DATA CAPSTONE PROJECT

Part I: Business Problem

Group: Insight Ink Crew

Professor: Zeyad Azem

Date: 3rd November, 2023

Submitted by:

Baskaran, Asmita

Coimbatore Jayaraj, Pavithran

Kaur, Harmanjot

Keshav, Pooja

Nalchigar, Niloofar

Pham, Thi

Executive Summary

Toronto, like many urban areas, faces the challenge of addressing homicides within its jurisdiction. This report presents a comprehensive analysis of homicides in Toronto, utilizing data from the Toronto Police Department's open data source. The report identifies key stakeholders, defines the project's scope, outlines its potential impact, and establishes success criteria. Through data analysis, advanced analytics and generative AI, we aim to provide actionable insights that will assist law enforcement, community organizations, and policymakers in addressing the issue of homicides in the city.

Introduction

The Toronto Police Department serves as the central authority for maintaining public safety in the city, with responsibilities encompassing crime prevention, investigations, emergency response, and community engagement. The increasing population of Toronto has a multifaceted impact, affecting public services, housing, diversity, and socio-economic dynamics. The recent economic crisis, such as the COVID-19 pandemic, has introduced challenges related to job loss and financial stress, potentially influencing crime rates. These factors collectively contribute to the complex landscape of public safety in Toronto, necessitating a data-driven analysis to address the evolving needs of the community. Homicides are a critical concern for any community, and Toronto is no exception. The purpose of this project is to harness the power of data analytics to understand the patterns and dynamics of homicides in the city. By doing so, we aim to contribute to the overall safety and well-being of Toronto's residents. This report outlines the business problem and the steps involved in addressing it.

Problem Statement

The increasing population and evolving socio-economic dynamics in Toronto, exacerbated by external factors such as the COVID-19 pandemic, have led to a rising concern over homicides within the city's jurisdiction. There is a critical need for a data-driven analysis to understand the patterns and dynamics of homicides in order to provide actionable insights for law enforcement, community organizations, and policymakers to effectively address this pressing issue and ensure the safety and well-being of Toronto's residents. In light of these complexities, addressing the concern of rising homicides in Toronto requires a focused approach that takes into account several critical challenges:

- Challenge of efficiently allocating police resources, especially in response to the changing demographics and the potential influence of immigration on crime rates
- Data Availability and Reliability in Homicide Cases: The challenge lies in obtaining comprehensive, up-to-date, and accurate data crucial for effective analysis of homicide cases.
- Privacy Compliance in Handling Criminal Case Data: Strict adherence to privacy laws and regulations is imperative when dealing with sensitive information from criminal cases. This involves the meticulous anonymization of data and ensuring full compliance with relevant legal frameworks for the project.

- Mitigating Data Biases for Trustworthy Analysis: Tackling potential biases within the data and analysis process is critical. It's essential to guarantee that the project's findings and recommendations are fair and impartial to uphold trust and credibility.

Stakeholder Analysis

Toronto Police Department:

- The Toronto Police Department (TPD) is the primary law enforcement agency responsible for maintaining public safety in Toronto.
- TPD aims to leverage data analytics to enhance their crime prevention strategies. They would be particularly interested in identifying patterns related to the time, location, and demographics of homicide incidents to allocate resources more effectively.
- The department could also seek insights that can help improve the accuracy and efficiency of investigations, ultimately leading to quicker resolutions of homicide cases.

Community Organizations:

- Various community organizations are committed to reducing homicides and ensuring the safety and well-being of residents.
- These organizations often work at the grassroots level and seek data-driven insights to guide their community engagement efforts. They would be interested in identifying factors that contribute to homicides, such as socioeconomic disparities or gang-related activities, to develop tailored intervention programs.

Policymakers and Government:

- Policymakers at the municipal and provincial levels play a critical role in shaping policies related to public safety, criminal justice, and social services.
- They rely on data to design and implement effective policies that address the root causes of homicides. The analysis will assist them in understanding how policies can impact crime rates and community safety.

