

Socio-Economic Disparities and Law Enforcement in Toronto: An Analysis of Arrest Trends and Median Household Income Across City Wards

David James Dimalanta

January 24, 2024

Code and data from this analysis are available at: https://github.com/davidjamesdimalanta/toronto_arrests_analysis

“This study investigates the correlation between socio-economic status and law enforcement activities in Toronto by analyzing arrest count data and median household income across the city’s wards. Utilizing data from the Toronto Police Service and the 2021 Toronto Census, the research reveals significant disparities in arrest counts between wards categorized as ‘low-income household neighbourhoods’ and ‘high-income household neighbourhoods.’ Notably, it was found that some low-income areas, such as Ward 20 (Scarborough-Agincourt), experience higher arrest counts, while certain high-income wards like Ward 10 (Spadina-Fort York) also report substantial arrest figures, illustrating that socio-economic status is not a singular predictor of law enforcement activity. The findings of this study are crucial for urban policymakers and planners, as they underscore the complex interplay between socio-economic factors and crime within urban settings, highlighting the need for nuanced approaches to law enforcement and social policy in diverse urban communities.”

Table of contents

1	Introduction	2
2	Data	3
2.1	Arrest Count Data	3

2.2	Census Data	4
2.3	Summary Statistics and Relationships	4
3	Results	5
3.1	Arrest Count by Ward	5
3.2	Median Household Income by Ward	6
4	Discussion	7
4.1	Socio-Economic Disparities and Law Enforcement	7
4.2	Arrest Patterns in High-Income Neighbourhoods	7
4.3	Median Income and Its Implications	8
4.4	Limitations and Further Research	8
5	Conclusion	8

1 Introduction

This paper presents a detailed analysis of arrest counts and median household income across the various wards of Toronto, drawing on data from the ‘Police Annual Statistical Report - Arrested and Charged Persons’ and the ‘Ward Profiles (25-Ward Model)’. The broader context of this study is to understand how socio-economic factors, particularly income levels, correlate with law enforcement activities within an urban landscape. This research addresses a critical gap in understanding the dynamics of crime and socio-economic status within a city, offering insights into the intricate relationship between median household income and arrest frequencies across different wards.

The methodology involved collecting and analyzing data using R programming software and several R libraries, including tidyverse, ggplot2, dplyr, and others. The arrest data was meticulously cleaned and reorganized into a ward-based structure to enable effective cross-referencing with the Toronto Census Data, thereby enhancing the potential for nuanced analyses. One of the key findings of this study is the notable variance in arrest counts between “low-income household neighbourhoods” and “high-income household neighbourhoods.” For instance, Ward 20 (Scarborough-Agincourt), classified as a low-income area, exhibited the highest arrest count, while high-income areas like Ward 10 (Spadina-Fort York) also reported substantial arrest figures.

The importance of this study lies in its potential to inform policy-making and urban planning by highlighting the socio-economic disparities that exist within the city and their association with law enforcement. This paper is structured to first present the data and results, followed by a detailed discussion of the findings, their implications, and the limitations of the study. The conclusion synthesizes these insights, emphasizing the need for further research and a holistic approach to addressing the challenges faced by diverse neighbourhoods in Toronto.

2 Data

This study employs an in-depth analysis of the socio-economic factors influencing law enforcement activities in Toronto, utilizing two primary datasets: the ‘Police Annual Statistical Report - Arrested and Charged Persons’ from the OpenDataToronto Library (Gelfand 2022) and the ‘Ward Profiles (25-Ward Model)’ containing 2021 Toronto census data. The statistical programming software R (R Core Team 2023) was used to collect and analyze the data, utilizing R libraries such as `tidyverse` (Wickham 2023b), `ggplot2` (Wickham, Chang, et al. 2023), `dplyr` (Wickham, François, et al. 2023), `readr` (Wickham, Hester, and Bryan 2024), `knitr` (Xie 2023), `stringr` (Wickham 2023a), `tibble` (Müller and Wickham 2023), `janitor` (Firke 2023), and `testthat` (Wickham 2011).

2.1 Arrest Count Data

The arrest count dataset, provided by the Toronto Police Services, offers a comprehensive view of law enforcement activities within the city. This data is aggregated by multiple dimensions, including division, neighbourhood, sex, age, crime category, and crime subtype. The latest refresh of this dataset was on November 28, 2023, and the data used in this study was captured on January 16, 2024.

