



CRIME DATA REPORT

Presentation for Data Scientist #23-18

JOHN Q LI
JOHNWILL4G@GMAIL.COM



CONTENT

01

About

02

Goal &
Objective

03

Datasets

04

Aproach

05

Obj #1
Analytics

06

Obj #2
Analytics

07

Optimization
Strategies

08

Q & A

ABOUT



Data Scientist #23-18 - Business Intelligence & Data Analytics
to analyze and forecast crime trends to facilitate the operation of the York Reginal Police Department

Individul Contributor:

- a decade of Automation and Cluster Support experience. Obtained
- Master's Degree in Data Science.

GOALS AND OBJECTIVES

Goal

Demo this presentation with advanced Data analytic skill sets and techniques

Obj #1

Understand the locations in York Region that are the highest risk for increases in criminal activity

Obj #2

Understand projected calls for service trends in their District and the impacts on front line staffing.



DATASETS

crime historical
data

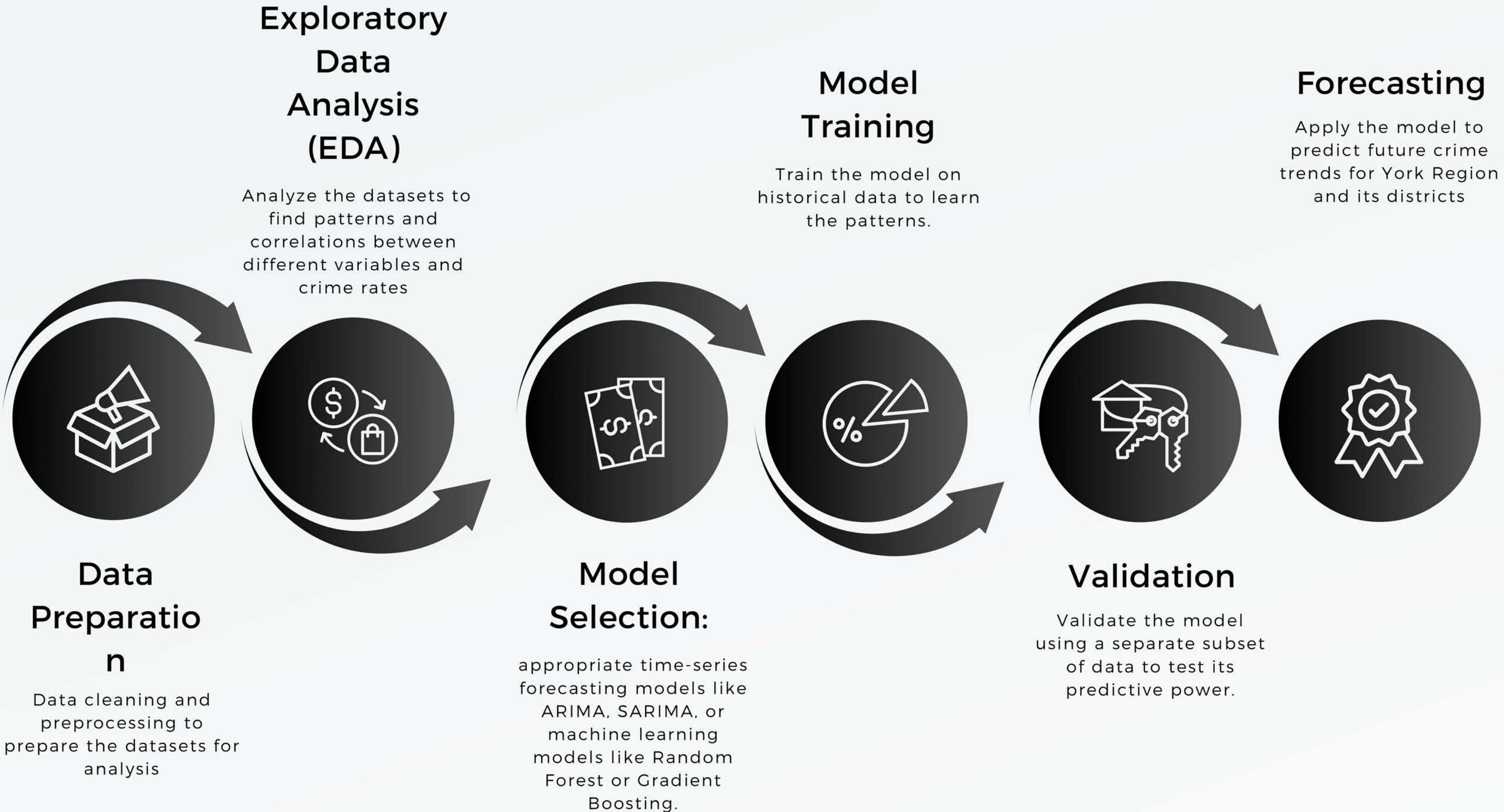


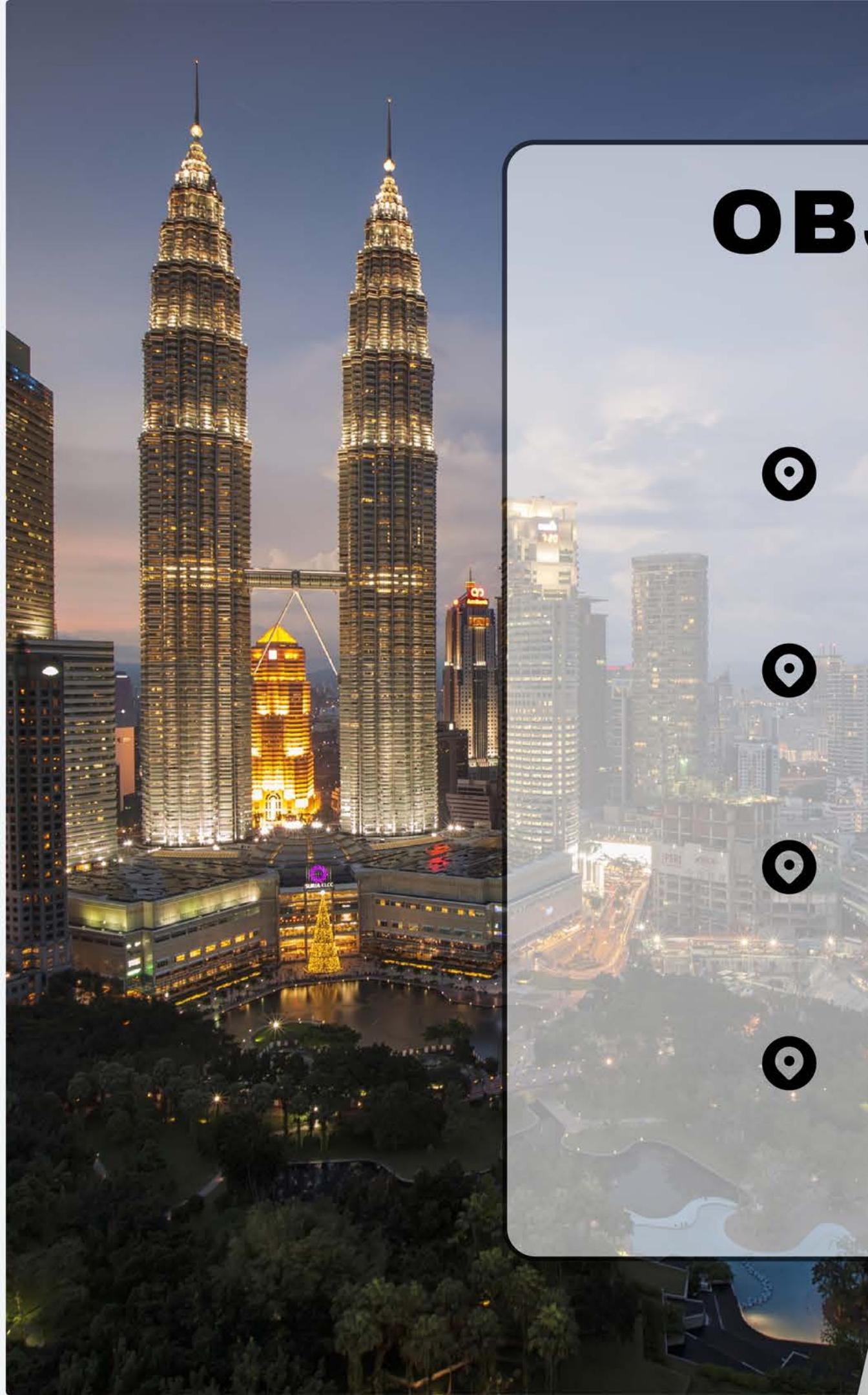
Calls for Service



District Platoon
Staffing
Exceptions







OBJ #1 - CRIME HISTORICAL FORECAST

- 📍 A crime forecast for York Region and at one of its Police Districts
- 📍 Identify the municipality in York Region at the highest risk for increases in criminal activity or public safety concerns and why
- 📍 EDA to identify insights
- 📍 Integrate a publicly available dataset not provided to you into your crime forecast model. Be creative and think of external factors that could impact crime

CRIME HISTORICAL DATA

01

Planning

Analysis of historical crime data from the dataset "Year_to_Date_Community_Safety_Data_.csv."

