CRIME AND PUBLIC HEALTH: A COMPARATIVE ANALYSIS OF CRIME RATES AND COVID-19 DATA IN CHICAGO (2020 - 2022)

BY:

OLOLADE AKINSANOLA

COURSE CODE:

DSCI 401

INSTRUCTOR'S NAME:

DR MENA WHALEN

UNIVERSITY NAME:

LOYOLA UNIVERSITY CHICAGO

DATE:

NOVEMBER 2024

OUTLINE

INTRODUCTION

- Briefly describe the dataset(s) and context (COVID-19 and crime data in Chicago).
- State the main research questions and objectives.

DATA OVERVIEW

- Explain the data cleaning process and describe the dataset's structure.
- Include important variables and their significance.

ANALYSIS AND EXPLORATION

- Present the analysis for each of your major questions (e.g., changes in crime rates, correlation with COVID-19 cases).
- Discuss visualizations and key findings.

DISCUSSION

- Interpret the results, offer insights into any patterns you discovered.
- Discuss limitations and potential biases in the data.

CONCLUSION

- Summarize the findings.
- Mention any potential future research or follow-up analysis.

INTRODUCTION

Description

For this project, I will analyze crime and COVID-19 data from the City of Chicago to examine the impact of the COVID-19 pandemic on crime rates between 2020 and 2022. The dataset, sourced from the Chicago Data Portal, includes detailed records of reported crimes across various neighborhoods, spanning a wide range of offenses from theft and burglary to violent crimes like arson and assault. It also includes COVID-19 case counts, testing rates, deaths, and community-level demographics to provide context.

I have processed, cleaned, and joined the datasets to include 498,713 crime records. Key variables include crime type, community area, ZIP code, population, as well as COVID-19-related data like weekly cases, deaths, and positivity rates. This data is essential in understanding how the pandemic may have influenced crime patterns in different parts of the city.

The City of Chicago, with a population of over 2.7 million, experienced significant changes during the pandemic, which could have impacted crime rates. This analysis is important for policymakers, law enforcement, and public health officials to understand the relationship between COVID-19 surges, lockdowns, and crime patterns.

My interest in how social and economic factors affect crime and public safety gives me the opportunity to apply data science techniques to explore the intersection of between them, focusing on how crime rates evolved in different neighborhoods and whether COVID-19 case rates correlate with shifts in criminal activity. This project aims to provide insights into how cities can better manage public safety during future public health crises.

Major Questions

I was interested in exploring the relationship between crime and the COVID-19 dataset to understand their correlation. I wanted to determine whether the COVID-19 period had an impact on crime in Chicago and how long these effects lasted. Additionally, I aimed to identify which types of crimes were more prevalent and which community areas experienced higher rates of crime. I also wanted to examine the impact of COVID-19 on these communities. Therefore, I will be focusing on the following questions:

1. Is there a correlation between COVID-19 cases rate and the crime rates?
2. How did crime rates vary by community area during the COVID-19 period?
3. How did crime rates change from 2019 to 2020-2022 (COVID-19 era)?
4. Did certain areas experience more severe COVID-19 impacts (in terms of positivity rate and deaths)?
5. What is the relationship between population size and crime rates during COVID-19?
6. How did testing rates correlate with COVID-19 case rates across zip codes?
7. How did weekly testing and positivity rates compare across different years (2020-2022)?

DATA OVERVIEW

Data Cleaning Process

The data preparation involved integrating crime records and COVID-19 statistics from Chicago for analysis during the 2019–2022 period. To ensure consistency and relevance, the datasets were filtered, merged, and transformed as follows:

COVID-19 Data Processing:

1. Merged COVID-19 case data by ZIP codes with community area mappings.
2. Extracted the year and week numbers from the dataset's weekly date range to enable temporal analysis.
3. Removed extraneous columns, focusing on relevant metrics like case counts, positivity rates, and deaths.

Crime Data Processing:

1. Cleaned crime data by removing non-essential attributes (e.g., geographic coordinates and administrative information).
2. Extracted year and week numbers from crime incident dates.
3. Grouped crime records by year, week, type, and community area to compute weekly crime counts.

Data Integration:

1. Crime and COVID-19 datasets were joined on community area and week numbers, aligning crime counts with pandemic metrics.
2. Additional filtering was applied to focus on the 2020–2022 pandemic period, with data from 2019 used as a baseline for comparison.
3. Any rows with missing data were removed to ensure analytical accuracy.

Dataset Structure

The cleaned and merged dataset includes weekly records of crimes alongside COVID-19 statistics, grouped by community area. Key variables include:

1. Crime Data: Crime type, count, and community area.
2. COVID-19 Metrics: Weekly case counts, positivity rates, deaths, and testing rates.
3. Temporal Indicators: Year and week number.

This structured dataset enables the examination of crime trends alongside pandemic dynamics, providing insights into their interplay across Chicago's neighborhoods.

ANALYSIS AND EXPLORATION

The analysis aimed to examine the interplay between COVID-19 and crime rates in Chicago, highlighting significant trends and disparities during the pandemic (2020–2022). By exploring various facets of the data—correlations, community impacts, and temporal patterns—the results shed light on how the pandemic influenced social and public health dynamics. Below, the findings are presented systematically to address the research questions.

Variations in Crime Rates by Community Area During COVID-19

To explore how crime rates varied by community area during the pandemic, I calculated total crime counts for each area and visualized the results using a bar plot (Figure 3). The findings revealed significant differences:

- Areas with the highest total crime counts included Near North Side (221,078), Austin (206,823), and Near West Side (193,380).
- In contrast, Montclare (667), Riverdale (1,414), and Edison Park (2,303) recorded the lowest total crime counts.

These results suggest that urban centers and highly populated areas experienced higher crime rates during the pandemic.

Crime Rates: Pre-Pandemic vs. Pandemic Era

To understand the influence of COVID-19 on crime, I compared crime rates from 2019 (pre-pandemic) to 2020–2022. Using bar plots (Figures 6a and 6b), I found that total crime counts increased significantly during the pandemic in several areas:

- Austin (14,795 in 2019 to 206,823 during the pandemic)
- Near North Side (12,603 to 221,078)
- Loop (10,607 to significantly higher counts)

Overall, the results suggest that socio-economic disruptions, such as increased unemployment and remote work, may have contributed to these patterns.

Correlation Between COVID-19 Case Rates and Crime Rates

To determine if there was a relationship between COVID-19 case rates and crime rates, I plotted weekly crime counts and COVID-19 case rates against the start of each week (Figures 1a and 1b). The results revealed that

during the COVID-19 period, weekly crime counts appeared to decline compared to pre-pandemic levels. However, COVID-19 case rates showed large weekly fluctuations, with a notable spike observed from late 2021 to 2022. This suggests that while overall crime rates dropped, the pandemic's timeline may have influenced these patterns in complex ways.

Testing and Positivity Rate Trends Across Years

To assess how testing and positivity rates changed over time, I grouped the data by year and summarized it to calculate total weekly tests, positivity rates, cases, and deaths. Violin and bar plots (Figures 4a–4d) illustrated these trends:

- Testing Rates: Testing increased steadily from July 2020 to January 2021, followed by a sharp decline. Testing rates fluctuated throughout 2021 and 2022, with a general downward trend in the latter half of 2022.
- Positivity Rates: Positivity rates were highest between April and July 2020, decreased towards the end of 2020, and showed slight fluctuations from 2021 to 2022.

