Trends and Patterns in Bicycle Thefts: A Temporal Analysis of Incident Data*

Zitong Guo

September 25, 2024

This paper investigates patterns of bicycle thefts and recoveries using incident report data from multiple years. By analyzing temporal trends, offense types, and bicycle characteristics, significant fluctuations in theft rates were observed over time, with certain months experiencing higher incidents. Additionally, certain types of bicycles, such as mountain and road bikes, were more frequently stolen, while higher-cost bicycles had a slightly higher recovery rate. Understanding these trends is crucial for improving theft prevention strategies and informing both policy and law enforcement efforts aimed at reducing bike-related crimes.

Table of contents

*A GitHub Repository containing all data, R code, and other files used in this investigation is located here: https://github.com/jennygzt/Bicycle-Theft-.git

1 Introduction

Bicycle theft remains a significant issue in urban environments, causing financial strain, personal frustration, and discouraging the use of bicycles as a mode of sustainable transportation. Despite improvements in security measures and heightened public awareness, bicycles continue to be vulnerable to theft. Understanding the factors influencing theft rates such as the timing, location, and types of bicycles stolen can provide crucial insights into preventing such crimes. Furthermore, the low recovery rates of stolen bicycles raise critical questions about how characteristics like bike cost and type impact the likelihood of recovery. Addressing these factors is essential for developing more effective crime prevention strategies and improving recovery efforts.

This paper seeks to bridge the gap in understanding temporal and spatial patterns of bicycle theft and recovery by analyzing detailed incident data over several years. We investigate theft trends over time, focusing on variations by month and specific bicycle types. Additionally, exploring whether the cost of the bicycle influences recovery rates and whether certain bike characteristics affect the likelihood of recovery.

The findings in this paper reveal notable trends, including a seasonal peak in thefts during the summer months, with mountain and road bikes being the most frequently targeted. It is also observed that higher-cost bicycles show a slightly increased recovery rate, but overall recovery rates remain low. These insights are critical in addressing the challenges of bicycle theft and formulating evidence-based recommendations.

The remainder of this paper is structured as follows. In Section 2, the paper focuses on describing the dataset and outlining the methodology for analyzing theft and recovery patterns. In Section 3, key observations and findings are analyzed, including temporal trends, bicycle-specific data, and recovery outcomes. Finally, in Section 4, further examination on the implications of these findings and providing recommendations for policymakers, law enforcement, and bicycle owners to help mitigate bicycle theft and enhance recovery efforts.

2 Data

2.1 Raw Data and Cleaning

The data used in this analysis is sourced from publicly available datasets on reported bicycle thefts in Toronto, obtained from the City of Toronto's Open Data Portal (Gelfand 2022). The datasets include detailed incident reports of bicycle thefts across various neighbourhoods, spanning multiple years. The key variables in the dataset include the date and the time of theft, the location of the incident, the type of bike stolen, and the recovery status when available.

The `R` programming language (R Core Team 2023) and `tidyverse` (Wickham et al. 2019) packages were used to simulate the dataset and generate tests for it. The `opendatatoronto` (Gelfand 2022) and `tidyverse` (Wickham et al. 2019) packages were then applied to download the raw Toronto Bicycle theft dataset. Next, the `tidyverse` package (Wickham et al. 2019), janitor (Firke 2024), ggplot2 (Wickham 2016), here (Müller 2020), and knitr (Xie 2024) were used to clean the raw dataset and test the raw as well as the cleaned dataset. The data cleaning process involved filtering specific columns from the raw dataset and renaming certain entries for clarity and simplicity. In this analysis, the missing values are not removed as they are crucial for ensuring the integrity and preserving sample size.

