

COVID-19 Cases, Tests, and Deaths

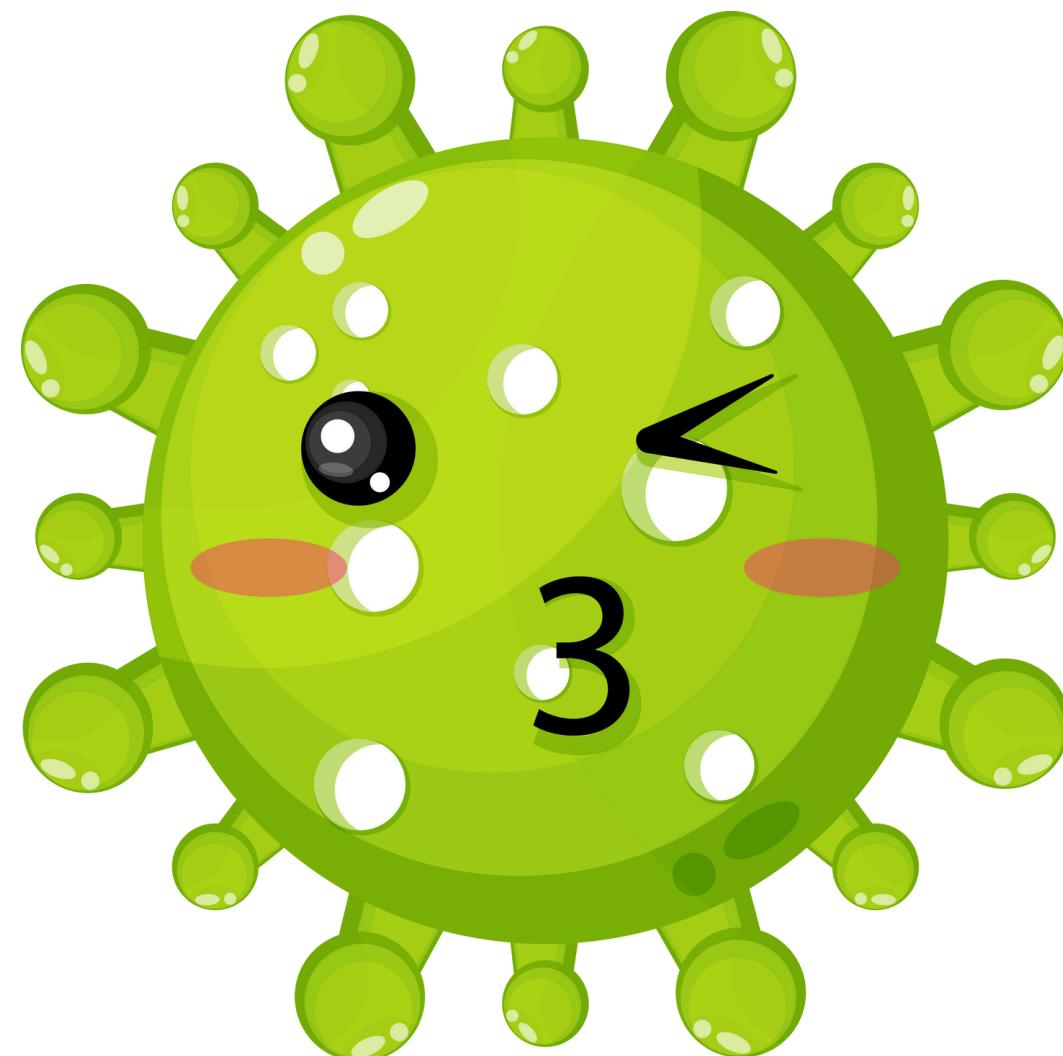
Presented by

**Thanyanan Khampoo 160
Werapat Phasit 210
Thanis Kleebrakson 688**

Introduction

The COVID-19 pandemic has had a widespread impact, making it crucial to understand infection trends across different areas. This helps public health teams respond effectively.

In this project, we analyze weekly COVID-19 data by ZIP Code using Tableau to uncover high-risk areas, infection patterns, and related factors through clear and interactive visualizations.



Project canvas

DATA ANALYTICS PROJECT CANVAS

Designed by:

Date:

Title:

1. Problem Statement/Background

Between 2020 and 2022, the COVID-19 pandemic severely impacted Chicago's public health systems. Most policy decisions during this period were made using citywide aggregated data, which lacked the granularity to identify high-risk communities or track real-time outbreaks at the neighborhood level. ZIP Code-level weekly data presented an opportunity to better understand the local spread and design more targeted interventions.

