IS525 Final Project Report

Junjie Lai Youdong Lu Richie Li Emily Chang

Introduction

Chicago, as one of the largest cities in the U.S., faces significant crime-related challenges that affect public safety and community trust. This project utilizes the Chicago Crime dataset (2020–2024) to provide actionable insights through interactive dashboards. Designed for the Chicago Police Department and city residents, these dashboards offer district-level, temporal, and seasonal crime analysis. They not only assist law enforcement in resource allocation and crime prevention but also serve as a tool for public information, helping residents better understand their city's safety landscape.

Dataset Overview

The Chicago Crime dataset from the City of Chicago Data Portal provides detailed information on crimes reported in Chicago, Illinois, from 2001 to the present. With over two decades of records, this dataset is a valuable resource for analyzing crime patterns, understanding community safety, and examining the effectiveness of law enforcement in various neighborhoods.The final focus of this project is on data for 2020-2024

Variables

- **Incident Identification:** ID, Case Number
- **Time Attributes:** Date, Year, Updated On
- **Location Attributes:** Block, Location Description, Community Area, Ward, Beat, District, Location (Latitude and Longitude)
- **Crime Classification:** Primary Type, Description, IUCR, FBI Code
- **Crime Details:** Arrest (True/False), Domestic (True/False)
- **Geospatial Coordinates:** Latitude, Longitude, X Coordinate, Y Coordinate

Dashboard Designs and Analysis

District-Level Insights (Junjie Lai)

1. Purpose of the Dashboard

The primary goal of this dashboard is to assist the Chicago Police Department in gaining a comprehensive understanding of crime patterns across different districts. By providing interactive visualizations and key insights, the dashboard enables police officers and policymakers to analyze the crime situation in each district,

identify high-risk areas, and implement targeted measures for deploying law enforcement resources or formulating strategic policies. The tool facilitates data-driven decision-making, allowing stakeholders to address crime issues more effectively.

2. Data Preparation

2.1 Data Integration

The preparation phase began with integrating crime and geographical data. Crime data was cleaned to include key fields such as the date, primary type, district , and IUCR. Geographical data was structured to provide district boundaries in WKT format, district numbers, and district labels. This integration ensured that both datasets were compatible and provided meaningful spatial and statistical insights.

2.2 Data Join

An inner join was performed between the crime data (District) and the geographical data (DIST_NUM). This step allowed for the seamless merging of crime statistics with their corresponding geographical context, enabling spatial analysis and visualization.

2.3 Calculated Fields

Several calculated fields were created to enhance the analytical capabilities of the dashboard.

- **Arrest rate:**
  ```
  SUM(IF [Arrest] THEN 1 ELSE 0 END) / COUNT([Case Number])
  ```
- **Time periods:**
  ```
  IF DATEPART('hour', [Date]) >= 5 AND DATEPART('hour',
  [Date]) < 12 THEN "Morning"
  ELSEIF DATEPART('hour', [Date]) >= 12 AND
  DATEPART('hour', [Date]) < 17 THEN "Afternoon"
  ELSEIF DATEPART('hour', [Date]) >= 17 AND
  DATEPART('hour', [Date]) < 21 THEN "Evening"
  ELSE "Night"
  END
  ```
- **Felony:**
  ```
  IF [IUCR] IN ("0110", "0130", "0260", "0310", "0410",
  "0610", "0625", "0650", "0910", "1010", "1025", "1813",
  "1822", "1910", "1920")
  THEN "Felony" ELSE "Non-Felony" END
  ```
- **Felony rate:**
  ```
  SUM(IF [Felony] = "Felony" THEN 1 ELSE 0 END) /
  COUNT([Case Number])
  ```

2.4. **Parameter for View Control**:

A parameter named “View” was designed to toggle between viewing overall crime data and district-specific data. This parameter was paired with a calculated field District:

```
CASE [View]
WHEN "Overall" THEN "Overall"
WHEN "By District" THEN STR([Dist Num])
END
```

This setup provided users with flexibility in choosing their desired perspective, ensuring the dashboard could cater to various analytical needs.

3. Dashboard Components

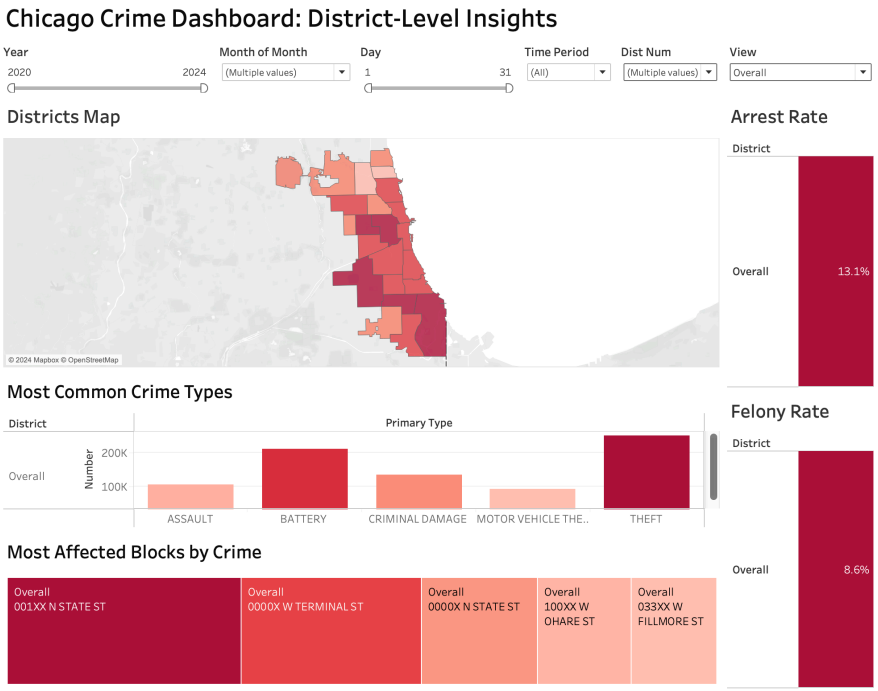
The dashboard includes five visualizations with interactive filters for year, month, day, time period, and district to dynamically explore data. A gradient red color scheme represents crime intensity, emphasizing high-crime areas.

