Chicago Crime AnaLysis

Olachi Mbakwe and Shafwat Mustafa

2024-04-07

Chicago Crime Analysis

Olachi Mbakwe and Shafwat Mustafa#1

#College of Information Sciences/Technology, The Pennsylvania State University, United States of America

Orm584@psu.edu and Sbm5973@psu.edu

Abstract—

In recent years, Chicago has been characterized by its high crime rates, particularly in violent offenses such as homicide, assault, and robbery. This study aims to provide a comprehensive analysis of crime in Chicago by examining the interplay between socio-economic factors, geographical distribution, and temporal patterns. Utilizing data from the City of Chicago Data Portal, we employ a combination of Geographic Information Systems (GIS), machine learning algorithms, and statistical analysis to identify significant predictors of crime and pinpoint hotspots of criminal activity. Our research reveals a strong correlation between crime rates and socio-economic indicators such as poverty, unemployment, and educational attainment. Geographical analysis indicates that certain neighborhoods exhibit higher crime rates, potentially due to historical, social, and economic conditions. Temporal patterns suggest seasonal and time-related trends in criminal activity, such as increased theft during holiday seasons and higher rates of violence in summer months. By integrating these findings, our study aims to contribute to the development of more effective public safety strategies and data-driven policy-making, ultimately enhancing community resilience and well-being in Chicago.

Keywords — Chicago, Chicago neighborhood, GIS, XG boost, , Decision Tree, Heat Map, Crime, Theft, Crime Analysis,

I. INTRODUCTION

Chicago, one of the largest cities in the United States, has long been associated with high crime rates.The city's complex urban landscape, characterized by diverse neighborhoods and socio-economic disparities, presents a challenging environment for law enforcement and public safety.The historical context of crime in Chicago, including the notorious era of Al Capone and the Prohibition, has left a lasting impact on the city's collective memory and continues to influence contemporary perceptions of safety and criminal activity.

In recent years, the city has struggled with persistent issues related to violent crimes, such as homicides, assaults, and robberies, which have raised concerns among residents and policymakers alike. The urgency to address these issues is underscored by the human cost of crime, including loss of life, physical and emotional trauma, and the erosion of community trust and cohesion.

The primary goal of this research is to analyze crime data in Chicago to identify patterns, hotspots, and potential predictors of criminal activity. By understanding the underlying factors that contribute to crime, we aim to inform more effective strategies for crime prevention and public safety.

Our study is motivated by several key objectives:

1. `Enhancing Public Safety`: At the heart of our research is the desire to contribute to the safety and well-being of Chicago's residents. By identifying patterns and trends in criminal activity, we hope to provide insights that can guide law enforcement and community efforts to reduce crime.

2. `Efficient Allocation of Resources`: Law enforcement agencies face the constant challenge of allocating limited resources effectively. Our analysis seeks to identify areas of the city that require heightened attention, enabling more targeted and efficient deployment of law enforcement personnel and resources.

3. `Data-Driven Policy Development`: In an era of evidence-based policymaking, our research aims to provide a solid empirical foundation for the development of public safety policies. By understanding the dynamics of crime in Chicago, policymakers can craft strategies that are grounded in data and tailored to the city's unique challenges.

4. `Community Engagement and Empowerment`: Engaging communities in crime prevention and public safety efforts is crucial for long-term success. Our research seeks to empower communities with knowledge about crime trends and factors, fostering a collaborative approach to enhancing safety.

The purpose of this section is to review, discuss, and assess the results of a set of three (3) research papers that are all related to the analysis of crime data in Chicago. These papers focus on various aspects of crime analysis, including spatial distribution, predictive modeling, and the use of Geographic Information Systems (GIS) and visualization tools. We will briefly introduce each of the three (3) research papers to be discussed in this section.

The first research paper, titled "***Crime Analysis in Chicago City***" by Ayidh Alqahtani et al. (2019), explores the design of a Crime Data Information System and the use of data mining techniques, including clustering and spatial mining, to identify crime hotspots and patterns in Chicago's open dataset.

The second paper, ***"Combining the richness of GIS techniques with visualisation tools to better understand the spatial distribution of data – a case study of Chicago City crime analysis***" by Omar Ibrahim Bani-Taha and M. Omair Shafiq (2020), investigates the benefits of adding a spatial GIS layer of analysis to other existing visualization techniques to provide a better understanding of crime patterns and prediction within Chicago.

The third research paper, titled "***Spatial Patterns of Crime and Its Relationship with The Physical Environment: Chicago Case***" by Elif Kırpık, investigates the intricate interplay between crime patterns and physical environmental elements in Chicago. Utilizing exploratory spatial statistical methods and GIS, the study aims to uncover the spatial distribution of crimes and examine the correlation between recurrent crimes and aspects of the physical environment. The research provides valuable insights into how environmental factors such as street lighting, vacant buildings, and sanitation complaints are associated with different types of crimes. By shedding light on these relationships, the study contributes to a deeper understanding of the environmental dimensions of urban crime and offers empirical evidence to support crime prevention strategies that focus on improving the physical environment.

Our fascination with the subject of crime analysis, particularly in the context of Chicago, stems from a deep-seated concern for public safety and a desire to contribute to the well-being of urban communities. The alarming crime rates in Chicago, especially regarding violent offenses, have not only captured my attention but also underscored the urgent need for a comprehensive understanding of crime patterns and effective prevention strategies.

The choice of Chicago as the focal point for this study is driven by the city's notorious crime statistics and the wealth of available data. Chicago's transparency in sharing detailed crime records offers a unique opportunity for an in-depth examination of spatial and temporal trends, demographic influences, and the distribution of various crime types. This rich dataset serves as a cornerstone for developing predictive models and analyzing crime patterns, which can potentially be applied to other cities facing similar challenges.

The interdisciplinary nature of crime analysis, which encompasses elements of criminology, sociology, geography, and data science, further fuels my interest in this research area. The prospect of employing machine learning techniques and statistical methods to uncover insights from complex crime data is particularly intriguing. By leveraging these tools, we can move beyond traditional reactive approaches to a more proactive stance in crime prevention, ultimately enhancing public safety.

Moreover, my personal commitment to improving community well-being motivates me to delve into this research. By exploring innovative approaches to crime analysis and prevention, I aim to contribute to the development of safer, more resilient urban environments. The potential to make a tangible impact on public safety and quality of life in Chicago and beyond is a driving force behind my interest in this topic.

In summary, my interest in crime analysis in Chicago is rooted in a desire to address pressing public safety concerns, leverage the extensive crime data available, and apply interdisciplinary methods to develop effective solutions. This research endeavor aligns with my aspirations to contribute to meaningful advancements in urban safety and community well-being

II. LITERATURE REVIEW —

In this Literature Review section, we will delve deeper into each of the three selected research papers, expanding upon our initial introduction to explore their relevance to the theme of crime analysis in Chicago. This section aims to dissect the methodologies, findings, and implications of each study, providing a comprehensive understanding of how these investigations contribute to our knowledge of crime patterns, hotspots, and the potential predictors within urban environments. Additionally, we will examine how the insights derived from these studies could enhance public safety initiatives, inform law enforcement strategies, and ultimately contribute to the reduction of crime in urban settings.

[1] ***Crime Analysis in Chicago CityCrime Analysis in Chicago City"*** by Ayidh Alqahtani, Ajwani Garima, and Ahmad Alaaid expands on the concept of utilizing data mining techniques for crime analysis. This paper explores the design of a Crime Data Information System capable of processing large datasets to identify crime hotspots and predict future incidents. The authors' approach to data preprocessing and the employment of clustering techniques underscore the significance of sophisticated data analysis in unveiling crime patterns.

The dual methodologies of K-means clustering and spatial mining adopted in this study underscore the utility of diverse analytical tools in crime analysis. By integrating traditional data mining with spatial clustering, the research offers a comprehensive view of crime hotspots, facilitating a more effective deployment of law enforcement resources. The comparative analysis of clustering results with ground truth data further validates the efficacy of these methods in identifying high-risk areas and informing preventive strategies.

