Project, STAT 650

Visualization of COVID-19 data worldwide using R

Gyubeen Park(si5722)

1. Introduction

In response to the pandemic of COVID-19, humans must make efforts to solve problems and share knowledge. Data-based decision-making is very important in a world that has been completely transformed into a covid-19 pandemic. In this research, we will use global COVID-19 data to identify the composition and basic trends of the data and visualize them. In this project, we check the status of COVID-19 outbreaks in many countries around the world and deaths and vaccines in the United States. It also examines the relationship between vaccines and COVID-19, the relationship between the elderly population and death, and the relationship between the human development index and COVID-19.

2. Data Description

The data set comes Our World in Data: https://ourworldindata.org/covid-vaccinations. The complete COVID-19 dataset is a collection of the COVID-19 data maintained by Our World in Data. Data has 221,268 rows and 67 columns. The variables are described in detail in: https://github.com/owid/covid-19-data/tree/master/public/data. This data is updated daily, and this analysis used updated data until October 4, 2022.

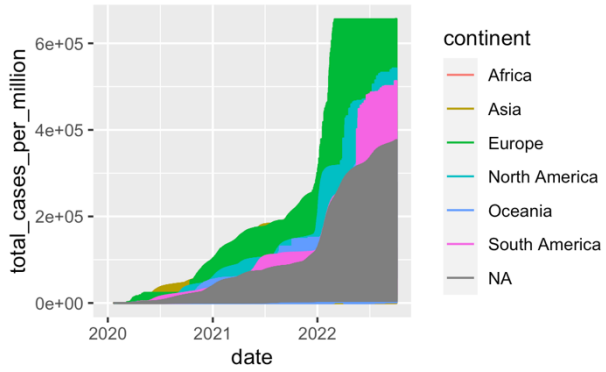
Variable	Description
total_cases	Total confirmed cases of COVID-19. Counts can include probable cases, where reported.
total_deaths	Total deaths attributed to COVID-19. Counts can include probable deaths, where reported.
total_vaccinations	Total number of COVID-19 vaccination doses administered.
aged_65_older	Share of the population that is 65 years and older, most recent year available
human_development_index	A composite index measuring average achievement in three basic dimensions of human development—a long and healthy life, knowledge and a decent standard of living. Values for 2019, imported from http://hdr.undp.org/en/indicators/137506
…	…

3. Results

High to low		Low to high	
location	Avg_cases	North Korea	0.00
Nauru	2002.86	Yemen	0.40
Cook Islands	748.52	Niger	0.40
Palau	740.66	Chad	0.47
Cyprus	701.86	Tanzania	0.66
Faeroe Islands	693.45	China	0.71

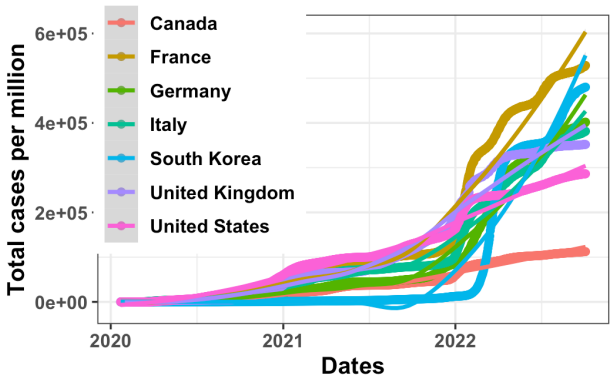
<Table 1. COVID-19 outbreak chart by country>

<Table 1> shows the top countries with the greatest number of COVID-19 outbreaks and the bottom countries with the least number of COVID-19. In the case of communist countries, there are fewer cases of COVID-19.

