Mortality Trends Among Shelter Residents in Toronto from 2007 to 2024*

Fangning Zhang

September 27, 2024

This study conducts a series of analyses of mortality among shelter residents in Toronto from 2007 to 2023, focusing on overall trends, gender differences, and seasonal variations. The results indicate a concerning increase in deaths starting from 2019, particularly among males. The number of deaths during the winter months is higher compared to other months. Understanding the mortality of shelter residents is crucial for formulating effective social policies, improving shelter services, and enhancing the health and well-being of vulnerable groups.

Table of contents

1 Introduction **2**

2 Data **3**
2.1 Data Source and Measurement 3
2.2 Data Tools 3
2.3 Data Cleaning 4
2.4 Data Visualization and Result 5
2.5 Other Datasets 7

3 Discussion **8**
3.1 Summary of Key Findings 8
3.2 Interpretation of Findings 8
3.3 Limitations 8
3.4 Suggestions for Future Research 8

References **10**

*Code and data are available at: https://github.com/FangningZhang81/Deaths-of-Shelter-Residents.

1 Introduction

Shelter residents are one of the most vulnerable groups in urban areas, facing numerous challenges such as housing instability, health issues, and social exclusion. Numerous studies have shown that shelter residents are at a significantly higher risk of premature death compared to the general population due to factors such as poor health, lack of access to medical care, and exposure to unsafe environments(Hwang et al. (2009)). Understanding the mortality of shelter residents is crucial for formulating effective social policies, improving shelter services and also improving the health and well-being of shelter residents.

Although many existing researches highlight the higher mortality rates among the homeless population, there is still a lack of analysis specifically targeting shelter residents. In this study, a comprehensive analysis of mortality data is made among shelter residents in Toronto, focusing on the following key aspects. First, we analyze the overall trend in the number of deaths over the 16-year period from 2007 to 2023. Second, we explored gender differences in mortality. Third, we analyzed the monthly distribution of deaths to understand whether different climates or seasons have varying impacts on shelter residents.

The analysis revealed a concerning upward trend in the number of deaths starting from 2019, despite a slight decline in the following two years. Male deaths number have consistently been higher than female deaths, which is more and more significant after 2019. Additionally, there was an increase in deaths during the winter months, suggesting a link between extreme weather conditions and mortality risk may exist. This may consistent with previous research indicating that bad weather conditions and overcrowding in shelters can exacerbate health risks for homeless individuals (Vuillermoz et al. (2016)).

The rise in the number of deaths starting in 2019 worths attention. However, this increase does not necessarily indicate worsening shelter conditions or declining health among residents. The rise could be due to an increase in the number of shelter residents, changes in admission policies, or external factors such as the impact of COVID-19 and other social and public health issues (Ramin and Svoboda (2009)). This is discussed in detail in the discussion section.

The remainder of this paper is structured as follows: Section 2 describes the data sources, tools, cleaning process, visualization, result and related datasets; and Section 3 summarizes the key findings, interprets the results, discusses limitations, and suggests directions for future research.

2 Data

2.1 Data Source and Measurement

The dataset used in this study is titled "Deaths of Shelter Residents". This dataset is published by the Toronto Shelter and Support Services Department and is obtained from OpenData-Toronto (Gelfand (2022)). The Toronto Shelter and Support Services Division has collected this information since 2007 and conducts an annual review of the data. Besides, the death of shelter residents data is a subset of data shared by Toronto Public Health (TPH) on its Deaths of People Experiencing Homelessness dashboard. The TPH dashboard includes figures on deaths of individuals experiencing homelessness staying in shelters, as well as people living outdoors or at other locations. Therefore, the dataset is accurate and reliable. The dataset covers the deaths of shelter residents from 2007 to the present. Variables such as the total number of decedents and the breakdown of deaths by gender (male, female, transgender/non-binary/two-spirit) are included. All death counts are reported on a monthly basis and is still updated monthly. For privacy and ethical reasons, the dataset does not contain any personally identifiable information.

A Table 1 with first few rows is shown below, which can give an overview of the raw data and variables.

