



January 2025

# PROJECT proposal

Team Members  
**Swati Mehta**  
**Jaskaran Singh**  
Prepared for:  
**CS2545**

# Crime Analysis in Canada

## Toronto vs Vancouver

Team Member & Roles		
Swati Mehta	Tech Lead and Project Manager	<ul style="list-style-type: none"><li>• Oversees the technical development of the project, ensuring the implementation of data analysis techniques and visualization tools.</li><li>• Manages project timelines, task assignments, and overall coordination among team member.</li></ul>
Jaskaran Singh	Tech Lead and Doc Lead	<ul style="list-style-type: none"><li>• Responsible for writing and organizing all project documentation, including the proposal and final report.</li><li>• Assists in coding and implementation, particularly in data preprocessing and visualization.</li></ul>

# Issues

Crime is a persistent challenge that impacts the safety, security, and well-being of individuals and communities across Canada. However, the nature and frequency of crimes vary significantly across different cities and regions. Vancouver and Toronto, two major urban centers with distinct population sizes, income distributions, and employment rates, experience crime patterns influenced by these factors. Analyzing crime trends in these cities is crucial for understanding regional disparities and the relationship between crime and socioeconomic conditions such as income levels, education, and housing conditions. By identifying high-crime areas, examining trends over time. This project aims to analyze crime rates in Vancouver and Toronto, highlight regional differences, and provide insights that can contribute to improved public safety in both cities.

## Key Highlights of the Project:

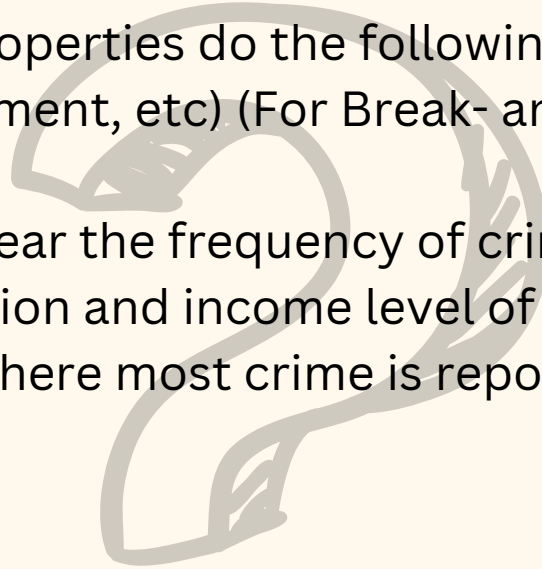
- **Crime Rate Analysis:** Examine crime rates across Toronto and Halifax to identify high-risk areas.
- **Trend Visualization:** Use data visualization techniques to track changes in crime patterns over time.
- **Socioeconomic Correlations:** Investigate how factors such as income levels, education, and employment influence crime rates.

# Methodology

- **Data Collection:** Utilize open data sources such as Canada's Open Government Portal and Toronto's and Halifax's city-specific open data portal. The dataset will include crime statistics from different neighbourhood and socioeconomic indicators.
- **Visualization:** Develop charts and GRAPHS to illustrate crime trends over time and across regions
- **Data cleaning :**Apply data wrangling techniques such as handling missing values, and filtering for relevant parameters to maintain data accuracy.
- **Socioeconomic Correlation Study:** Perform statistical correlation analysis to examine relationships between crime rates and socioeconomic factors like income disparity, unemployment, and education levels.
- **Crime Rate Classification by Region:** Identify neighborhoods with higher and lower crime rates in Toronto, analyzing patterns based on geographic and socioeconomic factors.
- **Insights and Recommendations:** Derive meaningful conclusions from the analysis and propose data-driven policy recommendations to aid crime prevention .

# RESEARCH QUESTIONS

1. What are the most common / least common crimes ?
2. What are the total number of crimes till 2025 and any trends?
3. What time of the year the frequency of crime is highest?
4. Relationship between crime and time of the day
5. In what kinds of properties do the following crimes occur? (house, commercial, apartment, etc) (For Break- and-enter, Assault and Robbery)
6. What time of the year the frequency of crime is highest?
7. What is the education and income level of the people living in neighbourhoods where most crime is reported?



# Approach to my Research

## 1. Most Common / Least Common Crimes

- Approach: Analyze the dataset for crime types and their frequency.
- Generate bar charts or pie charts to visualize crime distribution.
- Methods: Use Pandas for data aggregation.
- Use Matplotlib for visualization.

## 2. Total Number of Crimes Till 2025 and Trends

- Approach: Aggregate crime counts over time to detect trends.
- Identify fluctuations and analyze if crime rates are increasing, decreasing, or stable.
- Methods: Use time-series analysis in Python.
- Visualize trends using line graphs and moving averages.

## 3. Time of the Year with Highest Crime Frequency

- Approach: Analyze crime occurrences by month.
- Identify peak crime periods.
- Methods: Group data by month.
- Generate bar graphs.

## 4. Relationship Between Crime and Time of Day

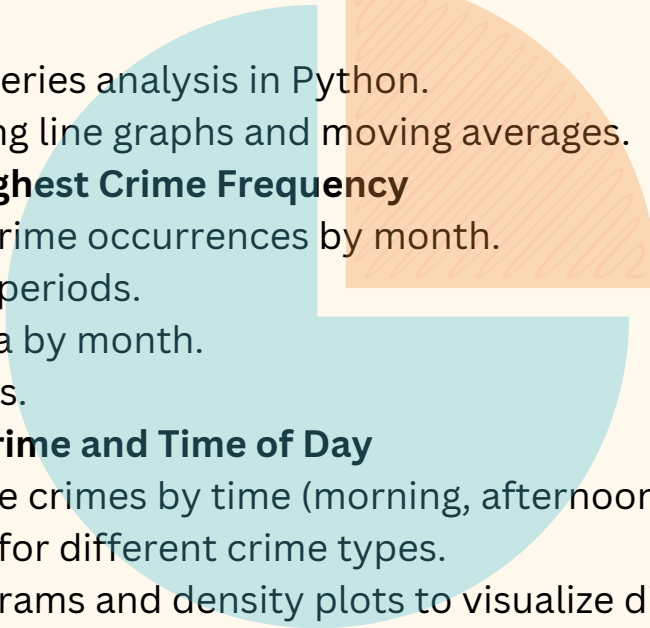
- Approach: Categorize crimes by time (morning, afternoon, evening, night).
- Identify peak hours for different crime types.
- Methods: Use histograms and density plots to visualize distributions.

## 5. Crime Occurrence by Property Type

- Approach: Analyze which types of properties are most affected by these crimes.
- Compare property types (house, apartment, commercial, public spaces).
- Visualize using bar charts.

## 6. Education and Income Levels in High Crime Neighborhoods

- Approach: Correlate crime rates with census data (income, education).
- Identify disparities between neighborhoods.
- Methods: Use Statistics



# Data Sources

## **Vancouver Crime Data:**

<https://opendata.vancouver.ca/pages/home/>

## **Toronto Crime Data:**

<https://www.kaggle.com/datasets/kapastor/toronto-police-data-crime-rates-by-neighbourhood>

<https://www.toronto.ca/city-government/data-research-maps/open-data/>