Homeless Shelter Usage Trends and Mortality: A Correlation Analysis*

Maroosh Gillani

April 18, 2024

This paper investigates the relationship between homeless shelter usage trends and mortality among those experiencing homelessness in Toronto. Using datasets from Open Data Toronto, I analyzed patterns of shelter occupancy and transitions alongside occurrences of death among homeless individuals on a yearly basis. The findings reveal a significant correlation between shelter use patterns such as returns to shelters, and an increased risk of mortality in those months. Understanding these correlations showcases the dynamics between shelter access, economic instability, and mortality risks faced by those experiencing homelessness in Toronto, and helps illuminate these issues, raise awareness, and to prompt reflection on potential avenues for assisting these individuals.

1 Introduction

In Toronto, homelessness remains a persistent social issue with significant implications for public health and social welfare (Ireland 2023). Amidst efforts to address homelessness, understanding the dynamics between shelter usage patterns and mortality risks among individuals experiencing homelessness is crucial. This paper looks into this relationship, aiming to analyze the underlying factors that contribute to mortality within this vulnerable population.

Using data sets sourced from Open Data Toronto (Gelfand 2022), this study examines the interplay between homeless shelter usage trends, particularly returns to homeless shelters per month, and monthly mortality among individuals experiencing homelessness in Toronto. By analyzing patterns of shelter occupancy, transitions, and occurrences of death on a monthly basis, I aim to highlight the connections between shelter access, economic instability, and

*Code and data are available at: https://github.com/Maroosh-Gillani/homeless_mortality.

mortality risks faced by those without stable housing. Additionally, I will employ a poisson regression model to further explore the relationships between shelter return patterns and mortality outcomes, providing statistical insights into this dynamic.

The findings of this study reveal a notable correlation between the pattern of returning to shelters, and an increased risk of mortality during certain months. This contributes to the understanding of challenges faced by individuals experiencing homelessness and emphasizes the importance of targeted interventions to address mortality risks within this population.

By discussing these issues and raising awareness of the interconnected factors influencing homelessness and mortality, this paper aims to encourage reflection on potential ways to support individuals experiencing homelessness in Toronto. Through this exploration, I aim to underscore the significance of addressing the underlying causes of homelessness and advocating for approaches to safeguard the well-being and dignity of all members of society.

2 Data

This paper uses data sets from OpenDataToronto (Gelfand 2022), which provides access to freely available digital data from the city of Toronto. In particular, the "Deaths of People Experiencing Homelessness" and "Toronto Shelter System Flow" data sets are being utilized for the purposes of analyzing the number of people returning to homeless shelters, and the number of homeless people who passed away over the time period of 2018 to 2022.

The "Deaths of People Experiencing Homelessness" is very simple, and consists of the year and months (starting January 2018), a unique id for each month, and the count of homeless persons who passed away in that month in Toronto. All columns in this dataset are numeric, with the exception of 'Month of death', which consists of characters. For this analysis, I have utilized the columns of 'Year of death', 'Month of death', and 'Count' only.

The "Toronto Shelter System Flow" is a more complex dataset, consisting of many columns such as 'date(mmm-yy)', and 'population_group', which contain characters, and details regarding the population relevant to homeless shelters in Toronto, such as 'returned_from_housing' or 'became_inactive', which are numeric in nature.

All data collection, analysis, modeling and visualization was done using R (R Core Team 2023), with the aid of the RStudio IDE (RStudio Team 2020) to streamline the process. In particular, the following packages were utilized: tidyverse (Wickham et al. 2019), ggplot2 (Wickham 2016), here (Müller 2020), knitr (Xie 2023), arrow (Richardson et al. 2024), dplyr (Wickham et al. 2023), and haven (Wickham, Miller, and Smith 2023).

3 Measurement

The "Deaths of People Experiencing Homelessness" data set is maintained by the Toronto Public Health, particularly in regards to data collection, reporting and analysis (Gelfand 2022). It is refreshed semi-annually. Their information is supplied to them by agencies such as the SSHA, alongside other social service agencies relevant to homelessness.

On the other hand, the "Toronto Shelter System Flow" gets its data from an information management system that is used to operate shelter related services that are funded by the City of Toronto, and it is updated on a monthly basis (Gelfand 2022).

