

# **TOPIC: CRIME ANALYSIS IN CHICAGO CITY**

## **ABSTRACT**

Security has always been one of the most significant concerns. Government and security agencies are working hard to prevent crimes and protect their people. However, the challenge of dealing with large amounts of data has become a major issue for all organizations and therefore, a crime information system that is able to process large amounts of data in a short period of time is needed. News and records indicate that there have been many crimes in Chicago city. It is called “Murder Capital” or “Crime Capital” of the U.S in 2012. Chicago had more murders in 2012 than any other city in the country. Therefore, there is a requirement to make more informed decisions and analyze crimes occurring in different regions of Chicago city.

Crime analysis plays a crucial role in understanding and addressing the complexities of crime patterns within urban environments. This project focuses on Crime Analysis in the City of Chicago, a metropolitan area known for its diverse neighborhoods and complex crime dynamics.

This study is to explore and analyze Crimes in Chicago City using a data-driven approach. Data Sources, including the official Chicago Data Portal, and Boundaries data, are integrated to identify the criminal activities within the City. This project aims to investigate Crime Analysis in Chicago City. Chicago Crime Data for 3 years 2021–2023 were analyzed. The research findings contribute to decision-making for Law Enforcement Agencies, Policymakers, and Urban Planners.

## **INTRODUCTION**

There is a rapid increase in crime in almost every country. There is a strong need to identify crime patterns and analyze different areas of Crime. Chicago is one of the largest cities in the United States, it faces unique complexities in understanding and addressing criminal activity within its urban landscape.

Security agencies all over countries are working hard to reduce these crimes, however, size of crime information is increasing rapidly, and it becomes difficult to manage such a huge amount of data and to keep record of crimes that are geographically widespread and at different time. Thus, it is very necessary to have a crime information system which can process large amount of data in short period of time. To tackle this issue effectively, crime analysis emerges as a vital tool, providing valuable insights into Crime Patterns or Trends, and Crime Hotspots.

The New York Times ranked Chicago's rate of homicide 7th behind Detroit, Milwaukee, Atlanta, and St. Louis. Crime in Chicago overall has increased during the past five years by nearly 20%, according to an official report by the Chicago Police Department.

Chicago Crime Dataset has 30 features which includes Primary Type, Latitude, Longitude, Arrest, Description, Location Description etc. The project focuses on analyzing crime in Chicago by investigating Crime Hotspots, Identifying Crime Patterns, and Developing a Dashboard displaying the Count of Crimes in each Neighborhood. We collected Crime Datasets from the Chicago Data Portal and preprocessed the data to ensure accuracy and reliability. Through Exploratory Analysis, we visualized relationships between variables to uncover meaningful patterns. Using the Folium package and K-means clustering, we identified Crime Hotspots and Patterns, providing actionable information for law enforcement. The Streamlit library is used to create an interactive Dashboard for visualizing crime counts by neighborhood. Overall, our project aims to provide valuable insights and tools to address crime effectively and enhance public safety in Chicago.

## **METHODOLOGY**

### **Load Packages**

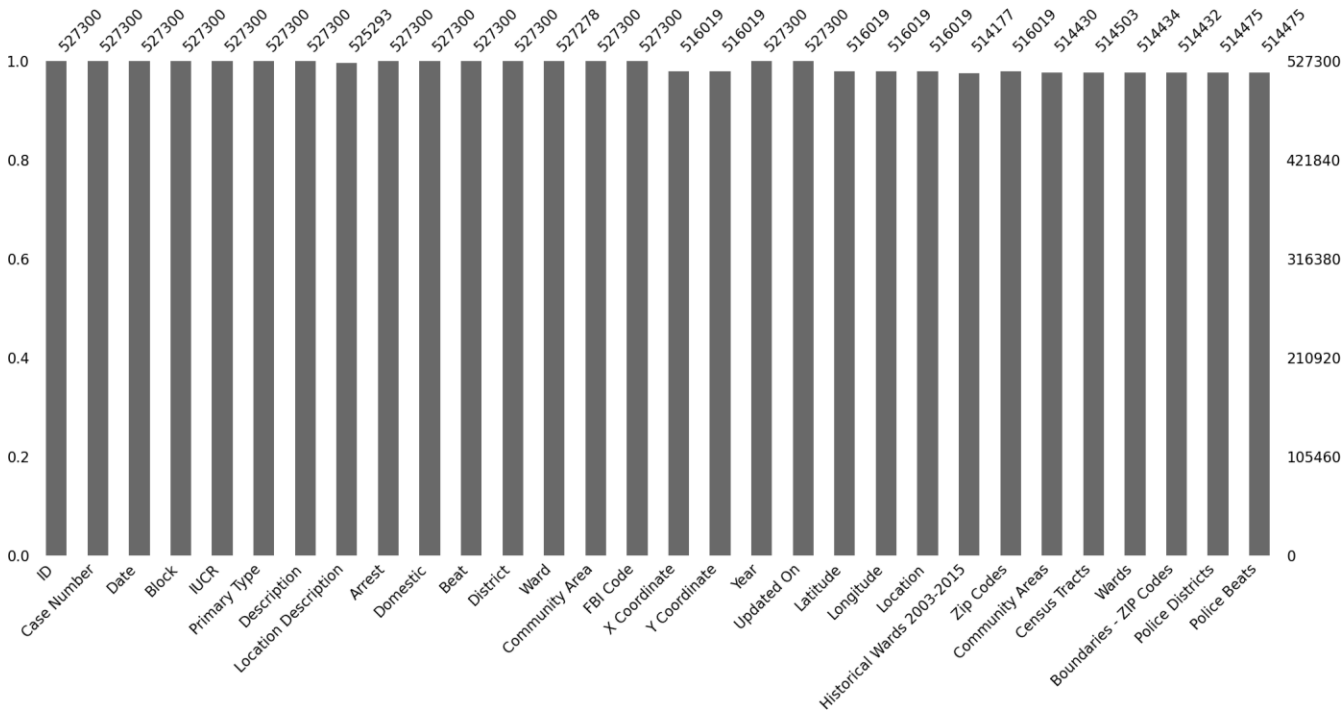
```
import pandas as pd
import numpy as np
from sklearn.cluster import KMeans
import folium
import missingno as msno
import matplotlib.pyplot as plt
import geopandas as gpd
import holoviews as hv
import geoviews as gv
import geoviews.tile_sources as gts
from folium.plugins import MarkerCluster
from matplotlib.cm import get_cmap
from geopy.geocoders import Nominatim
import seaborn as sns
```

## **PREPROCESSING**

### **Display of Dataset**

														Historical Wards			
		Case Number	Date	Block	IUCR	Primary Type	Description	Location Description	Arrest	Domestic	...	Longitude	Location	2003-2015	Zip Codes	Community Areas	Census Tracts
0	13061882	JG247197	05/03/2023 11:58:00 PM	039XX W 83RD ST	1310	CRIMINAL DAMAGE	TO PROPERTY	RESIDENCE	False	True	...	-87.720375	(41.742201266, -87.72037516)	6.0	4,300	69.0	199.0
1	13062468	JG247914	05/03/2023 11:53:00 PM	024XX N LARAMIE AVE	0910	MOTOR VEHICLE THEFT	AUTOMOBILE	STREET	False	False	...	-87.756195	(41.925405035, -87.756194727)	7.0	22,615	19.0	80.0
2	13061897	JG247243	05/03/2023 11:51:00 PM	080XX S BISHOP ST	141A	WEAPONS VIOLATION	UNLAWFUL USE - HANDGUN	STREET	False	False	...	-87.659740	(41.748132542, -87.659739879)	18.0	21,554	70.0	558.0
3	13061896	JG247190	05/03/2023 11:47:00 PM	043XX N CENTRAL AVE	0860	THEFT	RETAIL THEFT	SMALL RETAIL STORE	False	False	...	-87.767223	(41.959891679, -87.767222504)	25.0	22,254	15.0	260.0
4	13061999	JG247280	05/03/2023 11:45:00 PM	048XX N KILDARE AVE	0486	BATTERY	DOMESTIC BATTERY SIMPLE	STREET	False	True	...	-87.732779	(41.873723427, -87.732778965)	36.0	21,572	27.0	717.0
5	13061886	JG247202	05/03/2023 11:45:00 PM	015XX S PAULINA ST	0460	BATTERY	SIMPLE	SIDEWALK	False	False	...	-87.668706	(41.859679422, -87.668705919)	48.0	14,920	29.0	96.0
6	13061933	JG247201	05/03/2023 11:39:00 PM	002XX S FRANKLIN ST	0460	BATTERY	SIMPLE	STREET	False	False	...	-87.635213	(41.878506591, -87.635213191)	48.0	14,914	38.0	92.0

