Comparison of Toronto Number of Deaths of Shelter Residents before COVID-19 and after COVID-19*

Wen Han Zhao

September 24, 2024

In this report, Toronto's Number of Deaths of Shelter Residents is analyzed from January 2007 to August 2024. The Coronavirus pandemic started in 2019 and has caused an outbreak globally. By comparing before and after COVID-19, this research aims to identify any significant trend and changes in deaths among Toronto shelter residents. The findings will contribute to future precautions by understanding the effects of COVID-19 on shelter's residents and inform public health organizations in shelter's environments.

Table of contents

*A GitHub Repository containing all data, R code, and other files used in this investigation is located here: https://github.com/younazhao/Death-of-shelter-residents.git

1 Introduction

According to World Health Organization, COVID-19 pandemic was first detected in China in December 2019 "COVID-19 - World Health Organization (WHO)" (2024) and spread rapidly to other countries. In 30 January 2020, WHO has declared Public Health Emergency of International Concern on coronavirus. As worldwide governments imposed lockdowns and protection measures to avoid the spreading of the virus, shelter's residents have been a population of high risk during this difficult time. As June of 2023, over 4 million cases of COVID-19 were reported in Canada and causing nearly 40,000 deaths "Wastewater Surveillance for COVID-19 in Shelters" (2024). People experiencing homelessness are more susceptible to COVID-19 infections due to inaccessibility to healthcare, overcrowded living conditions, and lack of protection measures.

The shelter's residents living environment, interaction with others, and accessibility to resources can increase their chance of exposing to the virus. The shelters' condition has a high density of people, which is difficult to practice social distancing. In addition, poor ventilation and insufficient infection control accelerated the spread of COVID-19 among shelter's residents. The Canadian government has imposed a wastewater surveillance strategy to detect the presence of any virus. Hopefully with the help of this analysis, the government could develop more strategy to prevent future outbreak and access to more public resources.

To do this, Toronto Number's of Death of Shelter Residents from 2007 to 2024 was analyzed in Section 2. The provided dataset from Toronto Public Health has shown a significant increase in trend near October of 2021 and has spiked to a maximum of total 19 deaths. Section 3 Section 3 focus on the weakness of this study and further possible improvements. Lastly, Section 4 Section 4 contains several steps performed on the raw data.

2 Data

2.1 Cleaned Data

" Deaths of Shelter Residents" is a dataset on Toronto Open Data Gelfand (2022) that provides monthly records of deaths of Toronto shelters' residents. Toronto Public Health started to track this data since January of 2007 until now. The dataset categorizes death by gender, and transgender/Non-binary/Two-spirit Gelfand (2022). In addition, I have included a new column of date for graphing purpose and classify Transgender/Non-binary/Two-spirit as other. Table 1 contained a detailed description of the dataset and variables.

Table 1: Dataset descriptions of the analysis data

Variables	Descriptions
Year	The calendar year being reported on.
Month	The month being reported on.
Total_Deaths	The total number of shelter residents who died in the reported month/year.
Male_Deaths	The total number of male shelter residents who died in the reported month/year.
Female_Deaths	The total number of female shelter residents who died in the reported month/year.
Others	The total number of Transgender, non-binary, and Two-Spirit shelter residents who died in
Date	Date combined Year and Month and adding in the 1st day of the month

2.2 Overview

Using the R programming language R Core Team (2023), the packages `janitor` Firke (2023) and `tidyverse` Wickham et al. (2019a) packages were used to simulate and testing the dataset. The dataset was download from Toronto Public Health by using `opendatatoronto` Gelfand (2022) and `tidyverse` Wickham et al. (2019a). The pacakages `janitor` Firke (2023) and `tidyverse` Wickham et al. (2019a) were used to clean the raw dataset and `lubridate` Grolemund and Wickham (2011) `dplyr` Wickham et al. (2023) were used to create new variables. Lastly, the packages `tidyverse` Wickham et al. (2019a) were used to perform tests on the clean dataset. `ggplot2` Wickham (2016) were used to generate plots for the report. `scales` Wickham and Seidel (2022), `knitr` Xie (2014), and `kableExtra` Zhu (2024) were used to generate tables and formatting.

