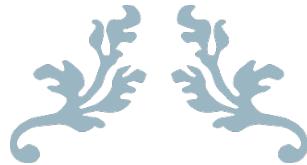


SEP 6DA3: Data Analytics and Big Data



RETAIL ACCESS VS. SOCIAL COST

Store Locations, Sales Trends, and the Measured Social Impact of Legal Cannabis in Canada.



Submitted by: Group 4

400587184 - Moosani, Iliyan
400678333 - Parekh, Aayushi Chintan
400550277 - Patel, Jainish Shaileshbhai
400552080 - Patel, Mayur Kalabhai
400676356 - Patel, Priyaben
400591994 - Virani, Jenil Salimbhai

Under the guidance of:

Dr. Pedro Tondo
McMaster University

Table of Contents

Abstract	1
1. Introduction	2
2. Methodology	4
2.1 Data Acquisition and Project Scope	4
2.2 Data Publishing.....	6
2.3 Tools, Frameworks and Software Development	7
3. Data	8
3.1 Data Preprocessing	8
3.2 Exploratory Data Analysis (EDA)	10
3.3 Feature Extraction.....	11
4. Experiments & Results	12
Canada Wide Cannabis Sales Analysis Dashboard.....	12
Cannabis Store Locations Dashboard.....	13
Cannabis Retail Sales by Province Dashboard.....	14
Cannabis Retail Sales by Cities Dashboard	15
Cannabis Retail Sales by Type Dashboard.....	16
Canada Wide Crime Analysis Dashboard	17
Major Cities Crime Analysis Dashboard	18
Toronto Crime Analysis Dashboard.....	19
Vancouver Crime Analysis Dashboard	20
Edmonton Crime Analysis Dashboard.....	21
5. Discussion	22
Interpretation	22
Challenges	22
6. Conclusion & Future Work.....	23
Summary.....	23
Future work	23
References	24

Abstract

Canada legalized recreational weed on October 17, 2018, when the federal Cannabis Act took effect made it legal to buy, possess, and grow non-medical cannabis in Canada created an unprecedented natural experiment in drug policy. This study examines how cannabis retail expansion relates to economic performance and community safety across Canada's provinces and cities.

We combined several Statistics Canada datasets such as stores locations, monthly sales data, retail sales, and statistics on crime. By applying data processing with Python, main data quality issues were dealt with inconsistent standards in reporting, incomplete location data, and classification changes over time. Cleaned datasets were then analyzed in Jupyter notebooks and visualized using Power BI dashboards.

Our analysis puts the annual cannabis market totalled \$5.2 billion for the fiscal year that ran from April 1, 2023, to March 31, 2024, with large regional variations. Ontario, Alberta, and British Columbia are leading total sales, with Ontario holding 40% of the national market share. Alberta posts the highest per-capita consumption. Dried cannabis dominates with 65% market share, although other products are growing. Crime analysis indicates that Toronto, Edmonton, and Vancouver rank among the top cities in Canada in overall crime rates, but we find only weak links between store density and crime levels, suggesting other factors drive these patterns. The study provides a reproducible framework for ongoing policy evaluation, demonstrating that the impacts of cannabis legalization vary significantly across regions and time.

Results will support improvements in regulation, public health planning, and evidence-informed policy development over the course of ongoing evolution in Canada's cannabis market.

1. Introduction

Background and policy context

Canada became the first G7 country to legalize recreational cannabis when the Cannabis Act took effect on October 17, 2018.^[1] The law created a split system in which the federal government oversees cultivation and processing, while the provinces control the retail stores, as well as basic consumption rules and additional taxes.

This created sharply different results across various provinces and territories of Canada. Alberta opted for a free market with few restrictions and saw more than 600 stores pop up for 4.4 million people. Several Atlantic provinces maintained tighter government control and limited private retailers. Ontario, home to 38% of Canada's population, opted initially for a lottery system that slowed early growth before moving to open applications in 2019.

The economic impact has been substantial. By 2024, legal cannabis sales were at \$5.2 billion a year. Ontario, Alberta, and British Columbia accounted for most of the national sales, with Ontario making up over a third of the market on its own. Dried cannabis remains the top product, taking more than 65% market share, but edibles and extracts continue to gain popularity.

Community Safety and Crime

The social effects of legalization, especially with regard to crime, are a leading policy concern. Crime statistics indicate that Toronto, Edmonton, and Vancouver invariably yield the highest crime rates in the country. These patterns reflect not only population size but also complex social factors, economic conditions, and policing approaches in major cities.

National trends in crime rates demonstrate a general decline or leveling off since legalization, though with a lot of regional variation depending on local demographics and policy. Our research found only weak links between cannabis store density and local crime rates, suggesting that crime patterns depend on much broader social and economic factors.

Motivation and Goals

Despite the ongoing policy debates, very few studies have comprehensively investigated market development together with social outcomes. Most of the existing analysis are narrowly focused on health impacts or impaired driving, missing broader spatial and temporal patterns that define the cannabis sector in Canada.

This goal of the study is to fill these gaps through comprehensive data integration and analytics. We harmonize government datasets from disparate sources, explore cannabis market trends at multiple levels, and investigate the relationship between retail expansion and community safety.

Objectives

Our main focuses are:

- Integration of national, provincial, and municipal datasets, standardization, and assurance of data quality
- Documenting market development patterns, including regional differences in sales, store access, and product trends
- Analyzing the relationship between cannabis retail infrastructure and crime in major cities
- Develop transparent, reproducible methods and interactive tools for the continuous monitoring of policies.

This framework showcases pragmatic methods for the integration of sophisticated government data. Our report represents one of the most in-depth analyses on the Canadian cannabis market and its community impacts. The tools and methods we developed can serve to assist government agencies, researchers, and advocacy groups in monitoring these trends as other jurisdictions around the world consider similar legalization policies.

2. Methodology

This section provides detailed information on the methodology of data acquisition, decision factors, benefits and drawbacks of the dataset and tools, frameworks and software's utilized during the project.

2.1 Data Acquisition and Project Scope

The process of gathering the datasets was a lengthy process as we required multiple datasets for Crime, Sales and Store locations. There were specific requirements to be fulfilled by the data to be collected for the project. The timeline of the data held the at most importance. We required dataset to be having data for more than 5 years so that we visualize the trend over the years. We also required the locations for the Stores and Crimes to have an address so that we could pinpoint the locations during visualization process.

Here is the detailed process of data collection for each of the requirement:

Store Locations

We evaluated several datasets with data on the locations of cannabis shops. During this process, we learned that Health Canada releases province-specific information on approved store counts through distinct links for each provincial data portal.[\[2\]](#) The files differed in format and organization because each province has its own dataset; some were CSVs, while others were published on website. Additional discrepancies were also evident. For instance, some datasets included postal codes that needed to be extracted from an address field, others had no postal codes at all, and none had latitude or longitude coordinates.

These differences hinted to a tedious data cleaning and processing, but the total number of data points surpassed 4000 when counted together. Therefore, we decided to move forward with these datasets.

Type of Cannabis Sales Data

While searching for sales data for Cannabis, we came across a dataset which showed Sales of cannabis by liquor and cannabis authorities and other retail outlets, by cannabis type on Statistics Canada.[\[3\]](#) There are various types of Cannabis such as Dried cannabis, Inhaled cannabis extracts, Ingested cannabis extracts, solid cannabis edibles, cannabis beverages, Topicals, seeds and other cannabis products. The dataset contained annual sales for all the provinces which was beneficial.

This dataset also required some cleaning process as it had few extra columns which did not provide any vital information and could be dropped, but it was important to understand the sales of cannabis by its type and understand the trends. Therefore, we decided to move forward with the dataset.

Retail Trade Data

While we had annual sales data, it was lacking depth as we did not have information on monthly and quarterly trends. We found a retail trade dataset (Monthly retail trade sales by province and territory) which measured monthly retail trade sales across all industries by province and territory.[\[4\]](#) This was an important finding as the dataset provided detailed information on monthly sales from 2017 for each province. It defined entire retail trade sector defined by North American Industry Classification System (NAICS).

This vast amount of data had to be filtered as we only required trade information for the Cannabis Retailers and had various missing data which had to be taken care of. As this dataset provided key insights into the trades, we decided to utilize the information.

Crime data

After wrapping upon store locations and sales part of the project, we had to look for crime data. We started with searching on Statistics Canada website and found a dataset that provided comprehensive annual police reported crime statistics for Canada, with 300 detailed violations and geography.[\[5\]](#) The timeline of this dataset was long as the reference date started from 1998 which was beneficial as we could observe the trend over the years. With crime data per province the project had the potential to study the provinces with highest crimes over the years and the increase and decrease of crimes.

Toronto Crime data

Along with province level crimes, we wanted to observe city level crimes as well and we wanted to observe the most popular cities for the same. Having thought of Toronto first led us to find crime dataset from the Municipal Data portal.[\[6\]](#) This portal had subsets of data based on Persons in Crisis with reference date starting from 2014, Calls for Service with reference date starting from 2019 Attended and Traffic collisions with reference date starting from 2019. These datasets did not contain addresses due to privacy restrictions but contained neighbourhood information.

