Analysis on Trend of Death of Shelter Residence in Toronto*

focus on gender distribution from 2007 to 2024

Xinqi Yue

June 21, 2025

Trough the analysis on the data about deaths of shelter residents, this paper found that the total number of deaths of shelter residences is increasing year by year. The proportion of male deaths is relatively higher. And seasonal factors does not affect much on it. The findings provide insights into the factors affecting well-being of shelter residents and can guide policies to address homelessness, improve living conditions, and tackle health-related issues in this vulnerable population.

1 Introduction

The rising number of deaths in Toronto's shelter system is a growing area of concern, reflecting broader societal challenges including homelessness, mental health crises, and inadequate access to health care. Understanding the dynamics of these deaths is critical to developing effective policies and interventions aimed at reducing mortality in the vulnerable population. This paper focuses on analyzing the dataset related to deaths in Toronto's shelter residence, and seeks to identify the demographic trends and factors contributing to these outcomes.

Despite the limited scope of the dataset due to privacy concerns, which excludes identifiable information such as names, dates of birth, and specific causes of death, and the fact that non-binary genders were not included until 2020, the aggregated data allows for the examination of key trends. The paper finds that the total number of deaths from January 2007 to August 2024 showed an overall increasing trend, and that male deaths accounted for the largest share of all genders during this period. Besides, although the data shows that death occurs in winter most frequently, the result can be attributed to bias of the pandemic.

These findings are important because they provide a clearer picture of the extent and nature of mortality in Toronto's shelter population, which can be used to inform future policy decisions

---

and the allocation of resources. By identifying key trends and limitations within the dataset, this study provides a basis for more targeted and informed interventions to address the underlying causes of death in this population. This paper begins with a comprehensive overview of the dataset, including its variables (Section 2.1), limitations (Section 2.2), the rationale for its selection (Section 2.3), and the data collection process (Section 2.4). Following this, the data is presented with accompanying graphs and analysis (Section 2.5). The subsequent section outlines key findings (Section 3), followed by a discussion of their broader implications (Section 4). All visualizations were generated using R Core Team (2023), with the data sourced from the City of Toronto (2024a) dataset.

2 Data

The dataset used for this analysis relates to deaths within Toronto's shelter system, from City of Toronto (2024a). It contains key demographic variables and aggregates information about the number of deaths reported each year, categorized by gender. However, due to privacy concerns and respect for the confidentiality of individuals who have passed, the dataset excludes personal identifying information including names, age, or specific locations of death. This limitation necessitated the use of data for analysis focusing on factors including gender distribution and time. There are 212 observations in the original dataset in total and there is no special cleaning methods needed as the data is already clear enough. See Appendix Section A for the summary of the dataset.

2.1 Variables

The dataset include the following variables:

1. **Year**: The year in which the death occurred at a Toronto shelter, ranging from 2007 to 2024. This variable aids in analyzing trends over time.

2. **Month**: The month in which the death occurred, corresponding to each year from January to December. Data beyond August 2024 has not yet been recorded.

3. **Total Deaths**: The total number of deaths recorded in the given month.

4. **Male Deaths**: The total number of male shelter residents who died in the given month.

5. **Female Deaths**: The total number of female shelter residents who died in the given month.

6. **Other Deaths**: The number of deaths of individuals identified as Transgender, Non-binary, or Two-Spirit in the given month. Note that data for this category was not recorded prior to 2020.

The variable follows the rule: Total Death = Male Deaths + Female Deaths + Other Deaths

2.2 Limitation

The dataset used for this analysis presents several limitations that must be acknowledged to provide a comprehensive understanding of the findings.

Privacy and Respect for the Deceased: Due to ethical considerations, certain sensitive information such as age and location of death was not collected. This lack of data may limit the ability to perform more granular analyses related to demographic factors influencing mortality rates among shelter residents.

Gender Data Constraints: Information on gender identity, specifically regarding transgender, non-binary, and two-spirit individuals, was not systematically recorded until 2020. As a result, prior to this year, any data related to these gender identities is missing. This gap restricts the ability to fully analyze trends and disparities in mortality rates among diverse gender identities over time.

