Toronto's Uptick in Hate Crime: Time and Group Trends*

An Exploratory Analysis of Hate Crime in Toronto

Justin Klip

September 24, 2024

Using data from Open Data Toronto, this paper the tracks the evolution of hate crime in Toronto from 2018 to the end of 2023. Third sentence. Fourth sentence.

1 Introduction

In 1988 the Canadian Parliament passed the Canadian Multiculturalism Act, an act that served to recognize and preserve Canada's multicultural heritage Canada (2024). Toronto serves as a shining example of Canada's multiculturalism, with over half of the population of Toronto being foreign-born (McNeil 2023). Yet, despite Toronto's reputation as an international city with large widespread cultural heterogeneity, hate crimes have seen an uptick in recent times. Hate crimes in Toronto have increased nearly 55% in 2024 in comparison to 2023. A hate crime is defined as a "criminal act done by a person who is motivated by an extreme bias or hatred towards a particular social group" (Government of Canada 2023). Acts like these are fundamentally at odds with the sentiment expressed by the Canadian Multiculturalism Act and raise questions about Toronto's history with hate crimes and its multicultural status. In this paper, I explore the dynamics of Toronto's Hate Crimes from 2018 to 2023, before the current uptick. In particular, I describe how the types of hate crimes, the types of victims, and the quantity of victims have changed over time.

The rest of the paper is split into three main parts Section 2, Section 3, and **?@sec-limitations**. The Section 2 section describes how the data is generated in the **?@sec-measurement** section. It then moves to the **?@sec-cleaningandpackages** which describes the data cleaning process and libraries used. The **?@sec-observations** section demonstrates what a typical observation in the data set looks like, and the **?@sec-plotsandanalysis** section displays the data to describe how it has varied over time and group.

*Code and data are available at: https://github.com/justinklip/hate_crime_toronto_paper.

2 Data

2.1 Measurement {##sec-measurement}

To run my analysis, data was collected from Open Data Toronto's Open Data Portal, through the R library opendatatoronto (Gelfand 2022). The data source used is their Hate Crimes Open Data data set (**open-data-set?**). This data of about 1400 observations has every crime that has been classified as a 'hate crime' by Toronto Police from the beginning of 2018 to the end of 2023. The data is generated as follows: a crime is committed in Toronto, and generally within a month that crime is reported by the victim. If a hate-motivated offence is suspected, the investigation is then led by a divisional investigator of the Hate Crime Unit (HCU) or by the HCU itself (Toronto Police Service). If hate-motivated offences are found, then this case will be added next year to the data set. Details of the crime are documented – including what social group they were biased against – these bias variables include ethnicity, race, gender, age, language, religion, and sexual orientation. Date and time (of both the crime and the reporting of it), offence committed, neighborhood, and place (such as apartments, malls, parks) were also recorded. Every unique case number was also attributed to each observation in the data set. There were almost no empty observations, except for a few cases where place data was missing. (**cleaning-packages?**) discusses this missing data further.

2.2 Cleaning and Packages

To download, clean, and analyze the data, the statistical software R was used (R Core Team 2023). Other packages such as lubridate (Grolemund and Wickham 2011), tidyverse (Wickham et al. 2019), dplyr (Wickham et al. 2023), janitor (Firke 2023) aided in this process. In the data cleaning process, about 30 missing observations were found. These 30 observations were all under the variable of location type, demonstrating the report writer had a hard time describing the location of the incident. To deal with this, all the location was set to unknown which can still be interpreted in the graphs. ## Observations {#sec-observations}

A typical observation of the cleaned data has the all data on the

2.3 Data Analysis {sec-data-analysis}

?@fig-1 documents the evolution of hate crimes in Toronto over time. It plots the monthly quantity of hate crimes committed for all months from January 2018 to December 2023.It also plots the 12 month moving average the demonstrate the trend over time. As seen, number of hate crimes go from averaging 12 a month in a given year, to nearly 30 by the end of the dataset. The catalyst for this upward trend seems to start during the early Covid-19 period in 2020, demonstrating that COVID-19 could be a possible explanatory factor this uptick. Interestingly, hate crimes still remained rather elevated post lockdown, and in 2023 climbed

