

# Crime Forecasting and Trend Analysis using **FBI Crime Data**



**SECTOR:** Public Safety and Law Enforcement Analytics

## **TEAM DETAILS:**

**Team ID:** Group 18

### **Team Members:**

- Agrima Gusain: Project & Dashboard Lead
- Adil Mirza: Data & Dashboard Lead
- Polana Rakshita: Strategy Lead
- Sanchit Garg: Analysis Lead
- Mishti Sharma: PPT Lead
- Om Mishra: Analysis Lead
- Ujjwal Bhardwaj: PPT & Quality Lead

**FACULTY:** Mr. Archit Raj

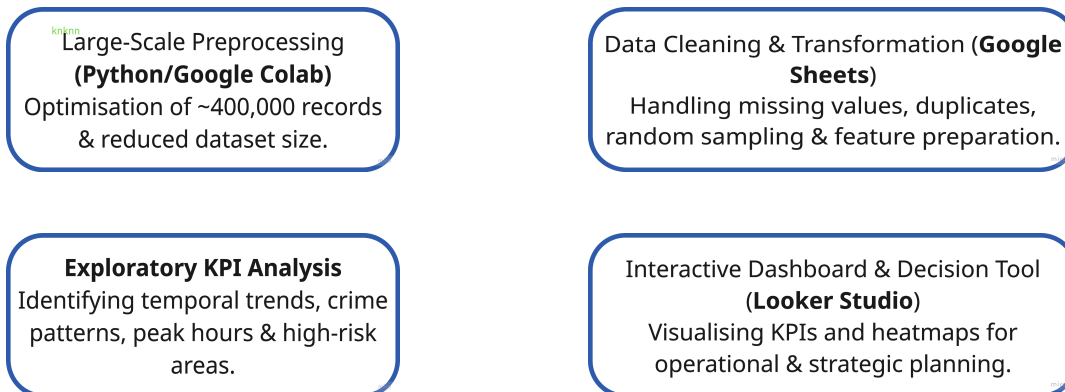
**INSTITUTE:** Newton School Of Technology

# EXECUTIVE SUMMARY

## PROBLEM

This project transforms historical **crime data** into an analytical dashboard that identifies **high-risk zones** and **peak crime periods**. The insights support evidence-based operational deployment and long-term strategic planning for crime prevention.

## APPROACH



## KEY INSIGHTS

- **Crime is structurally concentrated in commercial hubs**, with the Central Business District emerging as a dominant hotspot, driven primarily by vehicle-related offences.
- **Evening hours (3 PM–6 PM, peak at 6 PM) represent the highest operational risk window**, indicating mobility-driven urban crime patterns rather than purely late-night activity.
- **Property and vehicle-related crimes account for the majority of incidents (~37% Theft from Vehicle alone)**, significantly outweighing violent categories, with forecast signals suggesting renewed growth risk in commercial zones.

## KEY RECOMMENDATIONS

- **Adopt hotspot-focused patrol deployment**, prioritizing CBD and high-density commercial corridors during 3 PM–6 PM peak escalation windows.
- **Implement targeted vehicle-crime prevention strategies**, including parking surveillance, awareness drives, and environmental design improvements in commercial areas.

# SECTOR AND BUSINESS CONTEXT

## SECTOR

### ***Public Safety & Law Enforcement Analytics***

Law enforcement agencies collect large volumes of **incident-level data**, including **crime type**, **location**, and **time details**. By analyzing spatial and temporal patterns, authorities can shift from reactive to **proactive policing**, optimize **patrol deployment**, and implement targeted prevention strategies. With rapid urbanization and population growth, predictive crime intelligence systems are increasingly essential for ensuring public safety and operational efficiency.

## CHALLENGES

Fragmented  
Pattern  
Recognition

Despite large volumes of crime data, agencies lack structured dashboards to clearly identify **high-risk zones**, **peak crime hours**, and emerging **spatial-temporal patterns**.

Limited  
Analytical  
Visibility

There is insufficient integration between seasonal trends, yearly variations, and geographic concentration, making it difficult to uncover meaningful correlations.

Reactive  
Decision-  
Making

**Resource allocation** and **patrol planning** often rely on static reports rather than dynamic, data-driven insights, resulting in **reactive instead of predictive policing strategies**.

## WHY THIS PROBLEM?

Urban crime is influenced by both **spatial concentration (where)** and **temporal patterns (when)**. However, raw datasets alone do not provide actionable intelligence.

This project was selected to:

- Transform raw crime data into KPI-driven insights
- Support patrol planning, shift allocation, and preventive policy design

The final interactive dashboard bridges the gap between raw data and strategic decision-making by enabling crime trend exploration through filters, KPIs, and spatial heat mapping.

