**ANALYZING AND PREDICTING CRIME TRENDS IN CHICAGO**

Fan Yang

Big Data Analytic-Fordham University

April 30, 2024

**Introduction**

In the urban landscape of Chicago, the issue of crime is prevalent. The situation in Chicago is prompting a need for a deeper dive and analyzing the patterns to enhance security throughout the city. There is a demand to reveal the dynamics of criminal activity within the city which results in ongoing incidents. The patterns and forces behind the trends in crime are the primary objectives of this study. This project aims to create a model that can predict where crime might happen next, helping law enforcement and policymakers act before it occurs.

**Problem Statement**

In Chicago, the challenge of crime demands a transformation from traditional measures to a more predictive approach. Recent Studies, including those published by MDPI, highlight the significant evolution in the field of predictive policing over the last decade(MDPI). Furthermore, a comprehensive review of predictive methods in law enforcement showcases the importance of predicting crimes, offenders and victims. These advancements underscore the potential to significantly enhance the effectiveness and efficiency of police interventions​ (RAND)​. The urgent need for this project arises from the opportunity to make use of these predictive technologies to reduce crime in Chicago.

**Data source and preparation**

The data for this project comes from the Chicago Police Department’s Citizen Law Enforcement analysis and reporting(CLEAR) system. This dataset reported incidents of Crime in Chicago from 2001 to February 2024 with a total of 7 million records. The dataset includes details such as Date, Type, Description, and location along with others which describe the crime incidents comprehensively. For additional research, supplementary datasets can be explored through platforms such as Kaggle and Wikipedia.

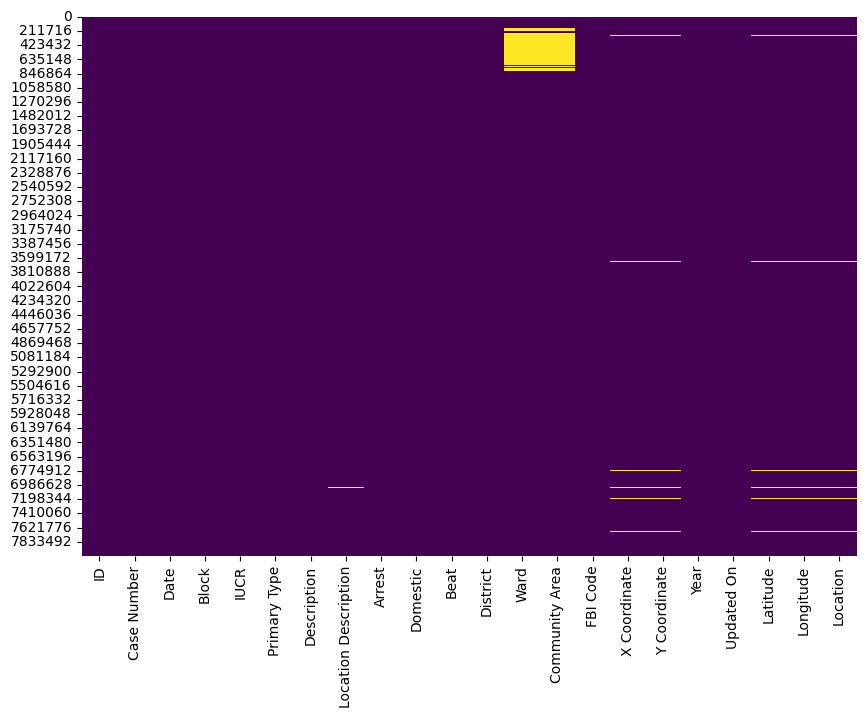
.

Figure 1. Null values count; Heat map analysis

To enhance the dataset’s usability for predictive modeling, the initial step is to identify and address missing values, particularly in critical location-based columns. This was visualized using heatmaps(Figure1). Further data preparation included narrowing down to relevant columns, converting data formats for time series analysis, and normalizing features to ensure data quality. Additionally, feature selection was performed to prioritize relevant variables to enhance model efficiency.

