BST260 Final Project

Alice Zhang

Abstract

This project investigates the impact of COVID-19 across the United States from January 2020 to December 2024. The purpose of the study is to provide perspectives on COVID-19's spread, and findings could be helpful for future research in public health and policy development. Case data and population statistics across U.S. states were used and data was divided into pandemic waves based on epidemiological trends. Analysis results revealed significant variability of COVID-19 cases across states and waves. The analysis highlighted states that performed better and/or during the pandemic reflected by case rate. These results shows the dynamic nature of the pandemic and the association between COVID-19 spread and geographical factors or public health interventions.

Introduction

The COVID-19 pandemic - the spread of SARS-CoV2 virus, significant impacted global health and economy since the beginning of 2020. The virus began causing illnesses in late 2019, then spread hitting countries without warning. Throughout the next 3 years and until now, aftermaths of the pandemic is still impacting the society in health and socio-economical aspects. With the evolution of viral variants, development of vaccines and social awareness developed to prevent viral spread, the pandemic hit the U.S. in waves throughout 2020 to 2024.

Understanding the spread and impact of COVID-19 across the country is crucial for effective strategies and interventions to be developed in preparation for next pandemics or epidemics within the country. Investigating COVID-19 trends across different state and identifying well-performed states is necessary for reflecting and adapting to suitable/successful healthcare policies.

We identified five pandemic waves within the time period January 2020 to December 2024. The goal is to explore factors that led to successful management of viral spread.

Methedology

The data set used for this project were publicly available state level SARS-CoV2 cases data from from Centers for Disease Control and Prevention (CDC), and 2021 population data from the U.S. Census. Data analysis was performed in R, employing packages including *tidyverse* for data wrangling, *ggplot2* for visualization, *knitr* and *httr2*. The study assumed that the quality and procedure of data collection across different states were consistent. The study also extrapolated for population data that were unreported and based result interpretations on case rates for better/worse performance of states.

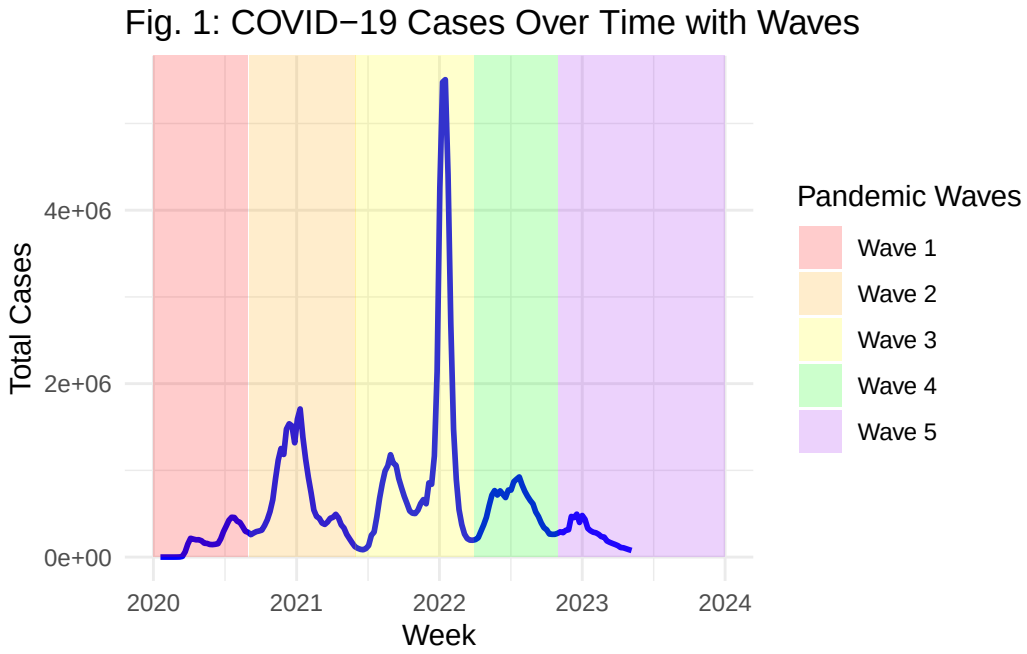
The time period that the data spans was divided into five waves aided by data visualization: Wave 1 (January 2020 to August 2020), Wave 2 (September 2020 to May 2021), Wave 3 (June 2021 to March 2022), Wave 4 (April 2022 to October 2022), Wave 5 (November 2022 to 2024). See Figure **number**

Case data were cleaned to include only relevant features: state, date, and daily new cases. Cases were summarized by waves and case rates per 100,000 people were computed by joining with the census population data. States with highest and lowest case rates for each wave were identified.

