My title*

Reem Abughazaleh

06 February 2022

Abstract

For many years there have been disparities in the treatment of people in minority communities by actors in the criminal justice system, particularly police officers. Data from the Toronto Police Service Annual Statistics Report was obtained and analyzed. It was found that neighborhoods with higher percentages of people of color had higher counts of criminal incidents and lower rates of charges being cleared. This provides further support for increased funding for community and social services in minority communities.

Introduction

In recent years the public eye has been on the practices of local police departments. In particular the public has been critical of how police treat minority communities. It has been documented that there are higher crime rates in minority communities than majority white communities in cities all over North America (Fagan and Meares 2008). Additionally, disparities in the treatment of minority communities have also been cited throughout the justice system, not just in policing (Rapping 2013). After the death of George Floyd in 2020, there has been increased demand for the implementation of public policy to demilitarize the police, decrease funding for the police and reallocation of funds to other community and social services (Weichselbaum and Lewis 2020).

Data from the Toronto Police Service Annual Statistics Report (ASR), which provides a comprehensive overview of police related statistics including reported crimes, was used for this report (Toronto Police Service 2020). The data was analyzed by a divisional basis, i.e. geographical location. Additionally, I analyzed the amount of charges that were cleared, For a charge to be "cleared" it means that "An offense can be considered cleared when a charge is laid, recommended or the person(s) who committed the offense has been identified and no charge has been laid for some reason" (Service 2020). It was found that minority communities had higher counts of charges, and a lower rate of charges cleared.

This provides further support for the need for public policies that support minority communities. Funds need to be redistributed to social services that can act as a preventative measure, instead of further arming the police to respond to the crime in minority communities.

The reminder of the paper: the Data section discusses the data used for this report.

Data

The data in the ASR has been aggregated and reported by the Toronto Police Service. The ASR is publicly available and can be accessed through the City of Toronto Open Data Portal. According to the Toronto Open portal "this data includes all crimes reported to the Toronto Police Service, including, but not limited to, those that may have been deemed unfounded after investigation, those that may have occurred outside the City of Toronto limits, or those that have no verified location" (Toronto Police Service 2020). The ASR contains data about the number of charges laid, how many of those charges were cleared, what category and subtype of crime was committed and which division the particular charges were reported.

*Code and data are available at: LINK.

The database used for this report was obtained in csv format from the City of Toronto Open Data Portal using the R package opendatatoronto (Gelfand 2020). The data is of all reported offenses from 2014-2020 and has been aggregated by year, category, subtype and geographic division. The data includes 2369 observations and 10 variables; id, index, Reported Year, GeoDivision, Category, Subtype, Count, CountCleared, ObjectId and geometry. The dataset was processed and analyzed in R (R Core Team 2020). The additional packages used are tidyverse (Wickham et al. 2019) and dplyr (Wickham et al. 2021) packages. Graphs and tables were created with ggplot2 (Wickham 2016) and knitr (Xie 2017). A sample view of the data can be seen below.

```
## # A tibble: 2,369 x 9
##    ReportedYear GeoDivision Category      Subtype   Count_ CountCleared ObjectId
##           <int> <chr>       <chr>         <chr>      <int>        <int>    <int>
##  1         2014 D11         Controlled D~ Other        201          195        1
##  2         2014 D11         Crimes Again~ Auto The~    119           42        2
##  3         2014 D11         Crimes Again~ Break & ~     85           37        3
##  4         2014 D11         Crimes Again~ Break & ~     58           18        4
##  5         2014 D11         Crimes Again~ Break & ~     89           34        5
##  6         2014 D11         Crimes Again~ Break & ~     23            7        6
##  7         2014 D11         Crimes Again~ Fraud        232           83        7
##  8         2014 D11         Crimes Again~ Other        628          230        8
##  9         2014 D11         Crimes Again~ Theft Ov~     36           12        9
## 10         2014 D11         Crimes Again~ Theft Un~   1774          790       10
## # ... with 2,359 more rows, and 2 more variables: geometry <chr>, Index_ <lgl>
```