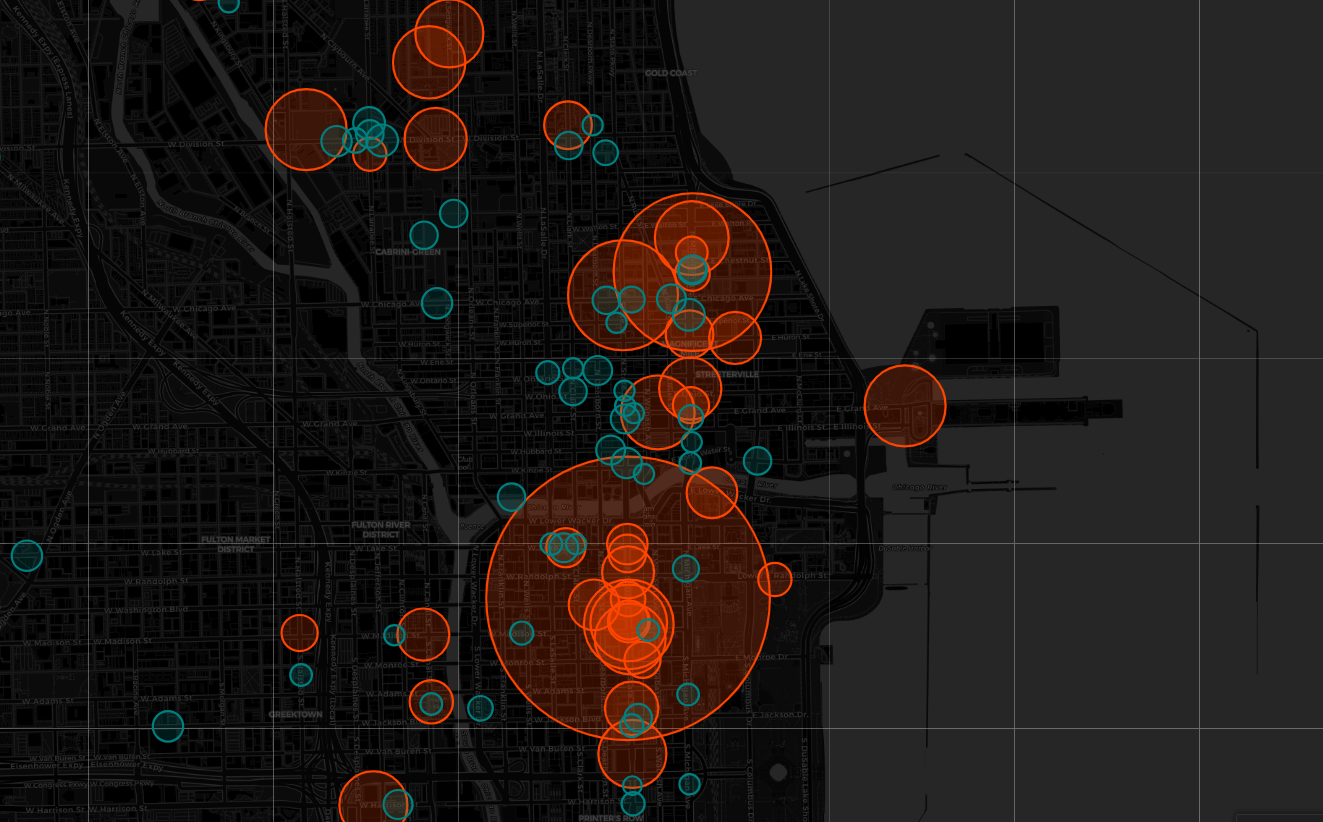
