Reported Toronto Bike Theft Recovery is Abysmal*

Reported Bicycle Theft to the Toronto Police Service, 2014-2020

Shae-Linn Davies

6 February 2022

Abstract

The accurate reporting of crime in cities is of the utmost importance to understand civic safety, and where to best allocate police funding. Individuals who are victims of crime may be motivated to report crime, especially if they are seeking the recovery of stolen items such as their mode of transportation. While reports of bicycle theft occuring at apartments have doubled in a seven-year period, reports of stolen bikes from other locations remain largely consistent. Recovery of stolen bikes is consistently reported as extremely low, or is of unknown status, in comparison to the number of bikes reported as stolen. This report utilizes data on reported bicycle theft in Toronto from the City of Toronto Open Data Portal (Gelfand 2020) to: (1) assess reported bike theft trends in Toronto and the success rate of bike recovery and (2) critically discuss factors related to the aforementioned, including potential challenges that come from and with the reporting and analysis of bicycle thefts.

Contents

1 Introduction

Since the launch of the Open Data Toronto Portal in 2009, (Gelfand, n.d.), municipal data has become freely-accessible to the public, assuming they know how to access, read and understand it. In the context of managing a complex city, this supports accountability and transparency, especially on the part of groups who hold social power, such as the police. In 2017, the Toronto Police Service released their The Way Forward Plan (Service, n.d.) as an effort to modernize community safety in Toronto. Part of this plan was to revamp the way data related to crime in Toronto is shared.

Regarding transportation data and research, there have not been many studies to evaluate bicycle thefts, nor overall security and availability of facilities to safely store bicycles (Van Lierop, Grimsrud, and El-Geneidy

*Code and data are available at: https://github.com/shaeplays/bicycle_thefts_TO

2015). This is an area of interest as "cycling is one of the fastest growing transportation modes in Toronto" (City of Toronto 2021). The City of Toronto has launched the *StreetSmartsTO* initiative, which highlights the importance of road safety through the prioritization of the expansion of bike lanes and creation of bike maintenance and road safety workshops (Cycle Toronto 2021). There exists some direction on their website related to theft prevention and what to do if one's bike is stolen, however it is not easily accessible there, nor on the Toronto Police Service website.

There also may exist gaps in reporting information by the Toronto Police, overall. While it is clear apartment bike thefts have doubled since 2014, thefts reported as originating at other locations remain consistent. The more worrisome trend is the lack of data regarding recovered bikes, and the disparity between the number of reported stolen bikes and reported recovered bikes. This is of concern because this means the data is either not accurately recorded, or, that there is a staggering amount of bikes that are never recovered. Often, bicycle crimes are viewed as low-priority crimes, and thus are not properly recorded to the police (Gamman, Thorpe, and Willcocks 2004). If there are inaccurate records, or lack of demonstration of ability to follow up on reported thefts, this has the potential to breed doubt in the police. In addition, this means that bike crimes may be much more prominent than expected.

In this report, I use `R` (R Core Team 2021) to identify trends in reported bike thefts between 2014-2020. This timeframe is explored as this is the timeframe for which there was available data at the time of writing this report however, it is understood that `opendatatoronto` (Gelfand 2020) regularly updates their datasets. I will first analyze trends in location of theft. Then I will compare how many bikes are reported stolen in comparison to recovered. Most importantly, I will discuss issues of bias, missing reporting data, the potential for misleading data, and what real-world implications this may have, through the lens of considering how the police is sharing data with the public.

2 Data

2.1 Data Source and Collection

I have imported and processed the Bicycle Thefts `opendatatoronto` dataset (Gelfand 2020) and analyzed it using `R` (R Core Team 2021) and its related packages. The dataset was last refreshed March 23, 2021 and is of "Gold" Data Quality Score. Packages that were used for analysis and preparation include `janitor` (Firke 2021), `lubridate` (Spinu, Grolemund, and Wickham 2021), `tidyverse` (Wickham 2021), `tidyr` (Wickham and Girlich 2022) and `dplyr` (Wickham et al. 2021). I created a graph using `ggplot2` (Wickham 2016) and created tables using `kableExtra` (Zhu 2021). I used `knitr` (Xie 2021b) and `bookdown` (Xie 2021a) to generate the final `RMarkdown` (Allaire et al. 2021) report.