The dataset was cleaned and restructured into a ward-based format, aligning it with the Toronto Census Data’s structure. This transformation facilitates a more effective comparison and cross-referencing between socio-economic factors and law enforcement data. The cleaning and restructuring process involved advanced functions from R libraries like `dplyr` and `stringr`, adhering to the spatial classifications of the City of Toronto’s ‘Neighbourhoods & Communities’ (2017) map.

A sample of the cleaned arrest data, as shown in Table 1, highlights the initial dataset’s structure:

Table 1: Table 1: Sample of Cleaned Arrest Count Data

x_id	crime_neighbourhood	arrest_count
1	Dufferin Grove (83)	1
2	Brookhaven-Amesbury (30)	2
3	Harbourfront-CityPlace (165)	1
4	New Toronto (18)	3
5	Kensington-Chinatown (78)	46

2.2 Census Data

The census data, sourced from OpenDataToronto’s ‘Ward Profiles (25-Ward Model)’, details the 2021 Toronto census data by ward. This dataset was chosen for its comprehensiveness and alignment with the city’s current ward structure. Alternative datasets, such as previous census data, were not used due to the relevance and recency of the 2021 census, providing the most up-to-date socio-economic insights.

The census data underwent a significant cleaning process, drawing methodologies from a colleague’s repository (Fox 2024). This process included data transformation techniques, utilizing R’s `janitor` and `tidyverse` libraries, to ensure the data’s usability and integrity.

A snippet of the census data, represented in Table 2, showcases the median income across different wards:

Table 2: Table 2: Sample of Census Data by Ward

ward	med_income
Toronto	84000
1	81000
2	100000
3	90000
4	85000
5	72000

2.3 Summary Statistics and Relationships

In this study, summary statistics of both datasets are examined to understand the distribution and relationships between variables. Bar charts are used to visualize the arrest counts across wards and to compare these with median household incomes. These visual aids, along with summary statistics, provide a comprehensive overview of the data, revealing patterns and correlations that are pivotal to the study’s findings.

The relationship between socio-economic status, as indicated by median household income, and arrest counts is a focal point of the analysis. This relationship is explored to understand the impact of economic factors on law enforcement activities and to identify any significant trends or anomalies within the data.

In conclusion, the datasets used in this study offer a rich source of information for examining the interplay between socio-economic status and law enforcement in Toronto. The careful selection, cleaning, and analysis of these datasets provide a solid foundation for the insights and conclusions drawn in this research.