02

Analysis

Identification of the district with the highest risk of crime increase, supported by data insights.

03

Optimization

optimize with analysis results

04

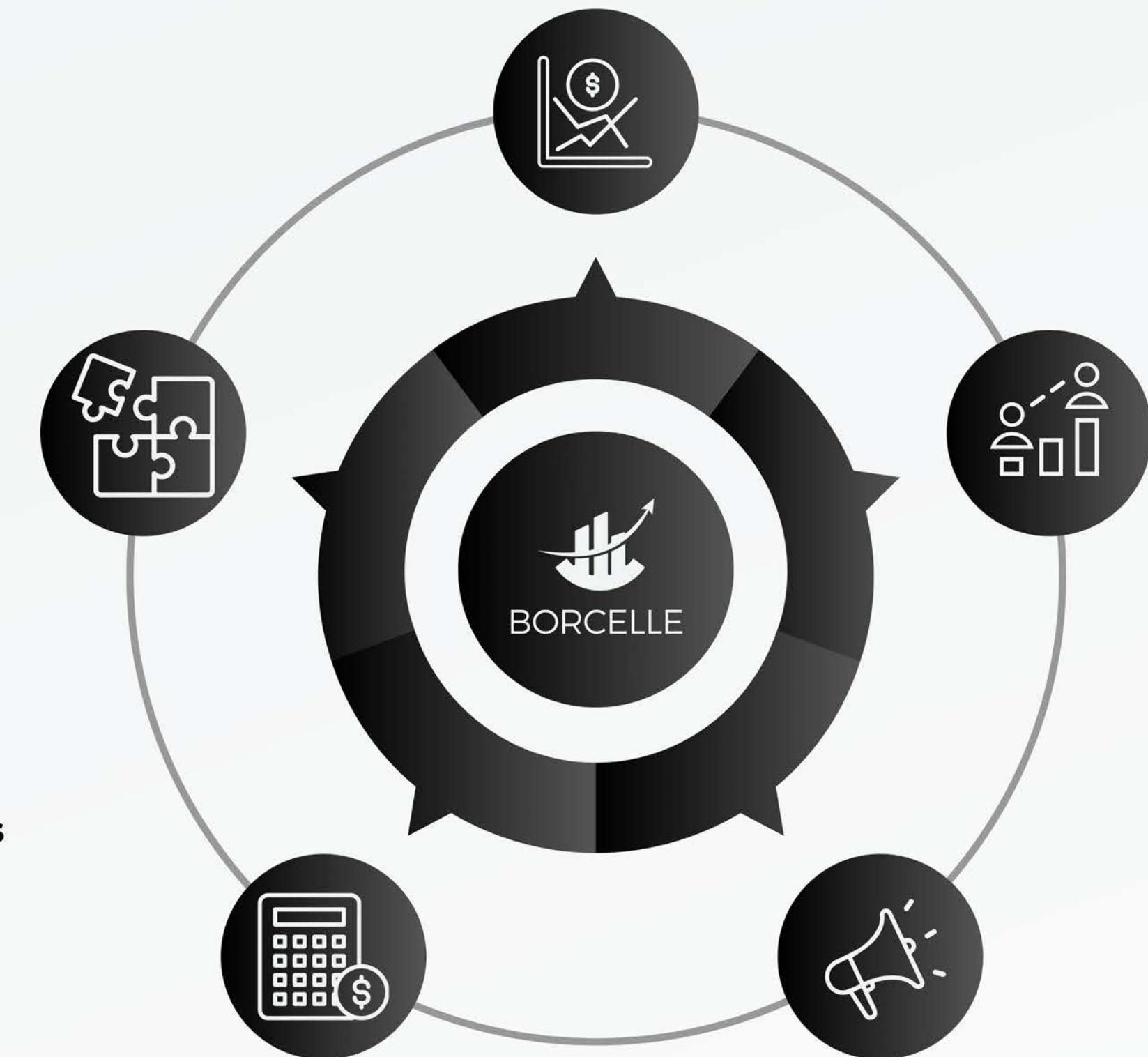
Insights

conclude findings as insights

05

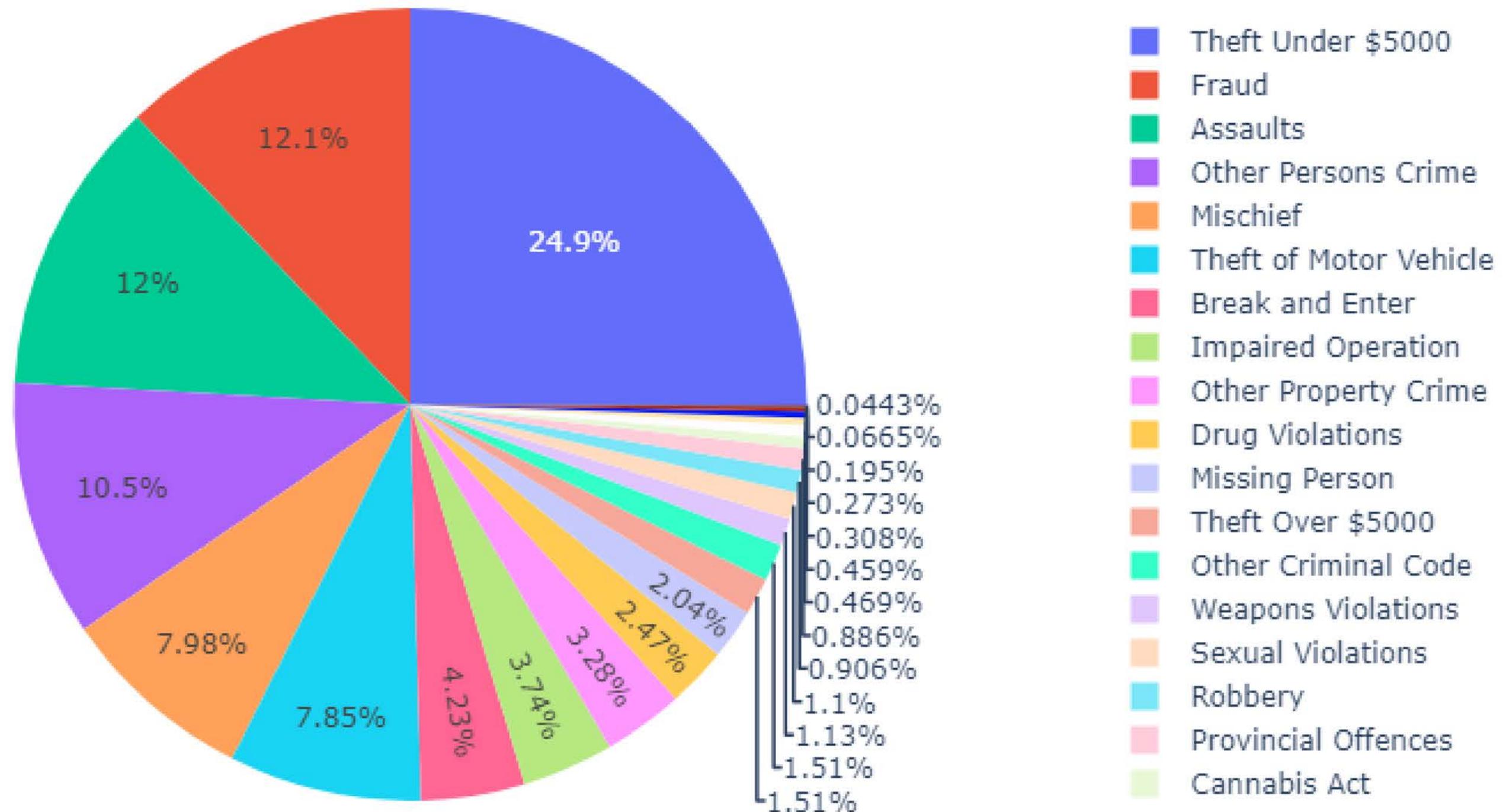
Additional datasets

Integrate with additional datasets for more advanced analysis

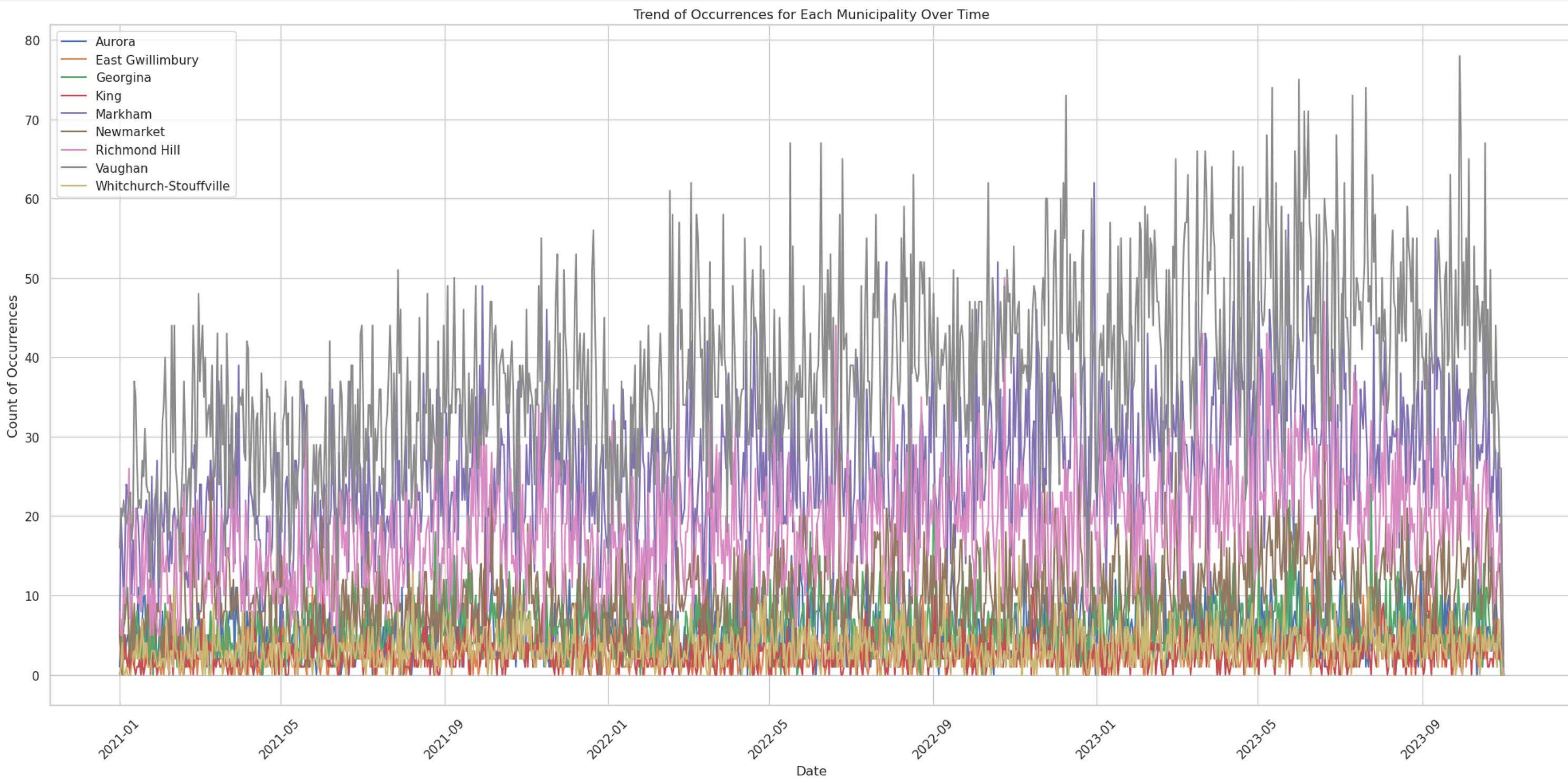


Distribution of Occurrence Types

CRIME TYPE