to even higher levels in comparison to the 2020 period. May 2023 had the highest amount of hate crimes committed in Toronto since the start of this data, with 44 being committed.

```
#| label: fig-1
#| echo: false
#| fig.align: center
#| fig.width: 10
#| fig.height: 7

# Extract year and month from the 'occurrence_date' column
hate_crime_analysis_data <- hate_crime_analysis_data %>%
  mutate(year_month = floor_date(occurrence_date, "month"))

# Group by year and month, then count observations
monthly_counts <- hate_crime_analysis_data %>%
  group_by(year_month) %>%
  summarise(count = n())

# Calculate a 12-month (yearly) moving average
monthly_counts <- monthly_counts %>%
  mutate(moving_avg = rollmean(count, 12, fill = NA, align = "right"))

# Reshape the data for easier plotting of multiple lines
monthly_counts_long <- monthly_counts %>%
  pivot_longer(cols = c(count, moving_avg),
               names_to = "Line_Type",
               values_to = "Value")

# Plot with legend for monthly total and moving average
ggplot() +
  geom_line(data = monthly_counts_long, aes(x = year_month, y = Value, color = Line_Type, li
  geom_point(data = monthly_counts, aes(x = year_month, y = count), color = "black") +  # Po
  scale_color_manual(values = c("count" = "black", "moving_avg" = "blue"),
                     labels = c("Monthly Total", "12 Month Moving Average")) +  # Define colo
  scale_linetype_manual(values = c("count" = "solid", "moving_avg" = "dashed"),
                        labels = c("Monthly Total", "12 Month Moving Average")) +  # Define
  labs(
    title = "Number of Hate Crimes Per Month in Toronto",
    x = "Month",
    y = "Reported Hate Crimes",
    color = "Legend",
    linetype = "Legend",
    caption = "Figure 1: Monthly Hate Crime Data from 2018 to 2024"
```

```
  ) +
  theme_minimal() +
  scale_x_date(date_labels = "%b %Y", date_breaks = "3 months") +
  theme(
    axis.text.x = element_text(angle = 90, hjust = 1),
    plot.caption = element_text(hjust = 0.5)  # Center the caption
  )
```

```
Warning: Removed 11 rows containing missing values or values outside the scale range
(`geom_line()`).
```

