Bike thefts in Toronto: Diffrences by neighbourhood type

Michael Shmelev

Table of contents

Crime has always been a topic in the public conscience in Toronto. To find out how crime rates differ between suburban areas and cities we compared the two. To do this we have taken 10 geographic areas from open data Toronto's collected data on various crime statistics in the 2014-2023 time span. Five of these are in the downtown core, while the other five are located in the Beaches and Danforth suburban area east of the downtown core. The findings from organizing data and comparing the two region types supports the hypothesis that more urbanized areas have a higher level of crime in them compared to suburban areas.

Introduction

In Toronto crime has always been a daily factor of life, as in any other large city. This paper is about examining the occurrence of crime in urban and suburban areas in Toronto, and determining if the hypothesis of there being more crime in urban cities over suburban ones is true. The data analyzed specifically examined bike theft, to the exclusion of other crimes recorded by residents living in Toronto. The findings were that there is more bike theft in downtown Toronto when compared to the suburbs east of the downtown core. By

focusing on bike thefts a clear gap is being filled as this is a crime that is often overlooked and underestimated on account of comparison to more high profile crimes such as automobile theft. By researching and evaluating bike theft crime, the intention is to raise awareness to bike theft prominence as an important issue. Bike usage is crucial to the life of a city, as bikes fill a gap left by cars not being affordable to those with a lower income. The paper follows the Introduction, data, results, discussion, and conclusion structure.

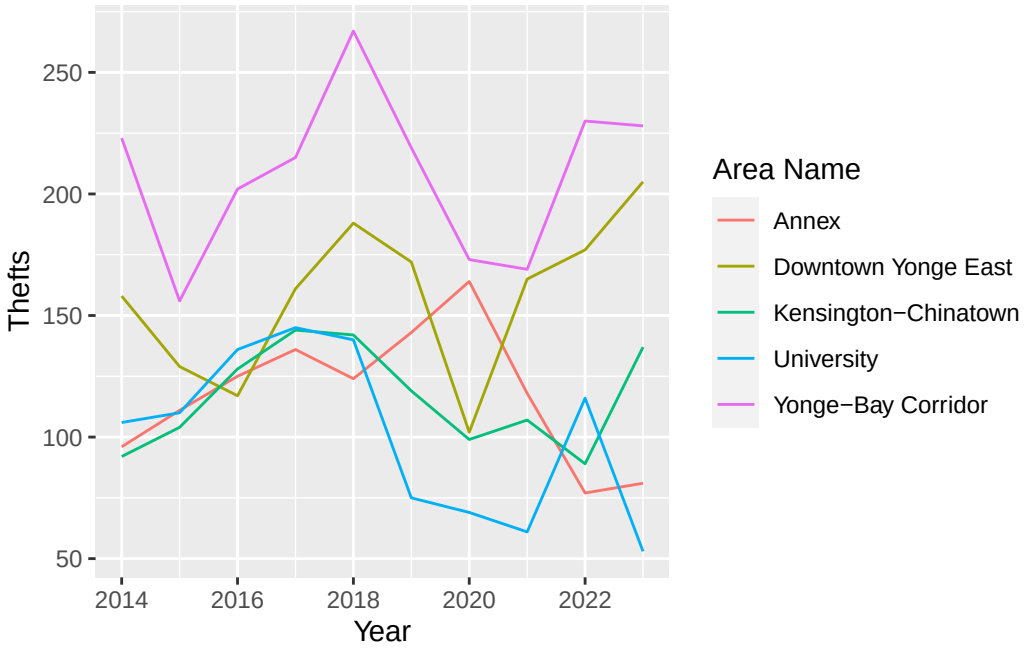
Data

The data examined in the paper is bike theft between the urban Toronto core, and the eastern suburbs in Toronto. The analysis focuses on the Yonge-Bay Corridor, Annex, University, Kensington-Chinatown, and Downtown Yonge East for the five downtown core neighborhoods. The east Toronto suburbs examined were Danforth, The Beaches, East End-Danforth, Danforth East York, and the Woodbine Corridor. These areas were located in the open data Toronto file, and identified on google maps in order to ensure that they were all close to proximity with one another and fit the suburbs or urban center classifications.

The dataset that was used in this paper was the 2014-2023 crime statistics for Toronto. The variables that were considered include the area, the number of thefts, and the year. This selection of variables provided a sound criteria for comparing and contrasting the two area types.

##move to results What really stood out was the sheer numbers difference when comparing the amount of stolen bikes between the urban core, and the suburban east. The downtown areas had significantly higher theft rates than the suburban area. The highest rate of theft for the suburban area was 89 compared to the highest urban bike theft number over 250. Another feature is that bike thefts during the covid lockdowns in Toronto, which happened in 2020, impacted theft in different ways. Theft increased in the suburban areas through 2020, which is contrasted by the sharp decrease of theft in the urban core during the 2020 year. It can be speculated that this decrease in the urban core was due to stricter enforcement of lockdown policies, but proper research should be conducted before making any claims.

Urban Area Theft by area, by yearly theft total value, and by area



The reason for choosing open data Toronto was because it seemed like the most credible collection of data available. Open data Toronto is a source that is maintained by the Toronto government, meaning that it is the most reliable source and up to date resource available. The advantage of using open data Toronto is that it is reputable, and collects data frequently.