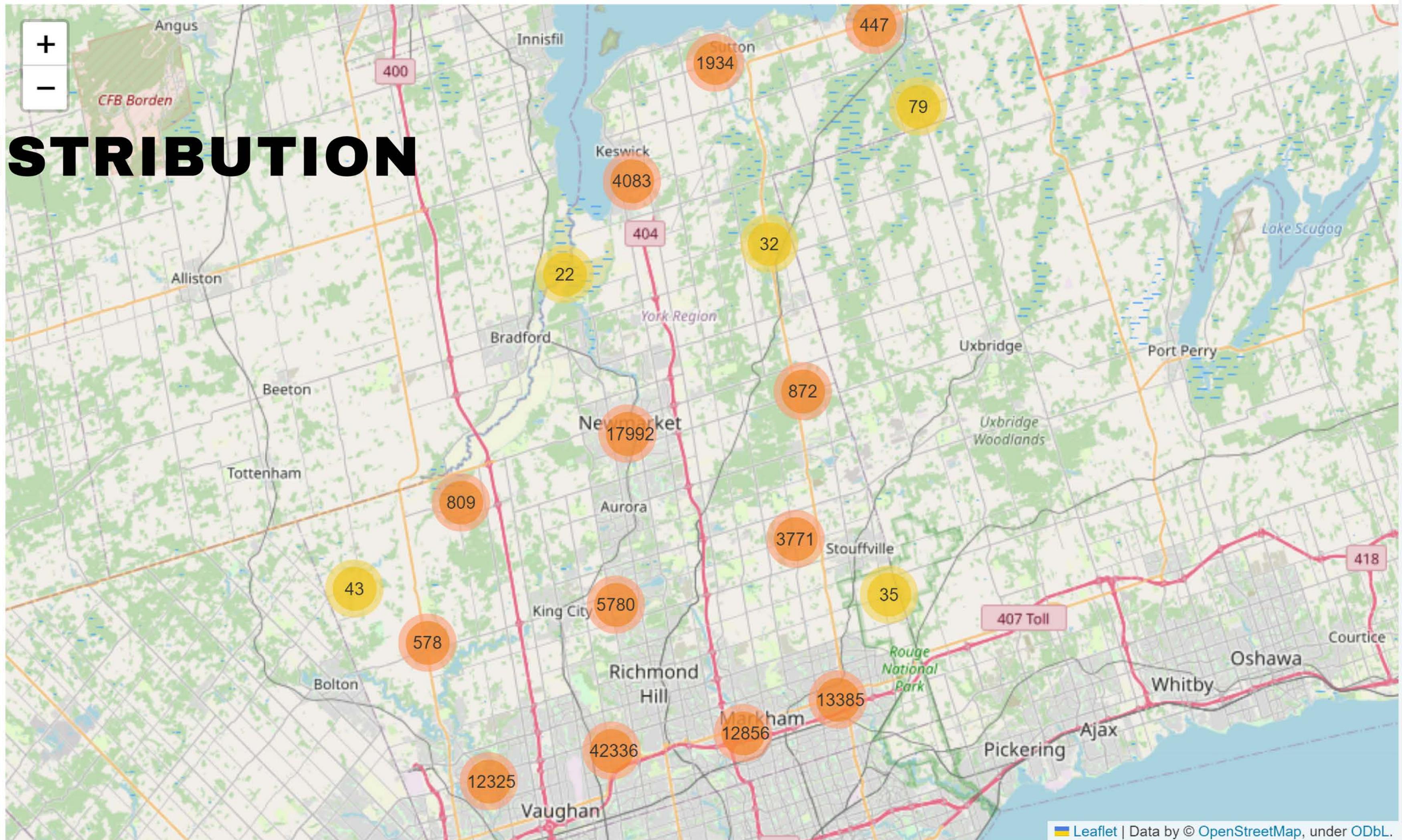


MUNICIPALITIES COMPAIRSON

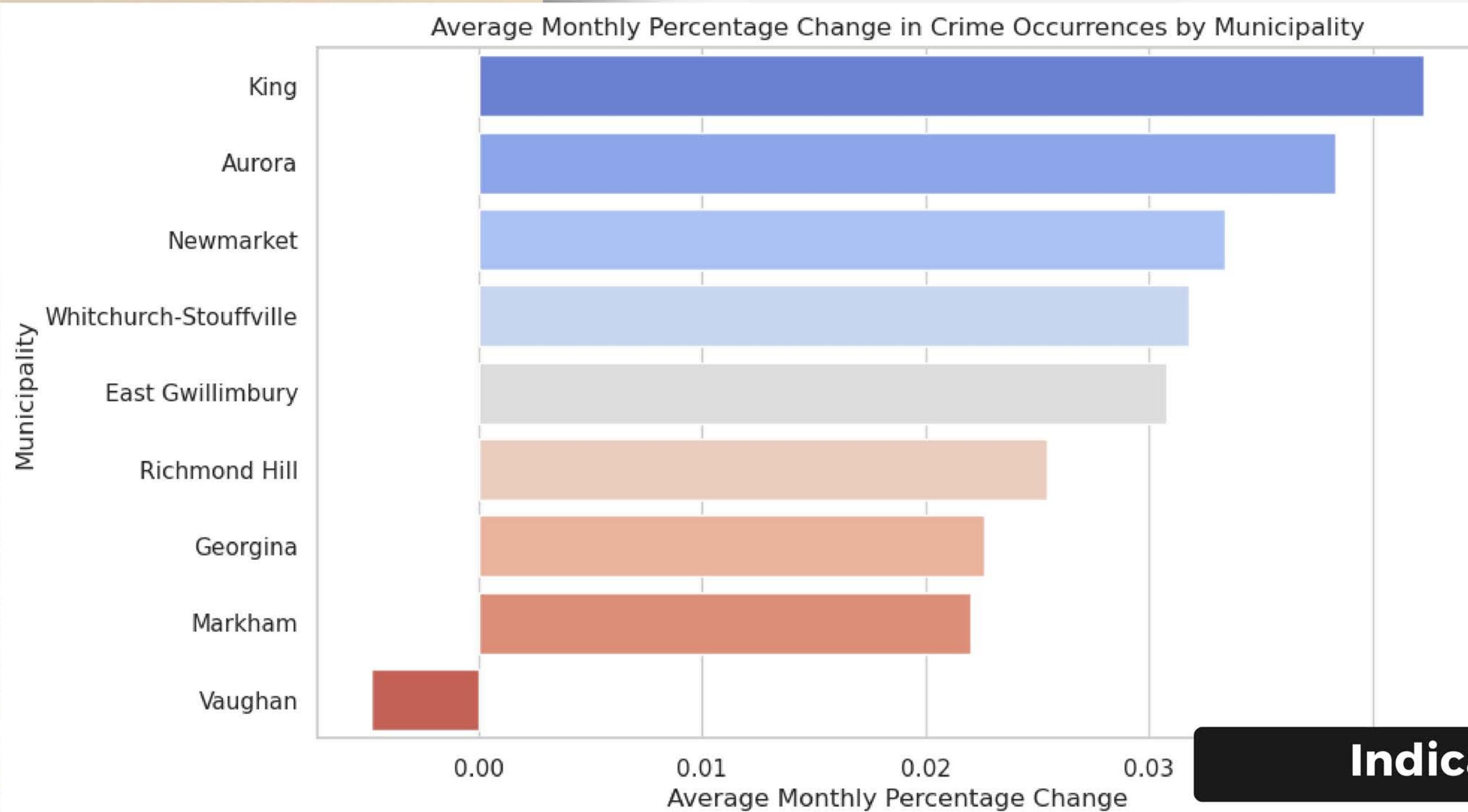




DISTRIBUTION



INSIGHTS



Indicator

average monthly percentage increase

Key findings

- Aurora: 3.83% average monthly change
- Markham: 2.20% average monthly change
- Richmond Hill: 2.55% average monthly change
- King: 4.23% average monthly change
- Vaughan: -0.48% average monthly change

ADDITIONAL PUBLIC DATASETS

1. Markham

- **Average Household Income:**