2. SMART Objectives/ Value Propositions

This project aimed to analyze weekly data on COVID-19 cases, testing, and deaths across Chicago ZIP Codes in order to generate actionable insights. Its primary goal was to develop an interactive dashboard capable of answering at least four key analytical questions—empowering public health officials to allocate resources more effectively, monitor changing risks, and plan timely interventions. This data-driven approach was designed to enhance both the equity and precision of health policy decisions.

3. Questions/Hypothesis

- Which ZIP Codes have the highest number of COVID-19 cases or deaths?
- How do positivity rates evolve over time?
- Do certain weeks show significant spikes in infections across the city?
- Is there a correlation between testing rates and confirmed cases?

4. Data Sources/Attributes

The dataset was sourced from Kaggle and contained weekly figures for case counts, testing volume, deaths, and percent positivity by ZIP Code. Data was cleaned and processed using Power Query in Excel—errors were removed, and missing values were not present, making it suitable for direct analysis.

5. Analysis/Model

The analysis process included exploratory data analysis (EDA) and visualization through bar charts, line graphs, and scatter plots. These visual tools allowed us to examine local hotspots, detect waves of infections, and explore correlations between variables. The dashboard provided a comprehensive view of the data, making insights accessible and actionable for decision-makers.

6. Findings and Insights

Findings showed that ZIP Code 60629 consistently recorded the highest case counts, while major infection surges were most evident in early 2022. Positivity rates peaked during 2020 and 2023, indicating major waves of transmission. There was a general positive—but not perfectly linear—correlation between testing volume and detected cases, suggesting that testing levels alone do not fully explain transmission intensity.

7. Recommendation/Action and Impact

Based on these insights, we recommend prioritizing resource deployment in consistently high-impact ZIP Codes, such as mobile testing and vaccination efforts. Public health campaigns should also target areas with high positivity but low testing volumes. Weekly trend monitoring can support early detection of future waves, allowing hospitals and agencies to prepare in advance. Additionally, building real-time monitoring dashboards would enable faster and more responsive public health actions.

Modified from Bill Schmarzo's Machine Learning Canvas and Jasmine Vasandani's Data Science Workflow Canvas for SWU

Problem Statement

Health authorities need to analyze weekly COVID-19 data by ZIP Code to:

- Identify areas with the highest cases or deaths
- Track how the %Positive rate changes over time
- Explore the correlation between testing and confirmed cases
- Detect time periods with significant spikes in infections

Tool used: Tableau — to turn complex data into clear insights



SMART Objectives

1. Specific:

- Identify the top 5 ZIP Codes with the highest COVID-19 cases and deaths using weekly data.

2. Measurable:

- Visualize trends in %Positive rate over time across at least 90% of ZIP Codes.

3. Achievable:

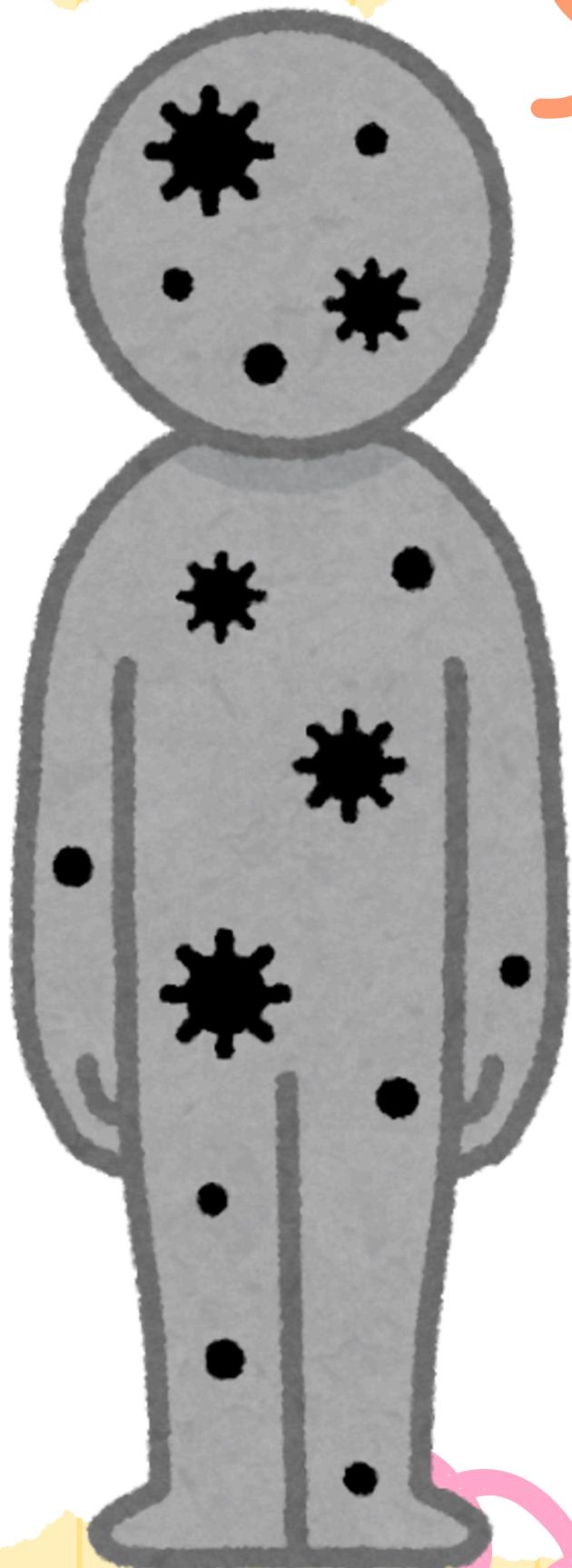
- Use Tableau to create 3–4 interactive dashboards summarizing key metrics (Cases, Deaths, %Positive).