4 Model

4.1 Model justification

To model the relationship between the number of homeless deaths and the number of homeless individuals returning to shelters on a monthly basis, a Poisson regression model was utilized. Such a model is suitable for this purpose because the number of deaths, as well as the number of people returning to shelters, are both instances of count data. The outcome variable here represents the counts of events occurring within a fixed interval.

4.2 Model set-up

The model is specified as follows:

$$log(ExpectedNumberofHomelessDeaths) = \beta_0 + \beta 1(NumberofIndividualsReturningtoShelters)$$

Here, β_0 represents the intercept, or the expected log count of homeless deaths when the number of people returning to shelters is 0.

β_1 represents the coefficient for the predictor variable, indicating the change in the expected log count of homeless deaths for each unit increase in the number of individuals returning to homeless shelters.

5 Results

Figure 1 shows that deaths among homeless people in Toronto, while fluctuating significantly over the months, are following a general trend of growth. Each year, the deaths have a consistent peak around winter time, especially near the month of January. The highest number

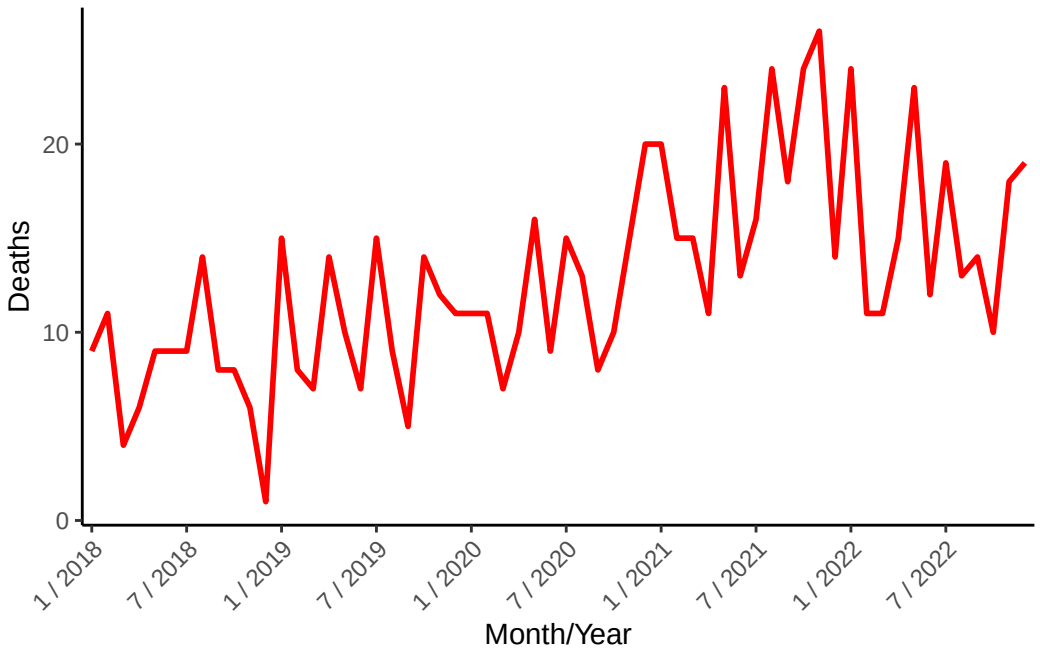

Figure 1: Deaths of Homeless Individuals in Toronto per month (2018 - 2022)

