Professional Business Problems for COVID-19 Data Analysis

1. Global & Regional Pandemic Burden

Which regions (continent/country) experienced the highest COVID-19 burden (cases, deaths per million) relative to their population size and healthcare capacity (hospital beds, ICU availability)? This helps international organizations allocate aid to most vulnerable regions.

2. Case Fatality Rate (CFR) Monitoring

How did the case fatality rate (CFR = total_deaths / total_cases) evolve across countries, and which factors (age structure, life expectancy, comorbidities like diabetes prevalence, smoking rates) best explain differences? This identifies high-risk populations and health system weaknesses.

3. Vaccination Rollout Effectiveness

Did countries with faster vaccination rollouts (total_vaccinations_per_hundred) experience quicker declines in new cases and deaths per million? This evaluates the global efficiency of vaccination programs.

4. Vaccine Inequality & Economic Divide

Is there evidence of inequality in vaccine distribution, where high GDP per capita and high HDI countries achieved higher vaccination coverage sooner than low-income nations? This provides insights for international equity policies.

5. Stringency vs Outcomes

Did stricter government measures (stringency_index) correlate with lower transmission (reproduction_rate, new_cases_per_million) and fewer deaths? This evaluates effectiveness of lockdowns and policy interventions.

6. Hospital Pressure & Death Outcomes

How strongly are icu_patients, hosp_patients correlated with new_deaths, and do countries with higher hospital_beds_per_thousand show lower mortality rates during peaks? This highlights the role of healthcare infrastructure.

7. Testing & Surveillance Effectiveness

Does higher testing intensity (tests_per_case, total_tests_per_thousand) correlate with lower mortality and case spikes? This tests whether proactive testing prevents uncontrolled outbreaks.

8. Lagged Vaccine Impact

What is the time lag (in weeks) between increased vaccination coverage and observable reductions in deaths or cases? This helps policymakers set realistic expectations for vaccine rollouts.

9. Demographic Risk Factors

Do countries with older populations (aged_65_older, aged_70_older, median_age) have disproportionately higher deaths per million, even after vaccination campaigns? This supports prioritization of vaccines/boosters.

10. Lifestyle & Comorbidities

Is there a correlation between lifestyle factors (smoking prevalence, diabetes prevalence, cardiovascular death rates) and COVID mortality rates? This highlights vulnerabilities beyond just age and vaccines.

11. Anomalies & Data Integrity

Are there countries showing suspicious anomalies — e.g., very low reported deaths but also very low testing (possible underreporting)? This flags data reliability issues for health agencies.

12. Comparative Effectiveness of Measures

Between vaccination rate and stringency index, which factor had a stronger impact in reducing new cases and deaths? This provides evidence on whether vaccines or lockdowns were more impactful.

13. Long-Term Vaccine Effectiveness

Do new cases and deaths rise again after a period of high vaccination coverage (waning immunity, variants)? This informs booster dose policies.

14. Global Inequality in Pandemic Outcomes

Did countries with higher HDI and life expectancy experience better COVID outcomes (lower deaths per million) compared to less developed nations, even after adjusting for vaccination coverage? This measures socioeconomic impact.

15. Continent-Level Storytelling

Which continent managed COVID most effectively (lowest deaths per million vs highest vaccination rates), and what lessons can be replicated globally? This provides high-level insights for global health strategy.