March 10, 2024

SafeDine Toronto

DSMM - Maple Mapping

PREPARED BY:

Auradee Castro (c0866821)

Bhumika Rajendra Babu (c0867081)

Lakshmi Kumari (c0867090)

Maricris Resma (c0872252)

TABLE OF CONTENTS

I. Project Overview

The SafeDine Toronto initiative, led by Maple Mapping in collaboration with McDonald's, is dedicated to ensuring the safety and security of McDonald's restaurant locations across Toronto. By leveraging Maple Mapping's expertise in geographic analysis and digital mapping services, this project aims to assess the risk of robbery incidents at McDonald's locations and provide actionable insights to mitigate potential risks.

II. Data Requirements

The datasets to be utilized for analysis are sourced from the Toronto Police Service's open data platform and Open Data Toronto. The project aims to assess the safety of McDonald's restaurant locations in Toronto regarding robbery incidents. The analysis will focus on identifying high-risk zones for robbery incidents to enable McDonald's to prioritize safety measures effectively. Two key datasets will be utilized for this purpose:

a. **Robbery Incidents Data**: This dataset contains information on reported robbery incidents in Toronto, including the date, time, location type, premises type, and other relevant attributes. It provides insights into the frequency and distribution of robbery incidents across different neighborhoods and types of locations.
b. **Neighborhood Boundary Data**: This dataset includes boundaries for the neighborhoods within the Toronto. It serves as a geographical reference for analyzing robbery incident patterns in specific neighborhoods and identifying high-risk areas.

III. Data Cleaning

Data cleaning and preprocessing are essential steps in the data analysis pipeline that involve transforming raw data into a clean, structured format suitable for analysis. These processes are crucial for ensuring the accuracy, reliability, and usability of the data for downstream analysis tasks. Data cleaning involves identifying and correcting errors, inconsistencies, and missing values in the dataset, while preprocessing focuses on transforming the data into a format that is suitable for analysis.

- **Handling missing or duplicate records**: Identified and handled the missing and duplicate records in the Robbery dataset to maintain data integrity and ensure the accuracy of analysis.

- **Feature removal**: 11 columns were removed from our dataset as they were deemed irrelevant for our analysis.
- **Renaming columns**: The columns LAT_WGS84 and LONG_WGS84 were renamed as Latitude and Longitude respectively.
- **Outlier removal**: We detected outliers in the 'occurred year' column and subsequently removed them to focus our analysis exclusively on data from 2014 onwards.
- **Data filtering**: Eliminated rows where ‘latitude’ and ‘longitude’ column were recorded as ‘0’. Eliminated rows where ‘Neighborhood name’ and ‘neighborhood number’ were recorded as NSA.

Here’s the key information of the data after preprocessing:

1. **Object ID**: Serves as a unique identifier for each record in the dataset.
2. **Report Year**: Year the offence was reported.
3. **Report Month**: Month the offence was reported.
4. **Report Day**: Day of the month the offence was reported.
5. **Report Day of the Week**: Day of the Week offence was reported.
6. **Report Hour**: Hour the offence was reported.
7. **Occurred Year**: Year the offence occurred.
8. **Occurred Month**: Month the offence occurred.
9. **Occurred Day**: Day of the month the offence occurred.
10. **Occurred Day of the week**: Day of the week the offence occurred.
11. **Occurred Hour**: Hour of the day the offence occurred.
12. **Location Type**: Location type of the offence. Provides contextual information about the setting of where the incident took place.
13. **Premises Type**: Provides details about the establishment or structure involved in the incident,
14. **Offence**: Provides specific type or category of the offence committed during the robbery.
15. **Hood 158**: Identifier of Neighbourhood using City of Toronto's new 158 neighbourhood structure.
16. **Neighbourhood 158**: Name of Neighbourhood using City of Toronto's new 158 neighbourhood structure.
17. **Longitude**: Longitude coordinates of the location of incident.
18. **Latitude**: Latitude coordinates of the location of the incident.