- **District Map:** The district map highlights crime distribution across Chicago’s districts, using polygons to display boundaries and red gradients for intensity. It serves as a filter, allowing users to adjust other visualizations by selecting a district.
- **Top Crime Types Bar Chart:** This bar chart displays the top five crime types in each district based on incident counts. It updates dynamically with selected filters to provide a clear comparison.
- **Top Blocks Treemap:** The treemap identifies the top five blocks with the highest crime incidents in each district. Block sizes are proportional to crime counts, highlighting localized hotspots.
- **Arrest Rate Table:** The table shows arrest rates by district as percentages, providing a measure of law enforcement effectiveness. Interactive filters allow users to refine analysis for specific districts or timeframes.
- **Felony Rate Table:** This table displays felony rates by district, highlighting the severity of crimes. It supports interactive filtering for customized insights into crime patterns.

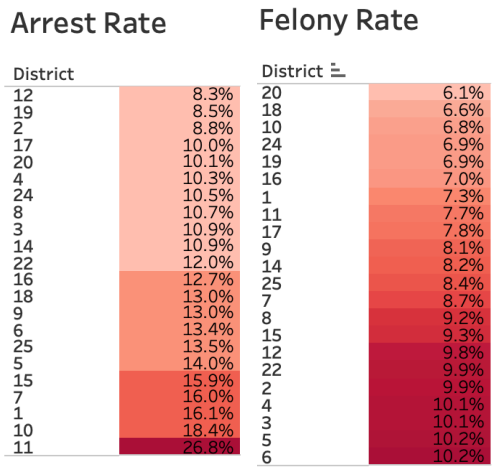
4. Findings

4.1 Overall Review

We first analyzed the overall performance of Chicago’s 25 districts. As shown in the visualization, the southern areas of Chicago (Districts 4, 6, 8) have the highest number of reported incidents, followed by the central areas, including Districts 1, 2, 3, 7, 9, and 10. In contrast, the northern districts (16, 17, 20, 24) demonstrate relatively better public safety. The top five crime types across the city are theft, battery, criminal assault, and motor vehicle theft. The city’s overall arrest rate is 13.1%, while the felony rate is 8.6%.



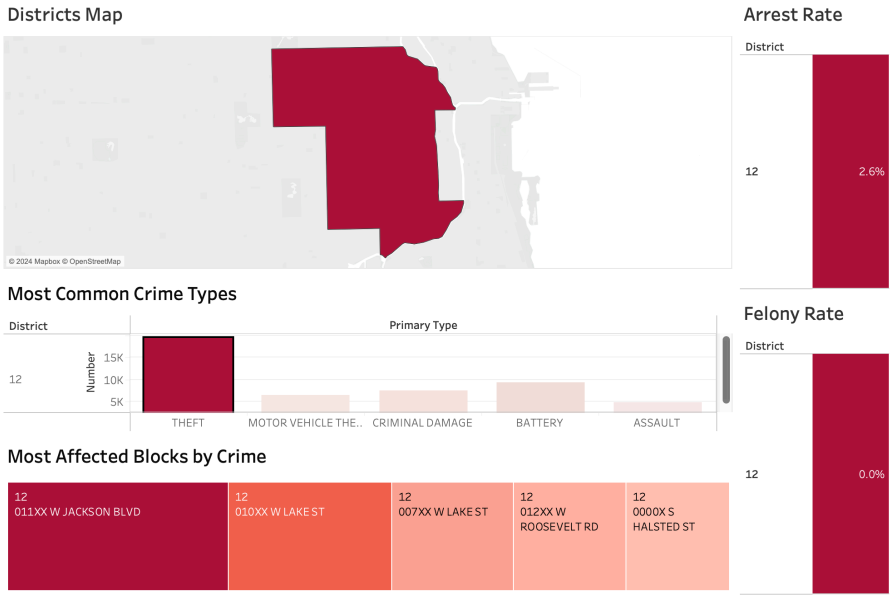
When focusing on district-level data, the findings reveal that District 12 has the lowest arrest rate at 8.3%, while District 11 has the highest at 26.8%. Regarding felony rates, District 20 has the lowest at 6.1%, and District 6 has the highest at 10.2%. Notably, District 12 shows both a low arrest rate and a high felony rate, indicating poor public safety in this area. Conversely, District 10 stands out as a safer district, with a relatively high arrest rate and a low felony rate.



