Analyzing Geographic and Demographic Patterns of Homelessness*

Exploring Spatial Variations in Homeless Shelter Demand and Gender/Age-Based Disparities

Jiwon Choi

January 25, 2024

This research investigates the difficulties encountered by the homeless, even with available shelters, motivated by a recent case (Barghiel 2024) of a homeless individual dying due to extreme cold. By conducting geographic and demographic patterns of shelter needs, using data from the Shelter Management Information System, it pinpoints areas with significant demand and uncovers trends in shelter use based on gender and age. The results assist policymakers, shelter operators, and advocacy groups in effectively distributing resources and actively addressing homelessness issues.

Table of contents

---

*Code and data are available at: https://github.com/jwonc4602/Analyzing-Geographic-and-Demographic-Patterns-of-Homelessness.

1 Introduction

Homelessness in Toronto presents a challenging and dynamic issue, influenced by demographic shifts, economic factors, and policy decisions. Despite a robust shelter system, designed to protect the city's most vulnerable, instances of individuals not benefiting from these services and facing tragic outcomes still occur. This paper aims to analyze demographic and geographic trends in shelter systems to leverage them to resolve and support future homeless shelter issues in Toronto.

This study is designed for an in-depth analysis of Toronto's shelter occupancy and capacity, utilizing data (Data 2024a) from the Shelter Management Information System (SMIS) database. While there are several papers on homelessness, this paper more focus on data-driven analyses that integrate both demographic and geographic dimensions within the context of urban shelter systems. Other datasets have often been limited either in data scope or timeframe, failing to provide a picture that reflects the current realities of urban homelessness. This paper addresses this gap better by offering an updated view of the shelter system in Toronto.

In detail, this study analyzes shelter occupancy trends using a dataset (Data 2024a) from the Shelter Management Information System (SMIS) (Data 2024a), employing R (R Core Team 2023) statistical tools like `tidyverse` (Wickham et al. 2019) and `ggplot2` (Wickham 2016). Findings reveal and a heavy reliance on bed-based capacities mainly for non-family groups but not for families, alongside high occupancy rates across Toronto. These results are playing key roles for guiding policymakers and social welfare organizations in enhancing shelter capacities and accessibility, and in reevaluating strategies to more effectively address urban homelessness.

To dissect the demographic and geographic patterns within the shelter system in Toronto, spotlighting trends in service utilization and capacity, this paper is organized into the following sections: Data (Section 2), and Discussion (Section 3). In the Data section, I discuss a detailed description of the data source and collection methods. It then progresses to an analysis of demographic and geographic patterns in shelter usage, followed by a discussion on the observed trends and their broader implications. The final sections of the discussion part critically evaluate the study's limitations and propose directions for future research.

2 Data

2.1 Data Source

This research is based on a dataset sourced from the Shelter Support and Housing Administration (SSHA) division's Shelter Management Information System (SMIS) database, which is accessible through the City of Toronto's OpenDataToronto Library (Gelfand 2022). The specific dataset employed is named 'Daily Shelter & Overnight Service Occupancy & Capacity' (Data 2024a). This database acts as a detailed collection of daily entries related to active overnight shelters and associated services.

2.2 Data Collection

In the process of selecting an appropriate data source for an in-depth analysis of current demographic and geographic trends within shelter services, several datasets were considered. Two other datasets considered are the 'Toronto Shelter System Flow data' (Data 2024b) and the 'Daily Shelter Occupancy' (Data 2022) datasets. However, the 'Daily Shelter & Overnight Service Occupancy & Capacity' (Data 2024a) dataset was ultimately chosen for the following reasons:

1. While the 'Toronto Shelter System Flow' (Data 2024b) data offers valuable insights into the flow of demography through the shelter system, it lacks geographic information. This limitation significantly hinders the potential for a detailed analysis of geographic trends and patterns within the shelter system, an aspect critical to my research objectives.

2. The 'Daily Shelter Occupancy' (Data 2022) dataset, although previously a proper source of information, has become less relevant due to its discontinuation. The most recent update for this dataset was recorded on March 10, 2022, therefore it is outdated for current analysis purposes. In contrast, the 'Daily Shelter & Overnight Service Occupancy & Capacity' (Data 2024a) dataset provides the most recent and thorough data available, making it a more suitable choice for analyzing up-to-date trends and developments in shelter usage and capacity.