\$104,000
- **Number of Crime Occurrences:**
26,749
- **Population:** 338,503

2. King

- **Average Household Income:**
\$141,000
- **Number of Crime Occurrences:**
2,802
- **Population:** 27,333

3. Georgina

- **Average Household Income:**
\$98,000
- **Number of Crime Occurrences:**
6,288
- **Population:** 47,642

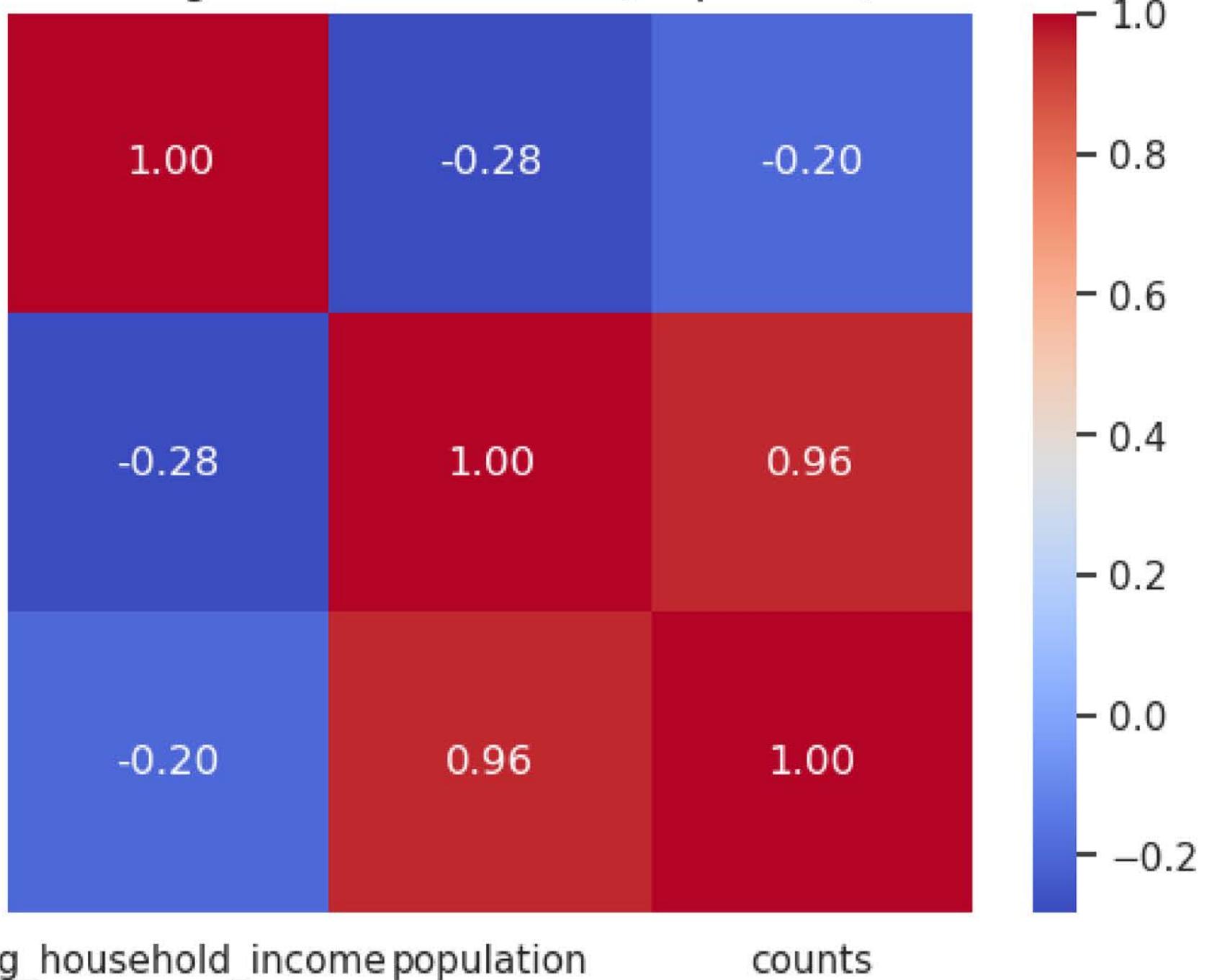
4. Aurora

- **Average Household Income:**
\$119,000
- **Number of Crime Occurrences:**
5,305
- **Population:** 62,057

