Capstone Project

Predict Home Price Index Based on Major Crime Incidents in Toronto

Allan Salamanca
June 09, 2025
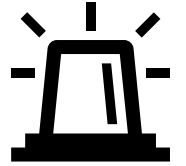
MLS Home Price Index (HPI)

- Established by the Canadian Real Estate Association (CREA) to gauge home price levels and trends.

- Based on four property types, **detached, semi-detached, townhouse and apartment**.

TPS - Major Crime Indicators (MCI)

- 5 crime categories that include: **Assault, Break and Enter, Auto Theft, Robbery and Theft**.

- Data is either provided by the victim, or as the offence occurs or is reported.
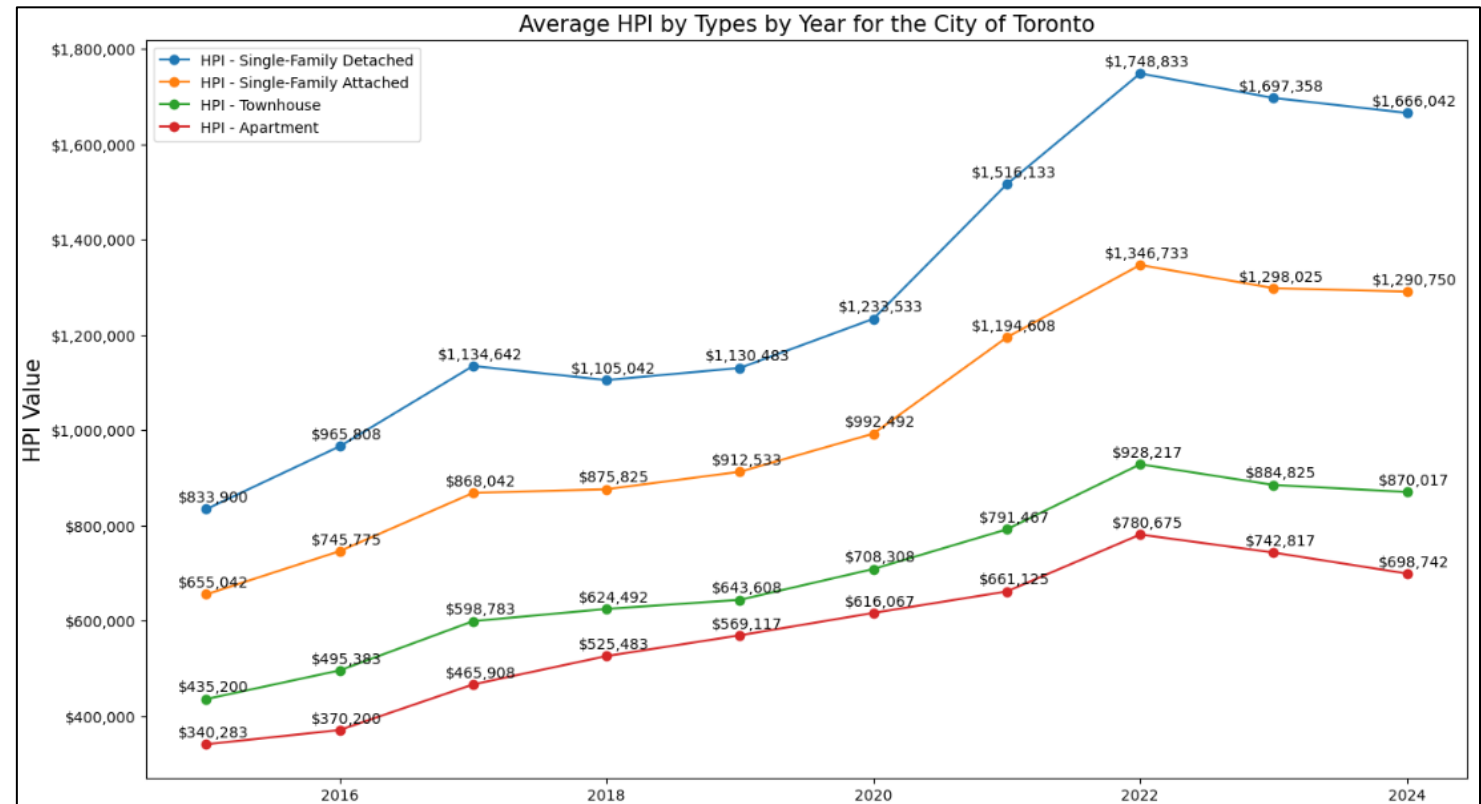
Problem Area

Home Prince Index in Toronto have increased 8% YoY from 2015 to 2024.