Table 1: First few rows of the raw dataset

_id	Year	Month	Total decedents	Male	Female	Transgender/Non-binary/Two-Spirit
1	2007	Jan	0	0	0	n/a
2	2007	Feb	3	3	0	n/a
3	2007	Mar	3	2	1	n/a
4	2007	Apr	1	1	0	n/a
5	2007	May	2	2	0	n/a
6	2007	Jun	3	3	0	n/a

2.2 Data Tools

The data was extracted, cleaned and visualize using R programming language and the whole coding process is done with R (R Core Team (2023)). Some packages are used throughout the paper, including tidyverse (Wickham et al. (2019a)), ggplot2 (Wickham (2016)), styler (Müller and Walthert (2024)), janitor (Firke (2023)), OpenDataToronto (Gelfand (2022)), knitr (Xie (2024)), and KableExtra (Zhu (2024)). Besides, the stater folder from Rohan (Wickham et al. (2019b))is used.

2.3 Data Cleaning

The raw data was first imported from the file and the column names were standardized using the `janitor` package(Firke (2023)) to ensure consistency and ease of use throughout the analysis. The next step involved converting the month abbreviations into numeric values to increase the ease of use and analysis. For example, "01" for January, "02" for February.

A new column named `year_month` was generated by concatenating the `year` and `month` columns making it easier to perform time-series analyses and visualizations.

The dataset also included a column named `transgender_non_binary_two_spirit`, which some rows contained non-numeric values (n/a). All `NA` values in this column were replaced with 0 to ensure the data was suitable for statistical analysis and visualization.

Finally, the cleaned dataset was saved to a new file named `analysis_data.csv` in the `analysis_data` directory. It includes the following variables:

- `year`: The year in which the death occurred.
- `month`: The month in which the death occurred, converted to numeric values for ease of analysis.
- `total_decedents`: The total number of deaths recorded in that month.
- `male`: The number of male decedents.
- `female`: The number of female decedents.
- `transgender_non_binary_two_spirit`: The number of decedents identified as transgender, non-binary, or two-spirit. In cases where this information is unavailable, the value is recorded as zero.
- `year_month`: A constructed variable representing the year and month combined, formatted as "year_month" (e.g., 2007_01 for January 2007).

The Table 2 below provides a summary of the cleaned data, showing key statistics (maximum, minimum, mean, median, and variance) for the total number of deaths, as well as for male and female decedents. It highlights that the average number of deaths per month is 4, with males averaging 3 deaths and females 1 death.

Table 2: Summary of the cleaned data

Statistic	Total	Male	Female
Max	19	16	5
Min	0	0	0
Mean	4	3	1
Median	3	2	0
Variance	13	8	1

2.4 Data Visualization and Result

Graphical representations, including line plot, pie chart and bar charts, are used to visualize the trends in the data.These visualizations are critical for understanding the distribution of deaths across different demographic groups and time periods.