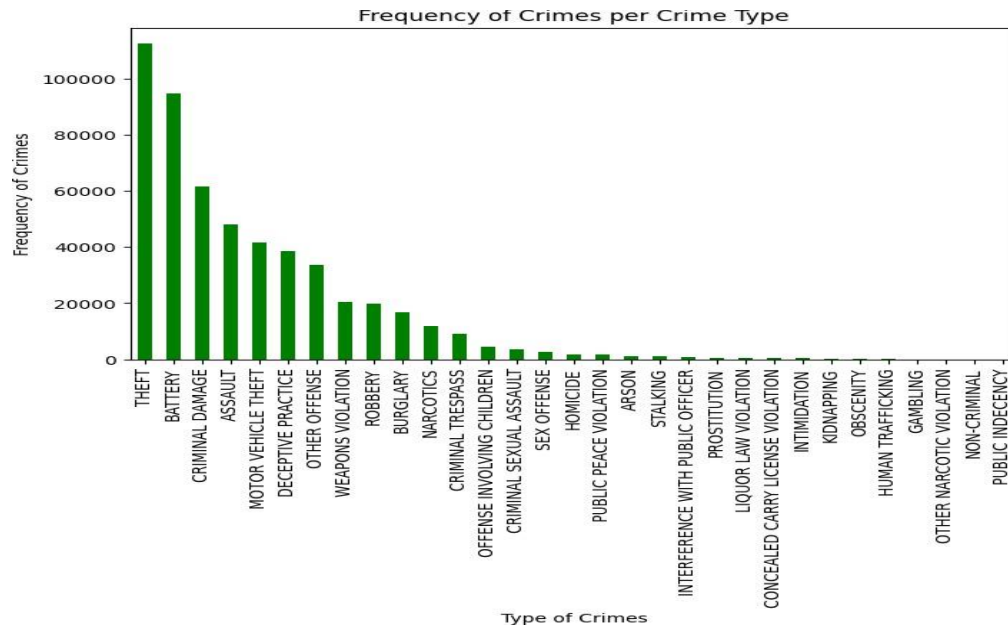
Check for Missing Values



Replacing Missing Values of Longitude, Latitude, X Coordinate and Y Coordinate to their respective median values and dropping of Missing Values for Location description, Ward etc.

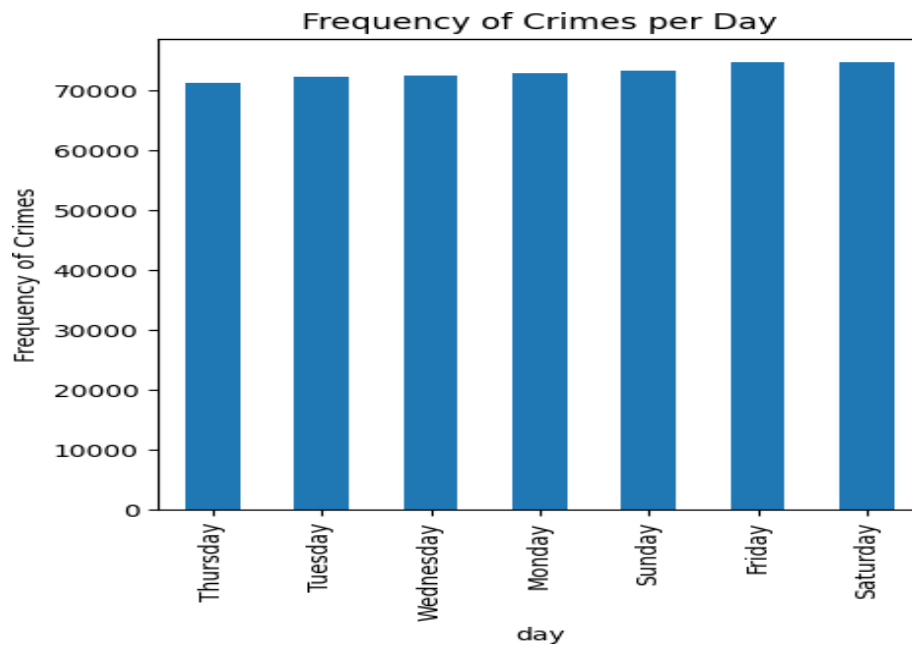
EXPLORATORY DATA ANALYSIS

Frequency of Crime per Crime Type



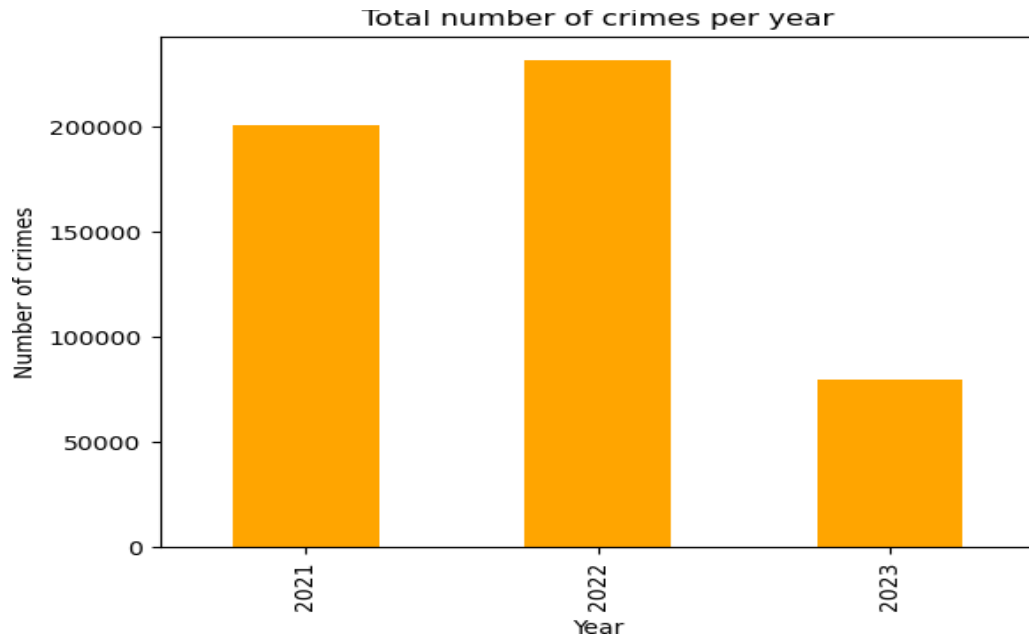
*This graph represents the Frequency of Crimes per Crime Type in the dataset, with Theft and Battery being the highest crime count.*