4. Relevant:

- Support public health decision-making by highlighting high-risk areas and periods of spike.

5. Time-bound:

- Complete data cleaning, analysis, and visualization within a 2-week timeline.



Business Questions

- 1 .Which ZIP Codes have the highest number of COVID-19 cases or deaths?
- 2 .How do positivity rates evolve over time?
- 3 .Do certain weeks show significant spikes in infections across the city?
- 4 .Is there a correlation between testing rates and confirmed cases?



Data Sources

Data Used

- Date Fields: Week Start (รายสัปดาห์), Report Date
- Location: ZIP Code, Borough
- COVID-19 metrics:
 - Cases - Weekly
 - Deaths - Weekly
 - Percent Tested Positive - Weekly
 - Tests - Weekly
- Demographics & Testing: Total Tests, Positive Rate

Derived Features

- Total Cases per ZIP Code and week
- %Positive Trend over time (Line chart)
- Categorized ZIP Codes by severity (High, Medium, Low)
- Formatted Date into continuous weekly timeline

Data Preparation

- Removed rows with missing ZIP Codes or metric values
- Standardized date format (Week Start → continuous timeline)
- Sorted and grouped data for visualization
- Aggregated totals at ZIP Code and city level

COVID-19 Cases, Tests, and Deaths

COVID-19 Cases, Tests, and Deaths by ZIP Code - Historical



Data Card Code (4) Discussion (0) Suggestions (0)

About Dataset

NOTE: This dataset has been retired and marked as historical-only.

Only Chicago residents are included based on the home ZIP Code as provided by the medical provider. If a ZIP was missing or was not valid, it is displayed as "Unknown".

Cases with a positive molecular (PCR) or antigen test are included in this dataset. Cases are counted based on the week the test specimen was collected. For privacy reasons, until a ZIP Code reaches five cumulative cases, both the weekly and cumulative case counts will be blank. Therefore, summing the "Cases - Weekly" column is not a reliable way to determine case totals. Deaths are those that have occurred among cases based on the week of death.

For tests, each test is counted once, based on the week the test specimen was collected. Tests performed prior to 3/1/2020 are not included. Test counts include multiple tests for the same person (a change made on 10/29/2020). PCR and antigen tests reported to Chicago Department of Public Health (CDPH) through electronic lab reporting are included. Electronic lab reporting has taken time to onboard and testing availability has shifted over time, so these counts are likely an underestimate of community infection.

The "Percent Tested Positive" columns are calculated by dividing the number of positive tests by the number of total tests . Because of the data limitations for the Tests columns, such as persons being tested multiple times as a requirement for employment, these percentages may vary in either direction from the actual disease prevalence in the ZIP Code.