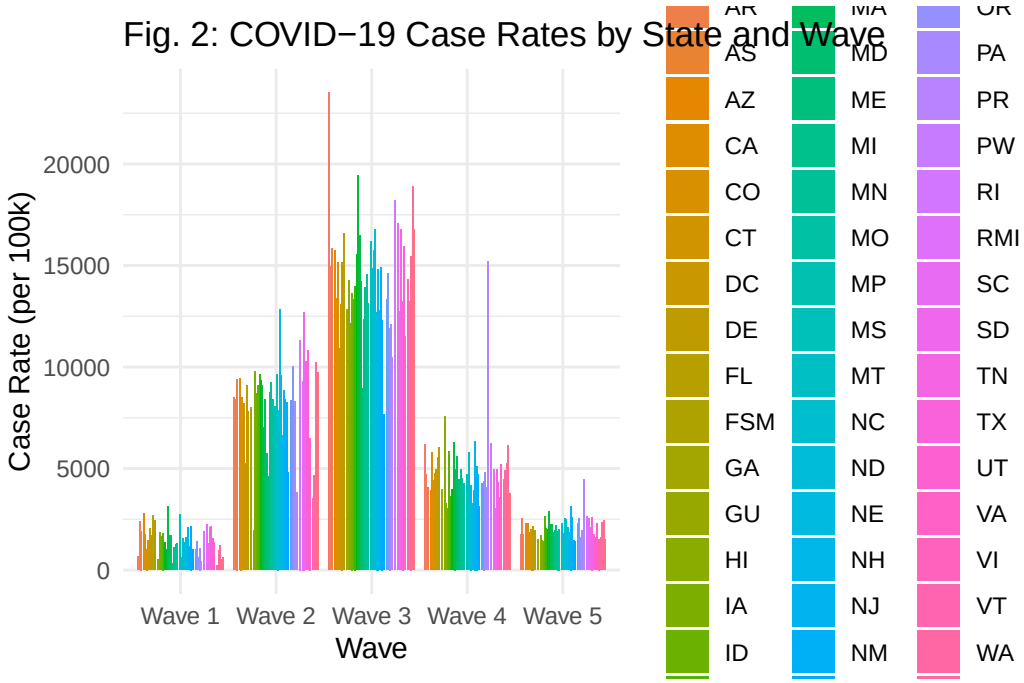
Results

We found that there were five waves of pandemic nationally, identified by peaks of case numbers. The span of each wave is represented by a different shading.

```
Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
i Please use `linewidth` instead.
```



For each period, state level case rate was calculated.



The states with highest and lowest case rates are highlighted below.

wave	max_rate_state	max_rate	min_rate_state	min_rate
Wave 1	LA	3151.395	VT	241.1822
Wave 2	ND	12815.828	HI	1919.5391
Wave 3	AK	23516.160	NY	7543.5180
Wave 4	PR	15217.687	ID	2941.4658
Wave 5	PR	4440.486	DC	1332.4291

From total cases per wave, we see that COVID got more virulent during the first three waves, then virulence decreased after the third wave.