# PROBLEM STATEMENT

Urban crime patterns are shaped by both **spatial concentration (where crimes occur)** and **temporal dynamics (when crimes occur)**. However, despite the availability of detailed incident-level data, the absence of structured analytical dashboards limits meaningful interpretation. As a result, decision-makers face challenges in:

- Identifying high-risk neighbourhoods
- Detecting peak crime hours
- Recognizing seasonal and yearly trends
- Linking geographic concentration with time-based crime patterns

Without integrated analytical visibility, crime management strategies remain reactive rather than proactive, reducing the effectiveness of operational planning and long-term prevention efforts.

## SCOPE

1. **Data Preparation & KPI Design:** Cleaning and structuring historical crime data and developing key metrics (**total crimes, peak hour, top neighbourhood, etc.**).
2. **Temporal & Spatial Analysis:** Analyzing hourly, monthly, and yearly trends along with **geographic clustering** using latitude and longitude.
3. **Dashboard & Insight Delivery:** Building an interactive dashboard (**Google Sheets & Looker Studio**) and generating actionable operational and strategic recommendations.

## OBJECTIVES

<b>Interactive KPI Dashboard Delivered</b> (Dynamic, filter-enabled crime intelligence system)	<b>High-Risk Zones &amp; Spatial Hotspots Identified</b> (Neighbourhood concentration + geo heatmap visualization)	<b>Peak Crime Patterns Detected</b> (Hourly, monthly, and yearly trend analysis)	<b>Actionable Operational &amp; Strategic Insights Generated</b> (3–5 data-driven recommendations for patrol and policy planning)
---	---	---	--

# RAW DATA DESCRIPTION

## SOURCE

- **Name:** FBI Historical Crime Incident Dataset
- **Source:** Prior Internship Model training Raw Data used.
- **Access Link:**  
<https://1drv.ms/x/c/832fe5385f70cdc9/IQDKXlt6KHi4RqcZwWayP74yAQfJsSK6XZhUwN3D5pLSK5k>

## STRUCTURE

**Granularity:** One row per reported crime incident

**Format:** Structured tabular dataset (CSV → Google Sheets)

**Total Records:** ~400,000 rows (initial dataset)

**Total Variables:** 13 columns

**Time Coverage:** 1999 – 2011

**Data Type Mix:**

- Categorical variables (Crime Type, Neighbourhood, Block)
- Numerical variables (X, Y, Latitude, Longitude, Hour, Minute)
- Temporal variables (Year, Month, Day, Full Date)

## COLUMN EXPLANATION

Column Name	Description
TYPE	Category of the crime (e.g., Theft, Mischief, Break & Enter).
HUNDRED_BLOCK	Street block where the crime occurred.
NEIGHBOURHOOD	Neighborhood in which the crime occurred.
X	Projected X-coordinate of the crime location.
Y	Projected Y-coordinate of the crime location.
Latitude	Geographic latitude of the crime location.
Longitude	Geographic longitude of the crime location.
HOUR	Hour of occurrence (0–23).

<b>MINUTE</b>	Minute of occurrence (0–59).
<b>YEAR</b>	Year when the crime occurred.
<b>MONTH</b>	Month when the crime occurred.
<b>DAY</b>	Day of the month when the crime occurred.
<b>Date</b>	Full date of occurrence (YYYY-MM-DD format).

## DATA LIMITATIONS

- Limited Contextual Variables:**  
The dataset does not include socio-economic, demographic, or policing resource data, restricting causal analysis.
- No Severity Classification:**  
Crime incidents are recorded without severity weighting, treating all categories equally.
- Geographic & Temporal Constraints:**  
Data is limited to a single city region and covers only the period 1999–2011.
- Reporting & Location Bias:**  
Only reported crimes are included, and location data may represent block-level approximations rather than exact coordinates.