### Frequency of Crimes per Day



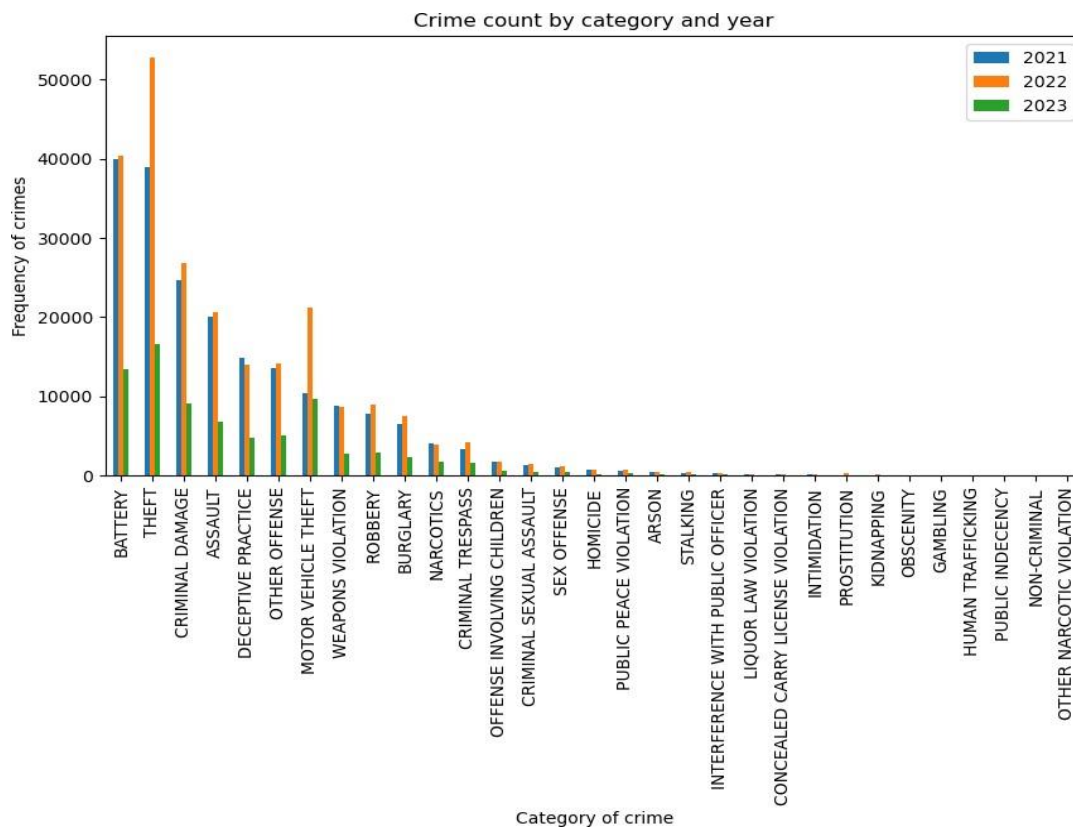
*The figure displays the Frequency of Crimes each day, with Friday and Saturday being the most active days for criminal activity.*

### Total Number of Crimes per Year



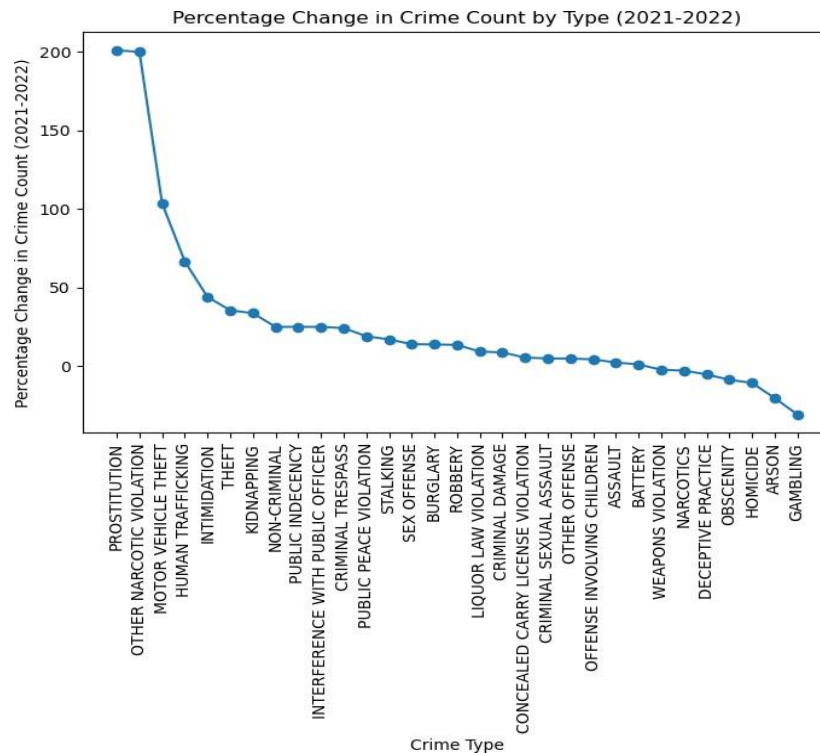
The graph depicts the Total Number of Crimes per Year, with 2022 having the highest number of crimes.

### Crime Count for Crime Types from 2021 to 2023



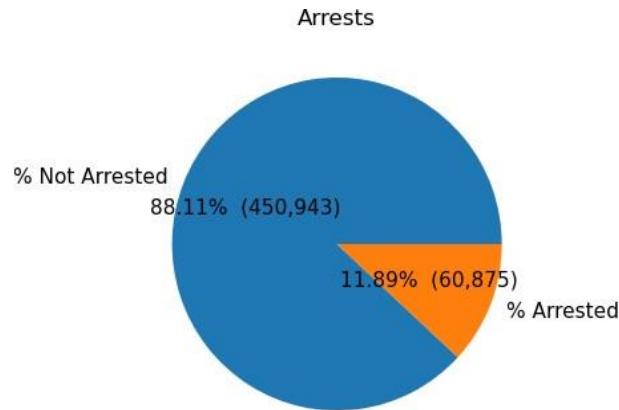
The graph represents the Crime Count by Crime type in each year (2021, 2022, and 2023). The graph allows you to compare the Crime Counts across different Crime types and observe any trends or changes in crime patterns over the years. Battery and Theft have the highest Crime count in 2022.

### Percentage change in Crime Count for each Crime Type from 2021 to 2022



*The graph provides a visual representation of how different Crime Types experienced changes in Crime Count from 2021 to 2022.*

### Percentage of Arrest



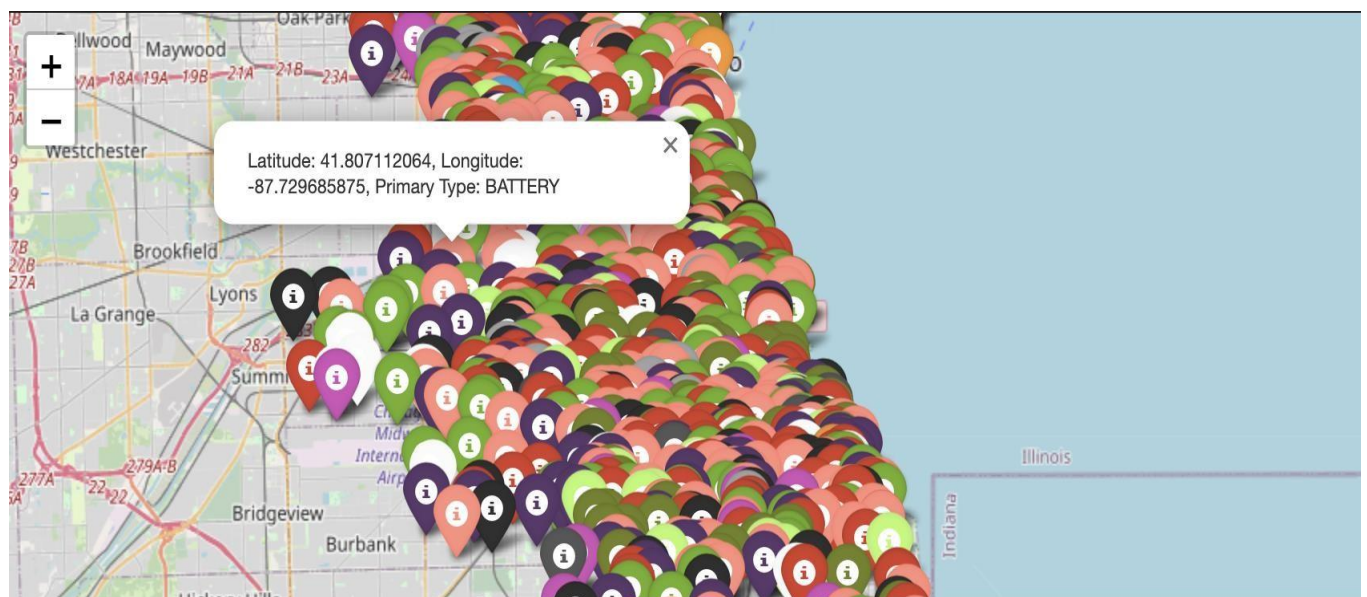
*The Pie Chart shows the Number of People (in % and count) who were Arrested and Not Arrested from the year 2021 – 2023.*

## VISUALIZATION OF VARIOUS PLOTS

### CRIME HOTSPOT

A Crime Hotspot refers to an area or location with a significantly higher crime rate compared to its surrounding areas. These Hotspots can vary in nature, including High Rates of Theft, Assault, Drug-Related Crimes, or other Criminal Activities. Chicago City has experienced certain areas with higher crime rates compared to others. It's worth noting that Crime Rates can be influenced by various socioeconomic factors, such as Poverty, Unemployment, and Gang Activity.