IV. DATA ANALYSIS

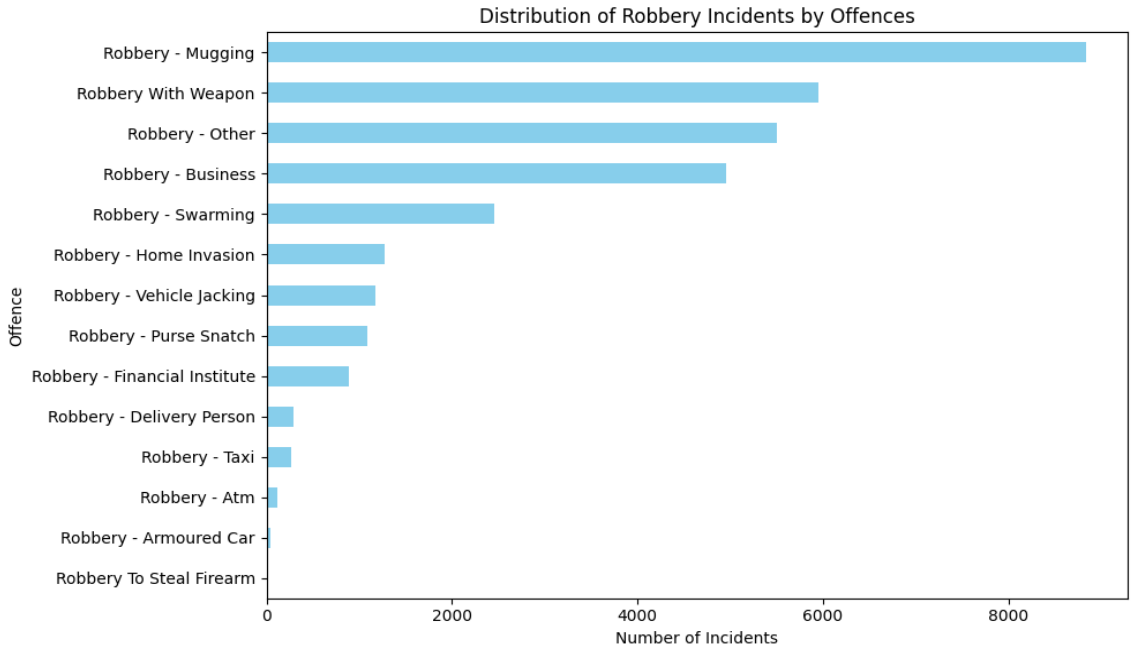
The dataset comprises over 30,000 rows of data spanning the period from **2014 to 2023**, reflecting reported robbery incidents in the Greater Toronto Area. These incidents are distributed across 158 distinct neighborhoods within Toronto, providing a comprehensive view of robbery occurrences throughout the city over the specified time frame.

Conducting statistical analysis involves examining data to uncover patterns, trends, and relationships, providing valuable insights for decision-making and problem-solving.

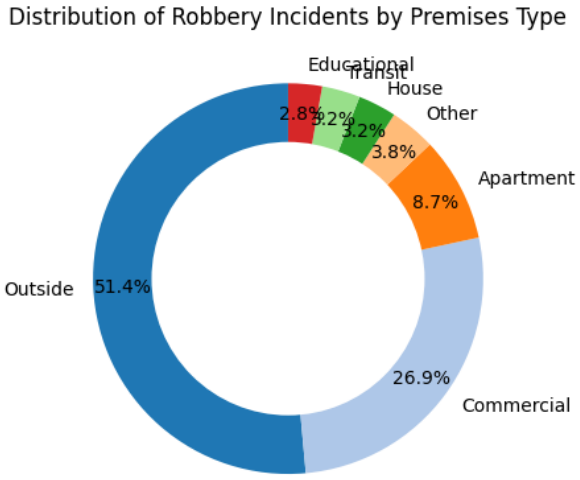
1. Descriptive Analysis

Descriptive analysis involves summarizing and describing the main features of a dataset. This type of analysis provides insights into the basic characteristics of the data, such as central tendency, dispersion, and distribution. Descriptive analysis is the first step in understanding the nature of the data before delving into more complex analyses.

