

Hate Crime in Toronto*

Deep Investigation in Hate Crime

Jingtao Luo

27 September 2024

This paper analyzed the Hate Crime Open Data, which is free and available to the public on Open Data Toronto. The paper investigated distribution of hate crime in Toronto based on year, reason and neighbourhood. In summary, the hate crime has increased in last 6 years, and the most often reason is religion bias, and the Yonge-Bay Corridor has the largest amount of hate crimes. These findings can help the government understand hate crime better, and thus protect the citizens.

1 Introduction

Toronto, often celebrated as one of the most multicultural cities in the world, stands as a beacon of diversity and inclusion. With over half of its population identifying as a visible minority and more than 180 languages spoken, the city reflects a rich tapestry of cultural, ethnic, and religious backgrounds. It has long been regarded as a model for successful urban integration, where immigrants and marginalized communities contribute to the city's vibrant social, economic, and cultural fabric. However, beneath this inclusive narrative, Toronto also faces challenges in addressing the rising threat of hate crimes.

According to government of Canada (Ndegwa and McDonald 2023), hate crime is defined as criminal acts done by a person who is motivated by an extreme bias or hatred towards a particular social group. Research studies show that hate crimes cause “disproportionate harm” to individual victims as well as other members of the community belonging to the targeted social group. These crimes lead to not only physical but also psycho-emotional hurt. For example, if the assault occurs because the victim's race, then victim may live in fear of being victimized again for the rest of his life because a person's race can never be changed.

*Code and data are available at: https://github.com/JerryLuoJT/Hate_Crime_Toronto

Despite Toronto’s inclusive policies and progressive reputation, the city has not been immune to the global rise in hate-fueled violence. In recent years, incidents targeting specific communities—including Muslim, Jewish, Black, and LGBTQ+ populations—have underscored the persistent challenges that marginalized groups face. While hate crimes constitute only a fraction of overall crime, their impacts are disproportionately harmful, not only to direct victims but also to the broader communities they belong to, fostering fear, division, and alienation.

This paper aims to analyze the hate crimes in the following perspectives: when, where and why. When means the amount of crimes happened in each year. Where means the neighborhood of the hate crime occurred. And why is the reason that triggered this hate crime.

The Data Section 2 of this paper explores the `opendatatoronto` data, with tools in R and methods used to analyze the data and some observations from the new dataset. The Results Section 3 introduces more observations found from the data analysis by visualizing the data. The Discussion Section 4 ties back the findings with the situation in real world. Lastly, the Conclusion Section 5 summarizes key findings and gives insight for further study.

2 Data

All relevant data can be found in the Open Data Toronto portal, and extracted using the `opendatatoronto` library for R (Gelfand 2022). In particular, the original dataset includes the information about each hate crime, investigated by the Hate Crime Unit by reported date since 2018, and its latest update was on 17 September, 2024. The information includes each crime’s offence number, the exact date of occurrence, the exact date of report, the police division where offence happened, the reason caused this hate crime, the neighborhood of such crime occurred, and if there is people got arrested.

2.1 Data Tools

The data was extracted, cleaned and analysed using R (R Core Team 2023), with functions from `tidyverse` (Wickham et al. 2019), `ggplot2` (Wickham 2016), `dplyr` (Wickham et al. 2023), `readr` (Wickham, Hester, and Bryan 2023), `opendatatoronto` (Gelfand 2022), `knitr` (Xie 2014), and `here` (Müller 2020).

2.2 Overview of Dataset

The selected dataset has selected some of the variables that will be discussed later in the paper. The selected variables are id, year of crime occurred; reasons of hate crime: age, mental or physical disability, race, ethnicity, language, religion, sexual orientation, gender; and the neighborhood of this crime occurred. See Table 1 for an overview of the data.

ID	Year	Bias Type	Neighbourhood
1	2018	Religion	Forest Hill North (102)
2	2018	Religion	Mount Pleasant East (99)
3	2018	Race	Annex (95)
4	2018	Religion	Rustic (28)
5	2018	Religion	Bay-Cloverhill (169)
6	2018	Religion	Oakdale-Beverley Heights (154)

Table 1: Sample of Hate Crime Open Data

The column “Bias Type” is not on the original dataset, instead, the original one has variables for each type of bias, I use the “Bias Type” variable which stores the name of bias(data type: character) so this is easier for later calculation. In the original dataset, there are one more bias called multiple bias, which indicates that the crime is caused by more than one bias. Multiple bias may cause over-counting in other bias, so for easier calculation I didn’t select crimes with multiple bias. Also, in the Neighbourhood variable, I find that some of the crimes didn’t have information about neighbourhood, so I removed crime with “NSA” in the ‘neighbourhood_158’ column.

