# The Impact of Seasonal Changes on Mortality Rates in Toronto*

## Death Registry Statistics in Toronto

Doran Wang

September 27, 2024

This paper analyzes death toll data across four civic centers (Scarborough, North York, Toronto and Etobicoke)) from 2011 to 2024, with the goal of identifying seasonal patterns and regional differences in mortality, which is made available to the public from Open Data Toronto. Utilizing death registry data, we categorized the deaths by season—Winter, Spring, Summer, and Fall—and aggregated the death tolls across the years for each region. The analysis revealed notable variations in mortality trends across both seasons and civic centers. The results show that Winter consistently reported higher death tolls compared to other seasons across most regions. Using visual representations generated through ggplot2, the findings highlight significant regional disparities, which suggest possible influences of environmental or socio-economic factors contributing to seasonal mortality trends.

## Table of Contents



*Code and data are available at: https://github.com/Wang20030509/Sta304-Term-Paper-1

# 1 Introduction

Seasonal variations in mortality have been a focus of public health research for many years, driven by fluctuations in temperature, disease prevalence, and environmental factors. In colder regions, such as Toronto, winter months often coincide with increased mortality, particularly among vulnerable populations like the elderly or homeless. Numerous studies have linked these seasonal changes to temperature variations, with colder months typically showing elevated mortality rates (Madaniyazi, Chung, et al. 2021) (Madaniyazi, Armstrong, et al. 2021)

This study examines the impact of seasonal changes on death tolls in Toronto between 2011 and 2024, utilizing death registry data from Open Data Toronto. By categorizing deaths by season—Winter, Spring, Summer, and Fall—and analyzing regional variations across the city's four civic centers (Scarborough, North York, Toronto, and Etobicoke), the paper aims to uncover significant trends in mortality rates throughout the year. Previous studies have shown that temperature plays a significant role in seasonal mortality (Martin et al. 2012) (Zhang et al. 2024), with cold winters and heat waves contributing to an increased death toll, but the exact influence of socio-economic and environmental factors at the local level remains under-explored (Madaniyazi, Chung, et al. 2021) (Madaniyazi, Armstrong, et al. 2021).

Past research has focused heavily on the role of extreme weather events and temperature variability in driving mortality rates. Findings suggest that colder months are associated with higher mortality, particularly due to cardiovascular and respiratory issues exacerbated by low temperatures (Madaniyazi, Chung, et al. 2021). This paper builds on those findings by exploring how these seasonal dynamics manifest in a large urban environment like Toronto, where varying environmental and socio-economic conditions across different neighborhoods might influence mortality outcomes (Martin et al. 2012).

The results of this study will offer new insights into seasonal mortality trends within Toronto and inform public health initiatives aimed at mitigating mortality spikes during colder months. Understanding these patterns is crucial for guiding policy decisions and healthcare resource allocation during high-risk periods (Zhang et al. 2024).

The remainder of this paper is structured as follows. Section 2 describes the dataset and cleaning methods used in the analysis. Section 3 presents the results and key findings, and Section 4 offers a discussion of the implications of these results.Then, **?@sec-conclusion** will be the conclusion

# 2 Data

The datasets used for this report are the latest delay statistics from the 2024 year, `Death Registry Data` (City of Toronto Open Data Team 2024), which was published by the City Clerk's Office from `Open Data Toronto` (Gelfand 2022).

These deaths registry statistics were downloaded, cleaned, parsed, analyzed, and visualized using R (R Core Team 2023), a statistical programming language, with package support from `tidyverse` (Wickham et al. 2019), a collection of libraries which included the following packages that were utilized:

- `ggplot2` (Wickham 2016)
- `dplyr` (Wickham et al. 2023)
- `readr` (Wickham, Hester, and Bryan 2023)
- `tibble` (Müller and Wickham 2023)

For additional assistance with cleaning, the `janitor` (Firke 2023) package was used. For additional assistance with report generation the `knitr` (Xie 2023) package was used.