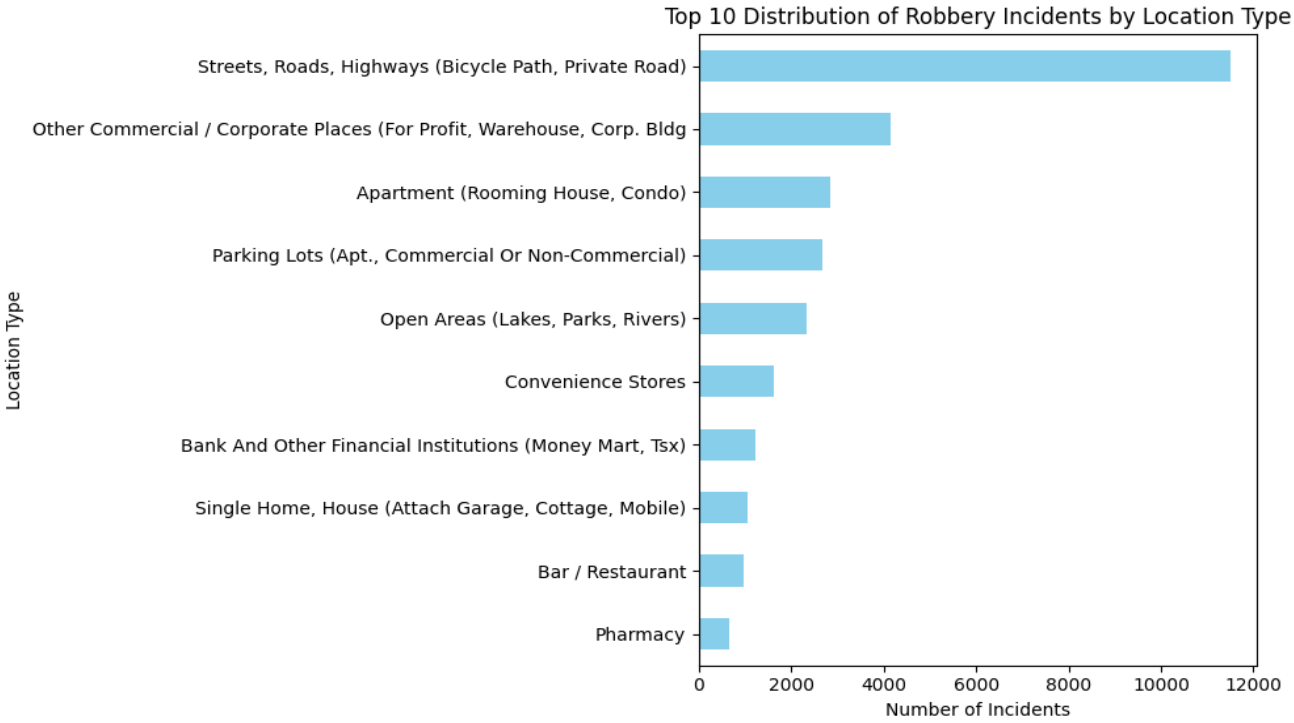
The robberies in the dataset are categorized into 14 distinct types of offences. Namely,



The analysis of premises type revealed that 51.4% of robberies occurred outdoors, while 26.9% took place in commercial spaces.