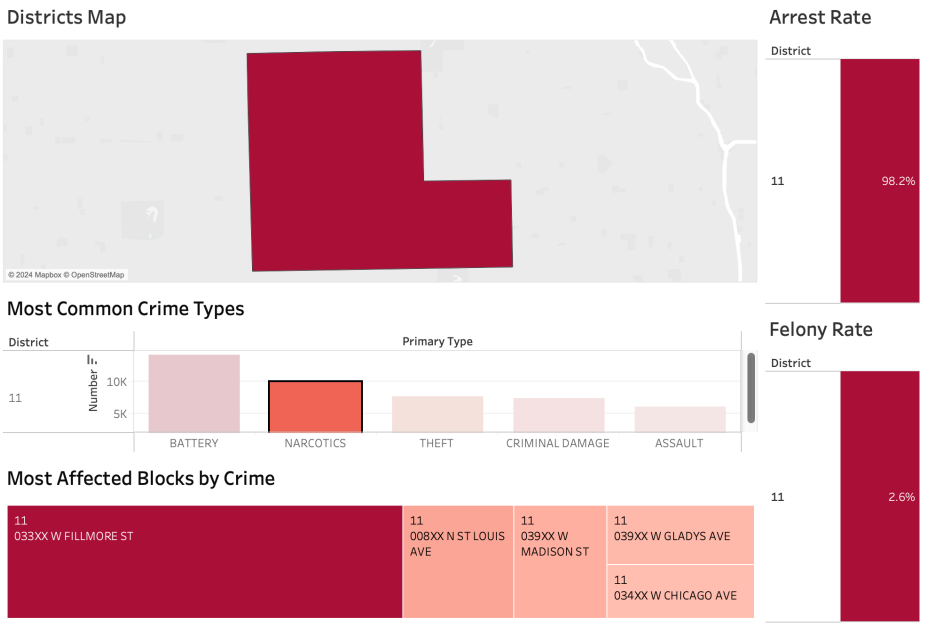
4.2 Specific District Analysis

We conducted a deeper analysis of Districts 11 and 12, representing the highest and lowest arrest rates. In District 12, theft dominates as the most frequent crime, with 19,404 reported incidents—far exceeding the other top four crime types. The arrest rate for theft in District 12 is only 2.6%, which helps explain the district's overall low arrest rate. This suggests that most crimes in District 12 involve petty thefts, which are challenging to solve. The police in District 12 could consider increasing patrols in

the blocks with the highest theft incidents, as highlighted by the dashboard, or promoting anti-theft systems in these communities to prevent further incidents.



In District 11, the crime composition is quite distinct, with battery and narcotics offenses as the leading crime types. Narcotics offenses have a high arrest rate, likely due to targeted police efforts, such as undercover operations and community informants. However, drug-related crimes still remain a major issue. Among the top five citywide crime blocks, 033XX W FILLMORE ST from District 11 stands out for narcotics activities. Authorities should intensify enforcement and consider large-scale crackdowns to address the problem.



Year-Level Insights (Youdong Lu)

1. Purpose of the Dashboard

Chicago is a vibrant and culturally rich city, but it also faces significant challenges related to crime. The problem we aim to address is understanding the distribution and patterns of criminal activity across the city to provide actionable insights for improving public safety. Our intended audience includes the Chicago Police Department, as well as Chicago residents and travelers. The data for this research was sourced from the Chicago Data Portal, which contains detailed records of reported crimes from 2001 to the present. By visualizing this information on a map of Chicago, we aim to assist law enforcement in making informed decisions to reduce crime and enhance safety, while also helping residents and visitors stay vigilant and avoid potential dangers.