Fig. 3: Average Case Rates by Wave

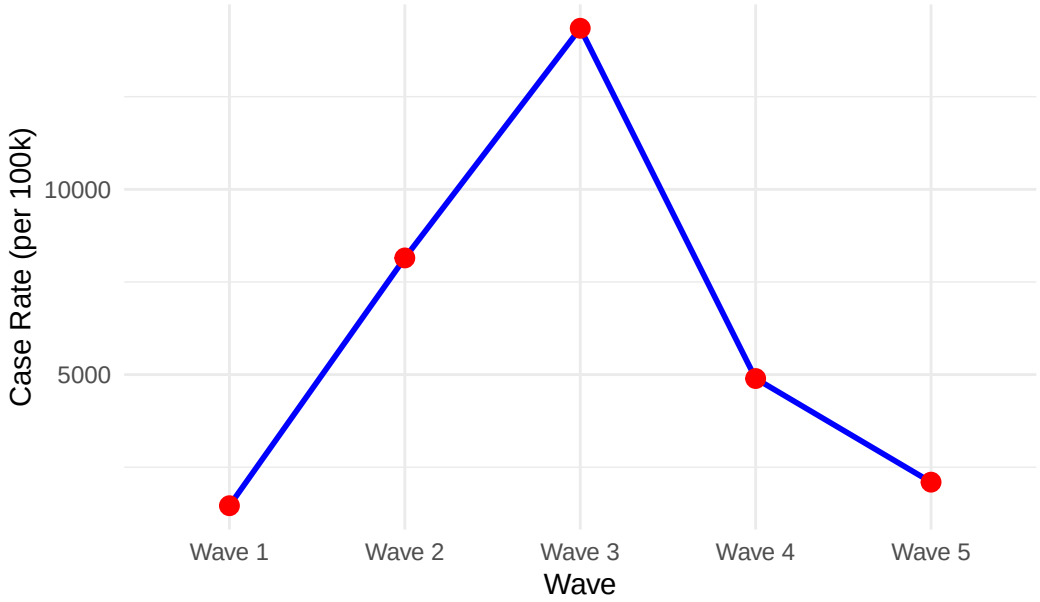
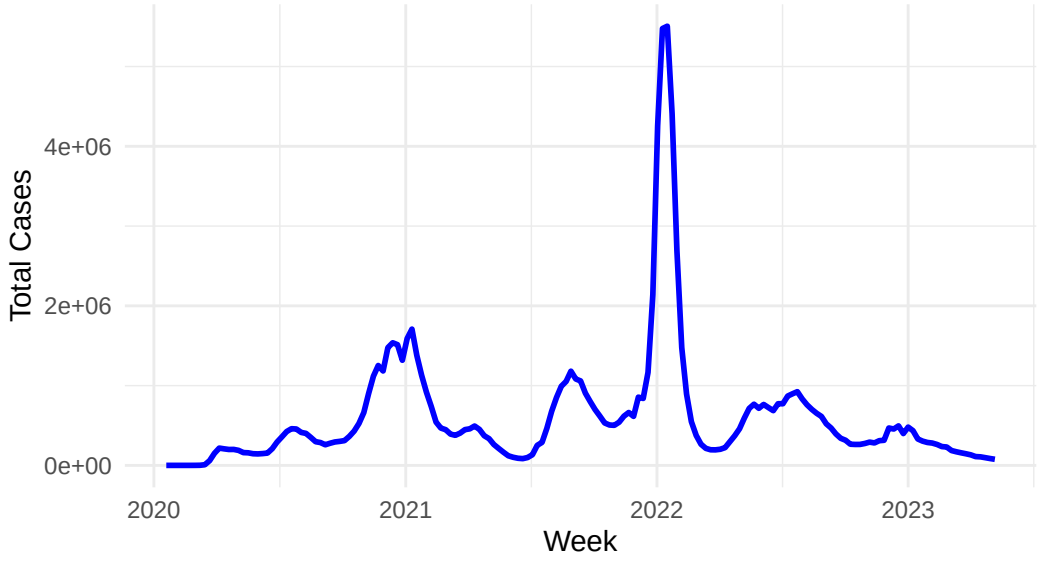


Fig. 4: COVID−19 Cases Over Time

Identifying Pandemic Waves (2020−2024)



Discussion

The findings revealed the dynamic nature of the COVID-19 pandemic throughout the four years, case rates across states and waves were highly variable. States that performed well

were Vermont, Hawaii, New York, Idaho, and the District of Columbia. According to Centers for Medicare and Medicaid Services (https://data.cms.gov/), Vermont, Hawaii, New York and District of Columbia were among the top states with highest vaccination rates. In contrast, Puerto Rico was among the states with lowest vaccination rates. Consistent with the challenges PR faced during multiple pandemic waves.

The reduction in case rates during the fourth wave (April 2022 to October 2022) suggests the combined effects of vaccination, gaining immunity from previous contractions, and the viral evolution into less severe variants. According to CDC, in April 2022, the Omicron subvariant BA.2 makes up more than 85% of all new COVID-19 infections in the U.S. and Omicron is less severe than previous strains. However, there was a upswing in Puerto Rico as shown in Figure 2, this consistent with the finding that PR had low vaccination rates and implies that vaccine hesitancy is a factor that influence how the pandemic trended.

Some future directions to be considered are: exploring the effect of socioeconomic factors, healthcare equity, and public health policies (e.g. social distancing, lock downs, quarantine) on pandemic outcomes. These finding allow policymakers and healthcare providers to better prepare for future nation-wide and/or global health crises. Public health strategies that are consistent and adaptive across different states and different legislative are the key to better prevention of disease spread.

In conclusion, this study conducted state-level analysis in understanding and managing pandemics. By examining case rates across waves, we gained knowledge on how to identify best practices for future public health challenges to be faced.

Citations

Centers for Disease Control and Prevention. (2023, March 15). *CDC Museum Covid-19 Timeline*. Centers for Disease Control and Prevention. https://www.cdc.gov/museum/timeline/covid19.html

Centers for Medicare and Medicaid Services. *COVID-19 Vaccination Rates - State and National Averages*. Centers for Medicare and Medicaid Services. (2024, December 8). https://data.cms.gov/provider-data/dataset/avax-cv19