### DISTRIBUTION OF CRIME IN CHICAGO CITY





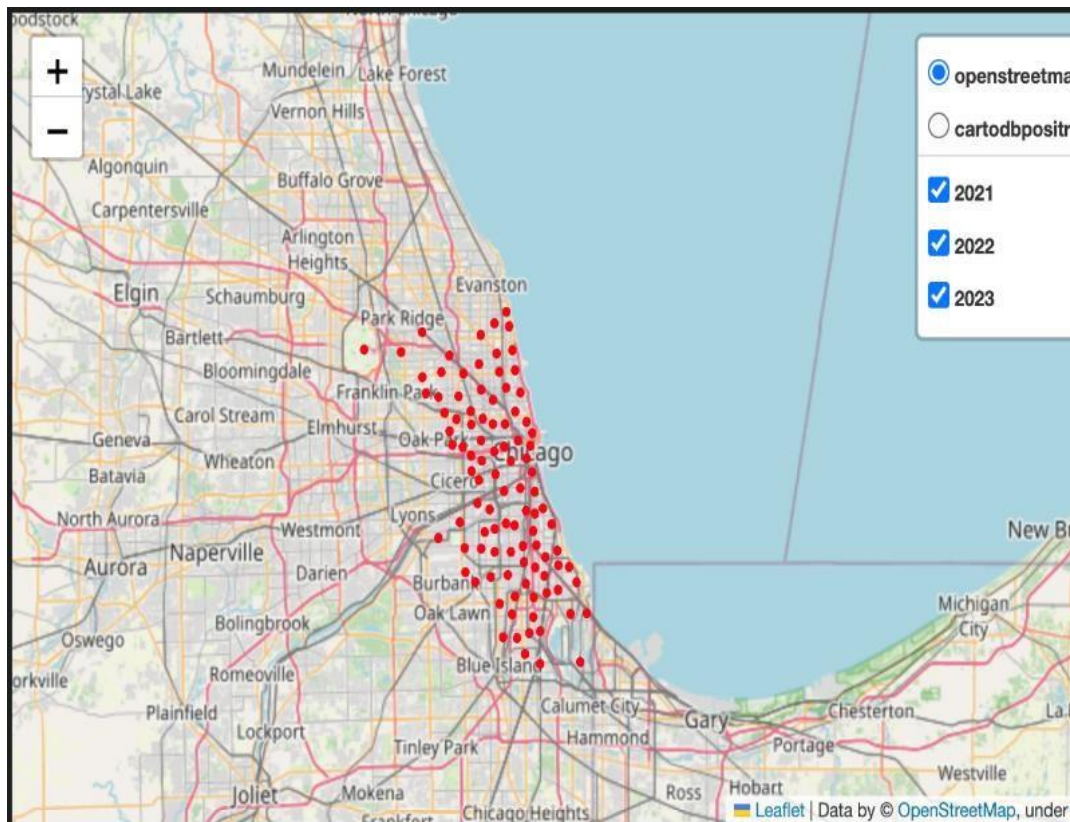
This map portrays the Distribution of various Crime Types in Chicago City using the folium.Marker. The pop-up text includes Latitude, Longitude of the Crime location and the type of Crimes with colour indicating different Crime Types.

## DISTRIBUTION OF BEATS

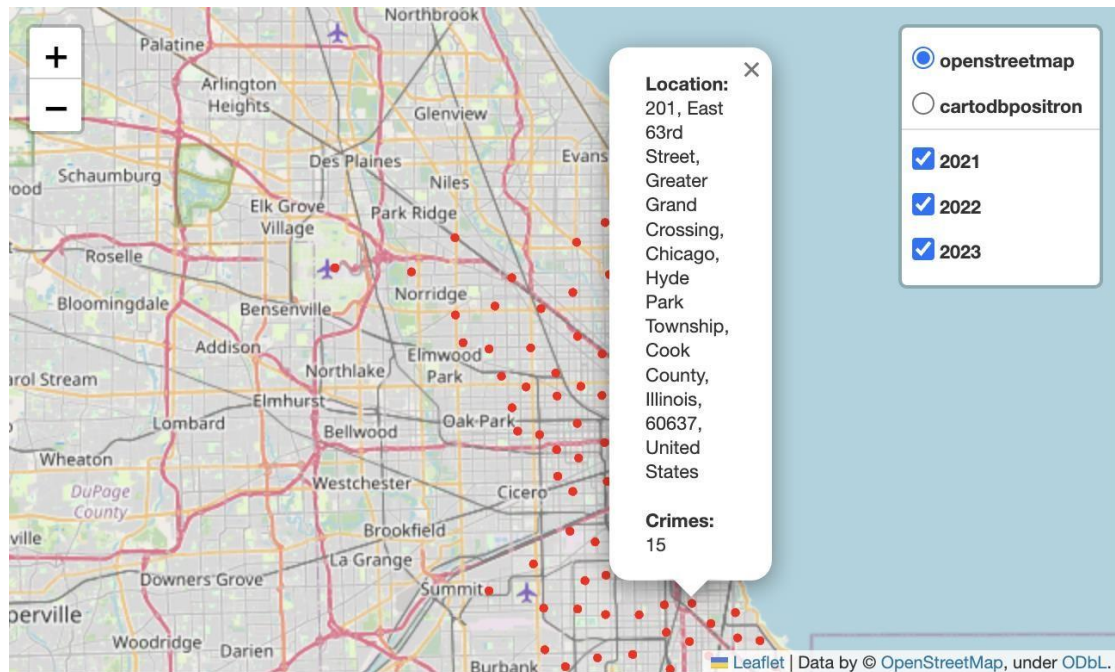
	beat_num	Total_Crimes
0	111	2979
1	112	2288
2	113	1352
3	114	2437
4	121	1173
5	122	2532
6	123	4077
7	124	1744
8	131	2873
9	132	1831

Beat indicates where an incident occurs and is the smallest Police geographic area – each Beat has a dedicated Police Beat Car. Three to five Beats make up a Police Sector, and three Sectors make up a Police District. From the above table, it indicates the Beat Number to its respective Total Number of Crimes occurred in a particular Police Sector.