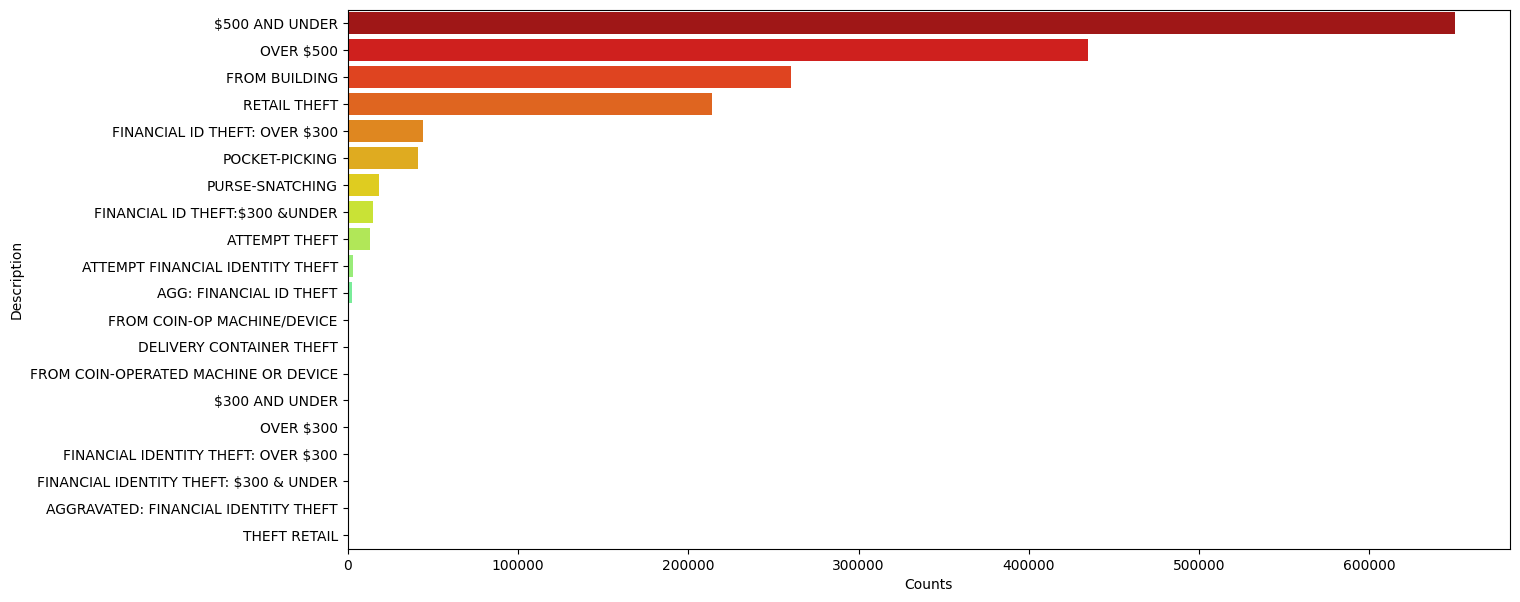
**Exploratory Data analysis**

