

My title*

Yiliu Cao

October 3, 2023

This study utilizes data on bicycle thefts sourced from the Toronto Open Data Portal. Preliminary findings indicate a increase propensity for bicycle thefts during summer months. While black mountain bicycles constitute a significant proportion of stolen bikes, it does not necessarily imply they are more susceptible to theft due to their overall prevalence. Moreover, the cost of a bicycle does not seem to influence its theft risk, thieves will not evaluate the price before they steal it. Future research should explore the role of neighborhoods as a potential factor influencing bicycle thefts.

Table of contents

1	Introduction	2
2	Data	2
2.1	The Time of Bicycle Thefts Occurrence	3
2.2	The Premises Place Where the Bike was stolen	3
2.3	The Stolen Bicycles' Own Features	4
3	Results	4
3.1	Compare the bicycle thefts in the recent six years	4
3.2	Compare the premises types and probability to be stolen during summer. . . .	7
3.3	Compare the top 4 colours and top 5 bicycle types from bicycle thefts	7
4	Discussion	7
4.1	Repeating patterns of bicycle thefts: expected to decrease in 2023.	7
4.2	Increasing number of bicycle thefts on the outside and educational areas. . . .	8
4.3	No evidence indicates that the cheaper/more expensive bicycles are less/more likely to be stolen.	8

*Code and data from this analysis are available at: https://github.com/yiliuc/Bicycle_Thefts_2023.git

1 Introduction

As the annual ranking by PeopleForBikes on 2022, Toronto has been named the 16th best major cycling city. In Toronto, there is about 300 kilometers on-street bike lanes and over 150 kilometers of off-street bike trails. Bicycle is becoming the main travel mode for the people living in Toronto. You can see people taking bicycle around each corner. However, as the number of bicycles increases each year, bicycles thefts had became a severe problem that needs to pay more attention on this. In Toronto, the bicycle thefts in Toronto surged during the past summer, nearly a 429 per cent increases. According to the neighborhood crime app by TPS, the downtown Toronto is place where the bicycle theft is mostly likely to occur, especially the U of T campus where the bicycle thefts is maximum.

Once the the bike is stolen, then the probability of getting it back is very low. However, even though the people lock their bikes, it may still be stolen as breaking the lock on the bicycle is a easy thing. You can even break the U lock with only a pen. Therefore, it is necessary for the bicycle riders to learn the feature influencing the possibility of bicycle to be stolen and estimate how likely their bicycle will be stolen. In this paper, I will use the historical bicycle thefts in Toronto to investigate the factors contributing to the bicycle thefts and tell the riders in what circumstances, the bicycle is most likely to be stolen.

In this paper, there will be four main parts: Data, Results, Discussion and Conclusion. In the Data section, I will introduce the data used in this paper and summarize the key variables. In addition to that, I will use the figures to show what kind of bicycles are most likely to be stolen as well as the places in the Results session. Moreover, in the Discussion session, I will emphasize the main findings and discuss how avoid the bicycle thefts. Finally, I will discuss the limitations and further conclude the entire paper in the Conclusion session.

2 Data

The data will be used in this paper was obtained from Open Data Toronto Portal by the library `opendatatoronto`. The name of the data file is **Bicycle Thefts**. The data contains above 30,000 rows and about 30 columns. The raw data was cleaned using the open source statistical programming language with packages 1, 2, 3. In the cleaned data, it contains 11 variables which can be divided into three main categories.

Table 1: Summary of Bicycle Thefts Occurrence Year

Occurrence Year	Count
2022	2433
2021	2623
2020	3253
2019	3089
2018	3407
2017	3301
2016	3241
2015	2708
2014	2503
2013	33
2012	2
2011	3
2010	2
2009	1

2.1 The Time of Bicycle Thefts Occurrence

The cleaned data contains very detailed information regarding the occurrence time of each theft. The smallest unit is the occurrence hour of each theft. The table below shows the summary table for the occurrence of bicycle thefts

Table 1 shows the number of bicycle thefts from 2009 to 2022. The difference between 2009 to 2013 and 2014 to 2022 indicate the missing data for latter years. From 2014 to 2022, we can see that number of bicycle thefts increased first from 2014 to 2015, then the bicycle thefts dropped gradually since 2018. The reported bicycle thefts for last year was about 2400.

2.2 The Premises Place Where the Bike was stolen

The second main category in the cleaned data is describing the places of bicycle thefts. There are six main premises type of bicycle thefts, which are: Commercial, Apartment, Other, House, Educational, Transit and Other. The premises type here represent the places where the bicycle was stolen.

Table 2 summarized the premises type of bicycle thefts. We can observe that premises types “Outside” and “Apartment” have more bicycle thefts than the rest places, meaning that people need to pay more attention on their bicycle if they park it near the condo or on the street.