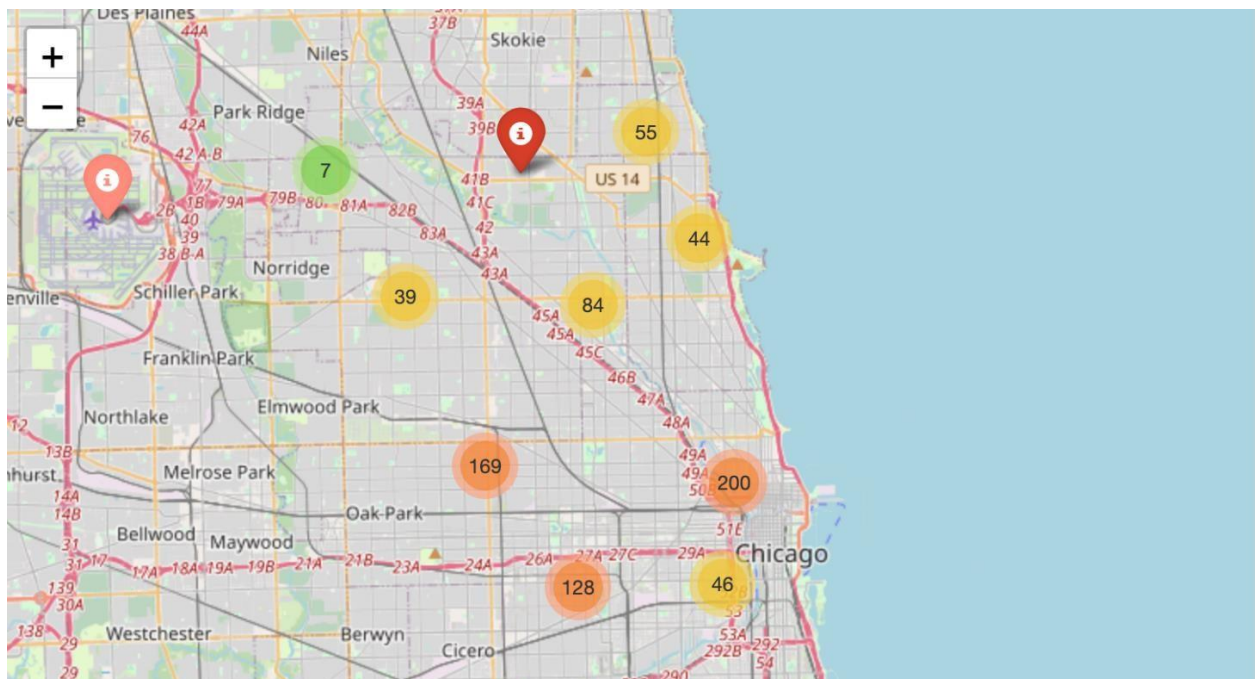
## DISPLAYING CRIME HOTSPOT

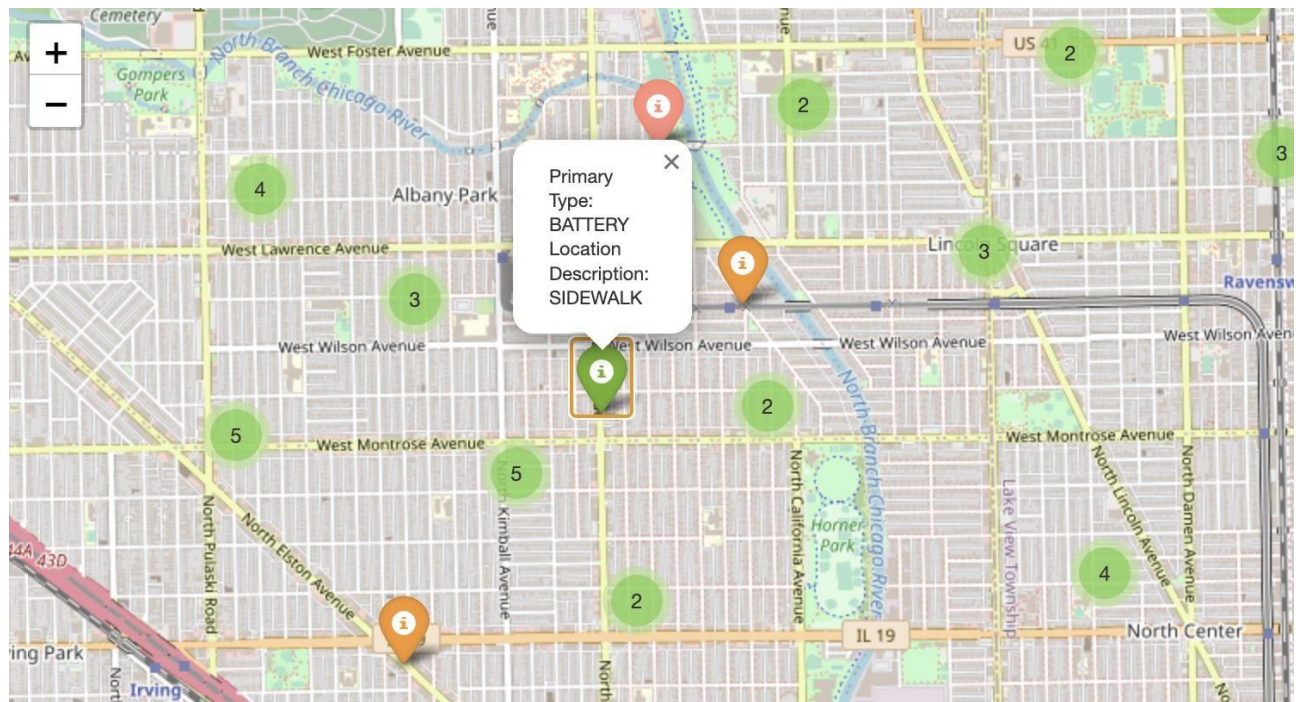






The above map shows the Crime Hotspots based on their location. It indicates the Total Number of Crimes that occur in a particular location. The map output shows a map of Chicago with circles of different sizes and Red Color representing Crime Clusters. The larger the circle, the more Crimes in the cluster. The map shows the Crime Clusters for each year in the dataset. The higher the Cluster Number, the better the visualization and widespread of Crime Clusters.





The above map shows a Cluster of Crimes with details of Primary Type and Location Description. Each Primary Type of Crime is indicated with different colours. The markers are clustered together at higher zoom levels for better visualization, and clicking on a marker reveals more information about the crime.

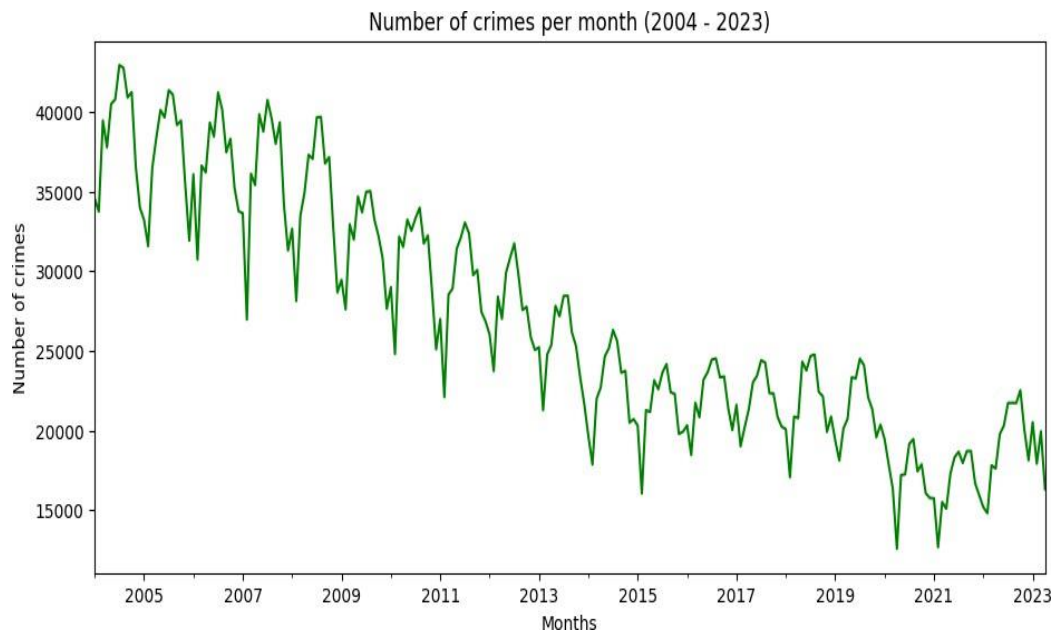
A marker cluster is a feature that groups multiple markers in a single, clickable marker that shows the number of markers it represents. It is useful when you have many markers to display on a map, as it makes the map less cluttered and more visually appealing.

## CRIME PATTERNS/TRENDS

Crime Patterns or Trends refer to recurring characteristics or behaviors observed in Criminal Activities over a period.

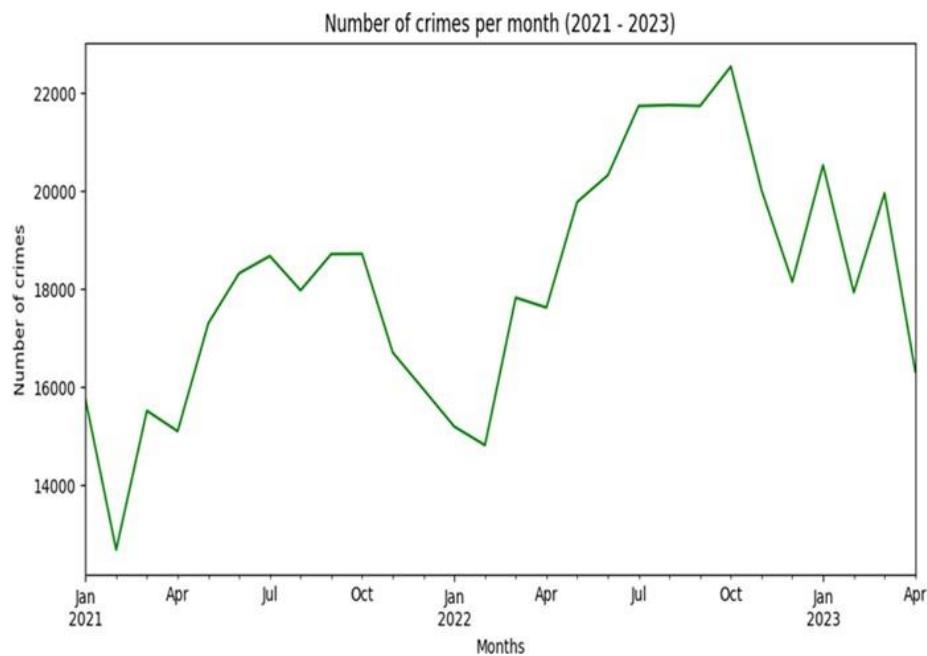
## CRIME PATTERNS/TRENDS PER MONTH

1.