At the end of 2024, HPI for properties in Toronto:

- Detached - $1.6M
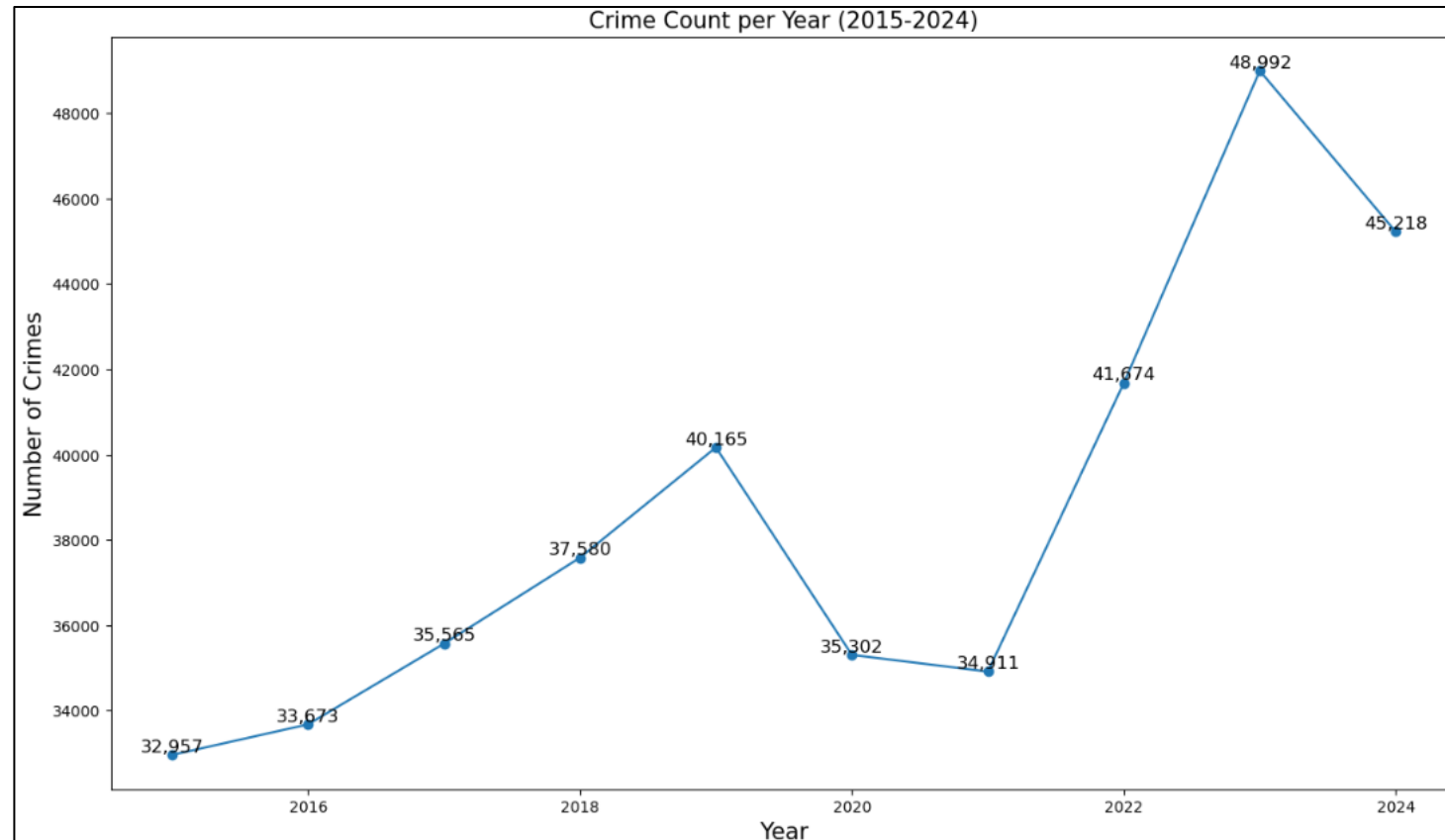- Semi-detached - $1.2M
- Townhouse - $850K
- Apartments - $670K

Average HPI by Types by Year for the City of Toronto

Problem Area

Major crime incidents have increased 4% YoY in Toronto from 2015 to 2024.

Crime Count per Year (2015-2024)
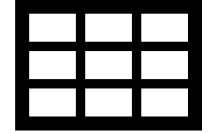
Problem Area: Data Solution