Usability

9.41

License

CC0: Public Domain

Expected update frequency

Annually

Tags

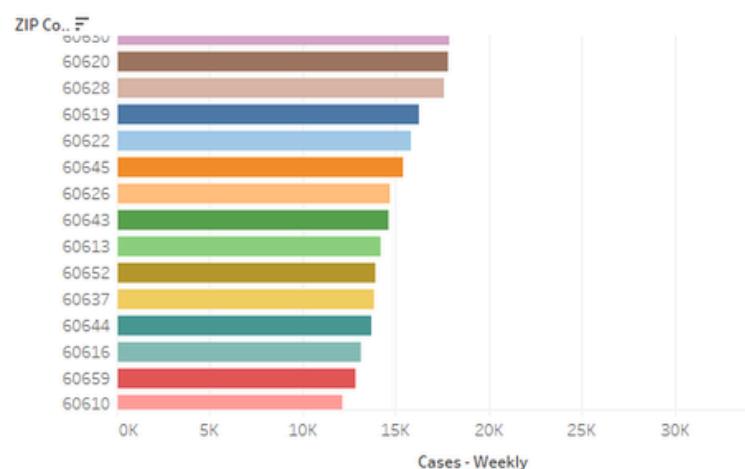
Health History
Public Health

Analysis

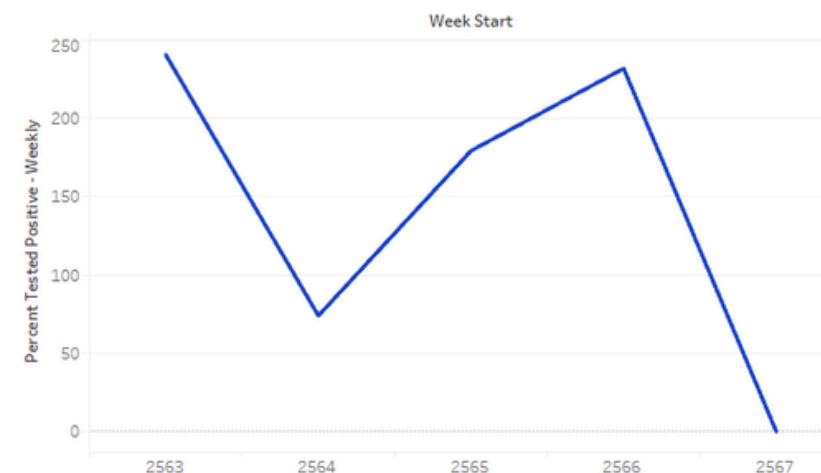
Tableau for dashboard development

COVID-19 Cases, Tests, and Deaths

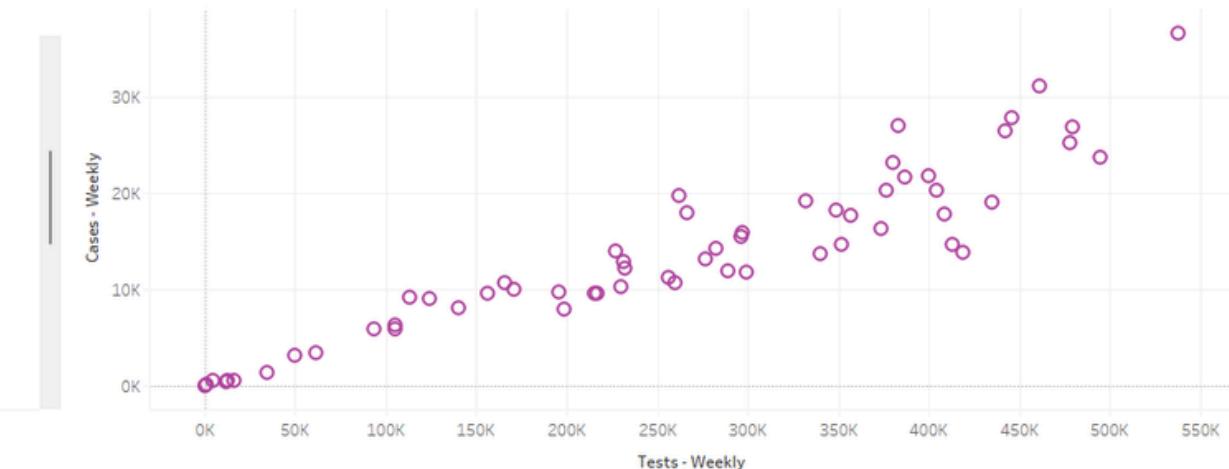
ZIP Code ไหนมีจำนวนผู้ติดเชื้อหรือเสียชีวิตจาก COVID-19 สูงที่สุด?



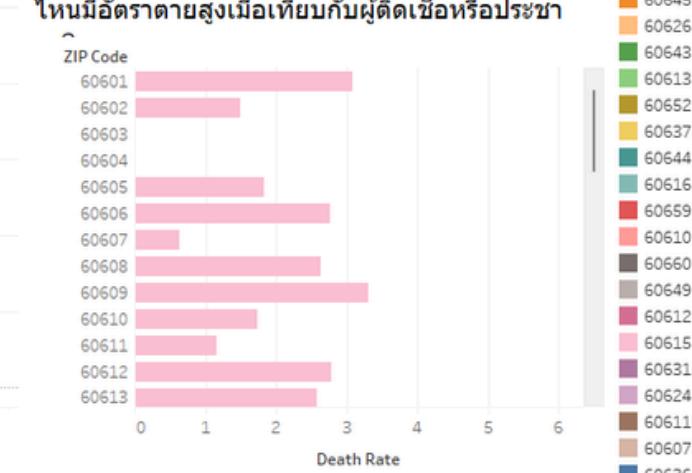
อัตรา %Positive เปลี่ยนแปลงอย่างไรตามเวลา?