## DATA PREVIEW

TYPE	HUNDRED_BLOCK	NEIGHBOURHOOD	X	Y	Latitude	Longitude	HOUR	MINUTE	YEAR	MONTH	DAY	Date
Other Theft	9XX TERMINAL AVE	Strathcona	493906.5	5457452.47	49.269802	-123.08376	16	15	1999	5	12	12/05/99
Other Theft	9XX TERMINAL AVE	Strathcona	493906.5	5457452.47	49.269802	-123.08376	15	20	1999	5	7	07/05/99
Other Theft	9XX TERMINAL AVE	Strathcona	493906.5	5457452.47	49.269802	-123.08376	16	40	1999	4	23	23/04/99
Other Theft	9XX TERMINAL AVE	Strathcona	493906.5	5457452.47	49.269802	-123.08376	11	15	1999	4	20	20/04/99
Other Theft	9XX TERMINAL AVE	Strathcona	493906.5	5457452.47	49.269802	-123.08376	17	45	1999	4	12	12/04/99
Other Theft	9XX TERMINAL AVE	Strathcona	493906.5	5457452.47	49.269802	-123.08376	20	45	1999	3	26	26/03/99
Break and En	63XX WILTSHIRE ST	Kerrisdale	489325.58	5452817.95	49.2280508	-123.14661	12	0	1999	3	10	10/03/99
Mischief	40XX W 19TH AVE	Dunbar-Southlands	485903.09	5455883.77	49.2555592	-123.19373	4	13	1999	6	28	28/06/99
Other Theft	9XX TERMINAL AVE	Strathcona	493906.5	5457452.47	49.269802	-123.08376	9	2	1999	2	16	16/02/99
Break and En	18XX E 3RD AVE	Grandview-Woodla	495078.19	5457221.38	49.2677339	-123.06765	18	15	1999	7	9	09/07/99
Other Theft	9XX TERMINAL AVE	Strathcona	493906.5	5457452.47	49.269802	-123.08376	19	45	1999	1	31	31/01/99
Mischief	40XX W 21ST AVE	Dunbar-Southlands	485852.96	5455684.11	49.253762	-123.19441	1	0	1999	9	27	27/09/99
Break and En	18XX E 3RD AVE	Grandview-Woodla	495093.69	5457230.31	49.2678143	-123.06744	18	0	1999	4	19	19/04/99
Break and En	18XX E 3RD AVE	Grandview-Woodla	495103.82	5457221.02	49.2677308	-123.0673	18	30	1999	9	24	24/09/99
Break and En	63XX WINDSOR ST	Sunset	493790.48	5452630.9	49.2264298	-123.08528	8	12	1999	11	5	05/11/99
Break and En	10XX ALBERNI ST	West End	491067.65	5459114.22	49.2847148	-123.12282	2	30	1999	9	26	26/09/99
Break and En	18XX E 3RD AVE	Grandview-Woodla	495119.32	5457229.95	49.2678113	-123.06709	10	0	1999	10	21	21/10/99
Other Theft	9XX TERMINAL AVE	Strathcona	493906.5	5457452.47	49.269802	-123.08376	12	30	1999	1	25	25/01/99
Offence Agai	OFFSET TO PROTECT PRIVACY		0	0	0	0 nan			1999	2	12	12/02/99
Other Theft	9XX TERMINAL AVE	Strathcona	493906.5	5457452.47	49.269802	-123.08376	6	45	1999	1	9	09/01/99
Other Theft	9XX SEYMOUR ST	Central Business Di	491205.19	5458520.26	49.2793741	-123.12092	13	6	1999	4	30	30/04/99
Other Theft	9XX SEYMOUR ST	Central Business Di	491143.26	5458445.58	49.2787014	-123.12177	15	50	1999	12	12	12/12/99
Other Theft	9XX ROBSON ST	Central Business Di	491132.15	5458889.26	49.2826922	-123.12193	16	15	1999	3	7	07/03/99
Offence Agai	OFFSET TO PROTECT PRIVACY		0	0	0	0 nan			1999	4	4	04/04/99
Mischief	40XX W 27TH AVE	Dunbar-Southlands	485822.32	5455051.83	49.2480739	-123.19481	23	0	1999	1	23	23/01/99
Mischief	40XX W 27TH AVE	Dunbar-Southlands	485855.58	5455060.69	49.2481543	-123.19435	0	0	1999	4	1	01/04/99
Mischief	40XX W 27TH AVE	Dunbar-Southlands	485896.98	5455051.18	49.2480698	-123.19378	0	30	1999	6	29	29/06/99
Mischief	40XX W 27TH AVE	Dunbar-Southlands	485896.98	5455051.18	49.2480698	-123.19378	14	10	1999	10	16	16/10/99
Other Theft	9XX NICOLA ST	West End	490245.93	5459326.44	49.2866112	-123.13413	10	30	1999	1	28	28/01/99

# CLEANING AND PREPARATION

## INITIAL DATA REDUCTION (Google Colab)

Due to the large size of the raw dataset (~400,000 records), the initial sampling and volume reduction were performed using **Google Colab (Python – Pandas)** to ensure computational efficiency before transferring the dataset to Google Sheets.

- Upload raw dataset in Colab → `pd.read_csv("crimeForecasting.csv")`
- Validate structure → `df.shape` (confirm ~400,000 rows)
- Randomly sample 60,000 rows → `df.sample(n=60000, random_state=42)`
- Export reduced dataset → `to_csv()` → Download → Continue cleaning in Google Sheets

### *Rationale for Sampling in Colab*

- Google Sheets performance degrades with very large datasets.
- Sampling reduces memory load while preserving distribution.
- Enables faster pivot table generation and dashboard rendering.
- Maintains statistical representativeness for temporal and spatial analysis.