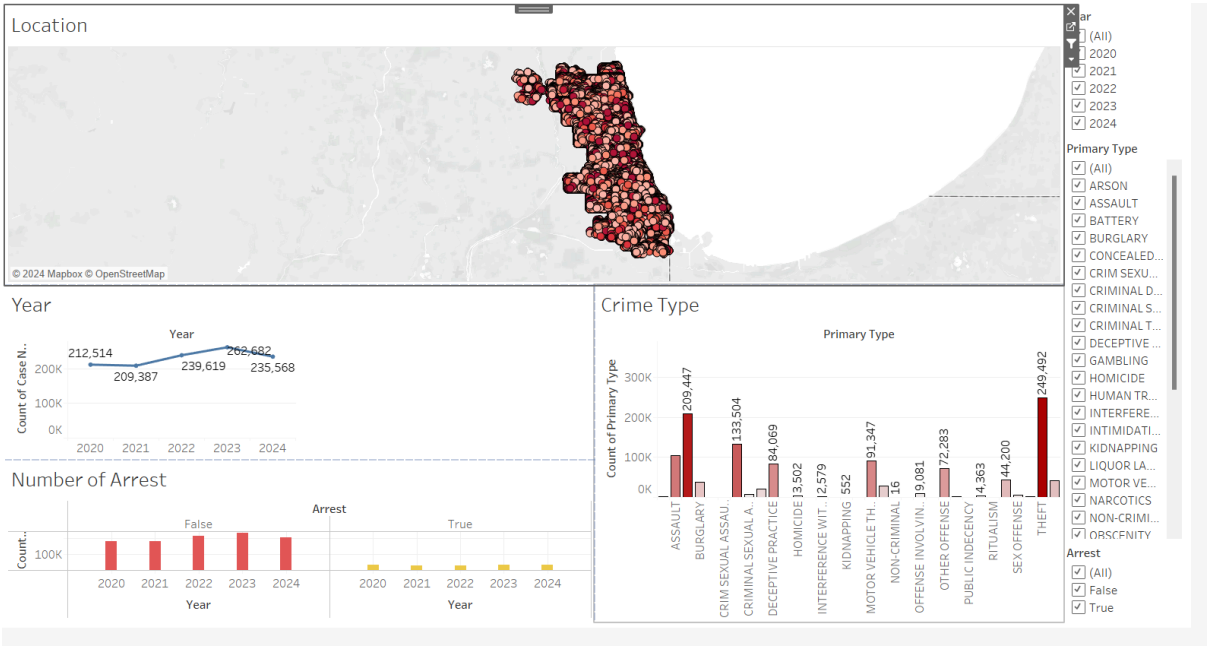
2. Data Preparation

For our analysis, we focused on data from 2020 to 2024, encompassing approximately 1,000,000 reported cases. This large dataset allows us to examine trends in crime types, arrest rates, and geographic distribution. Fortunately, the dataset is well-maintained, and with nearly 1,000,000 rows of data, the minimal presence of null values does not significantly impact the overall analysis.

3. Dashboard Components

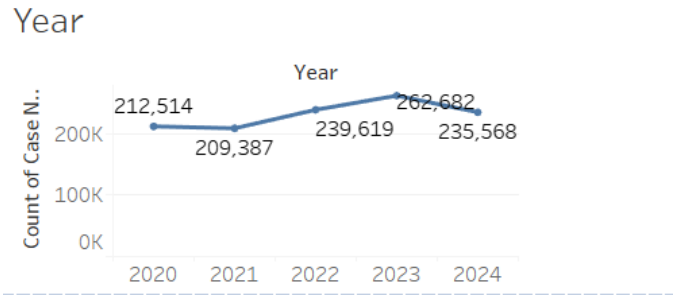
This dashboard has four visualizations with interactive filters for year, crime type, location, and arrest result to explore our data. The explanation of each dashboard will be demonstrated below.

- **Location of Crime Happen:** The map visualization provides a graphical representation of crimes across Chicago, with different kinds of red dots indicating a reported crime. This helps identify hotspots of criminal activity within the city. It also is a filter that can help users to find crime type, report year and arrest result happen at that spot.
- **Year and Crime Number:** The line chart displays the total number of crimes reported each year from 2020 to 2024. Users can use the year filters to focus on specific years to help them understand more information with other filters.
- **Crime Type:** The bar chart categorizes crimes based on their primary type, highlighting the frequency of each crime type during the analyzed years. Users can focus on a specific type of crime they are interested in by the filter.
- **Arrest Result:** The bar chart illustrates the proportion of crimes that resulted in arrests versus those that did not. False means fail to arrest, in contrast, True means arrest successfully.
- **Dashboard overview:**

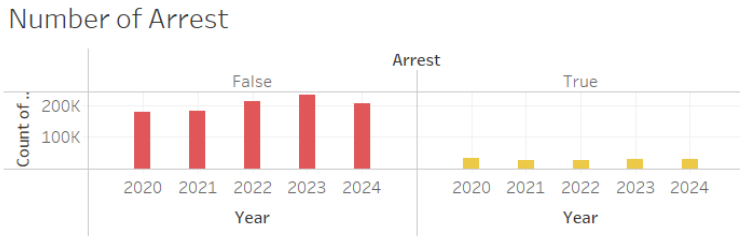


4.Findings

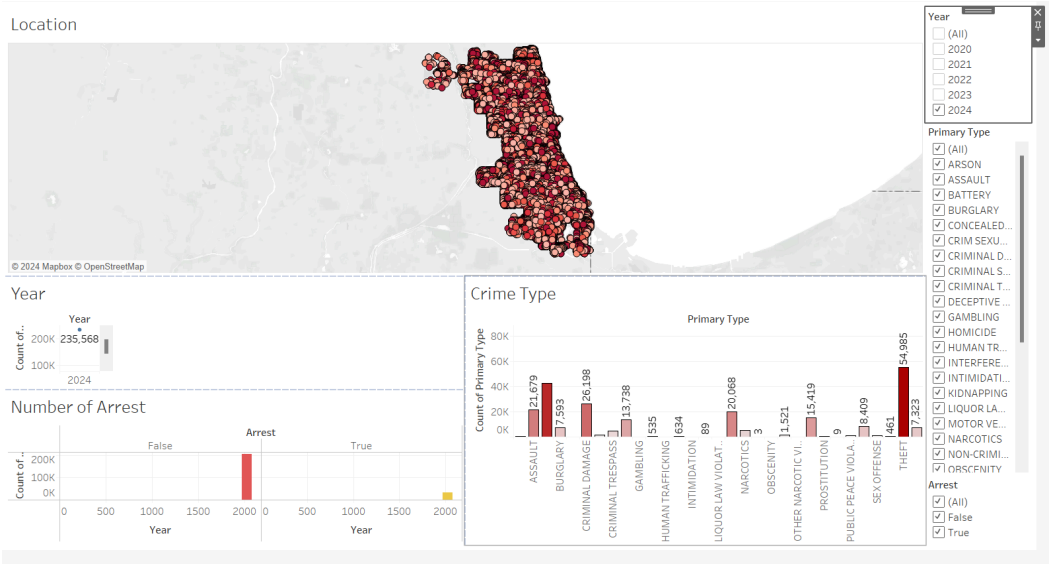
1. This graph illustrates the yearly trend in reported crimes in Chicago from 2020 to 2024, providing insight into how societal changes, we thought it particularly related to the COVID-19 pandemic, influenced criminal activity.
    - 2020-2021 (Pandemic Years): During this period, crime rates declined, with 212,514 cases in 2020 and 209,387 in 2021. This decrease can likely be attributed to the effects of the pandemic, including lockdowns, reduced public gatherings, and limited mobility, which lowered opportunities for crimes such as theft or public altercations.
    - 2022 (Post-Pandemic Transition): As restrictions were lifted and life began returning to normal, crime rates surged significantly, reaching 239,619 cases. This marked the largest year-on-year increase in the dataset, reflecting how increased public activity and the return to social interactions created more opportunities for criminal behavior.
    - 2023 (COVID Declared Over): Crime peaked in 2023, with 239,619 cases reported. This peak coincided with the full reopening of society after the World Health Organization declared the end of the global COVID-19 emergency. The sustained high crime levels during this period suggest a lag in law enforcement and social systems adapting to post-pandemic challenges.
    - 2024 (Stabilization Period): In 2024, there was a slight decline in crime, with 235,568 cases reported. This decrease may indicate efforts by law enforcement to address the rising crime trend or societal stabilization after the disruptions caused by the pandemic.