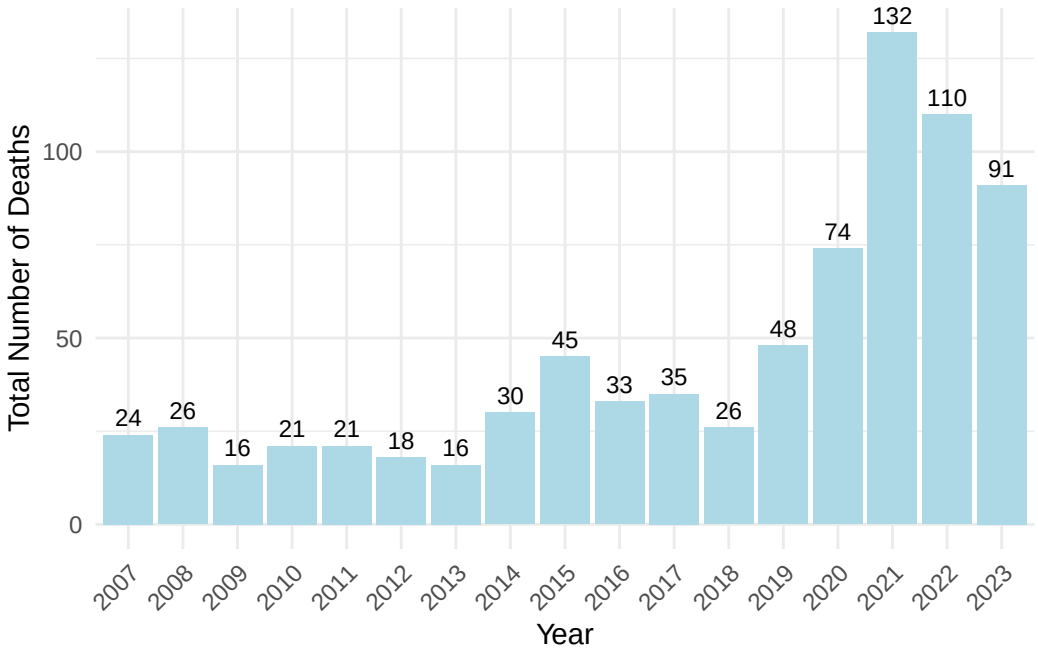

Figure 1: Total Number of Deaths by Year (2007-2023)

This bar chart(Figure 1) displays the total number of deaths among shelter residents in Toronto from 2007 to 2023. The data shows a significant increase in deaths starting from 2019, with a peak in 2021 (132 people). The upward trend continues into 2022 and 2023 (110 and 91 people), though slightly lower than 2021. The data for 2024 has been excluded from this chart as the dataset only covers information up to August 2024.

Figure 2 displays the death trend of each category among shelter residents in Toronto from 2007 to 2023. The data shows a significant increase in deaths starting from 2019 of all three lines, while the yellow line remains at a relatively low level. The total number of deaths equals the sum of male, female, and other gender deaths.The males death rate has consistently been higher than females and the differences is becoming increasingly obvious. The pie chart (Figure 3) shows the gender distribution of the deaths. 78.2% of the total deaths are from males, while females account for 65.3%. In contrast, individuals of other genders make up only 4.7%.

Figure 4 illustrates the total number of deaths by month among shelter residents. The highest number of deaths occurred in January which is 80 people, followed by relatively high numbers