2.3 Results

After loading the dataset using the R programming language R Core Team (2023) and ggplot2 Wickham (2016), and tidyverse Wickham et al. (2019a) package was used to generate graphs. In doing so, R code was adapted from Wickham et al. (2019b).

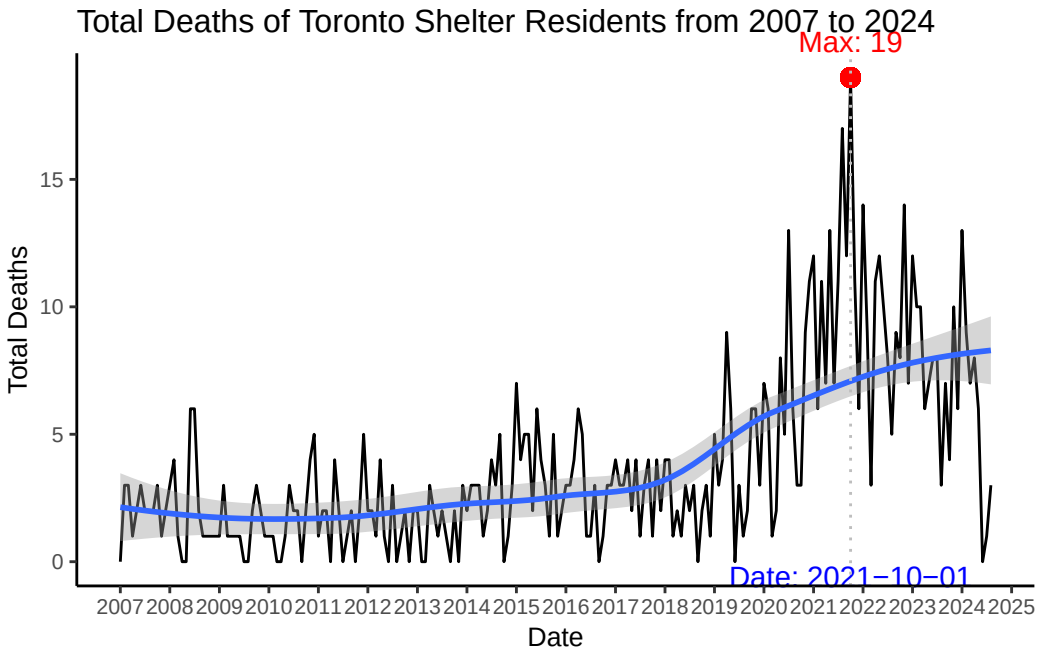

Figure 1: Deaths of shelter residents over time

The above graph illustrates the total number of deaths among Toronto shelter residents from January 2007 to August 2024. We could see a clear increasing trend over time and some cyclic observation across months and years. We could splipt the data in several phase. The 1st phase is from 2007 to 2015, where the number of deaths fluctuated but we the trendline (blue line) remain flat. Nearly all observations are under 5 with some spontaneous spikes. The 2nd phase starts around 2015 to 2020, where the frequency and magnitude of death amounts began to rise. This slowly increase could be affected by the rising of living expense in Toronto, causing the numbers of homeless to increase and therefore total deaths to increase. From 2020 to 2022, a sharp increase is noticeable from the graph, which corresponds to the COVID-19 pandemic. The numbers of deaths spikes significantly during this period almost reaching to 10-15 deaths per months. It has reached the maximum of 19 deaths in October 2021. This part suggest how COVID-19 has impacted the shelter residents enormously in Toronto. The 3rd phase is post 2022 which followed the pandemic's peak. The death numbers of residents shelter became more volatile but the trendline still shows a upward increasing trend. There are several small peaks during this phase but have decreased more than during pandemic.

From the trendline (blue line), we can observed an upward increasing trend with an acceleration during the pandemic period. This further aligns with our hypothesis where COVID-19 has significantly impacted mortality rate in shelter residents.