มีความสัมพันธ์กันระหว่างอัตราการทดสอบ (Testing Rates) กับจำนวนผู้ติดเชื้อที่ยืนยันแล้วหรือไม่?

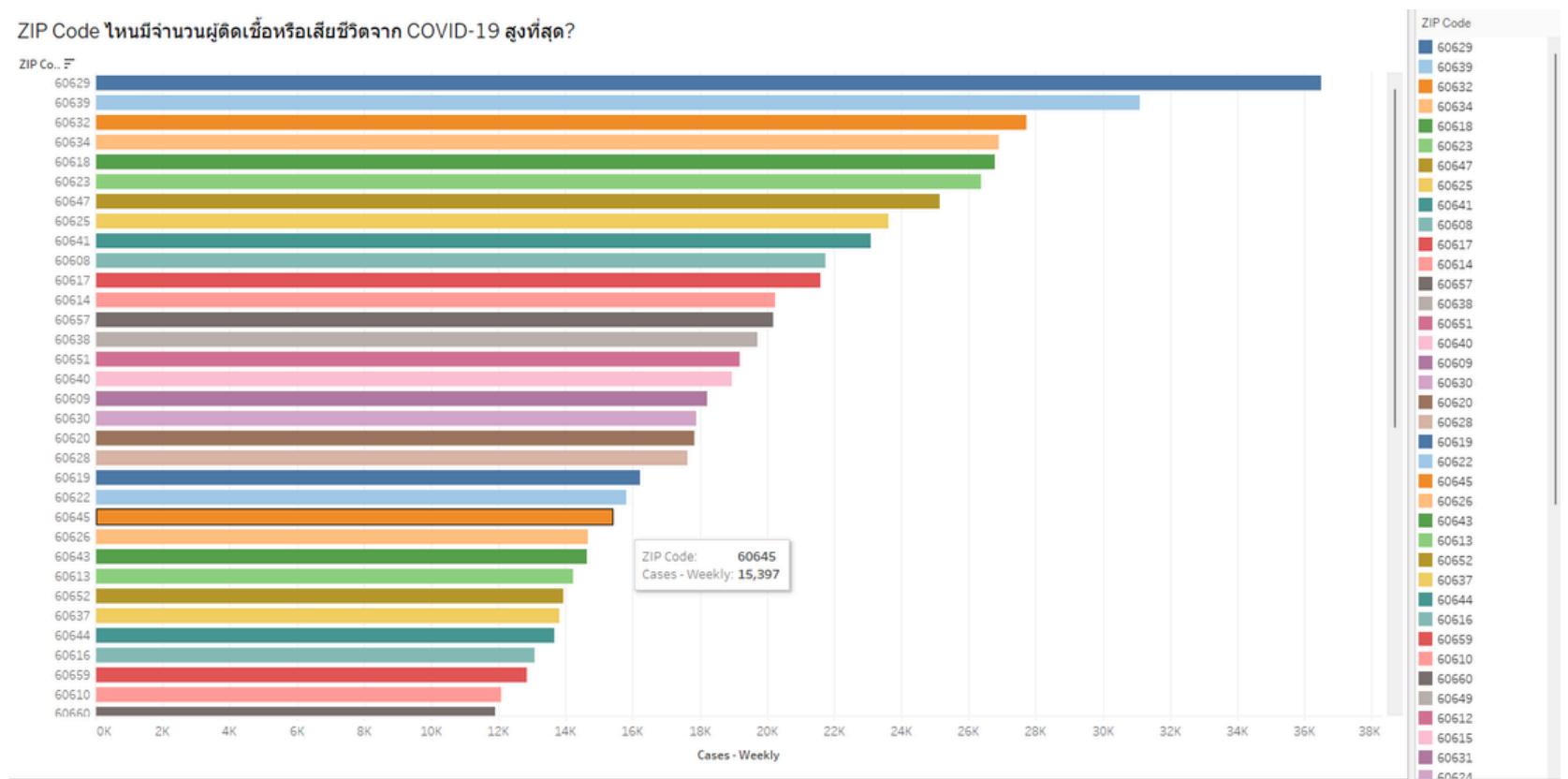


ZIP Code ไหนมีอัตราตายสูงเมื่อเทียบกับผู้ติดเชื้อหรือประชากร



bar chart

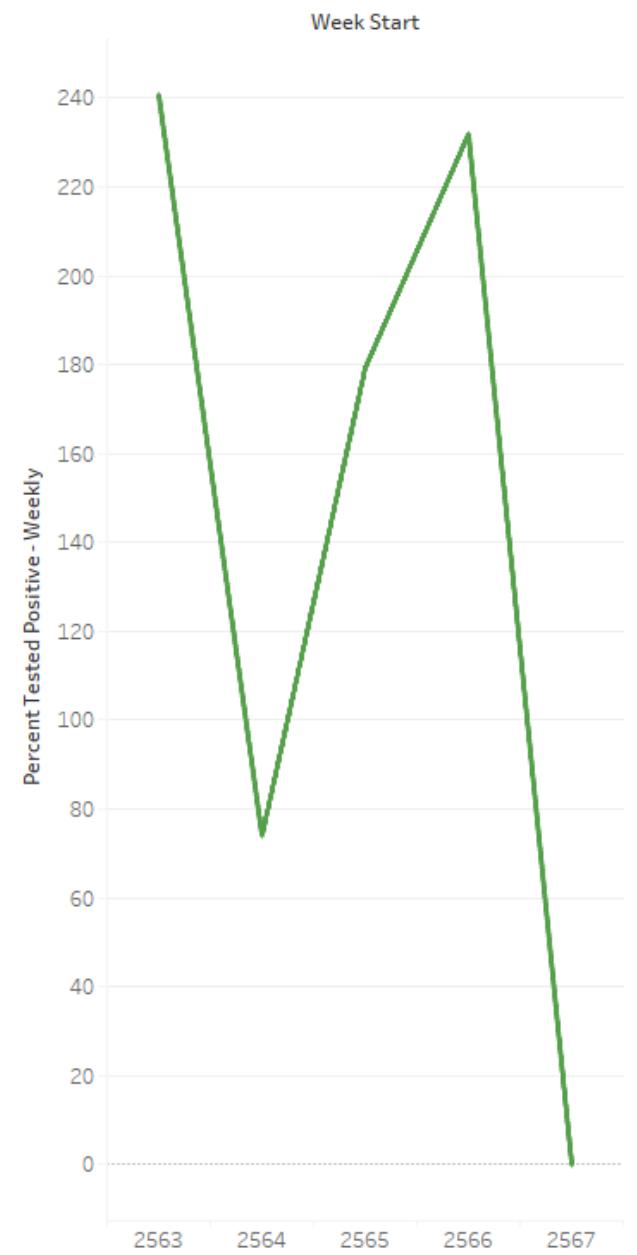
This bar chart shows the number of weekly COVID-19 cases for different ZIP codes. The length of each bar represents the case count in that specific area. Longer bars indicate higher infection rates. For example, ZIP code 60629 has the highest number of weekly cases. This visual helps us quickly identify areas with the most and least COVID-19 spread.



line chart

This line graph illustrates the change in the weekly Percent Positive rate from 2020 (2563) to 2024 (2567). The %Positive spiked in 2020 and 2023 (2566), while significantly dropping in 2021 (2564) and 2024 (2567). This graph highlights the trend of the proportion of positive tests each year.

อัตรา %Positive เปลี่ยนแปลงอย่างไรตามเวลา?