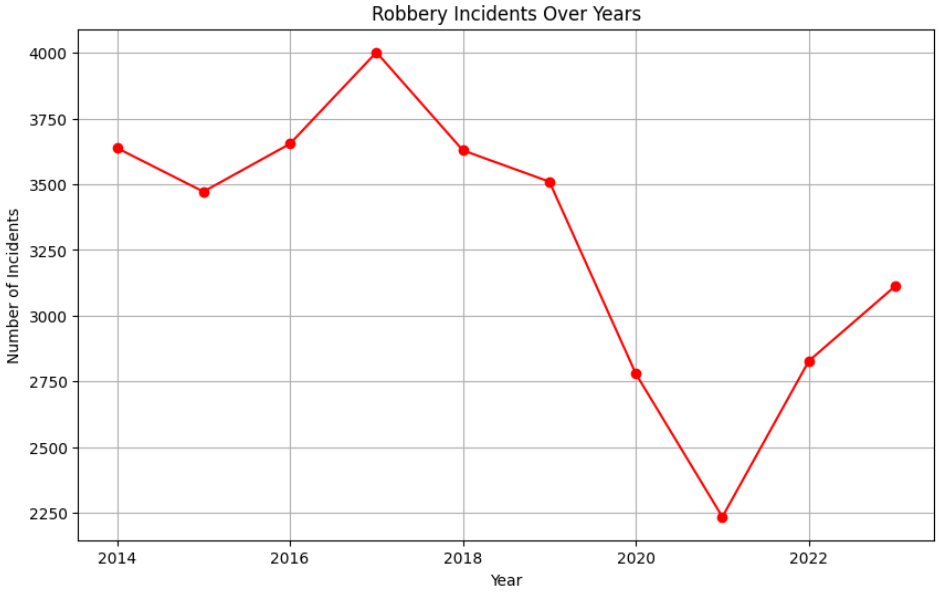
The distribution of location types provides insights into where robbery incidents commonly occur, offering valuable information for crime prevention strategies and resource allocation.



2. Temporal Analysis

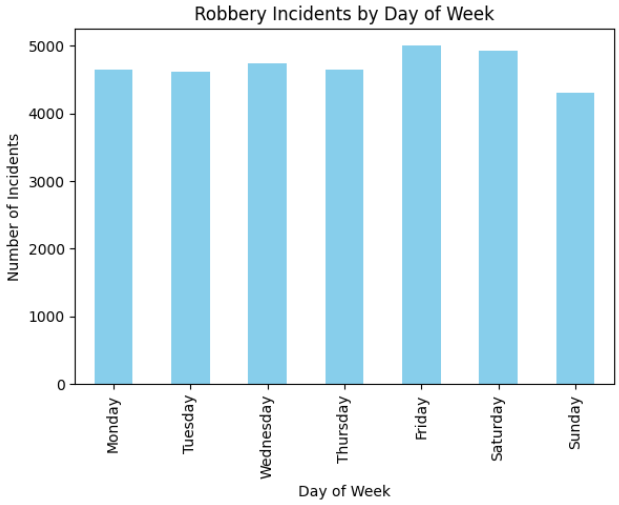
Temporal analysis involves examining patterns and trends in data over time. In the context of crime analysis, temporal analysis focuses on understanding how criminal activities vary over different time periods, such as days, weeks, months, or years.

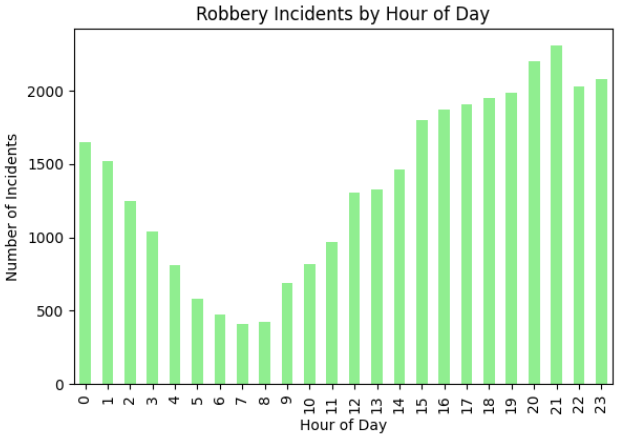
a. The line graph depicting robbery incidents across the time span from 2014 to 2023 reveals a peak in occurrences during the year 2017, with a notable decrease observed in 2020 and 2021.



b. Analysis based on the day of the week does not reveal significant variations in robbery incident patterns. This finding suggests that robbery incidents occur relatively uniformly across different days of the week, with no distinct peaks or troughs.

c. Analysis based on the hour of the day reveals interesting patterns in robbery incidents. It appears that certain hours are associated with higher frequencies of robberies, indicating potential temporal trends or patterns in criminal activity.