Community Areas Most Affected by COVID-19

I examined the average positivity rates and weekly death counts by community area to identify neighborhoods most severely impacted by COVID-19. After grouping and summarizing the data, I visualized the results using bar plots (Figures 2a and 2b).

- Positivity Rates: The top three community areas with the highest positivity rates were Montclare (12.99), Riverdale (10.53), and Chicago Lawn (10.08). Conversely, O'Hare, Washington Park, and Greater Grand Crossing exhibited the lowest positivity rates at 0, 4.45, and 4.49, respectively.
- Death Counts: For weekly death counts, the highest averages were in Gage Park (2.44), Chicago Lawn (2.43), and West Lawn (2.41), while O'Hare (0), Loop (0.05), and Hegewisch (0.20) recorded the lowest averages. These findings highlight disparities in COVID-19 impacts across different neighborhoods.

Correlation Between Testing Rates and COVID-19 Cases

I analyzed the relationship between weekly testing rates and COVID-19 case counts across ZIP codes using scatterplots (Figure 5). The analysis showed a strong positive correlation (0.903), indicating that higher testing rates were associated with higher reported case counts.

Severity of COVID-19 Impacts by Area

Lastly, I explored whether certain areas experienced more severe impacts in terms of positivity rates and deaths. The visualizations (Figures 7a and 7b) revealed notable variations, with some areas reporting higher deaths despite lower positivity rates, emphasizing localized differences in pandemic outcomes.