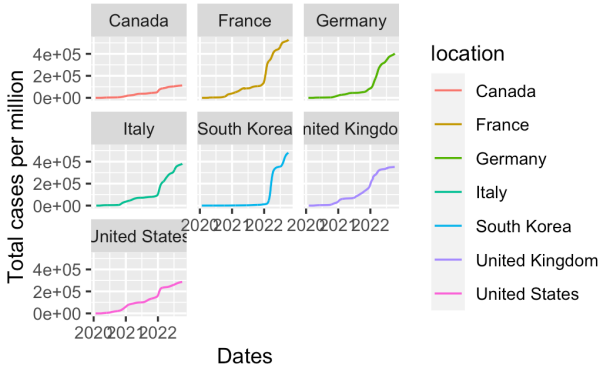


<Figure 1. Time series chart of the total covid-19 cases>

<Figure 1> is a time series chart of the total COVID-19, cases worldwide within the continent. In early 2022, the number of confirmed cases increased rapidly worldwide. COVID-19 occurred the most in continental Europe.

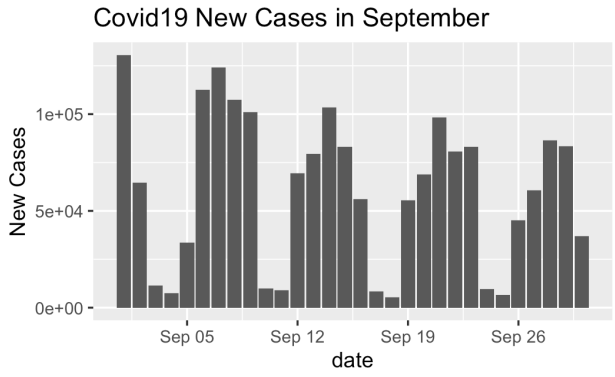


<Figure 2-1. COVID-19 Outbreaks in 7 Countries>



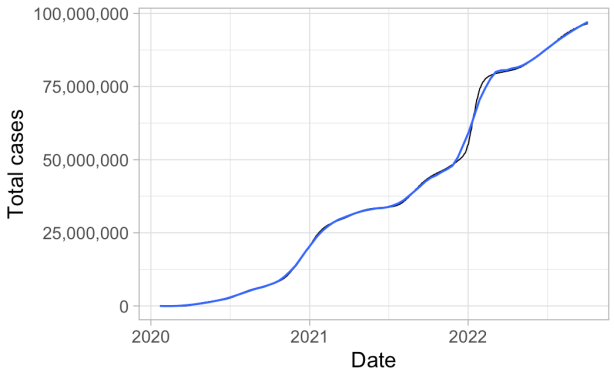
<Figure 2-2. Trends in COVID-19 in each country>

<Figure 2> below shows the trends of the seven countries (Canada, France, Germany, Italy, South Korea, United Kingdom, and the United States) in more detail. Of the seven countries, Canada has fewer coronavirus cases. In addition, it increased rapidly in early 2022, especially in Korea and France.

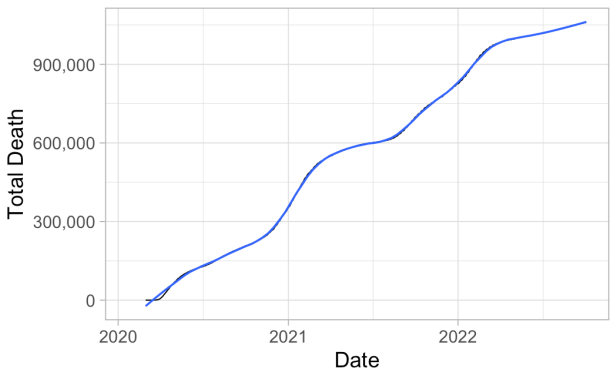


<Figure 3. U.S. COVID-19 outbreak status in September>

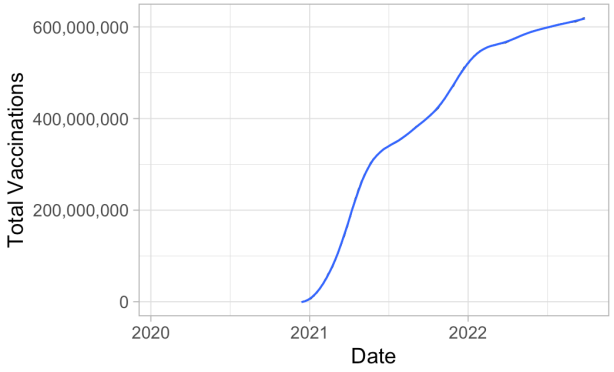
<Figure 3> shows the trend of confirmation within September. The graph is a multimodal distribution with several humps, and the number of confirmed cases is gradually decreasing.



