Time, Seasonal, and Locational Trends of Bicycle thefts in Toronto*

Dai Moroi

February 8th 2022

Abstract

In this paper, I examined time and locational trend of bicycle thefts in Toronto. The result shows that most of the cases are concentrated on the specific 10 neighborhoods, there are less cases in winter and weekends. It may implie that Toronto Police Service should keep in mind the result when strengthening the security of the city.

1 Introduction

Bicycles are very useful tool for commuting, workout, and travelling. Since the outbreak of COVID-19, bicycles has been gathering attention, which led to so-called "COVID-19 bicycle boom" (CBC 2021). That is because biking can be a alternative for workout in the gym; Or it can prevent people from using crowded public transportation, such as trains and buses. However, this boom has highlighted the serious problem as well, that is to say, bicycle thefts.

In a big city like Toronto, bicycle thefts can be a huge issue. In 2020, Toronto Police Service (Service 2020a) reported nearly 4000 thefts in the city. To prevent these thefts, it is important to analyze from data when and where it is likely to occur and take proper measures according to analysis. Thus, in this paper, I will examine the time, seasonal, and locational trends of bicycle thefts using the actual data, and try to make some suggestions to Toronto Police Service.

2 Data

The analysis for this paper uses the R statistical programming language (R Core Team 2021), primarily using the `tidyverse` package for data manipulation (Wickham et al. 2019). In addition, The data is imported from the Open Data Toronto Portal using the `opendatatoronto` package (Gelfand 2020). The packages `janitor` (Firke 2021) and `tidyr` (Wickham 2021a) are used to make figures and clean the data, and `bookdown` (Wickham 2021b) is used for making a R markdown report.

The source of the data is the Toronto Police Service Annual Statistical Report (ASR). ASR is a comprehensive overview of police related statistics including reported crimes, victims of crime, search of persons, firearms, traffic collisions, personnel, budget, communications, public complaints, regulated interactions and other administrative information (Service 2020b). The dataset contains bicycle thefts occurrences from 2014-2020 by reported date and details such as cost and types of bike, name of neighborhood, and premise types.

Collection of this data is based on the record of phone call; Those who get bicycle thefts call 416-808-2222 (details are available here), and then Toronto Police Service processes the data, making it available as open source data.

*Code and data are available at: https://github.com/moroidai/paper1

Bias that this dataset potentially hold can be divided into three aspects. First, the accuracy of location that the thefts occurred is not consistent. The location of crime occurrences have been deliberately offset to the nearest road intersection node to protect the privacy of parties involved in the occurrence (Toronto 2020). Also, it includes those where the location has not been able to be verified and occurrences where the coordinate location is outside the City of Toronto because it contains all bicycle theft occurrences reported to the Toronto Police Service (Toronto 2020). Thus, we have to be cautious when analyzing it in terms of locations or neighborhood. However, the data is still useful to grasp the big picture or trends. Second, we have to keep in mind that not all the crimes occurred in Toronto are reported to Toronto Police Service, so the actual number could be much higher than the data. People often don't report the crimes if it doesn't really matter to them, thus bicycle thefts that the data contains could be biased toward those of relatively expensive bikes. In addition, total cases can change a lot if Toronto Police Service changes the way of collecting report to easier one or they give some incentives for citizens to report crimes. However, I couldn't see such descriptions on the website, so we don't need to care about that. Third, the reason for some increase in the cases of bicycle thefts can be the increase in the total number of bicycle owned by citizens in Toronto. The more bikes there are, the more thefts can occur even if the crime rates are the same. I couldn't find the statistics about evolution of the number of uses of bicycles in Toronto, but we should keep in mind that the increase in cases can be due to the increase in total number of bicycles.

In order to understand the trend and characteristics of the data, I will analyze the data from these aspects below:

- Evolution of Cases
- Seasonal Trend
- Day of Week
- Location
- Neighborhood

2.1 Cases

Figure 1 shows the changes in cases of bicycle thefts over time.

Figure 2 shows the growth from previous year by percentage.

Looking at those two graphs, we can notice the rapid increase in 2015 and 2016. It is also noticed that there was some increase in 2020 along with the COVID-19 restrictions. This may show the effect of the bicycle boom.

2.2 Seasonal Trend

Next, we will see if there is a seasonal trend for bike thefts in Toronto.