1. Understand what factors affect HPI values.

2. Develop a linear regression model that can predict HPI values by Time Period, Number of Crime Incidents, Property Type and Neighborhoods (Municipality Code).
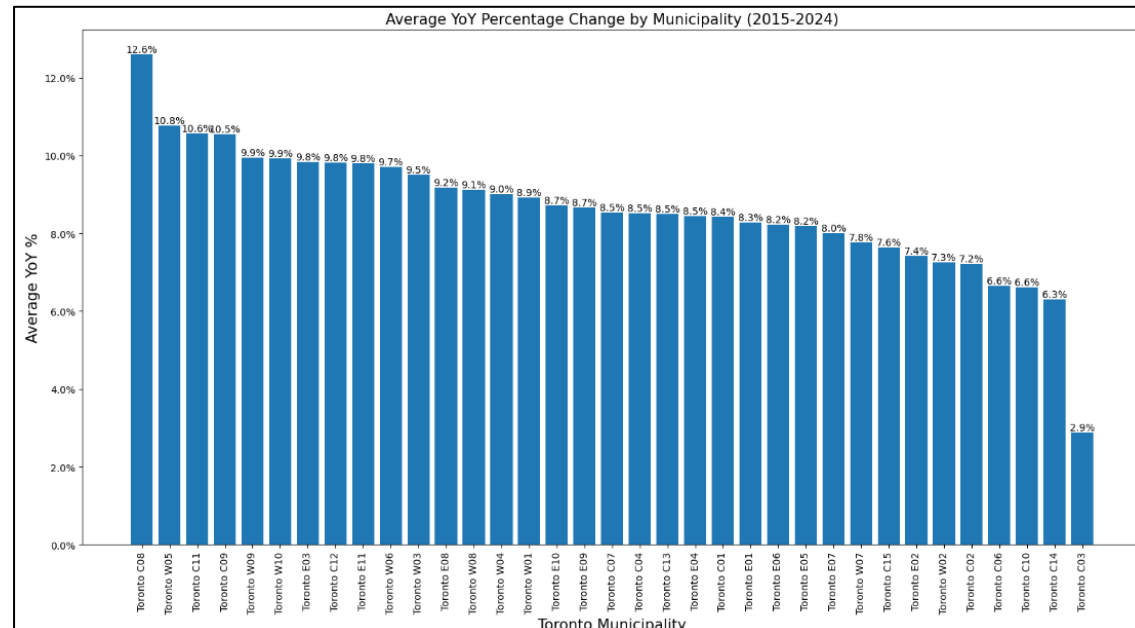
Objective:

Ability to predict and select the neighbourhoods with the best return on investment.
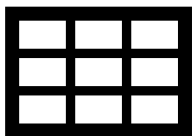
Data Set: Toronto Home Price Index

- Toronto Municipality Code (Neighbourhood)
- HPI - Single-Family Detached
- HPI - Single-Family Attached
- HPI – Townhouse
- HPI - Apartment
- Date (2015-2024)

	Toronto Municipality	HPI - Single-Family Detached	HPI - Single-Family Attached	HPI - Townhouse	HPI - Apartment	Effective Date
0	City of Toronto	787300	625000	413100.0	330200	2015-01-01
1	Toronto W01	849400	677500	415100.0	326300	2015-01-01
2	Toronto W02	805300	661100	410300.0	563000	2015-01-01
3	Toronto W03	509900	493700	NaN	257400	2015-01-01
4	Toronto W04	538000	501900	337000.0	206400	2015-01-01
...	...	...	...	...	...	...
4315	Toronto E07	1219200	1013700	824900.0	593600	2024-12-01
4316	Toronto E08	1219900	872600	752500.0	530000	2024-12-01
4317	Toronto E09	1026700	884600	656600.0	575000	2024-12-01
4318	Toronto E10	1171300	887000	620300.0	486100	2024-12-01
4319	Toronto E11	1062900	885800	709300.0	500300	2024-12-01

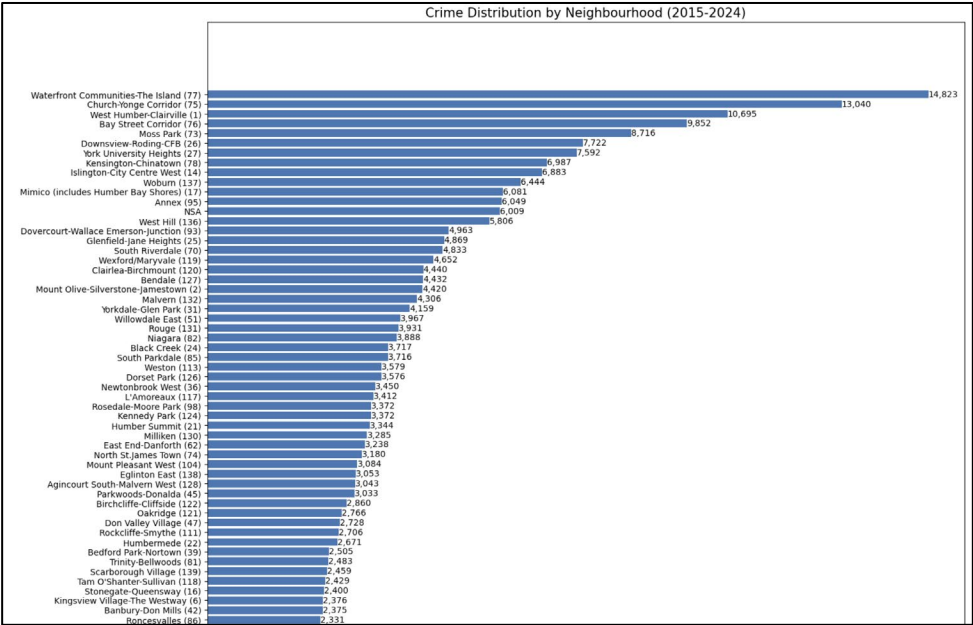Average YoY Percentage Change by Municipality (2015-2024)