These subsets of datasets had to be utilized separately as they provided different information that could not be clubbed. For example, Traffic Collisions dataset contained a column named collision type which had types such as Fatal & Injury, Injury and property damage only. Whereas Person in Crisis had event type column with types such as Overdose, Person in crisis and suicide related.

Vancouver Crime Data

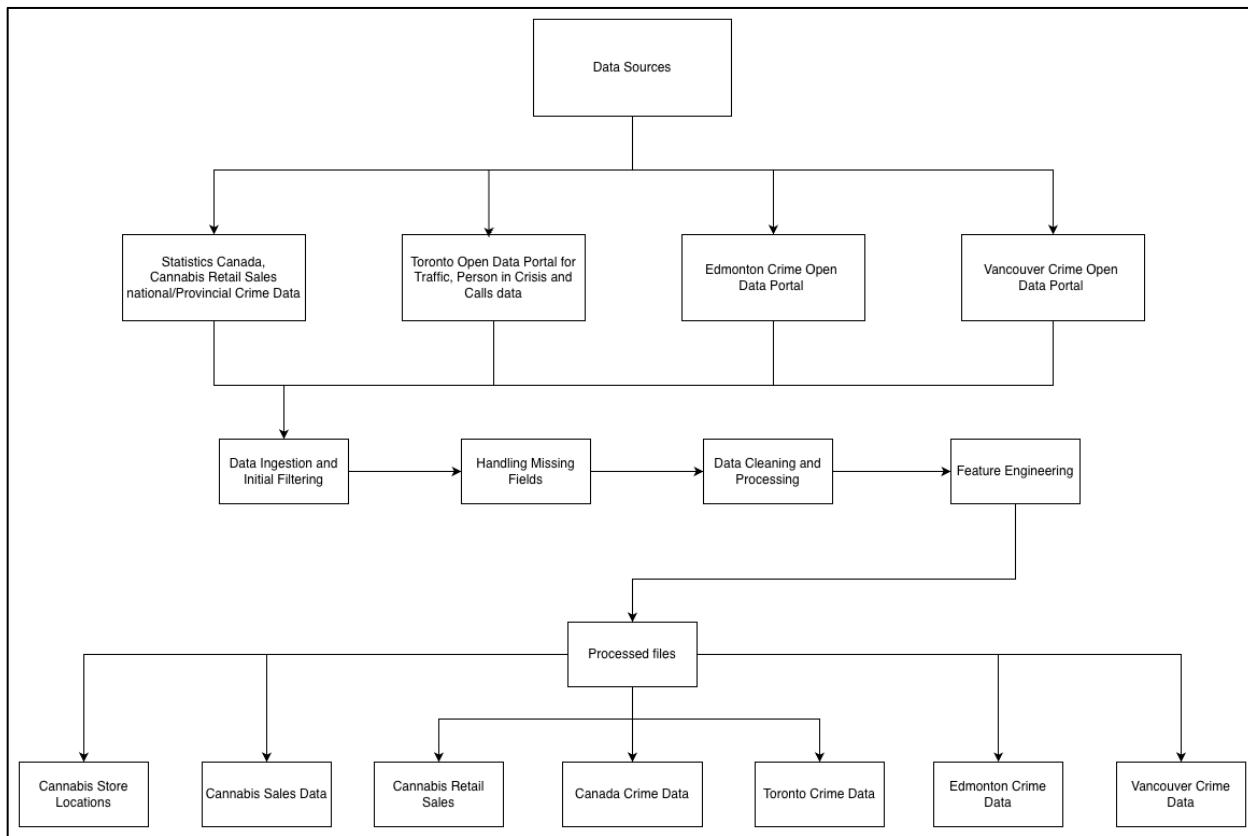
Another city we decided to consider was Vancouver and the Municipal data portal provided by this city had detailed dataset with reference dates from 2014.[\[7\]](#) The dataset contained important information such as the column named Category which contained types of crimes such as Break and Entering, Homicide, Mischief and types of Theft. With all the information in one place, we could visualize the crimes by its categories.

Edmonton Crime Data

Edmonton had a lack of data as the datasets ranged from 2022 to 2025 separated in different CSV files by the year. Although, each year data had about 80K rows of data classified into category and category by group. Categories included Disorder, Drugs, Traffic, Non-Violent, Violent, Weapons and Others. Categories by Group included Drug Violations, Mischief, Explosives, Personal Violence, Property, Sexual violence and many more. [8]

The categorization of the data was very detailed, and the format had been maintained among the years, so it was easier to combine the files into a single CSV file for Edmonton.

With the acquisition of these datasets our first step of the process was complete. The next steps included cleaning and preprocessing of the data according to our needs. An overview of this process has been put together in a flow chart below.



2.2 Data Publishing

Cannabis Store Locations datasets for each province were cleaned, processed and merged into a single comprehensive CSV file. Each data for store location was enhanced by adding Latitude, Longitude and Postal codes with the use of Google Geocoding API to ensure high accuracy of spatial coordinates.

As the dataset was newly created and enhanced, it was published as an open dataset to the Kaggle platform. This publication was followed by the complete Python code used for the data, cleaning, merging, geocoding information retrieval and final CSV generation, to make the methodology transparent and reproducible.

The screenshot shows a Kaggle dataset page titled "Cannabis Store Locations Across Canada". The page includes a search bar, a user profile icon, and a timestamp "JAINISH PATEL AND 3 COLLABORATORS - UPDATED 7 DAYS AGO". It features a "Code" button, a "Download" button, and a map titled "CANADA'S CANNABIS RETAIL MAP". Below the title, it says "Licensed Cannabis Store Locations in Canada (Geocoded)". A "Data Card" tab is selected, showing "About Dataset" details. The dataset is described as providing a comprehensive list of publicly licensed and operating retail cannabis store locations across Canada, with essential details like Store Name, City, Province, Full Address, Postal Code, and geocoded Latitude and Longitude coordinates. It also mentions the "Data Collection and Provenance" from official public listings provided by provincial and territorial regulatory bodies. The "Tags" section includes "Tabular", "Intermediate", "Data Cleaning", "English", and "Canada".

Here is the link to the Kaggle dataset we created:

<https://www.kaggle.com/datasets/jainishpatel31/cannabis-store-locations-across-canada/data>

2.3 Tools, Frameworks and Software Development

The project's execution was relied on a set of technology stack and tools to handle the complexity of data processing, analysis and results generation.

- **Programming Language:** The core programming language used was Python 3.x for all ETL operations, supported by libraries.
- **Core libraries:** Pandas, NumPy, Matplotlib, Seaborn, OS, Time, Requests, PyProj
- **APIs:** Google Geocoding API
- **Development Environment:** Google Collab, Jupyter Notebooks
- **Data Sharing Platform:** Kaggle
- **Business Intelligence:** Power BI
- **Version control system:** GitHub (Link to the repository:
https://github.com/polonium31/SEP_6DA3_Retail_Access_vs_Social_Cost)

3. Data

The project integrates multiple datasets representing cannabis retail sales, store locations, Canada wide crime analysis and major cities crime analysis. These data originate from federal, provincial and municipal open-data platforms. Analyzing the relationship between these datasets is key to the project. All raw data are stored in a consistent directory structure and retained in the original format and all the processed data are generated for analysis.

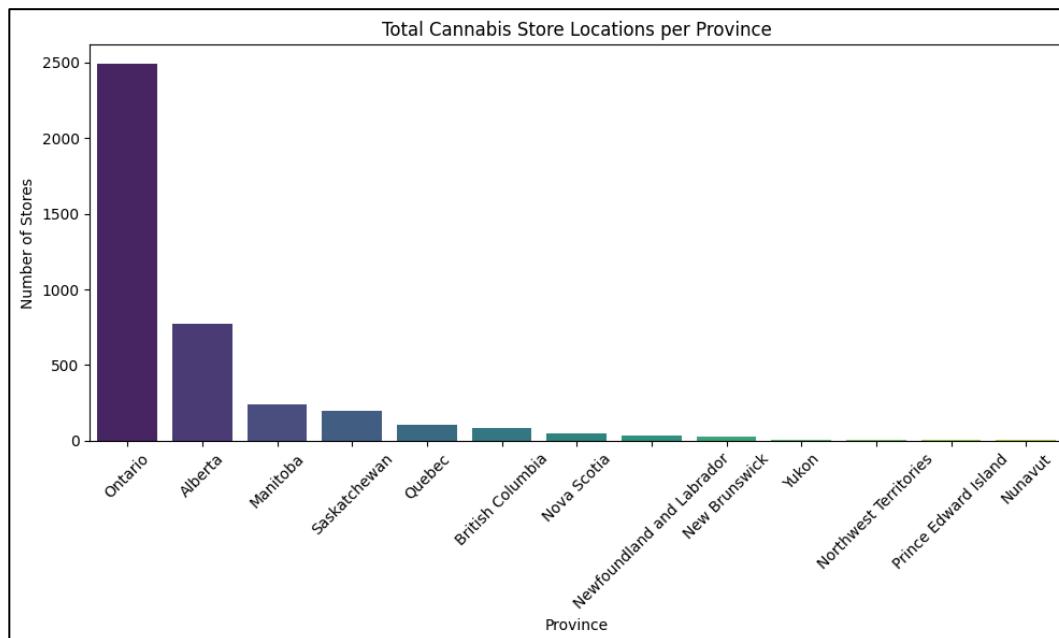
3.1 Data Preprocessing

Data preprocessing focused on creating consistent and analysis ready datasets applied over the original data which varied in structure, format and granularity.