[2] ***Combining GIS Techniques with Visualisation Tools for Crime Analysis***

Building upon the spatial analysis foundation laid by Kırpık, the second paper we review expands the toolkit for crime analysis by integrating advanced GIS techniques with visualization tools. This study demonstrates how combining these technologies can offer deeper insights into the spatial distribution of crime, enhancing the ability to identify hotspots and patterns that might not be immediately apparent through traditional analysis methods.

The incorporation of visualization tools allows for a more intuitive understanding of crime data, making it accessible not only to researchers but also to policymakers and law enforcement officials. This approach facilitates the identification of trends over time and across different types of crimes, enabling a more dynamic response to emerging threats. The paper's methodology, emphasizing the synergy between GIS and visualization technologies, represents a significant advancement in the field of crime analysis, offering new avenues for developing predictive models and strategic interventions.

[**3**] ***Spatial Patterns of Crime and Its Relationship with The Physical Environment: Chicago Case***

The foundational paper in our literature review, authored by Elif Kırpık, serves as the cornerstone for our exploration of crime analysis in Chicago. Kırpık's study, employing exploratory spatial statistics and GIS technologies, ventures into the intricate relationship between crime occurrences and physical environmental elements in Chicago. The research leverages comprehensive crime data to examine spatial patterns, revealing how certain physical characteristics—such as poor street lighting and the prevalence of vacant buildings—can significantly influence crime rates in particular neighborhoods.

This study's methodology, focusing on spatial analysis and the incorporation of environmental criminology theories, sets a precedent for our research. By identifying spatial concentrations of crime and correlating them with specific environmental factors, Kırpık's work underscores the potential of environmental modifications in crime prevention strategies. The implications for public safety, particularly in urban areas akin to Chicago with similar socio-economic complexities, are profound. The paper's findings highlight the necessity for targeted interventions and urban planning efforts that address the root causes of crime from an environmental perspective.

Our Study Design and Methods

To conduct our analysis of crime data in Chicago, we will employ a combination of Geographic Information Systems (GIS), machine learning algorithms, and statistical analysis. Our study design is structured as follows:

1. `Data Collection`: We will gather crime data from the City of Chicago Data Portal, which provides detailed records of reported crimes across the city. This dataset includes information on the type of crime, location, date, and time, among other attributes.

2. `Data Preprocessing`: Before analysis, we will preprocess the data to ensure its quality and usability. This includes handling missing values, filtering out irrelevant data, and converting data formats as needed.

3. `Exploratory Data Analysis (EDA)`: We will conduct an initial exploration of the data to identify patterns, trends, and anomalies. This step involves visualizing the data using histograms, scatter plots, and heatmaps to gain insights into the distribution and frequency of crimes.

4. `Geographical Analysis`: Using GIS, we will map the crime data to visualize the spatial distribution of crimes across Chicago. This will help us identify hotspots and areas with higher crime rates.

5. `Machine Learning Models`: We will apply machine learning algorithms, specifically XGBoost and Random Forest, to predict crime occurrences and identify significant predictors. These models will be trained and tested on the dataset to evaluate their performance.

6. `Evaluation`: The performance of the machine learning models will be assessed using appropriate metrics such as accuracy, precision, recall, and F1-score. We will also compare the results of XGBoost and Random Forest to determine the most effective algorithm for crime prediction.

About the Dataset

For our study, we are using the "Crimes - 2023" dataset from the City of Chicago Data Portal. This dataset contains detailed records of reported crimes in Chicago for the year 2023. The dataset includes the following attributes:

- **ID:** A unique identifier for each crime report.
- **Case Number:** A unique number assigned to each case.
- **Date and Time:** The date and time when the crime occurred.
- **Block:** The block where the crime occurred.
- **IUCR (Illinois Uniform Crime Reporting code):** A code that classifies the reported crime according to the Illinois Uniform Crime Reporting standards.
- **Primary Type**: The primary classification of the crime (e.g., theft, assault, robbery). Description: A detailed description of the crime.
- **Location Description:** A description of the location where the crime occurred (e.g., street, residence, school).
- **Arrest:** Indicates whether an arrest was made.
- **Domestic:** Indicates whether the crime was domestic-related.
- **Beat:** A geographic area within the Chicago Police Department's patrol system.
- **District:** The police district where the crime occurred.
- **Ward:** The city ward where the crime occurred.
- **Community Area:** The community area where the crime occurred.
- **FBI Code:** A code used by the FBI to classify crimes.
- **Coordinates:** The latitude and longitude coordinates of the crime location.
- **Year:** The year of the crime report.
- **Updated On:** The date when the record was last updated.

Our initial exploration of the data set revealed that it contains 197,683 records, providing a comprehensive overview of crime incidents in Chicago for the year 2023. This dataset will serve as the foundation for our analysis and modeling efforts.

During the preprocessing step, we identified and removed entries with missing data. This resulted in the removal of 5,878 records from the dataset, ensuring that our analysis is based on complete and accurate information.

Exploratory Data Analysis (EDA)

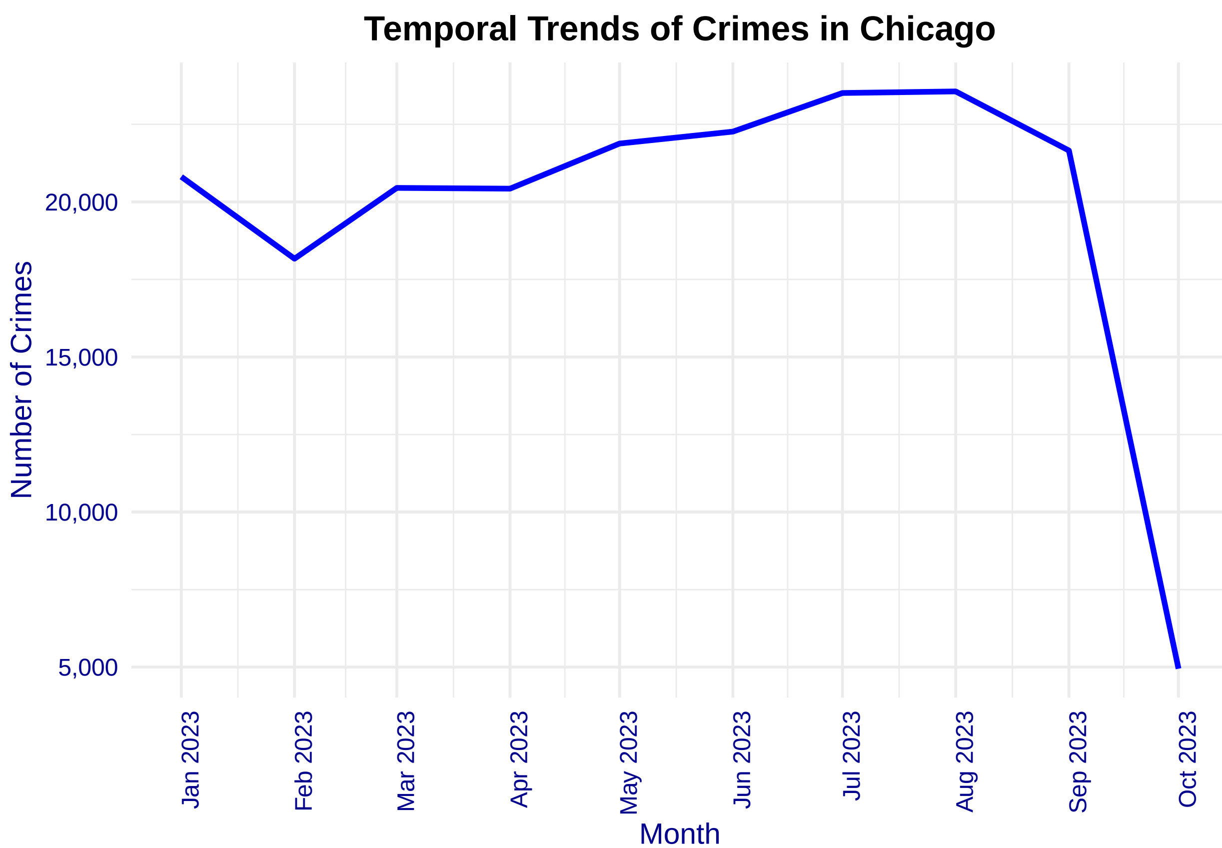
Exploration 1: Crime Type Distribution The table below presents a count of reported incidents across different crime types. By calculating the frequency of each Primary Type of crime and arranging them in descending order, we are able to discern the most common crimes reported in the dataset. This frequency distribution is crucial for understanding which crime types are most prevalent in the community and may require increased attention from law enforcement and public policy makers.