From Plot 1, we can see that there are Fluctuations in the Number of Crimes over time, with some Months having higher numbers of Crimes than others. There are also some peaks and valleys in the trend line, indicating periods of higher and lower Crime Rates in the Months.

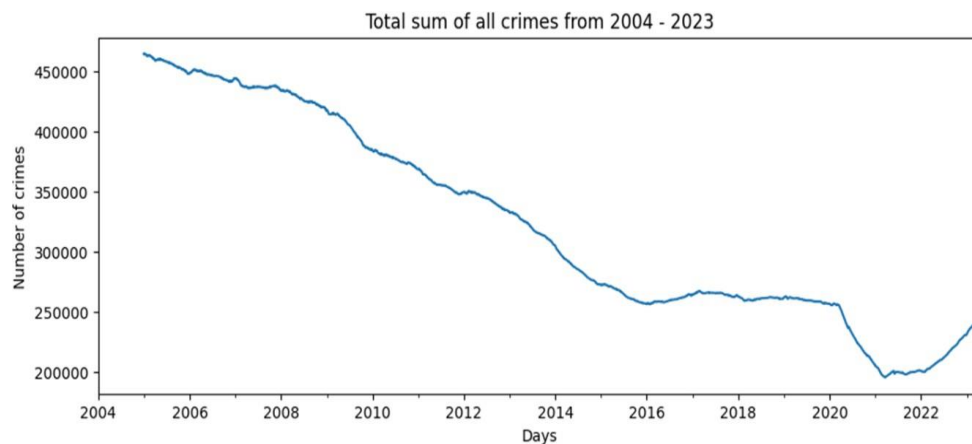
2.



From Plot 2, we delve to know the Type of Months contributing to the higher and lower Crimes Rates from 2021-2023. We can see that the first 3 Months (Jan, Feb, Mar) has lower Crime Rates due to a major factor like Weather - Chicago has very cold and snowy winters which make it more difficult for criminals to operate and people to move around the city. With fewer people on the streets, there may be fewer Crimes such as Theft, Robbery, Assault, and others from happening which can lead to lower Crime Rates.

Due to the Holiday Season, Crime Rates will be higher for the rest of the months. During the Summer and Christmas Seasons, Crime Rates may rise because of Shopping, Traveling, and Socializing.

### CRIME PATTERNS/TRENDS PER YEAR



*This chart shows the Total Sum of Crimes recorded from 2004 to 2023.*

The plot displays a trend line that shows the general trend of the Total Sum of Crimes over the selected period.

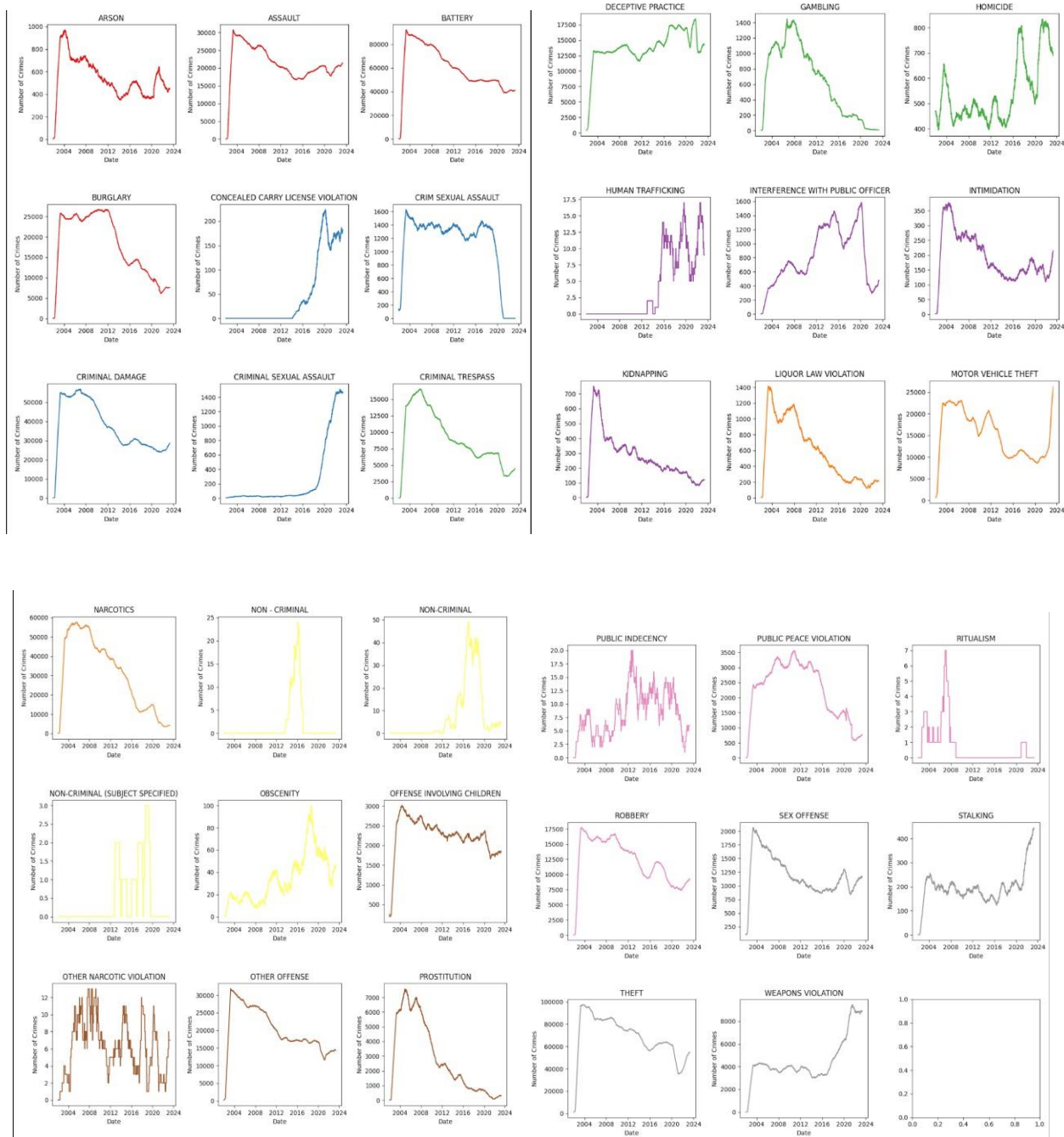
From the Plot, we can see that the Total Sum of Crimes has been slightly decreasing over time, with some Fluctuations in the trend. The line shows some peaks and valleys, indicating periods of higher and lower Crime Rates. We can see that 2021 – 2022 had lower Crime Rates.

Chicago, like many other cities around the world, experienced changes in Crime Patterns during the COVID-19 Pandemic. During the early stages of the Pandemic, Chicago, like many other cities, implemented stay-at-home orders and restrictions on public activities. These measures had an impact on Crime Patterns in the city. There was a decrease in certain types of crimes, such as Thefts, Burglaries, and Robberies, as people stayed home, and businesses remained closed.

However, it is important to note that not all types of crime decreased. Chicago, like some other cities, saw an increase in certain categories, including Homicides and Shootings. Some factors contributing to this increase may have been related to the economic and social effects of the Pandemic, strained community resources, and disrupted social networks.



## CRIME PATTERNS/TRENDS FOR VARIOUS PRIMARY TYPES



The plot shows subplots, with each subplot displaying the trend of the Total Sum of Crimes for a different Primary Type of Crime.

