

CRIME IN TORONTO



Arsalan Arif Radhu

Agenda

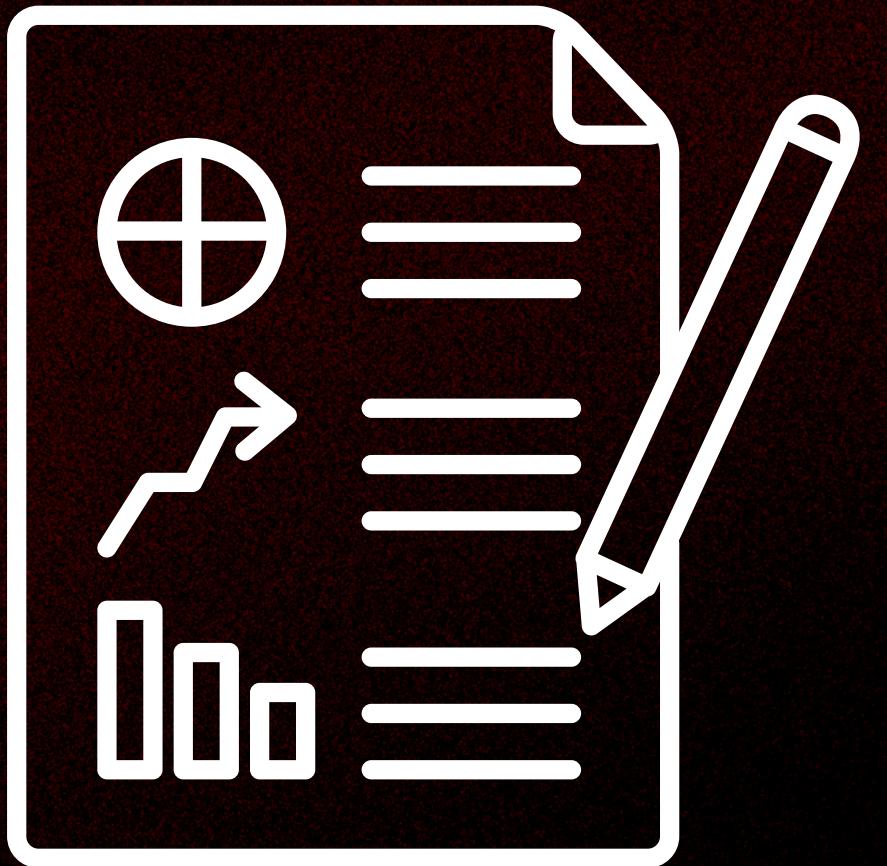
Overview	3
Data Set Overview	4
EDA Findings	5
Baseline Models	6
Next Steps	7



