Final Report: COVID-19 & Demographic Data Analysis

Overview

This report summarizes the key insights and findings extracted from the COVID-19 dataset used in the associated Jupyter Notebook. The data include daily records across countries, containing information about COVID-19 cases, deaths, testing, vaccinations, and demographic indicators such as population, life expectancy, and GDP per capita.

Data Description

- **Source**: Processed dataset from Our World In Data, containing 180,477 records.
- **Time Span**: Daily data entries for each country from the beginning of the pandemic.
- **Key Features**:
    - Epidemiological data: total cases, deaths, tests, vaccinations
    - Demographic and geographic: population, life expectancy, median age, GDP per capita, latitude, longitude

Key Findings

1. COVID-19 Cases per Capita (Map)

- A map-based visualization shows the ratio of confirmed cases to population.
- Some small countries (e.g., Gibraltar) display high ratios due to their population size.
- The map allows filtering by continent and metrics, improving regional comparison.

2. Vaccination Coverage – Total and Per Capita

- A dual bar chart reveals top countries by:
    - Absolute number of vaccinated people

    - Vaccination per capita

- In some countries, the vaccination ratio exceeds 1.0 due to booster doses being counted.

- A note was included to clarify this methodological nuance.

3. Time Series Comparison: Czechia vs. Slovakia

- This dashboard compares new daily COVID-19 cases in both countries over time.

- While the two nations follow similar pandemic waves, some temporal differences are evident.

4. Relationship Between Population and Life Expectancy

- A scatter plot shows the distribution of countries by population and life expectancy.

- Larger populations do not correlate directly with higher or lower life expectancy.

- High-income countries generally cluster with higher life expectancy.

5. Top 10 Countries by Population

- A bar chart visualizes the most populous countries, with China and India at the top.

- Other countries include the USA, Indonesia, Pakistan, and Brazil.

Conclusions

- The COVID-19 dataset allows for both regional and global insight into pandemic impacts.

- Visualization tools provide intuitive interpretation and dynamic interaction.

- Key takeaways involve understanding not only totals but also **relative metrics** (e.g., per capita, trends over time).

One key observation is that countries with smaller populations often display higher infection rates per capita. This does not always reflect a more severe epidemiological situation but may be a result of more accurate or frequent data reporting in those regions.

Recommendations

- Future analyses could include correlation and regression analysis.
- Machine learning could be applied to forecast case trends or identify clusters.
- Data updates and cleansing should be maintained for improved accuracy.

Prepared by: Denis Velcovsky
Date: 9/4/2025
Attached notebook: `Covid 19 Tracker.ipynb`