The following disclaimer is not included on the `opendatatoronto` dataset (Gelfand 2020) site, however, the specific Bicycle Thefts dataset is also accessible through the Toronto Police Service Publc Safety Data Portal, where it is included: "The reported crime dataset is intended to provide communities with information regarding public safety and awareness. The data supplied to the Toronto Police Service by the reporting parties is preliminary and may not have been fully verified" (Service 2022). This must be taken into careful consideration when conducting data analysis, as there is a margin of uncertainty that is unknown.

2.2 Data Characteristics and Methodology

The Bicycle Thefts `opendatatoronto` dataset (Gelfand 2020) includes bike theft occurrences reported by date, in Toronto's 140 neighbourhoods between 2014-2020. There are 25,569 observations across 33 attributes in the original dataset. The 33 attributes include:unique IDs for each occurrence, the primary offence, the occurrence date (i.e., when the presumed theft occurred), report date (i.e., when the occurrence was reported), the neighbourhood and police division the theft occurred in, the type of location the theft occurred at (e.g., outside, at an apartment, at an educational institution), as well as the bike's make, model, type, speed, colour, cost, and whether it had since been recovered or not, as information was available. An additional attribute, called *instances*, was created to count each occurrence, for efficiency. I primarily utilized attributes including the location of the bike theft, along with the occurrence and report dates of theft and the status of

whether the stolen bike had been recovered. Of note, the dataset included occurrence date, report date, and also provided separate month, day, time, hour, and day of year attributes which while I have not used, are appreciated as conveniences should I explore the dataset further at a later time.

2.3 Reported Bike Thefts

To begin exploring bike thefts in Toronto occurring between 2014-2020, I first considered bike thefts by reported location. Per *Breaking into bicycle theft* (Van Lierop, Grimsrud, and El-Geneidy 2015), lack or fear of theft while using bike parking facilities is of great concern thus, I used the "premises type" attribute to understand where bike theft was most common.

Of note, in the `opendatatoronto` Bicycle Thefts dataset (Gelfand 2020), there also exists a "location type" attribute. "Premises type" provides a higher level categorization, with types including: Transit, Educational, Commercial, House, Apartment, Outside, or Other. "Location type," is unnecessarily detailed for this level of report (i.e., types include more precise categorization such as "GO Bus" or "TTC Subway" instead of the higher level premise type, "Transit"). In addition, "location type" does not further illuminate the finding that there is an overwhelmingly evident disparity, seen in Figure 1, between bikes that are reported as stolen and bikes that have been recovered. Thus, utilizing "premise type" was sufficient and I have not included a more detailed graph by location type.