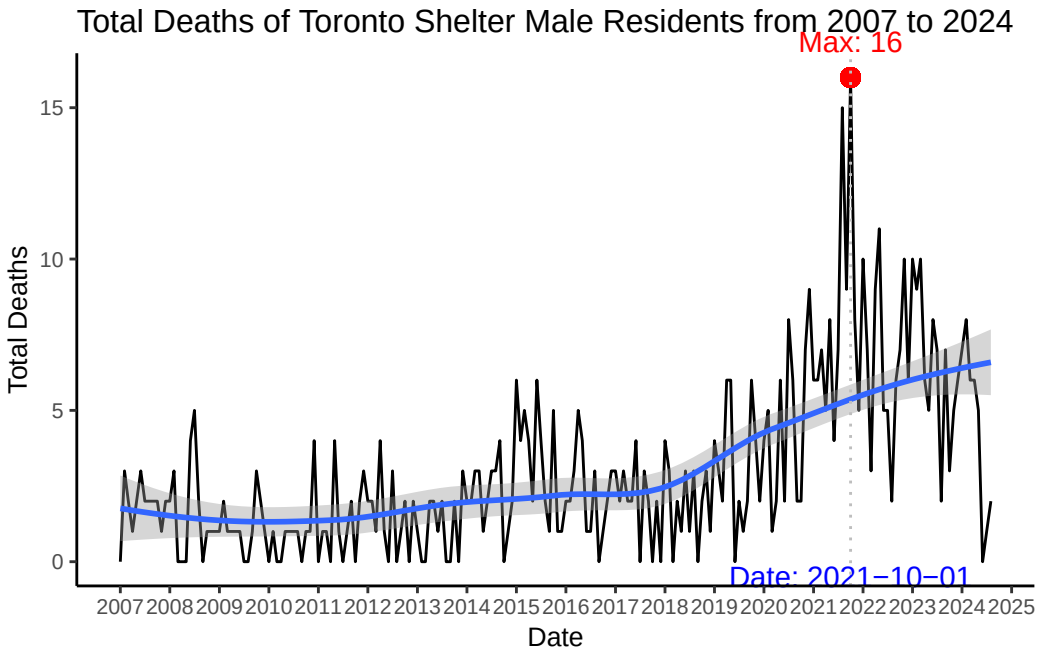

Figure 2: Deaths of male shelter residents over time

From the above 2 graphs illustrates the total number of deaths among Toronto male and female shelter residents from 2007 to 2024. Both graphs follow the same pattern as the number of total deaths. I have not included a seperate graphs on Others since the data is composed of many N/A and it is only a small proportion of the entire dataset. By comparing the 2 graphs, we can see that male shelter residents have a higher mortality compared to female shelter residents. During the first phase, we could still see volatility in the 1st phase. However, the number of deaths in female shelter's residents are more flat compared to the male and overall decedents. Especially in 2012 to 2013, where the number of female shelter's residents remained 0 for the entire year. During the pandemic, male residents shelter spiked to a maximum of 16 deaths in October 2021. Female residents shelter has a much lower mortality compared to male, with a maximum of only 5 deaths per several months starting in 2020 to 2022. From both of the graphs, male shelter residents' plot matches the overall plot more closely compared to the female's plot. This could be due to the numbers of death in male outcome female and the percentage of male in shelter residents is also greater than female in shelter residents. Overall, from the trenline and both graph's trend, it shows clear evidence an increase in deaths among shelter residents during the COVID-19 pandemic.