Figure 1a

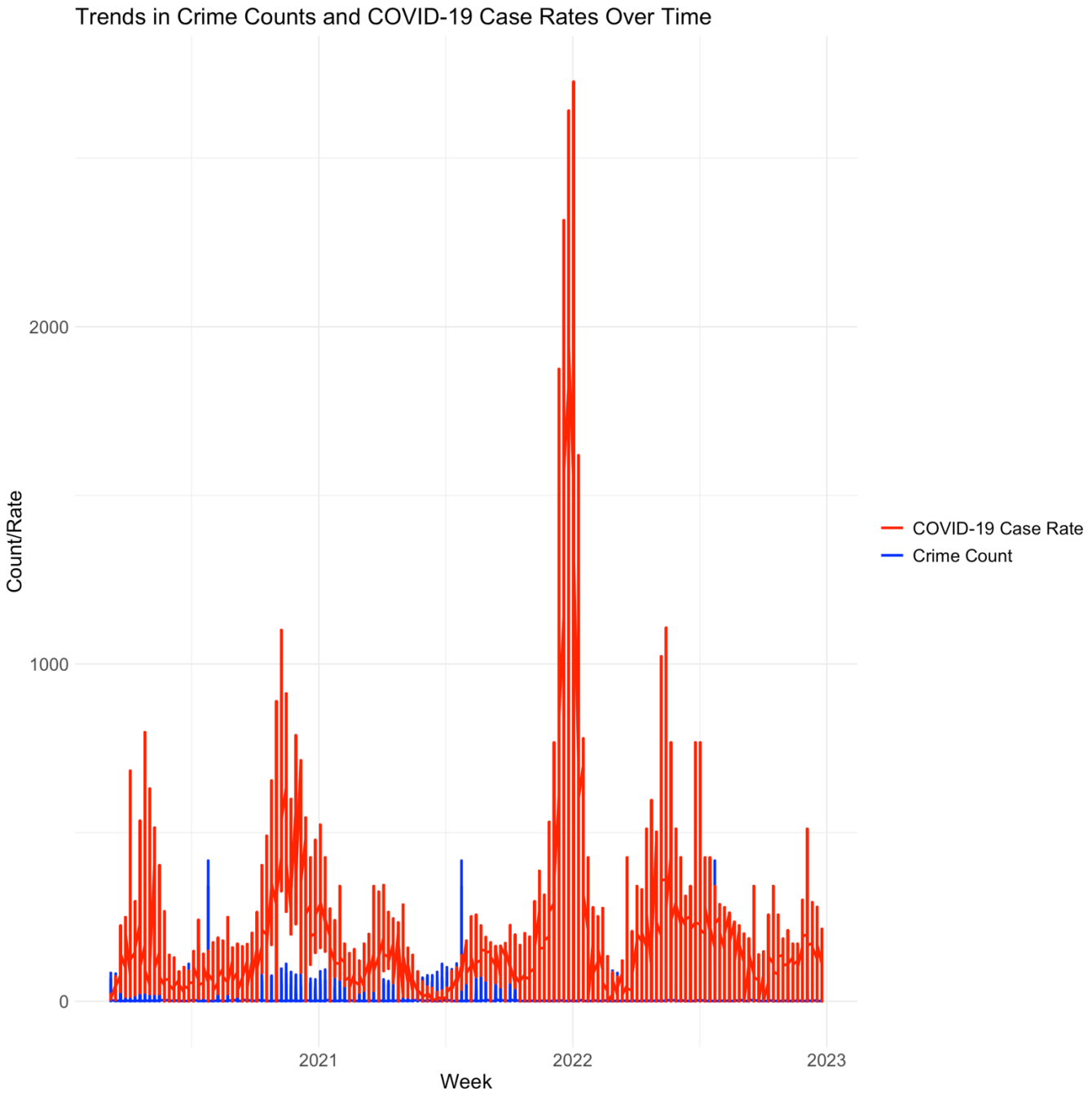


Figure 1b

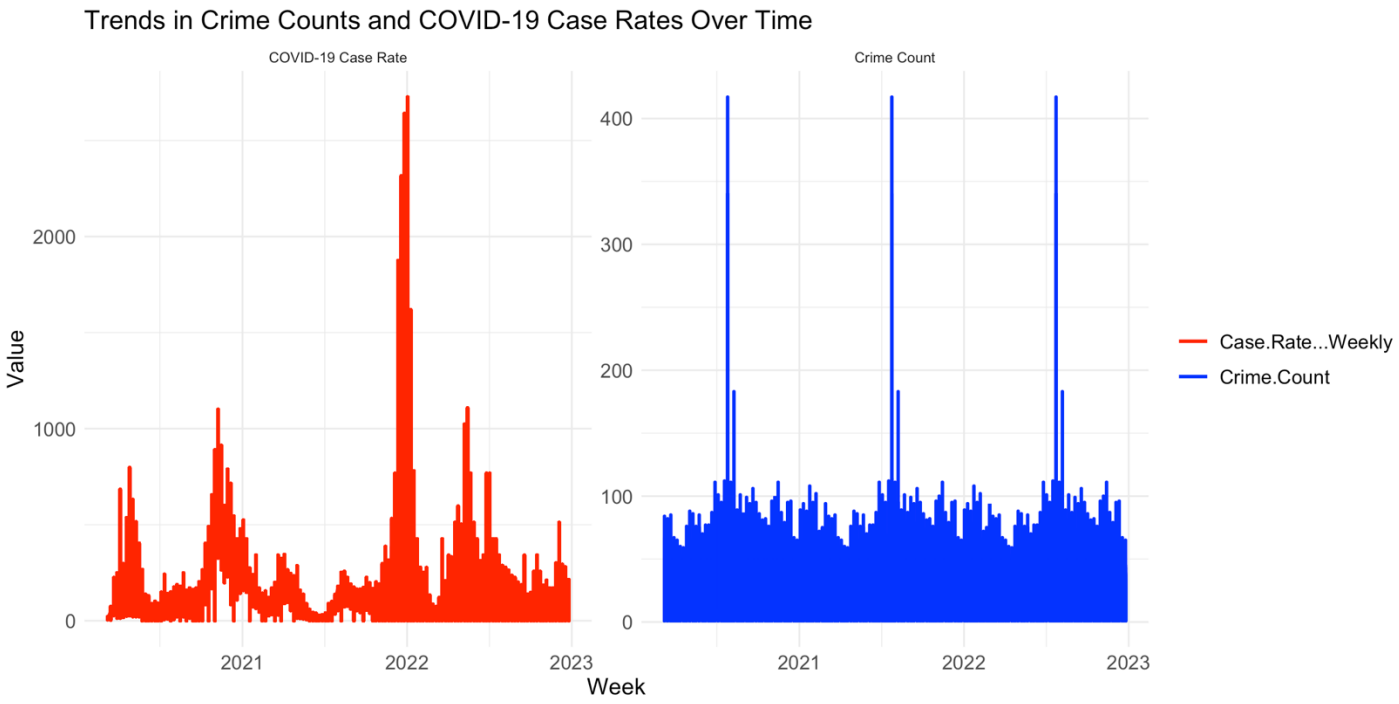


Figure 2a

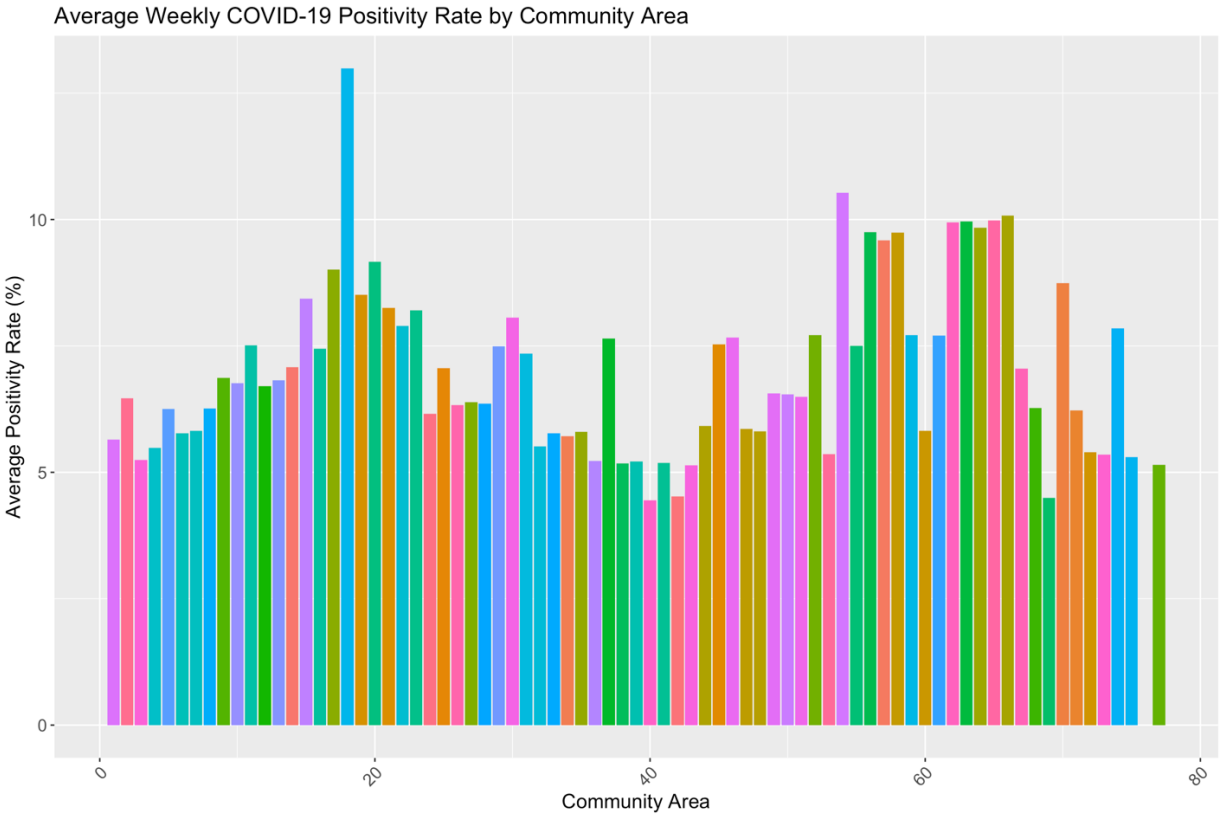
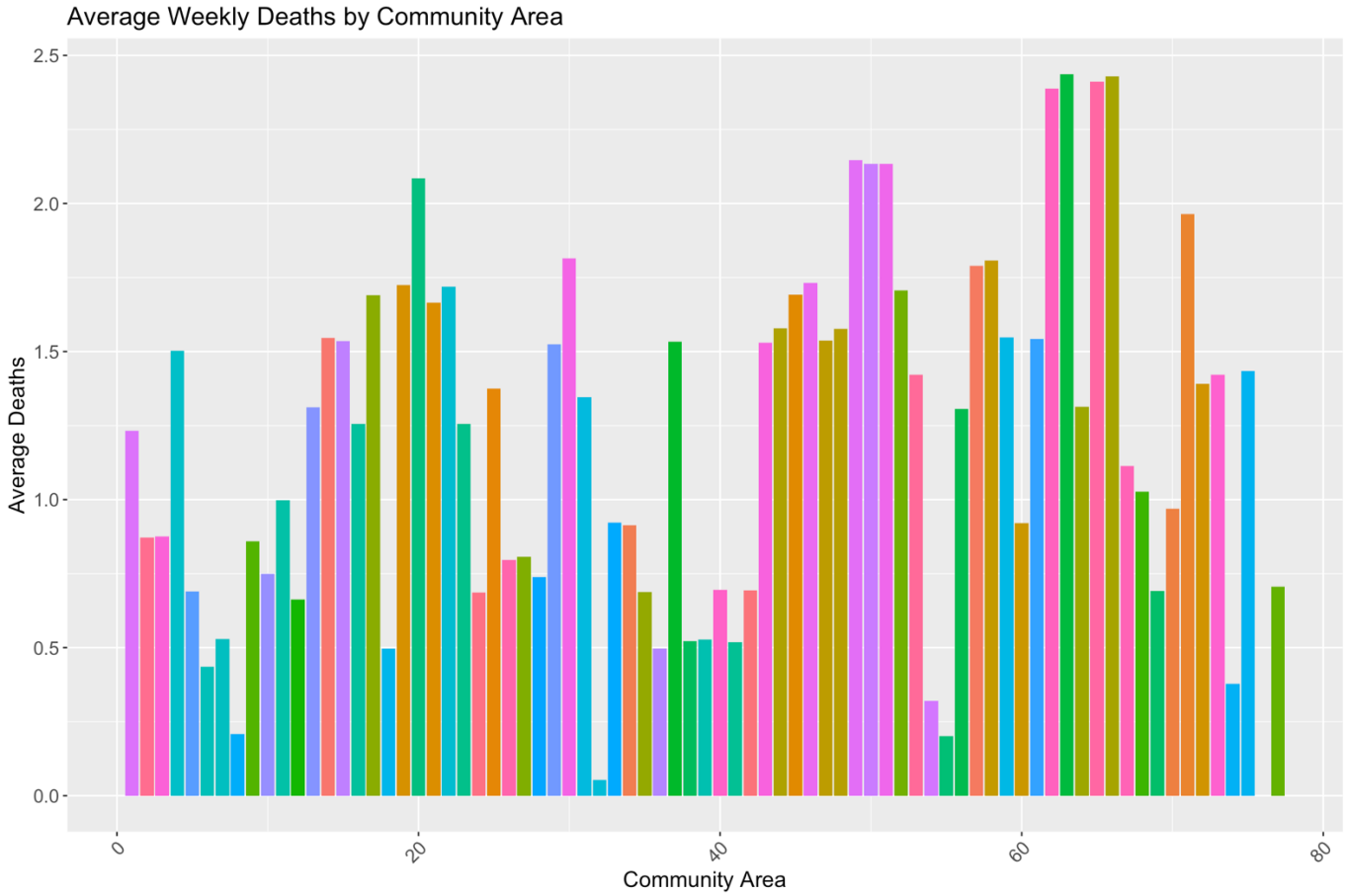