Table 2: Summary of Bicycle Thefts Occurrence Year

Premise Type	Count
Outside	8266
Apartment	6434
House	3809
Commercial	3195
Other	3043
Educational	1288
Transit	564

Table 3: Summary of Bicycle Thefts Occurrence Year

Mean	SD	Min	Median	Max
1020.818	1654.558	1	700	120000

2.3 The Stolen Bicycles' Own Features

The last category of variables in the cleaned data is the features of each stolen bicycle. The features of each stolen bike include its brand, type, color and cost.

Table3 shows the bicycle price of stolen bicycles. The mean price of stolen bicycles is around 1000 dollars. However, it is surprising that the most expensive stolen bicycles worth 120,000 dollars.

3 Results

3.1 Compare the bicycle thefts in the recent six years

Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
i Please use `linewidth` instead.

The Figure 1 compares the monthly bicycle thefts from 2017 to 2022 with each line with different colors representing different years. It seems that the summer has significant higher bicycle thefts than winter in each year. In addition, it seems that the July is the most severe month of bicycle thefts in each year, ranged from 350 to 550 from the previous years. On the beginning and ending of each year, the number of bicycle thefts were always stable, keeping constant around 100.

Table 4: Summary of Bicycle Thefts Occurrence Year

Bike Colour	Count
BLK	8154
GRY	2480
BLU	2418
WHI	1996
None	1861
RED	1744
SIL	1191
GRN	729
ONG	489
PLE	469
DBL	432
YEL	355
LBL	342
DGR	279
LGR	274
OTH	230
BLKRED	214
TRQ	184
BRN	175
MRN	175
BLKBLU	148
BLKWHI	136
PNK	131
BGE	117
BLUWHI	106
BLKGRN	100
GLD	96
REDWHI	81
CRM	79
BLKGRY	70
REDBLK	63
BRZ	62
BLKSIL	60
BLKYEL	56
BLKONG	53
BLUBLK	50
WHIBLK	45
CPR	32
WHIBLU	31
WHIRED	30
GRNBLK	29
GRYRED	27
GRYBLK ⁵	26
TAN	26
GRYWHI	24
BLURED	21
BLKGLD	20
BLUSIL	20
GRYONG	20