I. Store Locations Data Preprocessing

Each province in Canada maintained its own store-location dataset, typically provided in CSV format. These provincial files were consolidated into a single national store-location dataset containing 4,041 rows and 8 columns. Most of the original files lacked essential fields such as postal codes, longitude, and latitude. To address this, the Google Geocoding API was used to enrich each store address with its corresponding postal code and geographic coordinates.

After completing the geocoding and ensuring all missing attributes were populated, a final cleaned and standardized CSV file was generated as the processed national store-locations dataset. The spatial distribution of cannabis store locations across Canada is shown below.



Cannabis sales Data Preprocessing

The cannabis sales dataset originally contained 15 columns, of which 9 key columns were retained for analysis. Sales values were standardized by multiplying the raw figures by 100, and the corresponding store counts were included in the processed subset to provide context for sales distribution. The final cleaned dataset consists of 588 rows and 9 columns, ready for downstream analysis and visualization.

Retail Trade Data Preprocessing

The retail trade dataset was first filtered to include only records classified as “Cannabis retailers” based on the NAICS column. Columns with more than 50% missing values were removed, and rows with missing entries in the target variable VALUE were dropped. Data types were optimized by converting REF_DATE to datetime, setting VALUE as numeric, and casting text fields to categorical for memory efficiency. The resulting cleaned dataset, encompassing its full-time range and geographic coverage, was saved for downstream analysis and consists of 1,801 rows × 14 columns.

Crime Statistics Data Preprocessing

The crime statistics dataset was prepared for integrated analysis by aligning its geographic coverage with the processed cannabis sales data using unique location identifiers (DGUID). The Violations column was transformed to separate the Violation Name from its corresponding Violation Code. The dataset was then divided into two components: a provincial and national-level crime dataset containing only locations present in the sales data, and a city-level crime dataset covering all other detailed locations. Both datasets were cleaned by selecting key variables, renaming columns for clarity, and removing rows with missing sales data.

The processed datasets were saved as canada_and_province_crime_data.parquet and city_wise_crime_data.parquet, 345,2294 rows × 8 columns for city-level crime, and 1,414,386 rows × 8 columns for national/provincial crime.

City specific Crime Data Preprocessing

Detailed crime data for Toronto, Vancouver, and Edmonton were utilized and cleaned in preparation for analysis. For Toronto, traffic collisions, persons-in-crisis incidents, and calls-for-service records were standardized. For Edmonton, the initial geographic coordinates were incorrect. The Google geocoding API was used to derive the correct values. Finally, Vancouver's data was prepared by concatenating multiple years and standardizing the event dates into one consistent column. The cleaned and geocoded datasets obtained after these various steps for each city were written to their respective city-specific files for subsequent spatial analysis.

Toronto:

- Traffic collisions – 772,516 rows × 7 columns
- Persons-in-crisis – 344,935 rows × 4 columns
- Calls-for-service – 5,399 rows × 4 columns

Vancouver: Crime data (multiple years concatenated) – 462,390 rows × 6 columns
Edmonton: Crime data (geocoded to correct coordinates) – 313,547 rows × 11 columns

3.2 Exploratory Data Analysis (EDA)

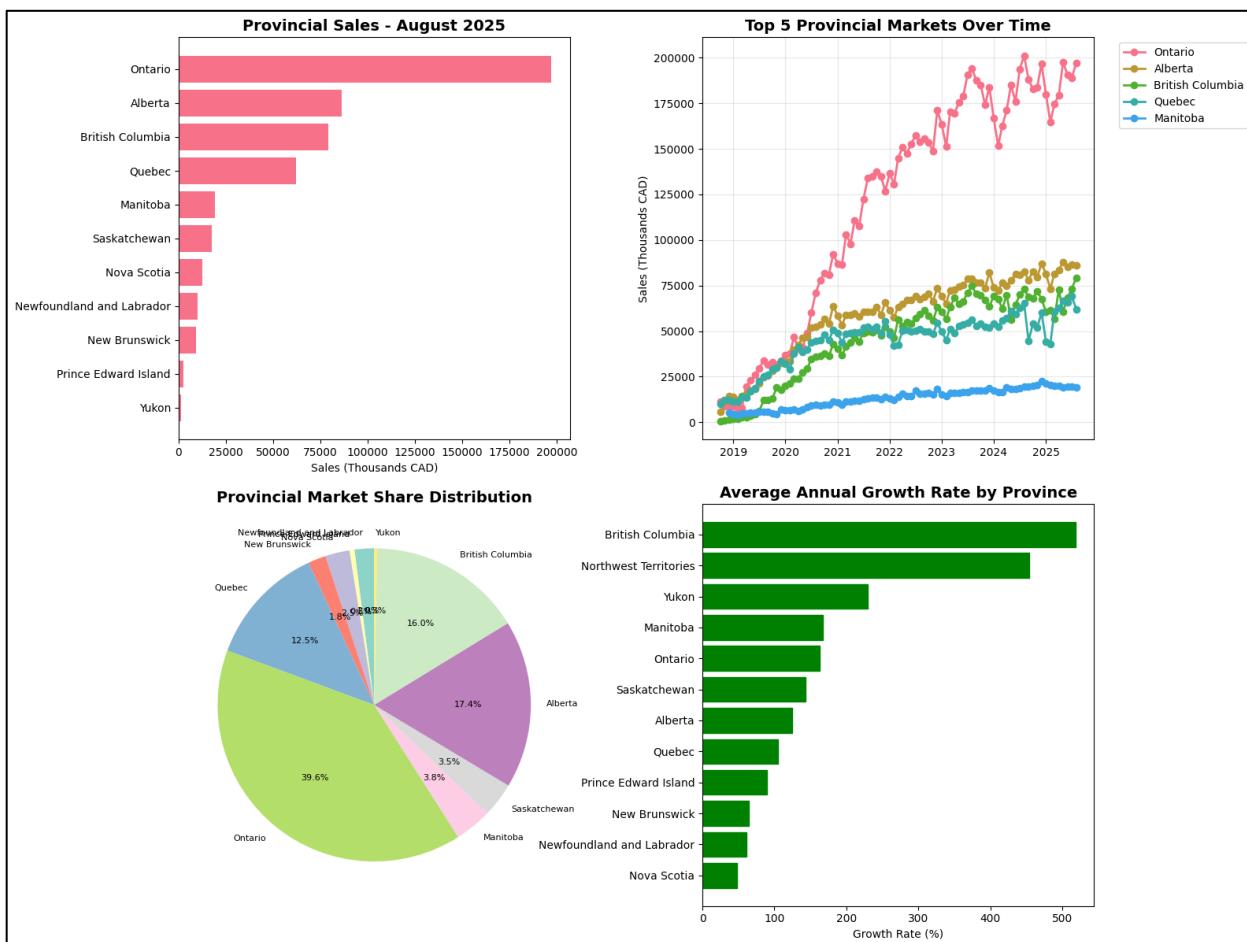
To understand the structure, distribution and initial patterns within the cleaned data, EDA was performed. The procedure included both statistical and visual analysis.

Cannabis Sales Data Exploration

Visualization of the breakdown of sales value by cannabis product type and consumption pattern is performed by plotting detailed bar charts for each geographic location and reporting period. Similarly, time series bar charts were generated for the visualization of the trend in the total cannabis sales across reporting periods of each location.