According to the table, the most common crime is Theft, with a staggering 42,425 reported cases. This is followed by Battery with 34,136 cases, reflecting a significant number of violent incidents. Criminal Damage is the third most frequent crime, consisting of 23,002 cases, indicating a considerable amount of property-related offenses. Motor Vehicle Theft and Assault round out the top five, with 22,649 and 17,394 cases, respectively.

The prevalence of theft and battery as the top crimes may indicate underlying social issues, such as economic disparity or a need for enhanced community programs and police presence.

Primary Type	Count
THEFT	42425
BATTERY	34136
CRIMINAL DAMAGE	23002
MOTOR VEHICLE THEFT	22649
ASSAULT	17394
OTHER OFFENSE	11799
DECEPTIVE PRACTICE	11398
ROBBERY	8105
WEAPONS VIOLATION	6879
BURGLARY	5614
NARCOTICS	3832
CRIMINAL TRESPASS	3588
OFFENSE INVOLVING CHILDREN	1377
CRIMINAL SEXUAL ASSAULT	1226
SEX OFFENSE	1035
PUBLIC PEACE VIOLATION	662
HOMICIDE	483
INTERFERENCE WITH PUBLIC OFFICER	450
ARSON	390
STALKING	389
PROSTITUTION	194
INTIMIDATION	190
CONCEALED CARRY LICENSE VIOLATION	160
LIQUOR LAW VIOLATION	141
KIDNAPPING	114
OBSCENITY	27
GAMBLING	11
HUMAN TRAFFICKING	6
PUBLIC INDECENCY	3
NON-CRIMINAL	2
OTHER NARCOTIC VIOLATION	2

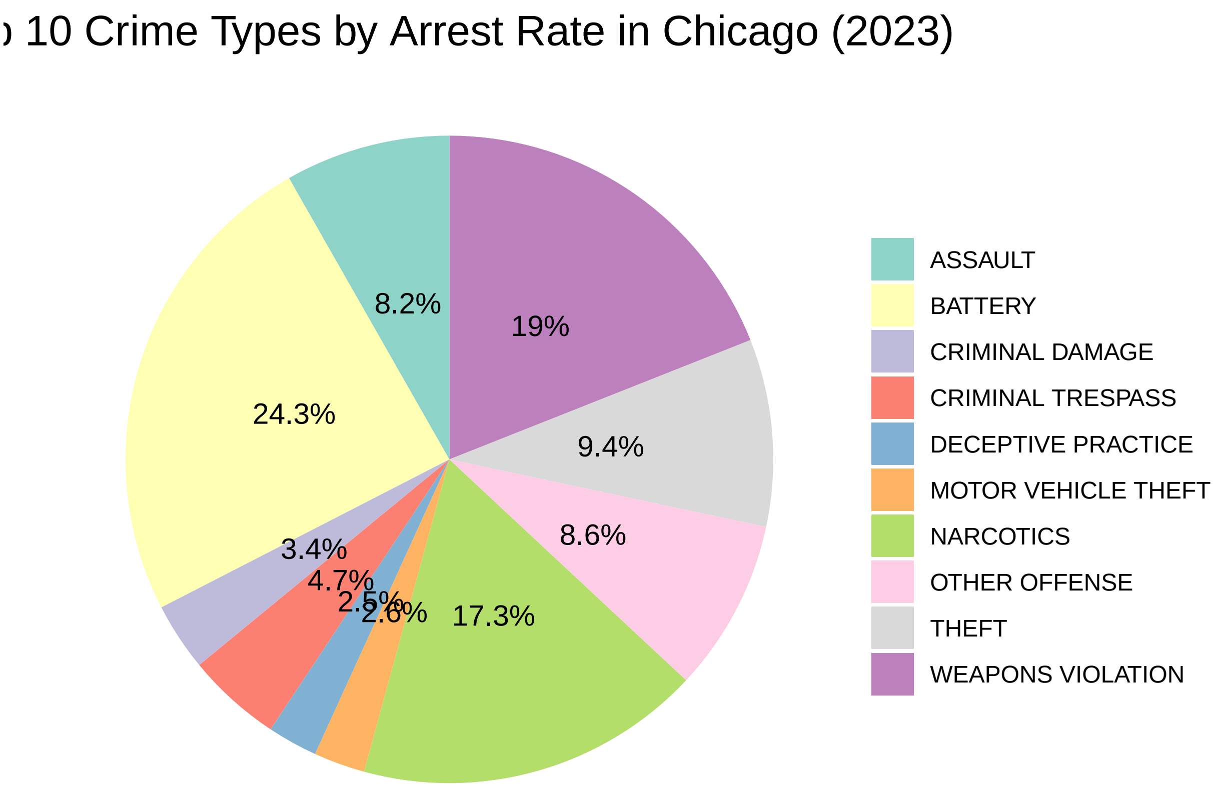
Exploration 2: Trend Analysis of Crime Occurrences In this segment, we delve into the temporal dynamics of criminal activity within Chicago, presenting a visualization of how crime trends fluctuate over the course of months. The line graph displayed elucidates the monthly progression of crime counts, offering stakeholders a visual narrative of the crime rate's ebb and flow within the city.



From the graph, we observe a steady trend in crime reports with subtle fluctuations across the months. However, there is a noticeable and sharp decline in reports as of October 2023. This abrupt drop prompts several questions: Is this a result of a new policy implementation, a significant city-wide event, a change in reporting mechanisms, or perhaps data yet to be reported for the month? Further investigation into these external factors would be necessary to contextualize this dramatic change in the data.

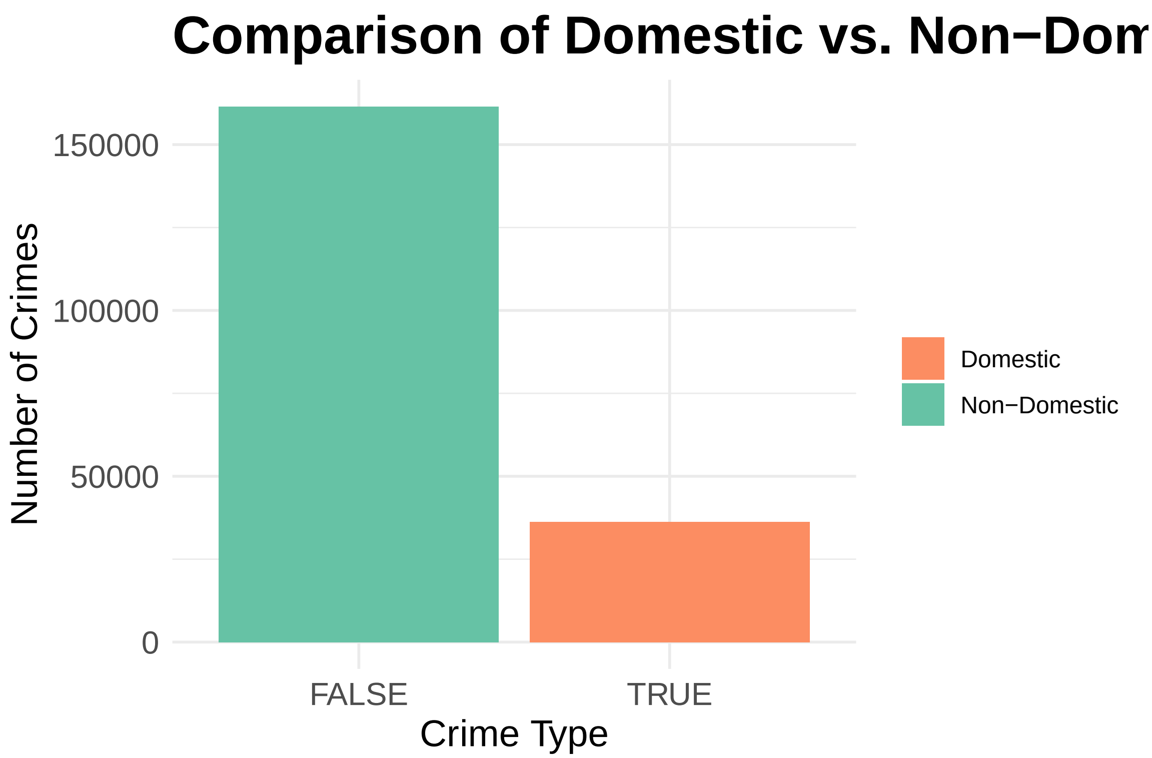
Exploration 3: Arrest Analysis by Crime Type This exploration focuses on dissecting the distribution of arrests across various crime categories. By examining the arrest data, we can infer which crimes have higher enforcement rates, which may reflect the severity of the crimes or the allocation of law enforcement resources.