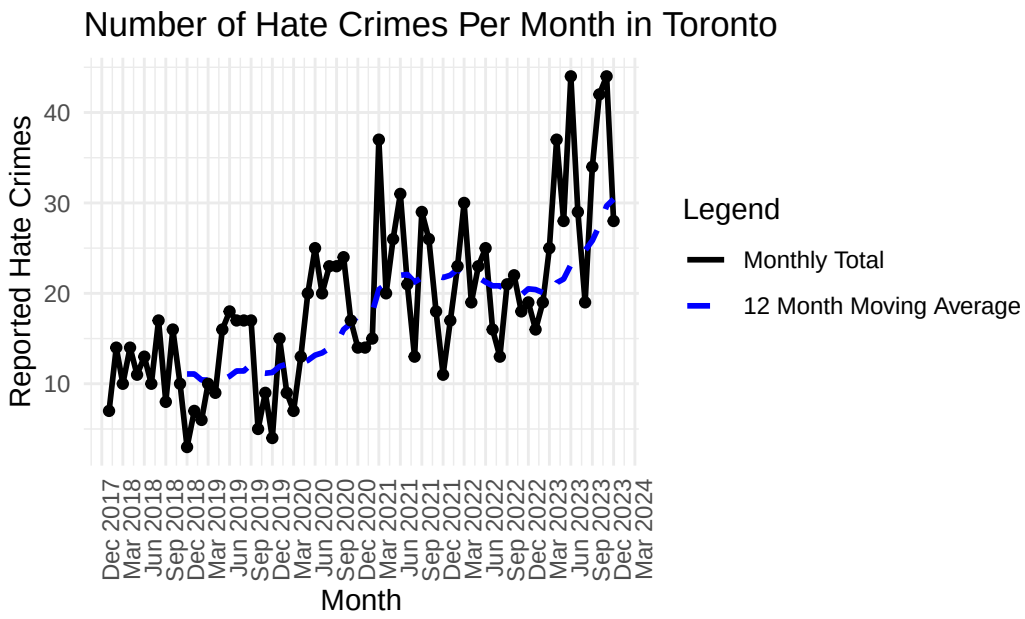


Figure 1: Monthly Hate Crime Data from 2018 to 2024

```
# Load required packages
#| label: fig-2
#| fig-cap: Types of Hate Crimes Committed, Excluding Multiple Bias Hate Crimes.
#| echo: false

# Step 1: Filter and reshape the data to long format
long_data <- hate_crime_analysis_data %>%
  filter(multiple_bias == "NO") %>%  # Exclude entries with multiple bias
  select(occurrence_date, race_bias, ethnicity_bias, religion_bias, language_bias, age_bias,
  pivot_longer(cols = c(race_bias, ethnicity_bias, religion_bias, language_bias, age_bias, se
```

```
                names_to = "bias_type",
                values_to = "bias_value") %>%
  filter(bias_value != "None", bias_value != "NO")  # Exclude entries with "None" or NA

# Step 2: Modify bias_type to just represent the category
long_data <- long_data %>%
  mutate(bias_type = case_when(
    bias_type == "race_bias" ~ "Race Bias",
    bias_type == "ethnicity_bias" ~ "Ethnicity Bias",
    bias_type == "religion_bias" ~ "Religion Bias",
    bias_type == "language_bias" ~ "Language Bias",
    bias_type == "age_bias" ~ "Age Bias",
    bias_type == "sexual_orientation_bias" ~ "Sexual Orientation Bias",
    TRUE ~ bias_type  # Fallback case, should not be reached
  ))

# Step 3: Count hate crimes by year and bias type
hate_crime_counts <- long_data %>%
  mutate(year = year(occurrence_date)) %>%  # Extract year
  group_by(year, bias_type) %>%             # Group by year and bias type
  summarise(count = n(), .groups = "drop")  # Count entries

# Step 4: Create the segmented bar chart
ggplot(hate_crime_counts, aes(x = factor(year), y = count, fill = bias_type)) +
  geom_bar(stat = "identity", position = "stack") +  # Use 'identity' for raw counts
  labs(title = "Hate Crimes by Bias Type and Year (Excluding Multiple Bias)",
       x = "Year",
       y = "Total Number of Hate Crimes",
       fill = "Bias Type") +
  theme_minimal() +
  scale_fill_brewer(palette = "Set1") +  # Optional: Set color palette
  theme(axis.text.x = element_text(angle = 45, hjust = 1))  # Rotate x-axis text for better r
```

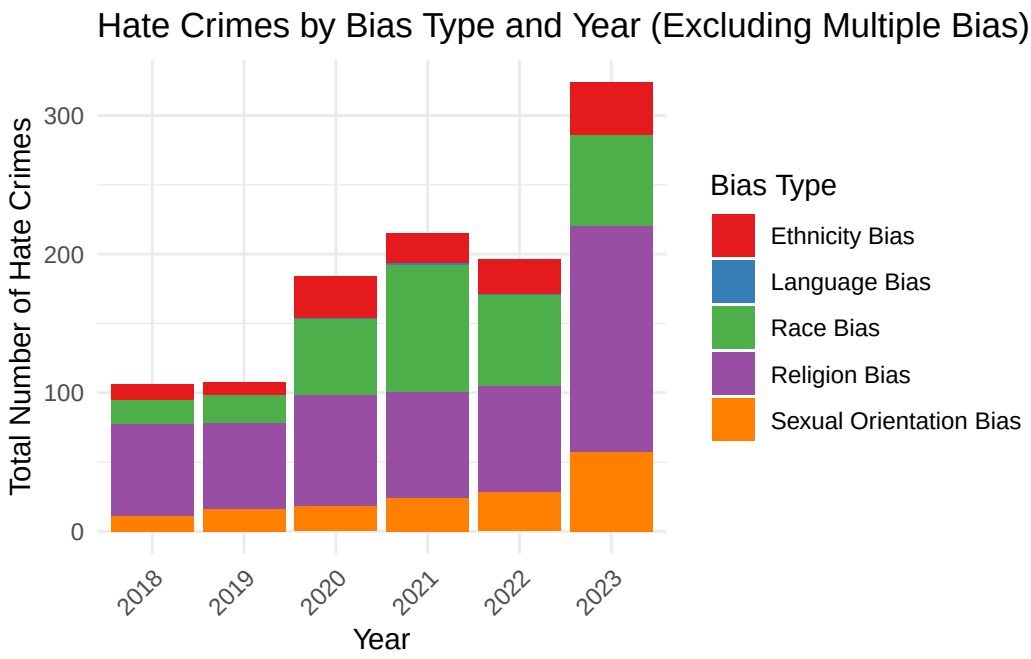


Figure 1 plots the year breakdown of race-motivated hate crimes. As seen, Black people are clearly the most targeted group for hate crimes in Toronto. They made up the largest proportion of hate crime victims in every single year, and were almost the only victims of race targeted hate crimes in 2018. Another significant subgroup were East/Southeast Asians. They are particularly notable in 2021 onward, suggesting that COVID may have played a role in East/Southeast Asians becoming victims of hate crimes; this particularly makes sense given the rhetoric at the time.