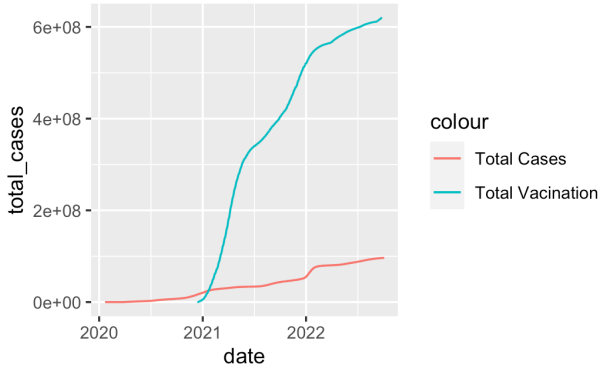
<Figure 4-1. U.S. COVID-19 Total cases Trend>



< Figure 4-2. U.S. COVID-19 Total Death Trend>

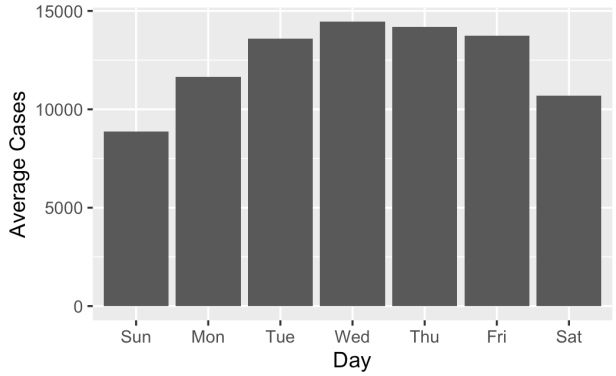


<Figure 4-3. U.S. COVID-19 Total Vaccinations Trend>



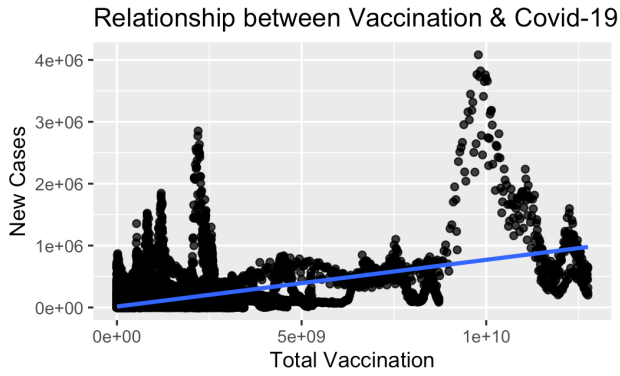
<Figure 4-4. U.S. Total Cases and Vaccinations>

<Figure 4> shows the overall trend of COVID-19 in the United States. The COVID-19 outbreak continues to occur over time, and the trend of confirmed cases and deaths is similar.



<Figure 5. Average number of Cases per day of the week>

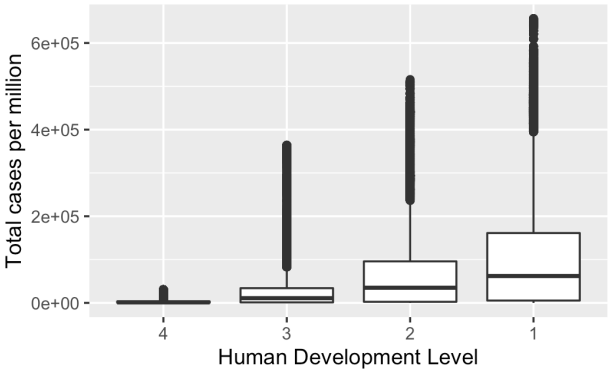
<Figure 5> is a graph to find out which days of the week COVID-19 cases occur a lot. There are fewer cases on weekends, and it occurs most frequently on Wednesdays. It is estimated that the confirmed cases are generated after an incubation period of 3-5 days after contact with the confirmed cases on weekends when the floating population is large.