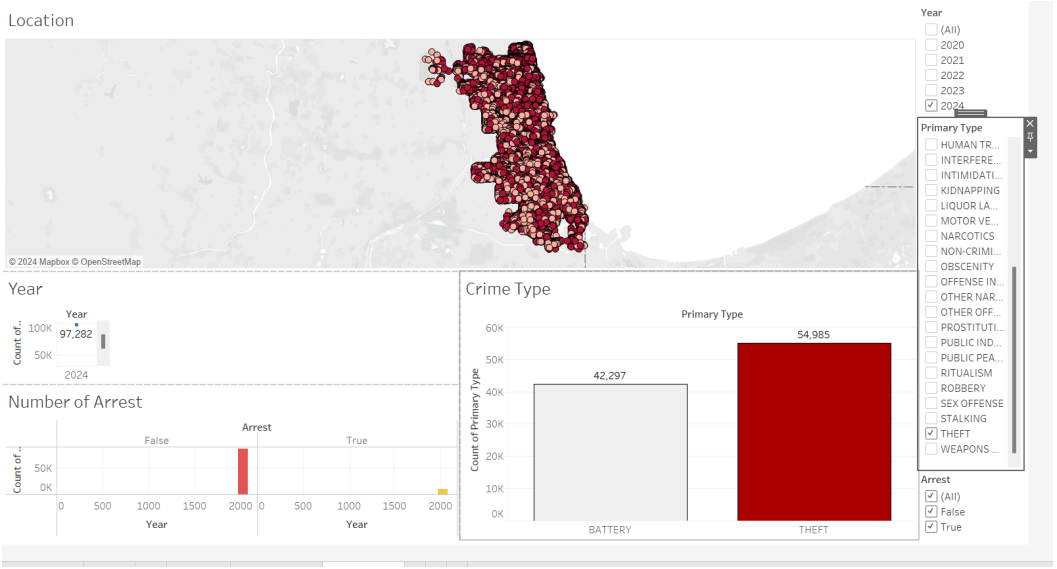


2. This graph displays the number of crimes in Chicago from 2020 to 2024, categorized by whether an arrest was made. The data is divided into two groups: **"False"** (no arrests made) and **"True"** (arrests made), represented in red and yellow, respectively. The interesting thing about this graph is color Choice**.** The red and yellow color scheme was chosen to ensure accessibility for individuals with color vision deficiencies like me (Youdong Lu). These colors provide a clear contrast, making it easier for everyone, including those who are colorblind, to distinguish between the two categories. This design consideration reflects an effort to make the visualization inclusive and accessible to a broader audience.



3. In 2024, crime remained a significant concern in Chicago, with battery and theft emerging as the two most common types of reported offenses. Battery, with its high frequency, often involves physical harm or threats, making it a top priority for law enforcement. Theft, particularly related to property, also continues to be a widespread issue, affecting both businesses and individuals. These crimes not only pose a danger to public safety but also strain the city's resources as law enforcement works to address them.





To address crime in Chicago, particularly battery, theft, and gang-related violence, both the police and residents must work together. The police should focus on disrupting gang activities, engage in community policing to build trust, and invest in crime prevention programs for at-risk youth. Residents can help by participating in neighborhood watch programs, supporting safe spaces for youth, and maintaining open communication with law enforcement. Additionally, individuals should take personal safety precautions, like improving home security and being aware of their surroundings. Together, these efforts can reduce crime and enhance public safety.