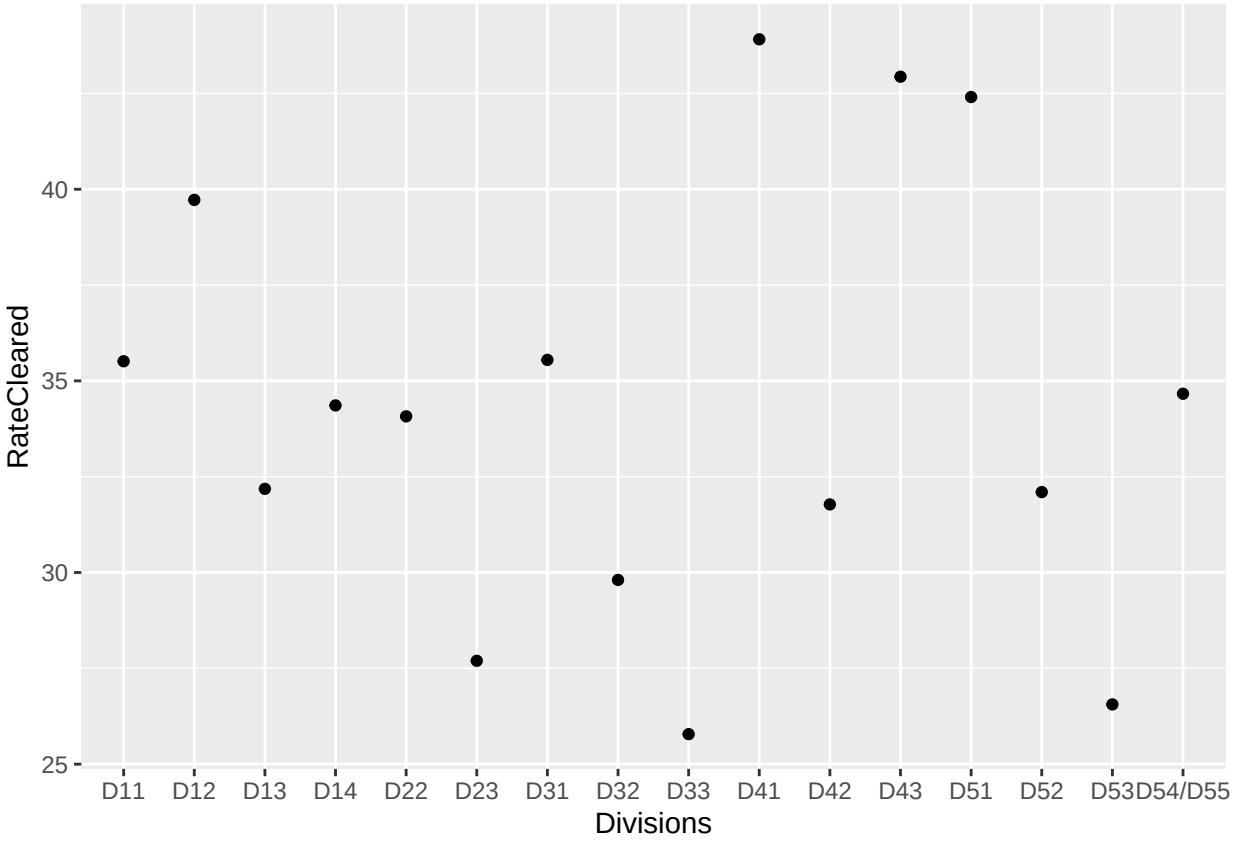
For my analysis I decided to sample just the data from the calendar year 2020. Additionally, some of the observations had not specified, "NSA," reported as their GeoDivision so I filtered those observations out of the data set. The variables of interest were kept in the data set while; id, index_,ObjectId and geometry were filtered out. This left us with 318 observations and 6 variables. The Count_ variable is the number of cases for a specific category, subcategory and division and is a numerical variable. The CountCleared is the number of cases that were cleared from the Count_ variable and is also a numerical variable. The GeoDivision variable is a categorical variable specifying the police division the charge was reported. The category variable is a categorical variable specifying the categories of crime reported. There are 5 categories, "Crimes Against Property," "Crimes Against the Person," Controlled Drugs and Substances Act," "Criminal Code Traffic," "Other Criminal Code Violations," and "Other Federal Statute Violations." Subtype is also a categorical variable specifying the sub-category of crime reported.

Table 1: Title

Division	Crimes Against Property	Crimes Against the Person	Controlled Drugs and Substances Act	Criminal Code Traffic	Other Criminal Code Violations	Other Federal Statute Violations
D11	3831	967	70	60	645	11
D12	2841	996	84	109	485	14
D13	3294	869	48	86	421	5
D14	6291	1718	137	106	922	14
D22	5151	1339	75	142	796	7
D23	3870	1169	35	97	329	6
D31	4538	1818	65	131	753	12
D32	7105	1685	38	114	858	10
D33	3638	954	33	74	306	6
D41	4547	1589	85	142	1205	4
D42	3914	1407	43	140	443	7
D43	4065	1974	74	170	916	14
D51	6719	2754	351	80	1684	14
D52	5627	1564	67	64	1641	28
D53	5239	1227	33	67	517	7

