Analysis of Crime Statistics from 2014 to 2023 in Neighborhoods of Varying Home Price*

A Statistical Basis for Policy Improvement

Chenika Bukes

September 24, 2024

This study analyzes crime rates across Toronto's neighborhoods from 2014 to 2023, focusing on the 10 neighborhoods with the highest home prices and the 10 neighborhoods with the lowest home prices. We investigated crime trends across nine categories, including assault, robbery, and autotheft, comparing the average crime rates between high-price and low-price neighborhoods. The analysis reveals a significant disparity in crime rates, with lower-priced neighborhoods consistently exhibiting higher crime rates, especially in violent crimes such as shootings and robberies. These findings underscore the importance of targeted crime prevention strategies in economically disadvantaged neighborhoods and suggest that housing affordability is closely linked to public safety concerns.

Table of Contents

---

*Code and data are available at: https://github.com/chenikabukes/TorontoDataset

1 Introduction

We run the model in R (R Core Team 2023). Understanding the relationship between housing prices and crime rates is critical for urban planning and public policy. Housing prices are often considered a key indicator of socio-economic status, with higher-priced neighborhoods typically associated with lower crime rates, greater economic stability, and improved social services. Conversely, lower-priced neighborhoods are often linked to higher crime rates, economic challenges, and reduced access to resources. However, while these patterns may hold in many cases, the relationship between housing prices and crime rates is complex and not universally consistent across all regions. This complexity raises the need for more detailed investigation into how these factors interact in specific urban settings, such as Toronto.

Despite numerous studies exploring socio-economic indicators and crime, there remains a gap in understanding how housing prices, particularly in high- and low-income neighborhoods, correlate with different types of crimes over time. Previous research often treats these neighborhoods as static entities, overlooking the dynamic nature of crime trends and their fluctuations over time. This paper aims to fill this gap by analyzing the crime rates across Toronto's neighborhoods from 2014 to 2023, specifically focusing on the 10 neighborhoods with the highest and the 10 neighborhoods with the lowest home prices.

In this analysis, we examined nine distinct crime categories—including assault, robbery, and autotheft—and compared the trends in both high- and low-priced neighborhoods. The goal was to identify significant patterns in crime rates, understand whether wealthier neighborhoods consistently experience lower crime rates, and investigate which types of crimes are most prevalent in different socio-economic contexts. Our findings reveal notable disparities, particularly in violent crimes such as shootings and robberies, which were found to be more prevalent in lower-income neighborhoods.

These findings are important because they offer insights that can help city planners and policymakers design targeted crime prevention strategies. By understanding how crime rates differ across neighborhoods with varying economic conditions, more effective resource allocation and policy interventions can be implemented. The structure of this paper is as follows: first, we review the existing literature on the relationship between housing prices and crime. Second, we present our methodology and data sources. Third, we provide a detailed analysis of the results, followed by a discussion of the implications for public policy and future research.

2 Data

2.1 Raw Data

The data for this analysis comes from two primary sources obtained from Toronto's Open Data portal:

Housing Price Data: This dataset includes information on average home prices across various neighborhoods in Toronto. It serves as an indicator of the socio-economic status of each neighborhood. The data is extracted from the "Wellbeing Toronto" dataset, which records various socio-economic variables across the city. For this analysis, the focus is on the average home price for each neighborhood, identified by the neighborhood name and ID. This variable acts as a proxy for socio-economic conditions, with higher home prices suggesting wealthier neighborhoods, and lower prices implying economically disadvantaged areas.

Crime Statistics Data: This dataset contains detailed annual crime statistics for each neighborhood between 2014 and 2023. It includes crime rates per 100,000 residents for a range of criminal activities: - Assault - Auto theft - Bike theft - Break and enter - Homicide - Robbery - Shooting - Theft from motor vehicle - Theft over $5,000 The crime data provides a robust view of criminal activity across various categories, offering insight into both violent and non-violent crime trends in different neighborhoods.