Data Set: Toronto Major Crime Indicators

- Occurrence Date

- Division

- MCI Category

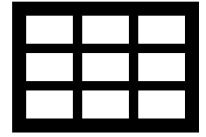- Offence (Subcategory)

- Neighbourhood

- Date (2015-2024)

	EVENT_UNIQUE_ID	OCC_DATE	DIVISION	PREMISES_TYPE	OFFENCE	MCI_CATEGORY	NEIGHBOURHOOD_140
26	GO-20151956	2015-01-01 05:00:00	D32	House	Assault With Weapon	Assault	Newtonbrook West (36)
27	GO-20151923	2015-01-01 05:00:00	D14	Outside	Assault	Assault	Waterfront Communities-The Island (77)
28	GO-20151955	2015-01-01 05:00:00	D42	Outside	Robbery - Mugging	Robbery	Malvern (132)
29	GO-20151945	2015-01-01 05:00:00	D41	Apartment	Assault	Assault	Wexford/Maryvale (119)
30	GO-20152177	2015-01-01 05:00:00	D14	Apartment	Assault	Assault	Dufferin Grove (83)
...	...	...	...	...	...	...	...
386176	GO-20242816627	2024-12-24 05:00:00	D32	Outside	Theft Of Motor Vehicle	Auto Theft	Clanton Park (33)
386177	GO-20242815828	2024-12-30 05:00:00	D31	Outside	Theft Of Motor Vehicle	Auto Theft	Glenfield-Jane Heights (25)
386178	GO-20242814680	2024-12-31 05:00:00	D51	Outside	Robbery - Mugging	Robbery	Church-Yonge Corridor (75)
386179	GO-20242815237	2024-12-31 05:00:00	D32	Commercial	B&E	Break and Enter	Englemount-Lawrence (32)
386180	GO-20242815267	2024-12-31 05:00:00	D14	Commercial	B&E	Break and Enter	Trinity-Bellwoods (81)

Crime Distribution by Neighbourhood (2015-2024)
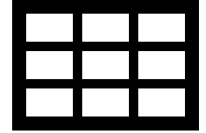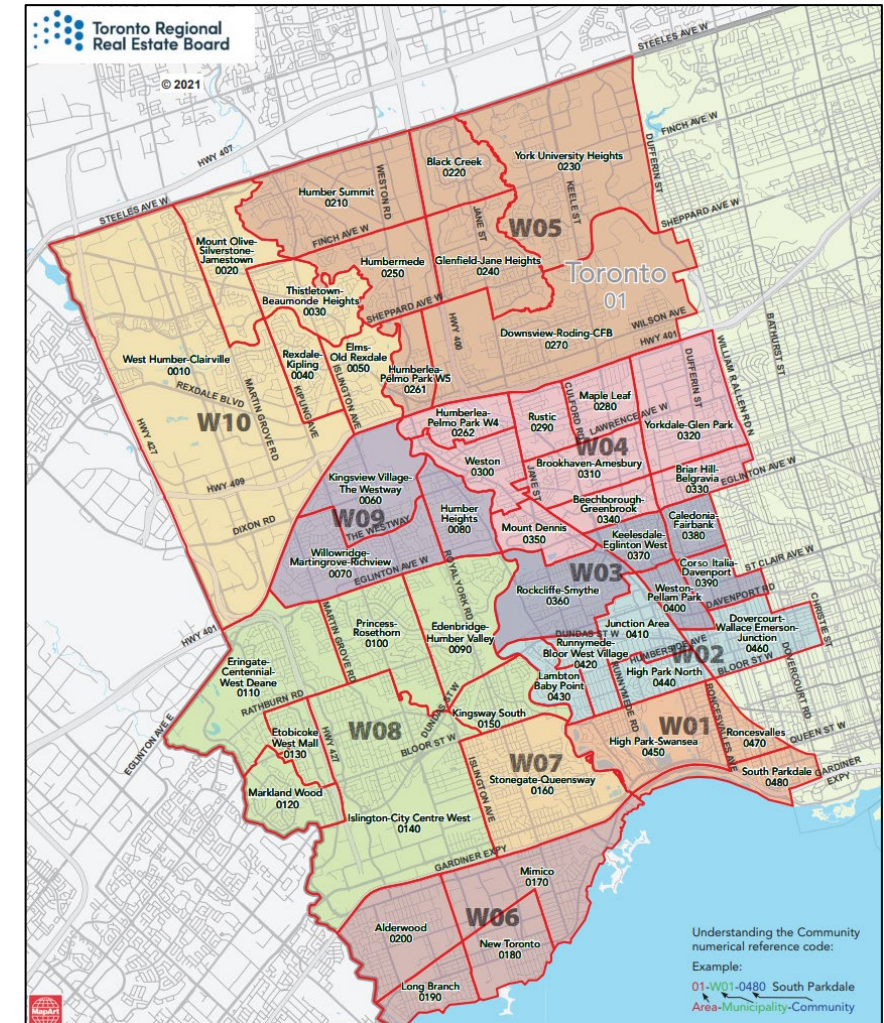
Data Set: Toronto Neighbourhood Index

- Manually created table.