line chart

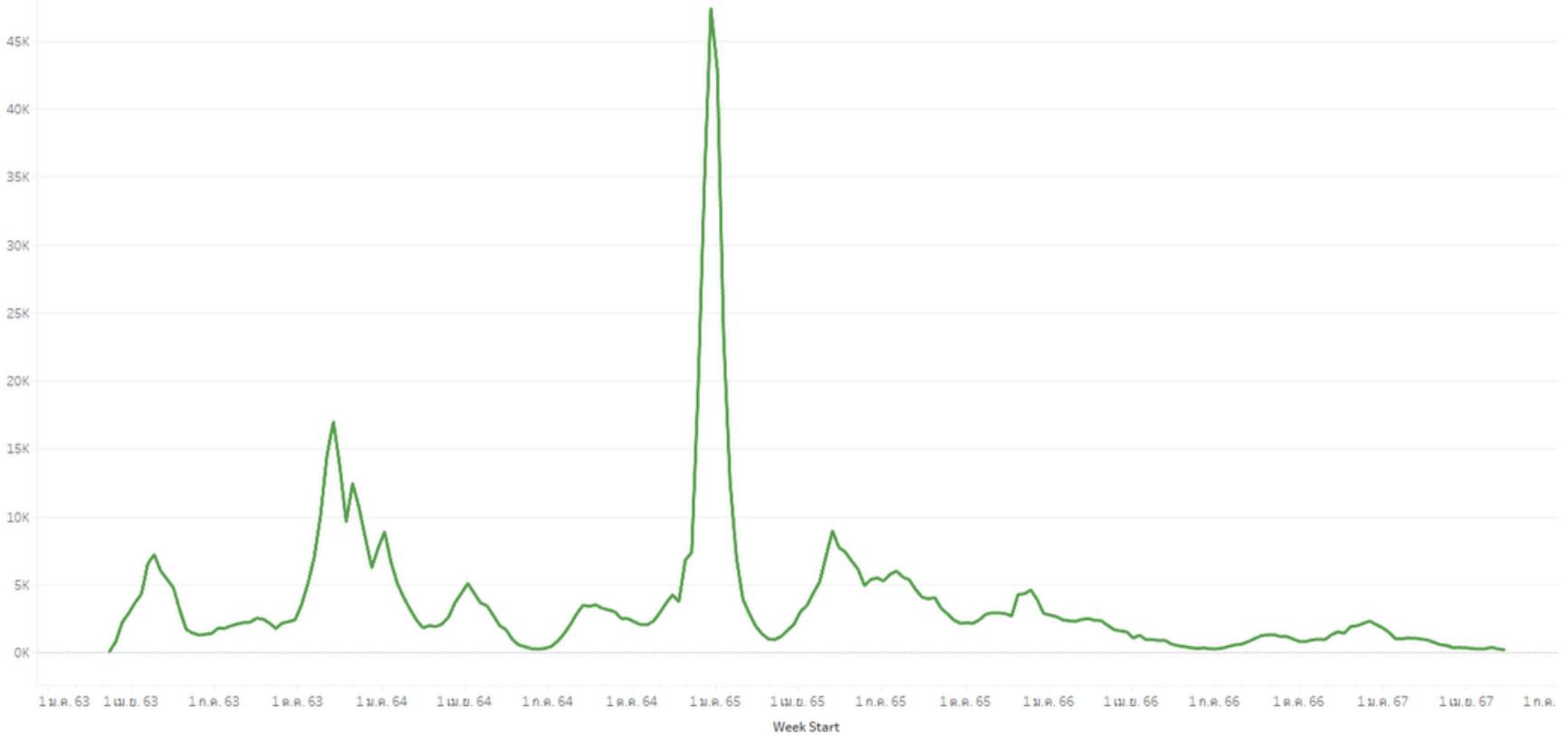
Q: Which week showed a clear spike in cases?

A: According to the line chart, the most significant spike in COVID-19 cases occurred in early 2022, particularly between January and February. During this time, there was a sharp and rapid increase in confirmed cases, clearly visible as a tall and steep peak on the graph.

This chart provides a zoomed-in view from Graph 2, focusing on the