Absence of Comparable Datasets: It is essential to consider that the dataset could be enriched by including broader contextual variables. Specifically, information about the total shelter population and its gender composition, age distribution, and health status of residents would provide a more nuanced understanding of the factors contributing to mortality in these settings. Such data would allow for deeper insights into the demographic profiles of those affected and could better inform policies aimed at addressing the challenges of homelessness and shelter living.

Data Quality and Cleaning: The dataset misses values of transgender/Non-binary/two-spirit decedents until 2020 due to lack of awareness of gender identification. The missing values only affect the research in gender vulnerability field and it is hard to be fulfilled. In our research, the missing value is replaced with 0 mostly. The value keeps being NA when calculate the death ratio of each gender.

By recognizing these limitations, the analysis aims to provide a clearer context for the findings and to encourage further research that can fill these gaps and enhance understanding of mortality trends among shelter residents.

2.3 Similar Datasets

The dataset titled Deaths of People Experiencing Homelessness, from City of Toronto (2023), provides information on the deaths of homeless individuals, which is somewhat similar to our current study of deaths among Toronto shelter residents. One notable advantage of this dataset is that it includes a detailed CSV file outlining the causes of death, offering valuable insights into the factors contributing to mortality in this population.

However, the dataset containing the causes of death cannot be downloaded directly, which limits the accessibility and availability of our analysis. The accessible data variables align perfectly with our research dataset but only cover the period from 2017 to 2023. The lack of

longitudinal data means it does not align with the comprehensive timeline of our study, which spans from January 2007 to August 2024. Therefore, although this dataset provides certain advantages, the limitations regarding its time span and accessibility led to its exclusion from our viable research options.

2.4 Measurement

Within this data set, the Structured Reporting System facilitates the collection of information regarding the death of shelter residents. Shelter operators are required to notify the City of Toronto within 24 hours of the death of a current or recent shelter resident and then submit a detailed written report within 30 days. (City of Toronto (2024b)) This systematic approach, in place since 2007 and reviewed annually by Toronto Housing and Support Services, ensures timely and consistent data collection. The structured reporting requirement imposes a level of discipline on shelter operators to ensure that deaths are not only reported in a timely manner, but also fully documented, thereby enhancing the integrity and credibility of the dataset.

However, certain limitations affect the completeness and reliability of the data. For example, due to privacy concerns, personal information such as names, dates of birth, and specific places of death are not collected.(City of Toronto (2024b)) This limits the granularity of the data and can hinder deeper analysis, such as correlations with demographic or location-specific factors. In addition, shelter operators often do not have access to death certificates or information on causes of death, so the dataset lacks this critical context that could be important for understanding the factors contributing to mortality among shelter residents.

Improvements to the measurement process could include better integration with public health and legal entities to access more detailed information about deaths, including cause of death, while still protecting individual privacy. This improvement would provide a more comprehensive view of the factors influencing mortality rates among shelter residents, allowing for more informed policy decisions to address health and safety issues in shelters.

2.5 Overall Visualization

In the following section, I will present a series of visualizations that analyze the number of deaths among shelter residents over time, highlighting the gender distribution and overall data trends. These charts will provide a clearer understanding of the data composition and allow for an insightful examination of how the number of deaths in this population has evolved.

2.5.1 Trend Over Time

Figure 1, created by R Core Team (2023), shows the total number of deaths of Toronto shelter residents each year from January 2007 to August 2024, divided into 15 sub-graphs, arranged

in 5 rows and 3 columns by year. This layout allows readers to clearly compare the trend of deaths from one year to the next and observe potential seasonal changes.

Based on the generated line graphs, the following conclusions can be drawn:

1. **Overall Increasing Trend in Deaths**: From 2007 to 2024, there has been a consistent increase in the number of deaths. This trend may be related to several factors, including changes in economic conditions, the impact of the COVID-19 pandemic, and overall population growth. These factors could contribute to a rise in the homeless population, leading to an increase in mortality rates.

2. **Peak Deaths in October 2021**: Notably, October 2021 recorded the highest number of deaths. This peak may be associated with the public health crisis stemming from the COVID-19 pandemic, social isolation policies, and their effects on mental health and living conditions. Further research could help clarify the specific contributing factors during this period.