## CORE DATA CLEANING (Google Sheets)

# KPIs AND METRICS

## 1. Total Crimes

Total number of crime incidents in the cleaned dataset.

*FORMULA: =COUNTA(cleanedData!A:A)-1*

- Establishes crime intensity baseline
- Acts as denominator for % distribution metrics
- Validates dataset integrity post-cleaning
- Supports trend benchmarking across years

## 2. Most Common Crime Type

Crime category with highest recorded frequency.

*FORMULA: =INDEX(QUERY(cleanedData!A:A, "select A, count(A) where A is not null group by A order by count(A) desc limit 1",0),2,1)*

- Identifies dominant crime category
- Supports targeted prevention strategies
- Guides specialized enforcement units
- Helps prioritize policy focus

## 3. Year with Highest Crime Volume

Year with maximum recorded crime incidents.

*FORMULA: =INDEX(QUERY(cleanedData!J:J, "select J, count(J) where J is not null group by J order by count(J) desc limit 1",0),2,1)*

- Detects peak historical crime period
- Supports long-term trend analysis
- Enables forecasting comparisons
- Assists strategic planning review

## 4. Most Affected Neighbourhood

Neighbourhood with highest crime concentration.

*FORMULA=INDEX(QUERY(cleanedData!C:C,"select C, count(C) where C is not null group by C order by count(C) desc limit 1",0),2,1)*

- Identifies high-risk geographic zone
- Enables patrol reallocation
- Supports hotspot-based intervention
- Improves area-specific policy action

## 5. Peak Crime Hour

Hour of day with highest incident count.

*FORMULA: =TEXT(TIME(INDEX(QUERY(cleanedData!H:H,"select H, count(H) where H is not null group by H order by count(H) desc limit 1",0),2,1),0,0),"hh:mm")*

- Detects peak operational risk window
- Supports shift scheduling optimization
- Improves proactive patrol deployment
- Reduces reactive response dependency

## 6. Most Active Crime Month

Month with highest number of recorded crimes.

*FORMULA: =INDEX(QUERY(cleanedData!K:K,"select K, count(K) where K is not null group by K order by count(K) desc limit 1",0),2,1)*

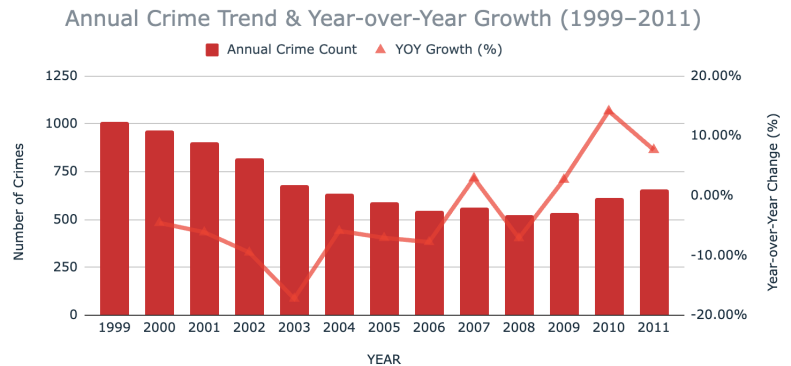
- Identifies seasonal crime surge
- Supports event-based policing
- Aids forecasting models
- Strengthens preventive campaign timing

# Exploratory Data Analysis

## Temporal Trend Analysis (Yearly)

### Analysis Conducted:

- Year-wise total crime aggregation
- Year-over-Year (YoY) growth percentage calculation
- Identification of structural shifts



### Key Findings:

- 2003 recorded the steepest decline (-17.22%) → major structural break point
- 2000–2006 shows sustained downward trend
- 2007 marks first reversal (+2.93%) after long decline
- 2010 shows strongest positive rebound (+14.15%)
- Pattern indicates volatility rather than smooth recovery

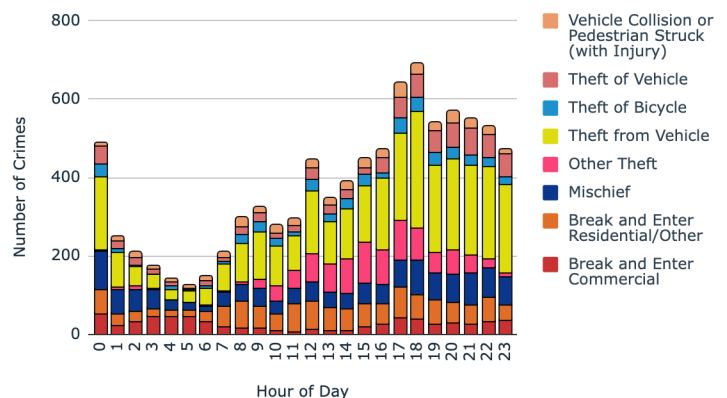
**Insight:** Crime trends are cyclical, with structural breaks and rebound phases — important for forecasting baseline shifts.