Sesonal-Level Insights (Emily Chang)

1. Purpose of the Dashboard

The primary goal of this dashboard is to help stakeholders—such as policymakers and the Chicago Police Department—understand seasonal crime patterns across

Chicago from 2020 to 2024. By analyzing seasonal crime volumes, felony rates, arrest rates, and geographic hotspots, the dashboard provides actionable insights to identify high-risk periods and locations, enabling more efficient allocation of law enforcement resources and strategic planning. This tool empowers data-driven decision-making to address crime more effectively throughout the year.

2. Data Preparation

2.1 Data Cleaning

The dataset included crime records from 2020 to 2024, with critical fields like `Date`, `Primary Type`, `Season`, `Arrest`, and `IUCR`. Data cleaning ensured that missing or inconsistent entries were removed, leaving high-quality data for analysis. Seasonal categorization was derived from the `Date` field, and new fields like `Felony Rate` and `Arrest Rate` were calculated for nuanced insights.

2.2 Calculated Fields

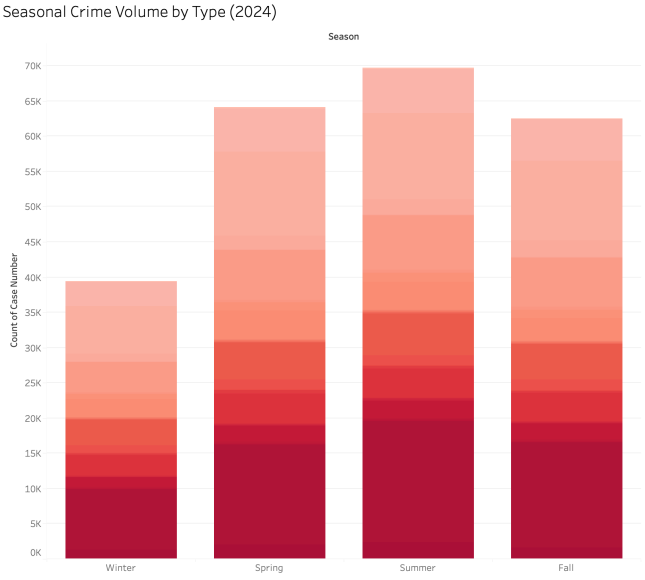
- **Season**

```
IF DATEPART('month', [Date]) IN (3, 4, 5) THEN "Spring"
ELSEIF DATEPART('month', [Date]) IN (6, 7, 8) THEN
"Summer"
ELSEIF DATEPART('month', [Date]) IN (9, 10, 11) THEN
"Fall"
ELSE "Winter" END
```

3. Dashboard Components

The dashboard consists of four visualizations designed to uncover seasonal crime patterns:

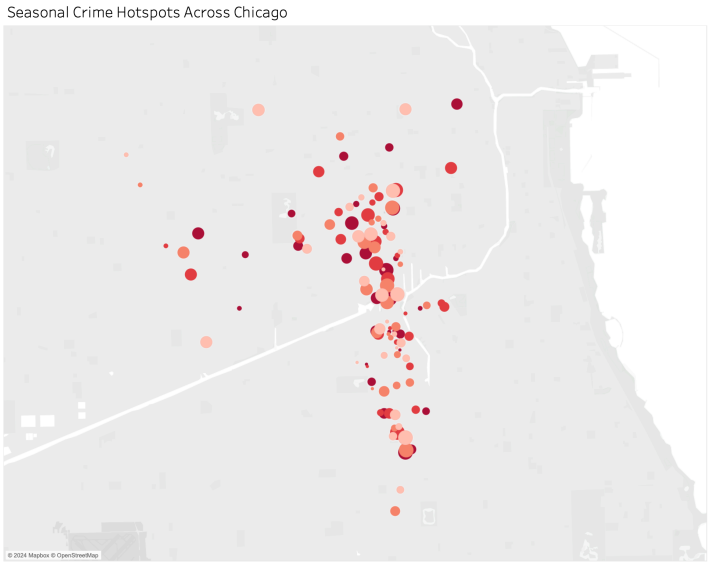
3.1 Seasonal Crime Volume by Type



- Purpose: Show the distribution of crime types across seasons.
- Details:

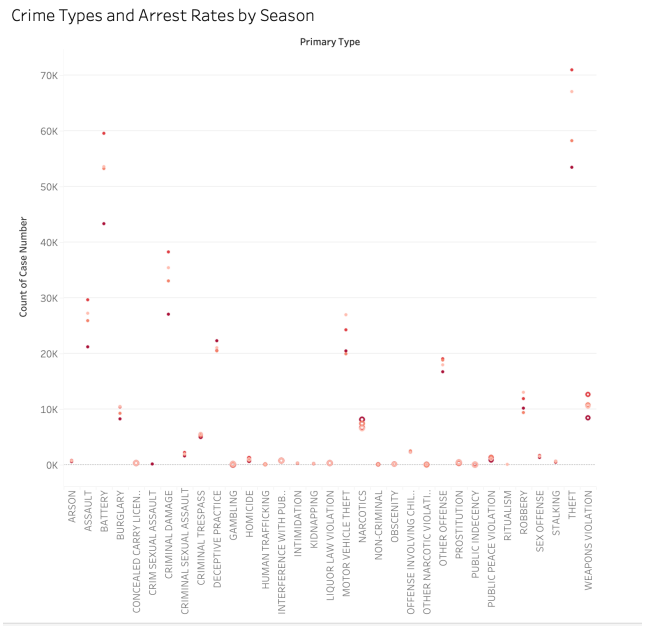
- Each bar represents a season.
- Crime types are color-coded.
- Insights: Summer has the highest crime volume, dominated by theft and battery, while Winter has the least.