Technical R studio considerations

The foundation for simulating data related to bike thefts in different neighborhoods over a specified time period is the inclusion of test code. The simulated data, stored in a tibble named simulated_theft_data, includes columns for the neighborhood, year, and the number of thefts. The neighborhoods A, B, and C each have 10 years of simulated data. The code then prints the first few rows of the generated data. Subsequent checks verify that the unique neighborhoods are "A," "B," and "C," the number of unique neighborhoods is three, the years range from 2014 to 2023, the minimum number of thefts is at least 10, the maximum number of thefts is no more than 100, and the data type of the thefts column is numeric. These checks ensure that the simulated data aligns with the specified criteria.

Additionally In order to make working with the graph and csv file more manageable, the csv file had to be modified to remove the geometry data. Before these changes were made, the table that was produced while running the code before graphing had a lot of negative values, and coordinates that made graphing difficult to do. Plotting and visualization functions did

not handle spatial data well because they were not be designed for geographic representations. Removing the "geometry" variable makes a simpler non-spatial dataset that is compatible with a broader range of visualization tools.

Results

The key findings were that theft rates in urban Toronto were going up in almost all areas until 2019, where they made a significant drop. However these rates of bike theft are once more growing to pre 2020 level with the exception of University. These findings differ from the findings for Toronto's eastern suburbs. In the suburban Toronto areas examined, the theft rate increased and was consistently growing before and during the covid-19 lockdowns in 2020.

Additionally, the amount of bike theft in the urban city eclipses the rate of bike theft at its peak in the suburban areas.

If these results are reliable, the research suggests that bike theft in urban areas is climbing back to pre 2019 rates, whereas suburban crime rates are declining after incremental growth and a surge in 2020 and 2021.

Discussion

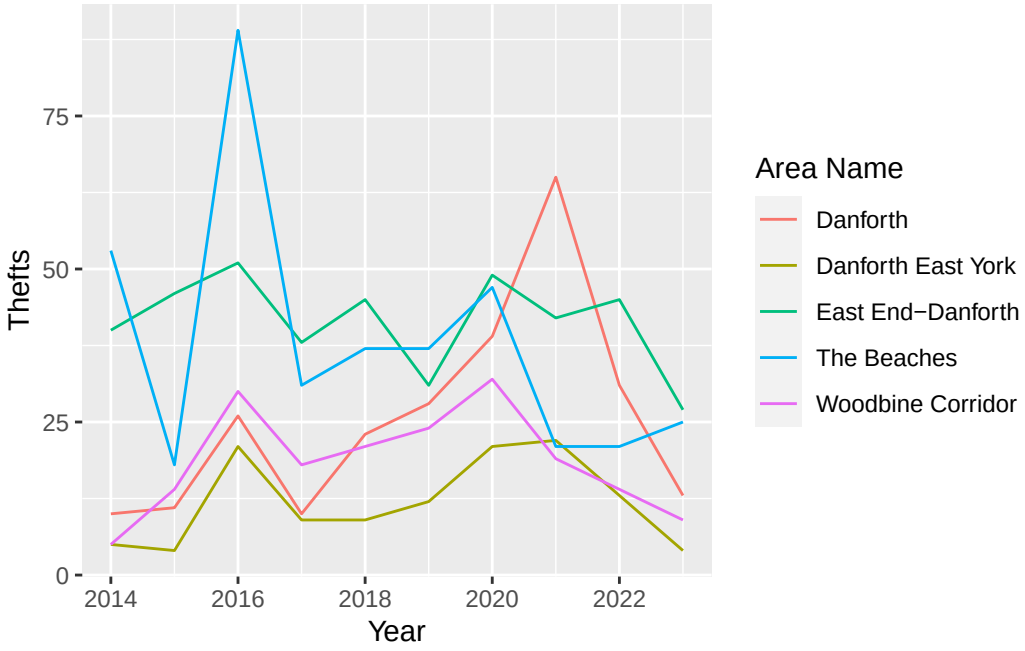
While answering the initial question of if suburban areas or urban areas experience more crime, some interesting findings were made along the way. While the graphs and data dictate that theft is higher in urban areas, further research needs to be done as to why numbers decreased in 2019 - 2021. The most obvious explanation is Covid-19 lockdowns and mandates affected the amount of bikes available in the city to be stolen, but this needs to be further explored with more research. Other considerations are that transit improvements affected the number of bike thefts by getting people to use transit over cycling.

One very important concern is the reliability of open data Toronto when it comes to the 2023 data. The data in the 2023 theft for suburban areas is substantially lower than the 2022 and 2021 data. This raises the question of validity regarding the 2023 sample. The concern is that the 2023 data might have been recorded part way through the year, instead of in December 2023 or January 2024. If this is true, assumptions will need to be reconsidered regarding the current data and conclusions.

Conclusion

In conclusion it is clear that bike related theft crime is more prevalent in urban centers than in suburban areas. This suggests but does not in any way prove that overall crime numbers

would be higher in urban centers compared to suburban areas. It is also clear that further research and a broader scope of data would benefit the process of discovery and knowledge aquisition.



Code and repo can be found following (link goes here)

(R Core Team 2021)

(Alexander 2023)

Alexander, Rohan. 2023. *Telling Stories with Data*. Chapman; Hall/CRC. https://tellingstorieswithdata.com.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.