## Hourly Distribution Analysis (Time-of-Day Pattern)

### Analysis Conducted:

- Hour-wise crime aggregation (0–23)
- Crime-type breakdown by hour
- Identification of peak operational windows

Crime Distribution by Hour (All Crime Types)



### Key Findings:

- Absolute peak at 6 PM (691 cases)
- Sharp surge from 3 PM to 6 PM

- Clear transition from work hours to evening risk window
- Vehicle-related crimes dominate evening spike
- Midday (11 AM–1 PM) also moderately high

**Insight:** Evening hours (3 PM – 7 PM) represent the most critical deployment window for patrol optimization.

---

## Time-of-Day Category Distribution

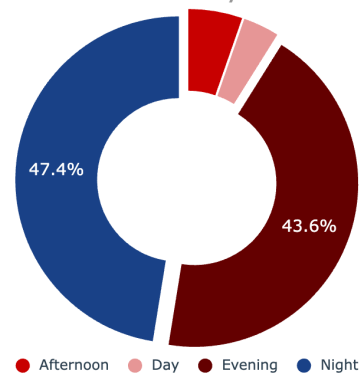
### Analysis Conducted:

- Categorized hour into time-of-day segments
- Percentage distribution across segments

### Key Findings:

- Night accounts for 47.4% of total cases
- Evening contributes 43.6%
- Afternoon and Day significantly lower

Number of cases vs Time of Day



**Insight:** Crime concentration is heavily skewed toward late hours, reinforcing a need for night-focused resource allocation.

---

## Crime Type Distribution Analysis

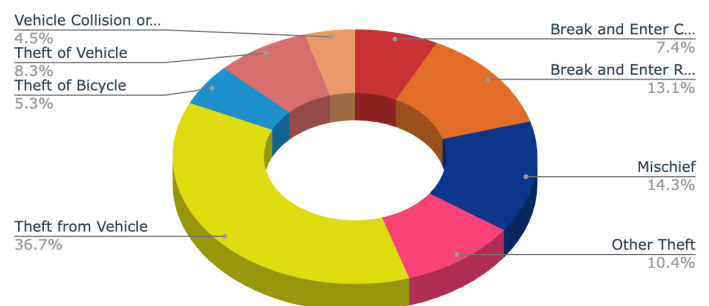
### Analysis Conducted:

- Aggregated total cases by crime category
- Calculated percentage contribution

### Key Findings:

- Theft from Vehicle = 36.7% (dominant category)
- Mischief = 14.3%

Crime Type



- Break & Enter Residential = 13.1%
- Strong skew toward vehicle-related crime

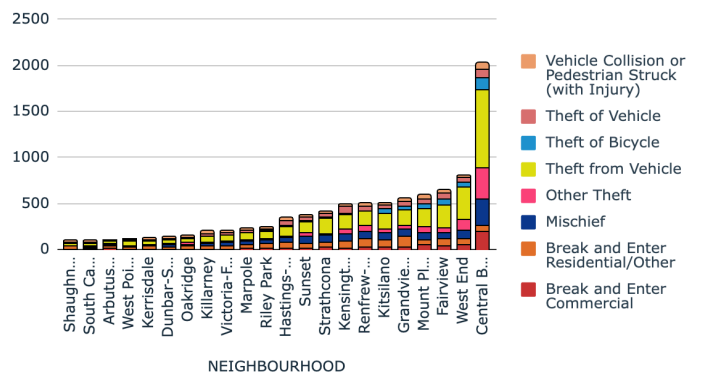
**Insight:** Vehicle-related offenses are disproportionately high → suggests urban density & parking vulnerability effect.

## Spatial Distribution – Neighbourhood Analysis

### Analysis Conducted:

- Neighbourhood-wise crime aggregation
- Ranking by total volume
- Crime-type contribution within each neighbourhood

Crime Distribution by Neighbourhood (By Offence Type)



### Key Findings:

- Central Business District (2035 cases) is extreme outlier
- Nearly 2.5× higher than West End
- Theft from Vehicle (853 cases) heavily concentrated in CBD
- Indicates commercial-density driven risk