Table 1: Sample of raw data

Occurrence Date	Location Type	Bike Cost	Status
2024-06-30	Apartment (Rooming House, Condo)	500	STOLEN
2024-06-30	Apartment (Rooming House, Condo)	950	STOLEN
2024-06-30	Bar / Restaurant	4000	STOLEN
2024-06-29	Ttc Subway Station	2000	STOLEN
2024-06-30	Streets, Roads, Highways (Bicycle Path, Private Road)	400	STOLEN
2024-06-29	Apartment (Rooming House, Condo)	550	STOLEN
2024-06-30	Streets, Roads, Highways (Bicycle Path, Private Road)	1500	STOLEN
2024-06-26	Streets, Roads, Highways (Bicycle Path, Private Road)	1000	STOLEN

Table 1 provides a snapshot of the cleaned data, showcasing key variables like Occurrence Date, Location Type, Bike Cost, and Status, which are central to the analysis. It helps to understand the structure and quality of the dataset. The table sets the context for further analysis, supporting discussions on theft patterns, trends, and the financial impact of stolen bicycles.

The dataset is published by the Toronto Police Service (Gelfand 2022), documenting bicycle theft incidents reported across the city. It provides detailed information about each theft incident, including the location where the bicycle was stolen (categorized as "Apartment/Condo," "Street/Highway," "Commercial," etc.), the estimated value of the stolen bicycle (in Canadian dollars), and whether the bicycle was recovered or remains stolen. The data is continuously updated, and the version used in this paper is current as of September 25, 2024.

The raw dataset contains 36,000 theft records, each capturing critical information such as the occurrence date, type of location, and the estimated bike cost. This dataset offers a comprehensive view of the prevalence and nature of bicycle thefts throughout Toronto, allowing for an analysis of patterns over time and across different parts of the city.

2.2 Dataset and Graph Sketches

Sketches depicting both the desired dataset and the graphs generated in this analysis are available in the GitHub Repository.

3 Results

3.1 Analysis of the Yearly Distribution of Bike Thefts Over Time

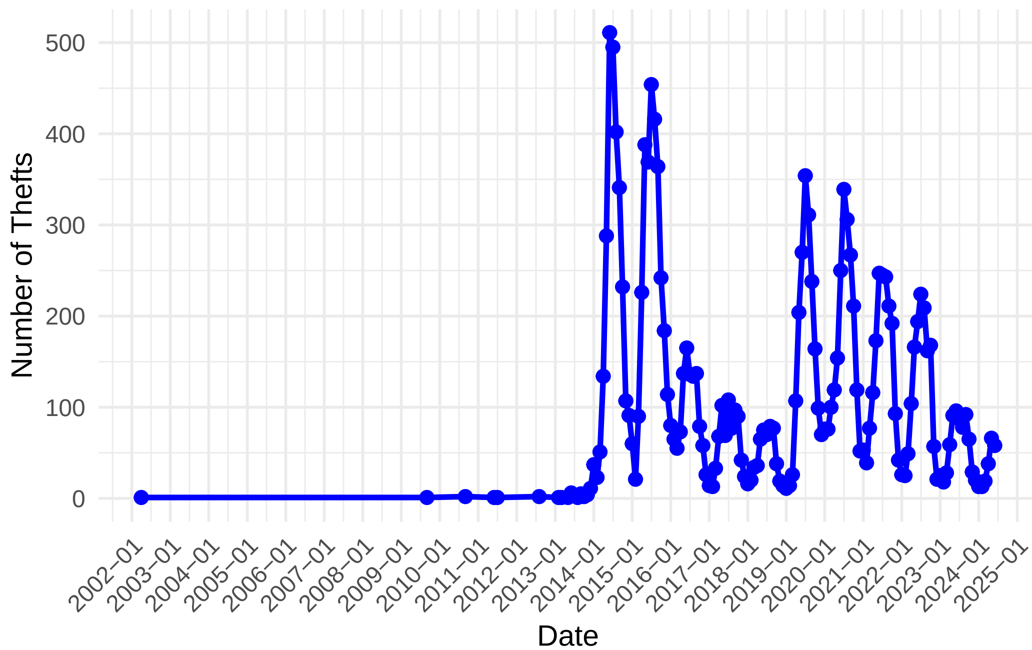

Figure 1: Yearly Distribution of Bike Thefts Over Time

Figure 1 illustrates the Yearly Distribution of Bike Thefts Over Time in Toronto. The x-axis shows the date of theft and the y-axis shows the number of thefts, allowing for an analysis of temporal patterns.