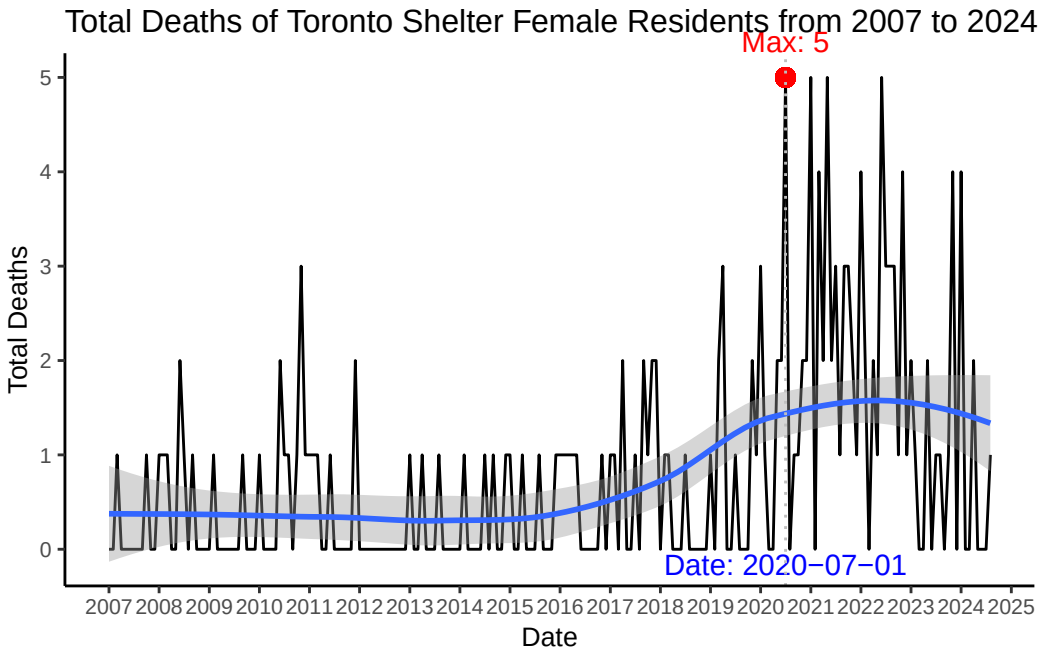

Figure 3: Deaths of female shelter residents over time

3 Discussion

3.1 Limitation

The analysis of "Deaths of Shelter Residents" dataset reveals a significant increase trend in mortality rates among shelter residents in Toronto during the pandemic. One limitation of this analysis could be the tracking of deaths number in shelters. Sometimes, this population might not always be able to have a spot in shelters and could wonder near the shelters. If a death occurred, Toronto Public Health might not be accurately report the number of deaths in that specific month. In addition, the dataset might also have gaps or inconsistencies during certain time periods, which could limit our ability to make a observation or judgement.

3.2 Suggestion for Future Step

For future consideration, it would be much more helpful to expand the analysis by combining other important datasets. For instance, the demographics of homeless population and shelters in Toronto, the cause of deaths in homeless population, or the population of shelter residents compared with Toronto overall population could all be helpful to identify our underlying assumptions. Lastly, we were able to observe a cyclic trend during in our graph. It would be a great if we could model the trend into a time series model. This could help warn the public

health institution for a cycle of peaks in shelter residents mortality rate and prepare for any upcoming disaster.

4 Appendix

4.1 Dataset and Graph Sketches

Sketches depicting in this analysis are available in the GitHub Repository.

4.2 Data Cleaning

The data cleaning process involved renaming column names and creating new variables from the raw dataset for clarity and simplicity.

References

"COVID-19 - World Health Organization (WHO)." 2024. 2024. https://www.who.int/europe/emergencies/situations/covid-19.

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.

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/.

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

"Wastewater Surveillance for COVID-19 in Shelters." 2024. 2024. https://www.canada.ca/en/public-health/services/reports-publications/canada-communicable-disease-report-ccdr/monthly-issue/2024-50/issue-1-2-january-february-2024/wastewater-surveillance-covid-19-shelters.html.

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

———, et al. 2019a. "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, and Dana Seidel. 2022. *Scales: Scale Functions for Visualization*. https://CRAN.R-project.org/package=scales.

Xie, Yihui. 2014. "Knitr: A Comprehensive Tool for Reproducible Research in R." In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC. http://www.crcpress.com/product/isbn/9781466561595.

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