Overview

- Problem Statement
- Proposed Solution
- Estimate of Potential Impact

Data Set Overview



The Dataset

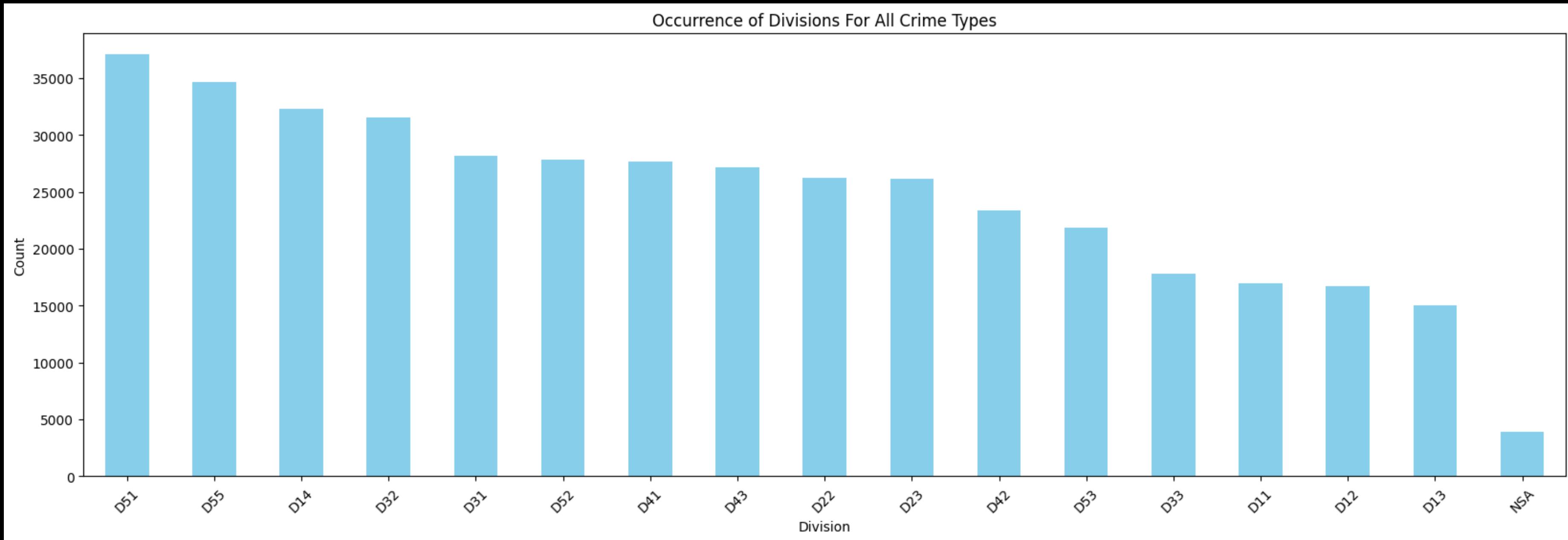
Pre-Processing

30 Predictors

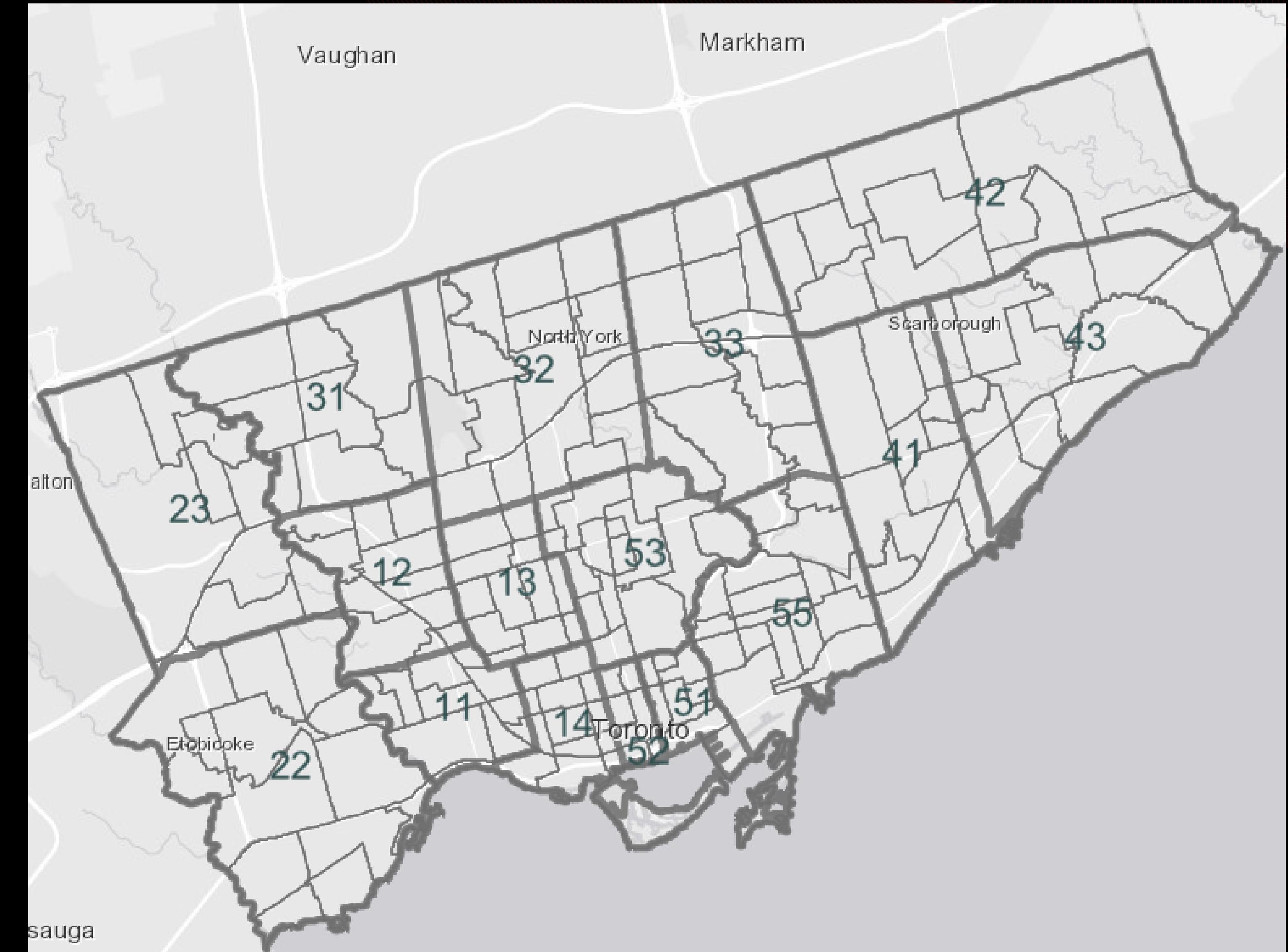