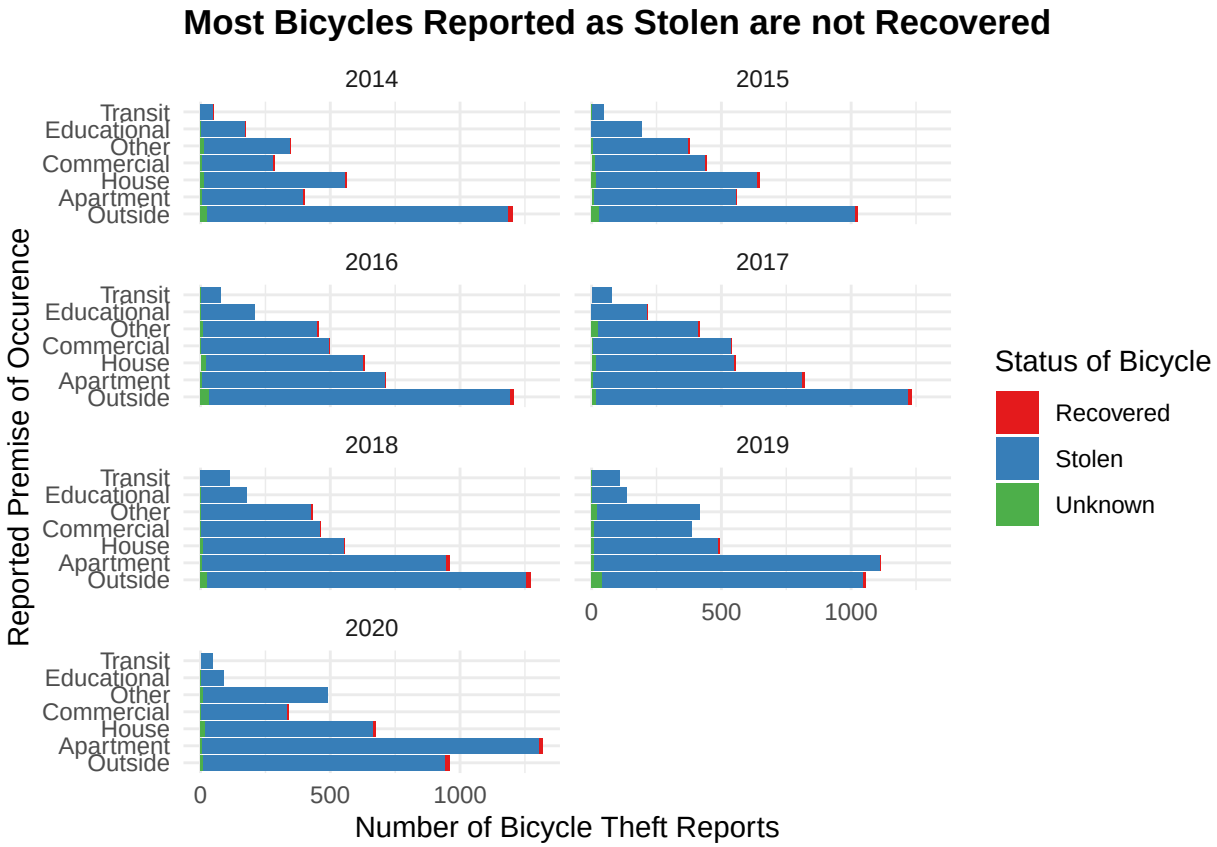

Figure 1: Reported Bicycle Thefts in Toronto by Premise of Occurence 2014-2020

Year over year, "Outside" is often the top location from which bikes were stolen. This is unsurprising as most bike racks are located outside (Gelfand 2020). Of interest, apartment thefts greatly increased over the six-year period, with it being the most common location for reported bike theft in 2019 and 2020. This may be due to the COVID-19 pandemic, where more people were in their place of residence more often and would be able to scout out and steal bikes. Similarly, the number of bikes stolen from a commercial

and educational locations also saw decreases in 2019 and 2020, most likely due to the COVID-19 pandemic, as less people we congregating in those spaces, thus, saw less occasions for bikes to be at those locations. Transit was consistently the least common location for bikes to be stolen from, followed by educational premises. Houses, commercial spaces, and other venues accounted for an overall average number of bike thefts. Unfortunately, most bikes' statuses remain stolen(24,807/25,569 or 97%), with only 308/25,569 (1.2%) reported as "Recovered." 454/25,569 (1.8%) reported bikes are of "Unknown" status, referring to bikes that had been reported as stolen, but it is unknown of whether they have been recovered or not. This may be due to lack of follow-up on the part of the police, the victim, both parties, or other factors.

Refer to Table 1, below, for a more detailed overview of reported bicycle thefts in Toronto from 2014-2020 by premise type.

Table 1: Reported Bicycle Thefts in Toronto by Premise of Occurence 2014-2020

Premise	2014	2015	2016	2017	2018	2019	2020
RECOVERED							
Apartment	6	3	5	9	12	5	15
Commercial	7	4	6	4	7	NA	6
Educational	1	2	NA	3	NA	NA	1
House	10	13	10	4	3	4	9
Other	4	9	7	9	6	1	2
Outside	22	11	18	14	21	12	18
Transit	1	NA	3	1	NA	NA	NA
STOLEN							
Apartment	389	550	703	805	941	1099	1297
Commercial	275	427	490	533	455	375	331
Educational	168	193	209	214	174	132	86
House	541	617	605	534	543	479	647
Other	329	364	441	383	423	396	480
Outside	1157	986	1157	1203	1227	1003	933
Transit	49	45	75	77	113	108	46
UNKNOWN							
Apartment	7	7	7	6	5	9	7
Commercial	4	11	2	3	2	9	2
Educational	3	NA	1	NA	3	1	2
House	13	17	19	15	9	8	18
Other	13	6	8	24	2	20	10
Outside	25	29	34	15	25	41	9
Transit	NA	1	1	NA	NA	1	NA

Note:

Graphed data from Figure 1 is provided in table format for readability and accessibility.