Figure 2b



Legend:

Albany Park
Archer Heights
Armour Square
Ashburn
Auburn Gresham
Austin
Avalon Park
Avondale
Belmont Cragin
Beverly
Bridgeport
Brighton Park
Burnside
Calumet Heights
Chatham
Chicago Lawn
Clearing
Douglas
Dunning
East Garfield Park
East Side
Edgewater
Edison Park
Englewood
Forest Glen
Fuller Park
Gage Park
Garfield Ridge
Grand Boulevard
Greater Grand Crossing
Hegewisch
Hermosa
Humboldt Park
Hyde Park
Irving Park
Jefferson Park
Kenwood
Lake View
Lincoln Park
Lincoln Square
Logan Square
Loop
Lower West Side
McKinley Park
Montclare
Morgan Park
Mount Greenwood
Near North Side
Near South Side
Near West Side
New City
North Center
North Lawndale
North Park
Norwood Park
O'Hare
Oakland
Portage Park
Pullman
Riverdale
Rogers Park
Roseland
South Chicago
South Deering
South Lawndale
South Shore
Uptown
Washington Heights
Washington Park
West Elsdon
West Englewood
West Garfield Park
West Lawn
West Pullman
West Ridge
West Town
Woodlawn

Figure 3

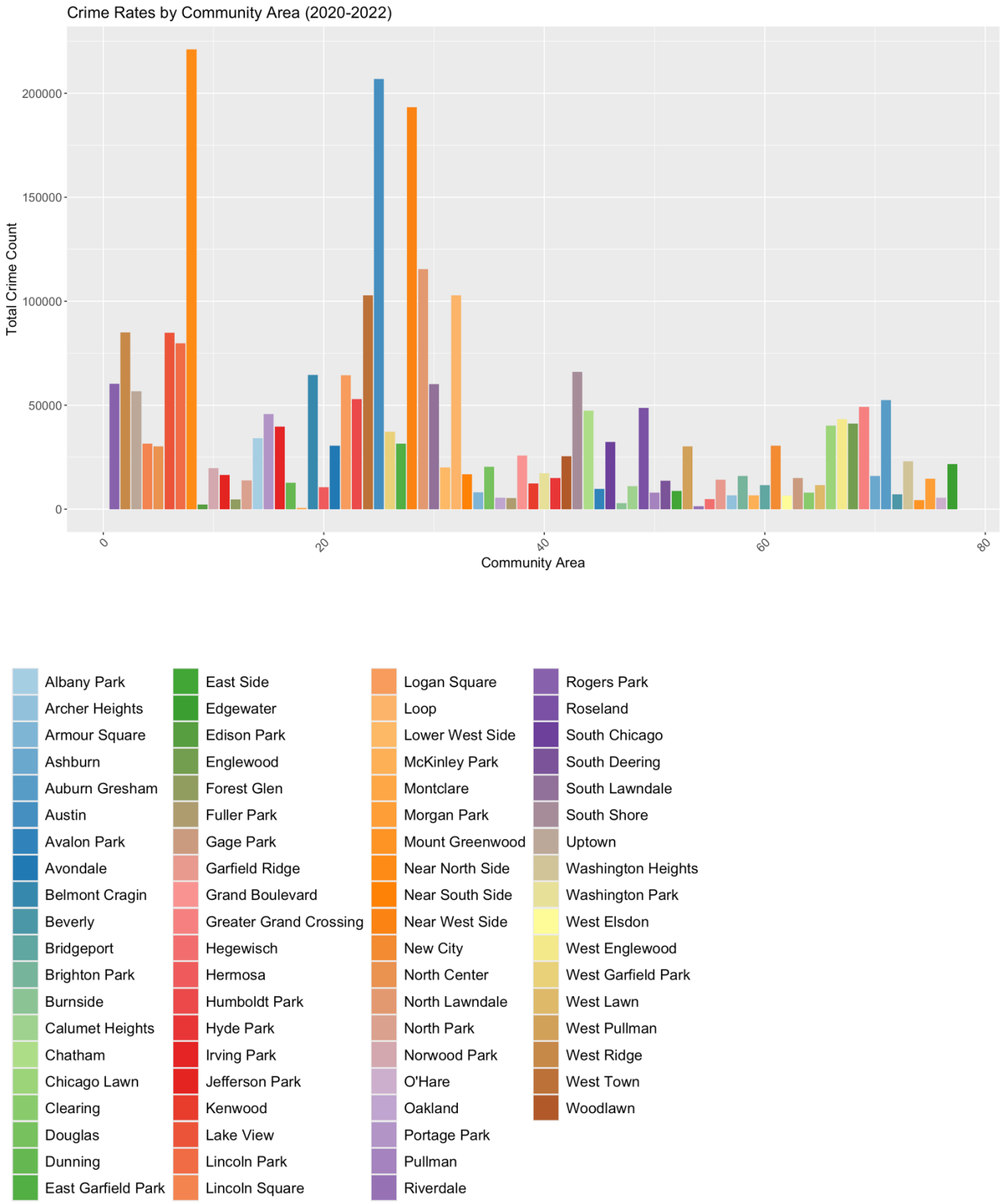

Crime Rates by Community Area (2020-2022)
Total Crime Count
200000
150000
100000
50000
0
0
20
40
60
80
Community Area
Albany Park
Archer Heights
Armour Square
Ashburn
Auburn Gresham
Austin
Avalon Park
Avondale
Belmont Cragin
Beverly
Bridgeport
Brighton Park
Burnside
Calumet Heights
Chatham
Chicago Lawn
Clearing
Douglas
Dunning
East Garfield Park
East Side
Edgewater
Edison Park
Englewood
Forest Glen
Fuller Park
Gage Park
Garfield Ridge
Grand Boulevard
Greater Grand Crossing
Hegewisch
Hermosa
Humboldt Park
Hyde Park
Irving Park
Jefferson Park
Kenwood
Lake View
Lincoln Park
Lincoln Square
Logan Square
Loop
Lower West Side
McKinley Park
Montclare
Morgan Park
Mount Greenwood
Near North Side
Near South Side
Near West Side
New City
North Center
North Lawndale
North Park
Norwood Park
O'Hare
Oakland
Portage Park
Pullman
Riverdale
Rogers Park
Roseland
South Chicago
South Deering
South Lawndale
South Shore
Uptown
Washington Heights
Washington Park
West Elsdon
West Englewood
West Garfield Park
West Lawn
West Pullman
West Ridge
West Town
Woodlawn