?@fig-3b plots the num

```
#| label: fig-3b
#| fig-cap: Yearly Breakdown of Religion-Motivated Hate Crimes
#| echo: false

ggplot(hate_crime_analysis_data %>% filter(religion_bias != "None"), aes(x = factor(year(occu
  geom_bar(position = "stack") +  # Stack the bars to show total counts by year
  labs(
    title = "Hate Crimes by Religion and Year",
    x = "Year",
    y = "Total Hate Crimes",
    fill = "Religion Bias"
  ) +
  theme_minimal()
```

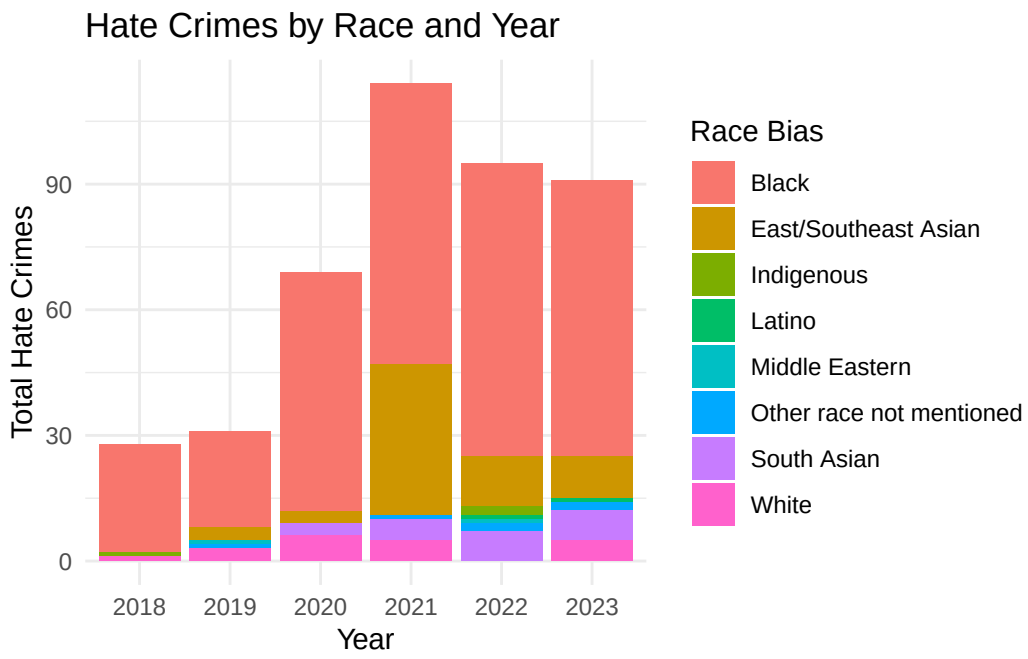
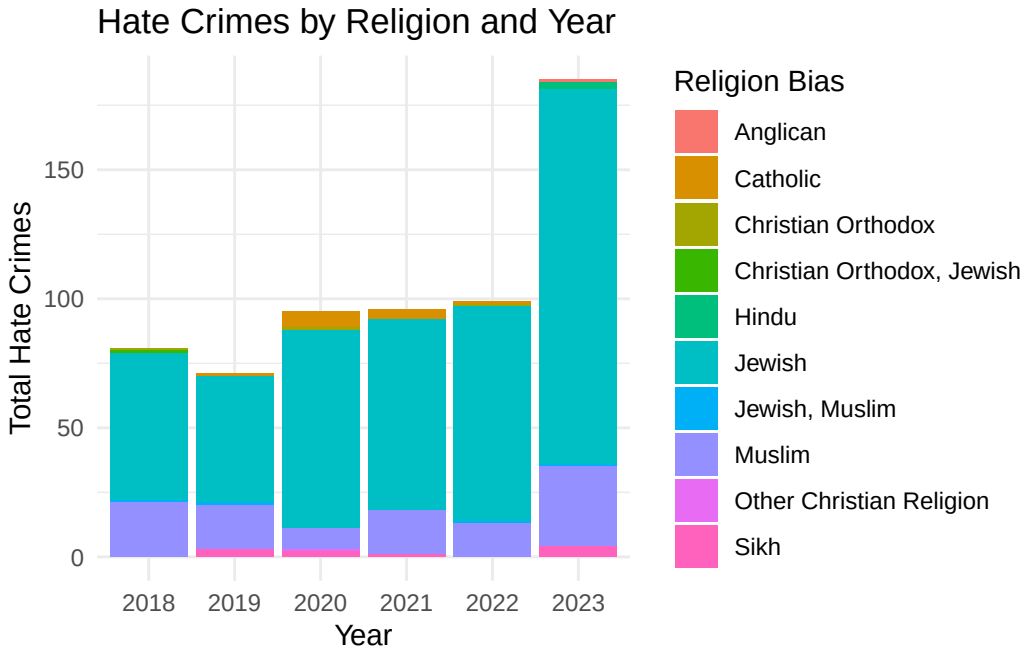