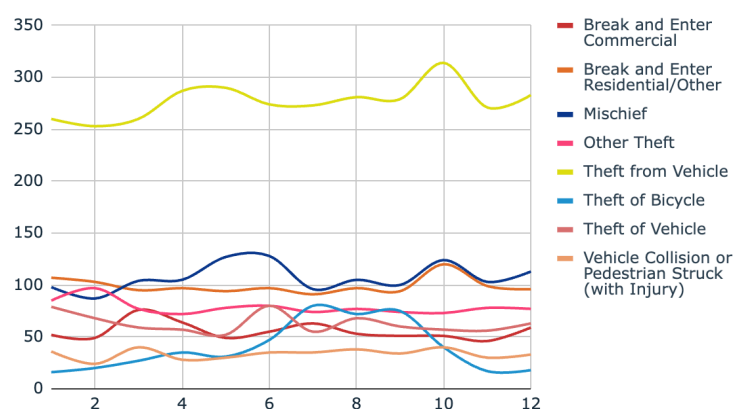
**Insight:** Crime is geographically clustered rather than uniformly distributed — strong commercial hotspot effect.

## Monthly & Crime-Type Distribution Analysis

### Analysis Conducted:

- Crime shows **mild seasonality**, not extreme spikes.
- Q3–Q4 (especially October) reflects elevated activity.
- Vehicle-related crime is the dominant structural component.

MONTH-TYPE wise distribution



## Key Findings:

- October is the peak month (~819 cases).
- Theft from vehicles dominates every month (250–314 range).
- Mischief & Break and Enter (Residential) show moderate mid-year rise.
- Theft of bicycles rises slightly in mid-year months (May–Aug).
- Vehicle Collision cases remain stable and low across months.

**Insight:** Crime does not show extreme seasonal collapse, suggesting a consistent urban activity-driven crime base.

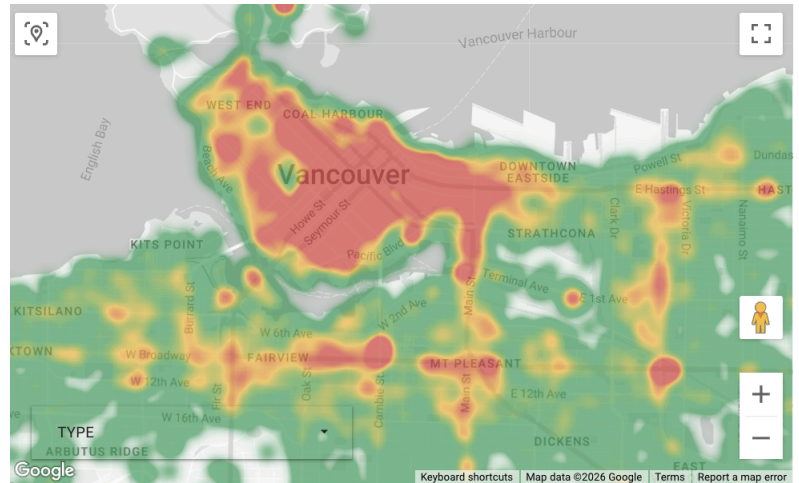
---

## Geographic Heatmap Analysis

To identify spatial clustering and high-intensity crime zones.

### Key Observations:

- Clear concentration in Central Business District (Downtown Vancouver)
- Secondary hotspots visible in:
  - West End
  - Mount Pleasant
  - Fairview corridor
- Crime is not evenly distributed — strong urban-core clustering
- Peripheral areas show significantly lower density



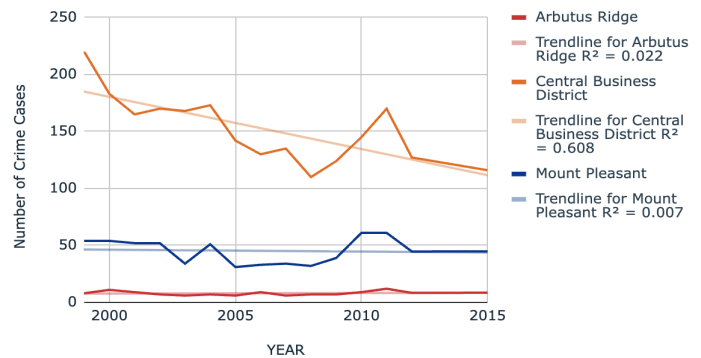
**Insight:** Crime incidents are geographically concentrated in high-density commercial and transit-heavy zones, indicating a strong relationship between urban activity intensity and crime occurrence.