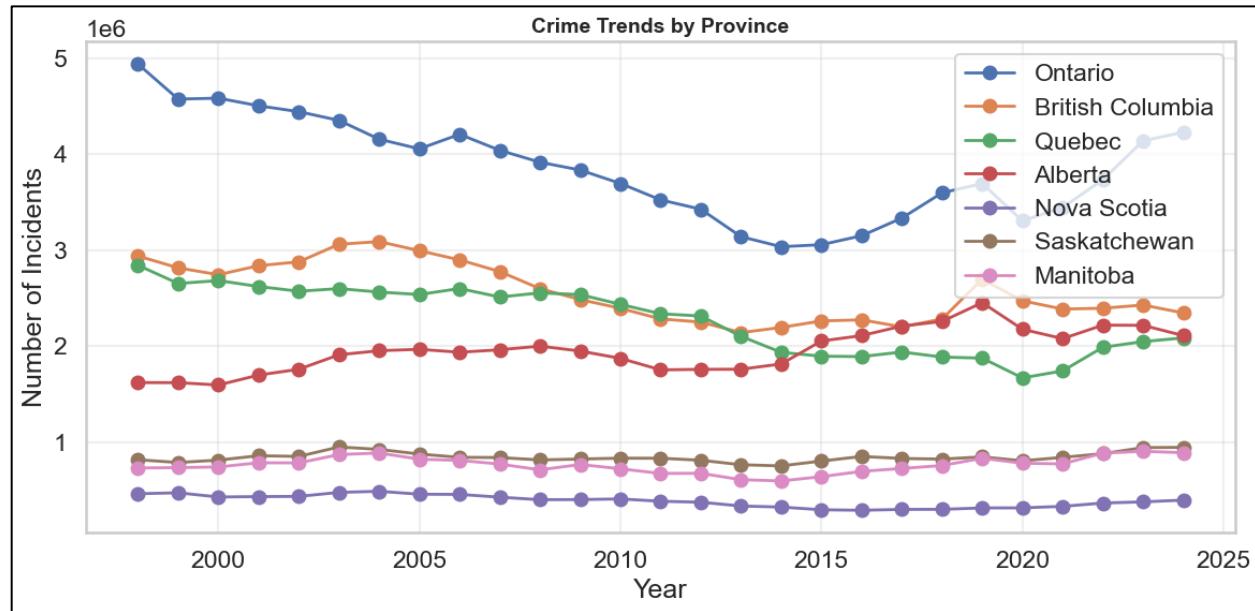
Retail Trade Data Exploration

The retail trade sales data is used to develop several visualizations and statistical metrics at the National, Provincial, and city levels. Geographic market share was established, temporal sales trends were assessed on a monthly and annual basis, while market concentration was looked at using Herfindahl-Hirschman Index to inform strategic insights.



Crime Statistics Data Exploration

Key insights visualized for crime statistics are a four-panel chart tracking total crime incidents over time for Canada, identifying the top 10 violation types, and comparing total crime volume across provinces and locations. The dedicated time-series analysis was done using a line plot to compare the crime trends across the major provinces that distinctly showed how different Canadian regions had evolved over the period under consideration.



City specific Crime Data Exploration

The specific analysis entailed in-depth EDA of city-specific crime and public safety data in Toronto, Edmonton, and Vancouver. This included feature extraction for temporal features along with visualizations that map annual trends, monthly seasonality, and hotspots of crime, for example, top neighborhood/intersections for different types of incidents.

3.3 Feature Extraction

Feature extraction focused on generating new variables that improved the correctness and analytical depth of the new data.

The preparation of data included necessary Feature Extraction to obtain analytical variables and allow the comparison of markets at many levels. This comprised the calculation of Actual_Sales_CAD through scaling raw sales values reported in thousands by 1,000, and the creation of temporal features such as Year, Month, and Quarter, derived from the date fields for time-series analysis. In crime and public safety datasets, parsing the complex Violations column into distinct

Violation Name and Violation Code features was performed, with binary collision indicators combined into a single categorical Collision_Type.

Importantly, raw geographic identifiers were converted to generate a geographic hierarchy: National, Province, and City. Where appropriate, coordinate systems were converted or geocoding APIs used to assign precise latitude and longitude to urban incident locations for spatial analysis.

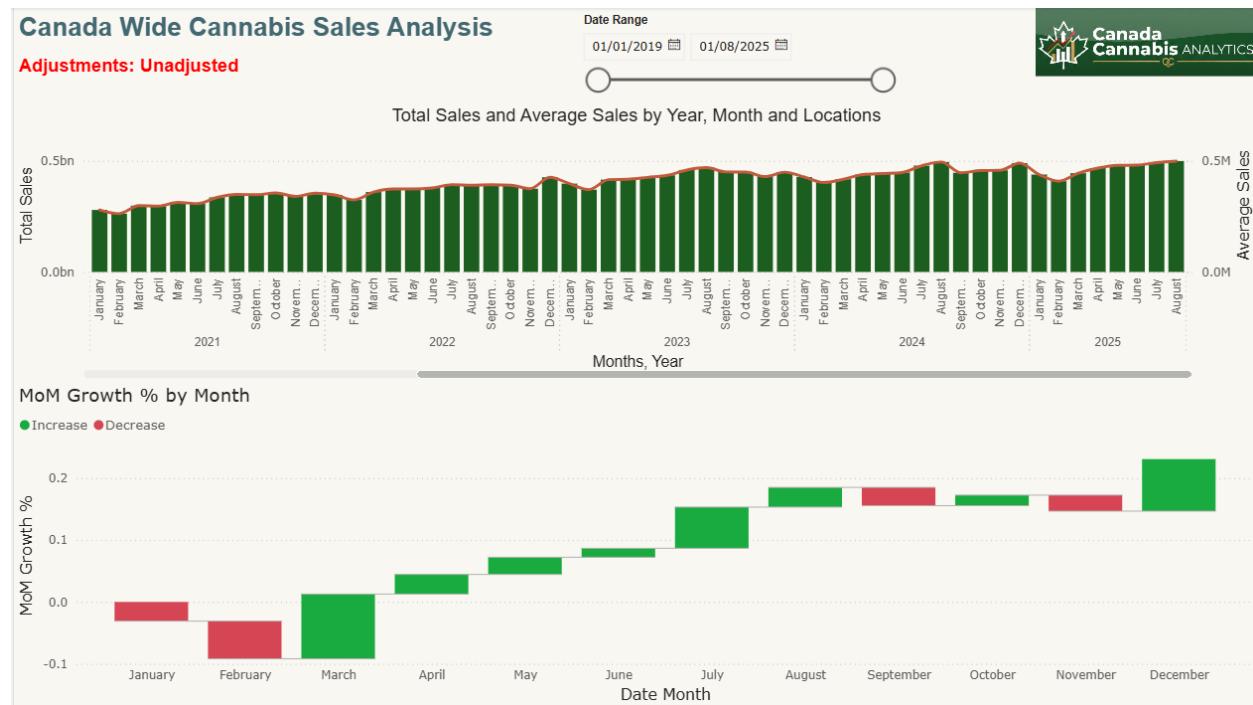
4. Experiments & Results

To quantify the scale of data handled in this analytics project:

- **Total Rows Processed:** *6.77 million* records across all sources
- **Number of Datasets:** *10 distinct datasets* spanning CSV and Parquet formats
- **Data Pipeline Runtime:** *3 hours, 30 minutes, 45 seconds* from raw ingestion to final processed outputs

Canada Wide Cannabis Sales Analysis Dashboard

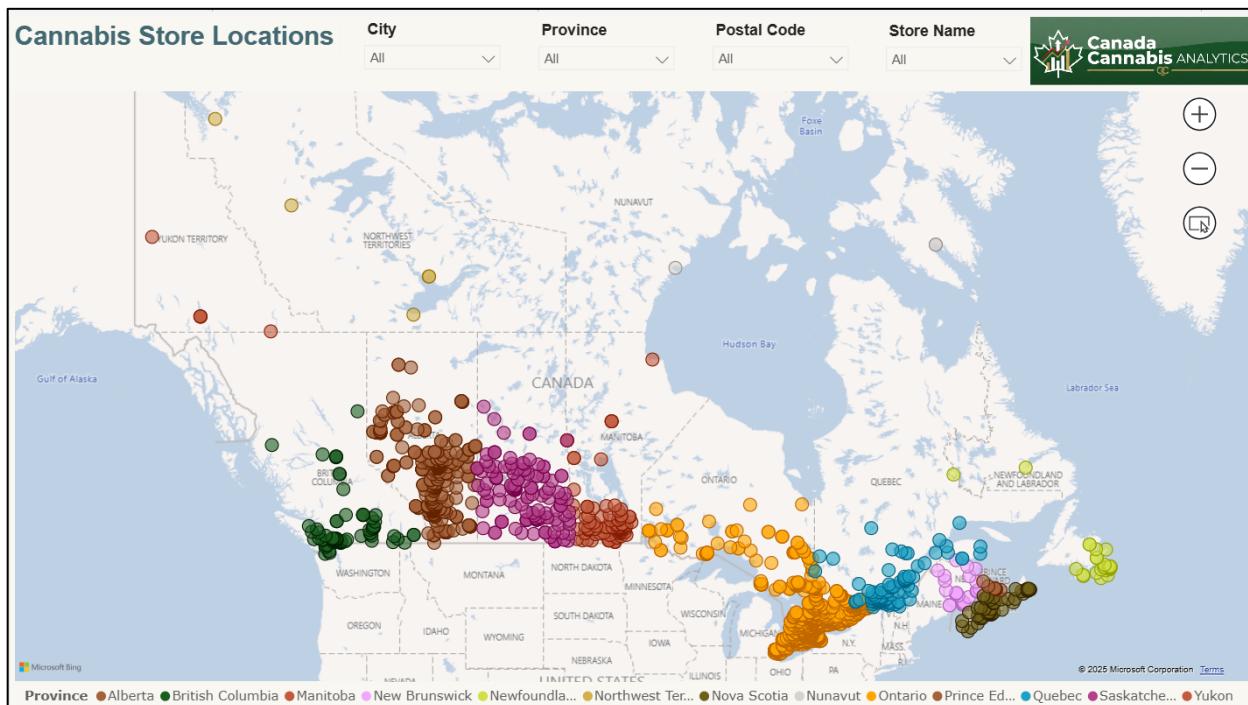
The Canada Wide Cannabis Sales Analysis view is tracking the changes in cannabis sales over the time since its legalization, showing monthly sales, trends, and growth patterns from 2019 to 2025.