From the plots, we can see that some of the Crime Types contributing to Lower Crime Rates from 2021 - 2022 include Burglary, Sexual Assault, Criminal Trespass, Gambling, Interference with

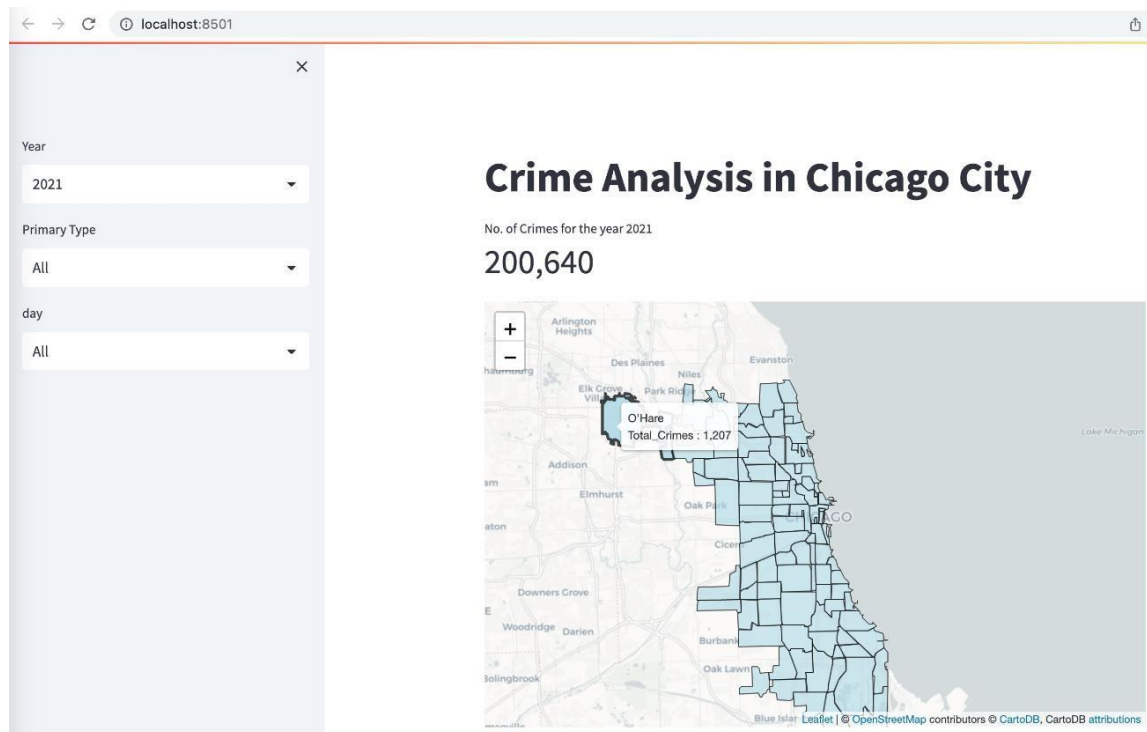
Public officer, Kidnapping, Liquor Law Violation Narcotics, Prostitution, Public Peace Evaluation, etc.

We can also observe that some Crime Types are at their peak, while others decline over time. For example, Burglary, Criminal Trespass, Sexual Assault, etc. declined from 2021. Some Crime Types, such as Homicide, Deception, and Human Trafficking, etc. are on an upward trend. Other Types of Crime, such as Theft, Robbery, Stocking, etc. were at their peak before 2016.

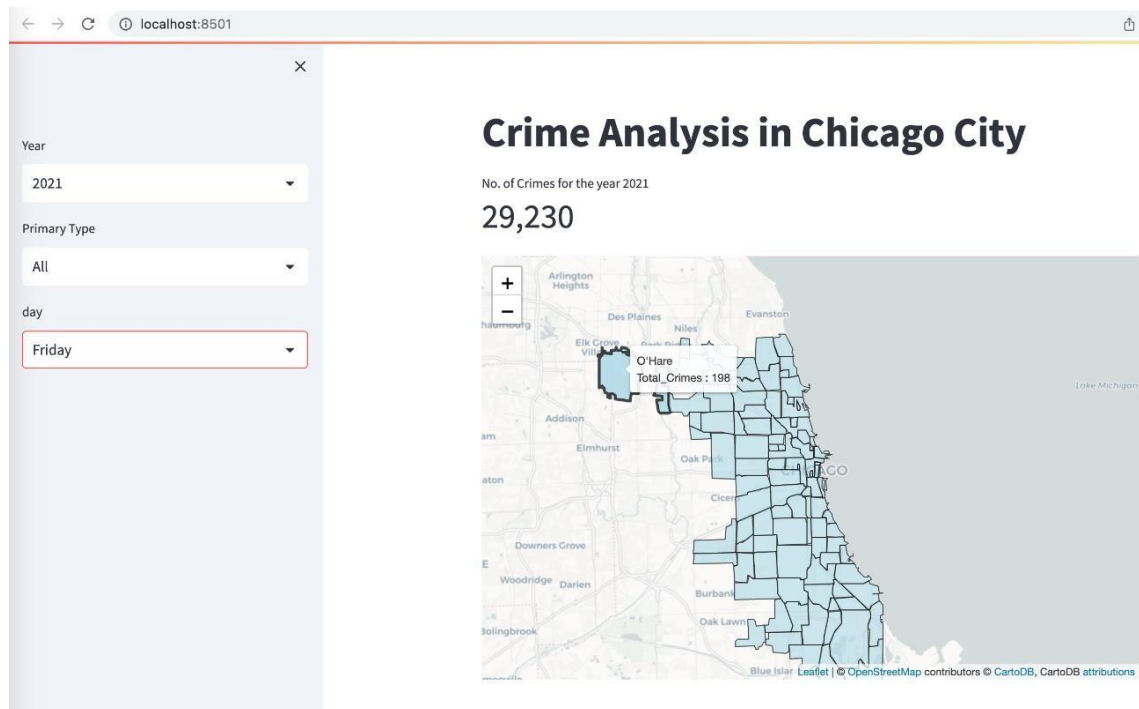
## DISPLAY OF CRIME RATES IN DASHBOARD WITH STREAMLIT

Streamlit is an open-source Python library for creating and sharing interactive web applications for data analysis and machine learning tasks. It provides a seamless way to convert Python scripts into interactive web apps, enabling users to explore and interact with data in a more intuitive and user-friendly manner.

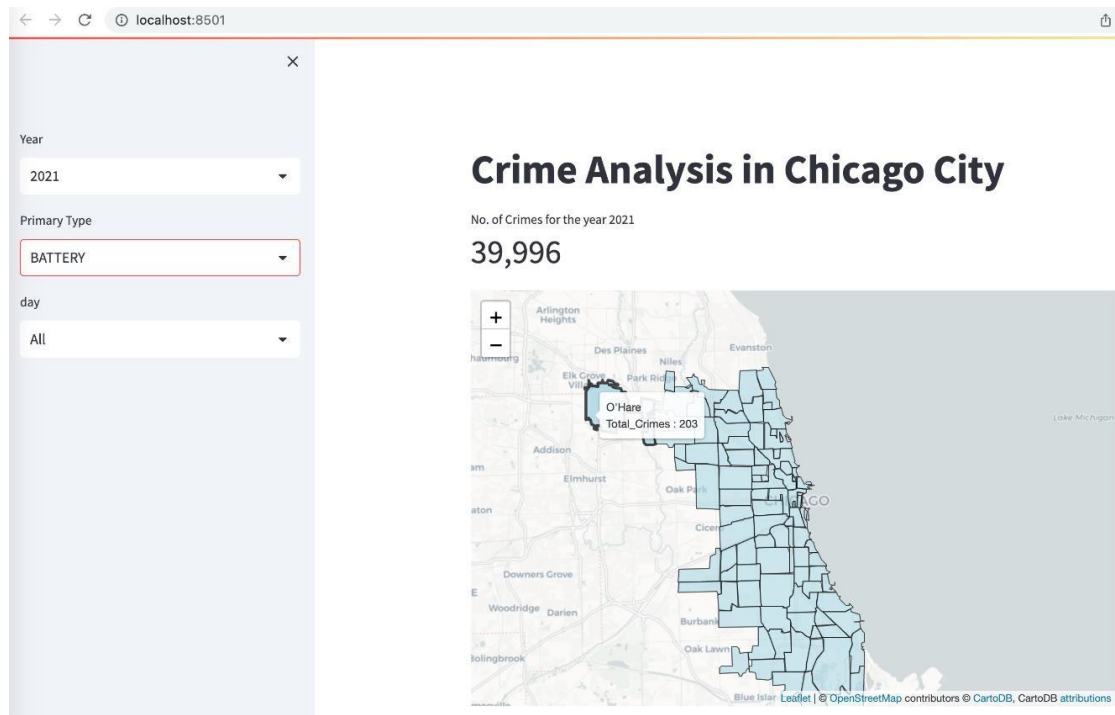
The dashboard helps filter the Total Number of Crimes per Year, Days, and Primary Types. Selecting the Days and Crime Type shows the Total Number of Crimes for each Neighbourhood per the selected Year.