- Helper table to join TPS and MLS neighbourhood to MLS municipality codes.

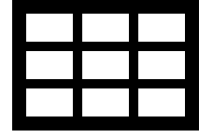|	Toronto Police Neighbourhood Categories	MLS Neighbourhood Categories	MLS Municipality Code
0	Agincourt North (129)	Agincourt North	Toronto E07
1	Agincourt South-Malvern West (128)	Agincourt South-Malvern West	Toronto E07
2	Alderwood (20)	Alderwood	Toronto W06
3	Annex (95)	Annex	Toronto C02
4	Banbury-Don Mills (42)	Banbury-Don Mills	Toronto C13
...	...	...	...
136	Wychwood (94)	Wychwood	Toronto C02
137	Yonge-Eglinton (100)	Yonge-Eglinton	Toronto C03
138	Yonge-St.Clair (97)	Yonge-St.Clair	Toronto C02
139	York University Heights (27)	York University Heights	Toronto W05
140	Yorkdale-Glen Park (31)	Yorkdale-Glen Park	Toronto W04

Data Preprocessing

- Data Sets:
  - Toronto Home Price Index
  - Toronto Major Crime Incidents
  - Toronto Neighbourhood Index

- Joined 3 data sets to create a master data frame based on unique neighbourhoods to a municipality code.

- Dummies created for categories: property type and neighbourhood.

- Imputed missing data using KNN, for Townhouses in specific neighbourhoods.
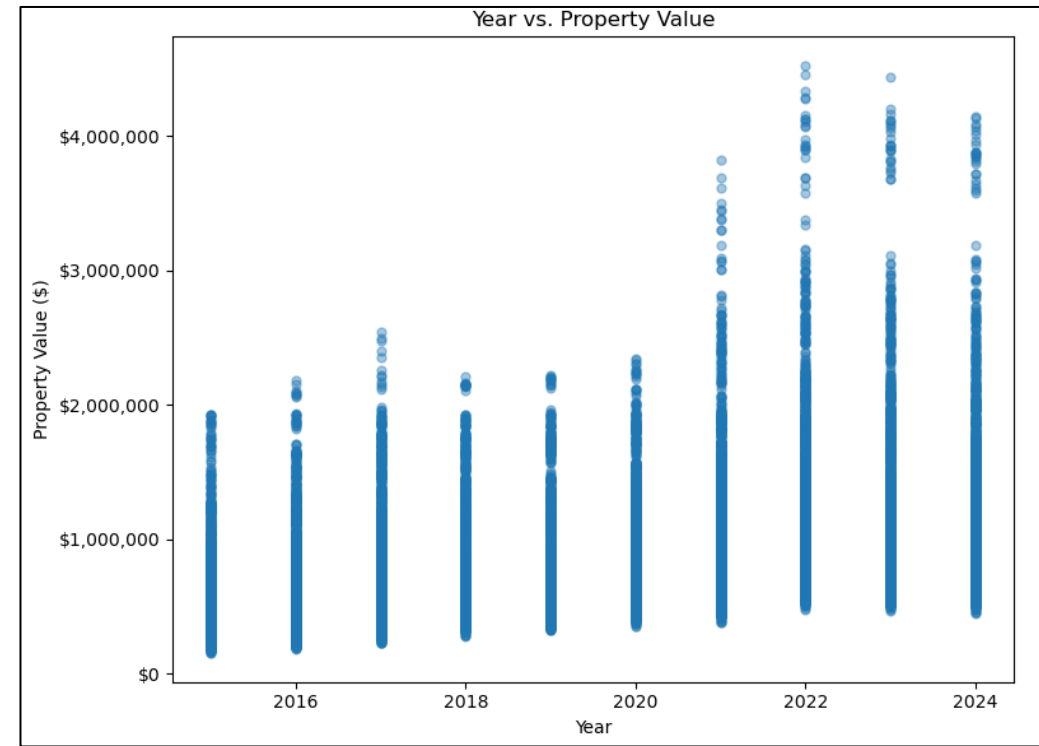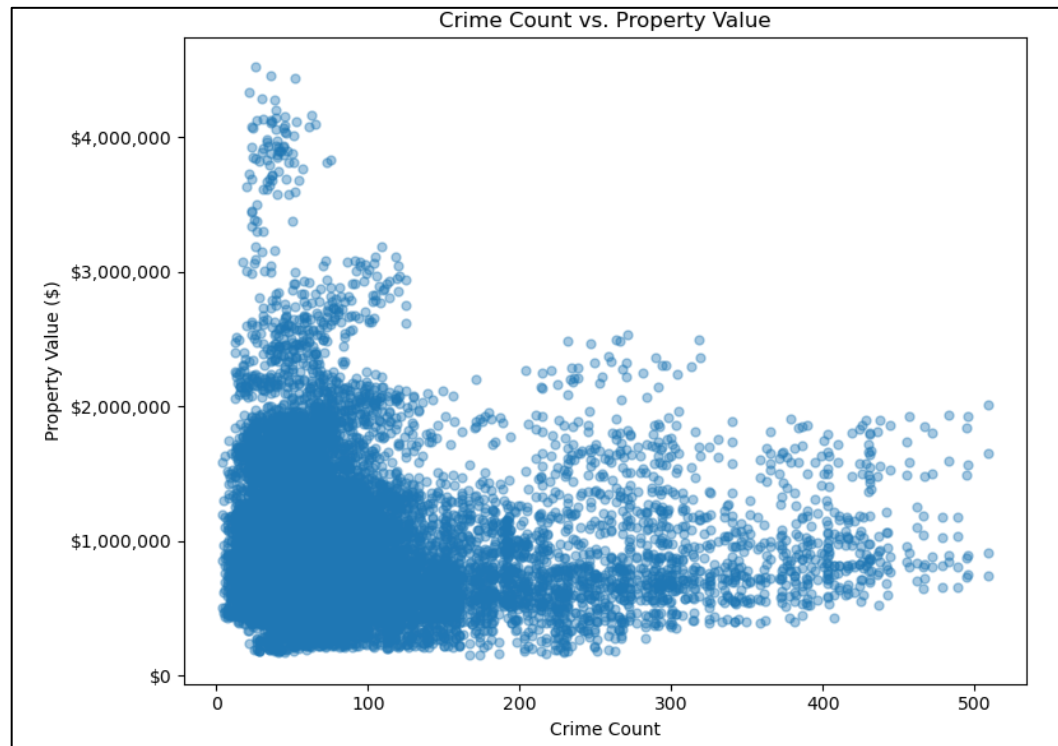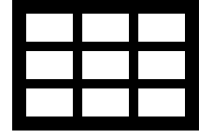
Data Preprocessing