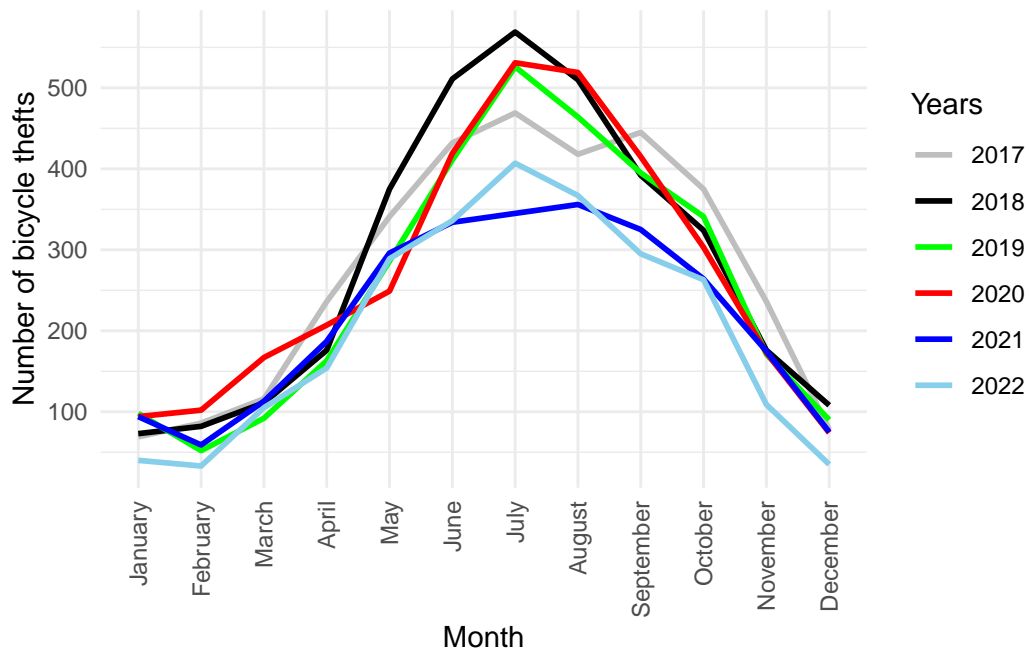


Figure 1: The number of bicycle thefts from 2017 to 2022 by month

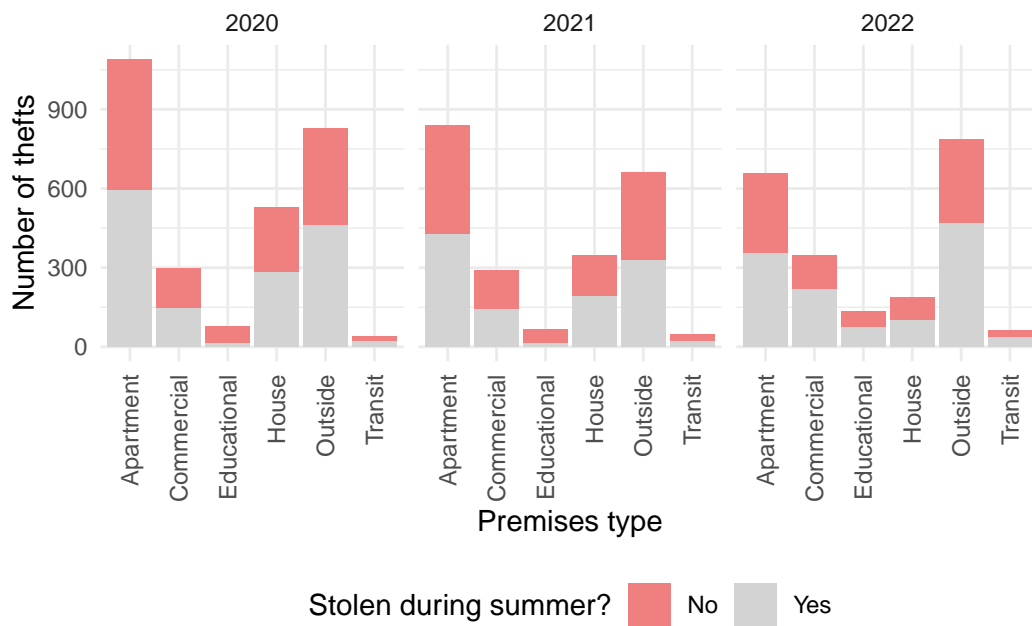


Figure 2: Summary of Bicycle Thefts Occurrence Year

3.2 Compare the premises types and probability to be stolen during summer.

Figure 3 compares the premises type of bicycle thefts in the recent three years and indicate the proportion of bicycles that are stolen during the summer. There are three facet indicating three different years. Each bar in each facet represent the number of bicycle thefts of a premises type. The filled color in each bar shows the proportion of bicycles that are stolen during the summer, with light coral meaning NO and light grey meaning Yes.

3.3 Compare the top 4 colours and top 5 bicycle types from bicycle thefts

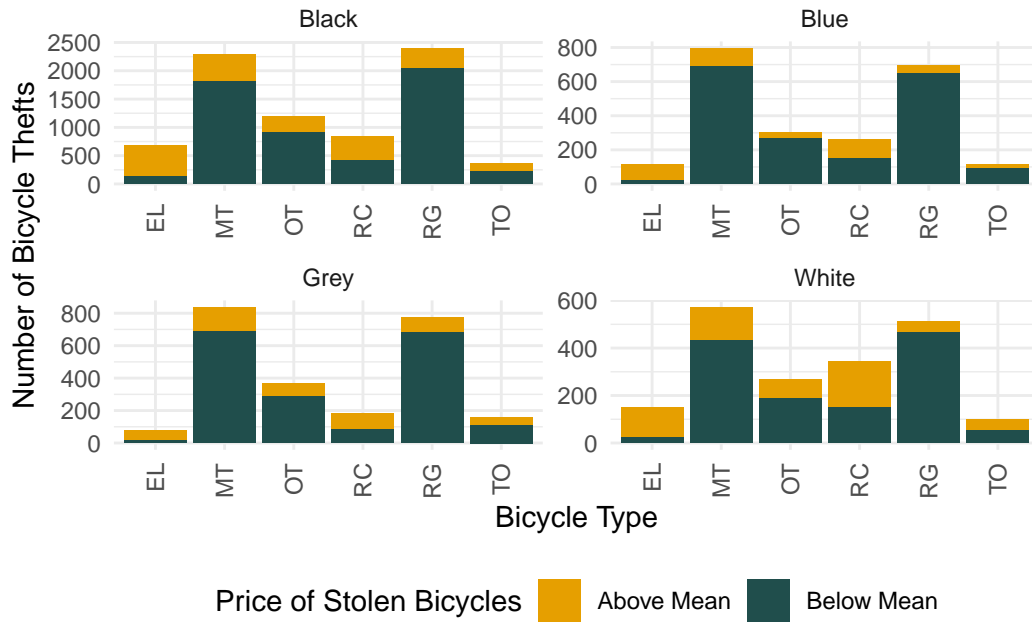


Figure 3: Summary of Bicycle Thefts Occurrence Year

Figure 3 compares the number of bicycle thefts for the top 4 colours and top 5 bicycle types. Each bar represent the corresponding number of bicycles of a bike type under different colors. In addition, each bar is filled by two colors orange and dark teal representing the proportion of bicycles that have a price above mean or below mean price.