The above Dashboard shows the Total Number of Crimes and Total Crimes in the Neighborhood (O'Hare) for the Year (2021) with respect to all Primary Types of Crimes and Days.

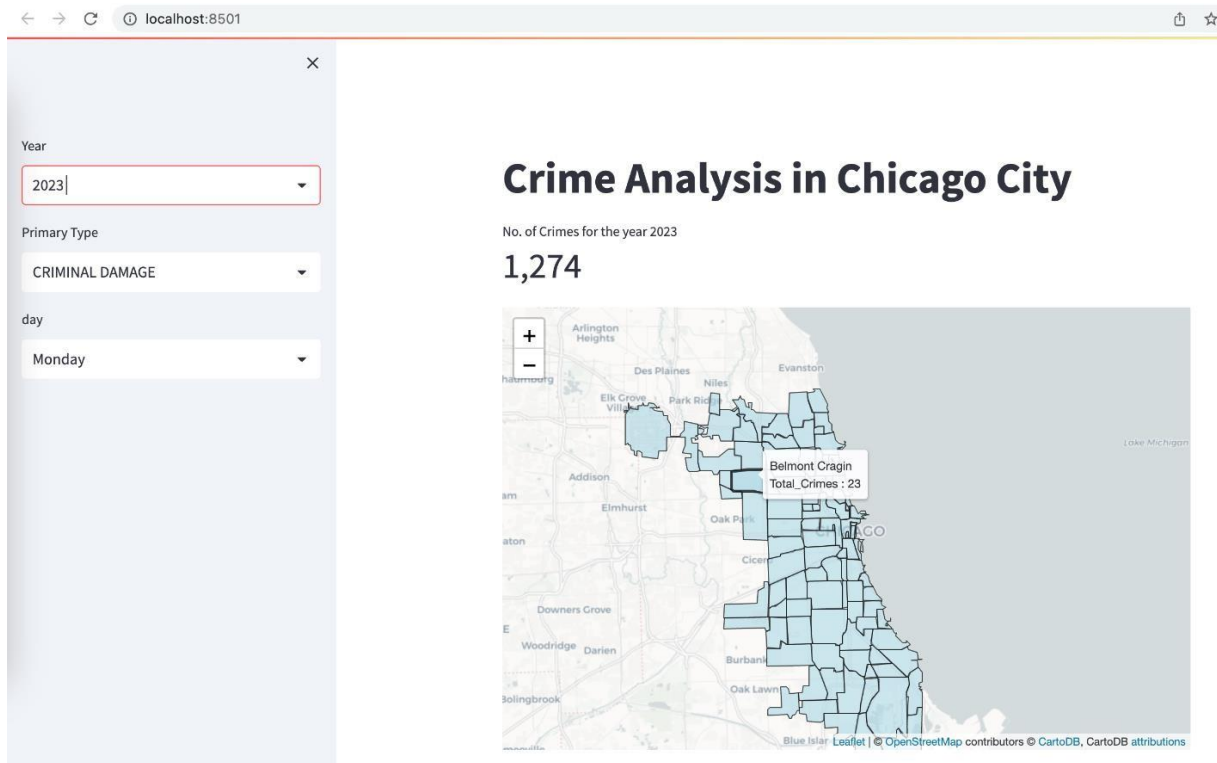


The above Dashboard shows the Total Number of Crimes and Total Crimes in the Neighborhood (O'Hare) for the Year (2021) with respect to all Primary Types of Crimes and the selected Day (Friday).





The above Dashboard shows the Total Number of Crimes and Total Crimes in the Neighborhood (O'Hare) for the Year (2021) with respect to selected Primary Type of Crime (Battery) and for all Days.



The above Dashboard shows the Total Number of Crimes and Total Crimes in the Neighborhood (Belmont Cragin) for the Year (2023) with respect to selected Primary Type of Crime (Criminal Damage) and for the selected Day (Monday).

From the Dashboard, the Neighborhood with the highest number of Crimes observed for all Years was Austin followed by Garfield Park, Humboldt Park, North Lawndale, and others.

## APPLICATION OF CRIME ANALYSIS

- With High-Crime areas, the number of Police Patrols must be increased or deployed into specialized units to address specific types of crime.
- The police can develop Crime Prevention Strategies to address the underlying causes of High Crime and it can include providing education and job training programs to at-risk youth.
- With the use of Crime Analysis, the Police can gather intelligence on criminal networks and organizations that operate in those areas. This includes gathering information on gang activity, drug trafficking, or other forms of organized crime.

- The Crime Analysis can help Police make more informed decisions about where to allocate resources and how to target enforcement efforts. This can help the Police to be more effective in reducing crime and improving public safety.

## **PROJECT LEARNING OUTCOME**

### **1. FOLIUM PACKAGE**

Folium is a Python library used for visualizing geospatial data. It is built on top of the leaflet JavaScript library, which is one of the most popular libraries for creating interactive maps on the web. Folium allows users to create maps with various features such as markers, popups, tooltips, polygons, and heatmaps. It also supports different tile sets such as OpenStreetMap, Mapbox (Limited levels of zoom for free tiles), and Stamen etc.

Folium provides an easy-to-use interface for creating interactive maps directly from Python code. It allows users to create maps with codes and customize the maps using various options and parameters. It also provides functionality to save the maps as HTML files that can be shared and viewed in any web browser.

### **INSTALLATION OF FOLIUM PACKAGE**

- `pip install folium`
- `import folium`

#### **To import necessary folium plugins**

- `from folium.plugins import MarkerCluster`
- `from folium.plugins import HeatMap, Fullscreen, DualMap`

### **2. STREAMLIT PACKAGE**

Streamlit Package is an open-source Python library used for creating interactive web applications. It helps gain knowledge and understand the Streamlit's features, syntax, and capabilities. It converts Python scripts into interactive web apps with ease, utilizing real-time updates, interactive widgets, and data visualization.

### **INSTALLATION OF STREAMLIT PACKAGE**

- `import streamlit as st`
- `from streamlit_folium import st_folium`

## CONCLUSION

This Crime Analysis presents multiple ways to visualize the Data. Data was depicted in an interactive map of Chicago, using folium library to find Hotspots which can help the police make more informed decisions about where to allocate resources. This can help the police to be more effective in reducing crime and improving public safety.

Furthermore, Crime Patterns or Trends were analysed per month and for the various years and there were fluctuations in the number of crimes over time, with some months having higher numbers of crimes than others. Months like Jan, Feb, Mar had lower crime rates due to some factors like weather etc, and higher crime rates for the rest of the months due to holiday seasons.

2021 – 2022 had lower crime rates and some of the Crime Types contributing to these include Burglary, Sexual Assault, Criminal Trespass, Gambling, Interference with Public officer, Kidnapping, etc. The Dashboard was used to filter the total number of Crimes per Year, Days and Primary Types. Selecting the Days and Crime Type shows the Total Number of Crimes for each Neighbourhood.

The City of Chicago has experienced significant challenges in addressing Crime and ensuring the safety of its residents. However, efforts to combat crime have shown some positive impact in recent years.

One significant finding is the decrease in Crime Rates in Chicago. Through various strategies and initiatives implemented by law enforcement agencies, community organizations, and local government, there has been a noticeable decline in Crimes such as Gambling, Criminal Trespass, Liquor Law Violation, etc. This reduction can be attributed to Police Efforts, Improved Community Engagement, and the use of Advanced Technologies for Crime Prevention and Investigation.