**A logo of a university

Description automatically generated**

**ALY 6015 – INTERMEDIATE ANALYTICS**

**CHICAGO POLICE DEPARTMENT: VIOLENCE REDUCTION - SHOTSPOTTER ALERTS**

**Capstone Group 4:** Devika, Jayakumar, Sagar

**Instructor**: Zhi (Richard) He

**15th March 2024**

**Table of Contents**

Introduction3

Analysis4

Conclusion 10

References11

INTRODUCTION

Given the rich yet diverse nature of the ShotSpotter alerts dataset, which encapsulates a plethora of categorical data (e.g., community areas, incident types) and continuous variables (e.g., time and location coordinates), it becomes imperative to refine our research inquiries strategically, aligning them with the analytical prowess of the listed statistical models and methodologies. It has 20 features with a minimum of 12 usable features, 3 million observations, and two observed variables.

# RESEARCH QUESTIONS

**Do specific community areas exhibit statistically significant disparities in the frequency of gunshot incidents compared to others?**

Delving into the spatial dynamics of gunshot occurrences, this question endeavors to unearth whether specific neighborhoods bear a disproportionate burden of gunshot incidents, potentially unveiling localized hotspots or identifying regions relatively insulated from such occurrences.

**Is there a discernible correlation between the temporal aspect, encompassing the time of day or week, and the probability of gunshot incidents materializing?**

By probing the temporal dimension of gunshot incidents, this inquiry aims to unravel any temporal patterns or fluctuations, shed light on whether particular temporal windows witness heightened activity, and inform strategic resource allocation and scheduling within law enforcement frameworks.

**Can we devise a robust predictive model capable of delineating between distinct incident types, such as gunshots and firecrackers, based on multifaceted factors, including location, time, and other variables?**Pioneering predictive analytics, this question seeks to harness the dataset's predictive power to develop a sophisticated model adept at discriminating between disparate incident types. Such a model can potentially revolutionize law enforcement strategies by facilitating swift and targeted responses to incidents, thereby bolstering community safety.

In this initial analysis, we can answer question 2. Since question 1 requires geographical mapping to spot the most affected area, we need more time, as the heatmap didn’t help. Question 2 involves the application of predictive modeling methods to find the answer, which will be done in the final report.

# EXPLORATORY DATA ANALYSIS

The dataset comprises approximately 200,000 entries across 20 columns, including date, location (block, zip code, community area), incident details (type, rounds fired), and geographical coordinates (latitude, longitude). This rich dataset allows for a comprehensive analysis of gunshot incidents over a defined period.

**Interpretations of Outputs and Results**

**Descriptive Statistics:** The summary statistics of the dataset would provide an overview of the distribution of incidents across different times, locations, and incident characteristics (e.g., rounds fired). For instance, a higher frequency of incidents in specific ZIP codes or community areas could indicate hotspots of gunfire incidents, guiding targeted interventions.

A screenshot of a computer

Description automatically generated

**Temporal Trends:** By analyzing incidents over time, we might observe patterns such as increased incidents during certain months or times of the day. For example, if there's an uptick in alerts during late-night hours, this could inform law enforcement and community organizations about when to increase patrols or community engagement efforts.

**A graph of a number of blue bars

Description automatically generated**

Based on the observations from the above graph, several shot spotter alerts happened during the night. The count gradually increases as sunsets go down. The primary reason for this increase is the “Fire Crackers,” which could be the reason for triggering shot spotter alerts.

**A graph of an object shot rounds

Description automatically generated**

**Incident Types and Rounds Fired:** Understanding the types of incidents (e.g., multiple gunshots, single gunshot) and the number of rounds fired can provide insights into the nature of gun violence in different areas. Areas with frequent incidents of multiple gunshots might be experiencing more severe forms of violence, necessitating focused intervention strategies.

A graph of blue rectangular bars

Description automatically generated

**Outliers:**

One of the observed variables, “Rounds,” which means the number of ShotSpotter alerts received by the Chicago Police Department, must be well analyzed. It has several outliers over 25. It is unrealistic to observe gunshots over 25-200. Hence, those are considered “Outliers” and could be “Fire Crackers.”

**A white box with black lines and dots

Description automatically generated**

**Spatial Distribution:**

The geographical data allows spatial analysis to identify where incidents are most concentrated. Areas with higher gunfire incidents can be prioritized for violence reduction programs, such as installing more surveillance cameras or increasing community outreach programs.

A blue map with white text

Description automatically generated

# MODELS CHOSEN

1. **Decision Tree:** Renowned for its innate ability to unravel intricate relationships between predictor variables and the target variable, decision trees emerge as a formidable ally in our quest for insights. By discerning pivotal splits within the data, decision trees enable the identification of nuanced patterns, empowering law enforcement agencies with actionable intelligence to combat crime effectively.
2. **Logistic Regression:** Famed for its versatility in handling binary classification tasks, logistic regression emerges as a stalwart analytical tool for dissecting the temporal intricacies of gunshot incidents. Logistic regression unveils temporal nuances crucial for informed decision-making within law enforcement circles by modeling the probability of incident occurrence across various temporal epochs.
3. **XGBoost:** Harnessing the formidable capabilities of ensemble learning, XGBoost emerges as a potent weapon in our analytical arsenal. Its adeptness in processing voluminous datasets with diverse features equips it to discern intricate patterns and subtle nuances, thus facilitating the creation of a robust predictive model capable of accurately delineating between distinct incident types.

**Preliminary Answers to Questions**

1. **What areas experience the highest frequency of gunshot incidents?**

Based on the analysis, areas with higher incident reports could be deemed critical zones for intervention. Identifying these areas is the first step in deploying targeted violence reduction strategies.

1. **Are there any temporal patterns in gunshot incidents?**

The analysis might reveal specific times of day, days of the week, or months of the year with increased gunshot activity. These patterns can help in more effectively planning law enforcement and community outreach efforts.

1. **What is the nature of gunshot incidents?**

By examining the incident type and rounds fired, stakeholders can better understand the severity and types of gun violence prevalent in the community, aiding in the development of tailored prevention and response strategies.

1. **How can the data inform violence reduction strategies?**

Insights from the dataset can guide the allocation of resources, such as focusing law enforcement efforts, community outreach, and prevention programs in areas and times most affected by gun violence. Additionally, understanding the nature of incidents can help design interventions to address the root causes of violence.

**Conclusion**

The preliminary analysis of the "Violence Reduction - ShotSpotter Alerts" dataset provides valuable insights into the spatial and temporal distribution of ShotSpotter incidents, as well as the nature of these incidents. These insights are instrumental in informing targeted violence reduction strategies, ultimately contributing to the safety and well-being of communities. Further in-depth analysis, possibly performing “Feature Engineering” to remove outliers considering logical analysis in the observed variables and testing predictive models such as Logistic Regression, Decision-Tree, and XGBoost, to predict gunshots using features such as time and location.

References

Data.gov, 2024, <https://catalog.data.gov/dataset/violence-reduction-shotspotter-alerts>