Figure 2. Geographical analysis on criminal incidents

In the exploratory phase, folium library was employed to construct an interactive map that provides a comprehensive visualization of crime data across Chicago. The map features a dark theme, centered on the city, enhancing the visual contrast and focus on the data points represented.

To interpret the distribution of criminal incidents, over 500 circles were plotted. Each circle’s size was designed to be proportional to the number of recorded criminal incidents at specific locations, allowing for an intuitive grasp of crime density. The color orange-red indicates whether incidents exceed a thousand and teal for those with fewer incidents, providing an immediate visual differentiation, aiding in the quick identification of high-crime areas. Interactive elements were integrated into the map to enhance user engagement.

By clicking on any circle, users can view a popup that displays the latitude, longitude and exact number of incidents at that location, offering detailed insights into specific areas. This analytic approach serves as a critical tool for policymakers and law enforcement agencies to allocate resources more effectively and design targeted interventions.

Figure 3. A bar chart analysis: Theft incident categories

The bar chart(Figure 3)presents a visual comparison of different categories of theft incidents, quantified by the number of cases reported in descending order. The category labeled “$500 AND UNDER” has the highest number of incidents indicating that smaller-scale thefts are the most frequent. With “Over $500” follows, suggesting that high-value thefts are also common. This chart gives a basic

understanding of the distribution of theft categories.

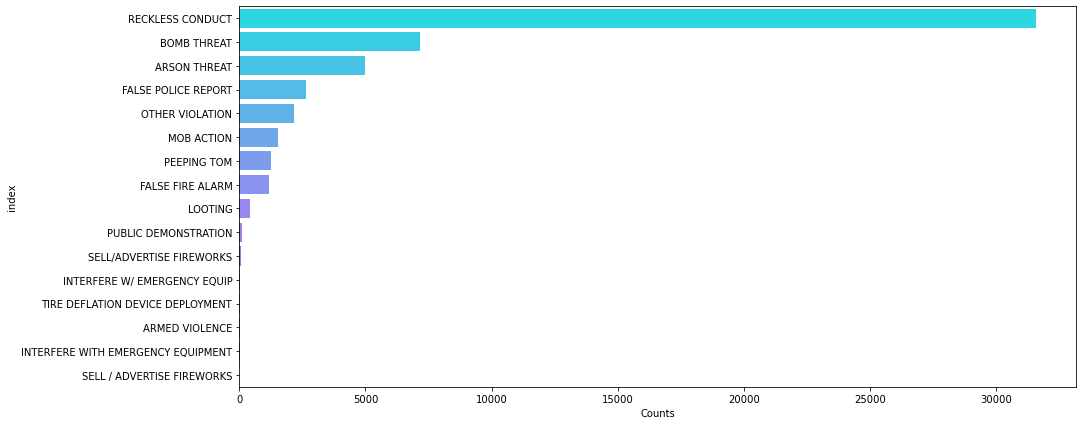


Figure 4. A bar chart analysis: Public Peace Violation

The bar chart(Figure 4)presents a visual comparison of different categories of public peace violation, quantified by the number of cases reported in descending order. The category labeled “RECKLESS CONDUCT” has the highest number of incidents indicating that smaller-scale thefts are the most frequent. With “BOMB THREAT” and “ARSON THREAT”follow, suggesting that bomb and arson threats are also common. This chart gives a basic understanding of the distribution of public peace violations.

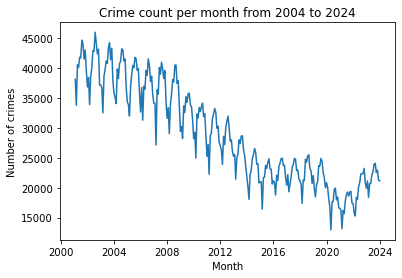
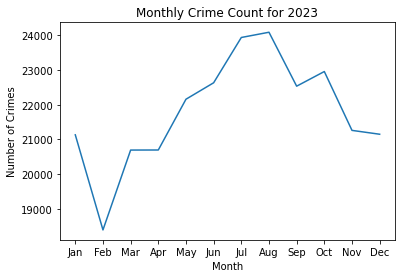


Figure 5.Analysis of monthly and yearly crime trends

The graph (Figure 5)on the left represents the monthly crime trends within 2023, the graph shows significant fluctuations within the year, with crime rates peaking around July and August. Indicating summer is a common period for higher crime rates. Summer brings warmer weather and with school breaks during the summer months, leading to more people spending time outdoors. This increase in social activity and public gatherings can create more opportunities for crimes such as theft, assault, and vandalism.

The graph (Figure 5)on the right represents the yearly crime trends from 2004 to 2024. There's a clear long-term downward trend in crime rates over the years displayed. And there’s a slight recovery seen in recent years. Economic downturns can lead to recent years rise in crime rates. Recovery from COVID-19 pandemic could also possibly impact on crime rates in Chicago.

Together these graphs provide a comprehensive view of the dynamics of crime over different time frames. Understanding these factors helps policymakers and law enforcement agencies to tailor their approaches to crime prevention. Effective measures such as enhancing police visibility during summer.