3. Association Analysis

Association analysis, also known as market basket analysis, is a data mining technique used to discover interesting relationships or associations between items in large datasets. In the context

of our project, association analysis can help identify correlations between various factors such as location type, premises type, and offense type, shedding light on potential relationships between these variables and robbery incidents.

a. **Contingency tables** and **chi-square tests** are statistical tools commonly used to explore relationships between categorical variables. In our analysis, we employ contingency tables to organize the counts of occurrences for different combinations of offenses and location types. The chi-square test is then applied to evaluate whether there is a significant association between these variables beyond what would be expected by chance.

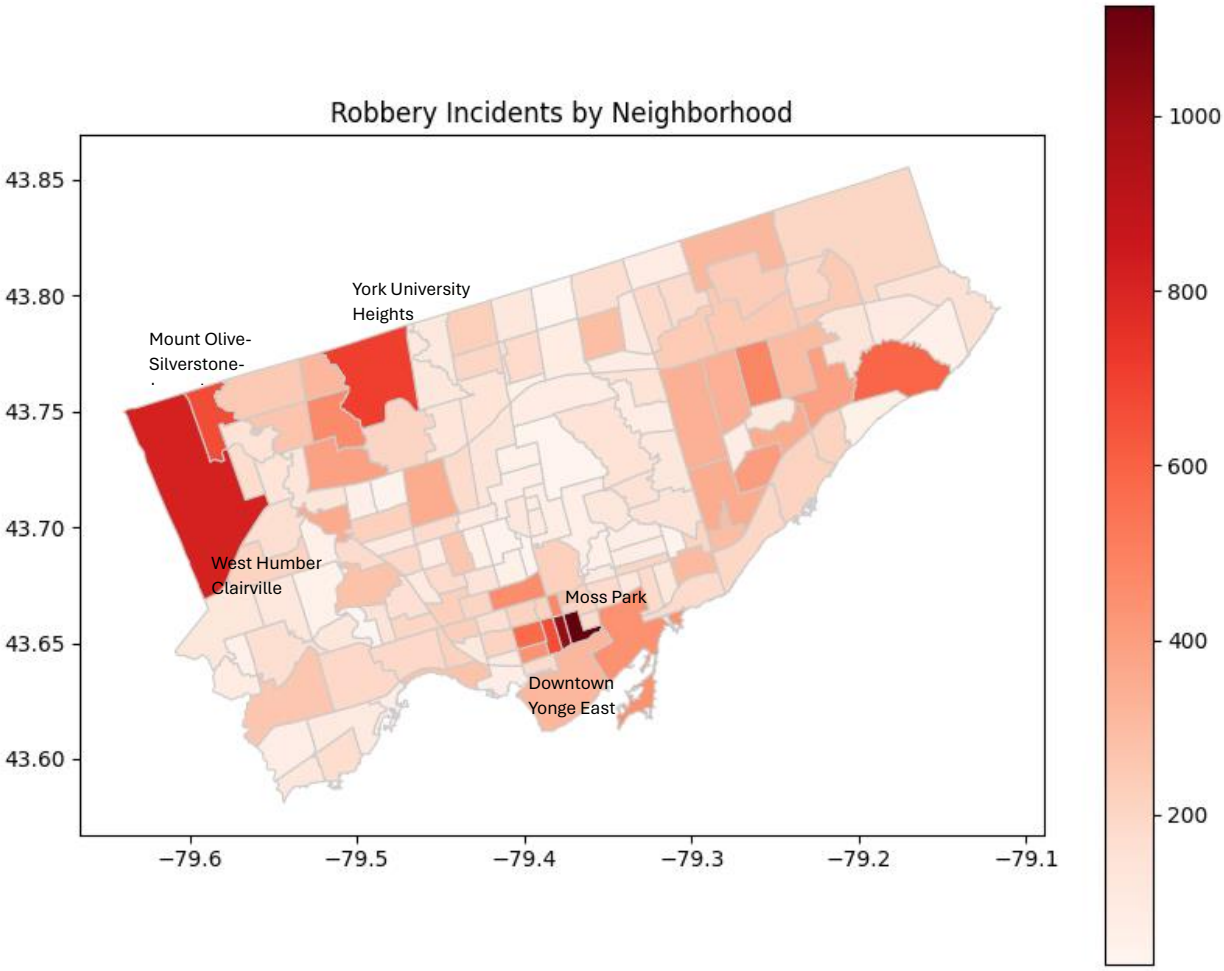
By examining the results of the chi-square test, we can determine whether there is evidence of a statistically significant relationship between offenses and location types. If the p-value associated with the chi-square test is below a predetermined significance level (e.g., 0.05), we reject the null hypothesis of independence and conclude that there is a significant association between the two variables.