3. **Lack of Significant Seasonal Impact on Deaths**: Overall, there appears to be no substantial seasonal variation in death rates, indicating that the fluctuations in mortality may be more closely linked to economic and social factors rather than seasonal influences.

These conclusions provide important insights into the overall mortality trends of Toronto shelter residents and suggest that relevant agencies need to consider multiple influencing factors when developing policies and interventions.

2.5.2 Annual Deaths with Gender Composition

Figure 3, created by R Core Team (2023), below illustrates the annual number of deaths from 2007 to 2023. And Figure 2, created by R Core Team (2023), illustrates the annual number of deaths of the same time period, with distinct color coding to represent gender distribution: pink for male, black for female, and gray for other gender identities, including transgender and non-binary.

Several important conclusions can be drawn from the visualization. First, the proportion of male deaths remains consistently the highest throughout the years observed, indicating a significant gender disparity in mortality rates among shelter residents. In addition, there is a noticeable upward trend in the overall number of deaths, suggesting a growing concern for the well-being of this vulnerable population. It can also be seen clearly from the line graph of Figure 3. Notably, 2021 stands out as the peak year for mortality. This can be seen clearly in Figure 3 too. This raises questions about the factors contributing to this spike and the broader implications for public health and social support systems.

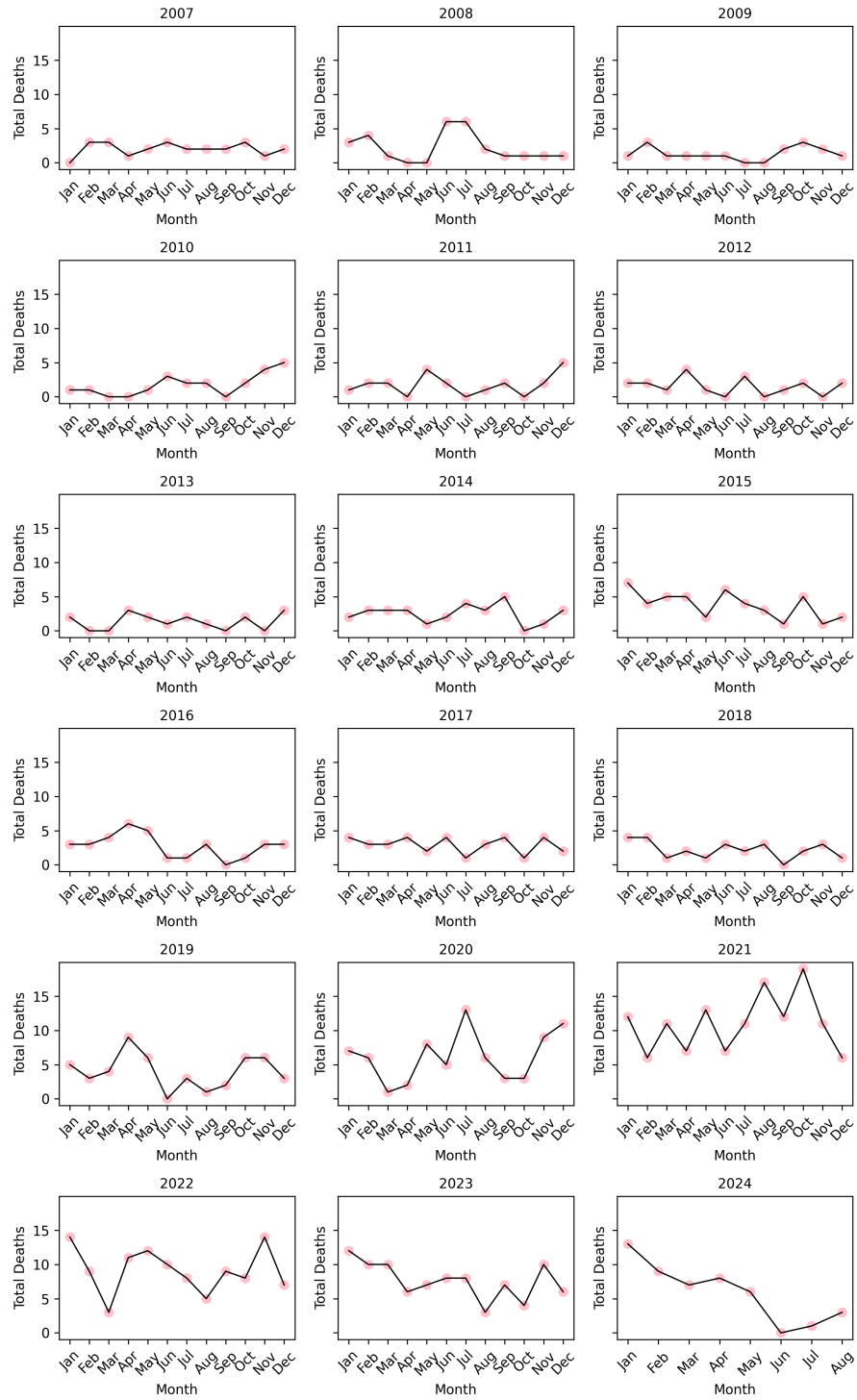**Deaths in Toronto Shelters Over Time (2007.01–2024.08)**