**Geographical Analysis**

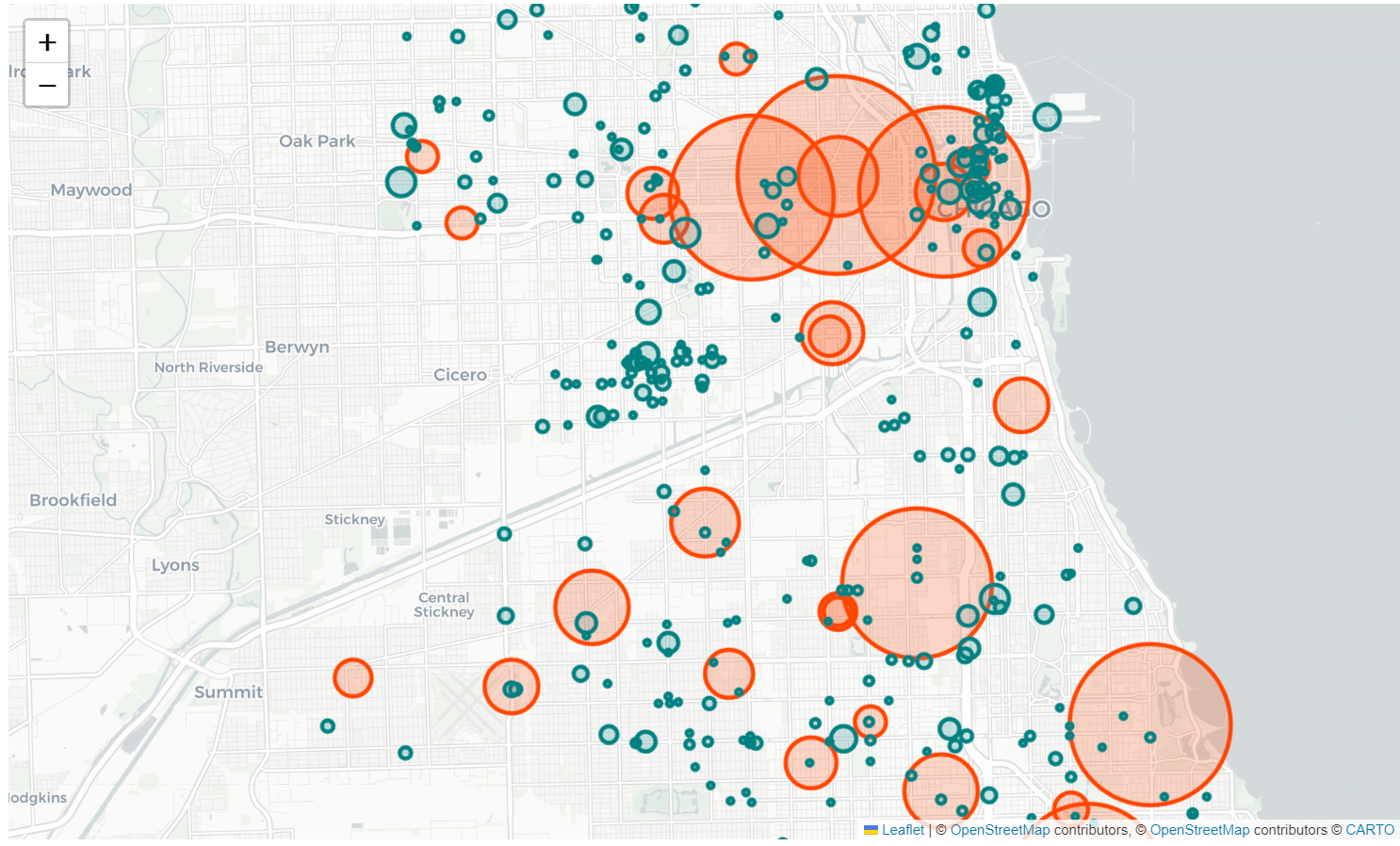
****

Figure 6. Geographical analysis on Bomb threats

To interpret the distribution of Bomb threats, over 500 circles were plotted. Each circle’s size was designed to be proportional to the number of recorded criminal incidents at specific locations, allowing for an intuitive grasp of crime density. The color orange-red indicates whether incidents exceed over 30 and teal for those with fewer incidents, providing an immediate visual differentiation, aiding in the quick identification of high-bomb threat areas. Interactive elements were integrated into the map to enhance user engagement.