2.4 Recovered Bikes

To take a more granular perspective, I next considered same-day reported bike thefts (refer to Table 2) and bike thefts that were reported at least 24 hours after the theft occurred (refer to Table 3). By comparing this factor, a better understanding of trends of rate of recovery of bicycles can be identified, along with the ability to hypothesize reasons why this might be.

Same-day reporting of a bike theft leading to the recovery of the bike was overall slightly more successful than recovery of a bike that had been reported as stolen 24 hours or more after the occurrence. 169/308 (54.9%) recoveries accounted for the former while 139/308 (45.1%) recoveries accounted for the latter.

Recovery data by premise type was fairly consistent whether the victim of crime had reported same-day or at least 24 hours after the theft had taken place. Bikes recovered were most commonly taken from outside, and from apartment buildings. Of note, these were also the most common places to have had a bike stolen from. Bikes recovered were least likely recovered if taken from transit, followed closely by educational premises. Table 2 and Table 3 can be fond below, further detailing this information.

The table below is a subset of data that corresponds to the data plotted in Figure 1, specifically depicting the number of bikes in Toronto between 2014-2020 that had been reported as stolen on the same day as the theft occurred, and were later recovered.

Table 2: Recovered Bikes in Toronto 2014-2020: Report Date Matched the Theft Occurence Date

Premise	2014	2015	2016	2017	2018	2019	2020
RECOVERED							
Apartment	1	2	1	2	5	3	7
Commercial	3	3	3	2	3	NA	5
Educational	1	1	NA	2	NA	NA	1
House	4	7	7	4	NA	2	5
Other	2	4	2	5	1	NA	1
Outside	17	7	14	7	13	7	11
Transit	NA	NA	3	1	NA	NA	NA

The table below is a subset of data that corresponds to the data plotted in Figure 1, specifically depicting the number of bikes in Toronto between 2014-2020 that had been reported 24 hours or more after the theft occurred, and were later recovered.

Table 3: Recovered Bikes in Toronto 2014-2020: Report Date was >24h After the Theft Occurrence Date

Premise	2014	2015	2016	2017	2018	2019	2020
RECOVERED							
Apartment	5	1	4	7	7	2	8
Commercial	4	1	3	2	4	NA	1
Educational	NA	1	NA	1	NA	NA	NA
House	6	6	3	NA	3	2	4
Other	2	5	5	4	5	1	1
Outside	5	4	4	7	8	5	7
Transit	1	NA	NA	NA	NA	NA	NA

2.5 Data Strengths and Weaknesses

It is quite impressive to have included 33 attributes in a dataset related to bike thefts. It is appreciated that the dataset includes both the date of the report, and when the bike was noticed missing, with a separate column for the occurrence date. This provides transparency as it is possible someone reported their bike as stolen much later than the date it was originally stolen. At the same time, there is no column to specify when bikes were recovered, if they were recovered, creating a source of ambiguity. Similarly, there are no columns holding the police accountable to maintaining up-to-date records, such as a column to specify when they last visited each theft case, or whether they followed up with the victim of crime to see if it had been found via other means. In addition, the data lacks contextual information that could support the recovery of more bikes in the future such as which police division recovered each bike if it was recovered, and what neighbourhood and premise type it was recovered from. This could be compared to which neighbourhoods and premises bikes are most often stolen from to further inform the police's bike recovery methodology. Data collection weaknesses are further discussed in the Discussion section, below.

3 Discussion

A closer look into reported bicycle theft and recovery highlights the unfortunate reality that bike recovery rates in Toronto are historically abysmal between 2014-2020. While there are minimal entries with unknown data, it is unfortunate that there are more bikes of "unknown" status than "recovered" status. Factors for the low rate of recovery potentially include the police's failure to follow-up with cases, lack of search leads, and/or lack of resources to update the status of occurrences. These factors are potentially worrisome, especially as failure to accurately keep and update records could create mistrust from the public, against the police's *The Way Forward* (Service, n.d.) plan. Worse, if the police were neglecting to update the data, this could create inflated data of open police cases. This would be extremely problematic if used as a basis to increase police presence and funding. While unlikely, these are still of note as potential factors. More likely, bikes are much harder to track as they are smaller, and easily taken apart. They can be swiftly be moved, and resold whole or sold for parts.

