Analysis on Trend of Death of Shelter Residence in Toronto*

focus on gender distribution from 2007 to 2024

Xinqi Yue

September 24, 2024

Trough the analysis on the data about deaths of shelter residents, we found that the total number of deaths of shelter residences is increasing year by year. The proportion of male deaths is relatively higher. The findings provide insights into the factors affecting shelter residents' well-being 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 within Toronto's shelter system is a growing area of concern, reflecting broader societal challenges such as 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 this 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 unfortunate outcomes.

Despite the importance of this issue, there is a clear gap in the public's detailed understanding of who is being affected and how patterns of mortality are evolving over time. Due to privacy concerns, the dataset is limited in scope and lacks identifying information such as names, dates of birth, and specific causes of death. Also, the Non-binary genders are not included until 2020. Nevertheless, these aggregate data allow for an examination of key trends, such as the gender distribution of deaths and the change over time. These trends highlight critical disparities, particularly the higher proportion of male deaths and the steady increase in overall mortality number.

*Code and data are available at: https://github.com/xinqiyue/ShelterDeath.

This study aims to provide a visual and statistical analysis of the available data and to draw attention to the factors that may be contributing to the increased mortality. The analysis reveals two important findings: first, the number of deaths has steadily increased over the years, indicating a deteriorating situation within the shelter system; Second, men make up a larger proportion of the deceased, suggesting potential gender-specific vulnerabilities.

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. The structure of this paper begins with a detailed description of the data@sec-data, followed by the statistic model@sec-model, results@sec-result, and discussion of implications@sec-discussion.

2 Data

The dataset used for this analysis relates to deaths within Toronto's shelter system. 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, birthdates, or specific locations of death. This limitation necessitated the use of data for analysis focusing on factors including gender distribution and time rather than locations and age.

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.

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 required careful cleaning to address issues such as missing values and inconsistencies, particularly in the reporting of deaths by gender. Although high-level cleaning was conducted to ensure the reliability of the data, the absence of certain variables may still introduce biases or limit the interpretability of the results.

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

2.4 Overall Visualization

2.4.1 Trend Over Time

Figure 1 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:

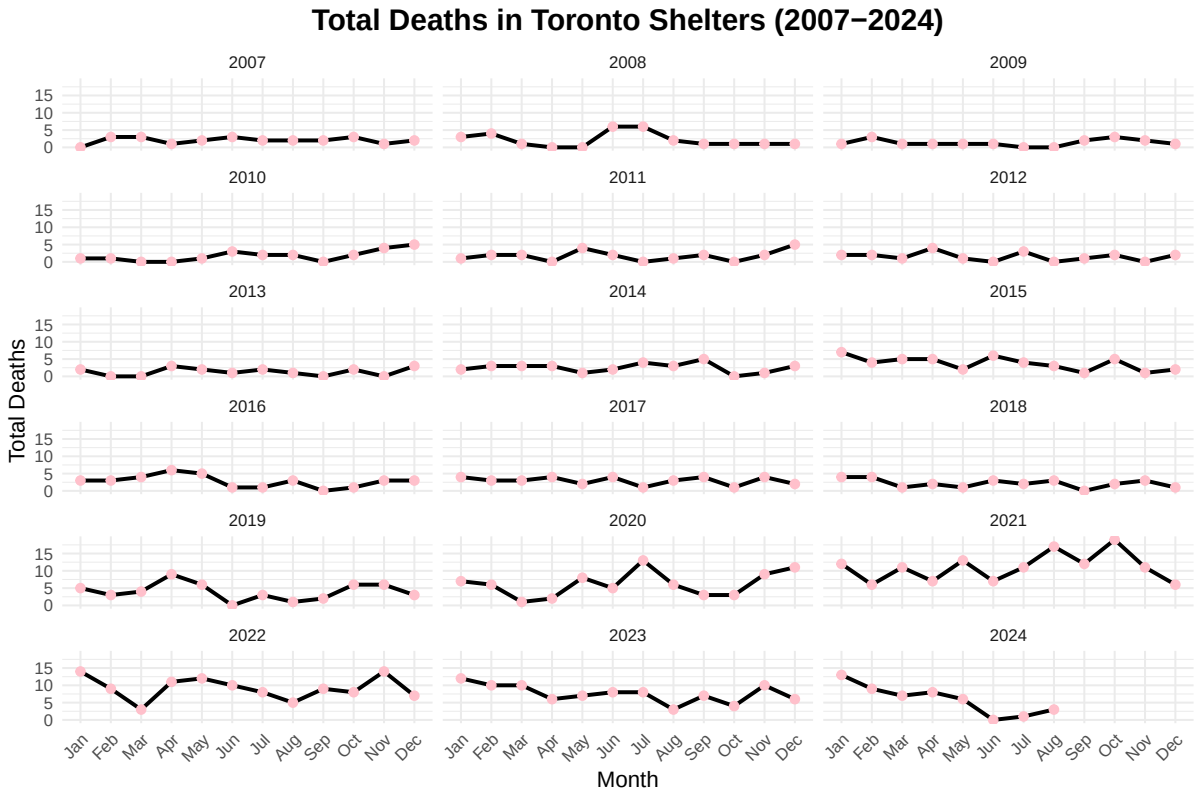

Figure 1: **?(caption)**

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.

2.4.2 Gender Composition

Figure 2

2.4.3 Compariasion Over Year

2.4.4 Compariasion Over Month

Talk way more about it.

3 Model

The goal of our modelling strategy is twofold. Firstly,…

Here we briefly describe the Bayesian analysis model used to investigate… Background details and diagnostics are included in Appendix B.