Data Selection Reasoning While Toronto's census data also contains valuable socio-economic information, the housing price dataset was chosen because it was updated more frequently and provides clearer spatial links between neighborhoods and socio-economic status. The census data, recorded in 2015 and 2021, presented two key limitations for this analysis: (1) It was not yet fully updated to 2021 data, and (2) the census dataset organizes

data by individual household income rather than neighborhood-level data, making it more challenging to compare neighborhoods directly. The housing price data, already grouped by neighborhood code, was easier to integrate with the crime statistics for this study's purpose.

Variables and their Roles The key variables under examination are: - Neighborhood Name and ID: Each neighborhood is identified by both a name and a unique ID code. - Home Prices: Average home prices (in CAD) across neighborhoods, reflecting the relative affluence of each area. - Crime Rates: Crime statistics for nine different crime categories (listed above), measured as the number of incidents per 100,000 residents annually. These crime rates provide a normalized metric for comparing areas with different population sizes. These variables allow for a detailed analysis of the relationship between socio-economic status (as proxied by home prices) and crime rates, both across crime categories and over time.

2.2 Data Cleaning and Preparation

The housing data was prepared by filtering out invalid entries and ensuring that all price data was numeric. Only neighborhoods with valid, positive home price values were included. The crime data was cleaned by removing any missing entries and focusing exclusively on the crime rates. Non-relevant columns (i.e., raw crime counts) were removed in favor of normalized rates per 100,000 residents, making comparisons across neighborhoods more meaningful.

Focus Neighborhoods To investigate the potential relationship between socio-economic status and crime, the analysis focuses on the neighborhoods with the 10 highest and the 10 lowest average home prices. These neighborhoods represent a wide socio-economic spectrum and are expected to exhibit varying patterns of criminal activity. By isolating these groups, we can assess whether wealthier neighborhoods experience lower crime rates and determine which types of crimes are more prevalent in less affluent areas.

2.3 Summary Statistics and Visualisation

Below, we provide an overview of the housing price data for the top 10 highest and bottom 10 lowest-priced neighborhoods, followed by visualizations that illustrate trends in crime rates across Toronto from 2014 to 2023.

2.4 Crime Rate Changes over Time

Crime rate trends over the years (2014-2023) were analyzed for each crime type to understand patterns and shifts. This is visualized in the following plots:
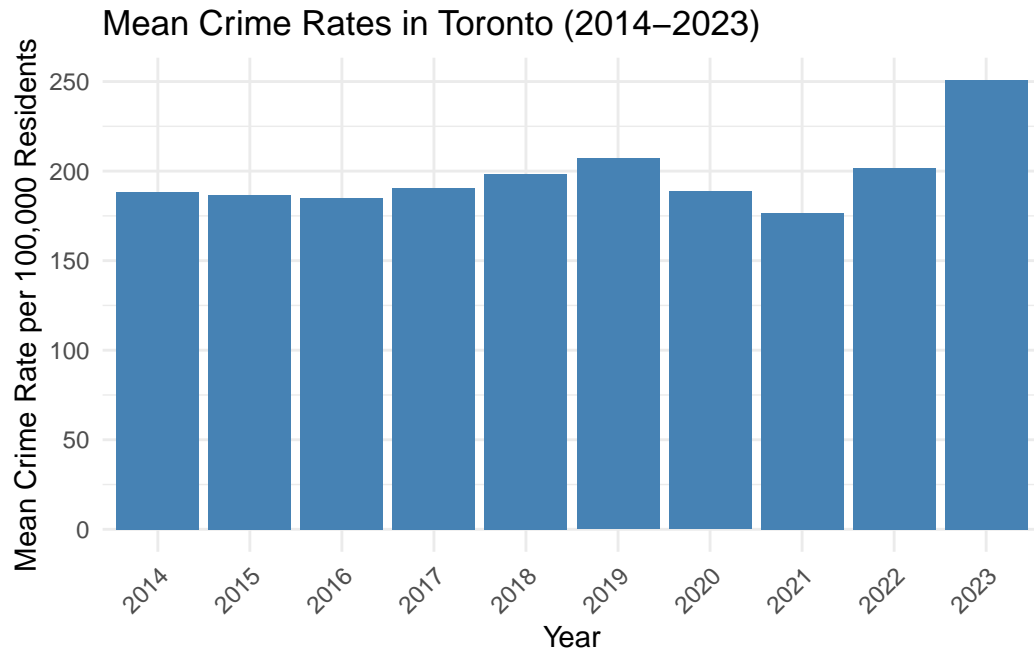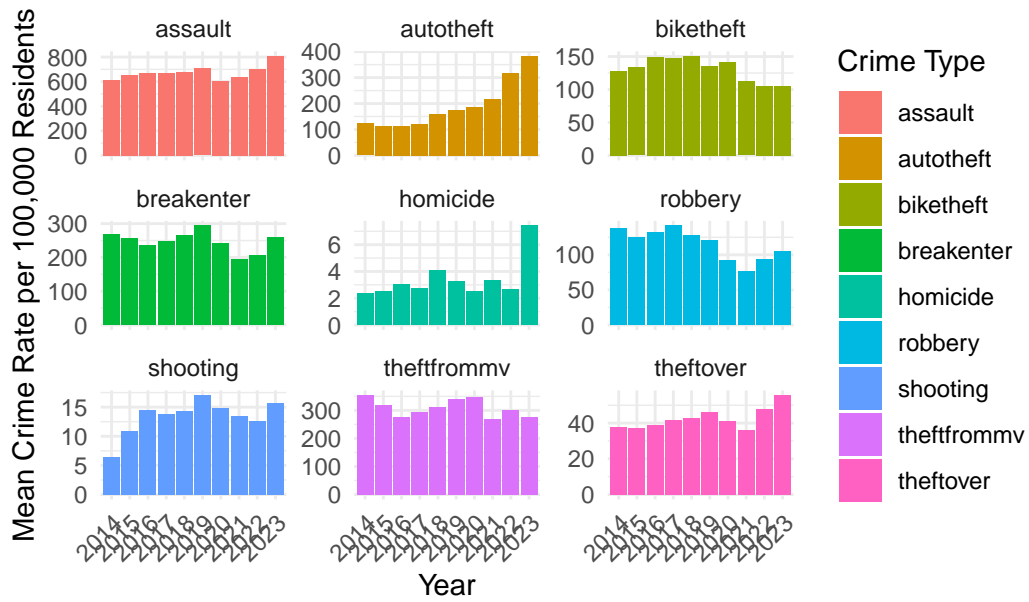
Figure 1: Figure 1: Total Crime Rates from 2014-2023