3.2 Seasonal Crime Hotspots Across Chicago



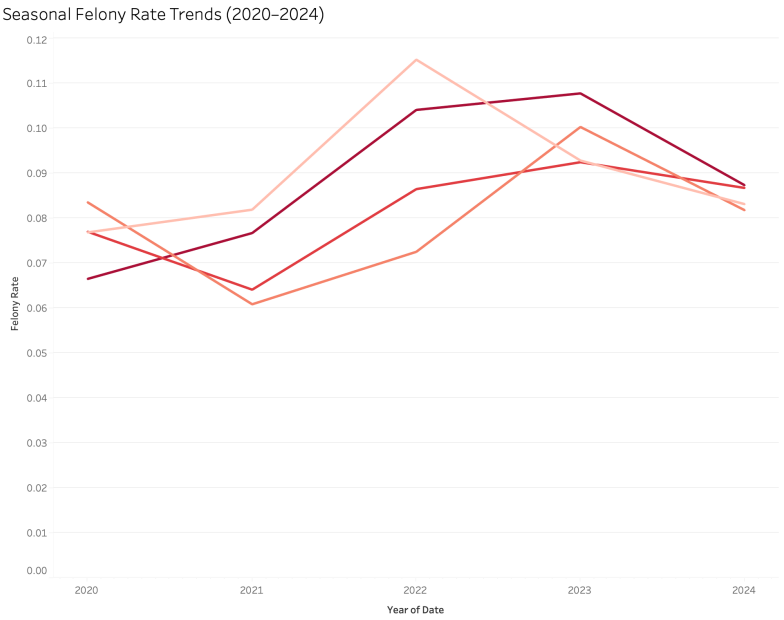
- Purpose: Highlight geographic areas with high crime activity in each season.
- Details:
  - Larger bubbles indicate higher crime counts.
  - Seasonal filters allow users to toggle between Winter, Spring, Summer, and Fall.
  - Insights: Central Chicago consistently shows high crime density, while southern regions have seasonally fluctuating hotspots.

3.3 Seasonal Felony Rate Trends (2020–2024)



- Purpose: Track how felony rates change over the years for each season.
- Details:
    - Line chart with separate lines for each season.
    - Seasonal filters to highlight specific trends.
    - Insights: Felony rates are highest in Summer and Fall, showing an upward trend post-2022.

3.4 Crime Types and Arrest Rates by Season



- Purpose: Examine the relationship between crime types and arrest rates across seasons.
- Details:
    - Dot plot comparing crime counts (`COUNT([Case Number])`) and arrest rates for various crime types.
    - Arrest rate is represented by dot size, with seasons differentiated by color.
    - Insights: Theft has high volume but low arrest rates across all seasons, while narcotics offenses have higher arrest rates due to targeted enforcement.

4. Findings

4.1 Overall Seasonal Patterns

1. **Crime Peaks**:
    - Summer consistently experiences the highest crime volume, driven by theft and battery.
    - Winter sees the lowest crime activity, likely due to reduced outdoor movement.

2. **Felony Trends**:
    - Felony rates are highest in Summer and Fall, suggesting more severe crimes occur during these periods.
3. **Geographic Insights**:
    - Central and southern regions of Chicago are persistent hotspots across all seasons.

4.2 Arrest Effectiveness

- Low Arrest Rates: Theft has low arrest rates year-round despite being the most common crime.
- High Arrest Rates: Narcotics crimes have high arrest rates, reflecting targeted law enforcement efforts.

4.3 Seasonal Differences in Crime Types

- Theft and Battery: Dominate during Summer.
- Narcotics and Assault: Significant in Fall and Winter, possibly due to indoor activities and enforcement focus.

Days and Times Insights (Richie Li)

1. Purpose of the Dashboard

This dashboard analyzes crime data in Chicago across different **days**, **times of the day**, and **community areas**. It will reveal peak hours or days when certain crimes are more likely to occur—for example, comparing crime trends on weekdays vs weekends or days vs nights. People will be able to know when and where to avoid certain areas from the history trend through this dashboard.

2. Data Preparation

2.1 Geological Data

To show different community areas in Chicago, this dashboard uses the geodata provided by the Chicago Data Portal, which allows Tableau to draw border lines on the map.

2.2 Calculated Fields

Several calculated fields were created to enhance the dashboard's analytical capabilities.