Monthly cannabis sales rose from about \$200 million in early 2019 to over \$500 million by 2022, driven by more store openings and growing consumer comfort with the legal market. The rolling 12-month trend shows steady growth until mid-2022, after which sales plateau and dip slightly, suggesting a maturing market where demand may be reaching saturation or the illegal market remains competitive. Month-to-month growth has also become more volatile, early years saw consistent gains, while recent years show frequent swings between growth and decline, another sign of market stabilization.

Cannabis Store Locations Dashboard

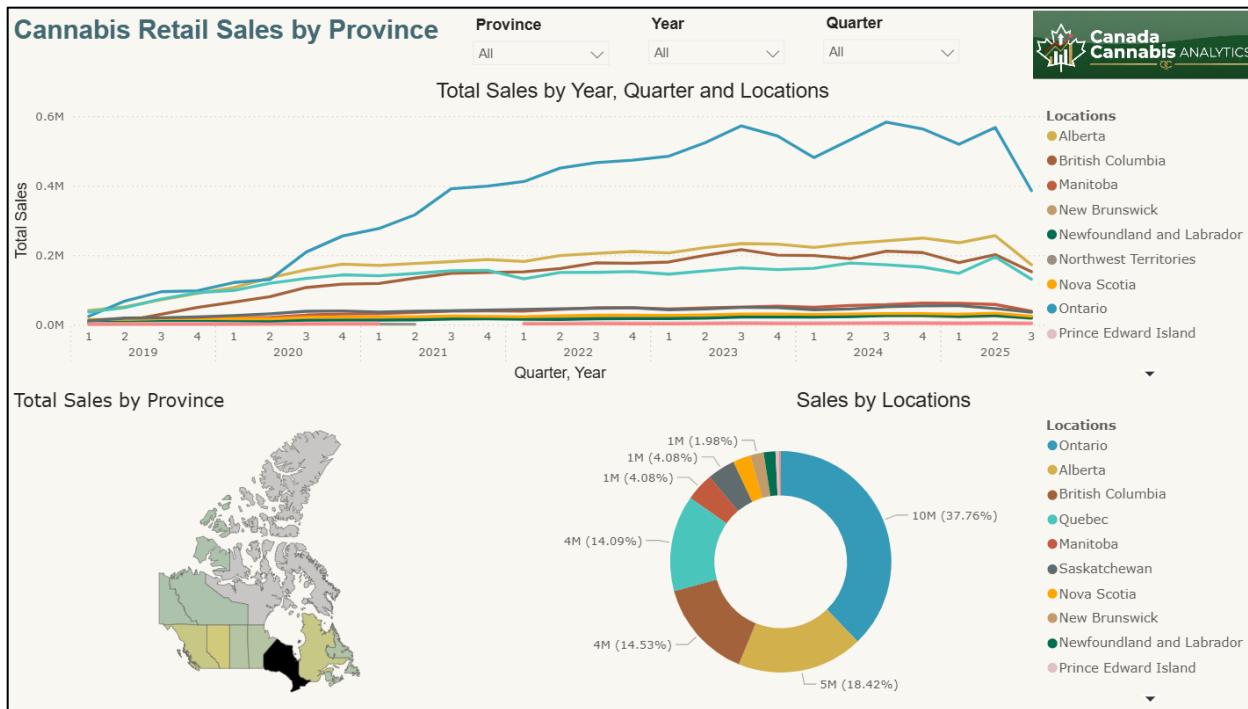
The Cannabis Store Locations view is using an interactive map that display where cannabis stores are located across Canada. You can filter by city, province, postal code, or even specific store names to explore different areas.



Ontario has the most cannabis stores, but they're concentrated in Toronto and southern Ontario where most people live. Alberta stands out with a very high number of stores relative to its population, especially clustered around Calgary and Edmonton due to its permissive retail policies. British Columbia's stores are focused on Vancouver and the coastal region, while Saskatchewan and Manitoba show more dispersed patterns. The Atlantic provinces have fewer stores overall, reflecting their more government-controlled retail models.

Cannabis Retail Sales by Province Dashboard

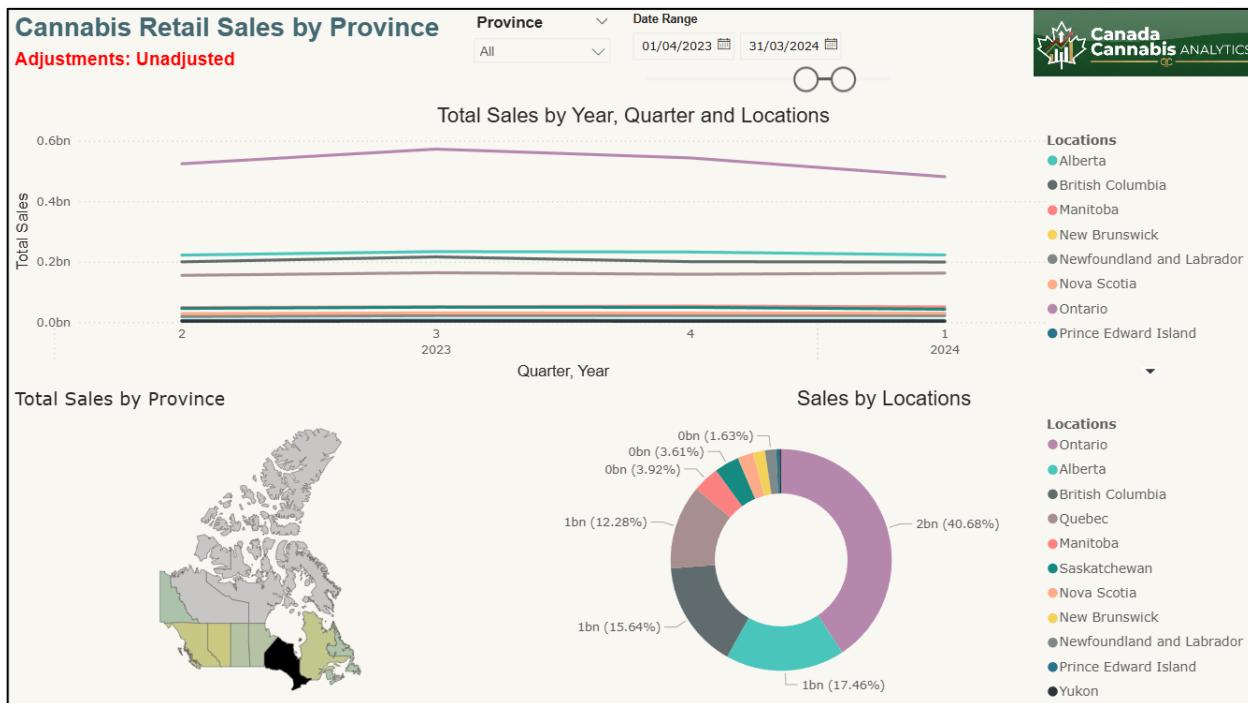
This view breaks down cannabis sales performance across Canada's provinces and territories, showing how the different regulatory approaches have translated into market outcomes.



Ontario drives the cannabis market with 37.76% of national sales roughly matching its share of Canada's population while Alberta stands out by generating 14.53% of sales despite having only about 11% of the population, highlighting how its open-market model boosts per-capita consumption; British Columbia contributes 14.09%, lower than expected given its strong cannabis culture, suggesting its cautious rollout may have allowed illegal markets to persist. The time-series trends show Ontario's steady dominance, with Alberta and B.C. growing more moderately, and most provinces peaking around 2022–2023 before flattening or declining, indicating market saturation nationally. Smaller provinces and territories remain near the bottom, reflecting their smaller populations and often more restrictive retail systems.