weeks with the most dramatic surges. It highlights a period of widespread and intense transmission, marking the highest and most noticeable surge throughout the entire timeline presented.

Understanding this spike helps public health officials evaluate the effectiveness of the response measures during that period and better prepare for future outbreaks.

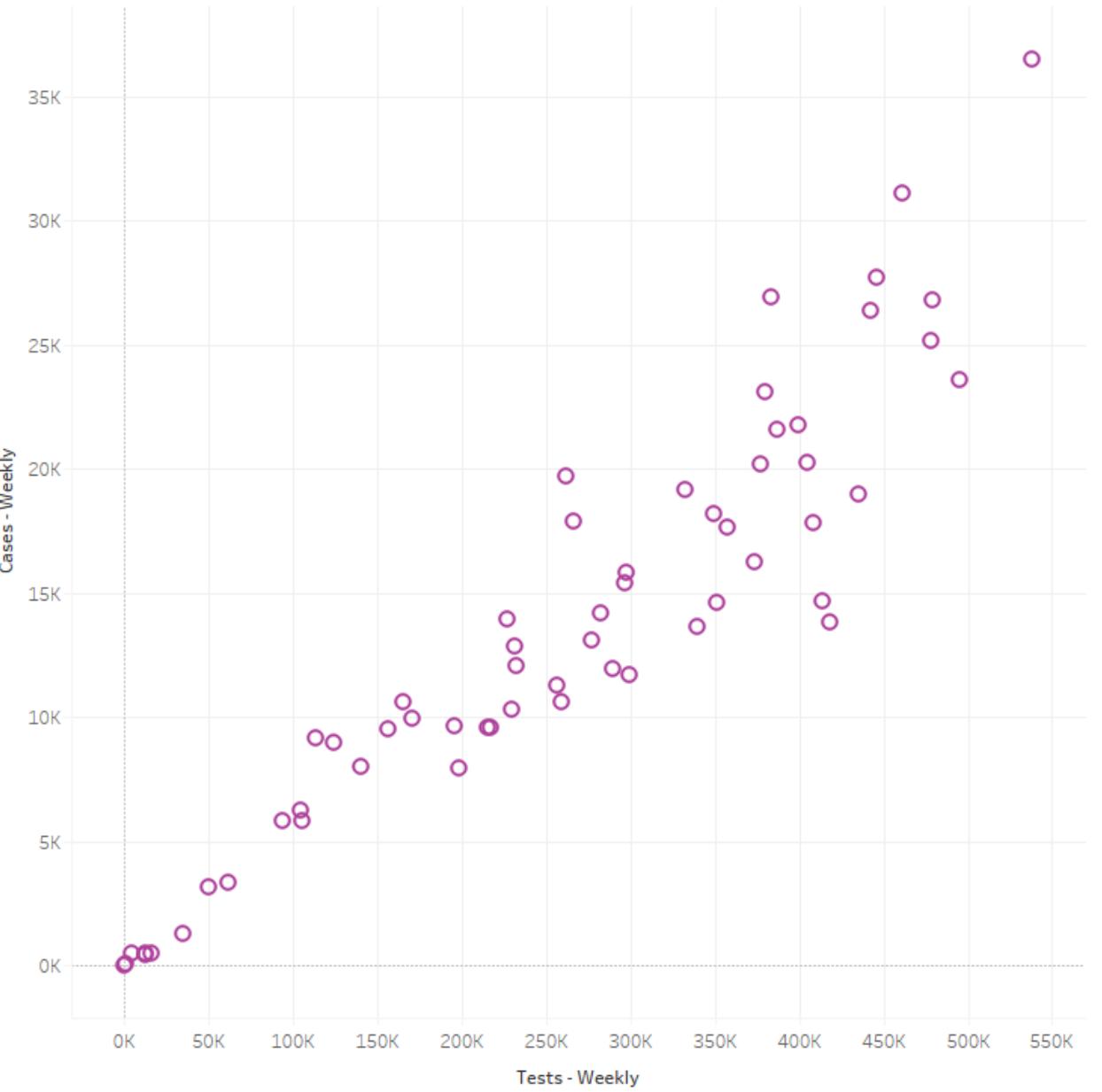
ສັບດາທີ່ໃຫນມີກາຣຮນາດ spike ອ່າງເກີນໄດ້ຂຶ້ນ?