5. East Gwillimbury

- **Average Household Income:**
\$119,000
- **Number of Crime Occurrences:**
3,328
- **Population:** 34,637

Correlation Matrix for Average Household Income, Population, and Crime Occurrences



ADDITIONAL INSIGHTS

New Indicator

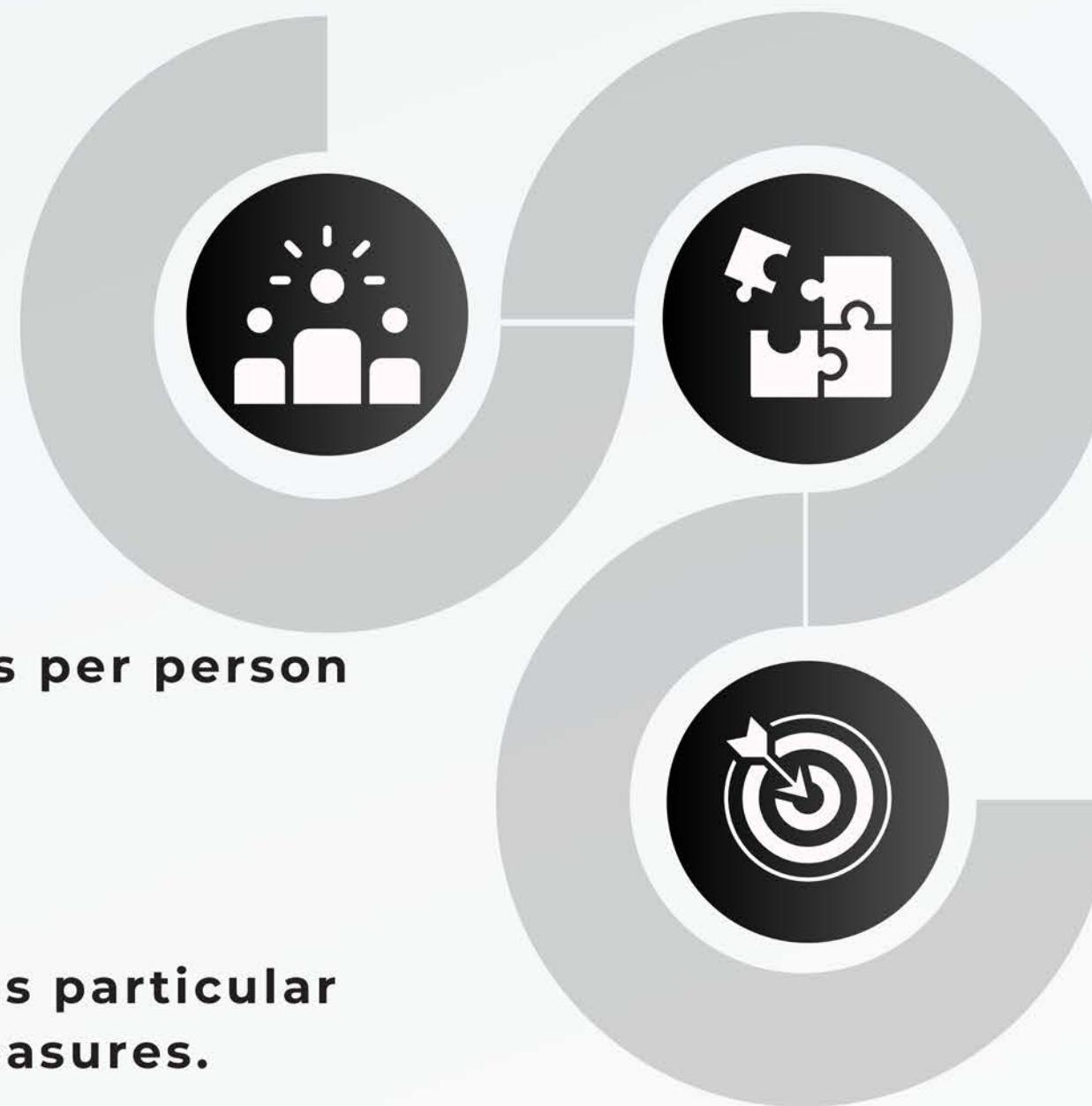
Crime Rate = Occurrence/Population

Key findings

- Average Household Income: \$98,000
- Population: 47,642
- Total Crime Occurrences (occ_type): 6,288
- Crime Rate: Approximately 0.132 occurrences per person
- Income Per Capita: Approximately \$2.06

recommendations

These insights suggest that Georgina requires particular attention for crime prevention and safety measures.



OBJ #1

CRIME DATA INSIGHTS

Crime Type

Understand timeline occurrence of all kind of Crime Type In York Region.

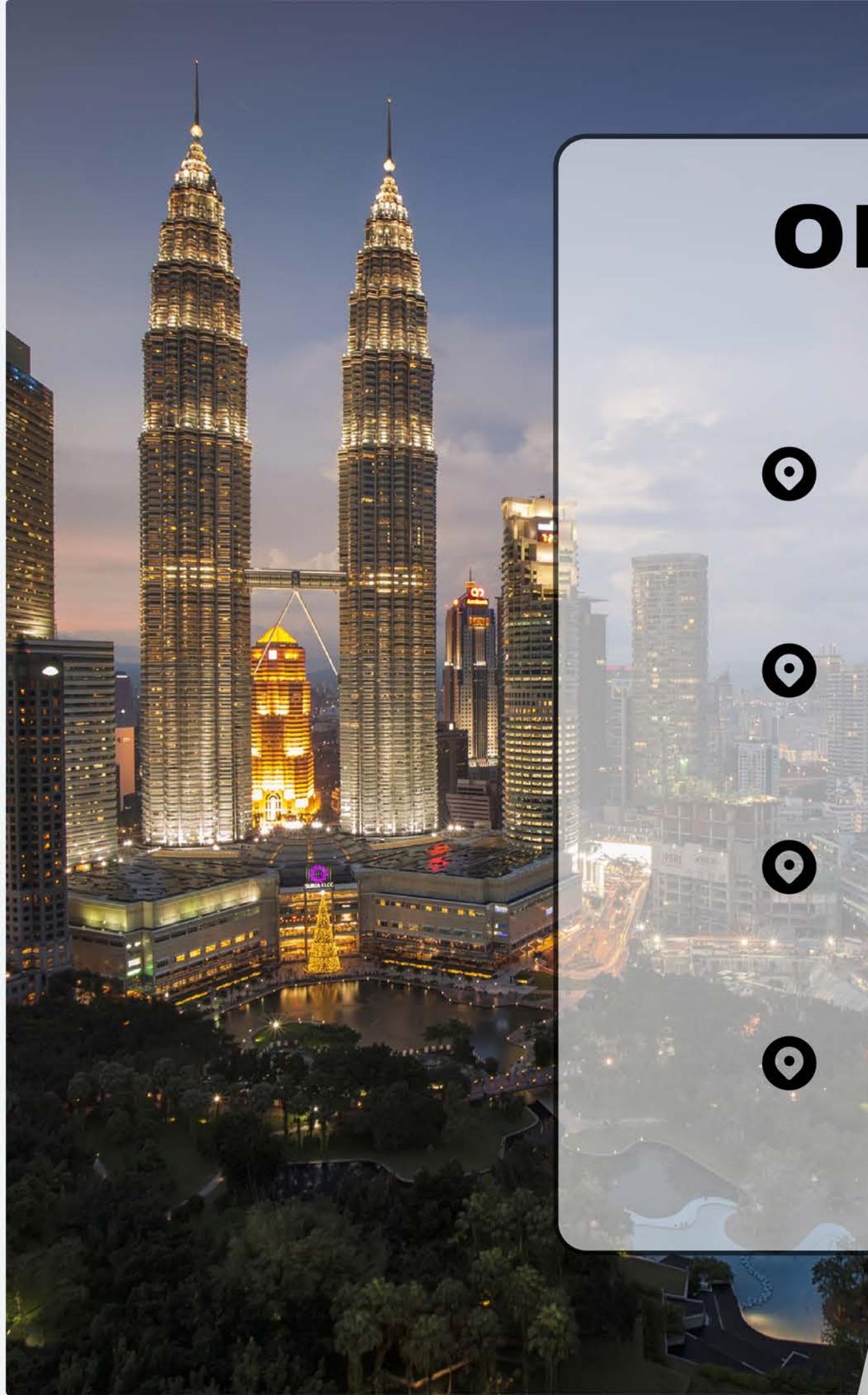
Municipality

Occurrence comparison among all Municipality; focus on the identification of the most risk distrit.