The line graph shows a sharp rise in bike thefts starting from around 2014 with a noticeable peak, where the number of thefts reached nearly 500 in a given month. This spike may have been influenced by factors such as increased bike usage and possible vulnerabilities in bike security systems during that period (Johnson, Sidebottom, and Thorpe 2008). Following this peak, thefts appear to gradually decline, although there are fluctuations. The drop observed during the pandemic aligns with findings that large-scale disruptions in mobility and economic activities lead to changes in crime patterns, including property crimes like bike theft (Nishisako and Fujino 2024).

Although the theft rate began to recover post-2020, it remained lower than the peak before the pandemic. This could be a suggestion of lasting effects of altered commuting patterns and security issues. Studies have indicated that disruptions like the pandemic can have lasting impacts on urban mobility and security strategies (Mburu and Helbich 2016). By 2024, the number of bike thefts drop significantly compared to earlier peaks.

3.2 Analysis of the Distribution of Theft Incidents by Location Type

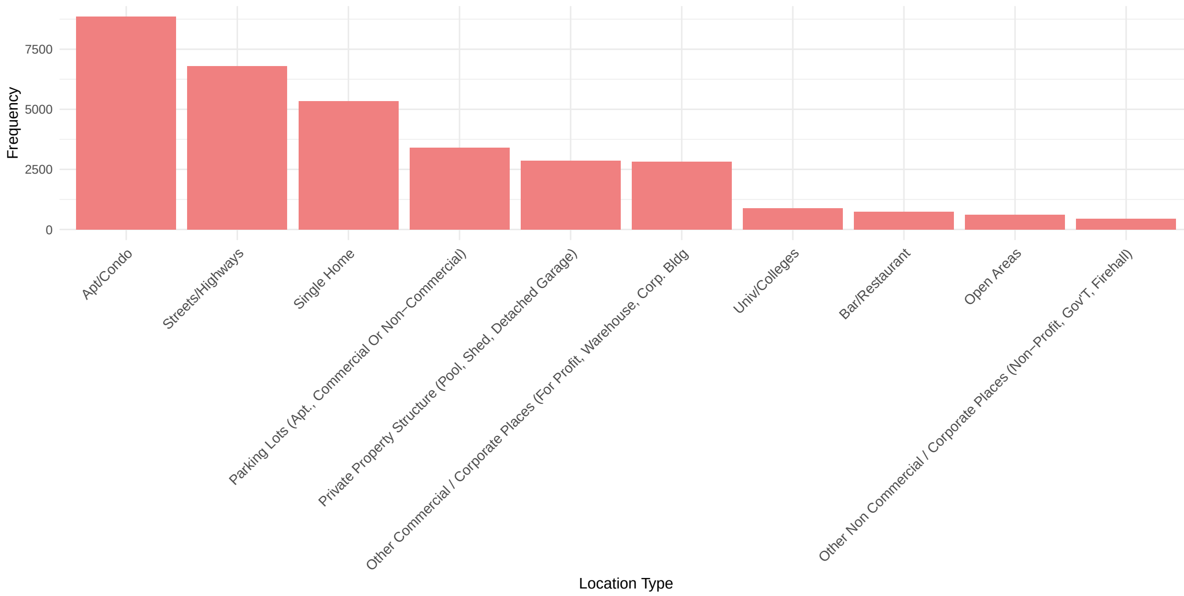

Figure 2: Top 10 Theft Incidents by Location Type

Figure 2 is a bar graph that illustrates the distribution of bicycle theft incidents by location type (x-axis), with the frequency of thefts (y-axis) across different categories.

The data reveals that apartments or condos have the highest frequency of thefts, aligning with the findings in the book Bicycle Theft (Johnson, Sidebottom, and Thorpe 2008), which highlighted that residential areas with minimal security such as outdoor spaces are common targets for theft.