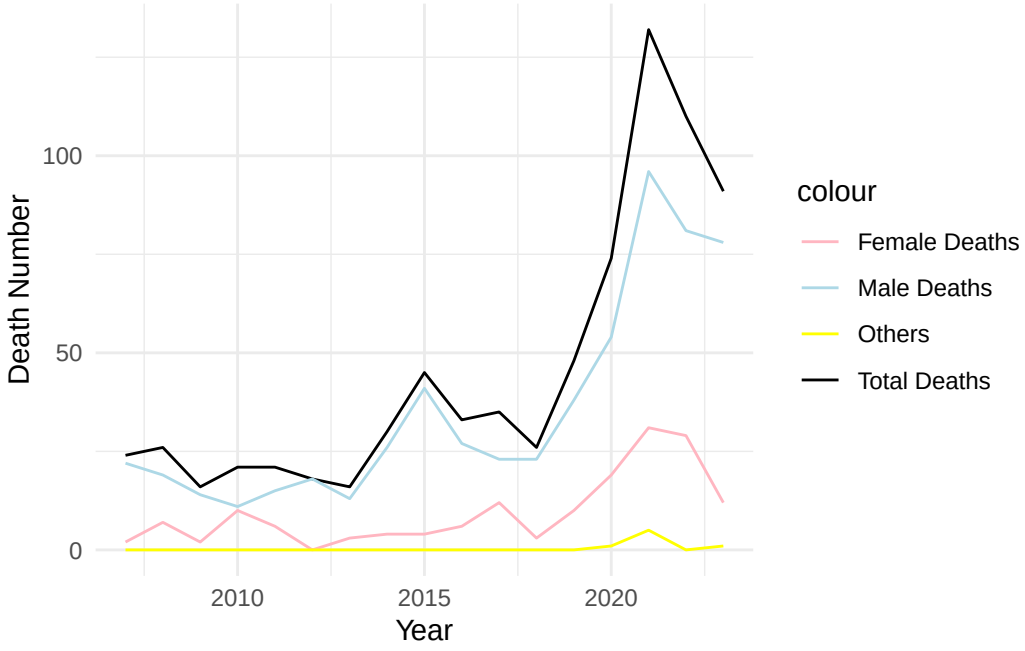

Figure 2: Death Numbers by Year and Category

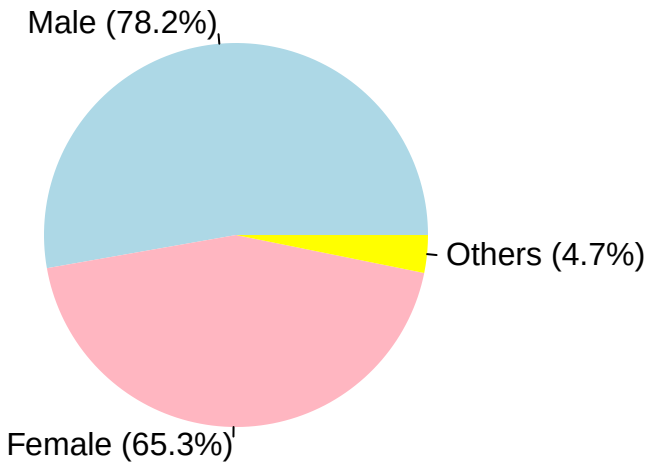

Figure 3: Gender Distribution of Deaths (2007-2023)

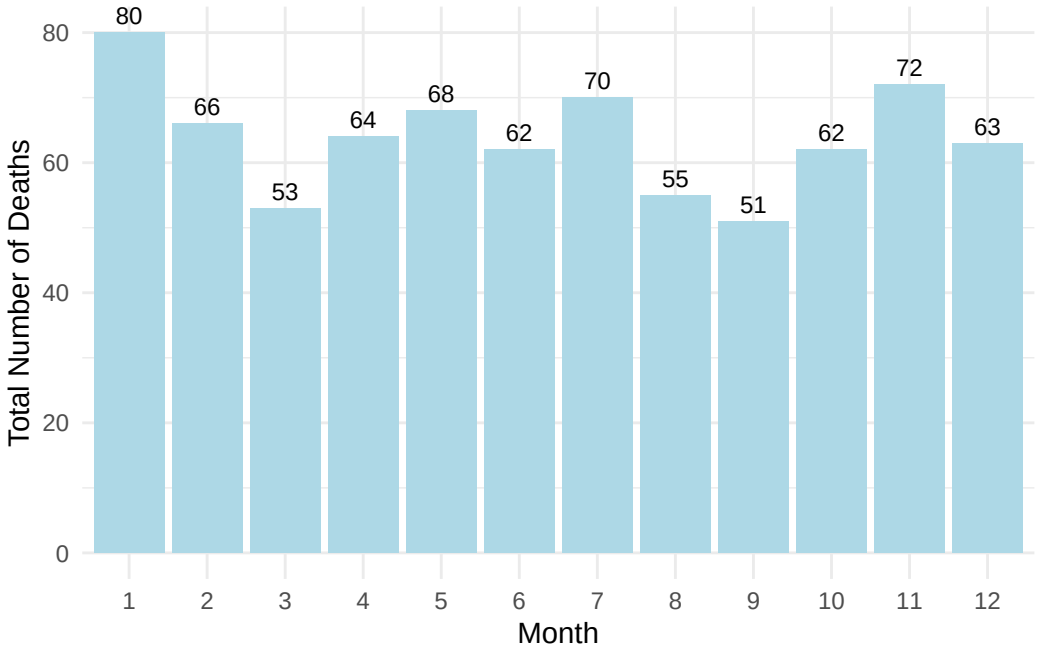

Figure 4: Total Number of Deaths by Month (2007-2023)

in November which is 72 people. Conversely, the lowest number of deaths occurred in March and September which is 53 and 51 separatly. There is a slight increase in deaths during the winter months (December to February).