A interactive solution is created for optimize visualization

Additional dataset

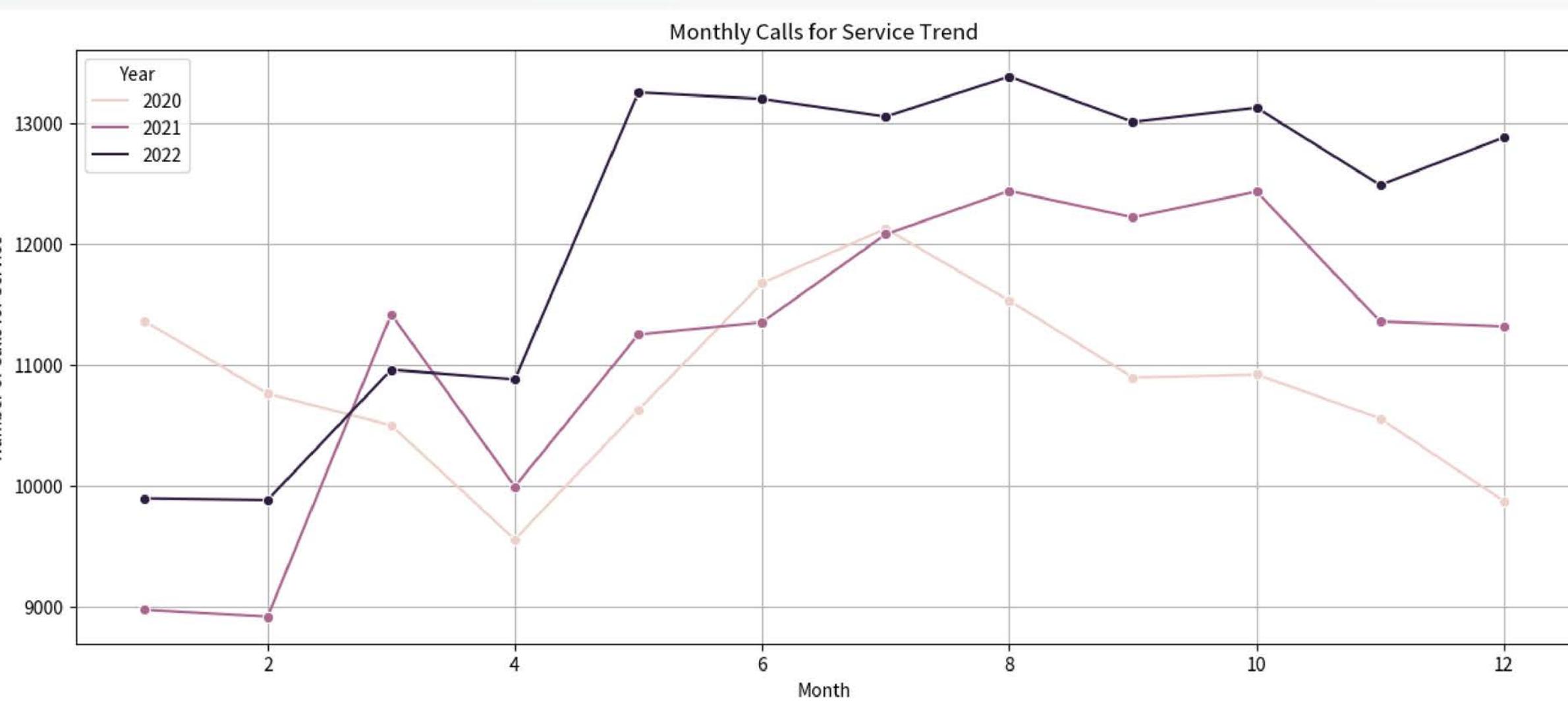
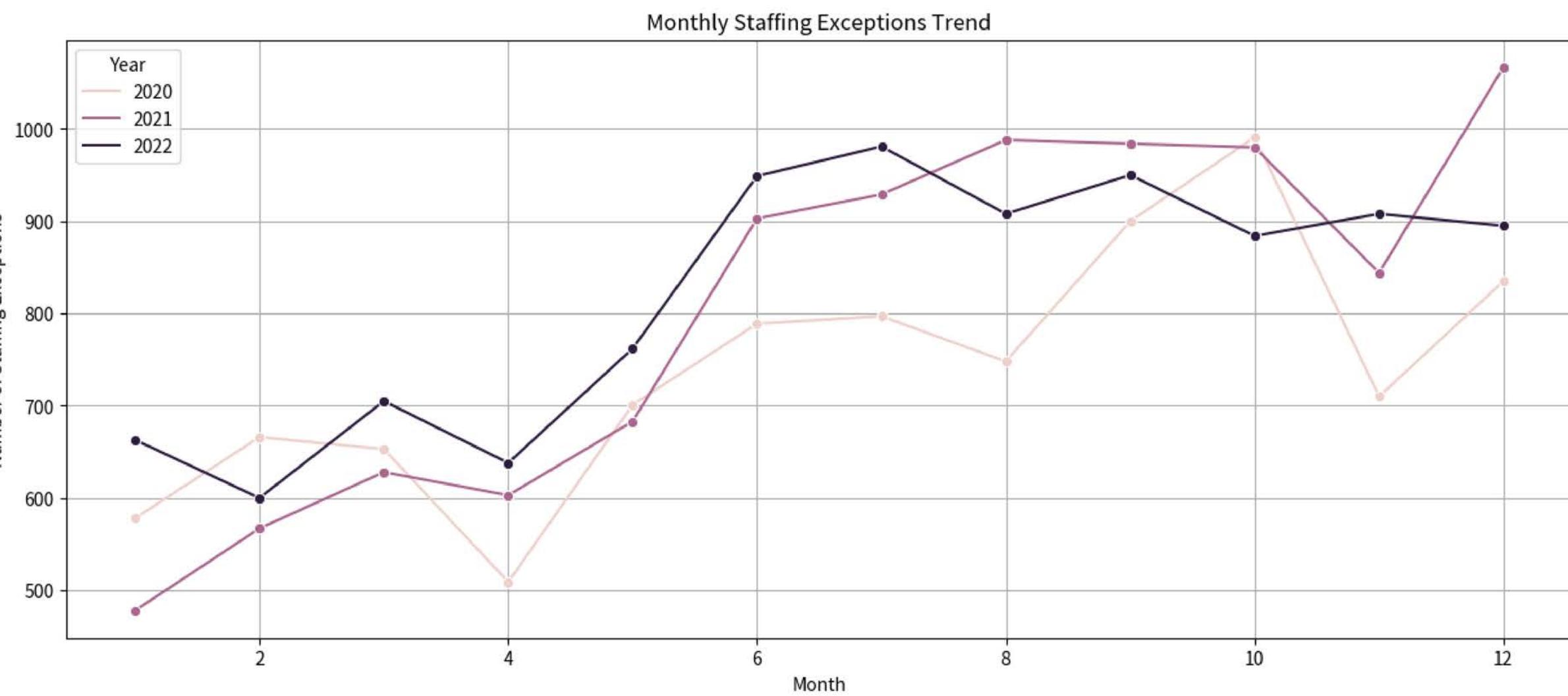
Integrate a publicly available dataset. and how the external factors impact the analysis