Figure 4 a

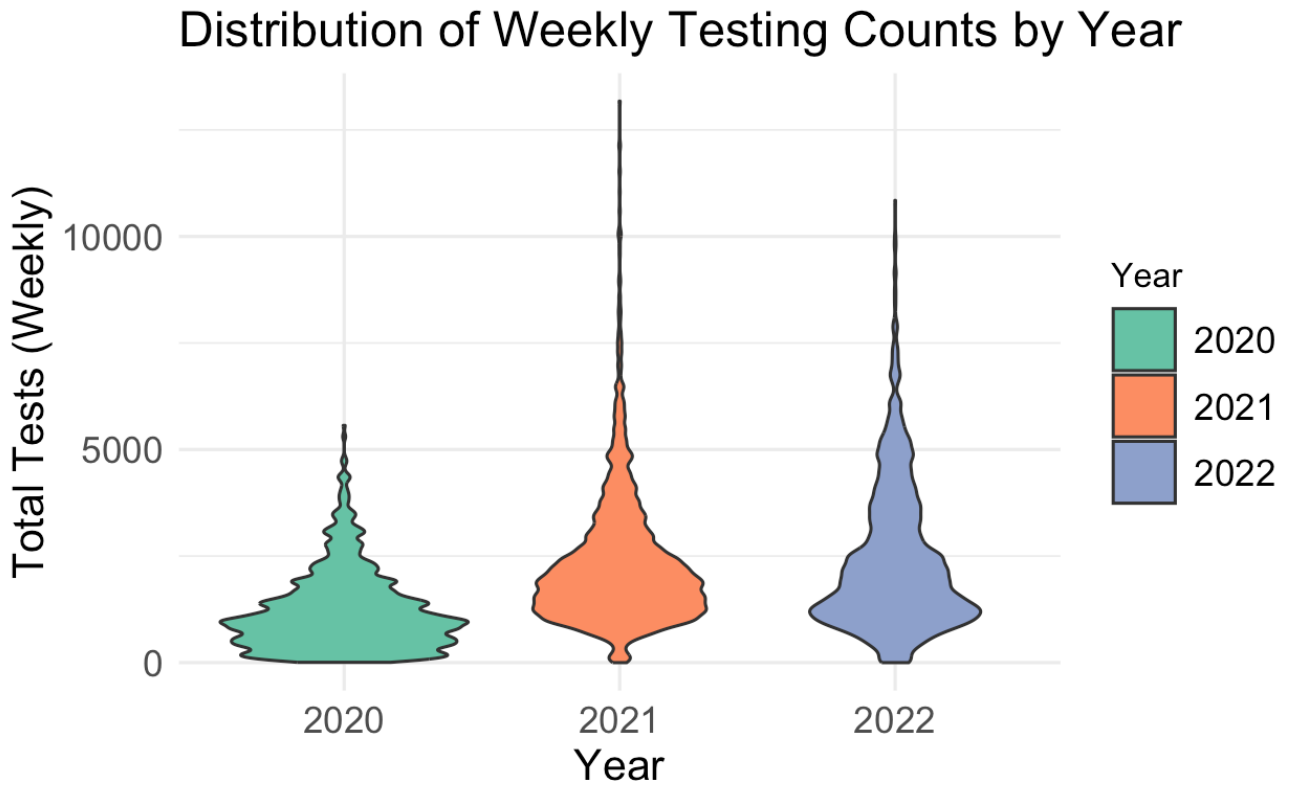


Figure 4 b

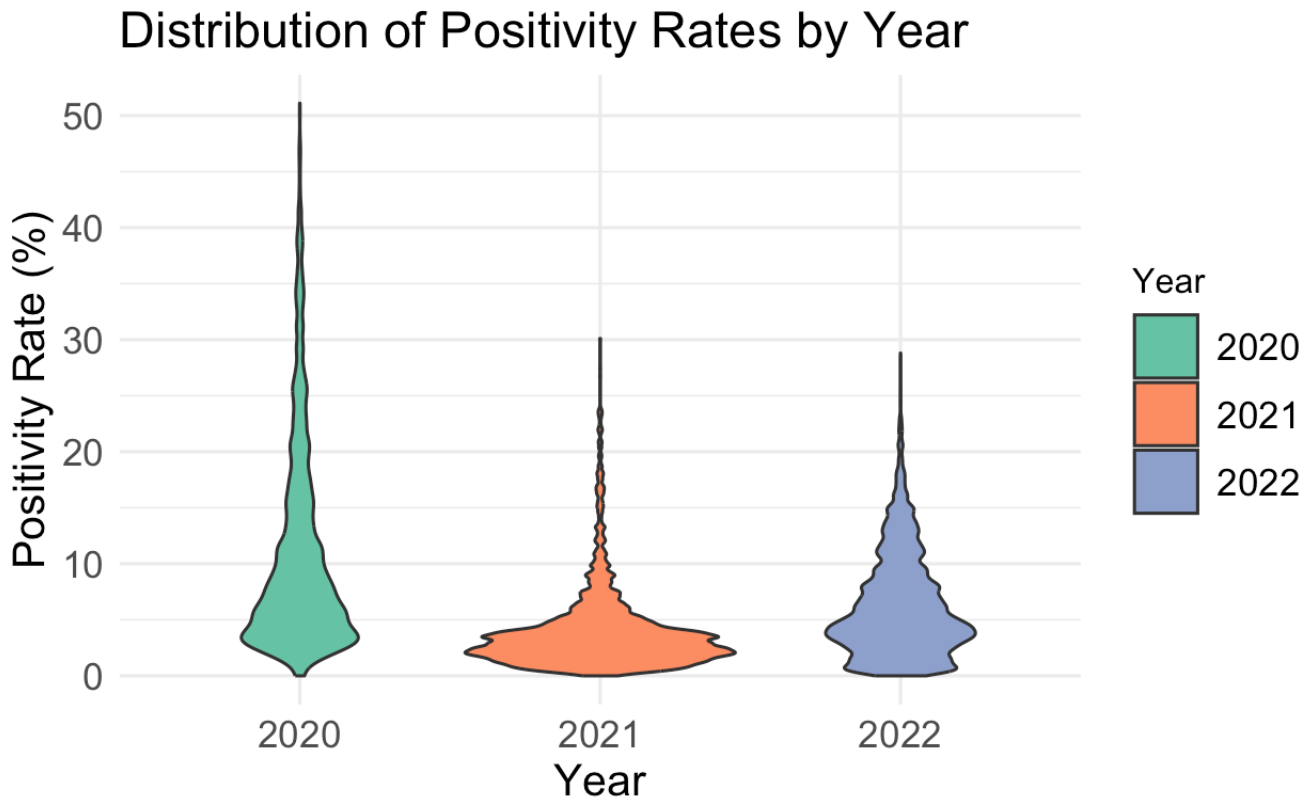


Figure 4 c

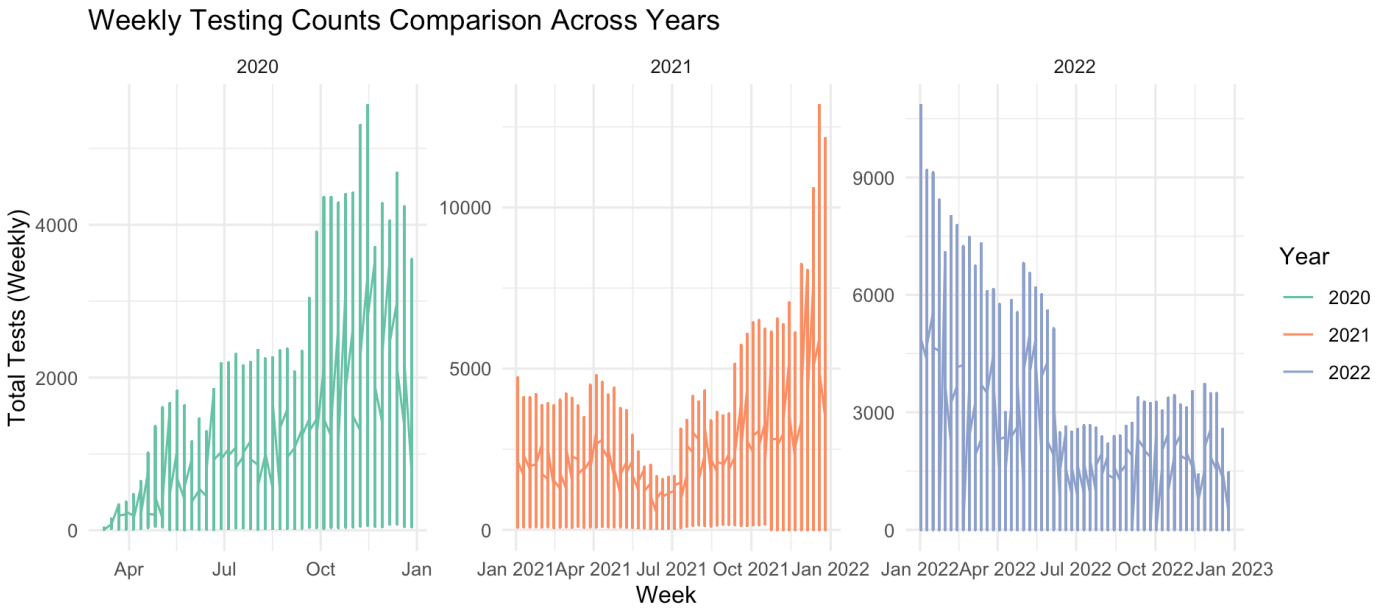


Figure 4 d

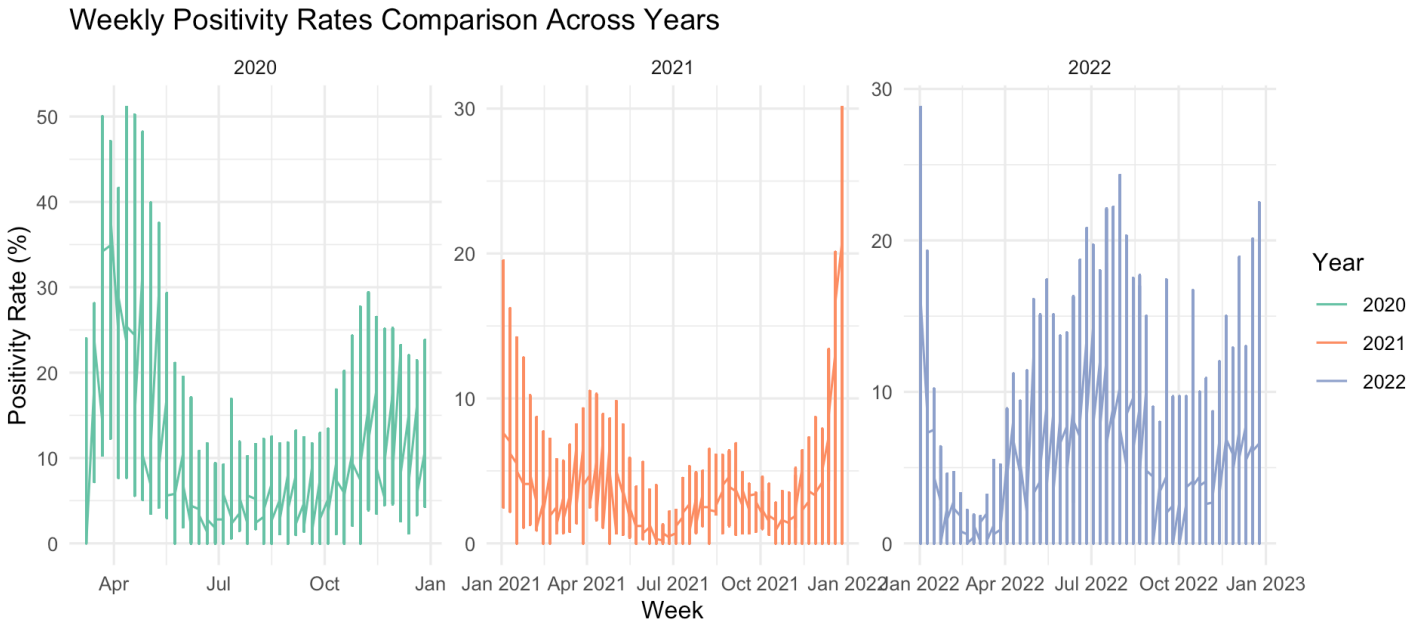


Figure 5

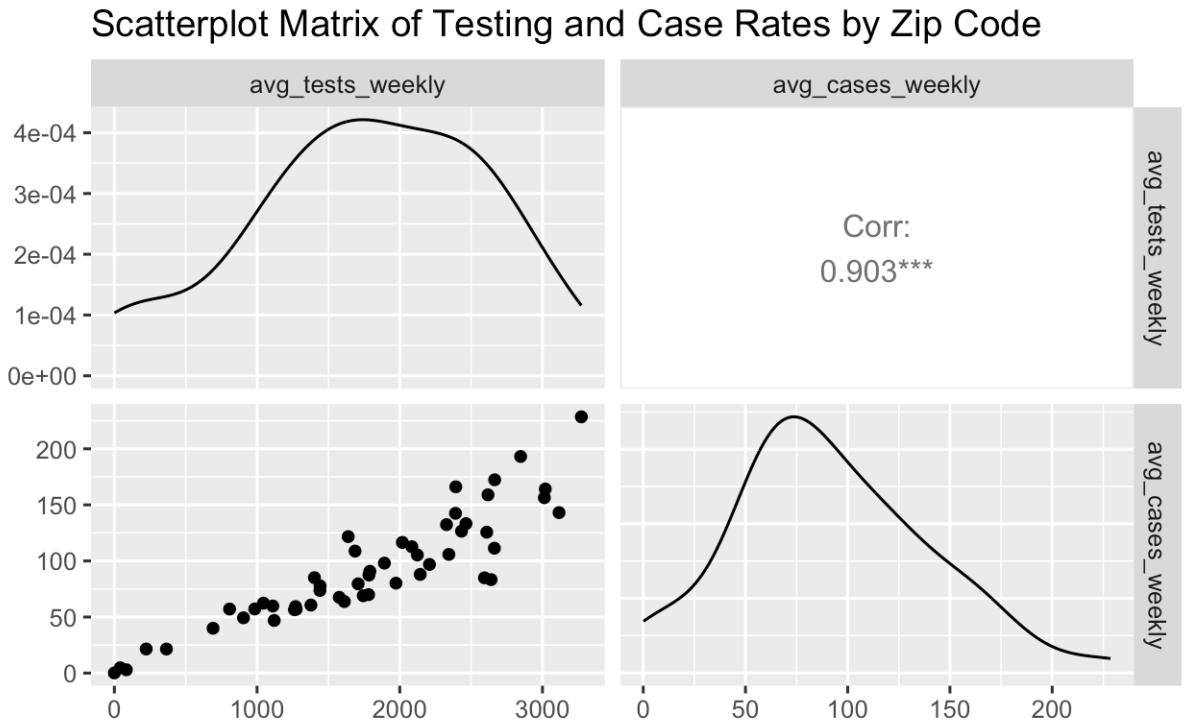


Figure 6a

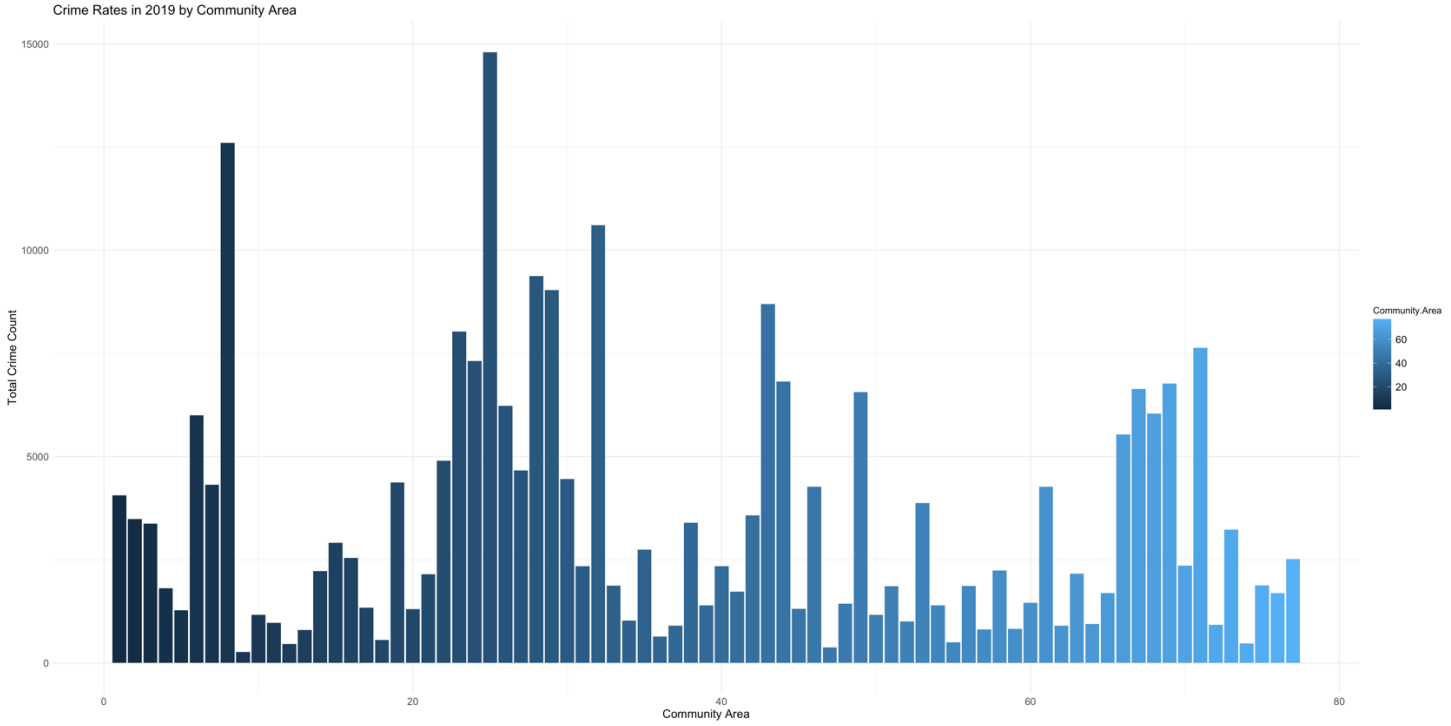


Figure 6b

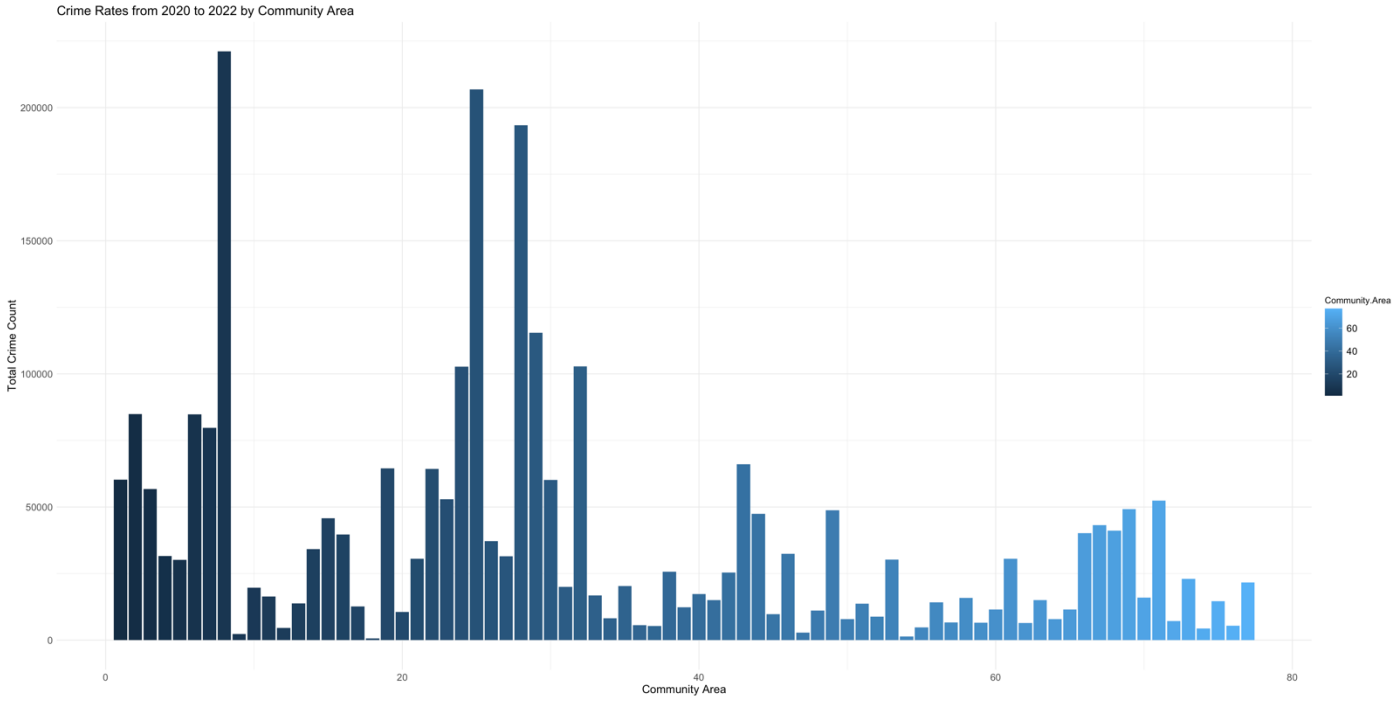


Figure 7a