Crime Type	Number of Arrests	Percentage of Total Arrests
BATTERY	5143	24.32
WEAPONS VIOLATION	4017	18.99
NARCOTICS	3658	17.30
THEFT	1978	9.35
OTHER OFFENSE	1820	8.61
ASSAULT	1743	8.24
CRIMINAL TRESPASS	1003	4.74
CRIMINAL DAMAGE	718	3.39
MOTOR VEHICLE THEFT	542	2.56
DECEPTIVE PRACTICE	528	2.50



From the exploration, we notice that 'Battery' has the highest number of arrests, followed by 'Weapons Violation' and 'Narcotics.' These categories suggest a strong law enforcement focus on violent crimes and drug-related offenses. The percentage column further illustrates that battery alone accounts for over a quarter of all arrests, indicating a significant issue with violent crime in the city.

Exploration 4: Domestic & Non-Domesti In this exploration, we are visualizing the frequency of domestic crimes versus non-domestic crimes in Chicago. The analysis will help to highlight the proportion of crime incidents that are reported as domestic-related.

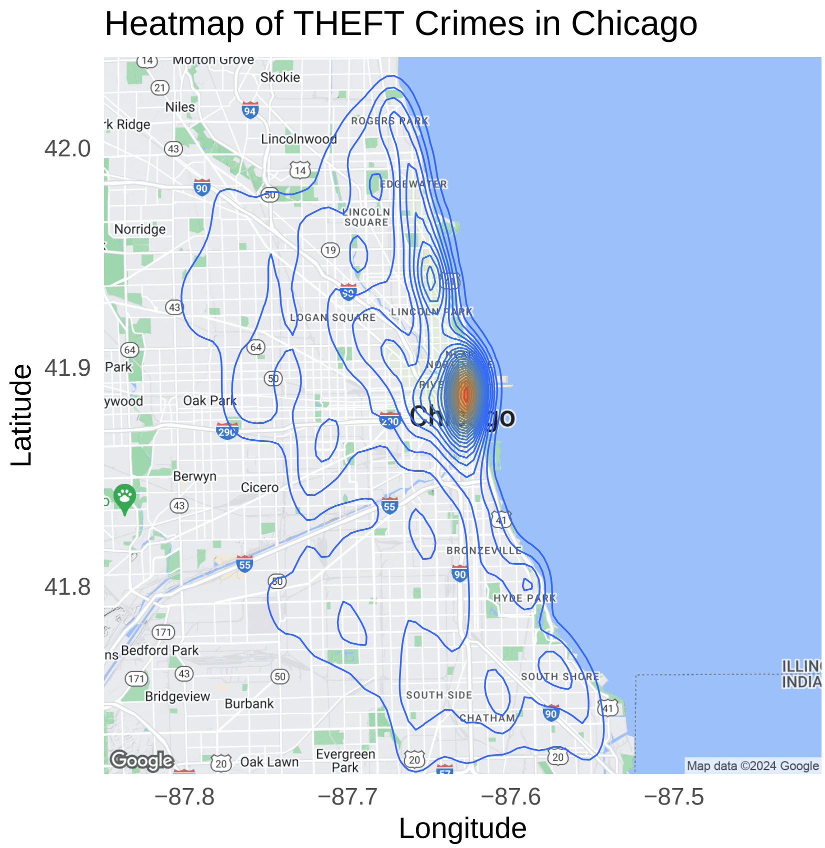


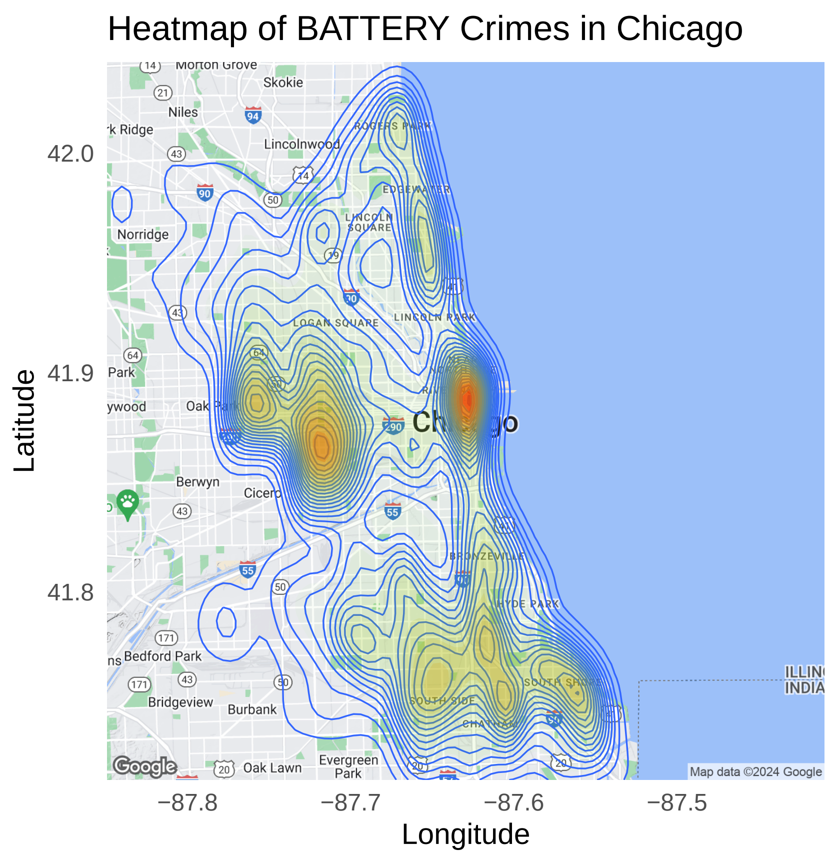
The bar chart reveals a substantial difference between the number of domestic and non-domestic crimes. Non-domestic crimes appear to be more frequent, with a count significantly higher than their domestic counterparts, as evidenced by the height of the bars.