By clicking on any circle, users can view a popup that displays the latitude, longitude and exact number of bomb threats at that location, offering detailed insights into specific areas. This analytic approach serves as a critical tool for policymakers and law enforcement agencies to allocate resources more effectively and design targeted interventions.

**Time Series Forecasting**

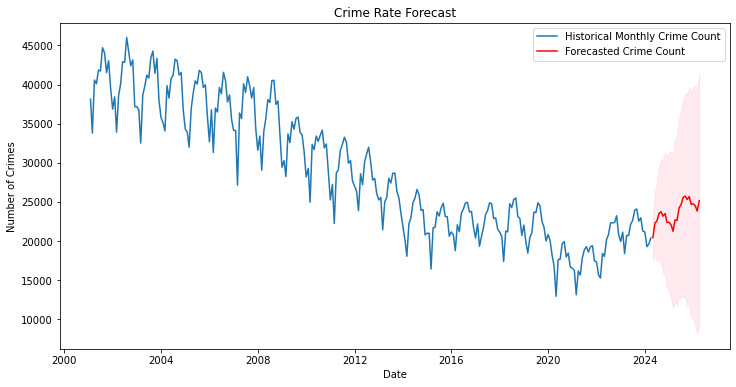
****

Figure 7. Crime count forecasting

Figure 7 provides a valuable representation of the historical and forecasted monthly crime counts, offering a clear picture of past trends and future expectations in crime rates.

The forecasted segment, shown in red, begins where the historical data ends, projecting future trends. The forecasting model predicts a slight increase followed by a stabilization in crime rates. The shaded area around the forecast line represents the confidence interval, providing an estimate of the uncertainty surrounding the forecasts. The forecast suggests an increase in crime counts in future years.

**Insights**

* Resource Allocation using interactive distribution map: by understanding where crimes are most concentrated, the department can deploy additional patrols, set up surveillance and initiate targeted community engagement programs more efficiently.
* Increase police visibility during summer: historical data indicates a spike in crime rates during the summer months. To address this seasonal increase, it is recommended to enhance police visibility during this period. This could involve increasing the number of patrols, particularly in public areas and neighborhoods that the data identifies as high-risk during these months.
* Boost public awareness and police budget in future years: Given the forecasted rise in crime rates in later years, there is a critical need to raise public awareness and allocate a higher budget to the police department. Investing in community outreach programs to educate the public on safety measures and crime prevention can empower citizens and help mitigate the risk of crime. Additionally, adjusting the police budget to accommodate advanced training, better equipment, and more comprehensive community services can prepare the department to handle the potential increase in crime effectively.

**References:**

Perry, Walter L., Brian McInnis, Carter C. Price, Susan Smith, and John S. Hollywood, Predictive Policing: Forecasting Crime for Law Enforcement. Santa Monica, CA: RAND Corporation, 2013. <https://www.rand.org/pubs/research_briefs/RB9735.html>.

Mugari I, Obioha EE. Predictive Policing and Crime Control in The United States of America and Europe: Trends in a Decade of Research and the Future of Predictive Policing. *Social Sciences*. 2021; 10(6):234. <https://doi.org/10.3390/socsci10060234>

City of Chicago. (n.d.). Crimes - 2001 to Present. Retrieved [April 2024], from https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-Present/ijzp-q8t2/data\_preview

Rob Cydzik, Hayley Demetres, Chris Forst, Kyle Freire, Arzoo Khan, Fan Yang