Table 1: Figure 2: Summary Statistics of Crime Rates (2014-2023)

Mean Crime Rates by Year (2014–2023) for Each Crime Type

3 Results

3.1 Average Crime Rate Comparison Between High and Low-Priced Neighborhoods over 2014 to 2023 for Individual Crimes
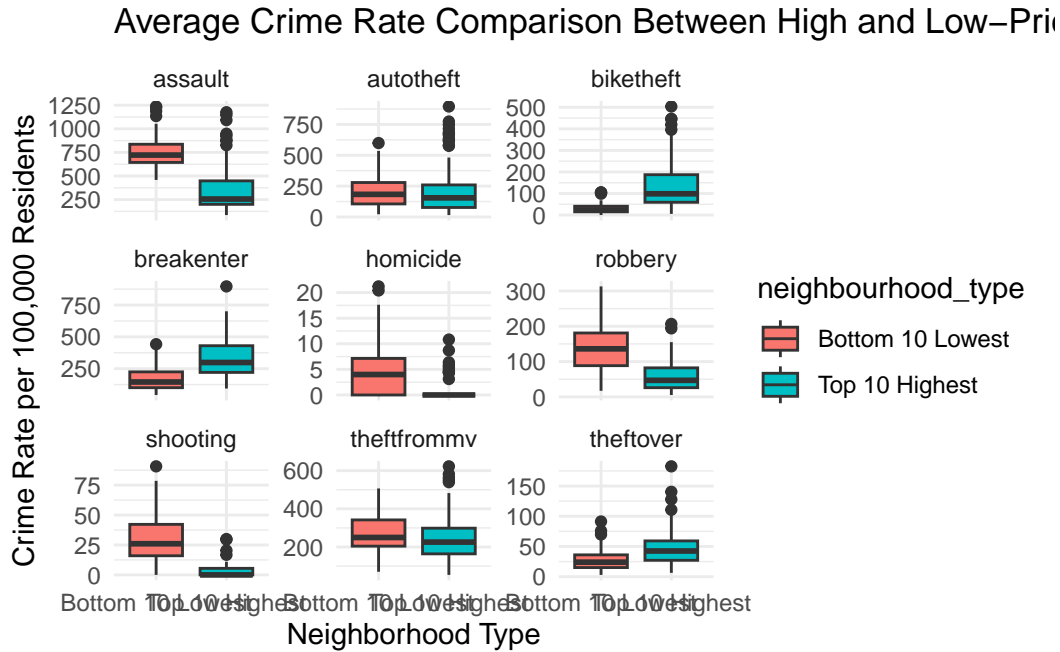