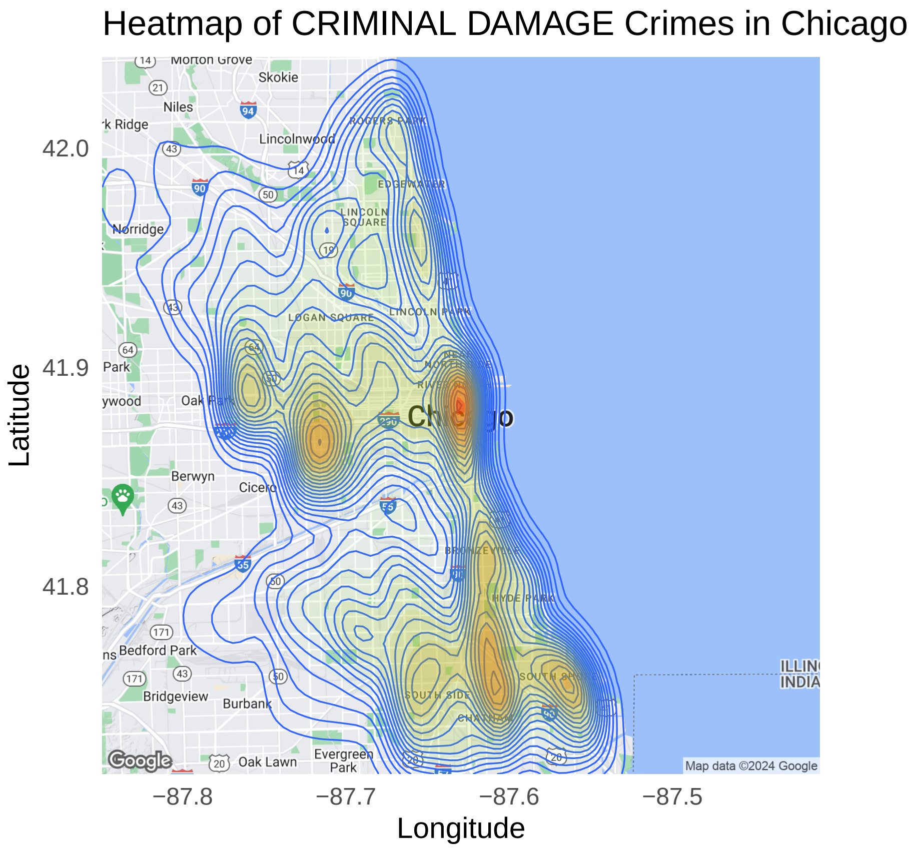
The visualization communicates the importance of recognizing domestic violence as a critical issue and may suggest the need for targeted interventions or resources to support victims of domestic crimes. Moreover, the comparison emphasizes the extent of non-domestic crimes, which encompasses a broad spectrum of offenses, reflecting the overall crime landscape in Chicago.

Geographical Analysis

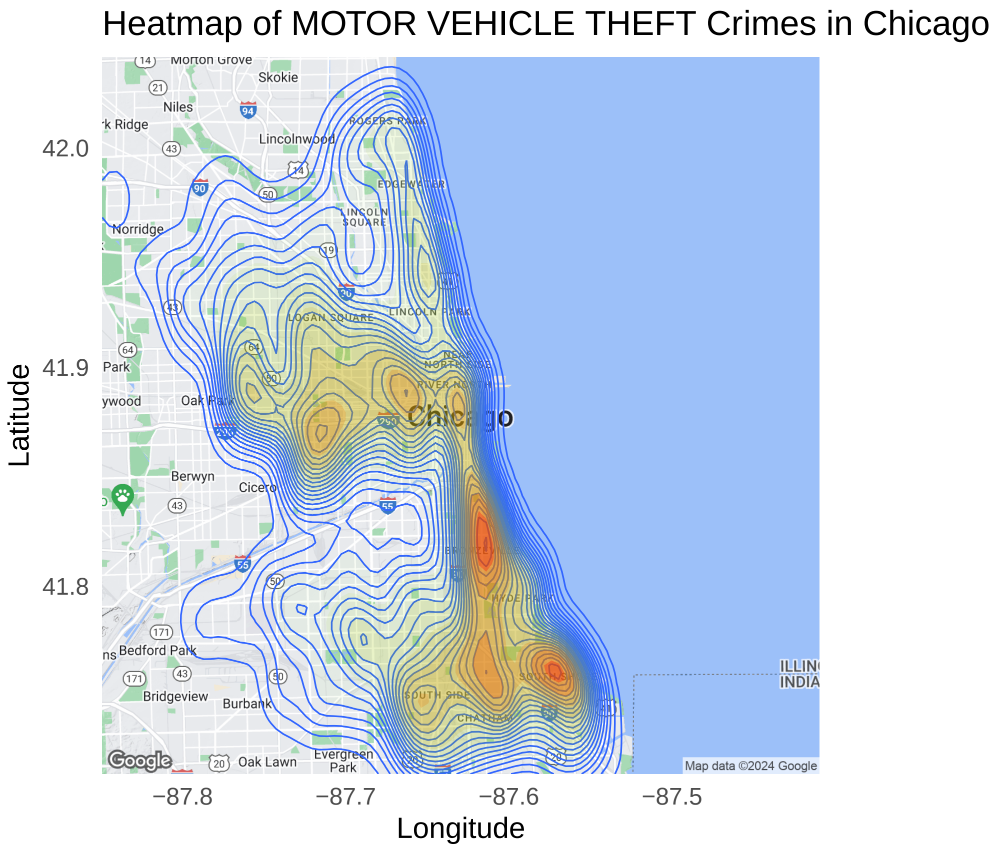
This section of our report provides a visual representation of crime density across different regions of Chicago, emphasizing the spatial distribution of the top five crime types. This analysis is pivotal for understanding which areas are most affected by specific crime types and can inform strategic decisions related to law enforcement and public safety.

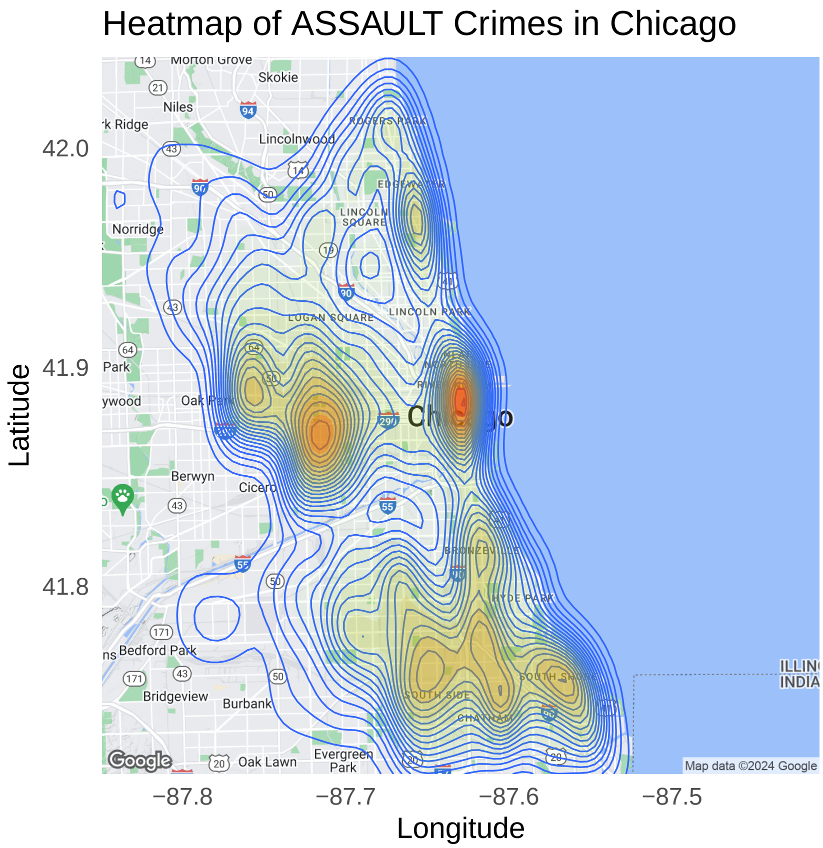

Heatmap of THEFT Crimes in Chicago
Latitude
42.0
41.9
41.8
Longitude
−87.8
−87.7
−87.6
−87.5
Morton Grove
Skokie
Niles
Lincolnwood
Norridge
Oak Park
Berwyn
Cicero
Bedford Park
Bridgeview
Burbank
Oak Lawn
Evergreen Park
ROGERS PARK
EDGEWATER
LINCOLN SQUARE
LOGAN SQUARE
LINCOLN PARK
Chicago
BRONZEVILLE
HYDE PARK
SOUTH SHORE
SOUTH SIDE
CHATHAM
Map data ©2024 Google