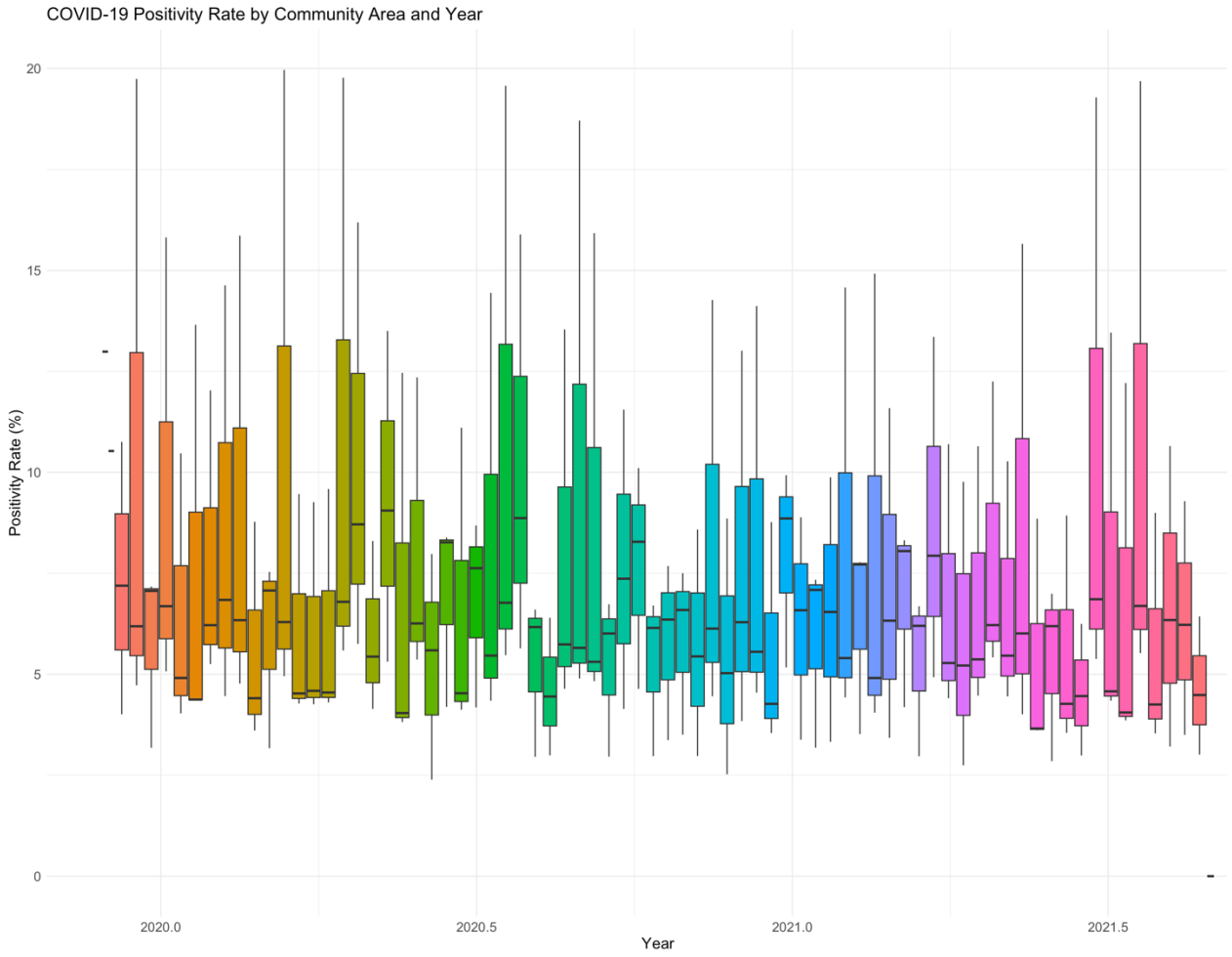
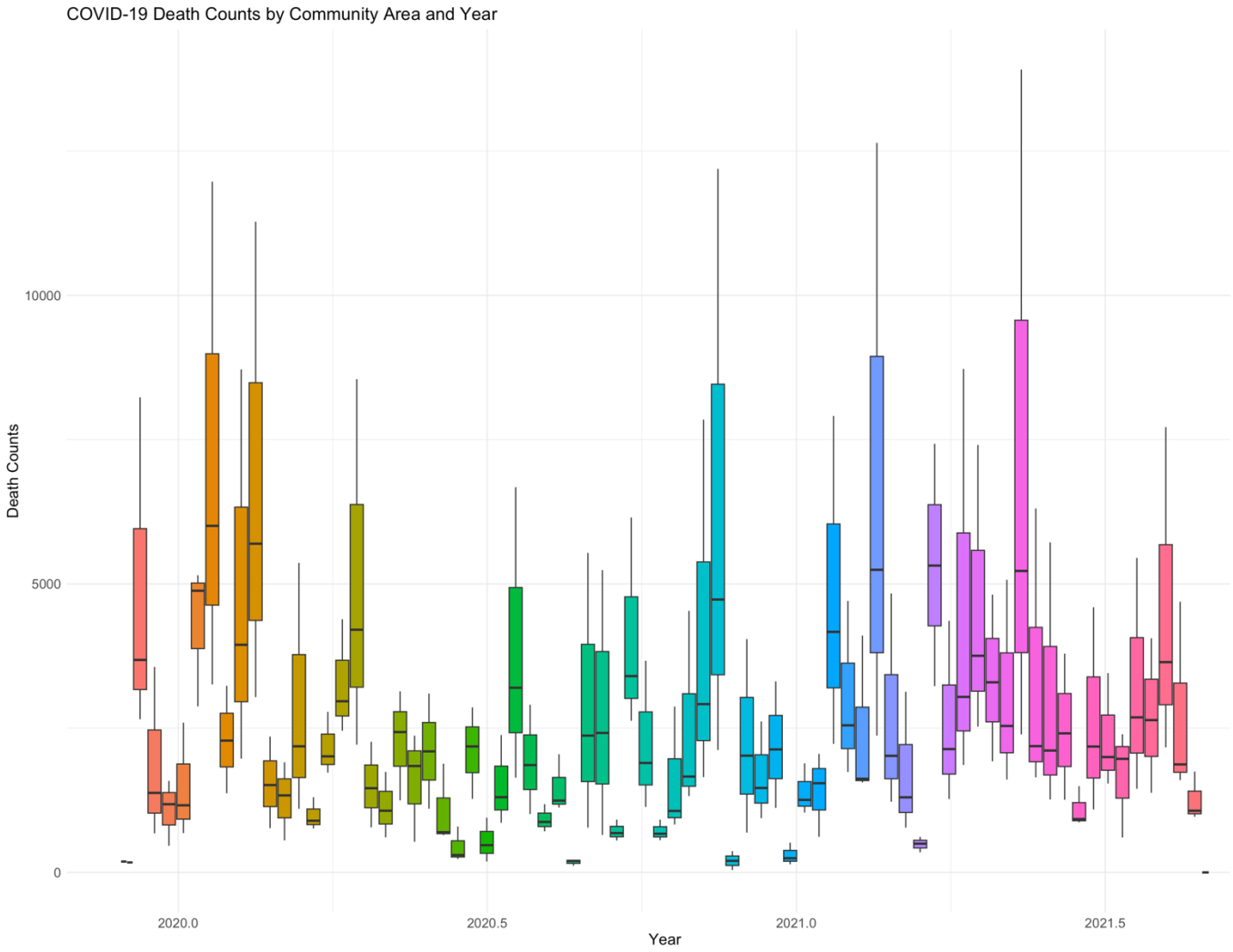


Figure 7b



Legend

CONCLUSION

This report examined the relationship between COVID-19 and crime rates in Chicago, exploring how the pandemic affected crime patterns, testing rates, and community impacts from 2020 to 2022. The findings indicate a complex interplay between public health and social dynamics: COVID-19 case rates fluctuated significantly, particularly during late 2021 to 2022. Community areas like Near North Side and Austin experienced higher crime rates, while others like Montclare and Riverdale showed lower rates. Positivity rates and deaths varied across communities, reflecting localized differences in pandemic impacts.

Potential future research could further investigate the socio-economic factors influencing these patterns, explore the effectiveness of specific public health interventions, and examine how demographic characteristics correlate with pandemic outcomes across different neighborhoods. Follow-up analyses could also focus on the long-term consequences of these trends on community safety and well-being.

The R codes for this analysis are available in the attached .qmd file.