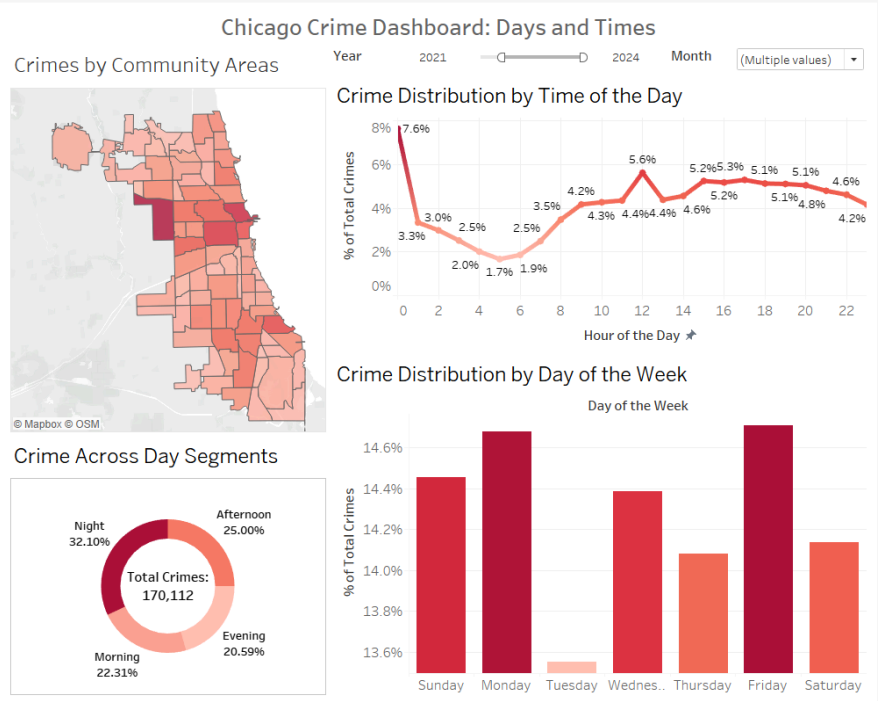
- **Day of the Week:**
  ```
  DATENAME('weekday', [Date])
  ```
- **Hour of the Day:**
  ```
  DATEPART('hour', [Date])
  ```
- **Time period:**
  ```
  IF DATEPART('hour', [Date]) >= 5 AND DATEPART('hour', [Date]) < 12 THEN "Morning"
  ```

```
ELSEIF DATEPART('hour', [Date]) >= 12 AND
DATEPART('hour', [Date]) < 17 THEN "Afternoon"
ELSEIF DATEPART('hour', [Date]) >= 17 AND
DATEPART('hour', [Date]) < 21 THEN "Evening"
ELSE "Night" END
```

3. Dashboard Components

The dashboard includes four visualizations with interactive filters for year and month to dynamically explore data. A gradient red color scheme represents crime intensity, emphasizing high-crime areas. The following is a brief introduction to the visualizations in this dashboard:

- **Crimes by Community Areas:** The community map highlights crime distribution across community areas in Chicago, using polygons to display boundaries and red gradients for intensity.
- **Crime Distribution by Time of the Day:** This line chart visualizes the hourly distribution of crimes throughout the day. It updates dynamically with selected filters.
- **Crime Distribution by Day of the Week:** This bar chart visualizes the percentage of total crimes by each day of the week. It updates dynamically with selected filters.
- **Crime Across Day Segments:** The donut chart segments crimes into Morning, Afternoon, Evening, and Night. Where a deeper color means a higher ratio of crime rate. It updates dynamically with selected filters.



4. Findings

4.1 Crimes by Community Areas

We first analyzed the overall performance of Chicago's 77 community areas. As shown in the visualization, crime is mainly concentrated in specific areas within the central and southern regions of Chicago. This result is in accordance with what we see and believe, the more southern a person goes, the more dangerous it is.

4.2 Crime Distribution by Time of the Day

We find several interesting observations when looking into the crime distribution of a day.

1. The distribution, which we expected to be smooth, has some abnormal peaks at 12 am/pm. This might result from the police department's upload time. We suspect that not every crime case has the correct time recorded. Especially lots of them are recorded in xx:00.
2. The time between 1 and 7 has the lowest crime compared to other times. It could be because fewer people are on the streets, which means there is a lower chance of a crime happening.

4.3 Crime Across Day Segments

This visualization is an alternative to the last one. This donut chart shows the crime rate in different time segments:

- Morning: 7 - 12 (5 hours)
- Afternoon: 12 - 17 (5 hours)
- Evening: 17 - 21 (4 hours)
- Night: 21 - 7 (10 hours)

Although Night time has the highest proportion in the donut chart, it spans 10 hours, while others only 4 or 5 hours. Leading to the conclusion that afternoons and evenings are peak times for crime to happen.

4.4 Crime Distribution by Day of the Week

There are seven days a week, so if the occurrence of a crime is independent of the day of the week, the distribution of crime should be 100/7 = 14.29% each day. Clicking through the filters, we do not see a significant difference in crime distribution throughout the week. No pattern helps us learn which day in a week has the most crimes committed. If we really need to make a comparison, the crime rates on weekends (Friday to Sunday) are slightly higher than on weekdays. Again, may be that more people are on the street which increases exposure to threats.

Summary

The analysis of Chicago's crime data from 2020 to 2024 reveals key patterns: southern districts like 4, 6, and 8 report the highest crime rates, while northern districts such as 16 and 17 are safer. District 12 shows both a low arrest rate and a high felony rate, highlighting the need for targeted interventions. Crime surged post-COVID-19, peaking in 2023, with slight stabilization in 2024. Summer emerges

as the most crime-prone season, dominated by theft and battery, while winter reports lower crime activity.

To address these issues, we recommend CPD focus patrols on high-crime districts, enhance theft arrest rates using technology and community partnerships, and maintain targeted narcotics enforcement. Seasonal insights can optimize resource allocation, especially during summer. Collaboration with residents through neighborhood watch programs and public safety initiatives can further enhance security and reduce crime rates citywide.