Look at Figure 3. We can see a very strong trend that there are higher number in summer and smaller number in winter. From December to March, lowest in February, there are very small number of thefts, and it gradually increases until July, then starting to decrease again. It is interesting that the months with very small number of thefts, from December to March, corresponds to the months that has snow. Because Toronto is very cold and it snows a lot during winter, the data shows that people may tend not to use bicycle in winter.

2.3 Day of week

Figure 4 shows the cases sorted by day of week.

It is noticed that bicycle thefts mostly occurred on weekdays, where Saturday has 3471 case and Sunday has 3323 cases compared to the average cases 3645. This is probably because people use bikes to go to their offices or schools on weekdays.

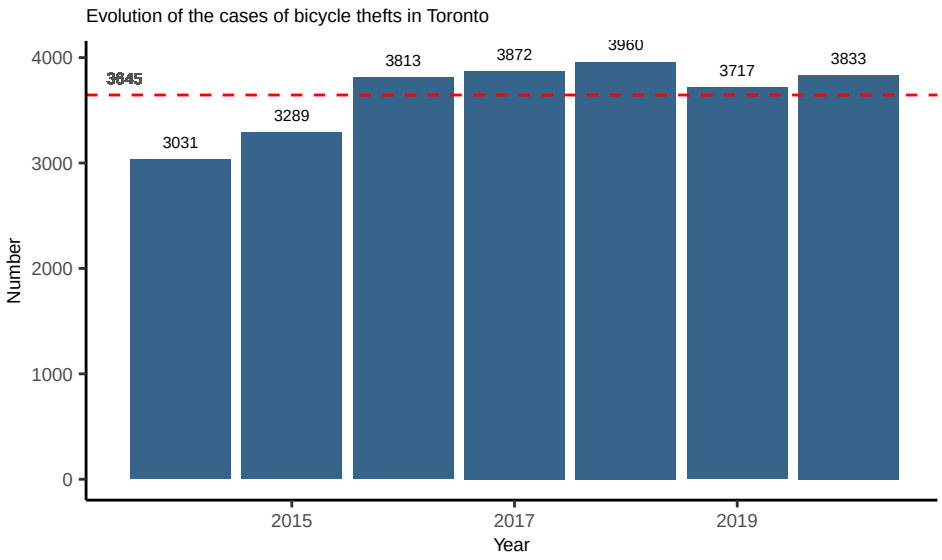

Figure 1: Cases over time

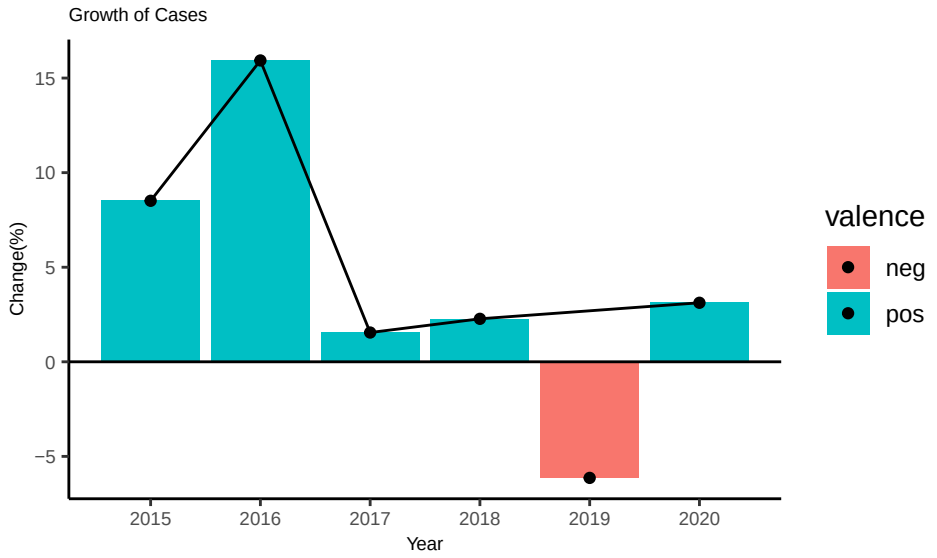

Figure 2: Growth over time

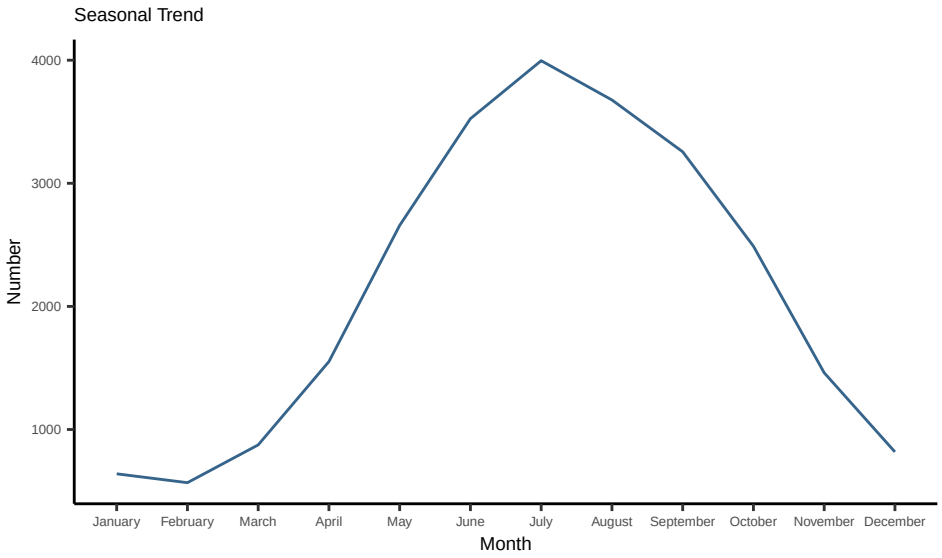

Figure 3: Seasonal trend

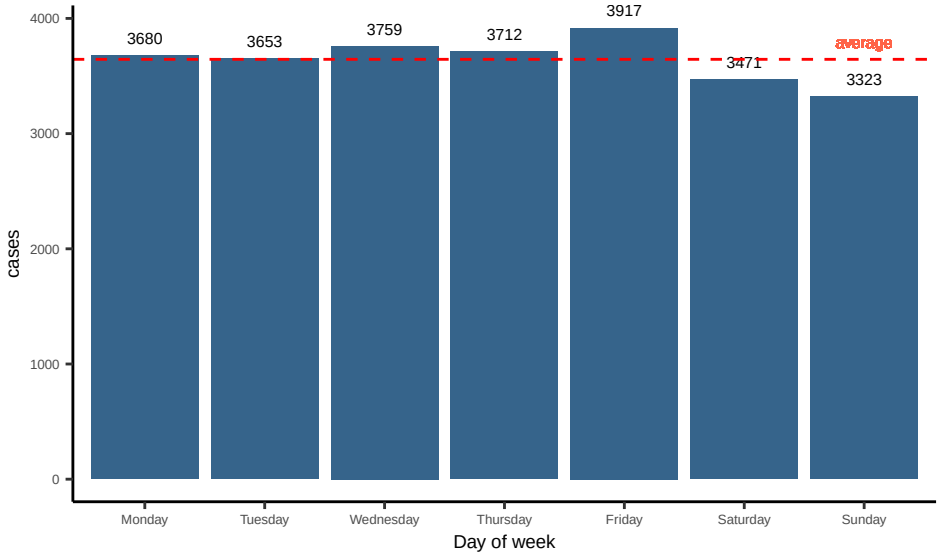

Figure 4: Day of Week

2.4 Trend of Hours

Figure 5 shows the cases sorted by hours.

The graph shows that bicycle thefts most occurred around 7pm with relatively high number from 8 a.m. to 0 a.m. It is noticed that arount midnight and early in the morning, 1 a.m. - 7 a.m..