Figure 1: Yearly Breakdown of Race-Motivated Hate Crimes



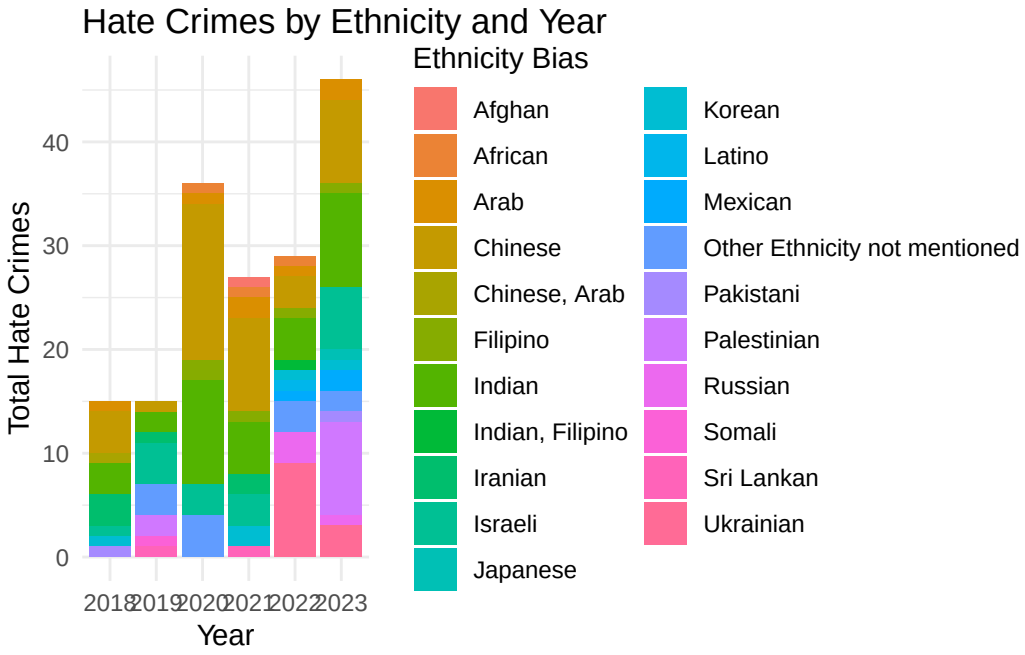


Figure 2: Yearly Breakdown of Ethnicity-Motivated Hate Crimes

3 Discussion

3.1 First discussion point

If my paper were 10 pages, then should be be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

3.2 Second discussion point

3.3 Third discussion point

3.4 Limitations

Weaknesses and next steps should also be included.

Appendix

.1 Diagnostics

References

Canada, Government of. 2024. "About the Canadian Multiculturalism Act." https://www.canada.ca/en/canadian-heritage/services/about-multiculturalism-anti-racism/about-act.html.

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://CRAN.R-project.org/package=janitor.

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

Government of Canada, Department of Justice. 2023. "Victims of Crime Research Digest No. 16." https://www.justice.gc.ca/eng/rp-pr/cj-jp/victim/rd16-rr16/p1.html.

Grolemund, Garrett, and Hadley Wickham. 2011. "Dates and Times Made Easy with lubridate." *Journal of Statistical Software* 40 (3): 1–25. https://www.jstatsoft.org/v40/i03/.

McNeil, Jaclyn. 2023. "How Multicultural Is the Toronto Region? Let Us Count the Ways… Again 2024." https://torontoglobal.ca/media-center/how-multicultural-is-the-toronto-region-let-us-count-the-ways-again/.

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, 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.

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.