Figure 2: Figure 3: Box Plots comparing Crime Rates in Top 10 vs Bottom 10 Home Price Neighborhoods

Box Plots Comparing Crime Rates for High vs. Low Price Neighborhoods: This plot shows a comparative analysis of crime rates between the Top 10 Highest and Bottom 10 Lowest home price neighborhoods in Toronto from 2014 to 2023. Each box plot represents the distribution of crime rates for different types of crime (e.g., assault, autotheft, biketheft) in these neighborhoods, with the y-axis representing the crime rate per 100,000 residents.

Key Observations: Assault Rates: Assault rates are notably higher in low-priced neighborhoods than in high-priced ones. The median assault rate in the bottom 10 neighborhoods is around 500, while in the top 10 neighborhoods, it is closer to 250. The top neighborhoods also show less variability in assault rates. Auto Theft: Auto theft rates are slightly higher in bottom-priced neighborhoods, but the difference between the two groups is not as pronounced as for assault. Both groups exhibit a broad distribution, with several outliers in low-priced neighborhoods indicating significantly higher rates in certain years. Bike Theft: High-priced neighborhoods have a notably higher rate of bike theft compared to low-priced ones. The

median bike theft rate in top neighborhoods is significantly larger, suggesting that wealthier neighborhoods may experience more of these types of non-violent crimes, possibly due to the greater number of bikes or higher-value targets. Break and Enter: The break and enter crime rate appears to be higher in low-priced neighborhoods, with a median crime rate around 400 incidents per 100,000 residents, compared to around 250 in high-priced neighborhoods. The spread of crime rates is more pronounced in low-priced areas. Homicide: Homicide rates in low-priced neighborhoods are significantly higher than in high-priced ones. While most of the values cluster around low rates (under 10 per 100,000 residents), there are some extreme outliers in low-priced neighborhoods. Robbery: Robbery rates are consistently higher in low-priced neighborhoods compared to high-priced ones. The difference in medians is noticeable, with the bottom 10 neighborhoods showing a broader range of robbery rates. Shooting: Shootings are more common in low-priced neighborhoods, with significantly higher variability and outliers. In contrast, high-priced neighborhoods show a relatively small number of shootings, with rates clustering below 25 incidents per 100,000 residents. Theft from Motor Vehicles: Theft from motor vehicles (theftfrommv) shows similar crime rates between both neighborhood groups. However, there is a slightly higher median and more variability in bottom-priced neighborhoods. Theft Over \$5,000: Both groups exhibit similar trends in "theft over \$5,000" rates. There is not a significant difference between the two groups, and both have fairly consistent rates over the years. Insights: Higher Crime in Low-Priced Neighborhoods: Violent crimes like assault, break and enter, homicide, robbery, and shooting tend to be more prevalent in neighborhoods with lower home prices. These areas also exhibit greater variability in crime rates, possibly reflecting more social instability or economic distress. Higher Non-Violent Crimes in High-Priced Neighborhoods: Interestingly, non-violent crimes like bike theft are more common in higher-priced neighborhoods, possibly reflecting the availability of more valuable goods in these areas. Socioeconomic Disparities in Crime: This analysis confirms the hypothesis that wealthier neighborhoods experience lower rates of violent crimes, while economically disadvantaged areas face greater exposure to violent crime. However, non-violent crimes (e.g., thefts) seem more prevalent in wealthier neighborhoods, suggesting different socio-economic dynamics at play.

This helps visualize how crime rates differ between the top 10 highest and lowest-priced neighborhoods.