The General Public:

- The general public in Toronto is a key stakeholder, as they are directly affected by the level of safety and security in their neighborhoods.
- Residents expect law enforcement and policymakers to use data to make informed decisions that enhance community safety. They would want to see tangible results in terms of reduced violence and improved living conditions.

Objectives

- A predictive model that can estimate resource requirements for Homicide incidents based on historical data and future immigration trends in Toronto.
- Uncovering trends to predict when homicide incidents are more likely to occur.
- Uncovering trends to predict areas of homicide incidents are more likely to occur.
- Predict the average response time of law enforcement to homicide incidents based on factors like location, time of day, and historical data. This can improve emergency response procedures.
- Examine how external factors like economic conditions, weather, and major events (e.g., festivals, protests) impact homicide rates and predict their effects in the future.
- Integrating Ottawa and Toronto datasets to extract meaningful insights informing Ontario Provincial Police (OPP).

Data Sources

The primary data source which will be used for this project is the open data available on Toronto Police Department's open data portal, which includes comprehensive details on Homicide incident data. Datasets from the Open Ottawa Portal will also be used for integration and comparison purposes. Data preprocessing will involve cleaning, aggregating, and merging data to ensure consistency and accuracy.

Data Analysis

Analysis Approach:

- The analysis approach will involve identifying patterns, correlations, and trends within the data to answer specific research questions.
- Exploratory data analysis will uncover insights, while statistical tests may be used to validate hypotheses.

Data Visualization:

- Visualizations will be created to illustrate key findings, such as time trends in homicides, geographical distribution, and demographic breakdowns.
- Data visualizations will enhance the interpretability of the analysis and reveal insights.

Hypotheses and Questions:

- Hypotheses and research questions will be focused on specific aspects of homicides, such as understanding if there are seasonal patterns, demographic associations, or geographic concentrations of incidents.

Project Scope

Crime Type: Homicide

Geography: The geographic scope of the project is the city of Toronto and Ottawa, which includes multiple neighborhoods, each with its own unique characteristics and challenges. By focusing on the two cities, the project will offer a comprehensive perspective on the distribution of homicides.

Timeline: The project will examine homicide data over the past nineteen years in Toronto and the past seven years in Ottawa, allowing us to understand long-term trends and assess the effectiveness of recent interventions. Recent data will be analyzed in greater detail to identify emerging patterns and hotspots.

Factors and Variables available to analyse: The analysis will include variables, type of homicide, date, occurrence day of the week, division, location (latitude and longitude), and any additional variables that might be relevant for understanding the patterns and causes of homicides.

Outputs:

- Actionable Recommendations - The project anticipates generating actionable recommendations for law enforcement, community organizations, and policymakers. These recommendations will be tailored to address specific challenges identified in the data analysis.
- Community Safety Improvements - The implementation of recommendations is expected to lead to visible improvements in community safety. This may manifest as reduced crime rates, especially in high-crime areas, and a greater sense of security among residents.
- Policy Implications - The project's findings will carry policy implications that can inform decisions related to crime prevention, social services, and community engagement.
- Resource Allocation - The optimization of resource allocation based on data analysis will enhance the efficiency and effectiveness of law enforcement efforts in the city.

Defining Success:

- **Reduction in Homicide Rates -** Success will be measured by a tangible reduction in the number of homicides in Toronto.
- **Response Times -** Improved response times to homicide incidents will signify success in terms of public safety and the police department's ability to swiftly address violent incidents.
- **Community Engagement Metrics -** Success in community engagement can be measured by increased participation in community safety programs and initiatives, reflecting a growing trust in law enforcement.
- **Policy Impact -** The project's success will be determined by the extent to which data-informed policies lead to a reduction in crime rates and the improvement of community safety.
- **Resource Allocation Efficiency -** Success will be indicated by the optimized allocation of resources, which will contribute to more efficient and effective law enforcement efforts.