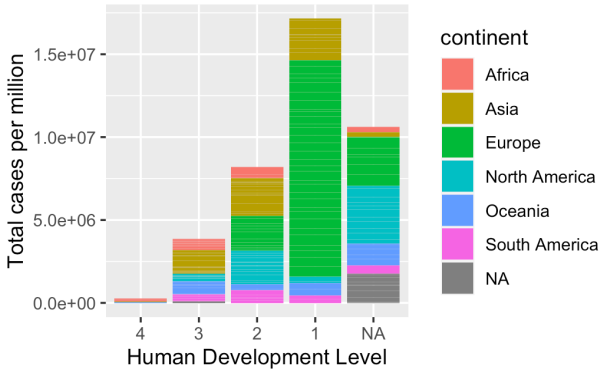
<Figure 6-1. Relation between Vaccination and COVID-19>

<Figure 6-2. Relationship between the Elderly Population>

<Figure 6-1> depicts the relationship between the two with the total vaccination as a predictor variable and the new case of COVID-19 as a response variable. Despite the increase in vaccination status, COVID-19 spreads more and eventually show a positive correlation. Figure 6.2 shows that the larger the elderly population (predictor variable), the more deaths from COVID-19 (response variable).

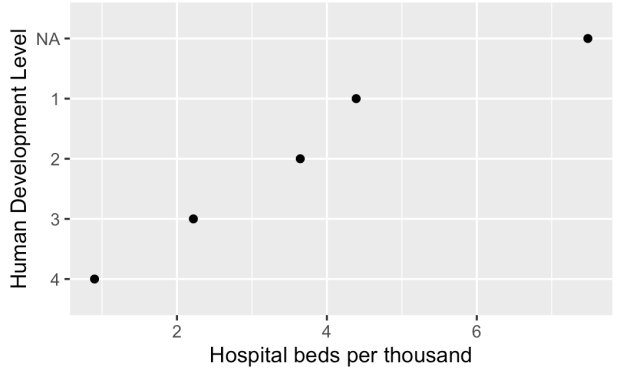


<Figure 7-1. Total cases for each Level>

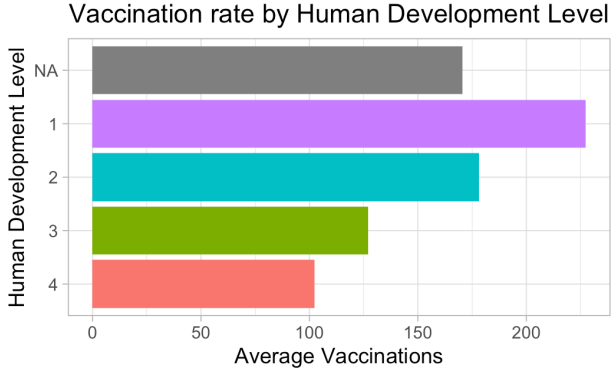


<Figure 7-2. Total Cases for each level with continent>

To confirm the relationship between the human development index and COVID-19, it was divided into four levels based on the quartile. In <Figure 7-1>, the higher the index, the more cases of COVID-19. It can be assumed that this is because the more advanced countries, the more active production activities and the larger the floating population. In <Figure 7-2>, the continents constituting each level can be known and the main composition of Level 1 in Europe.



<Figure 8-1. Hospital beds ratio for each of Level >



<Figure 8-2. Vaccination rate by each Level>

To see more details of the Human Development Index, <Figures 8-1>, <Figure 8-2> examine the average number of hospital beds and vaccination at each level. As a result, the higher the index, the more hospital beds, and vaccinations were.