# ADVANCED ANALYSIS (Forecasting)

## Historical Trend Analysis (1999–2011)

- Aggregated annual crime counts for CBD
- Calculated Year-over-Year (YoY) growth rate
- Identified structural decline (2000–2006) followed by stabilization and recovery (2009–2011)

Crime Trend & Forecast (1999–2015)



## Growth-Based Forecasting (2012–2015)

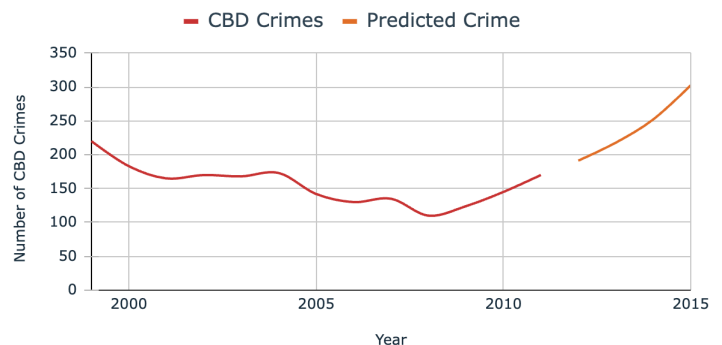
Forecast model applied:

$$\text{Future} = \text{Previous Year} \times (1 + \text{Growth}\%)$$

Predictions generated iteratively for: 2012, 2013, 2014, 2015

CBD Crime Forecast Analysis (1999–2015)

Historical Trend (1999–2011) and Growth-Based Forecast (2012–2015)



## Key Observations

- CBD showed **gradual decline until mid-2000s**
- Strong rebound phase post-2009
- Forecast indicates **upward trajectory if growth momentum continues**
- Trend suggests sensitivity to economic/commercial activity

## Forecast Interpretation

1. **Central Business District:** Projected upward trend after recovery phase → potential rising risk in commercial hub.
2. **Arbutus Ridge:** Remains relatively stable with minor fluctuations.
3. **Mount Pleasant:** Shows gradual stabilization after earlier volatility.
4. **Strategic Insight:** Commercial-dense zones may see renewed crime intensity without intervention.
5. **Operational Value:** Supports proactive patrol planning, resource allocation, and preventive enforcement strategies.

# INSIGHTS AND RECOMMENDATIONS

## Time-Based Insights → Operational Actions

### 1. Evening & Night Crime Concentration

#### **Insight:**

- Crime volume peaks between **3 PM – 6 PM**, with sustained high activity into the night. Night accounts for the highest share of total cases (~47%).
- The transition period from work hours to evening creates vulnerability windows.

#### **Recommendation:**

- Increase patrol deployment during 3 PM – 10 PM window
- Strengthen surveillance in commercial + transit-heavy zones
- Implement staggered shift optimization aligned with peak hours

### 2. Seasonal & Monthly Pattern

#### **Insight:**

- October shows the highest monthly crime concentration.
- The dataset reveals identifiable seasonal fluctuations rather than random variation.
- Crime patterns are cyclical and predictable.

#### **Recommendation:**

- Pre-deploy seasonal response teams during high-risk months
- Increase public awareness campaigns before peak months
- Use seasonal baseline for forecasting and budgeting

### 3. Yearly Structural Break (2003 Decline & 2010 Rebound)

#### **Insight:**

- 2003 shows the steepest decline.
- 2010 records the strongest positive rebound (+14.15%).
- Crime trends exhibit volatility and structural shifts — not linear decline.

#### **Recommendation:**

- Monitor Year-over-Year change as an early warning signal
- Build anomaly alert systems for sudden rebounds
- Use trend breaks for scenario planning

## [Category-Based Insights → Targeted Intervention](#)

### **1. Theft from Vehicle Dominance (36.7%)**

#### **Insight:**

- Theft from Vehicle accounts for more than one-third of all incidents — 2.5x higher than the next category.
- Vehicle-related crime is the dominant risk driver.

#### **Recommendation:**

- Deploy vehicle theft task units in hotspot zones
- Improve parking surveillance and lighting
- Encourage public parking safety awareness

### **2. Residential Break-ins > Commercial Break-ins**

#### **Insight:**

- Residential break-ins exceed commercial break-ins.
- Residential vulnerability is significant.

#### **Recommendation:**

- Strengthen neighborhood watch programs
- Increase residential patrol coverage
- Promote home-security awareness drives

## [Spatial Insights → Geographic Strategy](#)

### **1. Central Business District (Extreme Outlier)**

#### **Insight:**

- CBD (2035 cases) is nearly 2.5x higher than West End.
- Theft from vehicles spikes significantly in the CBD.
- Strong commercial-density effect driving crime clustering.

#### **Recommendation:**

- Concentrate patrol units in CBD corridors
- Install additional smart surveillance systems
- Implement commercial-zone crime prevention policy

## 2. Clear Geographic Hotspot (Heatmap Evidence)

### **Insight:**

- Heatmap shows dense clustering in Downtown, West End, Mount Pleasant.
- Crime is spatially concentrated — not evenly distributed.