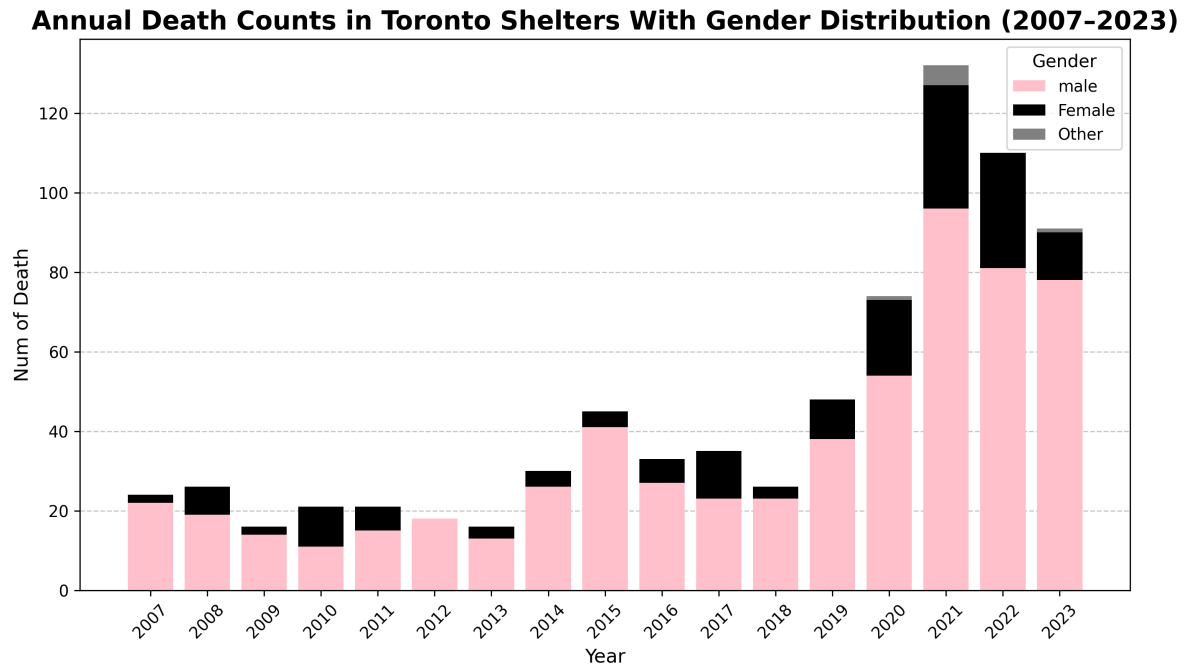Figure 1: Deaths in Toronto Shelters Over Time (2007.01–2024.08)

Figure 2: Annual Death Counts in Toronto Shelters With Gender Distribution (2007–2023)

Figure 3: Annual Deaths in Toronto Shelters (2007–2023)

2.5.3 Compariasion Over Month

Figure 4, created by R Core Team (2023), shows the aggregated monthly death counts for shelter residents from 2007 to 2023. This data was derived by summing the death counts for each January from 2007 to 2023, resulting in a comprehensive data set for January, and repeating this process for each of the other months, resulting in data for all twelve months.

