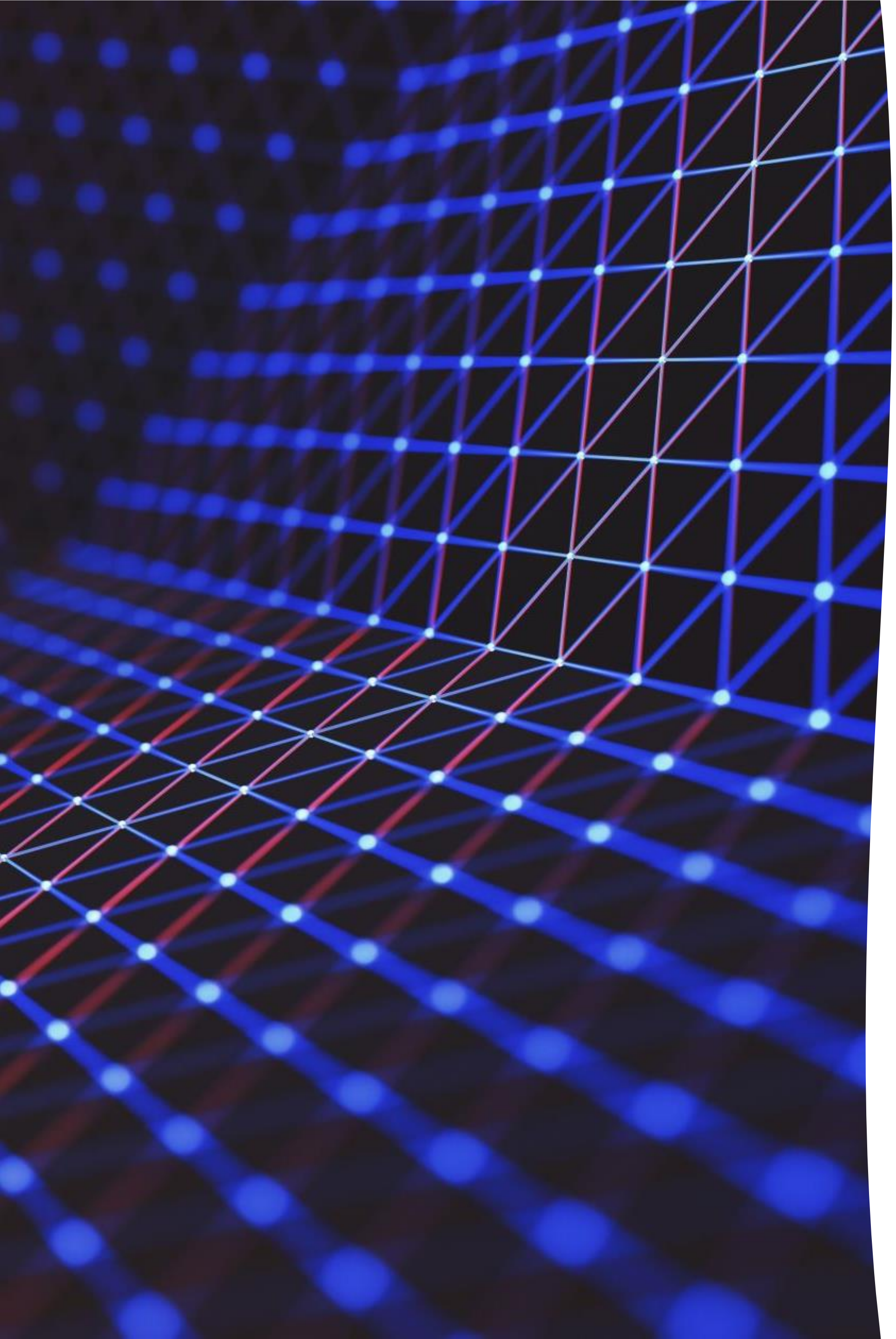


Optimizing Urban Safety in Fredericton

Advanced Data Science Capstone Project – Week 5

Utilizing data-driven insights to improve urban planning and safety strategies.

Executive Summary

Main Objective: To analyze and optimize urban safety in Fredericton using data science tools.

Key Tools: SQL for queries, Folium for geospatial visualization, and Plotly Dash for interactive analysis.

Highlighted Results: Identification of high- and low-crime zones, patterns in crime types, and strategic recommendations for safety improvements.

Additional Suggestions:

Include a brief visualization, such as a summary chart or a representative image.

State an expected impact: "These strategies could potentially reduce crime incidents by 15%."

