Covid-19 Insights Analysis: Data Visualization Project Report

Student: Jun Fu

Instructor: Prof. Ling Zhang

Course: ITDS 2371

Class Term: Fall 2025

Date: December 5, 2025

Description of Project

This project focuses on analyzing global COVID-19 data using Power BI to deliver a comprehensive, interactive overview of the pandemic's development. By examining essential parameters—including confirmed cases, deaths, recoveries, and government response indicators, the project uncovers key trends across countries and over time. After loading and preparing the dataset, a series of important metrics is calculated, such as active case rate, recovery rate, mortality rate, infection rate by population, and average Government Response Index.

The dashboard integrates a variety of visualization types, including bar charts, a doughnut chart, and a line-and-clustered column chart, to provide a high-level summary. Additional report pages highlight specific perspectives and detailed breakdowns of the global situation such as time-series patterns, geographic distributions, and government policy responses.

By leveraging Power BI's interactive features, the project allows users to easily navigate the data, compare countries, and identify meaningful relationships between case numbers, population size, and policy actions. Overall, this project not only summarizes the evolving pandemic situation but also offers valuable insights into how COVID-19 spread and how different regions responded over time.

Insights / Questions of Your Data Visualization

Question 1	What are the most recent global averages for the infection rate and the Government Response Index recorded on July 27, 2020? How do these two indicators vary across countries?
Answers	0.21% 43.72
Reference Visual	Comparison of Infection Rate and Government Response Index by Country (Page 1)

Question 2	What are the most recent global average recovery rate, active rate, and death rate among confirmed cases as of July 27, 2020?
Answers	57.43% & 38.60% & 3.97%
Reference Visual	Recovered, Active, and Deaths Proportions (Page 1)

Question 3	What are the numbers of recovered, active, and death cases compared to the number of confirmed cases across countries?
Reference Visual	Total Cases vs. Recovered Cases (Page 1) Total Cases vs. Active Cases (Page 1) Total Cases vs. Deaths Cases (Page 1)

Question 4	How are COVID-19 cases and deaths distributed geographically across the world?
Reference Visual	Geographical distribution of cases and deaths (Page 2)

Question 5	How does the daily number of new cases and recoveries progress over time for a selected country?
Reference Visual	Daily New Confirmed Over Time (Page 3) Daily New Recovered Over Time (Page 3)

Question 6	How are COVID-19 confirmed, recovered, and death cases broken down hierarchically by WHO Region, country, month, and date?
Reference Visual	Breakdown of Cases (Page 4)

Question 7	How have global confirmed cases and deaths trended over time?
Reference Visual	Global Confirmed and Deaths Trends Over Time (Page 5)

Question 8	How have confirmed cases and deaths trended over time in the United States?
Reference Visual	US Confirmed and Deaths Trends Over Time (Page 6)

Question 9	How does the Government Response Index relate to the mortality rate across countries over time?
Reference Visual	Government Response Index vs. Mortality Rate Across Countries Over Time (Page 7)

Things That Went Well in the Project

The most exciting part of this project was using ChatGPT throughout the entire process.

The first task was searching for the datasets I needed. To clearly understand how government response policies influenced the trends of infected cases, I needed data on government stringency levels. Since no single dataset on Kaggle contained this information, I turned to ChatGPT. With its help, I spent a full day searching for a suitable and reliable dataset. Eventually, I selected **OxCGRT_nat_differentiated_withnotes_2020**, which comes from the Oxford COVID-19 Government Response Tracker (OxCGRT), a project developed by the University of Oxford. The goal of this project is to systematically collect information on how governments around the world responded to COVID-19.

I also needed a dataset containing the population of each country. The population dataset available on Kaggle was missing several important countries, including China, so once again I used ChatGPT to help me locate a more complete dataset on GitHub. Compared with the previous search, this one was much easier.

Another major benefit of using ChatGPT was in creating DAX measures and building visualizations. This was my first time learning Power BI, and although I was impressed by its powerful and abundant features, I was also overwhelmed by the number of functions to learn in such a short time. Throughout the project, I interacted with ChatGPT frequently. Although it sometimes gave unhelpful or even incorrect answers, those responses helped me recognize the gap between my questions and their understanding. By asking more detailed follow-up questions, I was ultimately able to achieve satisfying results.