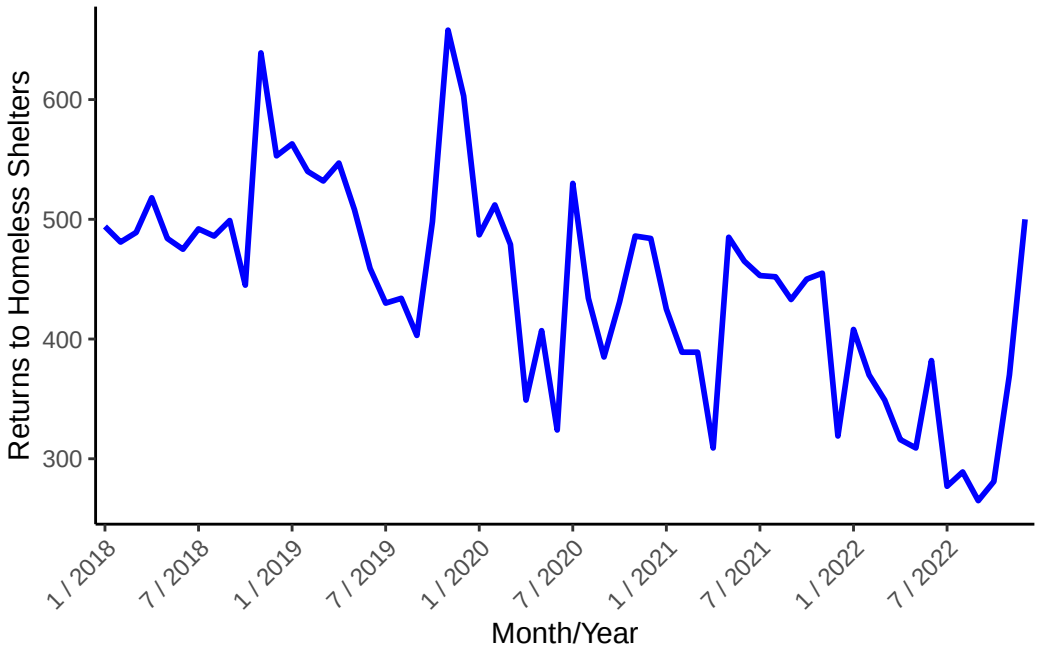

Figure 2: Number of Individuals Returning to Homeless Shelters per Month (2018 - 2022)

of homeless deaths in Toronto seem to have occurred in the November of 2021, while the lowest amount of deaths recorded are in December of 2018.

Figure 2, while also fluctuating significantly, seems to follow a negative pattern for returns of homeless people to shelters. The returns are at their peak in November of 2019. In contrast, September of 2022 saw the least amount of Toronto's homeless population return to a shelter.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	3.0368482	0.1827703	16.615656	0.0000000
Returned.to.shelter	-0.0010889	0.0004107	-2.651166	0.0080215

Figure 3: Summary of the Results from The Poisson Regression Model

Figure 3 shows the results from the Poisson regression conducted on the cleaned dataset. The aim of this model was to investigate the relationship between the number of homeless deaths and the number of individuals returning to homeless shelters. In particular, "Returned to shelter" was used as the sole predictor variable for monthly homeless deaths.

Firstly, the intercept represents the estimated log count of homeless deaths when the predictor variable is 0. Here, we can see that it is 3.0368 (with SE = 0.1828, z = 16.6157 and p = 0), which represents the base level of homeless deaths when no individuals are returning to shelters.

The estimated coefficient for Returned to shelter is approximately -0.0011 (with SE = 0.0004, z = -2.6511, and p = 0.0080). This represents the estimated change in the log count of homeless deaths in relation with a one unit increase in the number of individuals returning to homeless shelters per month. Since the coefficient is negative, the average increase of people returning to homeless shelters can be associated with a decrease in the number of homeless death.

6 Discussion

Limitations: cause of death not known in datasets.

References

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

Ireland, Nicole. 2023. "Growing Number of Homeless People Turning to ERs for Shelter and Warmth in Ontario, Study Says." https://www.cbc.ca/news/canada/toronto/homeless-people-emergency-room-shelter-1.7042041.

Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. https://CRAN.R-project.org/package=here.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Richardson, Neal, Ian Cook, Nic Crane, Dewey Dunnington, Romain François, Jonathan Keane, Dragoș Moldovan-Grünfeld, Jeroen Ooms, Jacob Wujciak-Jens, and Apache Arrow. 2024. *Arrow: Integration to 'Apache' 'Arrow'*. https://CRAN.R-project.org/package=arrow.

RStudio Team. 2020. *RStudio: Integrated Development Environment for r*. Boston, MA: RStudio, PBC. http://www.rstudio.com/.

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.

Wickham, Hadley, Evan Miller, and Danny Smith. 2023. *Haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files*. https://CRAN.R-project.org/package=haven.

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