## 2.1 Overview Data

Table 1: Sample Data

| _id | CIVIC_CENTRE | DEATH_LICENSES | PLACE_OF_DEATH | TIME_PERIOD |
|---|---|---|---|---|
| 27767 | ET | 69 | Outside City Limits | 2011-01 |
| 27768 | ET | 341 | Toronto | 2011-01 |
| 27769 | NY | 141 | Outside City Limits | 2011-01 |
| 27770 | NY | 540 | Toronto | 2011-01 |
| 27771 | SC | 129 | Outside City Limits | 2011-01 |
| 27772 | SC | 545 | Toronto | 2011-01 |

The Death Registry Statistics dataset provided by the City of Toronto Open Data team contains key information about death registrations across various civic centers in Toronto. The dataset includes the following columns:

- CIVIC_CENTRE: This column represents the civic center code, corresponding to one of Toronto's four main civic centers: Scarborough (SC), North York (NY), Toronto (TO), and Etobicoke (ET). These civic centers are used to geographically categorize the death registrations.
- DEATH_LICENSES: This column records the number of deaths registered within a given month at each respective civic center. It represents the official count of death licenses issued, reflecting the mortality data across the city over different time periods.
- PLACE_OF_DEATH: This column provides information on where the death occurred. It indicates whether the death took place inside the City of Toronto or outside its city limits. This information can help to identify trends in deaths occurring within the jurisdiction of the city as opposed to those occurring elsewhere.
- TIME_PERIOD: This column captures the specific month and year during which the death licenses were registered. The time period is formatted as "YYYY-MM", allowing for temporal analysis of death registrations over time. The combination of month and year in this column enables the dataset to be used for tracking changes in death registration patterns on a monthly and yearly basis.

The death registry that Table 1 samples contains 947 entries in total. The dataset provides valuable insights into the geographic and temporal distribution of death registrations across Toronto, which is the main dataset used to analyze seasonal mortality patterns and potential regional disparities from 2011 to 2024.

### 2.2 Cleaned Data

Table 2: Sample Cleaned Data

| ID | Civic Centre | Death Toll | Place of Death | Year | Month | Season |
|---:|---|---:|---|---|---|---|
| 27767 | ET | 69 | Outside City Limits | 2011 | Jan | Winter |
| 27768 | ET | 341 | Toronto | 2011 | Jan | Winter |
| 27769 | NY | 141 | Outside City Limits | 2011 | Jan | Winter |
| 27770 | NY | 540 | Toronto | 2011 | Jan | Winter |
| 27771 | SC | 129 | Outside City Limits | 2011 | Jan | Winter |

To prepare the Death Registry Statistics dataset for analysis, several steps were taken to clean and standardize the data. These steps are essential to ensure consistency and accuracy, enabling meaningful comparisons across civic centers and time periods. The key cleaning procedures applied to the raw data include:

1. Renaming Columns: All column names were converted to more formal ones (e.g., death_licenses was renamed to Death Toll) for easier understanding in the table to the reader.
2. Splitting the TIME_PERIOD Column: The time_period column, originally formatted as "YYYY-MM", was split into two separate columns: year and month. This allowed for more granular temporal analysis and easier data aggregation by season.
3. Categorizing MONTH into Seasons: Based on the month column, a new season column was created to categorize each record into one of the four seasons:
    - Winter (December, January, February)
    - Spring (March, April, May)
    - Summer (June, July, August)
    - Fall (September, October, November)
4. Standardizing PLACE_OF_DEATH: The place_of_death column, which indicates whether the death occurred within or outside Toronto, was standardized. The text was converted to title case to ensure consistency (e.g., "Outside City Limits" and "Toronto").
5. Handling Missing or Invalid Data: Any missing or invalid entries in the dataset were either filled or removed where necessary to maintain the integrity of the analysis. For example, rows with missing or invalid death license counts were excluded from the final dataset.

After applying these cleaning procedures, the dataset was saved in its cleaned form as cleaned_data.csv, and it is now ready for use in subsequent analysis. This cleaned version of the data contains reliable and well-structured information on death registrations by season and civic center from 2011 to 2024.

### 2.3 Summarized Data

After cleaning and processing the Death Registry Statistics dataset, the data was aggregated and summarized to provide insights into seasonal and geographic trends in mortality across Toronto from 2011 to 2024. The summary tables are Total Death Toll by Season and Civic Centre, Average Deaths by Month, Death Toll by Place of Death, Death Toll by Civic Centre. As shown in Table 3 and for more detailed information, please refer to the Appendix. These summary tables are primarily intended to facilitate the plotting and analysis in the Results section.

