No Correlations were Found between the efforts made by the Staff of Toronto's Central Intake Line and Toronto's Homeless Death Counts*

Sandy Yu

January 22, 2024

For the City of Toronto, the physical well-being of its local homeless population has always been one of its primary focuses. Numerous social efforts, funded by the City, were made in attempts to reduce the death rate of the local homeless population. In order to visualize the magnitude of social impact enforced by the City of Toronto, analyses were performed to investigate the correlation between the efforts made by the staff of Toronto's Central Intake Line and the monthly homeless death counts between November 2020 to June 2023 using data from OpenDataToronto. Staff efforts such as the number of calls coded, number of referral to shelter, and number of information for homelessness prevention, were documented and compared against the monthly homeless death counts. No discernible correlations were found between each of these efforts to the monthly homeless death counts.

Table of contents

*Code and data are available at: https://github.com/Jingying-yu/central_intake_calls_and_homeless_death_count.git

1 Introduction

The City of Toronto has been consistently providing funding and staff in an effort to reduce the active homeless population within its governing district. One of the key aspects of reducing homelessness is the assignment of sheltering space. Amongst numerous city efforts, the 24/7 helpline — Toronto Central Intake Line — was created to offer referrals to emergency shelters, sleeping accommodations, and provide general information about homelessness and prevention. (Shelter 2020). Starting in 2020, detailed data about the Central Intake Line became available on the OpenDataToronto portal (Sharla Gelfand 2022).

In this paper, we want to measure the magnitude of positive impact created by the Central Intake Line in relation to the monthly cumulative number of deaths in the homeless community. Using Central Intake Call Wrap-up Codes data (Shelter 2020) from OpenDataToronto (Sharla Gelfand 2022), three different aspects of the Central Intake Line efforts were isolated: total calls coded, referral to shelter, and homelessness prevention information provided. These three key aspects were measured against the monthly cumulative Deaths of People Experiencing Homelessness Data (Health 2017) documented by the Toronto Public Health (Toronto 2024). The analyses yielded no discernible correlation between each of the three efforts and the monthly death counts. Further investigation could be conducted in the future if more detailed tracking data of the individual callers became available.

Analyses and findings in this paper are structured into several sections: Section 2 – Data, Section 3 – Result, Section 4 – Discussion, and Section 5 – Conclusion. The Data section examines all datasets and variables kept for analysis, followed by an explanation of their data cleaning processes. The Result section focuses on visualizing and presenting the correlation between desired variables. The Discussion sections further evaluate the meaning behind the correlation presented in the previous section. Lastly, the arguments presented in all previous sections are wrapped up by the Conclusion section which summarizes the main findings in this paper.

2 Data

All data used in this paper are obtained through OpenDataToronto Portal (Sharla Gelfand 2022). Two different datasets, Central Intake Call Wrap-Up Codes Data (Shelter 2020) and Deaths of People Experiencing Homelessness(Health 2017), are retrieved to analyze the correlation between different factors of Toronto's Central Call Line efforts and the monthly cumulative death counts of homeless individuals in Toronto. Data was cleaned and analyzed using the open source statistical programming language R (R Core Team 2022) and supporting packages tidyverse (Wickham et al. 2019), janitor (Firke 2023), dplyr (Wickham et al. 2023), lubridate (Grolemund and Wickham 2011), zoo (Zeileis and Grothendieck 2005), ggplot2 (Wickham 2016), and knitr (Xie 2023). Detailed process of data extracting and cleaning can be found in the subsections below.

2.1 Central Intake Call Wrap-Up Codes

On the OpenDataToronto portal, there are several datasets that reflect the City's effort to shelter the local homeless population. The Central Intake Call Wrap-Up Codes Dataset (Shelter 2020), stored in the Central Intake Calls Catalogue (Shelter 2020), is one of the freshest and most detailed. Data was stored and published by the Shelter, Support & Housing Administration since November of 2020 and refreshes on a monthly basis. The latest refresh occurred on January 15th, 2024.

The data set provides a daily summary of the number of calls received, the number of calls classified into distinct wrap-up codes by the nature of the issue, and a count of calls under each wrap-up code. One of the example wrap-up codes in the original data set was: Code 1A - Referral to a Sleeping/Resting Space. The original data set includes 13 distinct wrap-up codes; only two codes, Code 1A — Referral to a Sleeping/Resting Space and Code 2C — Information - Homelessness & Prevention Services, were chosen for our analysis.

The process of data cleaning for this dataset started from renaming and selecting. Variables of relevance: Date, Total calls coded, Code 1A, and Code 2C, were renamed into appropriate column names and selected (Table 1).