The analysis of the contingency table and chi-squared test reveals a significant association between the type of robbery offense and the location type. Specifically, mugging emerges as the most prevalent form

of robbery across streets, roads, and highways. This finding suggests that these public thoroughfares are particularly susceptible to mugging incidents compared to other types of robbery offenses.

4. Spatial Analysis

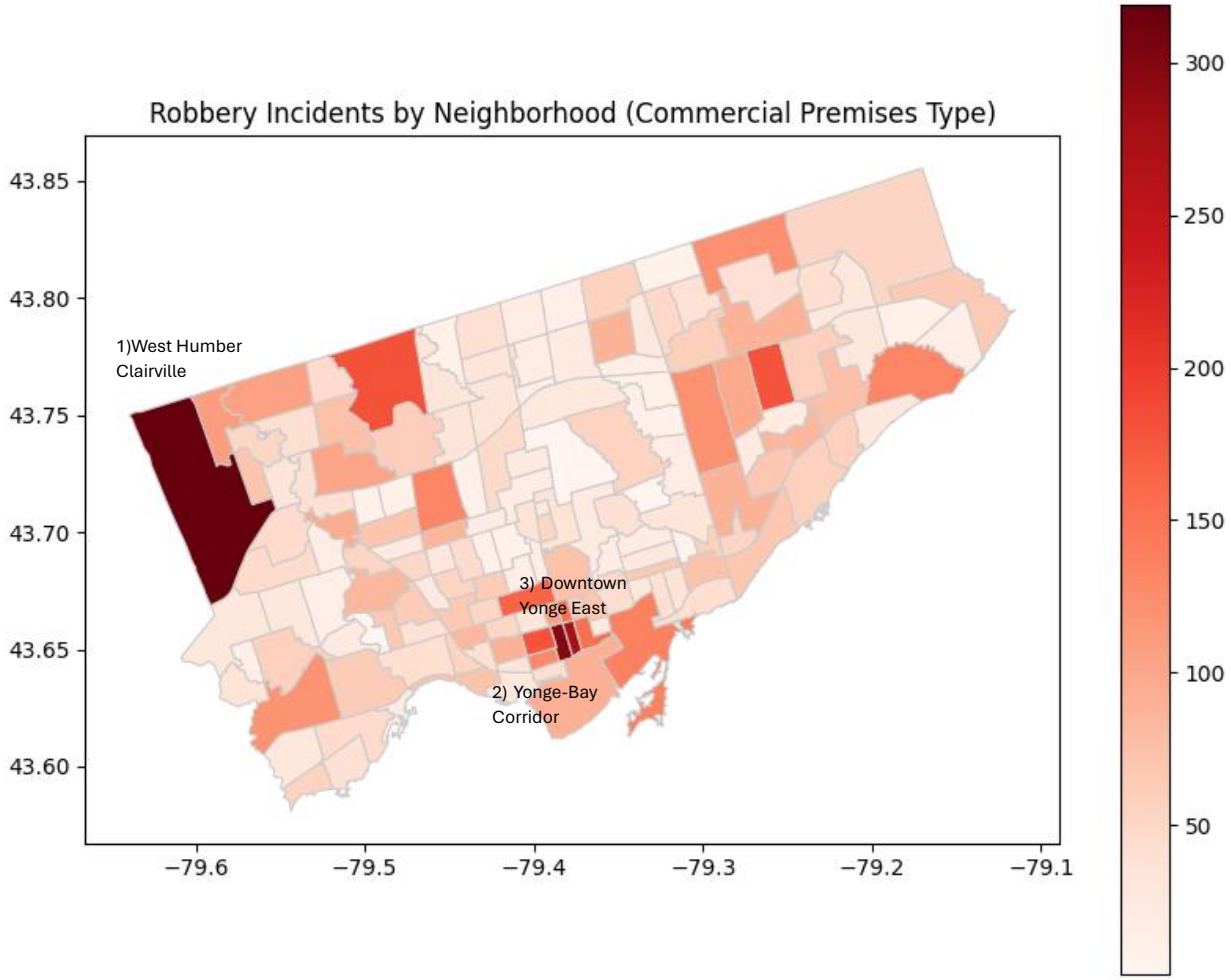
Spatial analysis is a powerful method used to examine geographical patterns, relationships, and trends within spatial data. By leveraging techniques such as geographic information systems (GIS), spatial statistics, and mapping, spatial analysis enables us to gain valuable insights into the spatial distribution of phenomena and understand the underlying spatial processes driving them.