Introduction

- This project explores how data can optimize safety in urban neighborhoods.
- It analyzes 2017 crime data in Fredericton and leverages advanced tools to uncover actionable insights.
- Key questions addressed: Where do most crimes occur? What are the safest areas? How do crime patterns vary geographically?
- **Additional Suggestions:**
  - Include an image or map of Fredericton.
  - Add a real or hypothetical example of how data impacts safety decisions.

Objectives

1

1. Identify neighbourhoods in Fredericton with the most and least reported crimes.

2

2. Analyze patterns of crime types across various regions.

3

3. Pinpoint geographic hotspots and safe zones using advanced geospatial tools.

4

4. Leverage location-based data for actionable insights.

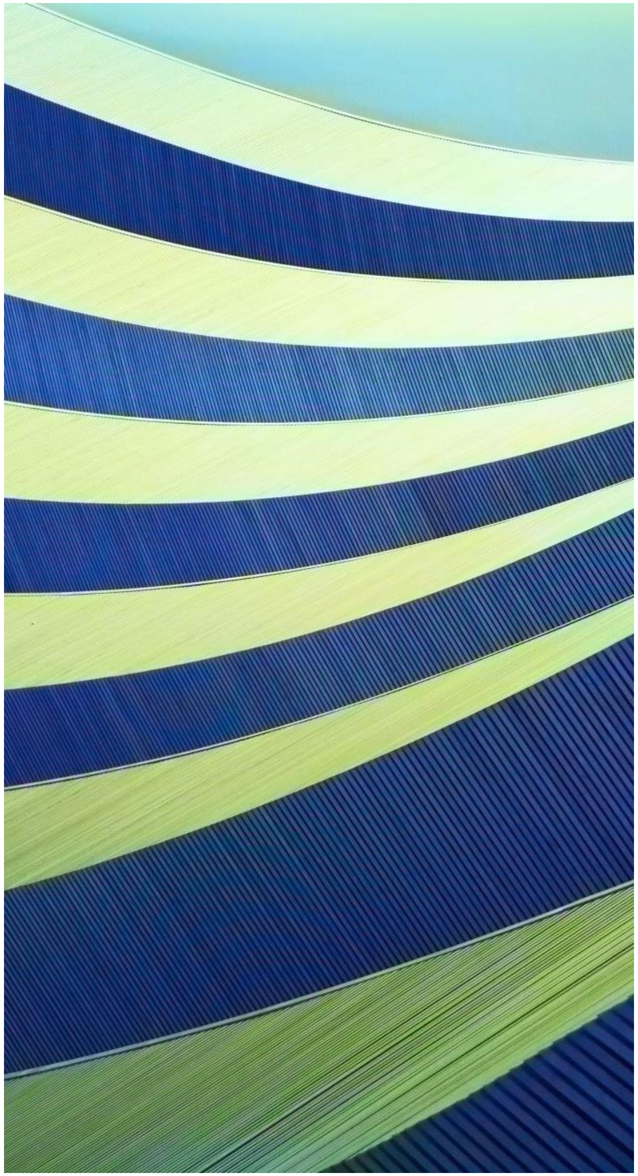
Data Sources

1. Fredericton City Open Data (Crime Statistics 2017).

2. Demographic and geographic data from public repositories.

3. Interactive data from Foursquare API.

4. Supplementary details from Fredericton City Wikipedia page.



Data Collection and Wrangling Methodology

Data Sources:
- Fredericton Open Data: Crime statistics from 2017.
- Demographic and geographic datasets from public repositories.
- Location-based data via Foursquare API.

Data Cleaning Steps:
- Addressed missing data by filling or excluding certain fields.
- Removed outliers to ensure consistent analysis.
- Consolidated datasets for compatibility with clustering and predictive modeling.

Additional Suggestions:
- Include a flowchart showing the data wrangling process.
- Add a table comparing raw and cleaned data.

