COVID-19 Global Monitoring Dashboard

1. Project Overview

This project aims to provide a real-time automated COVID-19 data tracking system. Using publicly available data from the disease.sh API, the dataset includes critical metrics like total and daily cases, deaths, recoveries, active cases, critical cases, and population by country.

The project is fully automated with a Python-based ETL pipeline that updates a SQLite database and exports the latest dataset into Excel. This Excel file is then connected to Power BI to produce interactive and insightful dashboards for global stakeholders.

2. Business Questions Addressed

- Which countries are the most and least affected by COVID-19?

- What is the global recovery rate vs. death rate?

- What is the distribution of active, recovered, and deceased cases?

- How are different countries managing daily new infections?

- How do today's cases compare across countries?

3. Data Analysis

Dataset Fields:

- country, cases, todayCases, deaths, todayDeaths
- recovered, todayRecovered, active, critical, population
- fetch_date (daily refresh)

Data Source:

- API Endpoint: https://disease.sh/v3/covid-19/countries
- Refreshed daily via a Python script and stored in a local SQLite DB + Excel

Transformations:

- Dropped unnecessary columns
- Normalized JSON response
- Filtered for latest fetch_date
- Summarized totals and trends for each country

4. Visualizations Overview (Power BI)

KPI Cards (Top Section)

- Total Global Cases
- Total Deaths
- Total Recovered
- Active Cases
- Total Population
- Last Updated Date

Global Trend of Daily Cases

- Line chart with fetch_date on the X-axis
- Values: SUM(todayCases)
- Filtered by Country slicer

Top 10 Countries by Total Cases

- Bar chart
- Country vs Total Cases
- Filtered to latest fetch_date

Recovery vs. Death Rate vs Active cases

- Clustered bar chart
- Country on X-axis
- Recovery Rate (%), Death Rate (%)
- Shows top 10countries

Global Map

- Filled Map of countries
- Color-coded by total cases or deaths
- Tooltip: Active, Recovered, Deaths, Critical

Pie Chart – Global Case Distribution

- Proportions of Active, Recovered, and Deaths globally

Slicers and Filters

- Country, fetch_date

5. Key Insights
- Some countries like USA, India, Brazil consistently show high total case counts.
- Despite high cases, recovery rates in most countries are high (above 80%).
- Countries with lower populations often show high case-per-capita impact.

- The map visually communicates global hot zones.
- The daily new cases trend helps observe outbreak patterns or declines.

6. Technical Stack

Tool	Purpose
Python	Automated ETL pipeline (API → SQLite → Excel)
SQLite3	Lightweight local database for daily historical data
Pandas	Data wrangling and transformation
Power	Dashboarding and business intelligence
Task Scheduler	Windows-based daily automation of the Python script

7. Automation & Maintenance
- Python script stored in a secure local directory (scrambled path for confidentiality).
- Scheduled using a .bat file via Windows Task Scheduler to run daily.
- Excel file is always updated and replaced to reflect only the latest day's data.
- Connected live to Power BI using this Excel file.

8. Planning for Further Improvement
- Add rolling average for daily cases and deaths
- Include per capita metrics (cases per million)
- Add trend arrows or indicators on KPIs
- Export Power BI dashboard as an interactive web dashboard (via Power BI service)
- Enable alerts when major spikes are detected

9. Conclusion
This project demonstrates end-to-end data automation, cleaning, analytics, and dashboarding. It showcases essential skills for a data analyst or BI developer including:

- API integration
- ETL pipeline creation
- Data modeling
- Insightful dashboard design
- Automation and maintenance