3.2 Average Crime Rate Comparison Between High and Low-Priced Neighborhoods from 2014 to 2023 for Overall Crime
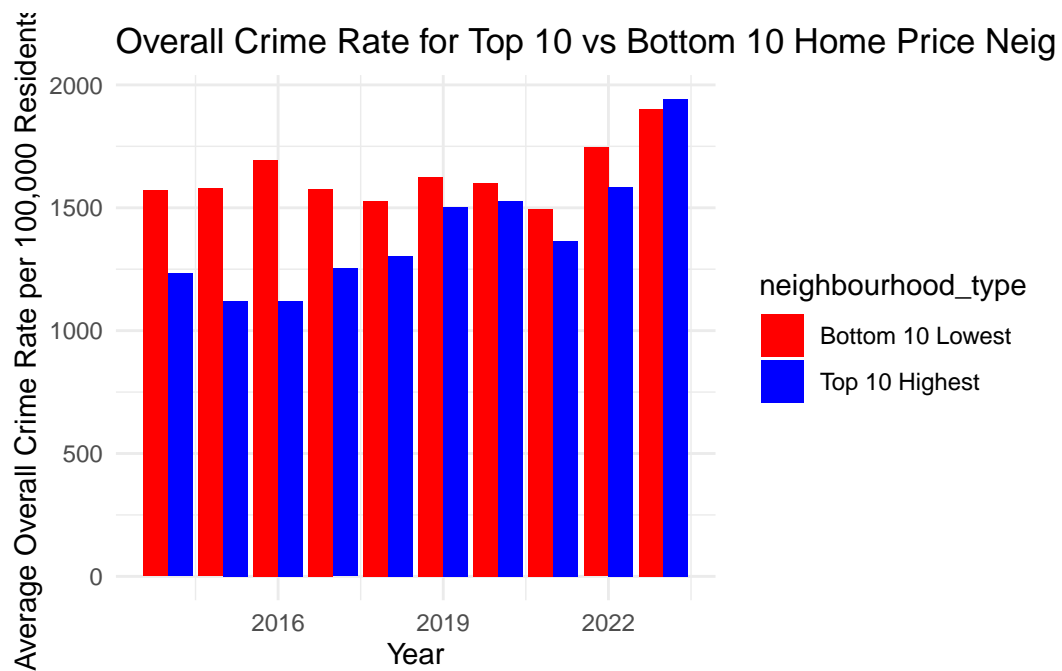

Figure 3: Figure 4: Crime Rate Time Series from 2014 to 2023 for Top 10 vs Bottom 10 Housing Price Neighbourhoods

3.3 Average Crime Rate Comparison Between High and Low-Priced Neighborhoods from 2014 to 2023 for Individual Crimes
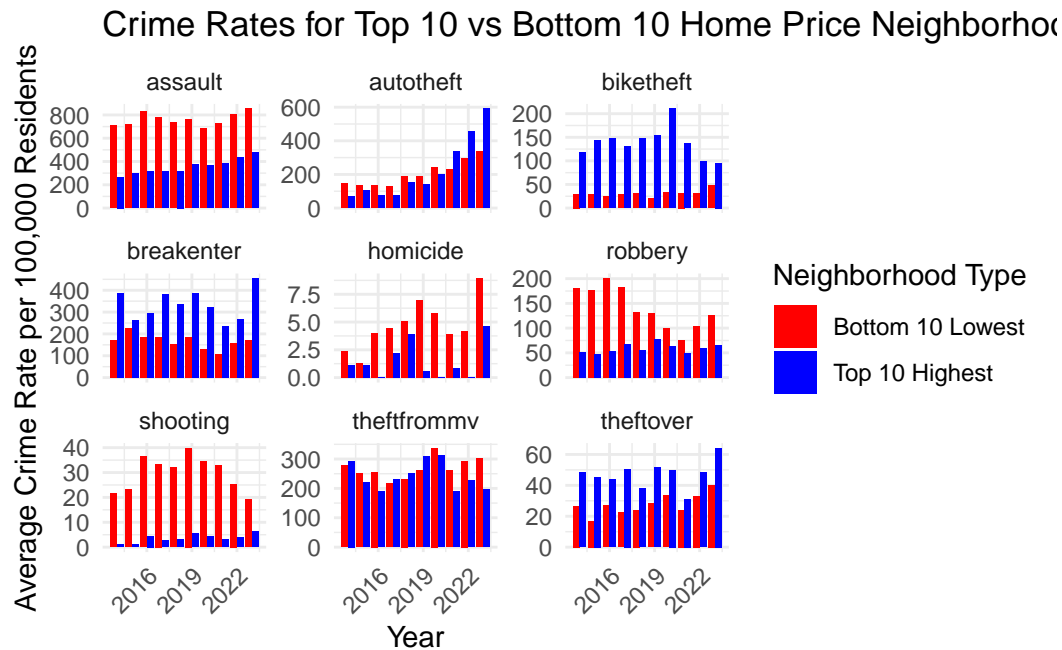