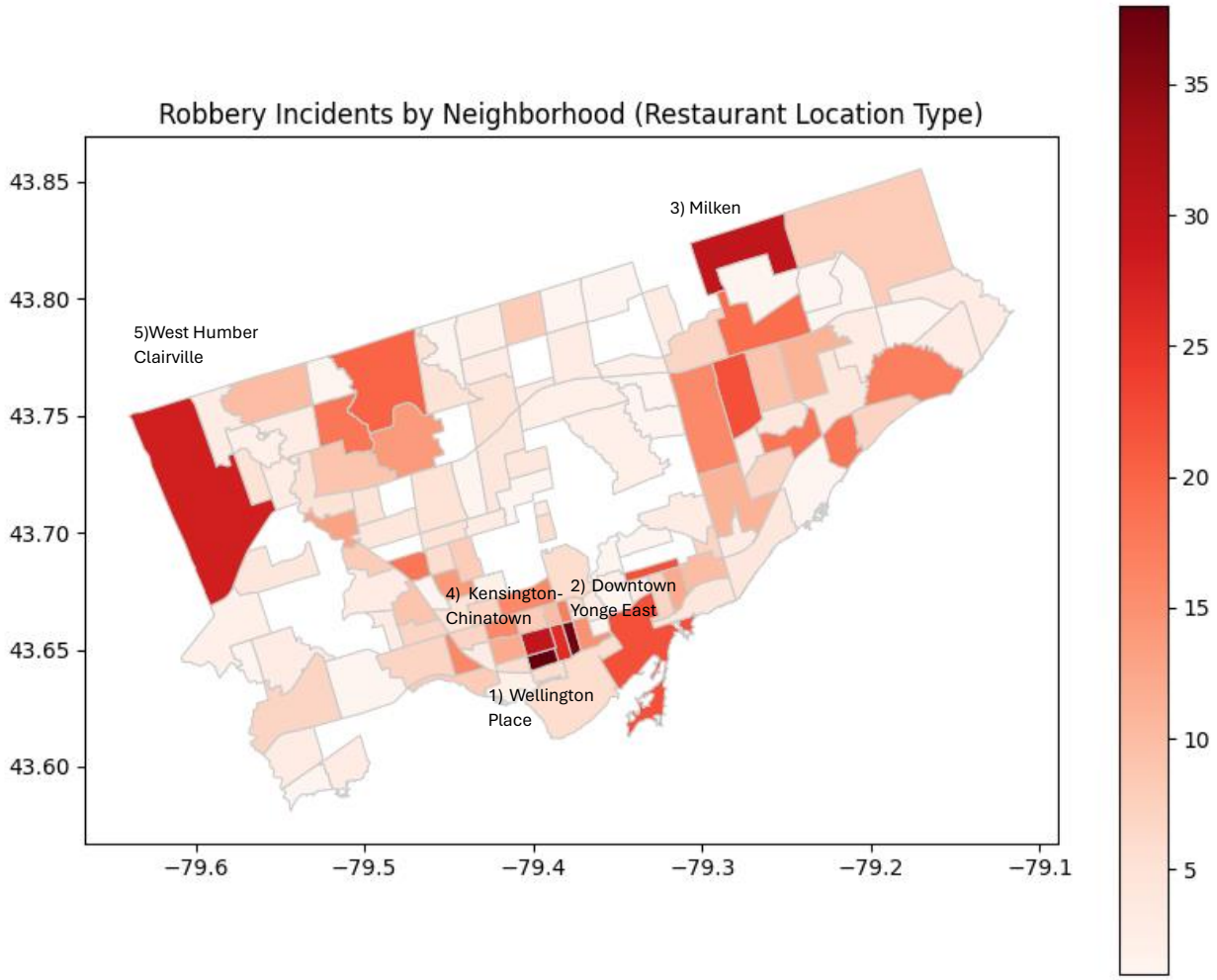
The choropleth map of Toronto neighborhoods represents the intensity of robbery incidents across different areas. Darker shades indicate higher incident densities, while lighter shades correspond to