2.5 Other Datasets

There are some related datasets found on OpenDataToronto (Gelfand (2022)), but they are not used for the following reasons.The "Death Registry Statistics" dataset contains variables that are too broad and do not provide specific information on deaths among vulnerable populations, making it unsuitable for detailed analysis of shelter resident mortality.The "Fatal and non-fatal suspected opioid overdoses in the shelter system" dataset, while relevant, only covers one specific cause of death (opioid overdoses) and therefore does not capture the full scope of mortality within the shelter system, making it insufficient as the primary dataset. The dataset "Daily Shelter Occupancy" can be a very helpful dataset for this study. As The number of deaths among shelter residents is closely related to the total number of people residing in the shelter. The "Daily Shelter Occupancy" dataset could be useful for understanding the relationship between death rates and the total number of shelter residents, but it was not included in this study because the primary focus is on the death statistics themselves, rather than occupancy rates. Future studies could consider integrating this dataset to explore such relationships in more depth.

3 Discussion

3.1 Summary of Key Findings

The study revealed an overall upward trend in the mortality among shelter residents, with a significant gender difference which is the death number for males is consistently higher than that for females. Additionally, there is a more deaths during the winter months. A particularly concerning finding is the sharp rise in mortality starting in 2019.

3.2 Interpretation of Findings

The observed seasonal variations in mortality, particularly the increase during winter months, could be influenced by several factors, including extreme weather conditions, overcrowding in shelters, and the availability of healthcare resources. Harsh winter conditions may exacerbate existing health issues.

The higher number of deaths among males compared to females may be attributed to various factors, including environmental and social factors within the shelter system. It is also possible that this difference is due to a larger number of male residents compared to female residents.

The underlying causes of these trends are important directions for further in-depth research in the future.

3.3 Limitations

One limitation of this study is that the original dataset does not account for Transgender/Non-binary/Two-Spirit individuals before 2020, which reduces the accuracy of gender-specific analyses and raises ethical concerns. Additionally, the number of deaths may depend on the number of shelter residents, but this information is not included in the dataset, making it impossible to calculate mortality rates. The rapid growth of the death number after 2019 may be influenced by the increasing number of shelters or Covid-19, etc. Furthermore, changes in shelter admission conditions and other policies, which could impact the number of deaths, are also not discussed in this paper.

3.4 Suggestions for Future Research

Future research should explore external factors such as economic conditions and the availability of healthcare resources that may further influence the number of deaths among shelter residents. It is recommended to utilize more diverse datasets, such as daily shelter occupancy rates which is mentioned earlier and other health-related data, to gain a deeper understanding of the health

and mortality trends of shelter residents. These approaches would offer a deeper understanding of the complex factors influencing the vulnerable population mortality.

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.

Hwang, Stephen W, Russell Wilkins, Michael Tjepkema, Patricia J O'Campo, and James R Dunn. 2009. "Mortality Among Residents of Shelters, Rooming Houses, and Hotels in Canada: 11 Year Follow-up Study." *Bmj* 339.

Müller, Kirill, and Lorenz Walthert. 2024. *Styler: Non-Invasive Pretty Printing of r Code*. https://CRAN.R-project.org/package=styler.

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

Ramin, Brodie, and Tomislav Svoboda. 2009. "Health of the Homeless and Climate Change." *Journal of Urban Health* 86: 654–64.

Vuillermoz, Cécile, Albertine Aouba, Lise Grout, Stéphanie Vandentorren, Fanny Tassin, Margarita Moreno-Betancur, Éric Jougla, and Grégoire Rey. 2016. "Mortality Among Homeless People in France, 2008–10." *The European Journal of Public Health* 26 (6): 1028–33.

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. 2019a. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

———, et al. 2019b. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Xie, Yihui. 2024. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. https://yihui.org/knitr/.

Zhu, Hao. 2024. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=kableExtra.