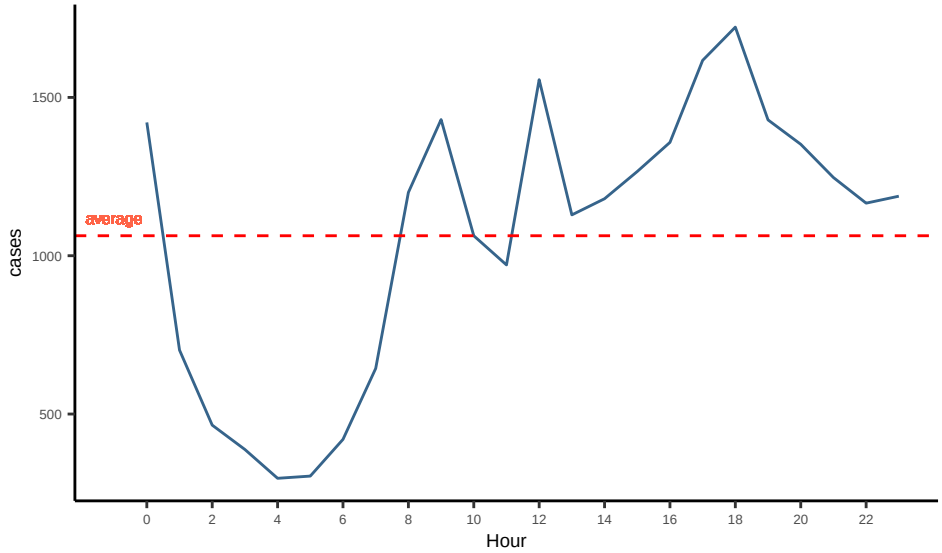

Figure 5: Time Trend

2.5 Location

Looking at the location where bicycle thefts occurred in Figure 6, we can see that place that it occurred most is outside, followed by apartment, house, commercial, and etc. It is surprising that many thefts are residential, where 39% of all bicycle theft occurred in apartments or houses.

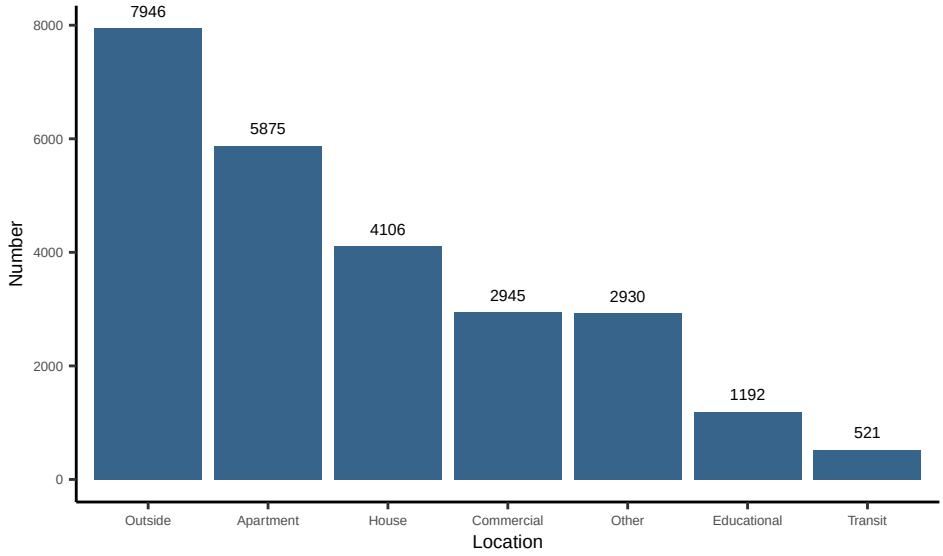

Figure 6: Location

Table 1: Top10 neighborhood population and cases per capita

Neighborhood	Number	Populattion	Cases per capita
Waterfront Communities-The Island (77)	2573	65913	0.03904
Bay Street Corridor (76)	2109	25797	0.08175
Church-Yonge Corridor (75)	1671	31340	0.05332
Niagara (82)	990	31180	0.03175
Annex (95)	946	30526	0.03099
Kensington-Chinatown (78)	836	17945	0.04659
Moss Park (73)	801	20506	0.03906
University (79)	779	7607	0.10241
South Riverdale (70)	715	27876	0.02565
Dovercourt-Wallace Emerson-Junction (93)	642	36625	0.01753

2.6 Neighbourhood

Figure 7 shows top 10 neighborhoods of 140 by number of bicycle thefts in Toronto and its number. Top 10 neighborhoods consist of "Waterfront Communities-The Island," "Bay Street Corridor," "Church-Yonge Corridor,""Niagara," "Annex,""Kensington-Chinatown," "Moss Park," "University," "South Riverdale," and "Dovercourt-Wallace Emerson-Junction." It shows that top 3 neighborhood has very high cases, 2573, 2109, and 1671 respectively.

Look at figure 8. It shows the ratio of cases occupied by top 10 neighborhoods compared to other 130 neighborhood. It is surprising that only 10 neighborhoods account for 47% of total cases. This number is significant and it may imply that Toronto Police Service should strengthen the security of those neibourhoods.

Because 7 and 8 don't take population into account, I made a table that includes population of those neighborhood. The data is from City of Toronto Open Data Portal (Toronto 2019) and it shows population in 2016. Total population of City of Toronto is 2,731,571, so average cases per capita is 0.00934. Thus those top 10 neighborhoods are much higher than average in cases per capita.