Figure 4: Figure 5: Crime Rate Time Series from 2014 to 2023 for Top 10 vs Bottom 10 Housing Price Neighbourhoods

Crime-Specific Analysis: Assault: The assault rates remain consistently higher in Bottom 10 Lowest neighborhoods than in Top 10 Highest neighborhoods across all years. There seems to be a slight downward trend in the top neighborhoods around 2019–2021, while rates in lower neighborhoods remain relatively stable. Autotheft: Bottom 10 Lowest neighborhoods show a steady increase in autotheft rates from 2015 onward, peaking around 2022. Top 10 Highest neighborhoods have significantly lower autotheft rates throughout the period, but they also show a slight upward trend, though the increase is far more gradual. Biketheft: In the earlier years (2015–2017), Bottom 10 Lowest neighborhoods have higher bike theft rates, but Top 10 Highest neighborhoods catch up around 2018–2020. From 2020 onwards, the bike theft rates in Top 10 Highest neighborhoods are actually higher than those in the Bottom 10 Lowest neighborhoods. Breakenter (Breaking and Entering): Rates for break and enter crimes fluctuate more in Bottom 10 Lowest neighborhoods, peaking around 2016–2017 and then decreasing slightly. Top 10 Highest neighborhoods show more stable but much lower break and enter rates overall. Homicide: Both neighborhoods show very low homicide rates (as expected), but Bottom 10 Lowest neighborhoods exhibit slightly higher rates. There is a gradual upward trend in homicides in Bottom 10 Lowest neighborhoods from 2016 to 2023,

although the scale is relatively small (0 to 7.5 per 100,000 residents). Robbery: Robbery rates in Bottom 10 Lowest neighborhoods are consistently higher than in Top 10 Highest neighborhoods. The gap between the two neighborhood types widens in 2020 and onwards, with a notable increase in robbery rates in Bottom 10 Lowest neighborhoods. Shooting: Bottom 10 Lowest neighborhoods experience a significant peak in shooting rates around 2018, after which they decline. Top 10 Highest neighborhoods have almost negligible shooting rates in comparison. Theftfrommv (Theft from Motor Vehicle): The theft from motor vehicle rates are much higher in Bottom 10 Lowest neighborhoods, and they fluctuate over the years, with no consistent trend. The rates for Top 10 Highest neighborhoods remain fairly stable and low across the entire time period. Theftover (Theft over $5,000): Theft over $5,000 shows stable but relatively low rates for both neighborhood types, though Bottom 10 Lowest neighborhoods have a slight edge. There's an increasing trend in Bottom 10 Lowest neighborhoods starting in 2020.

4 Results

Our results are summarized in **?@tbl-modelresults**.

5 Discussion

5.1 First discussion point

If my paper were 10 pages, then should be be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

5.2 Second discussion point

5.3 Third discussion point

5.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

Appendix

A Additional data details

B Model details

B.1 Posterior predictive check

In **?@fig-ppcheckandposteriorvsprior-1** we implement a posterior predictive check. This shows...

In **?@fig-ppcheckandposteriorvsprior-2** we compare the posterior with the prior. This shows...

Examining how the model fits, and is affected by, the data

B.2 Diagnostics

?@fig-stanareyouokay-1 is a trace plot. It shows... This suggests...

?@fig-stanareyouokay-2 is a Rhat plot. It shows... This suggests...

References

R Core Team. 2023. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.