The purpose of this visualization is to analyze whether certain months have higher mortality rates. It helps to exploring potential seasonal influences on mortality rates, such as the impact of extreme temperatures or adverse weather conditions, which could contribute to increased mortality among shelter residents.

The current findings, shown in the line graph, indicate that January and November have the highest mortality rates, while March and September have the lowest mortality rates, which was different in what we find in Section 2.5.1. These insights highlight the need for further investigation into the factors influencing these monthly variations to inform targeted interventions and improve the health and safety of shelter residents.
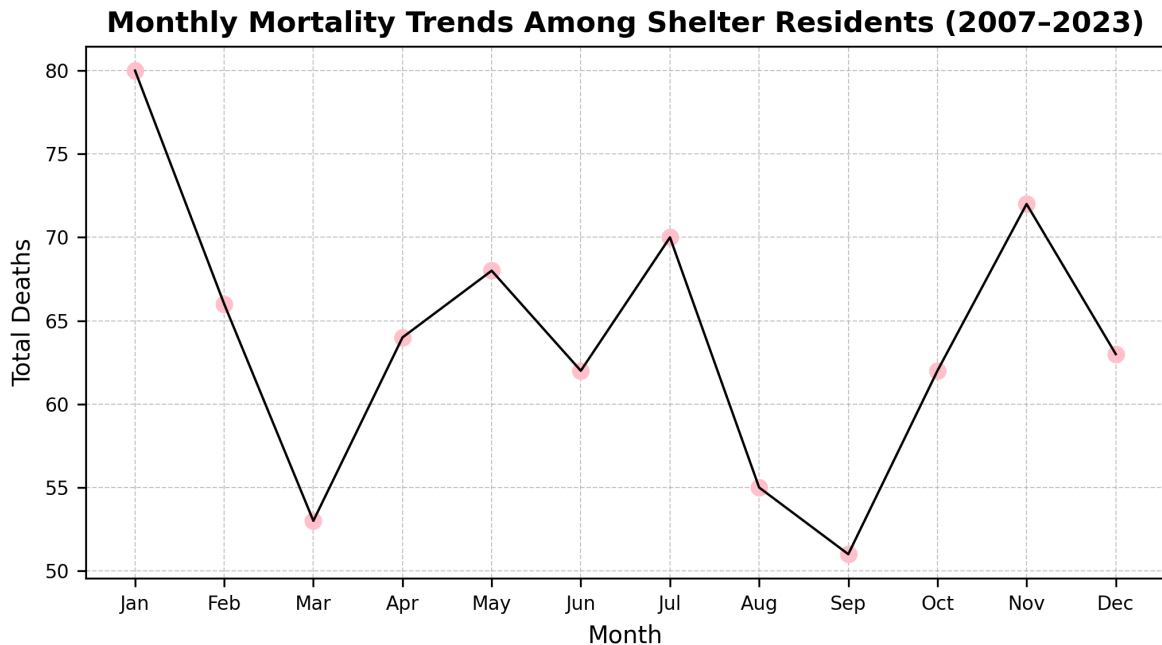

Figure 4: Monthly Mortality Trends Among Shelter Residents (2007–2023)

3 Results

As time progresses, it is evident from Figure 2 and Figure 3 that the number of deaths among shelter residents has increased. Potential reasons for this trend may include changes in the economic landscape, shifts in the overall population, policy modifications, and factors related to the pandemic.

Regarding monthly variations, while Figure 4 illustrates a noticeable peak in January and a low in September, a comprehensive view of the data in Figure 1 suggests that seasonal changes do not appear to have a significant correlation with the number of deaths. One possible explanation for this phenomenon could be related to the timing of the pandemic's outbreak.

In terms of gender composition, male deaths have consistently constituted the highest proportion from 2007 to 2023, indicating a potential underlying vulnerability. However, this trend may also be influenced by the higher proportion of males residing in shelters.