3 Results

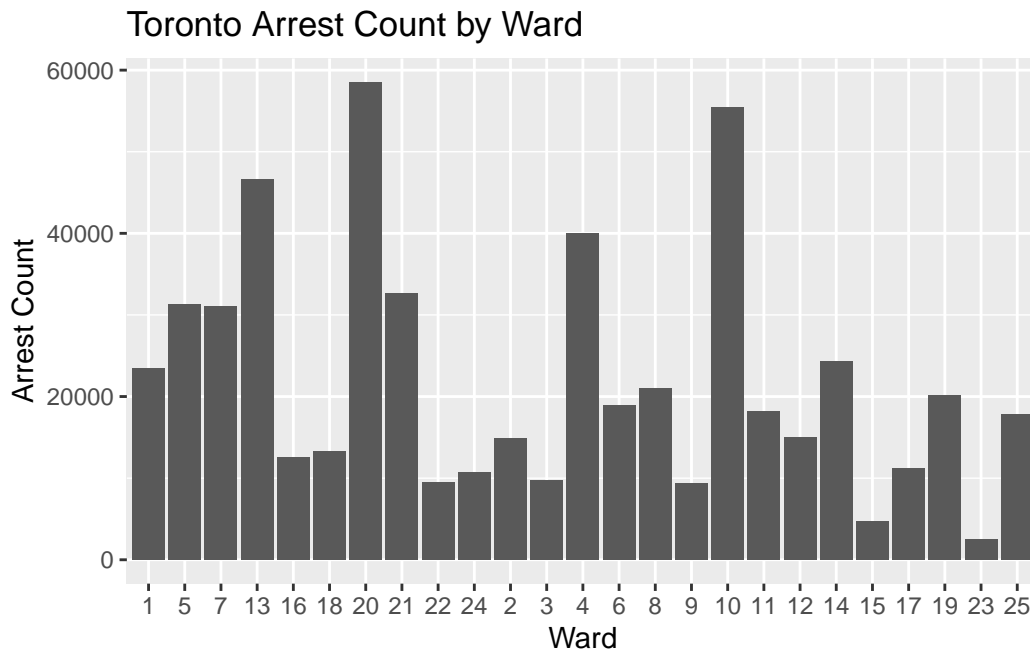
3.1 Arrest Count by Ward

In the “low-income household neighbourhoods” category, which includes Wards 1, 5, 7, 13, 16, 18, 20, 21, 22, and 24, there is a notable variance in arrest counts. Ward 20 (Scarborough-Agincourt) exhibits the highest arrest count in this category, with 58,505 arrests. Following closely is Ward 13 (Toronto Centre) with 46,599 arrests. In contrast, Ward 22 (Willowdale) shows a considerably lower count at 9,500 arrests.

Among the “high-income household neighbourhoods,” Ward 10 (Spadina-Fort York) stands out with the highest arrest count overall, at 55,391. This figure is particularly noteworthy given the contrasting socioeconomic backdrop of this ward. Other high-income wards like Ward 4 (Etobicoke North) and Ward 25 (Scarborough-Rouge Park) also report substantial arrest counts, with 39,972 and 17,798 arrests respectively.

The data also reveals that the lowest arrest count across all wards is in Ward 23 (Scarborough North), a high-income neighbourhood, with only 2,487 arrests.

To visually represent these disparities in arrest counts across the wards, the following graph has been created using the `ggplot2` package in R. This graph provides a clear and comprehensive illustration of the arrest patterns in Toronto’s wards:



3.2 Median Household Income by Ward

The citywide median income for Toronto stands at \$84,000, serving as a comparative baseline for the median incomes in individual wards. The citywide median income for Toronto stands at \$84,000, serving as a comparative baseline for the median incomes in individual wards.

Conversely, Ward 5 and Ward 7 are among the wards with the lowest median incomes, at \$72,000 and \$73,000 respectively. These figures suggest these wards are comparatively lower-income areas within the city.

Other wards like Ward 2 and Ward 11 also report relatively high median incomes, \$100,000 and \$84,000 respectively, indicating a prevalence of upper-middle-class households in these areas. Wards such as 13 and 24, with median incomes of \$65,000 and \$78,000 respectively, represent regions with more moderate median income levels.

The following table displays a subset of the census data, offering a snapshot of the median household income by ward. This table, built using the `knitr` package in R, aids in understanding the economic backdrop against which the arrest counts should be considered:

Table 3: Census Data by Ward

ward	med_income
Toronto	84000
1	81000
2	100000
3	90000
4	85000
5	72000
6	82000
7	73000
8	97000
9	85000
10	89000
11	84000
12	86000
13	65000
14	93000
15	102000
16	78500
17	84000
18	81000
19	89000
20	79000
21	78000

ward	med_income
22	77000
23	87000
24	78000
25	105000

4 Discussion

The analysis of Toronto’s census data juxtaposed with arrest counts across various wards presents a multifaceted understanding of the city’s socio-economic landscape and its correlation with law enforcement activities. This discussion aims to explore the underlying patterns and potential implications revealed by the results.

4.1 Socio-Economic Disparities and Law Enforcement

One of the most striking observations is the relationship between median household income levels and arrest counts in different wards. The data indicates that “low-income household neighbourhoods,” particularly Ward 20 (Scarborough-Agincourt) and Ward 13 (Toronto Centre), experience significantly higher arrest counts compared to some “high-income household neighbourhoods.” This disparity could be attributed to a range of factors, including but not limited to, socio-economic challenges, varying crime rates, or differences in policing strategies.

Ward 20, despite being classified as a low-income area, has the highest arrest count in the city. This could indicate a higher prevalence of factors leading to criminal activity or possibly more rigorous law enforcement efforts in these neighbourhoods. Conversely, Ward 22 (Willowdale), another low-income area, reports significantly fewer arrests, suggesting that low income does not uniformly correlate with high arrest counts.

4.2 Arrest Patterns in High-Income Neighbourhoods

In high-income neighbourhoods, the pattern of arrests presents a complex picture. For instance, Ward 10 (Spadina-Fort York) records a high number of arrests, second only to Ward 20, despite its high-income status. This could point towards a nuanced understanding of crime and law enforcement that goes beyond simplistic associations with economic conditions. It might also reflect a diversity of sub-communities within these wards, each with its unique socio-economic dynamics.