Streets and highways rank second in theft frequency, a pattern consistent with an article based on environmental risk factors influencing bicycle theft (Mburu and Helbich 2016), which identified public areas as hotspots for theft due to their open nature and lack of surveillance. Noticeably, universities or colleges and commercial places show lower theft rates, likely due to better infrastructure like bike racks and surveillance, which is supported by the article on designing against bike theft (Thorpe, Johnson, and Sidebottom 2012) that stressed the importance of secure urban design in reducing theft.

3.3 Analysis of the Relationship Between Bike Cost and Theft or Recovery Status

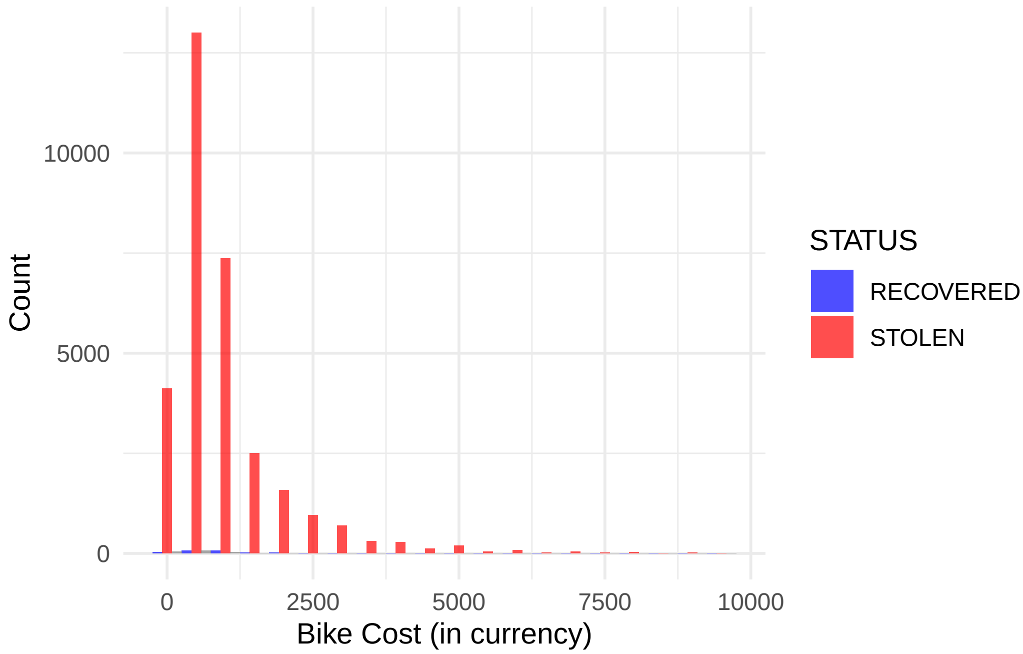

Figure 3: Bike Cost vs Theft Rate (Filtered Data)

Figure 3 shows the relationship between bike cost and theft rate, distinguishing between stolen and recovered bicycles. The x-axis shows the cost of bikes at different price levels and the y-axis shows the number of bikes been stolen. Moreover, the areas shaded in red are bikes that have been stolen and never recovered, and the area shaded in blue, which is hard to see, is the number of bikes that have been recovered.

The majority of the thefts occur with bikes valued at lower price points, typically below \$2500. This observation is consistent with research by Johnson, Sidebottom, and Thorpe (2008), which highlights that low-cost bicycles are often targeted due to their high availability and lower security measures. In addition, referring to (Van Lierop, Grimsrud, and El-Geneidy (2015)), it was noted that inexpensive bikes are frequently stolen in urban areas where they are parked in unsecured areas.

Interestingly, the graph shows that the recovery rates for stolen bikes are very low across all price ranges, as highlighted by the small count of recovered bikes. This reflects findings from the concentration of crime at places in Montreal and Toronto (Boivin and Melo 2019), which discussed the difficulties in recovering stolen property, particularly bicycles, which are often resold or dismantled quickly. In contrast, higher-value bikes, bikes priced over \$2500 are less likely to be stolen, likely because owners invest in better security measures as suggested by the article on designing against bicycle theft (Thorpe, Johnson, and Sidebottom 2012).