# 3 Results

Here are 3 part of result with figures

## 3.1 Distribution of deaths in different civic centers by season

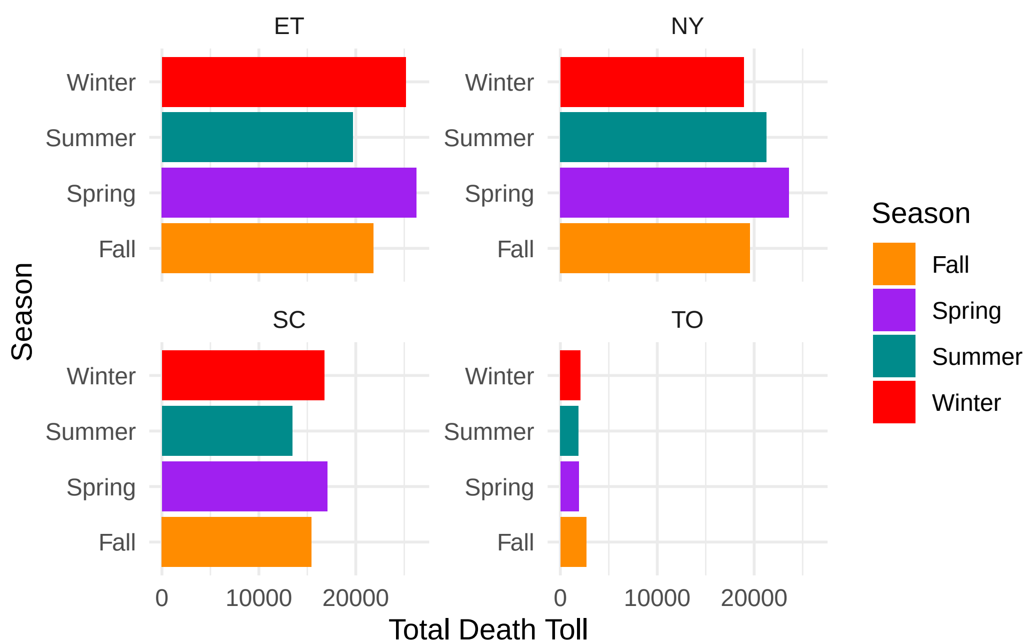


Figure 1: Total Death Toll by Season and Region (2011-2024)

As shown in Figure 1 The total death toll in each season reveals notable seasonal differences. Winter consistently reported higher death tolls, particularly in Etobicoke (ET) and North York (NY), compared to the other civic centers. This pattern suggests a potential link between colder weather and increased mortality, aligning with previous studies on the impact of temperature on health outcomes.

## 3.2 Distribution of mean of deaths overall months

As shown in Figure 2 January consistently reported the highest average death toll, with a mean of 336 (from Table 3, see more detail in Appendix)deaths. This could be attributed to the colder winter months in Toronto, which may exacerbate health conditions such as cardiovascular and respiratory issues. In contrast, the summer months, such as July and August, reported lower average death tolls.

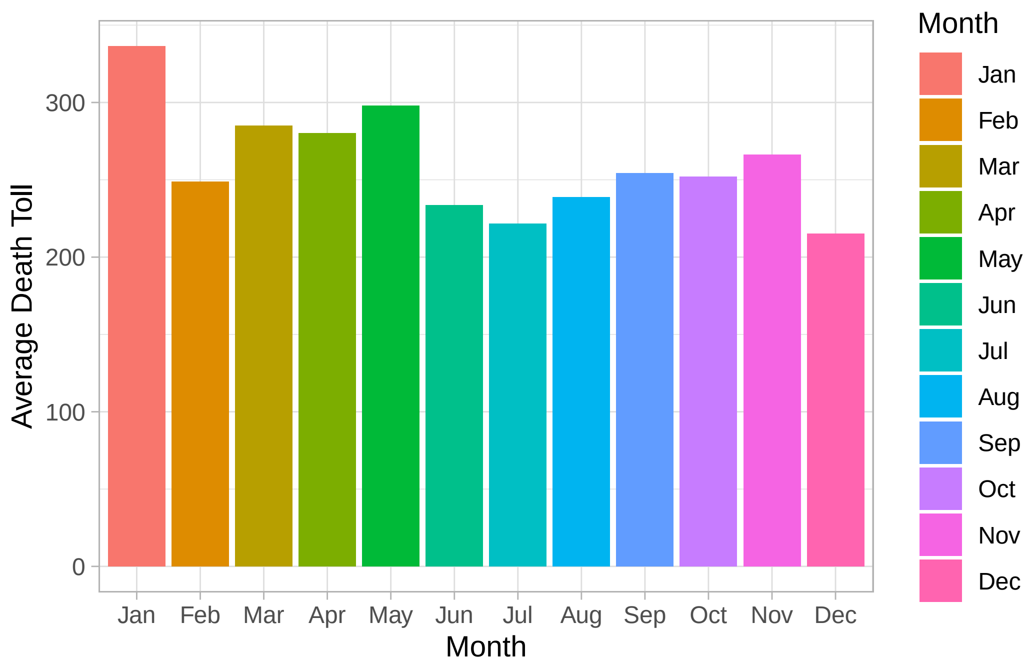


Figure 2: Average Death Toll by Month (2011-2024)