Table 1: Sample of Renamed Central Intake Call Wrap-up Codes Variables

Date	Calls Coded	Referral to Shelter	Homelessness Prevention Info
2020-11-03	301	44	66
2020-11-04	337	56	77
2020-11-05	315	66	92
2020-11-06	283	49	71
2020-11-07	168	52	41
2020-11-08	213	48	48

Cleaned data is then filtered to include only data before July of 2023. Remaining Data is then summarized into monthly cumulative counts using the group_by() function and summarise() function under the tidyverse package (Wickham et al. 2019) (See Table 2).The rationalization behind the filtering and summarizing actions will be explained in Section 2.3.

Table 2: Sample of Cleaned Central Intake Call Wrap-up Codes Dataset

Month	Calls Coded	Referral to Shelter	Homelessness Prevention Info
2020-11-30	8367	1433	2029
2020-12-31	10232	2027	2427
2021-01-31	12091	1726	2453
2021-02-28	10525	1662	2094
2021-03-31	12287	1771	2202
2021-04-30	12668	1232	1870

2.2 Deaths of People Experiencing Homelessness

The Deaths of People Experiencing Homelessness Dataset (Health 2017) contains monthly cumulative records of homeless deaths. The dataset is published by Toronto Public Health (Toronto 2024). The earliest data record started in January of 2017, and the latest record ends in June of 2023.

The original dataset contains three columns: Year of death, Month of death, and Count. The first step of data cleaning, after downloading the data, was to exclude any rows that contained an "unknown" value. Although this action created a source of error in future analyses, this is still a necessary step because we are plotting death counts against a timeline. A date column is then created to combine the information on the year and month of death. The format of this date-class column is "yyyy-mm-dd" where the "dd" is the last day of the month when the death count of the month is finalized. This design allows easier graphing of monthly cumulative counts as discrete variables against time. The date-class column is added to the original dataset by the mutate() function under the tidyverse package (Wickham et al. 2019).

Cleaned data is then filtered to include only data from November 2020 to June 2023 and a new column that contains the name of the "Month Year" is added to the dataset (Table 3). The rationalization behind the filtering will be explained in Section 2.3.

Table 3: Sample of Cleaned Deaths of People Experiencing Homelessness Dataset

Date Displayed	Last Day of Month	Death Count
Nov 2020	2020-11-30	15
Dec 2020	2020-12-31	20
Jan 2021	2021-01-31	20

2.3 Using Data from Different Datasets

During the data cleaning process, code chunks are written to ensure that both the Central Intake Call Wrap-Up Codes Dataset (Shelter 2020) and the Deaths of People Experiencing Homelessness Dataset (Health 2017) are filtered to include only data between November 1st, 2020 to June 30st, 2023.

The decision to only include this time period is reached by taking the common time period between the two datasets. In the latest refresh, the Central Intake Call Wrap-Up Codes Dataset (Shelter 2020) begins on November 3rd, 2020 and end on December 31st, 2023; whereas the Deaths of People Experiencing Homelessness Dataset (Health 2017) begins on January of 2017 and ends on June 2023.

The ultimate purpose of taking the common time period is to ensure that we can plot variables of interest in the two datasets onto the same graph.

3 Results

Our results are summarized in **?@tbl-modelresults**.

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.

Examining how the model fits, and is affected by, the data

Figure 1: **?(caption)**

5 Conclusion

?@fig-stanareyouokay-1 is a trace plot. It shows… This suggests…

?@fig-stanareyouokay-2 is a Rhat plot. It shows… This suggests…

Checking the convergence of the MCMC algorithm

Figure 2: **?(caption)**

References

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://github.com/sfirke/janitor.

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/.

Health, Toronto Public. 2017. "Deaths of People Experiencing Homelessness." https://open.toronto.ca/dataset/deaths-of-people-experiencing-homelessness/.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Sharla Gelfand, City of Toronto. 2022. "Opendatatoronto: Access the City of Toronto Open Data Portal." https://open.toronto.ca.

Shelter, City of Toronto, Support & Housing Administration. 2020. "Central Intake Call Wrap-up Codes Data." https://open.toronto.ca/dataset/central-intake-calls/.

Toronto, City of. 2024. "Toronto Public Health." https://www.toronto.ca/city-government/accountability-operations-customer-service/city-administration/staff-directory-divisions-and-customer-service/toronto-public-health/.

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://dplyr.tidyverse.org.

Xie, Yihui. 2023. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. https://yihui.org/knitr/.

Zeileis, Achim, and Gabor Grothendieck. 2005. "Zoo: S3 Infrastructure for Regular and Irregular Time Series." *Journal of Statistical Software* 14 (6): 1–27. https://doi.org/10.18637/jss.v014.i06.