Through this project, I truly experienced the power and convenience of AI tools. They helped me shorten much of the learning process and save programming time, allowing me to focus more on understanding the data, answering analytical questions, and presenting insights through visuals.

Challenges Addressed During the Project

The first challenge was my laptop. I use a Mac, but Power BI does not have a native Mac version. As a result, I borrowed an old Windows laptop from a friend, since none of the HCC campuses had one available. Because of this laptop's limited performance, the ArcGIS visual could not run the global daily progress of COVID-19 cases without crashing. Therefore, I had to adjust the time scale to weekly instead of daily. Although the laptop ran very slowly and sometimes became unresponsive, I was still able to complete the work.

Another challenge I did not expect was using Cards to switch between Bookmarks. On one report page, I wanted to create a feature where selecting different cards would trigger different Bookmarks. Although I believed I had configured everything correctly, it still refused to work. This was the only time ChatGPT could not help me identify the problem. It was only after I published the report to the Power BI Service that I realized this functionality does not work in Power BI Desktop but does work on the server. Unfortunately, it took me nearly an entire day to discover the reason.

Conclusion

In conclusion, this project gave me the opportunity to experience the powerful and diverse capabilities of Power BI. With Power BI, it becomes much easier to uncover relationships among key KPIs—such as confirmed cases, recovered cases, deaths, and the government response index—in an intuitive and interactive way. The project also provided meaningful insights into COVID-19 patterns across the world, enabling me to present a comprehensive and dynamic overview of the pandemic's development and to illustrate how government stringency levels influenced the trends of infected cases.

Questions and Reflections

Firstly, how can we ensure data consistency and standardization? As we know, data is the foundation of all calculations, visualizations, and analyses. Yet, statistical standards vary significantly across countries. For example, according to public information, China generally counts COVID-19 deaths only when the death is directly caused by COVID-19, such as pneumonia or respiratory failure. If a patient has underlying conditions (e.g., heart disease, diabetes), even if they die while infected with COVID-19, it may not be counted as a COVID-19 death. In contrast, the U.S. uses a broader standard — as long as COVID-19 is considered a contributing factor to the death, it is counted as a COVID-19 death. This means patients with underlying conditions are included if COVID-19 worsened their condition. Because of these differing definitions, comparing official death rates or deaths per 100,000 population between the two countries can be misleading. In fact, recent studies using excess mortality suggest that actual deaths in China after loosening restrictions may be far higher than official numbers. Therefore, it is critical to establish a universal standard when collecting and analyzing COVID-19 data.

Secondly, is the government response truly effective? According to common understanding — and my own assumption — one would expect that strict government restrictions would have a positive effect in reducing infection rates. However, in this project, whether using a line-clustered column chart or a scatter chart to compare infection rates with the government response index, the data did not support this expectation. Therefore, I believe that while strong government interventions may be effective at the very beginning of an outbreak, their long-term impact may not be as significant as we might imagine. Further analysis from multiple perspectives is needed to fully understand these effects.

Thirdly, are there other key factors that may be even more important when facing such serious disaster? I would like to highlight a special case: Taiwan. Taiwan has a total population of over 23 million, yet its infection rate was only 0.0019%, meaning that roughly 19 people per million were infected. Considering Taiwan's proximity to China and the extensive economic and personal exchanges between the two, this number is remarkable. At the same time, Taiwan's government response index was only 33.21, far below the global average of 43.69.

I believe that in addition to effective scientific governance, another important factor contributed to Taiwan's success: as early as December 2019, Taiwan's health authorities noticed reports from Wuhan, China about "pneumonia of unknown cause" and began monitoring the situation — even though the Chinese government at the time dismissed this information as a rumor. Taiwan implemented preventive measures at the end of 2019 and early 2020, several weeks to a month earlier than many other countries. This early action helped Taiwan successfully control the outbreak in its initial stages.

Therefore, while the government response index is an important metric, this example suggests that **transparent government policies and freely available information may be even more crucial for protecting public health.** I hope to find data in the future to further compare and explore this insight.