4 Discussion

4.1 Data Selection

The data for this study was obtained from Open Data Toronto (Gelfand 2022) and include variables essential for analyzing bicycle theft patterns: Occurrence Date, Location Type, Bike Cost, and Status. These variables were selected based on their relevance to understanding the temporal and spatial aspects of theft, as well as the financial impact and recovery rates.

Occurrence Date was chosen to examine trends in thefts over time, such as seasonal spikes during warmer months, which is consistent with findings by Van Lierop et al (Van Lierop, Grimsrud, and El-Geneidy 2015). Location Type provides spatial insights, helping to identify high-risk areas like public streets and commercial zones (Charron 2009). Bike Cost offers a view into the financial consequences of theft, while Status helps assess the effectiveness of recovery efforts (Boivin and Melo 2019).

Data from 2010 to 2024 were selected to ensure consistency and capture recent trends, while records with missing or incomplete entries, especially in Occurrence Date or Location Type, were excluded. This ensures the analysis focuses on well-documented incidents. Outliers in Bike Cost were also reviewed to avoid skewed results from extreme values.

4.2 General Analysis and Empirical Evidences

The three graphs—Yearly Distribution of Bike Thefts Over Time (Figure 1), Top 10 Theft Incidents by Location Type (Figure 2), and Bike Cost vs Theft Rate Using Filtered Data (Figure 3)-revealing important connections between time, location, and bike value in theft patterns. Together, they provide a broader understanding of how these factors influence bike theft risk.

Figure 1 shows that bike thefts peak in the summer when bikes are more frequently used and left in public places, which is supported by the second graph, Figure 2, showing that streets, roads, and apartments are high-risk locations for thefts. This suggests that both the seasonality and visibility of bikes in these areas play a major role in theft frequency. Figure 3, showing that lower-cost bikes are stolen more often but higher-cost bikes are more likely to be recovered, adds an economic dimension to this pattern. This implies that bike value influences both theft targeting and recovery efforts, likely due to better security measures for expensive bikes.

The findings from this study have significant implications for policy interventions aimed at reducing bicycle theft. One key area is expanding secure bike parking facilities in high-theft areas like streets and commercial zones. Previous research has highlighted the importance of urban design in crime prevention, particularly in ensuring that public spaces are well-lit and equipped with secure bike storage (Nettle, Nott, and Bateson 2012). By implementing more secure bike parking in these high-risk locations, cities can help deter theft and reduce the overall incidence of bicycle theft.

In addition to improving infrastructure, targeted theft prevention efforts during peak months are essential. As the data shows a seasonal trend in bicycle thefts, increasing public awareness and law enforcement patrols during the warmer months can be an effective deterrent. These interventions could help reduce the vulnerability of cyclists during the times of year when thefts are most common.

The low recovery rates for stolen bicycles, as indicated by the Status variable, suggest a need for improved recovery systems. More comprehensive bike registration programs, where bicycle owners can register identifying details, would make it easier to track and recover stolen property. This approach is supported by findings in broader crime recovery research, where stronger identification measures have been linked to higher recovery rates (Charron 2009). Additionally, enhancing community engagement in reporting stolen bicycles and improving coordination between law enforcement and local communities could also increase recovery rates.

Finally, it is essential to consider the broader social and environmental factors that influence theft. Research has shown that certain environmental conditions, such as high-traffic areas with low natural surveillance, can create opportunities for crime (Linden and Chaturvedi 2005a). Urban planners should account for these risk factors in designing safer public spaces and integrating crime prevention into urban development strategies.

4.3 Limitations and Next Steps

This study, while providing valuable insights into bicycle theft patterns in Toronto, has several limitations. Firstly, it relies on a dataset containing bicycle theft occurrences from 2014 to 2020. A significant limitation is the deliberate offsetting of crime location data to the nearest road intersection node. This was done to protect the privacy of individuals involved in the incidents, but it introduces a level of uncertainty regarding the exact locations of thefts. As a result, the reported counts of occurrences by Division and Neighbourhood may not accurately reflect the true distribution of bicycle thefts within these areas. Consequently, the Toronto Police Service does not guarantee the accuracy, completeness, or timeliness of the data, and it should not be compared to other crime data sources.