lower densities. This visualization allows us to quickly identify areas with a higher concentration of robbery incidents. The top neighbourhoods with most number of robberies are pin pointed.

To refine the analysis specifically for McDonald's, further examination was conducted focusing exclusively on commercial-type spaces, which are more relevant to McDonald's operations. By narrowing down the scope to commercial spaces, we could extract insights that are directly applicable to McDonald's business environment, ensuring the analysis aligns closely with their needs and objectives.



Moreover, recognizing the significance of fast-food chains like McDonald's within the commercial landscape, additional analysis was conducted specifically targeting bars and restaurants.



In the choropleth map representing robbery incidents at bar and restaurant locations, a striking difference emerges when compared to the overall robbery incident distribution. While certain neighborhoods remain significant hubs for criminal activity, the top places for bar and restaurant-related robberies vary considerably.

Interestingly, the analysis reveals numerous neighborhoods with notably low robbery incidents, some reporting as few as just one occurrence. This discrepancy underscores the diverse safety profiles across Toronto's neighborhoods and highlights the localized nature of criminal activity, particularly in establishments such as bars and restaurants.

V. Conclusion

Through this project, we aimed to assess the safety of McDonald's restaurant locations in Toronto regarding robbery incidents. By leveraging geographic analysis and digital mapping services, we analyzed historical data on robbery incidents spanning from 2014 to 2023, sourced from the Toronto Police Service Open Data, and neighborhood boundary data from Open Data Toronto.

Overall, this project underscores the importance of leveraging data-driven insights to enhance safety measures and make informed decisions. By collaborating with Maple Mapping, McDonald's demonstrates its commitment to creating a safe environment for customers and employees, utilizing advanced analytical techniques to identify high-risk zones and implement targeted interventions effectively.

VI. Appendices

Appendix	Item	Location
A	Codes for SafeDine	Codes available on link
B	Robbery Analysis Dataset	Dataset available on link