### 3.3 Total death toll overall seasons

We can see intuitively from Figure 3 that the death toll in winter and spring, two seasons with relatively low temperatures, is higher than that in summer and fall.Also, there are more deaths in fall than in summer, since the temperature gradually drops as fall begins.

### 3.4 Death toll trends vary by season

As Figure 4 and Figure 5 shows, regardless of whether the death was found outside the city or in the Toronto area, the death toll curve is lowest in the summer (June-August), and starts to rise in the winter until it reaches its highest point in January. Even when we distinguish by civic centre, the curve is similar and within some errors.

## 4 Discussion

The analysis of mortality trends in Toronto from 2011 to 2024 reveals significant seasonal and regional variations. Winter consistently shows the highest death tolls across all civic centers, particularly in Etobicoke and North York. This finding aligns with established research linking colder temperatures to increased mortality, especially among vulnerable populations such as

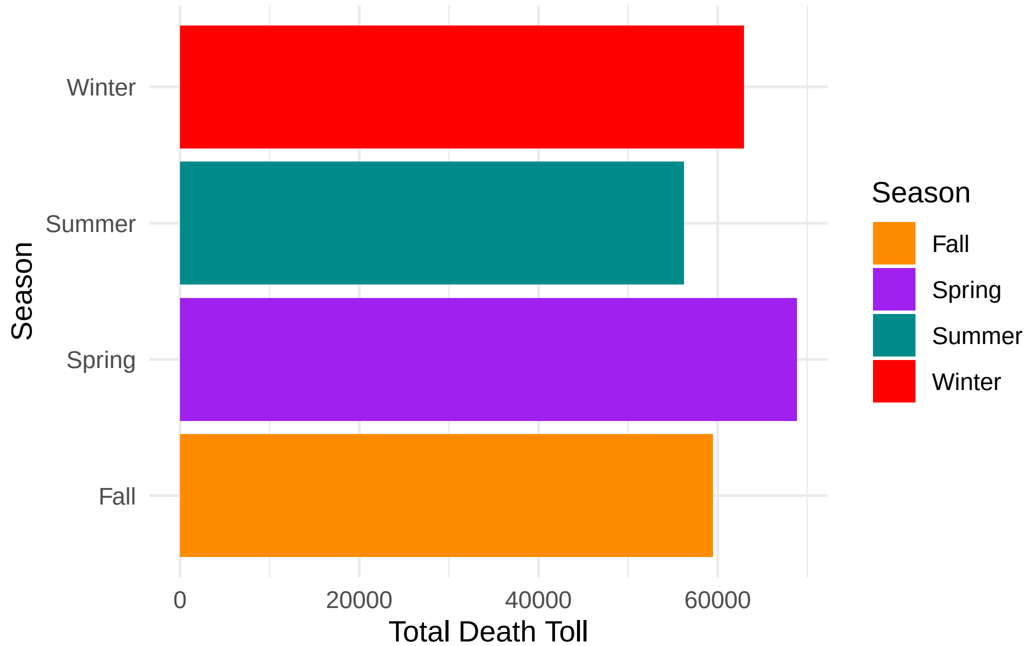


Figure 3: Total Death toll by Season (2011-2024)