Raw Dataset

[12]:

```
#import pandas as pd
crime_df = pd.read_csv('Crime_by_neighbourhood_2017.csv')
crime_df.head(10)
```

[12]:

	Neighbourhood	From_Date	To_Date	Crime_Code	Crime_Type	Ward	City	FID
0	Fredericton South	2017-01-05T00:00:00.000Z	2017-01-26T00:00:00.000Z	2120	B&E NON-RESIDNCE	7	Fredericton	1
1	Fredericton South	2017-03-04T00:00:00.000Z	2017-03-06T00:00:00.000Z	2120	B&E NON-RESIDNCE	7	Fredericton	2
2	Fredericton South	2017-05-07T00:00:00.000Z	NaN	2120	B&E NON-RESIDNCE	12	Fredericton	3
3	Fredericton South	2017-06-20T00:00:00.000Z	2017-06-21T00:00:00.000Z	2120	B&E NON-RESIDNCE	12	Fredericton	4
4	Fredericton South	2017-07-09T00:00:00.000Z	2017-07-10T00:00:00.000Z	2120	B&E NON-RESIDNCE	7	Fredericton	5
5	Fredericton South	2017-01-06T00:00:00.000Z	2017-01-09T00:00:00.000Z	2170	MISCHIEF TO PROP	7	Fredericton	6
6	Fredericton South	2017-02-04T00:00:00.000Z	NaN	2170	MISCHIEF TO PROP	7	Fredericton	7
7	Fredericton South	2017-03-08T00:00:00.000Z	NaN	2170	MISCHIEF TO PROP	11	Fredericton	8
8	Fredericton South	2017-04-03T00:00:00.000Z	NaN	2170	MISCHIEF TO PROP	7	Fredericton	9
9	Fredericton South	2017-04-13T00:00:00.000Z	2017-04-13T00:00:00.000Z	2170	MISCHIEF TO PROP	11	Fredericton	10

Cleaned Dataset

```
crime_df.drop(['From_Date', 'To_Date'], axis=1,inplace=True)
crime_df.head(10)
```

	Neighbourhood	Crime_Code	Crime_Type	Ward	City	FID
0	Fredericton South	2120	B&E NON-RESIDNCE	7	Fredericton	1
1	Fredericton South	2120	B&E NON-RESIDNCE	7	Fredericton	2
2	Fredericton South	2120	B&E NON-RESIDNCE	12	Fredericton	3
3	Fredericton South	2120	B&E NON-RESIDNCE	12	Fredericton	4
4	Fredericton South	2120	B&E NON-RESIDNCE	7	Fredericton	5
5	Fredericton South	2170	MISCHIEF TO PROP	7	Fredericton	6

EDA and Interactive Visual Analytics Methodology

1. Data cleaning and preprocessing to ensure reliability.

2. Exploratory Data Analysis (EDA) to uncover key trends.

3. Clustering analysis using K-means for grouping neighbourhoods.

4. Visualization of spatial and statistical insights.

5. Development of interactive dashboards and predictive models.

Predictive Analysis Methodology

- Objective: To predict crime types based on neighborhood characteristics.
- Approach:
- Selected a classification model (e.g., Logistic Regression, Decision Tree).
- Performed feature selection to identify key drivers (e.g., population density, proximity to hotspots).
- Split data into training and testing sets for validation.
- **Additional Suggestions:**
- Include a flowchart summarizing the predictive modeling process.
- Add reasons for choosing the specific model and comparisons with alternatives.

Highest crime sectors

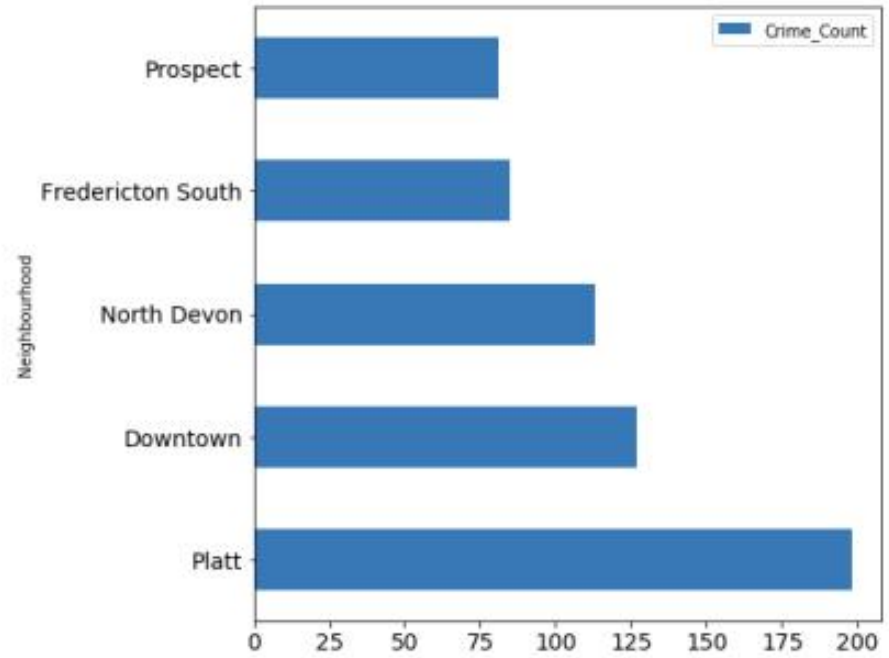
1. Platt= 198
2. Downtown = 127
3. North Devon = 113
4. Fredericton South = 85
5. Prospect = 81