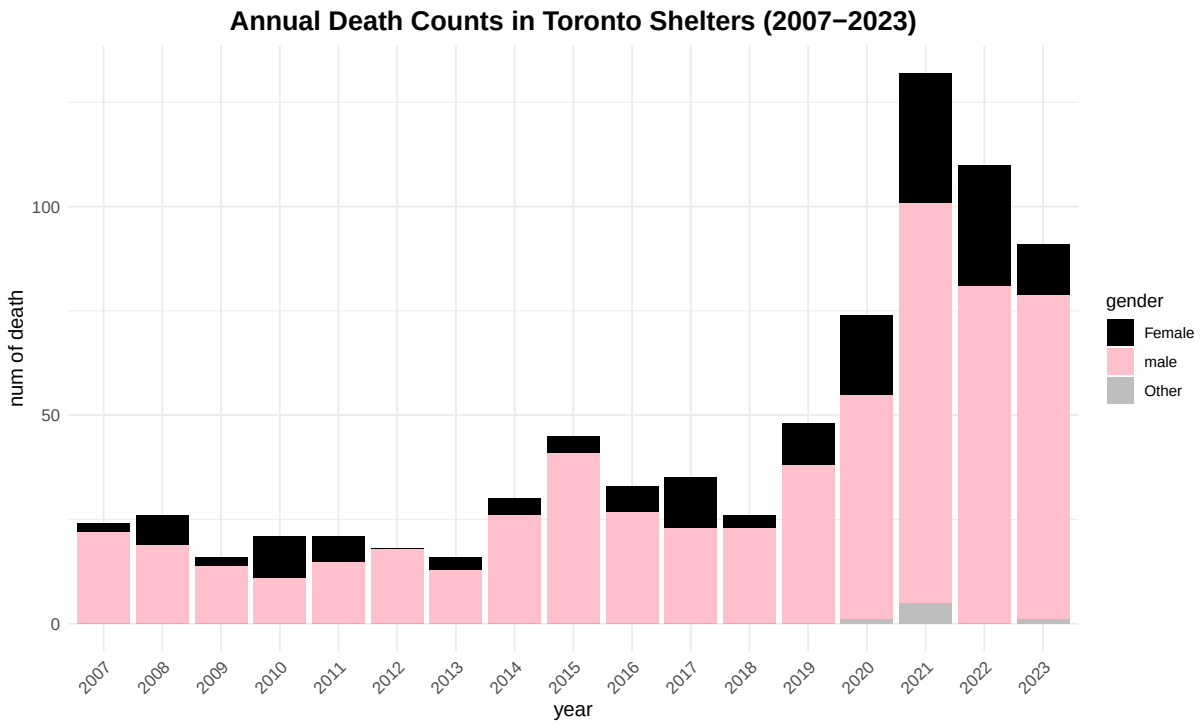

Figure 2: **?(caption)**

Figure 3: **?(caption)**

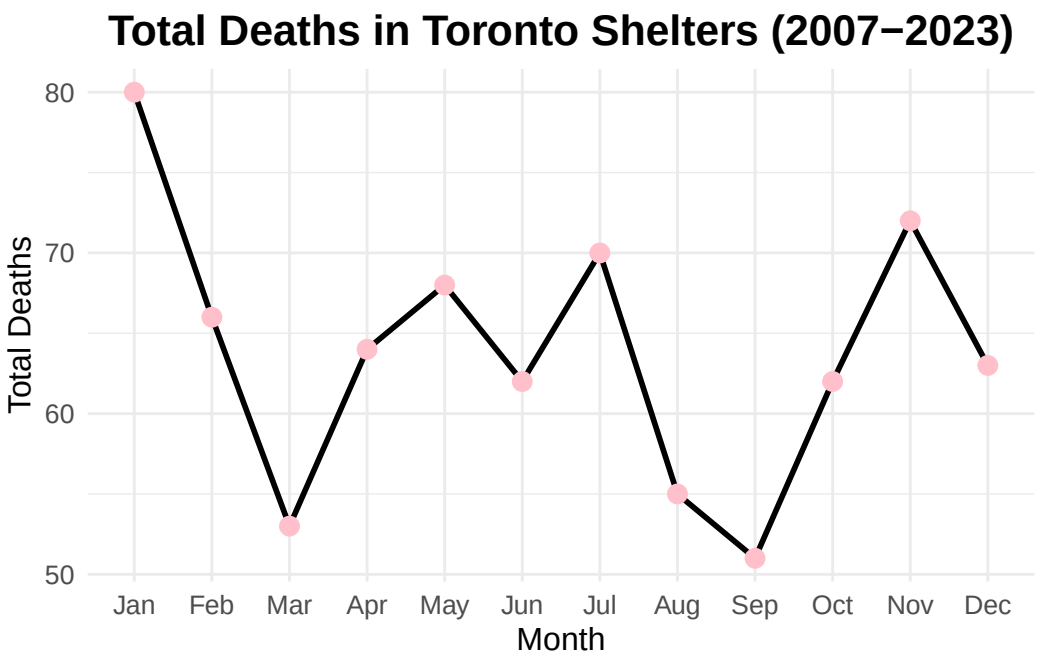

Figure 4: **?(caption)**

3.1 Model set-up

Define y_i as the number of seconds that the plane remained aloft. Then β_i is the wing width and γ_i is the wing length, both measured in millimeters.

$$y_i|\mu_i, \sigma \sim \text{Normal}(\mu_i, \sigma) \tag{1}$$
$$\mu_i = \alpha + \beta_i + \gamma_i \tag{2}$$
$$\alpha \sim \text{Normal}(0, 2.5) \tag{3}$$
$$\beta \sim \text{Normal}(0, 2.5) \tag{4}$$
$$\gamma \sim \text{Normal}(0, 2.5) \tag{5}$$
$$\sigma \sim \text{Exponential}(1) \tag{6}$$

We run the model in R (R Core Team 2023) using the `rstanarm` package of . We use the default priors from `rstanarm`.

3.1.1 Model justification

We expect a positive relationship between the size of the wings and time spent aloft. In particular...

We can use maths by including latex between dollar signs, for instance θ.

4 Results

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

5 Discussion

5.1 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.

5.2 Second discussion point

5.3 Third discussion point

5.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

Appendix

A Additional data details

B Model details

B.1 Posterior predictive check

In **?@fig-ppcheckandposteriorvsprior-1** we implement a posterior predictive check. This shows…

In **?@fig-ppcheckandposteriorvsprior-2** we compare the posterior with the prior. This shows…

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

Figure 5: **?(caption)**

B.2 Diagnostics

?@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 6: **?(caption)**

References

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