Data Summary - 16,800 rows & 45 columns
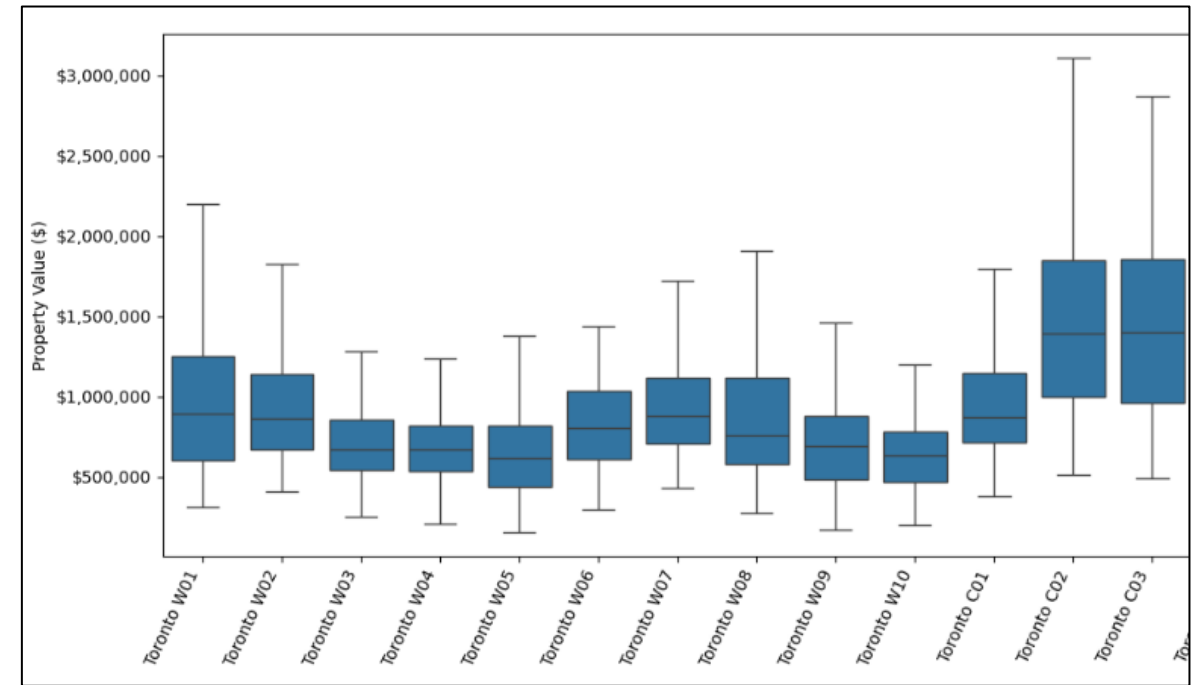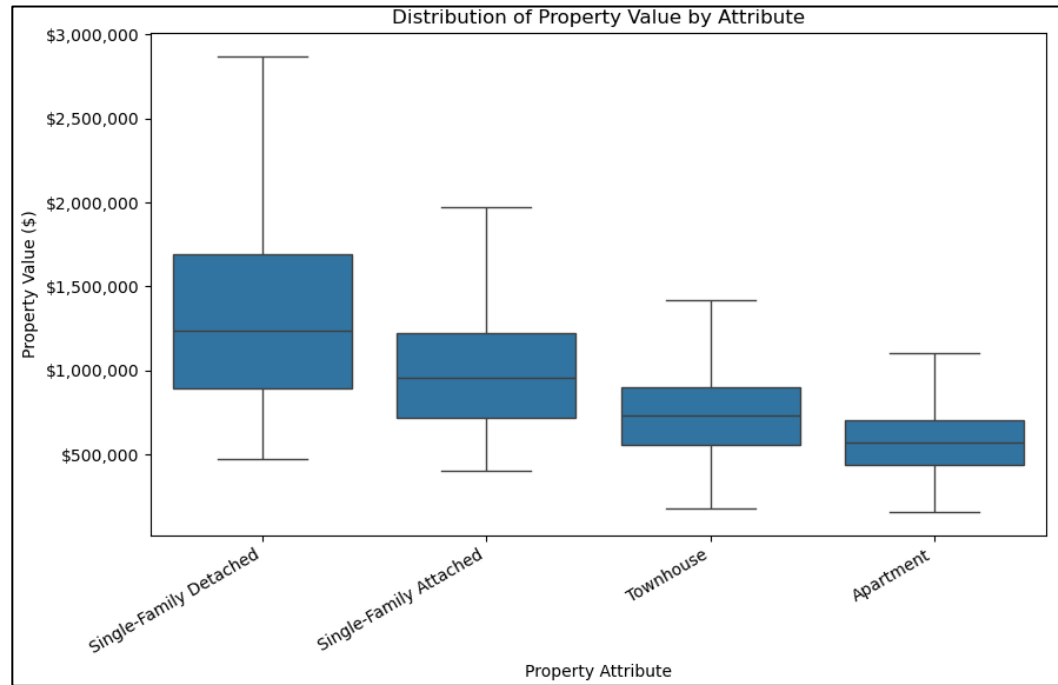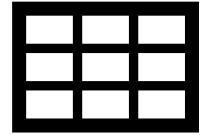
Dependent Variable - Value (HPI)

