My title*

My subtitle if needed

Hritik Shukla

January 22, 2024

First sentence. Second sentence. Third sentence. Fourth sentence.

1 Introduction

You can and should cross-reference sections and sub-sections.

The remainder of this paper is structured as follows. Section 2….

2 Data

The analysis for this project was done using R programming language (R Core Team 2022), and multiple libraries were used within R to help with tasks such as plotting graphs, data cleaning and manipulation. In particular, the opendatatoronto library (Gelfand 2022) was used for acquiring the dataset off of Open Data Toronto's servers, janitor (Firke 2023), tidyverse (Wickham et al. 2019) and dplyr (Wickham et al. 2023) were used for cleaning, manipulating and testing the dataset, and ggplot2 (Wickham 2016) was used for creating plots and graphs shown in the analysis.

2.1 Data Features

The data contains crime data for all 158 neighborhoods of Toronto, from the year 2014 to the year 2023. The data includes counts for individual categories of crimes, such as assault, auto theft, break and entering, robbery, homicide, theft over ($5000), shooting and firearm discharges. The data also includes the crime rate for each of these categories, calculated as crimes committed per 100,000 population. These crime rates were calculated using population estimates for each neighborhood in 2023, which were provided by Environics Analytics.

*Code and data are available at: https://github.com/hritikshuklas/toronto-neighbourhood-crimes

2.2 Data Collection

3 Results

4 Discussion

4.1 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.

4.2 Second discussion point

4.3 Third discussion point

4.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

References

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.

R Core Team. 2022. *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 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.