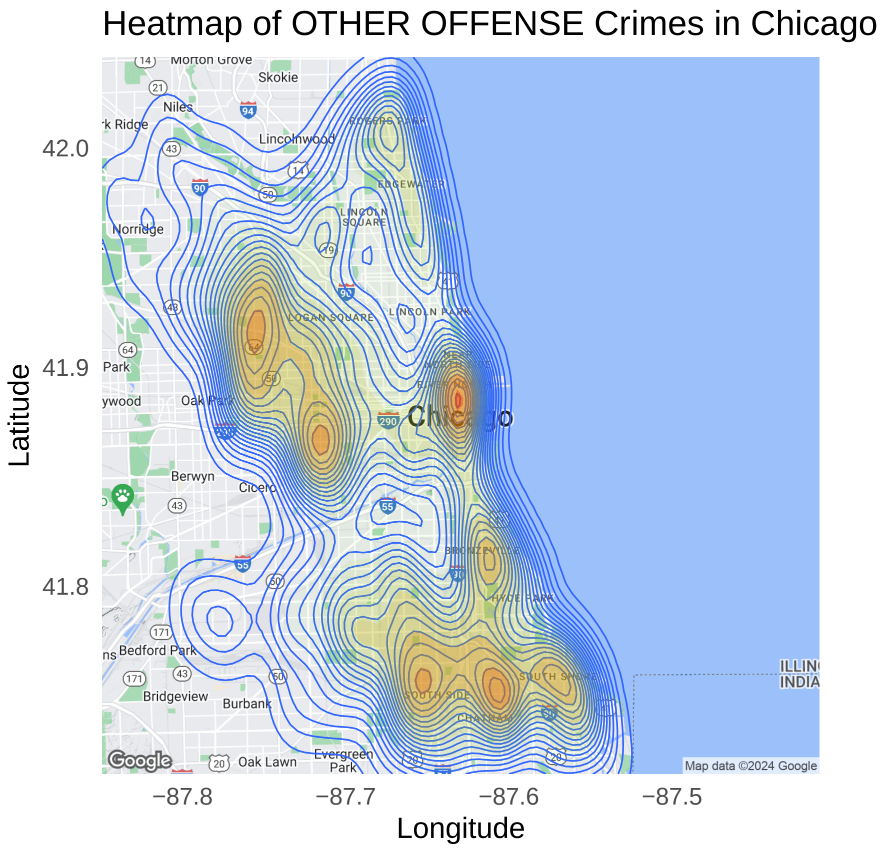

Heatmap of BATTERY Crimes in Chicago
Latitude
42.0
41.9
41.8
Longitude
−87.8
−87.7
−87.6
−87.5
Morton Grove
Skokie
Niles
Lincolnwood
Norridge
Berwyn
Cicero
Bedford Park
Bridgeview
Burbank
Oak Lawn
Evergreen Park
ROGERS PARK
EDGEWATER
LOGAN SQUARE
LINCOLN PARK
BRONZEVILLE
HYDE PARK
SOUTH SIDE
SOUTH SHORE
CHATHAM
Map data ©2024 Google

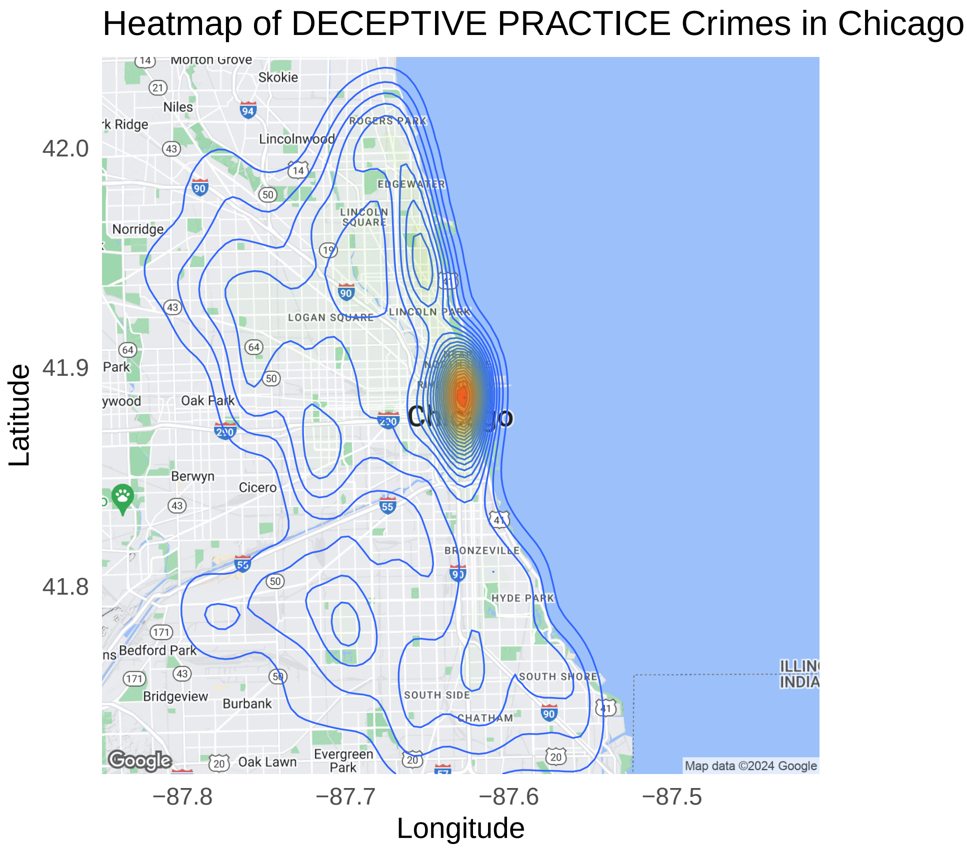


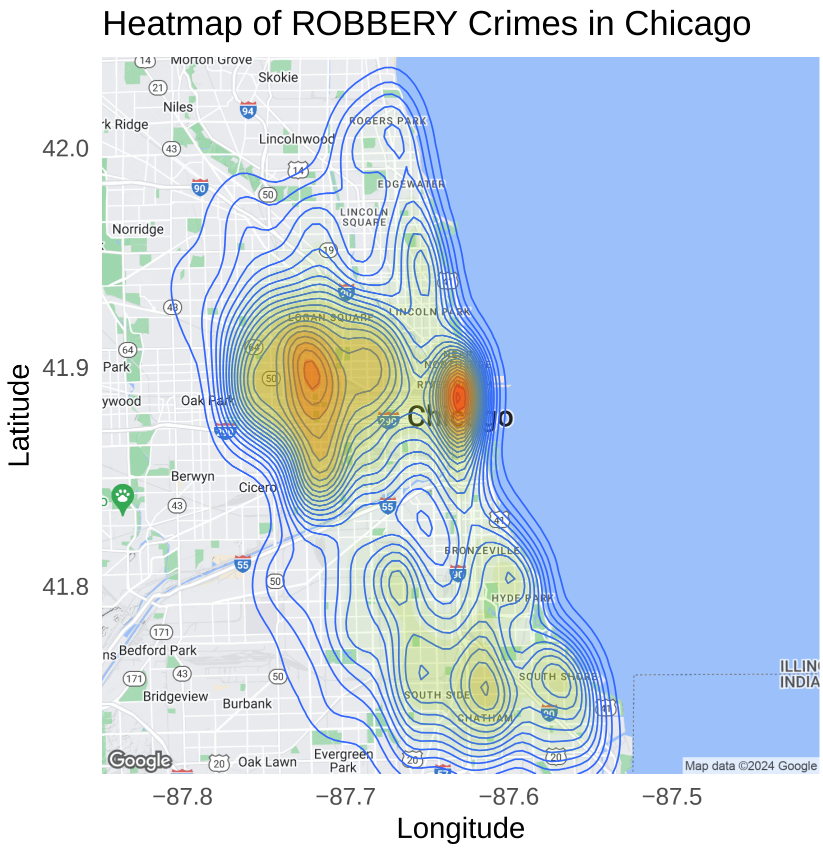
Heatmap of CRIMINAL DAMAGE Crimes in Chicago
Latitude
42.0
41.9
41.8
Longitude
−87.8
−87.7
−87.6
−87.5
Chicago
Skokie
Niles
Lincolnwood
Norridge
Oak Park
Berwyn
Cicero
Bedford Park
Bridgeview
Burbank
Oak Lawn
Evergreen Park
ROGERS PARK
EDGEWATER
LINCOLN SQUARE
LOGAN SQUARE
LINCOLN PARK
BRONZEVILLE
HYDE PARK
SOUTH SIDE
SOUTH SHORE
CHATHAM
Map data ©2024 Google