Cannabis Retail Sales by Cities Dashboard

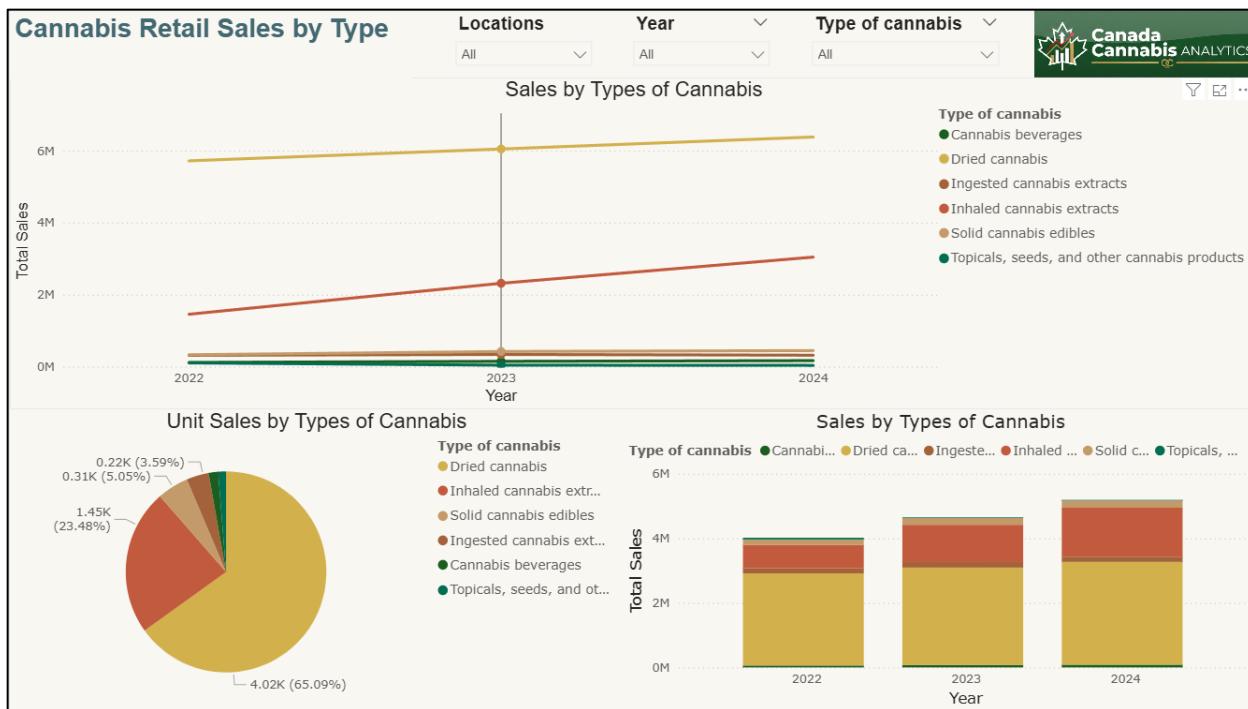
The city-level analysis shows which urban centers drive provincial performance and how the metropolitan markets compare against each other.



Toronto overwhelmingly leads national cannabis sales, far outpacing every other city, with Montreal a distant second and Edmonton, Calgary, Vancouver, and Ottawa following. This strong urban concentration is expected, since larger cities have dense populations, higher disposable incomes, and typically more cannabis-friendly attitudes. Over time, Toronto shows steady growth until about 2022 before dipping, while smaller Alberta cities like Edmonton and Calgary appear more resilient, likely due to a less saturated market. The pie chart further highlights Toronto's dominance at 29.38% of major city sales, followed by Montreal at 16.34% and Edmonton at 13.64%, reflecting Alberta's supportive retail environment.

Cannabis Retail Sales by Type Dashboard

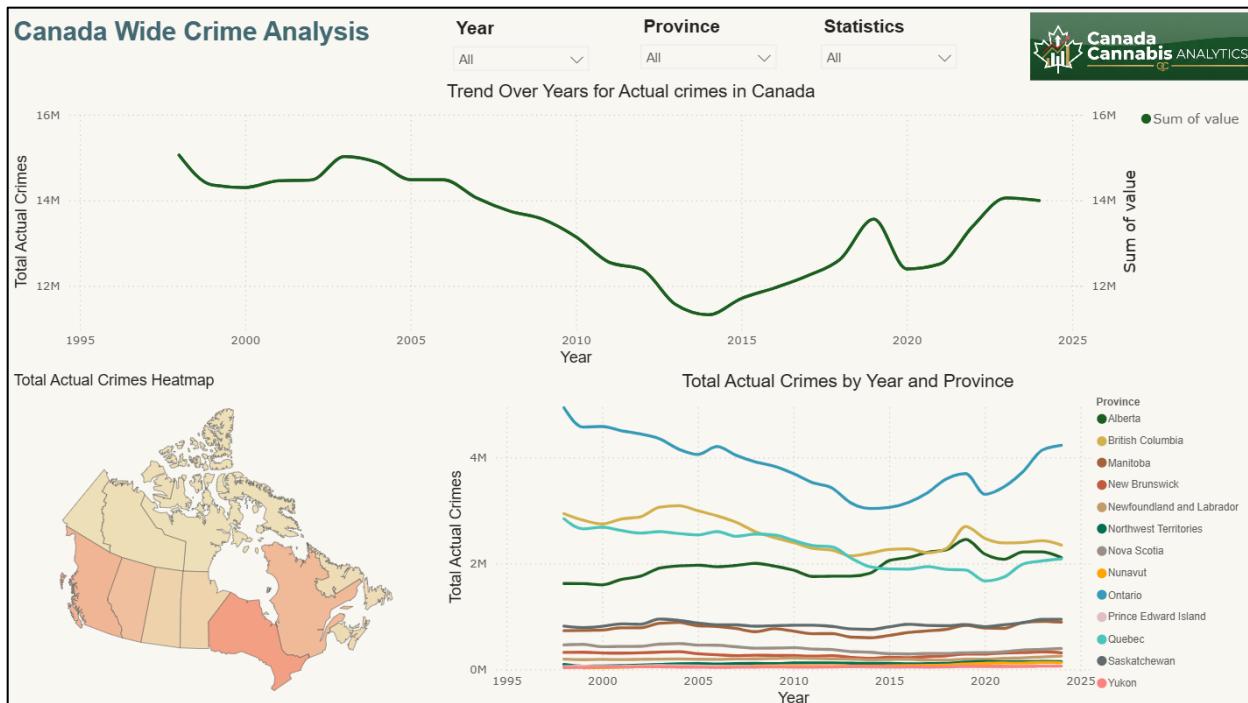
This view shows how consumer preferences are shifting across various cannabis product categories, with a demonstration of dominance from traditional products and growth among newer formats.



Dried cannabis dominates the market at about 65% of sales, reflecting its early availability and consumers' familiarity with traditional smoking; its trend line has been stable from 2022 to 2024, suggesting the category has reached maturity. In contrast, alternative products are growing: inhaled extracts have nearly doubled over the period as consumers adopt vaping and concentrates, and solid edibles show steady gains, now making up roughly 23.48% of unit sales despite being legalized later. Drinks, topicals, and other niche formats remain small but highlight the sector's diversification, appealing especially to nonsmokers and signaling a maturing, more segmented market.

Canada Wide Crime Analysis Dashboard

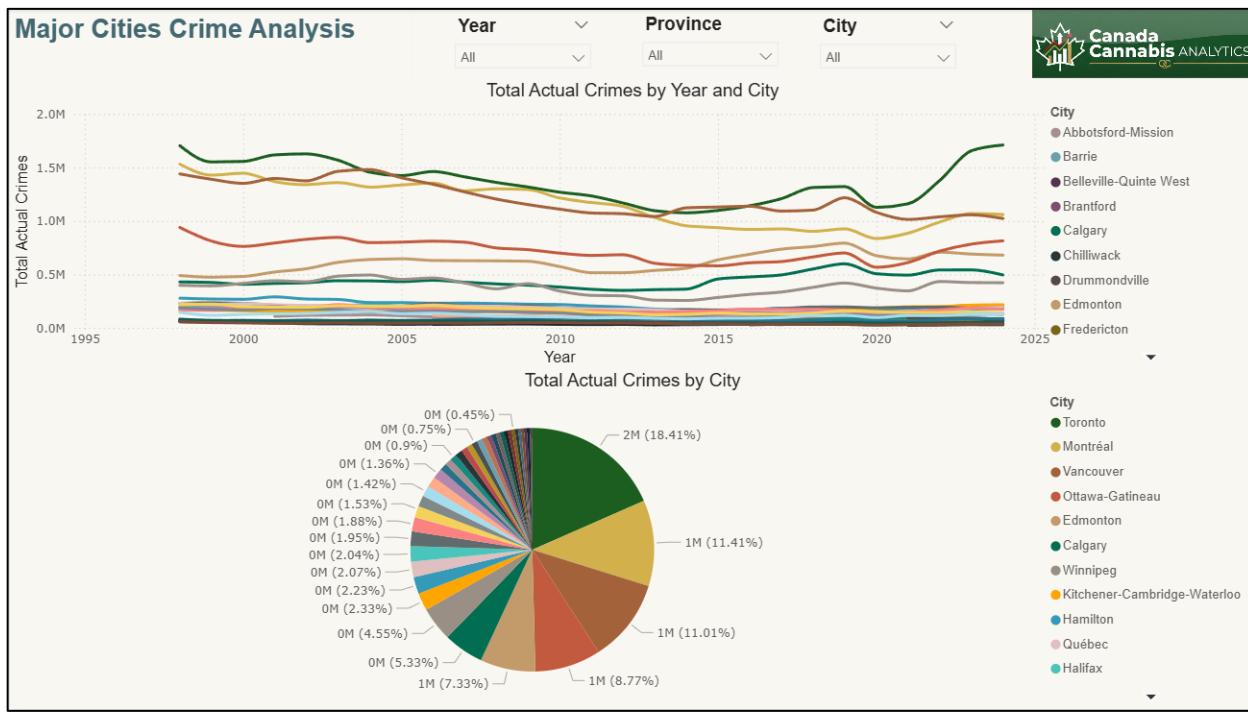
This view puts into a single place the trends in crime across Canada, both the long-term national patterns and provincial variations that help place the period of cannabis legalization into context.