**400,000+
rows**

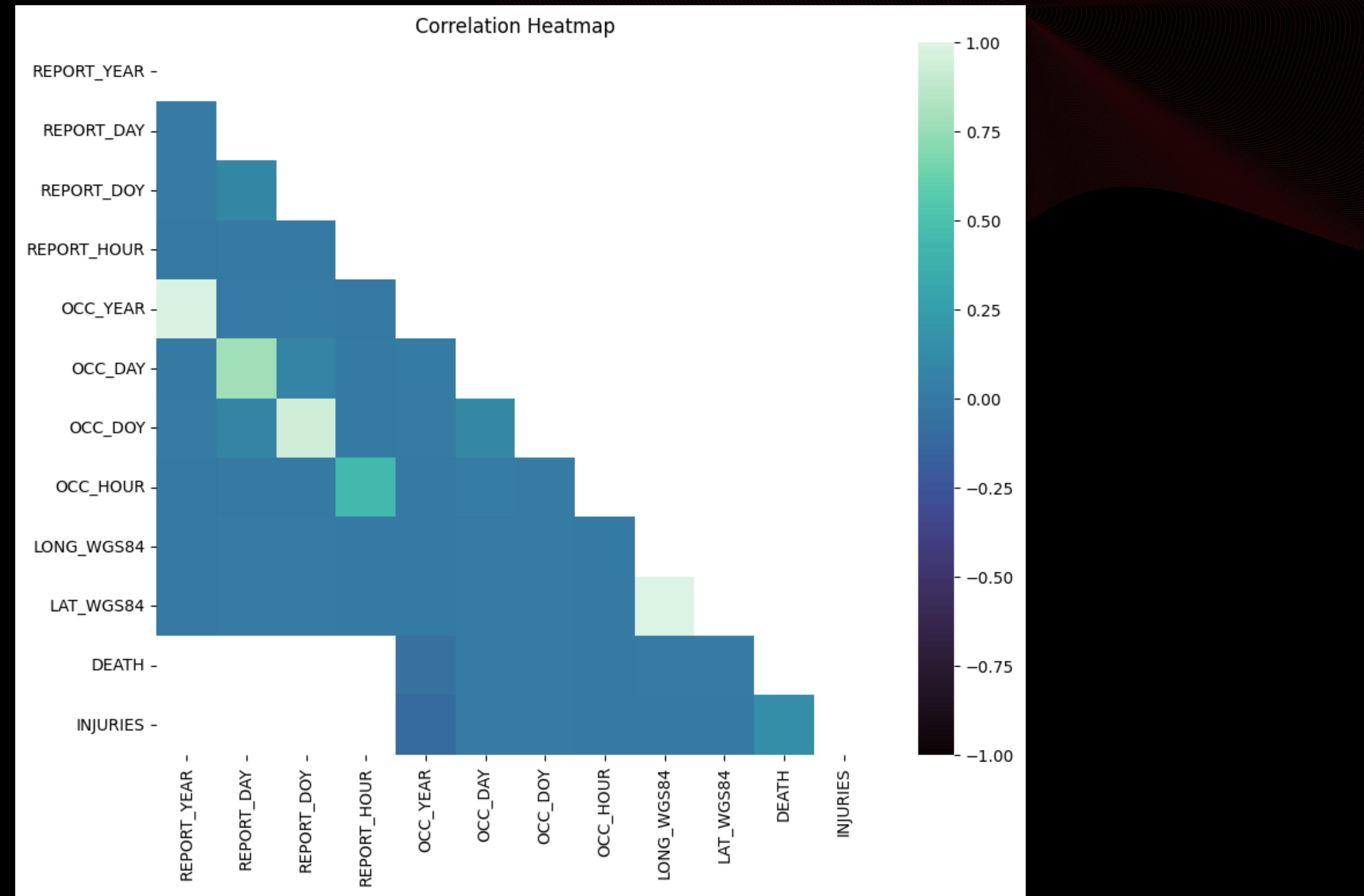
**Logistic
Regression**

EDA FINDINGS

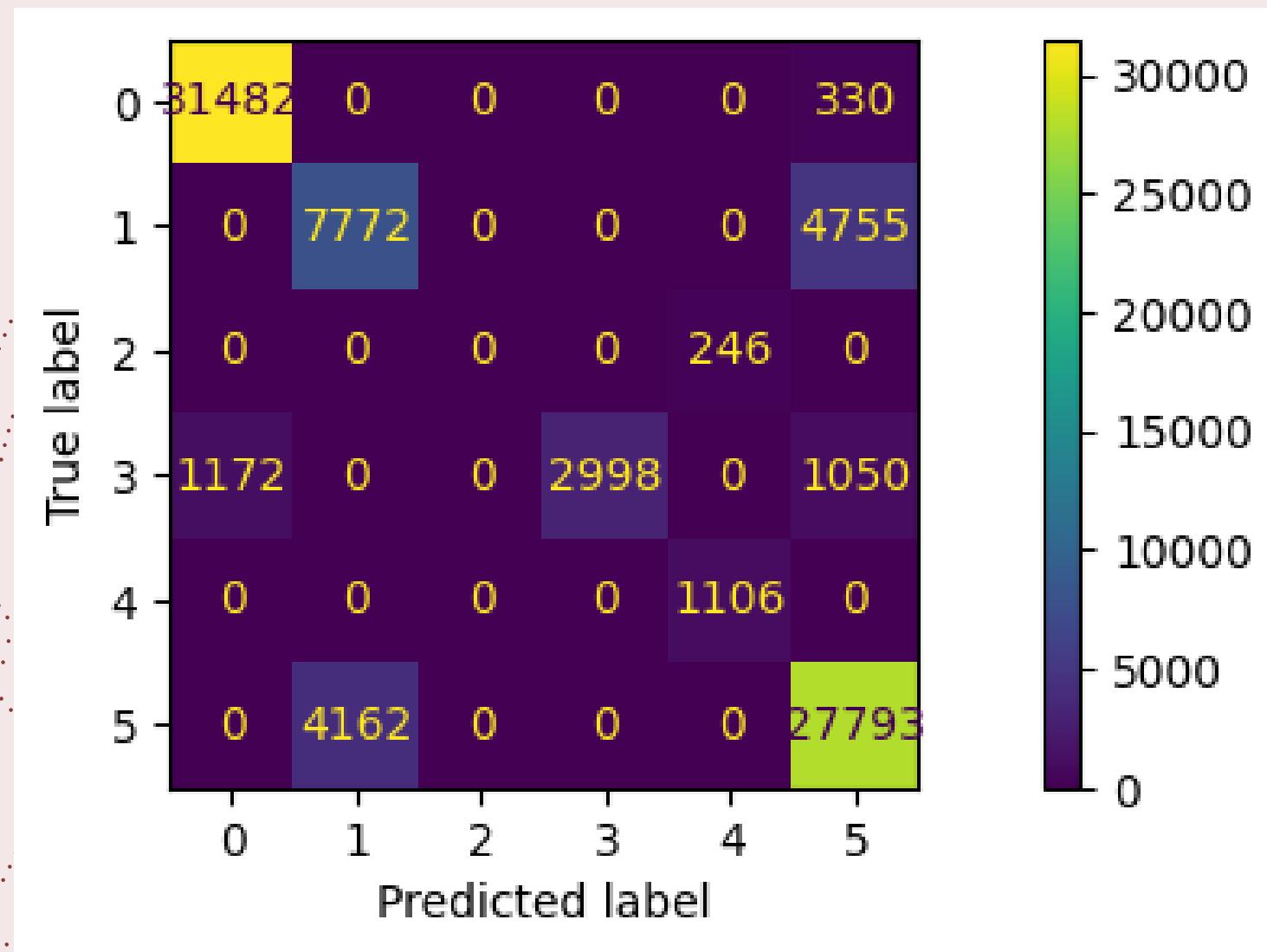


Police Divisions of Toronto





BASELINE MODEL



Logistic Regression:

- Quick Prototyping and Benchmarking
- Robustness to Irrelevant Features

Evaluation:

- Accuracy Score
- Confusion Matrix

NEXT STEPS

- Feature Engineering
- Scalability and Efficiency
- More complex model