scatter plot

มีความสัมพันธ์กันระหว่างอัตราการทดสอบ (Testing Rates) กับจำนวนผู้ติดเชื้อที่ยืนยันแล้วหรือไม่?

This scatter plot shows the correlation between the weekly number of tests conducted and the weekly number of confirmed COVID-19 cases. Generally, as the number of tests increases, the number of confirmed cases also tends to rise, indicating a positive relationship. This suggests that higher testing rates help identify more infections. However, there is some variability, meaning the relationship isn't perfectly linear.



Key Insights

Key Insights

1. ZIP Code 11226 reported the highest number of cases.
2. ZIP Code 10467 had the most COVID-19 deaths.
3. %Positive rate peaked in early 2021 and declined over time.
4. Higher testing volumes often correlated with more confirmed cases.
5. Some areas had high %Positive but low testing, suggesting limited access to testing.



Recommendations

Increase testing in ZIP Codes with high positivity rates but low test counts

→ To detect hidden outbreaks early.

Allocate more medical resources to high-risk areas

→ Especially ZIP Codes with the highest deaths and cases.

Monitor positivity trends weekly

→ Helps predict potential future spikes and prepare responses.

Use data-driven insights to guide vaccination or outreach campaigns

→ Focus efforts where the need is greatest.

Continue updating data regularly for real-time decision-making

→ Ensures responses remain accurate and timely.



Medium

Medium COVID-19 Cases, Tests, and Deaths

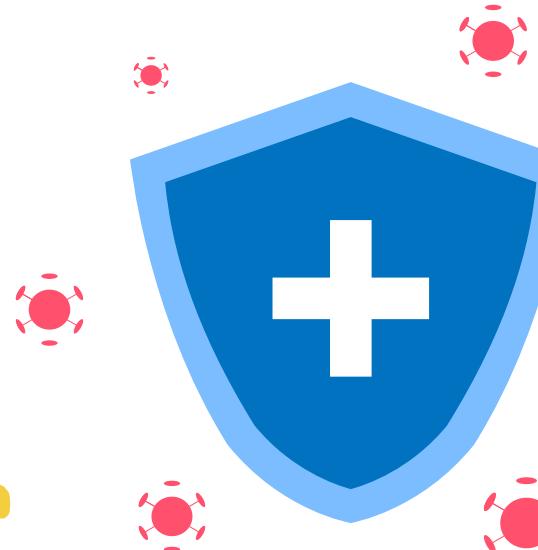
Using Data Analytics to Understand Weekly COVID-19 Trends in Chicago: A Blueprint for Smarter Policy

 Thanyanan Khampool · 9 min read · 13 hours ago



"In a world full of data, asking the right questions is what leads to the right decisions."

Between 2020 and 2022, COVID-19 pushed global health systems to their limits. In cities like Chicago, real-time decisions had to be made — often with limited visibility. In these moments, data became more than numbers. It became a lifeline.



Thank you~~