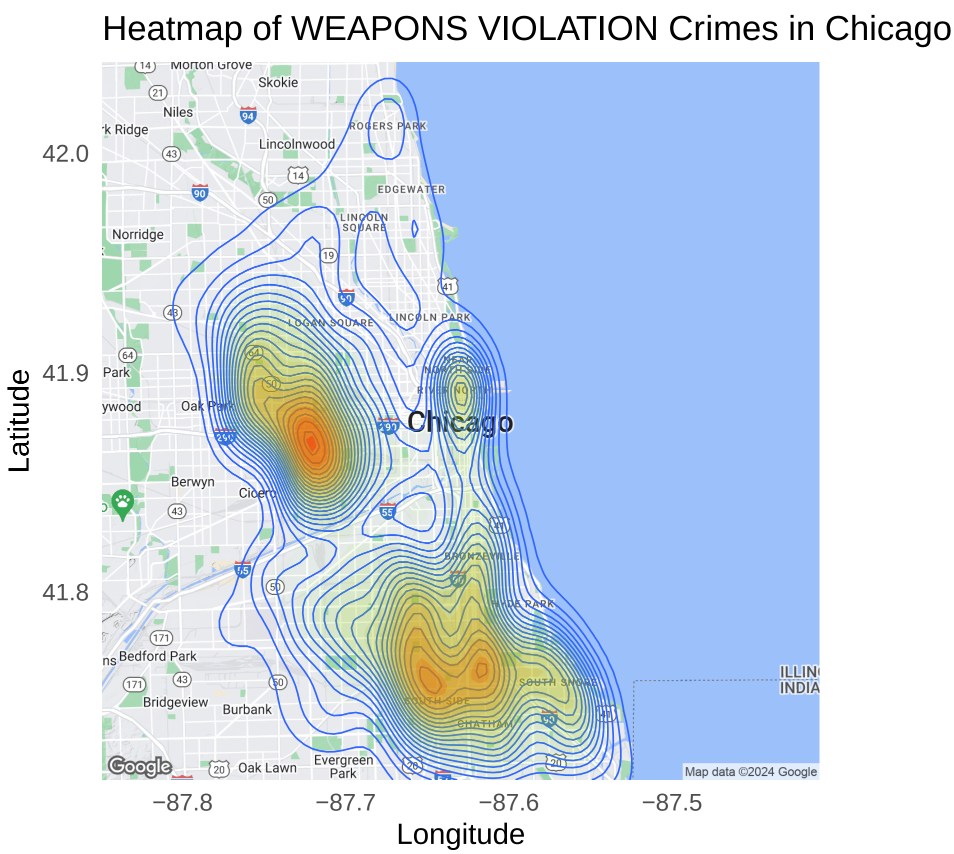

Heatmap of MOTOR VEHICLE THEFT Crimes in Chicago
Latitude
42.0
41.9
41.8
Longitude
−87.8
−87.7
−87.6
−87.5
Morton Grove
Skokie
Niles
Lincolnwood
Norridge
Oak Park
Berwyn
Cicero
Chicago
Bedford Park
Bridgeview
Burbank
Oak Lawn
Evergreen Park
ROGERS PARK
EDGEWATER
LINCOLN SQUARE
LOGAN SQUARE
LINCOLN PARK
HYDE PARK
SOUTH SIDE
CHATHAM
Map data ©2024 Google

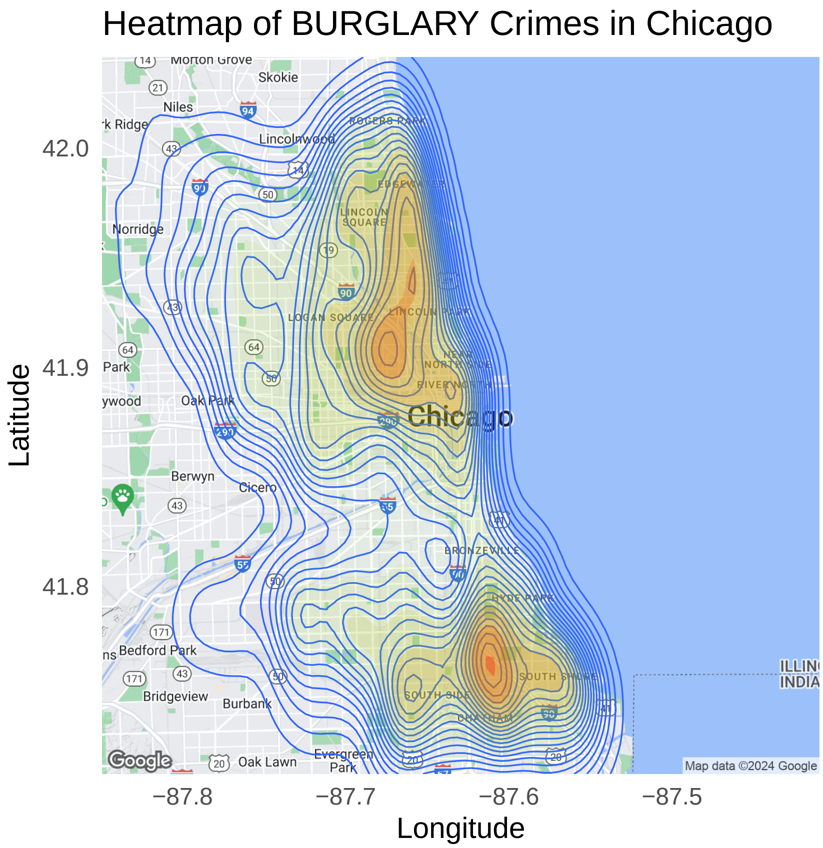

Heatmap of ASSAULT Crimes in Chicago
Latitude
42.0
41.9
41.8
Longitude
−87.8
−87.7
−87.6
−87.5
Skokie
Niles
Lincolnwood
Norridge
Oak Park
Berwyn
Cicero
Bedford Park
Bridgeview
Burbank
Oak Lawn
Evergreen Park
Chicago
Map data ©2024 Google


Heatmap of OTHER OFFENSE Crimes in Chicago
Latitude
42.0
41.9
41.8
Longitude
−87.8
−87.7
−87.6
−87.5
Chicago
Map data ©2024 Google


Heatmap of DECEPTIVE PRACTICE Crimes in Chicago
Latitude
42.0
41.9
41.8
−87.8
−87.7
−87.6
−87.5
Longitude
Skokie
Niles
Lincolnwood
Norridge
Oak Park
Berwyn
Cicero
Bedford Park
Bridgeview
Burbank
Oak Lawn
Evergreen Park
ROGERS PARK
EDGEWATER
LINCOLN SQUARE
LOGAN SQUARE
LINCOLN PARK
BRONZEVILLE
HYDE PARK
SOUTH SHORE
SOUTH SIDE
CHATHAM
Map data ©2024 Google


Heatmap of ROBBERY Crimes in Chicago
Latitude
42.0
41.9
41.8
Longitude
−87.8
−87.7
−87.6
−87.5
Morton Grove
Skokie
Niles
Lincolnwood
ROGERS PARK
EDGEWATER
LINCOLN SQUARE
Norridge
LOGAN SQUARE
LINCOLN PARK
Oak Park
Chicago
Berwyn
Cicero
BRONZEVILLE
HYDE PARK
Bedford Park
Bridgeview
Burbank
SOUTH SIDE
CHATHAM
SOUTH SHORE
Oak Lawn
Evergreen Park
Map data ©2024 Google


Heatmap of WEAPONS VIOLATION Crimes in Chicago
Latitude
42.0
41.9
41.8
Longitude
−87.8
−87.7
−87.6
−87.5
Chicago
Skokie
Niles
Lincolnwood
Norridge
Oak Park
Berwyn
Bedford Park
Bridgeview
Burbank
Oak Lawn
Evergreen Park
Map data ©2024 Google



For each of the top crime categories, a density plot is generated, applying color gradients from green (lower density) to red (higher density) to depict areas with varying levels of crime concentration. This visualization technique allows for the immediate identification of crime hotspots within the city. The use of transparency (alpha blending) ensures that areas with overlapping crime types can be visualized effectively, showing the compounding effects of multiple crime densities.