4 Discussion

4.1 suammry

This article examines trends in mortality among shelter residents in Toronto from 2007 to 2023, with a focus on gender distribution. Our analysis shows that there was an upward trend in deaths during this period, with a higher proportion of men dying.

In exploring the implications of these findings, we gain valuable insights into the complex interplay of economic, social, and health factors that affect shelter residents. The data suggest that economic downturns, policy changes, and global health events may exacerbate the vulnerability of this population. In addition, the lack of a clear seasonal pattern in mortality suggests that other factors, such as the pandemic, may be masking typical seasonal effects and warrant further investigation.

4.2 Weaknesses and next steps

While the results highlight important issues, certain limitations must be acknowledged. The dataset used for this analysis may not fully capture the demographics of the shelter population, such as its gender composition and health status. In addition, the analysis relied on reported deaths, which may be inaccurate or underreported. Future research should aim to incorporate more comprehensive demographic data, such as data on the total shelter population and its gender composition, causes of death, etc., to improve our understanding of the factors driving these trends.

Finally, this study highlights several areas for future research. Further research could focus on the specific causes of mortality among shelter residents and the effectiveness of current policies and programs designed to reduce mortality among this vulnerable population. By addressing these gaps, we can better inform policy decisions and improve outcomes for people experiencing homelessness.

A Appendix

Table 1, created by R Core Team (2023), shows the a sample how data is like in the original dataset from City of Toronto (2024a) provided by Gelfand (2022). It is the first 5 lines in the dataset.

Table 2, created by R Core Team (2023), shows the summary of our dataset. It contains the total decedents number, monthly average deaths, minimum, maximum and median of the each year. It also provide the ratio of different gender in the shelter decedents each year.

Table 1: Sample Table of Deaths in Toronto Shelters (2007–2023)

	ID	Year	Month	Total decedents	male	female	Others
0	1	2007	Jan	0	0	0	nan
1	2	2007	Feb	3	3	0	nan
2	3	2007	Mar	3	2	1	nan
3	4	2007	Apr	1	1	0	nan
4	5	2007	May	2	2	0	nan

Table 2: Annual Summary of Deaths in Toronto Shelters (2007–2023)

	Year	Total Deaths	Mean	Median	Min	Max	Male Ratio (%)	Female Ratio (%)	Other
0	2007	24	2.000000	2.000000	0	3	91.670000	8.330000	NA
1	2008	26	2.170000	1.000000	0	6	73.080000	26.920000	NA
2	2009	16	1.330000	1.000000	0	3	87.500000	12.500000	NA
3	2010	21	1.750000	1.500000	0	5	52.380000	47.620000	NA
4	2011	21	1.750000	2.000000	0	5	71.430000	28.570000	NA
5	2012	18	1.500000	1.500000	0	4	100.000000	0.000000	NA
6	2013	16	1.330000	1.500000	0	3	81.250000	18.750000	NA
7	2014	30	2.500000	3.000000	0	5	86.670000	13.330000	NA
8	2015	45	3.750000	4.000000	1	7	91.110000	8.890000	NA
9	2016	33	2.750000	3.000000	0	6	81.820000	18.180000	NA
10	2017	35	2.920000	3.000000	1	4	65.710000	34.290000	NA
11	2018	26	2.170000	2.000000	0	4	88.460000	11.540000	NA
12	2019	48	4.000000	3.500000	0	9	79.170000	20.830000	NA
13	2020	74	6.170000	6.000000	1	13	72.970000	25.680000	NA
14	2021	132	11.000000	11.000000	6	19	72.730000	23.480000	NA
15	2022	110	9.170000	9.000000	3	14	73.640000	26.360000	NA
16	2023	91	7.580000	7.500000	3	12	85.710000	13.190000	NA

References

City of Toronto. 2023. "Deaths of People Experiencing Homelessness." https://open.toronto.ca/dataset/deaths-of-people-experiencing-homelessness/.

———. 2024a. "Deaths of Shelter Residents." https://open.toronto.ca/dataset/deaths-of-shelter-residents/.

———. 2024b. "Deaths of Shelter Residents." https://www.toronto.ca/city-government/data-research-maps/research-reports/housing-and-homelessness-research-and-reports/deaths-of-shelter-residents/.

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

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