Independent Variables: Year, Crime Count, Property Type (Attribute) and Neighbourhood (Toronto Municipality)

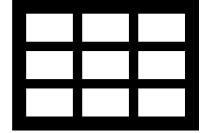| Attribute_HPI - Single-Family Detached	Attribute_HPI - Townhouse	Crime Count	Month	Toronto Municipality_Toronto C01	Toronto Municipality_Toronto C02	Toronto Municipality_Toronto C03	...	Toronto Municipality_Toronto W04	Toronto Municipality_Toronto W05	Toronto Municipality_Toronto W06	Toronto Municipality_Toronto W07	Toronto Municipality_Toronto W08	Toronto Municipality_Toronto W09	Toronto Municipality_Toronto W10	Value	Year
1.0	0.0	44.0	1.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	849400.0	2015.0
0.0	0.0	44.0	1.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	677500.0	2015.0
0.0	1.0	44.0	1.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	415100.0	2015.0
0.0	0.0	44.0	1.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	326300.0	2015.0
1.0	0.0	62.0	1.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	805300.0	2015.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
0.0	0.0	56.0	12.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	486100.0	2024.0
1.0	0.0	58.0	12.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1062900.0	2024.0
0.0	0.0	58.0	12.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	885800.0	2024.0
0.0	1.0	58.0	12.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	709300.0	2024.0
0.0	0.0	58.0	12.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	500300.0	2024.0

Exploratory Data Analysis

Exploratory Data Analysis

Baseline Modelling

Standard Scaler & Linear Regression

- Train score: 77%
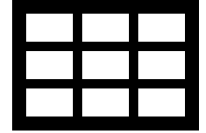- Test score: 77%
- Test RMSE: $247K
- Test MAE: $167K