To expand upon the police's potential challenges of lack of leads and motivation, factors could include lack of incentive or time to find the bike in comparison to larger or more dangerous crimes. It must also be taken into account that the dataset and my data representations do not and cannot include all stolen and recovered bikes, as mentioned in the disclaimer. This is because this data relies on reporting and follow-up to take place between the police and the public. In itself, this produces challenges due to potential for communication flow barriers between the parties, separate and in addition to data input and validation.

4 Conclusion

While there exists a large gap between the number of bikes reported as stolen and those reported as later recovered, it is important that this data is recorded, available and accessible to allow opportunity for the public to access information related to neighbourhood crime, and make more informed decisions. On the part of the police, analyzing this data is important to be able to recognize trends such as the large increase of reported bike thefts in apartments. Hopefully with more awareness regarding this gap, along with these additional insights on bike theft trends, more bikes are reported as recovered in the future. Further methods for bike theft prevention must be explored, along with the potential to streamline the reporting process to reduce the potential for reporting inaccuracies or ambiguity. Statistics of bike thefts in Toronto can be an indicator of city safety but should not be the only factor due to the reliance on public reporting and police follow-up. Overall, it is appreciated that the Toronto Police Service is working towards being more transparent to help build relationships of trust with the public.

5 References

Allaire, JJ, Yihui Xie, Jonathan McPherson, Javier Luraschi, Kevin Ushey, Aron Atkins, Hadley Wickham, Joe Cheng, Winston Chang, and Richard Iannone. 2021. *Rmarkdown: Dynamic Documents for r.*

Firke, Sam. 2021. *Janitor: Simple Tools for Examining and Cleaning Dirty Data.* https://github.com/sfirke/janitor.

Gamman, Lorraine, Adam Thorpe, and Marcus Willcocks. 2004. "Bike Off! Tracking the Design Terrains of Cycle Parking: Reviewing Use, Misuse and Abuse." *Crime Prevention & Community Safety* 6 (4): 19–36. https://search-ebscohost-com.myaccess.library.utoronto.ca/login.aspx?direct=true&db=i3h&AN=64003360&site=ehost-live.

Gelfand, Sharla. 2020. *Opendatatoronto: Access the City of Toronto Open Data Portal.*

———. n.d. "About City of Toronto Open Data." *City of Toronto Open Data Portal.* https://open.toronto.ca/about/.

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

Service, Toronto Police. 2022. "Bicycle Thefts." *Public Safety Data Portal.* Toronto Police Service. https://data.torontopolice.on.ca/datasets/TorontoPS::bicycle-thefts/about.

———. n.d. "The Way Forward." *Toronto Police Service.* https://www.torontopolice.on.ca/TheWayForward/.

Spinu, Vitalie, Garrett Grolemund, and Hadley Wickham. 2021. *Lubridate: Make Dealing with Dates a Little Easier.*

Toronto, City of. 2021. "Cycling in Toronto." *City of Toronto.* https://www.toronto.ca/services-payments/streets-parking-transportation/cycling-in-toronto/.

Toronto, Cycle. 2021. "Streetsmartsto." *Cycle Toronto.* https://www.cycleto.ca/streetsmartsto.

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. https://doi.org/10.1080/15568318.2013.811332.

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

———. 2021. *Tidyverse: Easily Install and Load the Tidyverse.*

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2021. *Dplyr: A Grammar of Data Manipulation.*

Wickham, Hadley, and Maximilian Girlich. 2022. *Tidyr: Tidy Messy Data.*

Xie, Yihui. 2021a. *Bookdown: Authoring Books and Technical Documents with r Markdown.* https://github.com/rstudio/bookdown.

———. 2021b. *Knitr: A General-Purpose Package for Dynamic Report Generation in r.* https://yihui.org/knitr/.

Zhu, Hao. 2021. *kableExtra: Construct Complex Table with Kable and Pipe Syntax.*