Extract top 5 neighbourhoods with high crime count

```
[28]: top_crime_count = crime_data.nlargest(5, "Crime_Count")

      top_crime_count
```

[28]:

	Neighbourhood	Crime_Count
40	Platt	198
10	Downtown	127
38	North Devon	113
13	Fredericton South	85
42	Prospect	81



Highest crime Mos frecuent sectors crimesectors

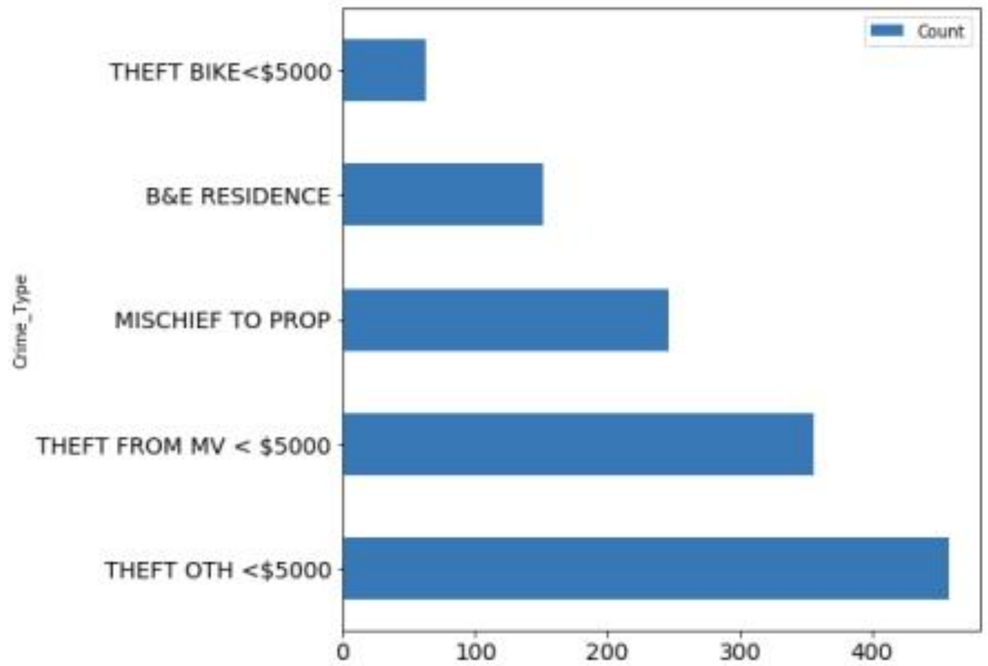
Extract top 3 crime types with highest frequency

[79]:

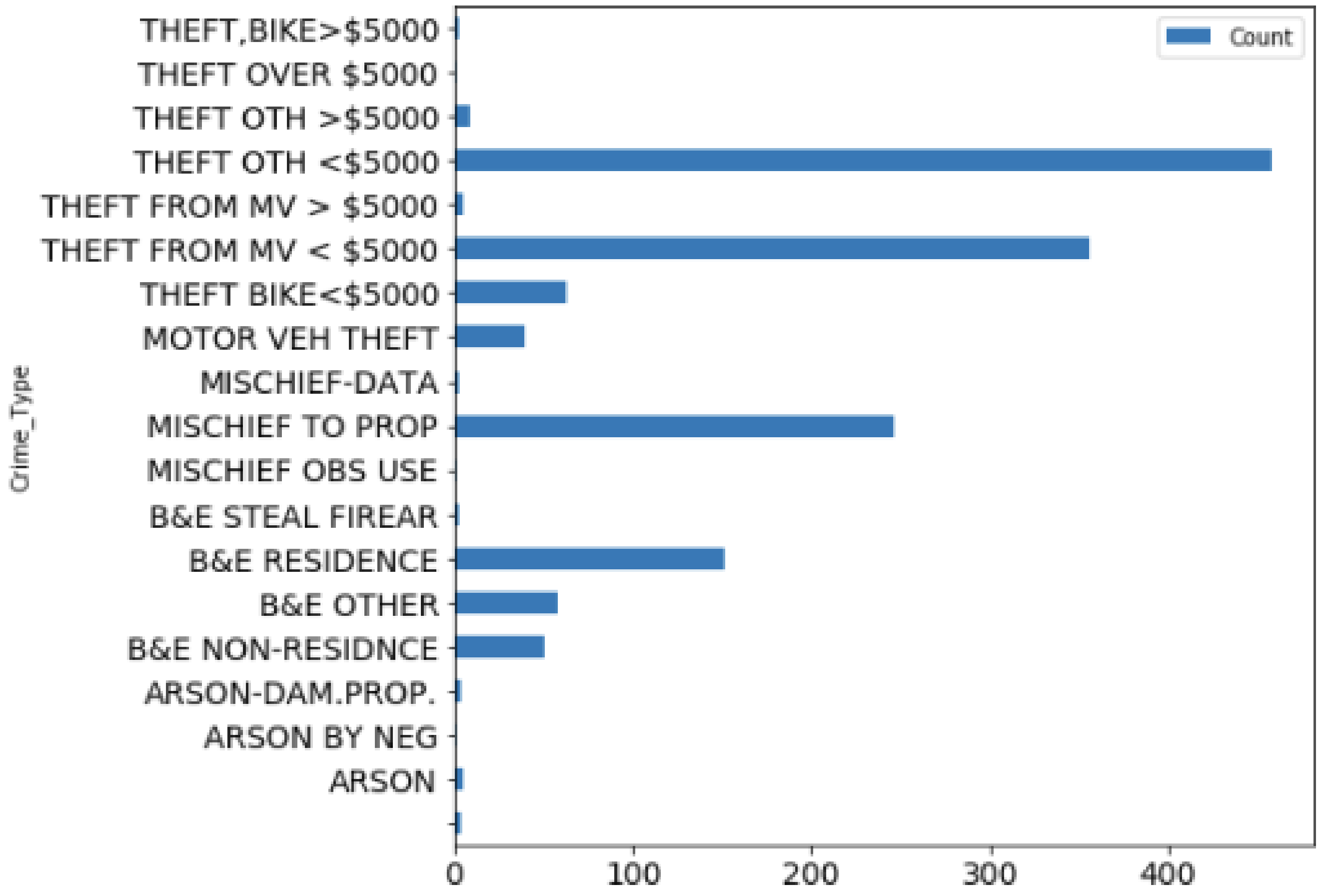
```
top_crime_type = crimetype_data.nlargest(5, "Count")

top_crime_type
```

[79]:

	Crime_Type	Count
15	THEFT OTH <$5000	458
13	THEFT FROM MV < $5000	356
9	MISCHIEF TO PROP	246
6	B&E RESIDENCE	151
12	THEFT BIKE<$5000	63



Proportion of different crime types that occurred in 2017



EDA with Visualization Results

Key Insights:

- Top neighborhoods with the highest crime counts.
- Patterns of frequent crime types and their geographic distribution.
- Statistical summaries, such as mean, median, and variance in crime data.

Additional Suggestions:

- Add well-labeled bar charts, histograms, or scatter plots.
- Highlight surprising insights (e.g., “Over 40% of incidents occur in just 3 neighborhoods”).

Specific locations in Fredericton & their geographical coordinat

EXAMINE SPECIFIC LOCATIONS IN FREDERICTION

[83]:
```
location_df = pd.read_csv('Fredericton_Locations.csv')
location_df
```

[83]:

	Location	Latitude	Longitude
0	Knowledge Park	45.931143	-66.652700
1	Fredericton Downtown	45.963026	-66.383550
2	Fredericton Hill	45.948512	-66.656045
3	Nashwaaksis	45.983382	-66.644856
4	University of New Brunswick	45.948121	-66.641406
5	Devon	45.968802	-66.622738
6	New Maryland	45.892795	-66.683673
7	Marysville	45.978913	-66.589491
8	Skyline Acres	45.931827	-66.640339
9	Hanwell	45.902315	-66.755113

Key Findings

- Neighbourhoods with the highest crime rates were analyzed and visualized.
- Certain regions exhibited unique patterns of crime distribution by type.
- Safe zones were identified, providing opportunities for policy reinforcement.

Closing Remarks

- The study demonstrates how data science can aid urban safety planning.
- Visualizations and clustering provide actionable insights for decision-makers.
- Future work: integrating real-time data for dynamic monitoring.