OBJ #2 - SERVICE CALLS & FRONT LINE STAFFING

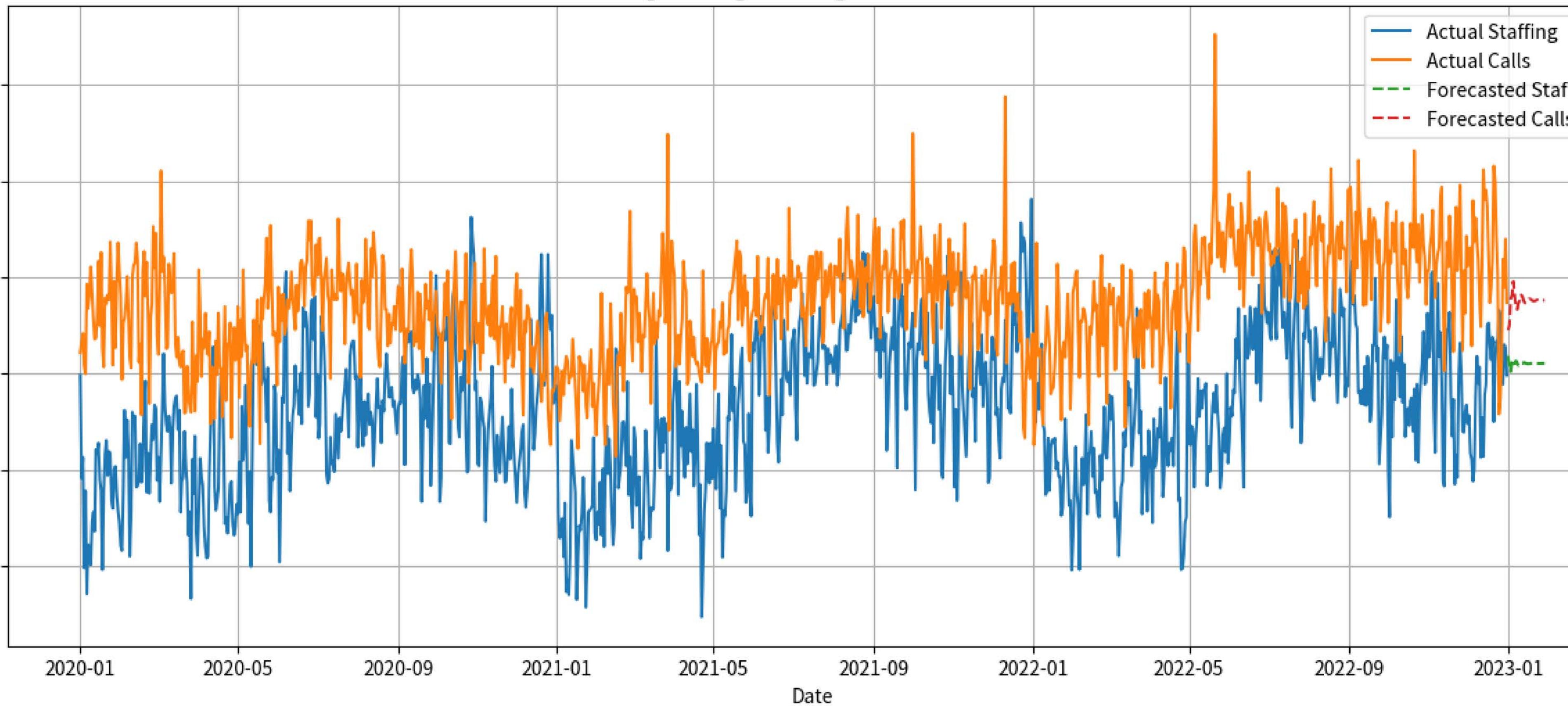
- identify projected calls for service trends in their District and the impacts on front line staffing.
- Identify the days of the year might have the most officers off and coincide with days that have high call demands
- EDA for service call and front line staffing data
- develop a Scheduling Optimization tool for District Command