2.3 Distribution of Hate Crimes per Year

The average number of hate crimes per year is 193. However, the range of crime is very wide: from 111 in 2018 up to 329 in 2023.

Year	Total Crimes
2018	111
2019	104
2020	186
2021	227
2022	204
2023	329

Table 2: Count of Hate Crime by Year

2.4 Distribution of Hate Crimes by Bias

Here we want to count the number of hate crime based on the bias that caused it. In the raw dataset, as I have mentioned before, we have 8 bias. However, after we delete the crimes with

multiple bias or don't have the neighborhood information, we only have 6 bias in the dataset: age bias and disability bias are not in the cleaned dataset anymore.

Bias Type	Total Crimes
Ethnicity	133
Gender	49
Language	3
Race	308
Religion	516
Sex	152

Table 3: Count of Hate Crime by Bias

The average number of crimes for each bias is 193.5. The range of number of crimes in different bias vary largely, from 3 to 516, this make me consider whether we have outlier. The lower bound and upper bound I calculated(See 'Table 4') indicating no outlier in this dataset.

Q1	Q3	IQR	Lower_Bound	Upper_Bound
70	269	199	-228.5	567.5

Table 4: Bounds of Table 3's Data

2.5 Distribution of Hate Crime by Neighbourhood

I am curious about which neighbourhood in Toronto has the most hate crimes, so I count the data by variable 'neighbourhood_158', which store the neighbourhood name of each crime based on City of Toronto's new 158 neighbourhood structure. See 'Table 5' for a sample of dataset.

Neighbourhood	Total Crimes
Agincourt North (129)	3
Agincourt South-Malvern West (128)	6
Annex (95)	35
Avondale (153)	7
Banbury-Don Mills (42)	7
Bathurst Manor (34)	19
Bay-Cloverhill (169)	12
Bayview Village (52)	4
Bayview Woods-Steeles (49)	1

Neighbourhood	Total Crimes
Bedford Park-Nortown (39)	16

Table 5: Sample of Count of Hate Crime by Neighbourhood

3 Results

The number of hate crime incidents has increased almost every year except 2019 and 2022. Last year's hate crimes surged to 329, nearly three times as much as the amount in 2018. The scatterplot intuitively shows a rapid growth rate(See 'Figure 1'). From the overall trend, the number of hate crimes is increasing every year.