Machine Learning Models

RandomForest The Random Forest model is trained with several predictors including Arrest, Domestic, Beat, District, Ward, Latitude, Longitude, TimeOfDay, Weekday, and Month. The ntree parameter is set to 100, indicating that 100 trees should be generated in the random forest.

After training, the model is used to predict crime types on the test dataset, and the predictions are compared against the actual crime types to construct a confusion matrix. The confusion matrix is a critical tool for evaluating the performance of a classification model; it shows the number of correct and incorrect predictions made by the model, categorized by type.

To provide a holistic view of the model's performance, several metrics are computed:

`Accuracy`: The overall correctness of the model, calculated as the number of correct predictions divided by the total number of predictions. `Precision`: The ratio of correctly predicted positive observations to the total predicted positives for each crime type. `Recall`: The ratio of correctly predicted positive observations to all actual positives for each crime type. `F1 Score`: The harmonic mean of precision and recall, providing a balance between the two for each crime type

```
## Accuracy:  0.3306129
```

```
## Mean Precision:  0.2207435
```

```
## Mean Recall:  0.1507919
```

```
## F1 Score:  0.1919534
```

The computed accuracy of 25.50% indicates that approximately one out of every four crimes is classified correctly. Precision, recall, and F1 score metrics are averaged across crime types to provide a summary of the model's performance. These metrics indicate how well the model is at identifying each crime type, with the mean F1 score providing an overall measure of the model's accuracy and robustness against imbalance in the dataset.

It is evident that while the model has predictive capability, there is substantial room for improvement. The relatively low values in precision, recall, and F1 score suggest that the model struggles to accurately classify all the different types of crimes.

XG BOOST Our XGBoost model targets the binary classification problem of predicting arrests based on crime features. By selecting pertinent variables such as Primary Type, Beat, District, Ward, Community Area, Latitude, Longitude, and Arrest, and transforming them into numerical formats, the model is trained to understand the underlying patterns that might influence an arrest's occurrence.

The data is randomly partitioned into training (80%) and testing (20%) sets to evaluate the model's predictive power accurately.

```
## [1]  eval-auc:0.783548
## Will train until eval_auc hasn't improved in 10 rounds.
##
## [2]  eval-auc:0.792463
## [3]  eval-auc:0.794270
## [4]  eval-auc:0.807695
## [5]  eval-auc:0.809053
## [6]  eval-auc:0.810883
## [7]  eval-auc:0.812713
## [8]  eval-auc:0.815138
## [9]  eval-auc:0.827262
## [10] eval-auc:0.828880
## [11] eval-auc:0.835592
## [12] eval-auc:0.838032
## [13] eval-auc:0.839254
## [14] eval-auc:0.840684
## [15] eval-auc:0.842585
## [16] eval-auc:0.844112
## [17] eval-auc:0.844962
## [18] eval-auc:0.845849
## [19] eval-auc:0.848264
## [20] eval-auc:0.850228
## [21] eval-auc:0.851146
## [22] eval-auc:0.851552
## [23] eval-auc:0.852525
## [24] eval-auc:0.852605
## [25] eval-auc:0.853661
## [26] eval-auc:0.853808
## [27] eval-auc:0.855054
```

```
## [28] eval-auc:0.855938
## [29] eval-auc:0.857469
## [30] eval-auc:0.857927
## [31] eval-auc:0.858612
## [32] eval-auc:0.859264
## [33] eval-auc:0.859966
## [34] eval-auc:0.860466
## [35] eval-auc:0.861002
## [36] eval-auc:0.861743
## [37] eval-auc:0.862504
## [38] eval-auc:0.863156
## [39] eval-auc:0.864014
## [40] eval-auc:0.864760
## [41] eval-auc:0.865097
## [42] eval-auc:0.865644
## [43] eval-auc:0.866246
## [44] eval-auc:0.866780
## [45] eval-auc:0.867331
## [46] eval-auc:0.867510
## [47] eval-auc:0.868289
## [48] eval-auc:0.868538
## [49] eval-auc:0.869070
## [50] eval-auc:0.869424
## [51] eval-auc:0.869875
## [52] eval-auc:0.870442
## [53] eval-auc:0.870836
## [54] eval-auc:0.871091
## [55] eval-auc:0.871515
## [56] eval-auc:0.871905
## [57] eval-auc:0.872165
## [58] eval-auc:0.872792
## [59] eval-auc:0.873068
## [60] eval-auc:0.873453
## [61] eval-auc:0.873750
## [62] eval-auc:0.873954
## [63] eval-auc:0.874297
## [64] eval-auc:0.874675
## [65] eval-auc:0.875333
## [66] eval-auc:0.875671
## [67] eval-auc:0.875925
## [68] eval-auc:0.876388
## [69] eval-auc:0.876773
## [70] eval-auc:0.876914
## [71] eval-auc:0.877213
## [72] eval-auc:0.877443
## [73] eval-auc:0.877518
## [74] eval-auc:0.877617
## [75] eval-auc:0.877999
## [76] eval-auc:0.878318
## [77] eval-auc:0.878446
## [78] eval-auc:0.878864
## [79] eval-auc:0.879105
## [80] eval-auc:0.879271
## [81] eval-auc:0.879623
```

```
## [82] eval-auc:0.879912
## [83] eval-auc:0.879941
## [84] eval-auc:0.880362
## [85] eval-auc:0.880672
## [86] eval-auc:0.881141
## [87] eval-auc:0.881423
## [88] eval-auc:0.881693
## [89] eval-auc:0.881909
## [90] eval-auc:0.882202
## [91] eval-auc:0.882270
## [92] eval-auc:0.882353
## [93] eval-auc:0.882683
## [94] eval-auc:0.883006
## [95] eval-auc:0.883297
## [96] eval-auc:0.883546
## [97] eval-auc:0.883850
## [98] eval-auc:0.883968
## [99] eval-auc:0.884055
## [100]    eval-auc:0.884260
```

```
## [1] "Accuracy: 0.90471296952516"
```

Author Contributions

The authors of this report would like to acknowledge their individual contributions to the report. Both authors contributed to ongoing discussions about study design and analysis.

- Olachi Mbakwe contributed to the Abstract, Introduction, Exploration of the Data,Comparing the results, coding, and writing of the report.

III. REFERENCES

[1] Bani-Taha, O. I., & Shafiq, M. O. (2020). Combining the Richness of GIS Techniques with Visualisation Tools to Better Understand the Spatial Distribution of Data – A Case Study of Chicago City Crime Analysis. Urban Academy | Urban Culture and Management, 2146-9229. [2] Alqahtani, A., Gharawi, S., & Ghanem, W. (2019). Crime Analysis in Chicago City. 2019 International Conference on Computer and Information Sciences (ICCIS). [3] Kırpık, E. (Year). Spatial Patterns of Crime and Its Relationship with The Physical Environment: Chicago Case. Journal Name, Volume(Issue), page range.