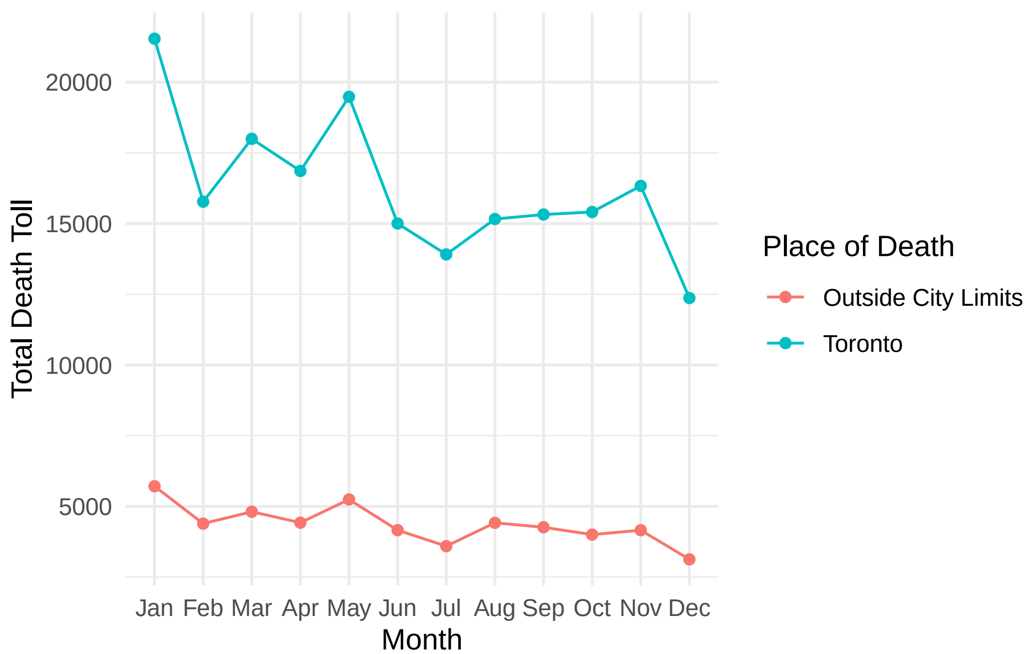


Figure 4: Death Toll by Place of Death Across 12 Months (2011-2024)

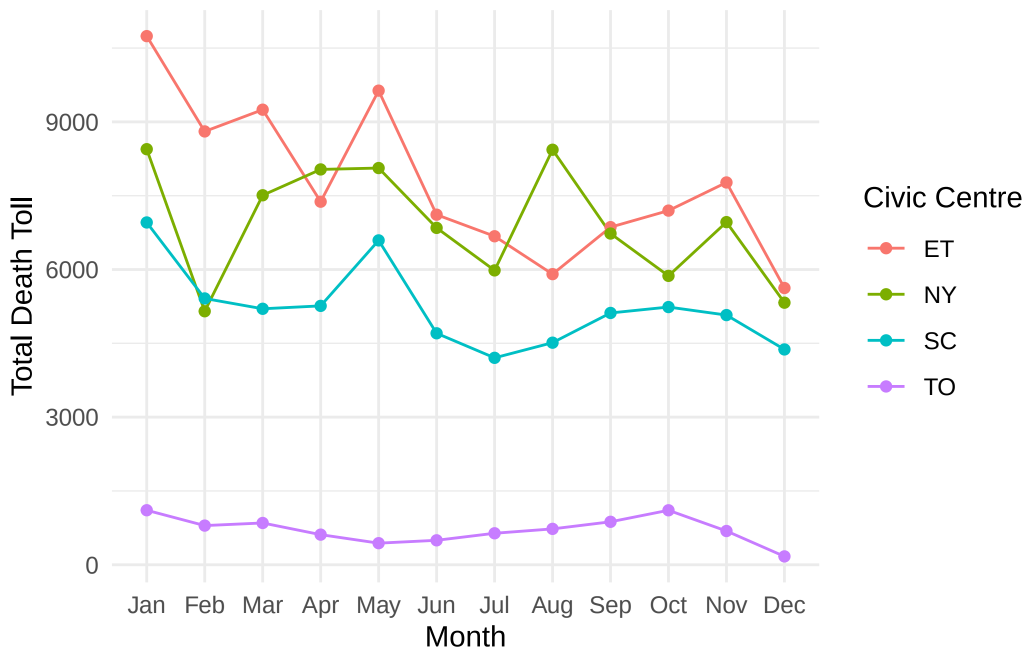


Figure 5: Death Toll by Civic Centre Across 12 Months (2011-2024)

the elderly and those with pre-existing health conditions. Cardiovascular and respiratory diseases, exacerbated by cold weather, likely contribute to these elevated death tolls in the winter months.

Spring also exhibits relatively high mortality rates, possibly due to the lingering effects of cold temperatures and the transition between seasons. In contrast, summer consistently reports the lowest death tolls, likely due to milder temperatures and lower incidence of seasonal illnesses. Interestingly, fall shows an increase in death tolls as temperatures begin to drop, signaling the onset of winter-related health complications.

Regional disparities in mortality were evident, with Etobicoke consistently reporting the highest death tolls across all seasons. This may be attributed to socio-economic and environmental factors specific to the region, such as access to healthcare, population density, and the presence of vulnerable populations. North York also showed similar trends, while Toronto and Scarborough generally reported lower death tolls.