Canada's crime trend shows a long, steady decline from a peak of about 16 million incidents in the late 1990s to roughly 11 million by 2014, followed by a modest rise to around 14 million by 2025, still far below levels from 20–30 years ago, highlighting that cannabis legalization in 2018 occurred during historically low crime levels. Provincial patterns largely mirror this national trend, with Ontario showing the highest crime counts due to its population, British Columbia reporting consistently high levels relative to its size, and Alberta and Quebec sitting in the middle while smaller provinces remain low. Across regions, crime declined through the 2000s and 2010s before a recent uptick. The geographic heat map reinforces that higher crime aligns primarily with population density and urbanization, not cannabis policy.

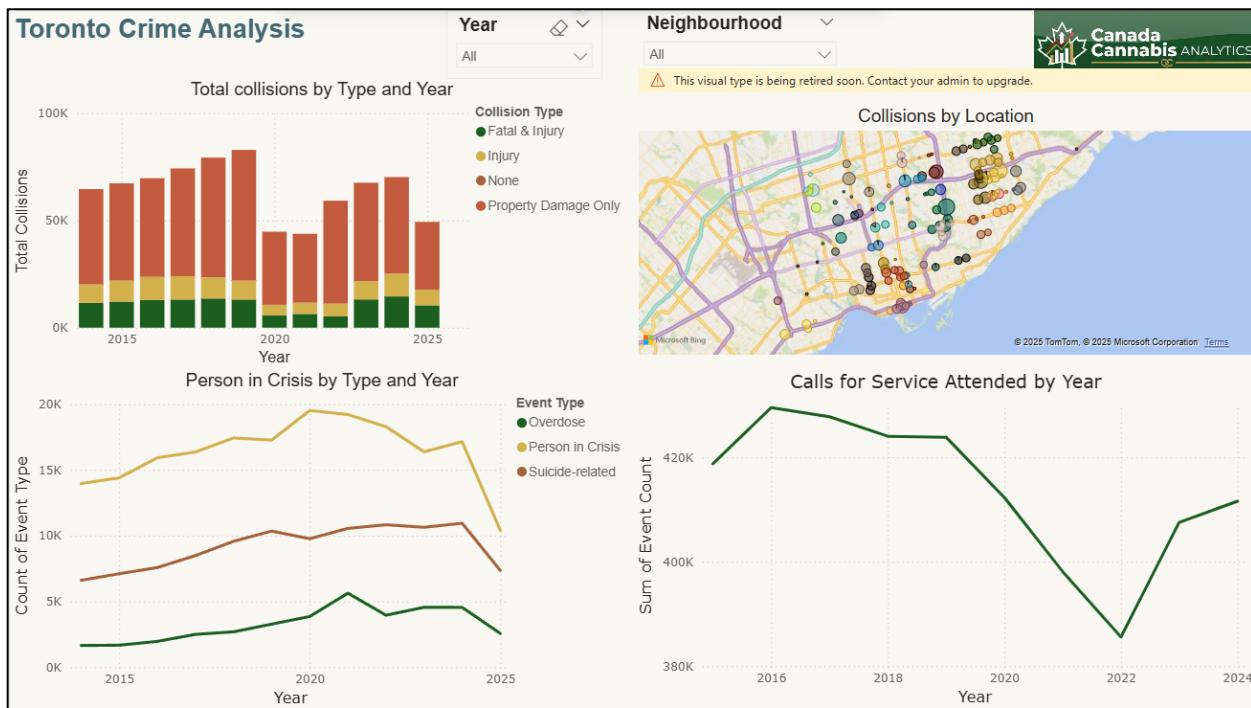
Major Cities Crime Analysis Dashboard

The following view zooms in on Canada's major urban centres, both where most of the country's crime occurs and where most cannabis retail happens.



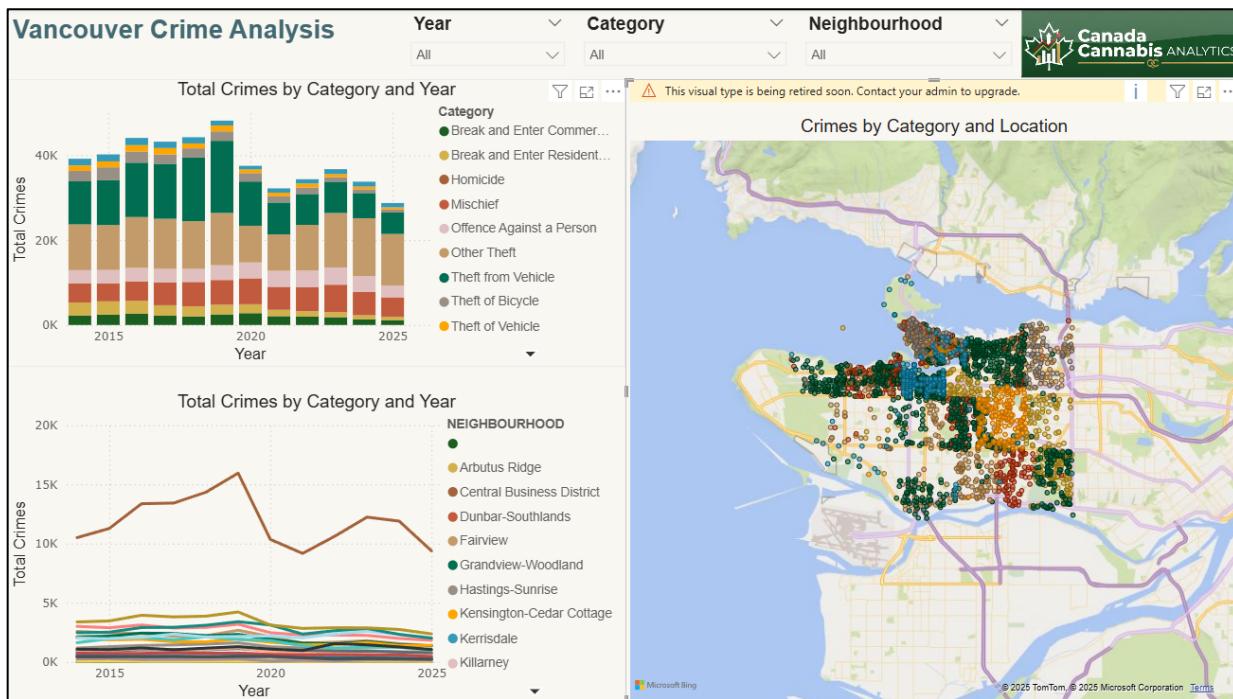
The pie chart shows crime heavily concentrated in major cities, with Toronto leading at 18.41% of incidents, followed by Montreal at 11.01% and Vancouver at 8.77%, reflecting their large urban populations; however, comparing this to earlier cannabis-sales data shows the proportions don't match perfectly, suggesting crime levels aren't directly tied to cannabis retail density. The long-term city-level trends mirror the national pattern high crime in the 1990s, steady declines through the 2000s and early 2010s, and some recent fluctuations most clearly seen in Toronto's trajectory, with Vancouver, Montreal, and other cities following similar shapes at different scales. Importantly, the small upticks seen between 2017 and 2020 don't align consistently across cities and don't indicate any meaningful link to cannabis legalization; cities with dense cannabis retail like Calgary and Edmonton still show stable or declining patterns, reinforcing those broader economic and social factors, not cannabis policy, are the main drivers of crime trends in Canada's major urban centres.