4.3 Median Income and Its Implications

The analysis of median household incomes across wards reveals significant disparities, with Ward 5 and Ward 7 at the lower end of the spectrum, and Wards 2 and 11 showing relatively higher median incomes. These disparities highlight the economic diversity within Toronto and underscore the need for tailored social and economic policies that address the specific needs of each ward.

4.4 Limitations and Further Research

While this analysis sheds light on certain patterns, it is important to note the limitations inherent in such a study. The data does not account for the complexities and nuances of individual neighbourhoods or the myriad factors that influence both crime rates and law enforcement activities. Further research, perhaps incorporating qualitative studies and more granular data, could provide a deeper understanding of the intricate relationship between socio-economic factors and law enforcement in urban settings.

5 Conclusion

In synthesizing the findings of this study, it becomes evident that Toronto's socio-economic landscape profoundly influences its law enforcement patterns. The detailed analysis of arrest counts and median household income across the city's wards reveals a complex interplay between socio-economic status and crime. Notably, the higher arrest counts in certain low-income neighbourhoods, such as Ward 20 (Scarborough-Agincourt), contrast sharply with the patterns observed in some high-income areas like Ward 10 (Spadina-Fort York). This disparity highlights the multifaceted nature of urban crime and its relation to economic factors.

The study underscores the necessity of considering socio-economic contexts in policy-making and urban planning, especially in addressing law enforcement and community welfare. It challenges simplistic notions about crime and income, suggesting that a ward's economic status does not singularly dictate its crime rate. Instead, a myriad of factors contributes to these patterns, necessitating a more nuanced approach in understanding and addressing urban challenges.

Furthermore, the variations in median household income across Toronto's wards paint a picture of a city with diverse economic realities. This diversity calls for targeted strategies that cater to the unique needs of each community, moving beyond one-size-fits-all solutions in urban governance.

While this research provides significant insights, it also acknowledges its limitations, particularly the need for more in-depth, qualitative analyses to fully capture the complexities of each neighbourhood. Future studies could delve deeper into the socio-economic fabric of Toronto's

wards, perhaps incorporating longitudinal data and qualitative narratives to enrich the understanding of urban dynamics.

Ultimately, this paper contributes to a broader discourse on urban socio-economic patterns and their impact on law enforcement, serving as a foundation for future research and informed decision-making. It highlights the importance of an integrated approach that combines socio-economic understanding with law enforcement strategies to create safer, more equitable urban communities.

2017. *City of Toronto*. City of Toronto. <https://www.toronto.ca/city-government/data-research-maps/neighbourhoods-communities/>.
- Firke, Sam. 2023. “Simple Tools for Examining and Cleaning Dirty Data [r Package Janitor Version 2.2.0].” *R-Project.org*, February. <https://doi.org/https://cran.r-project.org/package=janitor>.
- Fox, Thomas. 2024. “Toronto_child_care.” *GitHub*. https://github.com/ThomasWilliamFox/toronto_child_care/tree/master.
- Gelfand, Sharla. 2022. *Opendatatoronto: Access the City of Toronto Open Data Portal*. <https://sharlagelfand.github.io/opendatatoronto/>.
- Müller, Kirill, and Hadley Wickham. 2023. *Tibble: Simple Data Frames*. <https://tibble.tidyverse.org/>.
- R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Wickham, Hadley. 2011. “Testthat: Get Started with Testing.” *The R Journal* 3: 5–10. https://journal.r-project.org/archive/2011-1/RJournal_2011-1_Wickham.pdf.
- . 2023a. *Stringr: Simple, Consistent Wrappers for Common String Operations*. <https://stringr.tidyverse.org>.
- . 2023b. *Tidyverse: Easily Install and Load the Tidyverse*. <https://tidyverse.tidyverse.org>.
- Wickham, Hadley, Winston Chang, Lionel Henry, Thomas Lin Pedersen, Kohske Takahashi, Claus Wilke, Kara Woo, Hiroaki Yutani, and Dewey Dunnington. 2023. *Ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics*. <https://ggplot2.tidyverse.org>.
- Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. <https://dplyr.tidyverse.org>.
- Wickham, Hadley, Jim Hester, and Jennifer Bryan. 2024. *Readr: Read Rectangular Text Data*. <https://readr.tidyverse.org>.
- Xie, Yihui. 2023. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. <https://yihui.org/knitr/>.