## 5 Conclusion

This study provides valuable insights into the seasonal and regional patterns of mortality in Toronto over a 13-year period. The findings reinforce the significant impact of cold weather on mortality, particularly in winter and spring, and highlight the need for targeted public health interventions during these high-risk periods. The regional disparities observed suggest that localized factors play a crucial role in influencing mortality rates, indicating the need for region-specific strategies to reduce mortality.

# Appendix

## A Additional table details

As shown in Table 3, the dataset was grouped by season and civic center to examine the variations in mortality over time. As shown in Table 4, an analysis of the monthly average death toll reveals interesting patterns throughout the year. As shown in Table 5, the dataset was further analyzed to examine the location of death registrations. The summarized data by civic center (Table 6) shows that Etobicoke (ET) consistently recorded the highest death tolls in all months.

Table 3: Total Death toll by Season (2011-2024)

| Season | Civic Centre | Death Toll |
|---|---|---:|
| Fall | ET | 21827 |
| Spring | ET | 26261 |
| Summer | ET | 19696 |
| Winter | ET | 25173 |
| Fall | NY | 19565 |
| Spring | NY | 23605 |

Table 4: Average Deaths by Month (2011-2024)

| Month | Mean of the Death Toll |
|---|---:|
| Apr | 280.0658 |
| Aug | 238.8049 |
| Dec | 215.2222 |
| Feb | 248.9630 |
| Jan | 336.4444 |
| Jul | 221.5570 |

Table 5: Death Toll by Place of Death Across 12 Months

| Month | Place of Death | Death Toll |
|---|---|---:|
| Apr | Outside City Limits | 4424 |
| Aug | Outside City Limits | 4420 |
| Dec | Outside City Limits | 3128 |
| Feb | Outside City Limits | 4391 |
| Jan | Outside City Limits | 5713 |

Table 5: Death Toll by Place of Death Across 12 Months

| Month | Place of Death | Death Toll |
|---|---|---|
| Jul | Outside City Limits | 3594 |

Table 6: Death Toll by Civic Centre Across 12 Months

| Month | Civic Centre | Death Toll |
|---|---|---|
| Apr | ET | 7378 |
| Aug | ET | 5906 |
| Dec | ET | 5624 |
| Feb | ET | 8806 |
| Jan | ET | 10743 |
| Jul | ET | 6676 |

## References


City of Toronto Open Data Team. 2024. *Death Registry Statistics Dataset*. City of Toronto. https://open.toronto.ca/dataset/death-registry-statistics/.

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.

Madaniyazi, Lina, Ben Armstrong, Yeonseung Chung, Chris Fook Sheng Ng, Xerxes Seposo, Yoonhee Kim, Aurelio Tobias, et al. 2021. "Seasonal variation in mortality and the role of temperature: a multi-country multi-city study." *International Journal of Epidemiology* 51 (1): 122–33. https://doi.org/10.1093/ije/dyab143.

Madaniyazi, Lina, Yeonseung Chung, Yoonhee Kim, Aurelio Tobias, Chris Fook Sheng Ng, Xerxes Seposo, Yuming Guo, et al. 2021. "Seasonality of mortality under a changing climate: a time-series analysis of mortality in Japan between 1972 and 2015." *Environmental Health and Preventive Medicine* 26 (1): 69. https://doi.org/10.1186/s12199-021-00992-8.

Martin, Sarah L., Christopher Alan Hebbern, Mary-Luyza Avramescu, and Neil Tremblay. 2012. "Climate change and future temperature-related mortality in 15 Canadian cities." *International Journal of Biometeorology* 56 (4): 605–19. https://doi.org/10.1007/s00484-011-0449-y.

Müller, Kirill, and Hadley Wickham. 2023. *Tibble: Simple Data Frames*. https://CRAN.R-project.org/package=tibble.

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

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, Jim Hester, and Jennifer Bryan. 2023. *Readr: Read Rectangular Text Data*. https://CRAN.R-project.org/package=readr.

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

Zhang, Yuqing, Kai Wang, Junjie Ren, Yixuan Liu, Fei Ma, Tenglong Li, Ying Chen, and Chengxiu Ling. 2024. "Bivariate extreme value analysis of extreme temperature and mortality in Canada, 2000-2020." *BMC Public Health* 24 (1): 1344. https://doi.org/10.1186/s12889-024-18785-3.