Another limitation stems from the dataset's time frame, which may not capture recent trends, particularly in light of the impact of the COVID-19 pandemic on cycling behavior and theft rates. The focus on a specific period may overlook shifts in crime patterns and the effectiveness of recent prevention measures, referring to the article focusing on the need for comprehensive crime prevention planning (Linden and Chaturvedi 2005b).

Finally, the study is confined to Toronto, limiting the generalization of the findings. Theft patterns and effective prevention strategies may vary significantly in different urban contexts, depending on factors such as infrastructure, crime rates, and socio-economic conditions.

To address these limitations, future research should consider incorporating additional datasets that include unreported bicycle theft incidents. This could help provide a more comprehensive view of the extent of the problem. Additionally, researchers could explore alternative methods for reporting theft occurrences that do not compromise privacy while still providing more precise geographical data.

Expanding the analysis to include more recent data beyond 2020 would also be beneficial, especially given the changes in cycling patterns during and after the pandemic. A longitudinal study could examine how theft rates have evolved in response to shifting urban dynamics, infrastructure changes, and community safety initiatives.

References

Boivin, Rémi, and Silas Nogueira de Melo. 2019. "The Concentration of Crime at Place in Montreal and Toronto." *Canadian Journal of Criminology and Criminal Justice* 61 (2): 46–65.

Charron, Mathieu. 2009. *Neighbourhood Characteristics and the Distribution of Police-Reported Crime in the City of Toronto*. Statistics Canada Ottawa.

Firke, Sam. 2024. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://github.com/sfirke/janitor.

Gelfand, Sharla. 2022. *Opendatatoronto: Access the City of Toronto Open Data Portal*. https://CRAN.R-project.org/package=opendatatoronto.

Johnson, Shane D, Aiden Sidebottom, and Adam Thorpe. 2008. *Bicycle Theft*. Vol. 52. US Department of Justice, Office of Community Oriented Policing Services ….

Linden, Rick, and Renuka Chaturvedi. 2005a. "The Need for Comprehensive Crime Prevention Planning: The Case of Motor Vehicle Theft." *Canadian Journal of Criminology and Criminal Justice* 47 (2): 251–70.

———. 2005b. "The Need for Comprehensive Crime Prevention Planning: The Case of Motor Vehicle Theft." *Canadian Journal of Criminology and Criminal Justice/La Revue Canadienne de Criminologie Et de Justice Pénale* 47 (April): 251–70. https://doi.org/10.3138/cjccj.47.2.251.

Mburu, Lucy Waruguru, and Marco Helbich. 2016. "Environmental Risk Factors Influencing Bicycle Theft: A Spatial Analysis in London, UK." *PLoS One* 11 (9): e0163354.

Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. https://CRAN.R-project.org/package=here.

Nettle, Daniel, Kenneth Nott, and Melissa Bateson. 2012. "'Cycle Thieves, We Are Watching You': Impact of a Simple Signage Intervention Against Bicycle Theft." *PloS One* 7 (12): e51738.

Nishisako, Moe, and Tomokazu Fujino. 2024. "Predicting the Risk of Bicycle Theft Occurrence Considering Routine Activity Theory and Spatial Correlation." In *International Conference on Human-Computer Interaction*, 111–20. Springer.

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

Thorpe, Adam, Shane D Johnson, and Aiden Sidebottom. 2012. "Designing Against Bicycle Theft." *Design Against Crime: Crime Proofing Everyday Products* 27: 107–30.

Van Lierop, Dea, Michael Grimsrud, and Ahmed El-Geneidy. 2015. "Breaking into Bicycle Theft: Insights from Montreal, Canada." *International Journal of Sustainable Transportation* 9 (7): 490–501.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Xie, Yihui. 2024. *Knitr: A General-Purpose Package for Dynamic Report Generation in r.*

https://yihui.org/knitr/.