top 10 countries with highest total deaths  
- Filtered and visualized data to find meaningful insights

- Interactive dashboard with slicers and visual KPIs for real-time filtering  
 - Included charts and tables highlighting top countries and case trends

## Key Insights

- The top 3 countries (based on total cases) were automatically selected based on the latest data: e.g., United States, India, Brazil  
- These countries saw significant growth during pandemic peaks  
- The highest total deaths were concentrated in a few large-population countries  
- Consistent decline in new cases observed post-vaccination periods

## Conclusion

This project demonstrates practical skills in data wrangling, visualization, and interpretation using real-world pandemic data. The use of Python and Power BI highlights the ability to automate analysis and create interactive dashboards.

## Attachments

- Python code and charts (PDF)  
- Cleaned dataset (covid\_cleaned.csv)