4 Discussion

4.1 Repeating patterns of bicycle thefts: expected to decrease in 2023.

From Figure 1, it seems that the bicycle thefts happen most during the summer but less during the winter. According to Amanat, the number of bicycle thefts surged 429 per cent during

the summer month. The potential reason behind this pattern is that people are more likely to ride bicycle during the summer but due to the cold weather, less people take bicycle during winter and hence the occurrence of bicycle thefts decreases.

For the summer month, we can observe a “pendulum movement” in bicycle thefts. The number of thefts alternates annually, increasing one year and then decreasing the next. This trend repeats consistently from 2017 to 2022. Comparing 2022 to 2021, we see that the number of bicycle thefts during summer of 2022 started to increase again comparing to 2021. We may expect that the number of bicycle thefts decrease during the past summer. In addition, we can observe that most decrease is between 2020 and 2021, where there was about 200 decrease in the thefts during summer months. However, the potential reason behind this dramatic decrease is COVID, we can not use this to justify the number changed in bicycle thefts this year.

From 2017 to 2022, July is always the month with most bicycle thefts. Even though there will be an expected decrease in bicycle thefts in 2023 summer, the number of bicycle in July will still remain around 350, roughly 10 thefts per day. If you are a rider, then you really need to raise your awareness of bicycle thefts during July.

4.2 Increasing number of bicycle thefts on the outside and educational areas.

From Figure 2, we can observe that the distribution of number of bicycle thefts in each year is roughly same. However, we can observe that apartment was the premise type that have highest bicycle thefts. However, thefts from the apartment decreases gradually from 2020 to 2023, instead, the number of bicycle thefts on streets, roads or highways (Outside) seems to increase gradually and dominated in 2022. In addition, the number of thefts in educational areas like school or university also increased. This suggest to students who take bicycle to school also need to protect their bicycle more.

In addition, for all premises types in each year, we can observe that almost half was occurred during summer. Except for “Apartment”, we can observe that the proportion of bikes that were stolen during summer increases.

4.3 No evidence indicates that the cheaper/more expensive bicycles are less/more likely to be stolen.

From Figure 3, it seems that the black bikes is in favor of thieves. The amount of black bicycle lost is about the same as the other three colors combined. In terms of bicycle types, MT (mountain type) is bicycle type that is most likely to be stolen, followed by RG (Regular) which has almost same thefts with MT.

In terms of price of bicycles, it seems that except EL, majority of stolen bicycles for the rest bicycle types have a price below the mean price. For the most two “welcome” bicycle types

for thieves MT and RG, nearly eighty percent stolen bicycles for each color has a price lower than the mean price. This indicates that thieves will not evaluate the price of bicycles before they steal it. They pick a random bicycle to steal as long as the time is right.

5 Conclusion

This paper investigate the trends of bicycle thefts in Toronto in the perspectives of year trend, premises type and bikes' own features. By this paper, it shows that there was a reboundance of bicycle thefts in 2022 and it is expected that the occurrence of bicycle thefts in 2023 will decrease according to the historical data. In addition, the summer months are the months in a year that the bicycle thefts occurs most, nearly half bicycle thefts that are stolen during the summer. Moreover, the apartment such as condo and the outside areas such as streets or highways are the two most likely place to occur bicycle thefts. Besides, bicycle thefts is on the rise on the educational areas such as schools or universities have increase. Students should protect their bicycles.

In terms of bicycle own features. Even though one third of stolen bicycles are black, we can not say that black bicycles are more likely to be lost as number of black bicycles is high. However, one important trend is that the price of a bike will not influence its likelihood to be stolen. That said, a cheap and an expensive bike has the same probability to be stolen. Don't take chances that your bike won't be stolen even though it is cheap.

However, this paper has limitations. The data for this paper comes from `opendatatoronto` which is not a up-to-date data. The data of bicycle thefts for 2023 is missing. This is crucial as we can verify the "pendulum movement" I discussed once we have the data of bicycle thefts in 2023. The second limitation is the lack of information of neighborhood for each bicycle theft. If we have the neighborhood information of bicycle thefts, then this paper will be more informative as it can tell the people living in the certain neighborhood to take more care of their bicycles. This will more helpful for them to avoid bicycle thefts. The last one is the accuracy of data. Some people may not report the bicycle thefts to the police. Instead, they only report to the insurance company to get money from the converage. In that case, the actual bicycle thefts may be even more and makes it hard to track the patterns.

In the future study, instead of using data from open data toronto, using the data from toronto police service open protal or from the insurance companies regarding the bicycle thefts might be a better option. This will make the data more accurate and help us to find the true patterns of bicycle thefts in Toronto. In addition, the data from TPS contains the information of neighborhoods regarding the bicycle thefts. With the information of neighborhood, it can have better effects.

6 References