Division	Crimes Against Property	Crimes Against the Person	Controlled Drugs and Substances Act	Criminal Code Traffic	Other Criminal Code Violations	Other Federal Statute Violations
D54/D55	6276	1968	92	111	914	9

Table 1 shows an alternative view of the data broken down by division. It can be seen that regardless of division, crimes against properties are the most common and federal violations are the least common. It should also be noted that the ASR reports which division the crime was committed, not where the person charged with the crime lives. This is particularly useful to keep in mind for the criminal code traffic violations. Division 32 (D32) has the highest number of crimes reported and Division 12 (D12) has the lowest number of crimes reported.



Graph 1 plots the divisions against their computed clearance rates, i.e. the percentage of the charges cleared. The rate cleared variable was computed using the Count and CountCleared variables in the dataset.

There are some biases and ethical concerns to consider with this dataset. Firstly, not all crime is reported to the police, about two-thirds of Canadians stated that they did not report to the police when they were a victim of a crime (Wallace et al. 2009). Therefore the ASR is not entirely reflective of the amount of crime occuring in the city. Additionally, there is a substantial risk of bias as the police department is providing data about itself, as opposed to an external independent and impartial department doing the data collection and reporting. Finally, a substantial part of the year 2020 was spent in lockdown due to the COVID-19 pandemic and it is so far unclear how this factor has played a role in the amount of crime in the city.

Appendix

References

Fagan, J., & Meares, T. L. (2008). Punishment, deterrence and social control: The paradox of punishment in minority communities. Ohio St. J. Crim. L., 6, 173.

R Core Team. 2020. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rapping, J. A. (2013). Implicitly unjust: How defenders can affect systemic racist assumptions. NYUJ Legis. & Pub. Pol'y, 16, 999.

Toronto Police Service, Annual Statistical Report Open Data Documentation (2020). Toronto, ON.

Toronto, Police ServicePolice annual statistical report - reported crimes. CKAN. (n.d.). Retrieved February 6th, 2022, from https://ckan0.cf.opendata.inter.prod-toronto.ca/gl/dataset/police-annual-statistical-report-reported-crimes.

Weichselbaum, Simone, and Nicole Lewis. 2020. "Support for Defunding the Police Department Is Growing. Here's Why It's Not a Silver Bullet." The Marshall Project.

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, and Kirill Müller. 2021. Dplyr: A Grammar of Data Manipulation. https://CRAN.R-project.org/package=dplyr.

Wallace, M., Turner, J., Matarazzo, A., & Babyak, C. (2009). Measuring Crime in Canada: Introducing the Crime Severity Index and Improvements to the Uniform Crime Reporting Survey (Catalogue no. 85-004-X). Statistics Canada. Statistics Canada.

Xie, Yihui. 2021. Knitr: A General-Purpose Package for Dynamic Report Generation in R. https://yihui.org/knitr/.

Xie, Yihui. 2015. Dynamic Documents with R and knitr. 2nd edition. Chapman and Hall/CRC. ISBN 978-1498716963

Fagan, Jeffrey, and Tracey L Meares. 2008. "Punishment, Deterrence and Social Control: The Paradox of Punishment in Minority Communities." *Ohio St. J. Crim. L.* 6: 173.

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

Rapping, Jonathan A. 2013. "Implicitly Unjust: How Defenders Can Affect Systemic Racist Assumptions." *NYUJ Legis. & Pub. Pol'y* 16: 999.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Service, Toronto Police. 2020. "Annual Statistical Report Open Data Documentation."

Toronto Police Service. 2020. "Police Annual Statistical Report - Reported Crimes." Toronto,Canada: Toronto Police Service. https://ckan0.cf.opendata.inter.prod-toronto.ca/gl/dataset/police-annual-statistical-report-reported-crimes.

Wallace, M, J Turner, A Matarazzo, and C Babyak. 2009. "Measuring Crime in Canada: Introducing the Crime Severity Index and Improvements to the Uniform Crime Reporting Survey (Catalogue No. 85-004-X). Statistics Canada." *Statistics Canada.*

Weichselbaum, Simone, and Nicole Lewis. 2020. "Support for Defunding the Police Department Is Growing. Here's Why It's Not a Silver Bullet." *The Marshall Project*.

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, and Kirill Müller. 2021. *Dplyr: A Grammar of Data Manipulation*.

Xie, Yihui. 2017. *Dynamic Documents with R and Knitr*. Chapman; Hall/CRC.