MONTHLY TREND COMPAIRISON



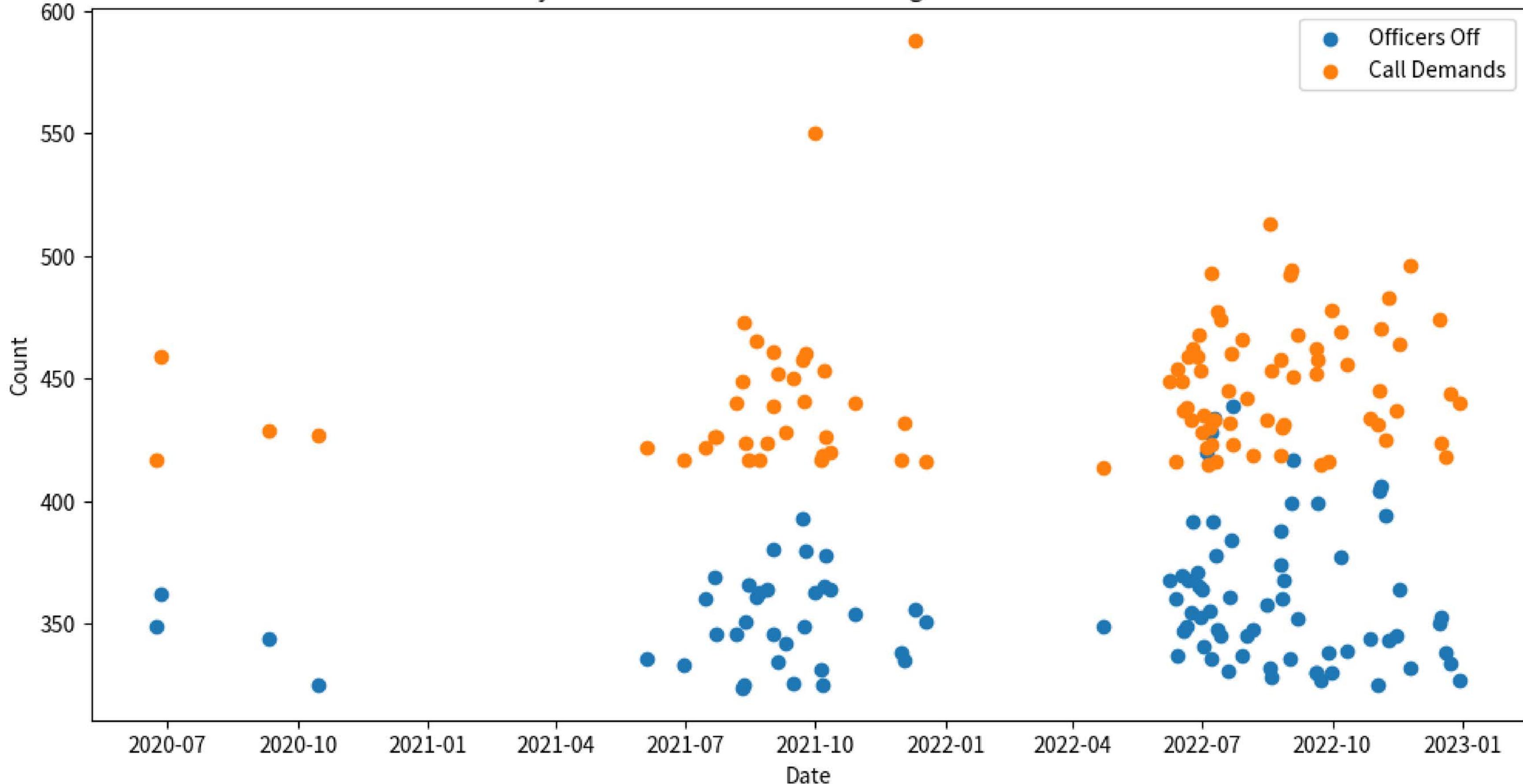
TREND AND FORECASTING

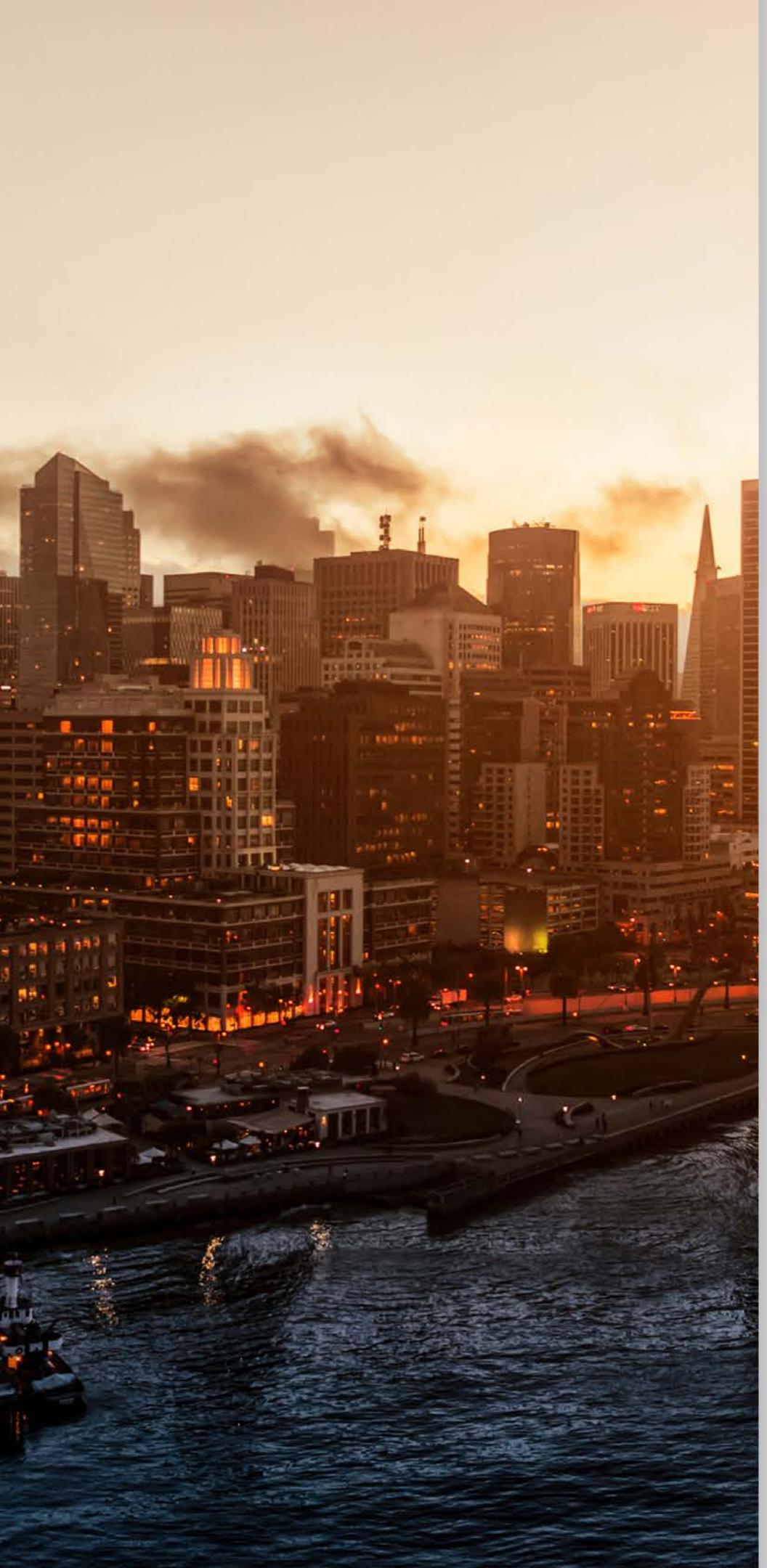
Forecasting Staffing Challenges and Calls for Service



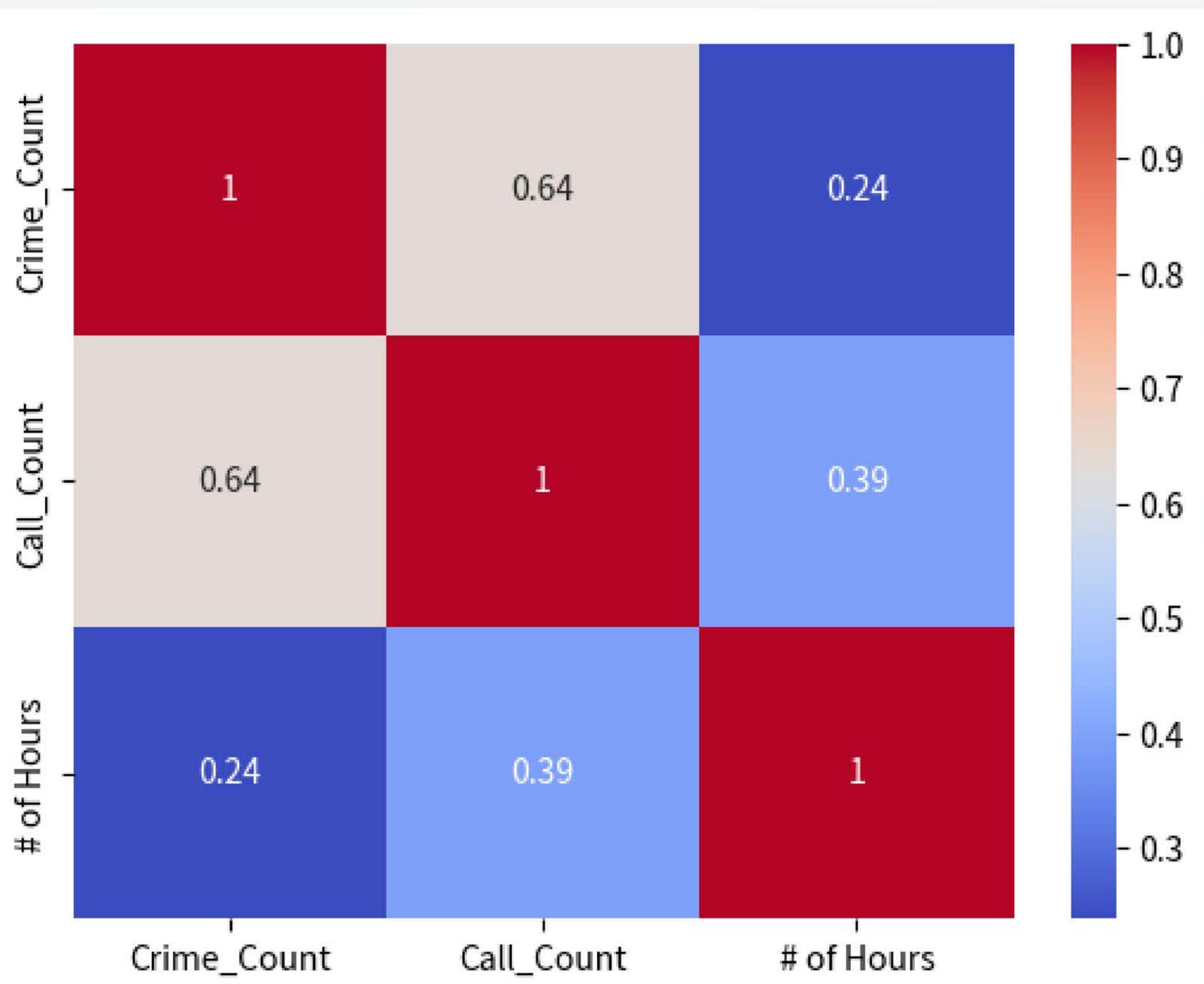
DATE OF HIGH DEMANDS

Days with Most Officers Off and High Call Demands





CORRELATION





SCHEDULING OPTIMIZATION



Available officers: 27



OBJ #2

SERVICE CALL & STAFFING INSIGHTS

Conclusion 1

Days with Most Officers Off & High Call Demands

Date # of Hours Call Type

630	2021-09-22	393.0	458
632	2021-09-24	380.0	460
905	2022-06-24	392.0	462
909	2022-06-28	371.0	459
919	2022-07-08	428.0	493
933	2022-07-22	384.0	460
968	2022-08-26	374.0	458
975	2022-09-02	399.0	494
994	2022-09-21	399.0	458
1010	2022-10-07	377.0	469
1038	2022-11-04	406.0	470

Over time Observation

This forecast can assist in predicting specific days that might present staffing challenges, allowing for better planning and resource allocation

Scheduling Optimization:

interactive tool or dashboard that allows the Deputy Chief to visualize the forecasts and adjust parameters to see potential impacts on staffing

CONCLUSION

Understand the locations in York Region that are the highest risk for increases in criminal activity



Understand projected calls for service trends in their District and the impacts on front line staffing.

Code Demonstration



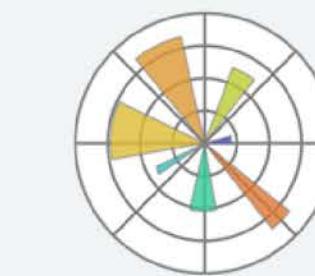
RESOURCE PAGE

Environment

Anaconda

Python 3

Jupyter Notebooks



Libraries

Pandas

Geopandas

Folium

Sci-kit Learn

XGBoost



SweetVIZ

Graphviz

Seaborn

Matplotlib

Pandas