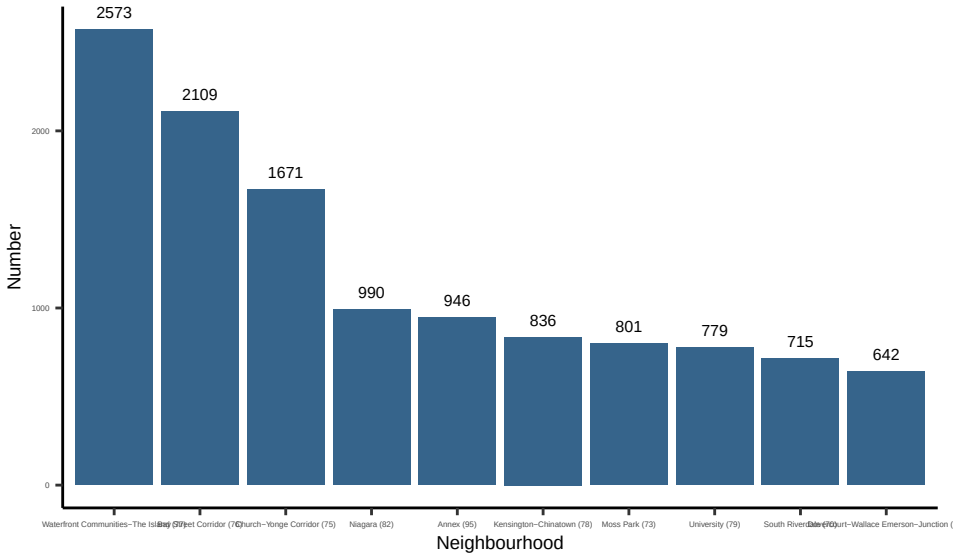

Figure 7: Neighbourhood

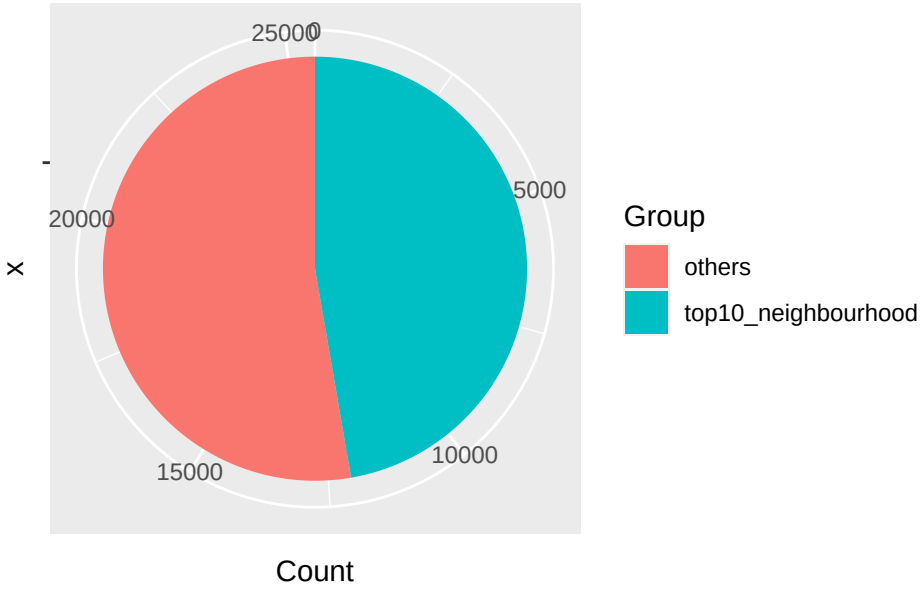

Figure 8: Top10 ratio

References

CBC. 2021. "When Will Bicycle Shops Be Fully Stocked Again? Try 2023 | CBC News." CBC News. https://www.cbc.ca/news/business/bicycle-boom-industry-turmoil-covid-19-1.5956400.

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

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

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. 2020a. "Bicycle-Theft." Toronto Police Service. https://open.toronto.ca/dataset/bicycle-thefts/.

———. 2020b. "Toronto Police Service Public Safety Data Portal." https://data.torontopolice.on.ca/pages/annualstatisticalreport.

Toronto, City of. 2019. "Neighbourhood Profiles." City of Toronto Open Data Portal. https://open.toronto.ca/dataset/neighbourhood-profiles/.

———. 2020. "Bicycle Thefts." City of Toronto Open Data Portal. https://data.torontopolice.on.ca/pages/bicycle-thefts.

Wickham, Hadley. 2021a. *Tidyr: Tidy Messy Data*.

———. 2021b. *Tidyr: Tidy Messy Data*.

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.