OLS Results

- R-Squared – 0.771
- No p-value > 0.05

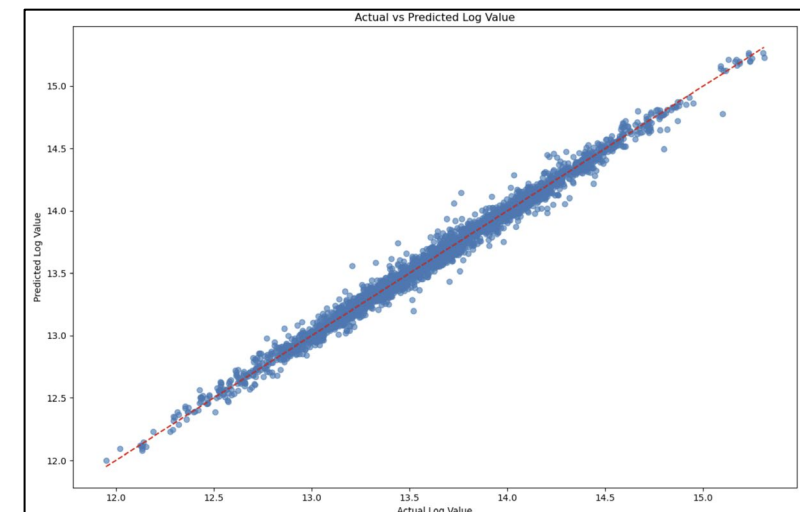Actual vs. Predicted Property Value
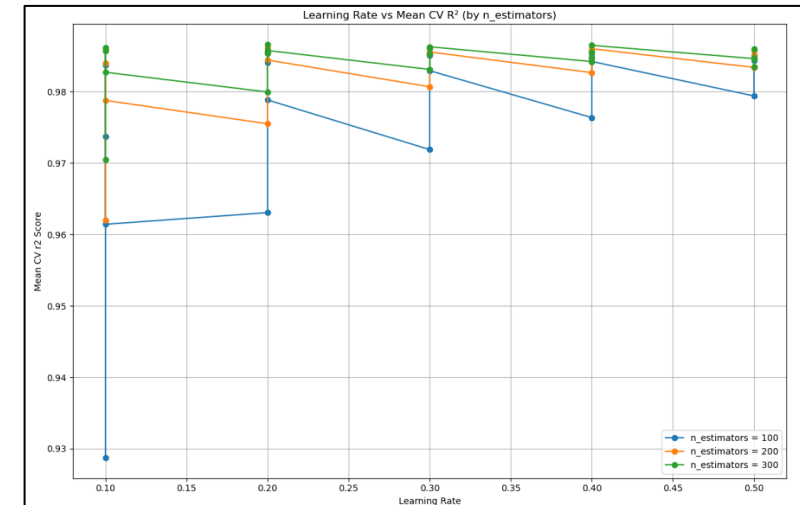
Advanced Modelling

- Compared three models: Decision Tree Regressor, Random Forest Regressor, and eXtreme Gradient Boosting.

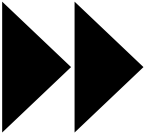- Utilized GridSearchCV to select the best model and parameters.

Best Model: eXtreme Gradient Boosting

Params: Learning Rate: 0.2 & Estimators: 300

- Test score: 98%

- Test RMSE: $61K

- Test MAE: $37K

Demo – Web App

Toronto Housing Price Index Predictor based on Crime Rates

Property Type

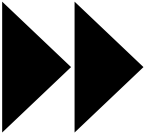Apartment

Toronto Municipality

C01

Year

2020

Crime Count

0

Predict Price

Municipality and Neighbourhood Reference Table 🔗

	Toronto Municipality	Neighbourhoods
0	C01	Bay Street Corridor, Dufferin Grove, Kensington-Chinatown, Little Portugal, Niagara,
1	C02	Annex, Casa Loma, Wychwood, Yonge-St.Clair
2	C03	Forest Hill South, Humewood-Cedarvale, Oakwood Village, Yonge-Eglinton
3	C04	Bedford Park-Nortown, Englemount-Lawrence, Forest Hill North, Lawrence Park Nor
4	C06	Bathurst Manor, Clanton Park
5	C07	Lansing-Westgate, Newtonbrook West, Westminster-Branson, Willowdale West
6	C08	Cabbagetown-South St.James Town, Church-Yonge Corridor, Moss Park, North St.Ja
7	C09	Rosedale-Moore Park
8	C10	Mount Pleasant East, Mount Pleasant West
9	C11	Flemingdon Park, Leaside-Bennington, Thorncliffe Park
10	C12	Bridle Path-Sunnybrook-York Mills, St.Andrew-Windfields
11	C13	Banbury-Don Mills, Parkwoods-Donalda, Victoria Village
12	C14	Newtonbrook East, Willowdale East
13	C15	Bayview Village, Bayview Woods-Steeles, Don Valley Village, Henry Farm, Hillcrest Vi
14	E01	Blake-Jones, Greenwood-Coxwell, North Riverdale, South Riverdale
15	E02	East End-Danforth, The Beaches, Woodbine Corridor
16	E03	Broadview North, Danforth, Danforth East York, O'Connor-Parkview, Old East York, F

Future State – Expanding the Model ▶▶

- Population Density
- Postal Code
- Crime Type
- Neighbourhood Ratings:
    - School
    - Hospitals
    - Community Centers
    - Police Station
    - Fire Station
    - Etc.
- Development
- # of Listings
- Bank Interest Rates