Toronto Crime Analysis Dashboard



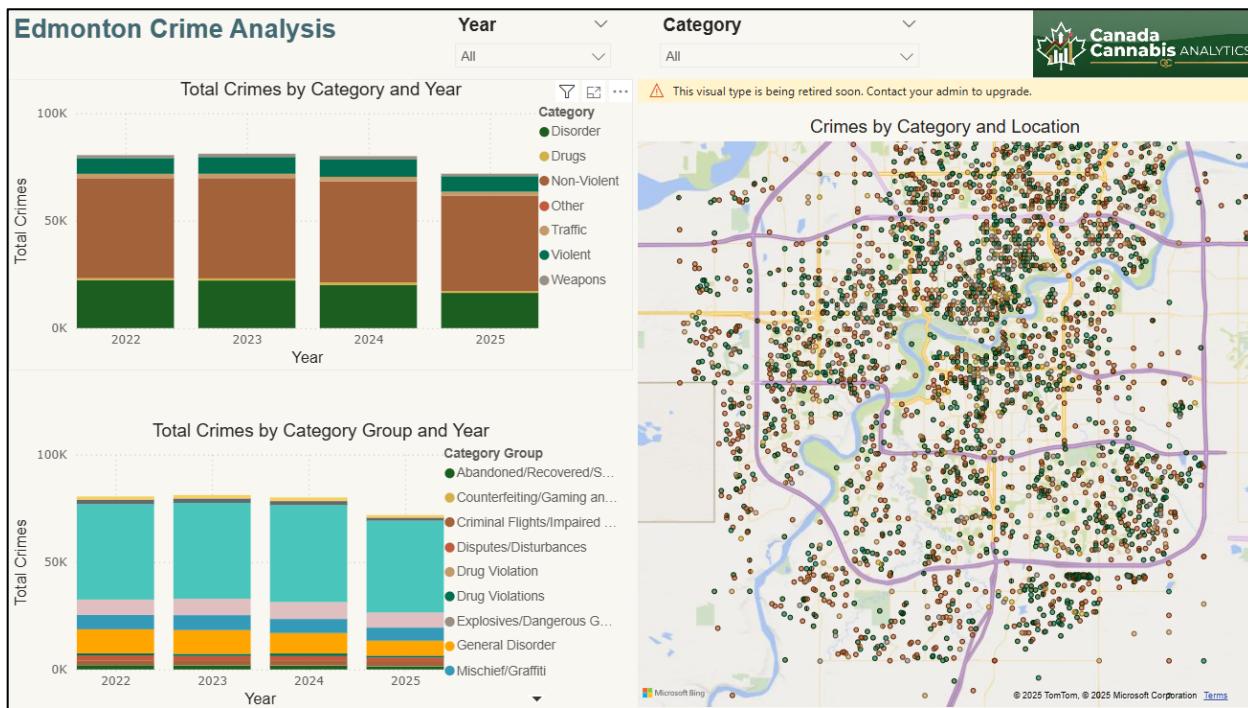
The collision data shows that most incidents in Toronto are traffic-related, with property-damage collisions dominating, followed by injury collisions, while fatal collisions remain rare. A clear dip appears around 2020 to 2021, likely due to reduced driving during the pandemic, with totals rebounding afterward but not fully returning to pre-COVID levels. The “Person in Crisis” trends highlight rising mental-health-related calls, peaking around 2020 to 2021, underscoring challenges that extend beyond traditional crime. Meanwhile, the “Calls for Service” data shows a sharp pandemic-era drop and later recovery, offering important context for interpreting crime and public-safety patterns during a period of major social disruption.

Vancouver Crime Analysis Dashboard



The stacked chart shows that theft-related crimes such as break and enter, vehicle theft, and other thefts dominate Vancouver's crime statistics, while violent crimes like homicide remain relatively low. From 2015 to 2025, crime peaked around 2017–2018, declined through 2020–2021, and has since fluctuated, showing no clear link to cannabis legalization in 2018. Geographically, crime is concentrated in the Central Business District, reflecting higher human, business, and nightlife activity, whereas residential neighborhoods have lower rates, though overall trends of decline and recent stabilization are similar across most areas.

Edmonton Crime Analysis Dashboard



Edmonton's crime data from 2022 to 2025 is notably stable, unlike the volatility seen in some other cities. Stacked bars show non-violent crimes, in brown, as the largest category, followed by violent and other offenses, while broader groupings highlight "General Disorder," mischief, and drug-related crimes as significant, and "Explosives/Dangerous Goods" and counterfeiting as minor. Geographically, crime is spread across the city rather than concentrated in the downtown core, reflecting Edmonton's more sprawling urban layout compared to Vancouver.

5. Discussion

Interpretation

1. **Market Growth Reaching Maturity:** Sales of cannabis increased quickly between 2019 and 2022, but they have since leveled down, indicating that the sector is maturing. Volatile month-to-month increase further implies stabilization rather than ongoing exponential expansion.
2. **Provincial and Urban Sales Concentration:** Ontario and Alberta are the primary market drivers, with Alberta having particularly high per-capita sales. Strong urban-center demand is demonstrated by the dominance of major cities, particularly Toronto, in national sales.
3. **Changing Product Preferences:** Although edibles and inhaled extracts are becoming more popular, dried cannabis continues to dominate the market. As usage habits change, these alternative formats are gradually gaining market share.
4. **Crime Trends Unaffected by Legalization:** National crime statistics show no major movement owing to cannabis legalization, maintaining consistent with long-term patterns. Crime rates are still mostly influenced by urbanization and socioeconomic variables.
5. **City-Specific Safety Insights:** Toronto observes high incident counts but pandemic-related declines in calls, Vancouver suffers concentrated stealing activity downtown, and Edmonton exhibits stable crime spread across its entire region. Crises pertaining to mental health are growing and influencing contemporary needs for public safety.

Challenges

1. **Determining the actual impact of retail on social cost:** Determining the impact of cannabis retail from other external factors, such as population growth, changes in enforcement, economic conditions, or reporting standards, is the most difficult task. Because social cost is a broad notion that covers crime, crisis calls, traffic incidents, health impacts, and community wellness.
2. **Spatial accuracy and geocoding issues:** Retail shop addresses must be precisely geocoded, but social cost events typically lack precise coordinates or contain rounding errors. These inconsistencies diminish the reliability of spatial proximity analysis between retail density and social consequences.
3. **Missing, inconsistent, and unstandardized datasets:** The column names, structure, and completeness of major cities related crime datasets vary significantly between provinces and years. It takes effort to clean up and reconcile these discrepancies, and thorough validation is necessary to prevent bias.
4. **Handling huge, high-volume datasets:** Social cost databases like as crime and collisions can surpass several hundred thousand rows, making merges, filtering, and

spatial joins computationally demanding. Efficient processing requires optimized pipelines, chunking, or memory-aware approaches.

5. **Choosing meaningful KPIs for both retail and social cost:** Retail indicators like sales and shop count are straightforward, while social cost KPIs such as severity, per capita burden, crisis intensity, or collision rates require careful formulation. Poorly defined KPIs might affect analysis and lead to erroneous conclusions.
6. **Identifying distinct correlations or trends:** The relationship between retail characteristics and social indicators may seem weak, noisy, or inconsistent even after data has been cleaned and aligned. Identifying true trends requires normalization, per capita adjustments, and rigorous comparison methodologies.

6. Conclusion & Future Work

Summary

The legalization of cannabis in Canada has created an industry that is worth \$4.5 billion that rapidly grew from its inception until 2022, then it started stabilizing ,a sign of a mature market. Various different provincial methodologies yielded strikingly different results: an open-market policy led to the highest per-capita consumption in Alberta, with only 11% of population, while Ontario's cautious lottery system eventually seized 37.76% of the nation's market share. Variations in the regulatory framework have thus demonstrated direct effects on market development and adoption rates.

While the initial study showed only weak associations between cannabis retail density and crime rates, further scrutiny into the temporal trends in 2018-2025 discloses a disturbing pattern. It evidences that cannabis sales have grown in the same period as the incidents of crime, indicating that cannabis can be one of the factors contributing to rising crime rates. Although this has been highlighted by the research as a weak association, the co-extensive upward trend calls for serious consideration for policy evaluation.

Future work

This analysis gives us a solid starting point for understanding cannabis policy impacts, but we clearly need to dig deeper into any potential connections between growing cannabis markets and rising crime rates. To really get to the bottom of this, future studies should bring in much richer data - things like detailed neighborhood demographics, economic conditions, how police are allocating their resources, and tracking patterns over many more years. These complex social issues don't happen in isolation, and we need to account for all the moving pieces.

With Canada's cannabis landscape still evolving and other jurisdictions around the world considering similar policy changes, continuing to collect comprehensive data and conduct thorough analysis will be crucial for making informed decisions that protect public safety while supporting legitimate business development.

References

- [1] Department of Justice Canada. (n.d.). Cannabis Legalization and Regulation. Retrieved December 3, 2025, from <https://www.justice.gc.ca/eng/cj-jp/cannabis/>
- [2] Health Canada. (n.d.). *Provincial and territorial regulations for cannabis*. Government of Canada. <https://www.canada.ca/en/health-canada/services/drugs-medication/cannabis/laws-regulations/provinces-territories.html>
- [3] Statistics Canada. (n.d.). *Table 10-10-0164-01: Sales of cannabis by liquor and cannabis authorities and other retail outlets, by cannabis type*. <https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1010016401>
- [4] Statistics Canada. (n.d.). *Table 20-10-0056-01: Monthly retail trade sales by province and territory*. <https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=2010005601>
- [5] Statistics Canada. (n.d.). *Table 35-10-0177-01: Incident-based crime statistics, by detailed violations, Canada, provinces, territories, Census Metropolitan Areas and Canadian Forces Military Police*. <https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3510017701>
- [6] Toronto Police Service. (n.d.). *Public safety data portal*. <https://data.torontopolice.on.ca/>
- [7] Vancouver Police Department. (n.d.). *Open data portal*. <https://geodash.vpd.ca/opendata/>
- [8] Edmonton Police Service. (n.d.). *EPS open data dashboard*. <https://experience.arcgis.com/experience/8e2c6c41933e48a79faa90048d9a459d/page/>