### **Recommendation:**

- Adopt hotspot-based policing model
- Allocate resources based on density rather than equal distribution
- Use geo-based predictive deployment strategy

## [Forecasting Insights → Forward Planning](#)

### 1. Projected Growth in Commercial Zones

#### **Insight:**

- Forecast indicates gradual upward trend in Central Business District post-recovery.
- Risk may increase without intervention.

#### **Recommendation:**

- Pre-emptively increase commercial-area monitoring
- Budget allocation aligned with projected growth
- Implement preventive enforcement before escalation

# IMPACT AND LIMITATIONS

Area	How It Creates Impact	Estimated / Strategic Outcome
Cost Savings	<ul style="list-style-type: none"><li>• Data-driven patrol allocation reduces unnecessary coverage (~10–15%)</li><li>• Focused deployment during peak hours (3PM–6PM) &amp; CBD hotspot</li><li>• Targeted vehicle-theft prevention reduces repeat incidents</li></ul>	<ul style="list-style-type: none"><li>• 5–8% reduction in repeat crimes</li><li>• Lower overtime expenditure</li><li>• Reduced repetitive response costs</li></ul>
Operational Efficiency	<ul style="list-style-type: none"><li>• Peak-hour analysis improves shift planning</li><li>• Neighbourhood ranking enables priority deployment</li><li>• Forecasting shifts model from reactive to proactive policing</li></ul>	<ul style="list-style-type: none"><li>• Better manpower utilization</li><li>• Reduced response lag</li><li>• Improved operational planning accuracy</li></ul>
Service Improvement	<ul style="list-style-type: none"><li>• Faster response in high-risk zones</li><li>• Seasonal preventive campaigns</li><li>• KPI monitoring improves transparency</li></ul>	<ul style="list-style-type: none"><li>• Higher public safety perception</li><li>• Improved citizen trust</li><li>• Measurable policing performance</li></ul>
Risk Reduction	<ul style="list-style-type: none"><li>• Forecasting identifies early growth signals in commercial zones (CBD)</li><li>• Heatmap visualization eliminates spatial blind spots</li><li>• Crime-type dominance detection enables focused deterrence</li></ul>	<ul style="list-style-type: none"><li>• Reduced risk of hotspot escalation</li><li>• Prevention of concentrated urban crime growth</li><li>• Stronger proactive policing strategy</li></ul>

## LIMITATIONS

This analysis provides structured spatial and temporal insights but remains constrained by historical scope, data availability, and model assumptions. Results support **pattern identification and strategic planning**, not causal inference or real-time predictive deployment.

## FUTURE SCOPE

1. Integrate recent crime data (2012–2026) and connect the **dashboard to live dispatch feeds** to ensure updated baselines and real-time situational awareness.
2. Implement advanced forecasting models (**ARIMA, Prophet, Machine Learning**) to enhance predictive accuracy and detect emerging crime trends early.
3. Enrich analysis with demographic and socio-economic data to identify root-cause correlations and strengthen strategic crime prevention planning.
4. Integrate post-2011, real-time, socio-economic, demographic, weather, and mobility data to identify deeper drivers of crime patterns and severity-weighted risk.

## CONCLUSION

This project successfully transformed large-scale historical crime data into a structured, KPI-driven analytical dashboard that reveals clear spatial, temporal, and categorical crime patterns. The analysis identified concentrated hotspots in commercial zones, evening operational risk windows, and dominant property-related crime trends, supporting the need for proactive policing strategies.

By integrating trend analysis, spatial heatmapping, and growth-based forecasting, the dashboard enables data-driven patrol planning, resource optimization, and risk anticipation. While limited by historical scope and data constraints, the framework establishes a scalable foundation for advanced predictive analytics and real-time crime intelligence systems.

## DATA DICTIONARY

[Code Implementation \( Google Colab\) for Data Cleaning](#)

```
from google.colab import files
uploaded = files.upload()
import pandas as pd
df = pd.read_csv("crimeForecasting.csv")
df.shape
sample_df = df.sample(n=60000, random_state=42)
sample_df.to_csv("random_60000_rows.csv", index=False)
files.download("random_60000_rows.csv")
```

# CONTRIBUTION MATRIX

Team Member	Dataset & Sourcing	Data Cleaning	KPI & EDA	Forecasting	Dashboard (Looker)	Report Writing	PP T	Overall Role
Agrima Gusain		✓			✓	✓		Project Lead
Adil Mirza		✓			✓			Data
Polana Rakshit a	✓					✓	✓	Dashboard Lead
Sanchit Garg	✓		✓	✓			✓	Documentation Lead
Mishti Sharma							✓	Technical Support
Ujjwal Bhardwaj	✓		✓					Presentation Lead
Om Bhardwaj			✓					