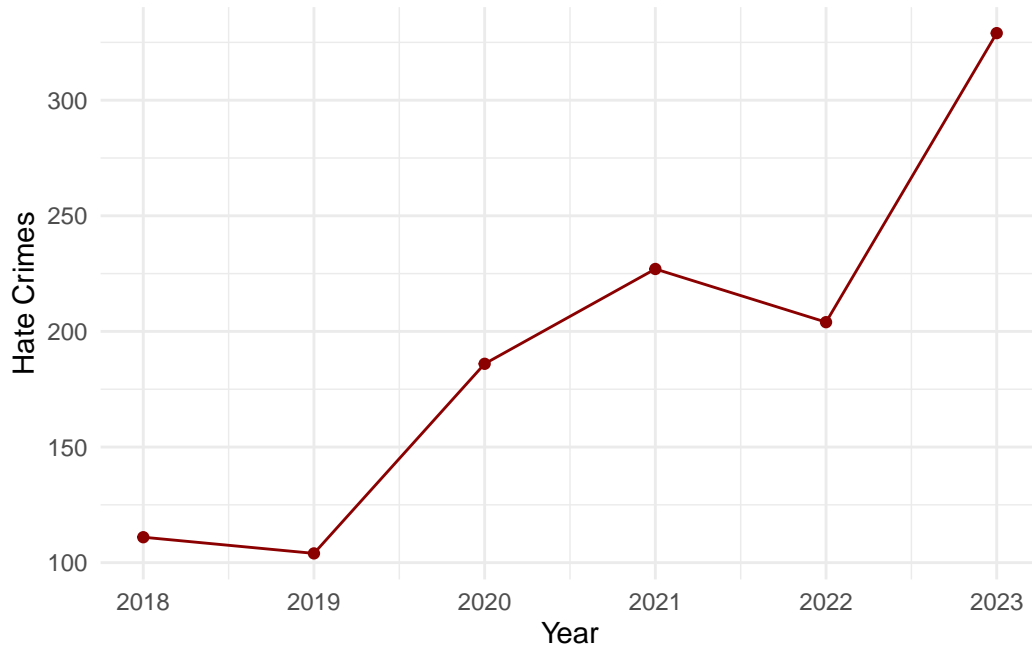


Figure 1: Number of Hate Crimes every year

The average of hate crime by each bias type is 193.5 and the median is 142.5. We have showed that there is no outlier based on the interquantile range test. From the graph (see Figure 2) we can say most of the causes are ethnicity, sex, race and religion. More particularly, religion bias is the number one cause for hate crime, and is way more than other bias.

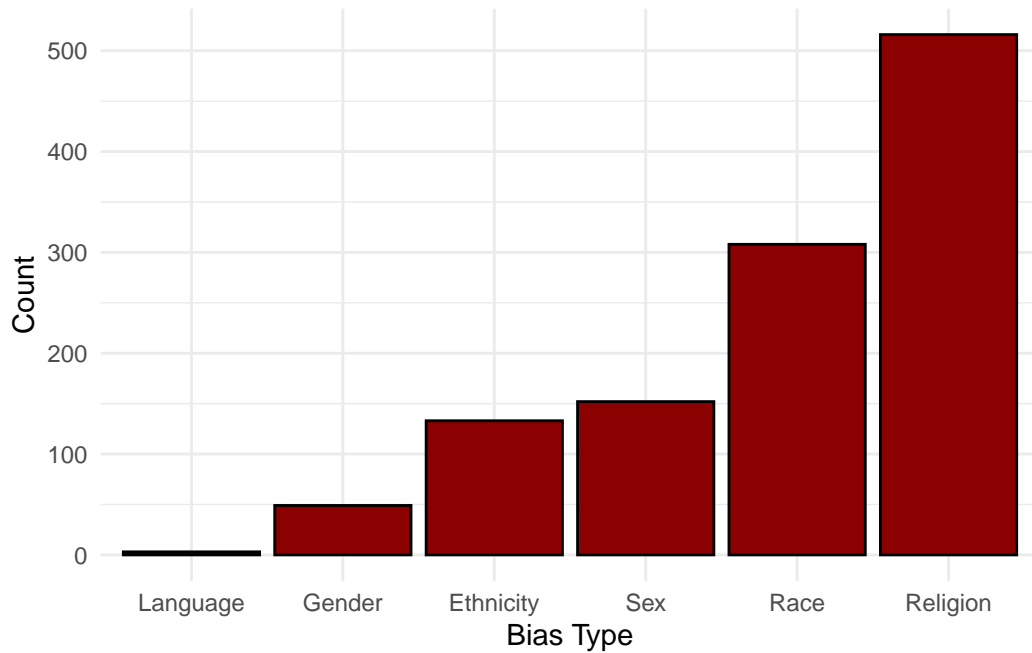


Figure 2: Hate Crime by Bias Type

The distribution by neighbourhood tells more stories. Since there are 158 neighbourhoods in Toronto, there are about 158 groups if I want to group by variable 'neighbourhood_158', which is too messy for a bar plot. So I use histogram to show the relationship(See Figure 3). The histogram is right-skewed and unimodal. From the graph we can find there are 69 neighbourhood that are have hate crime less than or equal to 5 from 2018 to 2023. There are 6 neighbourhoods have more than 25 crimes.

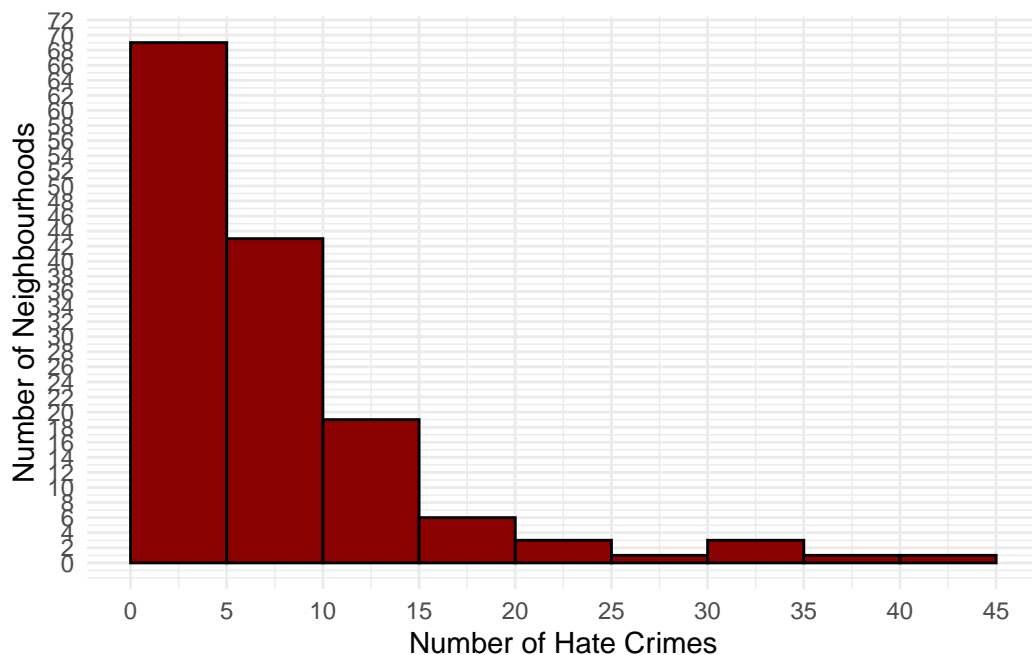


Figure 3: Histogram of Hate Crime by Neighbourhood

4 Discussion

These tables and figures give readers a visual and organized understanding of the development and causes of hate crimes. ‘Figure 2’ shows that most of the hate crimes are caused by religion bias, race bias and sex bias. ‘Figure 3’ shows that there are 6 neighbourhoods out of 158 are considered as dangerous neighbourhood while the rests have more than 25 hate crimes for recent 6 years. By revisiting the dataset ‘hc_by_nbh’ and sorting the dataset, I can find those 6 neighbourhoods are: Yonge-Bay Corridor, Church-Wellesley, Annex, Downtown Yonge East, York University Heights and Moss Park.

Algorithm-driven content recommendation systems often exacerbate polarization, leading users into place where different perspectives are diminished, and hate speech proliferates. This growing extremism can incite individuals or groups to escalate from online rhetoric to offline violence. As social media platforms continue to globalize and expand their user base, hate crimes may become more frequent and harder to control, driven by the influence of these platforms. In the near future, increasing in hate crime in Toronto is inevitable as more and more people of different nationalities and beliefs flock to this city, which is also showed in ‘Figure 1’. So, the government should try its best to prevent hate crime.

The analyzed data and figures can help the government of Toronto understand the composition of hate crimes and the main areas where they occur. For example, the police can increase the

number of patrols in dangerous areas and prevent crimes in time. The propaganda department can support views such as racial equality and religious freedom on social media and educate citizens to respect everyone.

5 Conclusion

This paper investigated the hate crimes happened in Toronto from 2018 to 2023. The paper discussed the hate crimes' distribution in three aspects: year of occurrence, reason of occurrence, place of occurrence. It demonstrates that, the number of hate crimes is increasing in recent years; the top-three reason for hate crimes are religion bias, race bias and sex bias; among the 158 neighbourhoods in Toronto, 69 of them had less than 5 hate crimes in last 6 years and 6 of them had more than 25 hate crimes in the same period. In future study, I can analyze each bias deeper e.g. which religion experienced most hate crime so that the research can be more detailed.

6 LLMs

Statement on LLM usage: no LLMs were used in the making of this paper.

References

- Gelfand, Sharla. 2022. *Opendatatoronto: Access the City of Toronto Open Data Portal*. <https://CRAN.R-project.org/package=opendatatoronto>.
- Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. <https://CRAN.R-project.org/package=here>.
- Ndegwa, Anna, and Susan McDonald. 2023. “Hate Crimes in Canada.” <https://www.justice.gc.ca/eng/rp-pr/cj-jp/victim/rd16-rr16/p1.html>.
- R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- 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 Grolemond, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.
- Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. <https://CRAN.R-project.org/package=dplyr>.
- Wickham, Hadley, Jim Hester, and Jennifer Bryan. 2023. *Readr: Read Rectangular Text Data*. <https://CRAN.R-project.org/package=readr>.
- Xie, Yihui. 2014. “Knitr: A Comprehensive Tool for Reproducible Research in R.” In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC. <http://www.crcpress.com/product/isbn/9781466561595>.