Consequently, the decision to utilize the 'Daily Shelter & Overnight Service Occupancy & Capacity' (Data 2024a) dataset was driven by its superior relevance, up-to-date, and the inclusion of both demographic and geographic data, which are essential for a thorough and current analysis of shelter services.

The methodology for gathering data is structured to ensure a detailed and current perspective of shelter and overnight service. The data was compiled and examined using the R statistical programming software (R Core Team 2023), supplemented by various tools such as `tidyverse` (Wickham et al. 2019), `ggplot2` (Wickham 2016), `dplyr` (Wickham et al. 2023), `readr`

(Wickham, Hester, and Bryan 2023), `tibble` (Müller and Wickham 2023), `janitor` (Firke 2023), `knitr` (Xie 2014), and `here` (Müller and Bryan 2020) (see Table 1).

Table 1: Sample of Cleaned Shelter Statistics Data

| occupancy_date	sector	capacity_type	location_city	occupancy_rate(%)
2024-01-01	Families	Room Based Capacity	Toronto	100
2024-01-01	Mixed Adult	Room Based Capacity	Toronto	100
2024-01-01	Mixed Adult	Bed Based Capacity	Toronto	100
2024-01-01	Families	Room Based Capacity	North York	100
2024-01-01	Families	Room Based Capacity	North York	100

2.3 Demographic Patterns of Homelessness
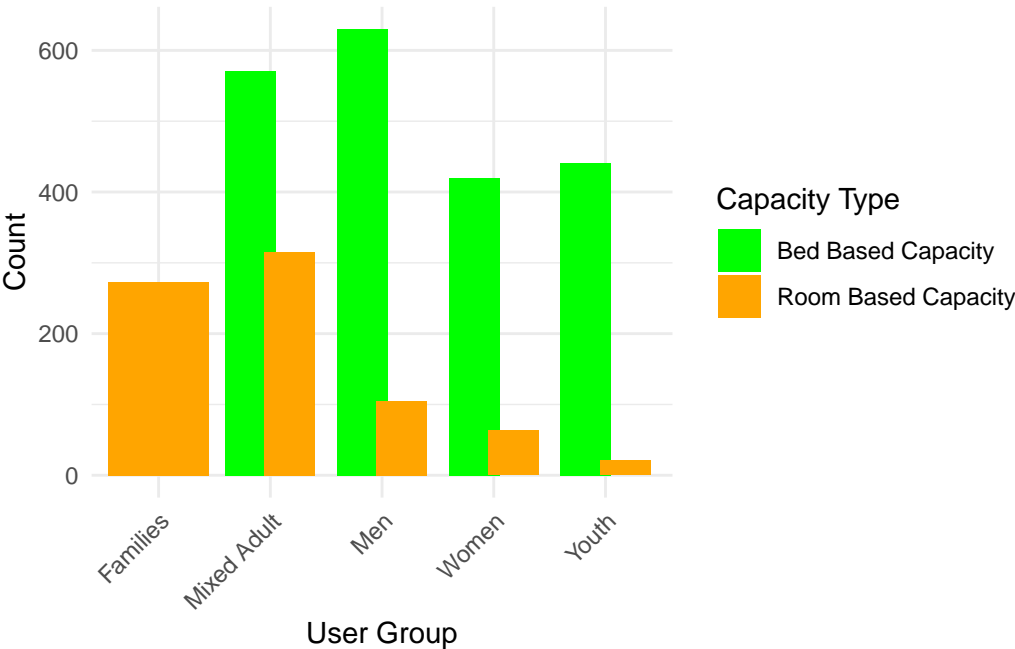
Figure 1: Number of Capacity Type by User Group

A focused analysis of the capacity types utilized in shelter services reveals distinct patterns in the allocation and utilization of resources, particularly when comparing bed-based and room-

based capacities across different user groups. Bed-based capacity is typically applicable for programs with common sleeping areas, while room-based capacity is measured at the room level where sleeping rooms are not shared by people from different households (see Figure 1).

A striking observation is the complete lack of bed-based capacity options for family groups. This suggests that the current shelter system does not offer bed-specific accommodations for families, possibly due to the unique needs and privacy concerns associated with family units.

In contrast to the situation with families, other user groups predominantly utilize bed-based capacity in shelters. This trend is significant, indicating that for individuals and non-family groups, bed-based accommodations are the primary form of shelter provided. This form of capacity is likely favored due to its efficiency in maximizing space and accommodating a larger number of individuals within the shelter system.

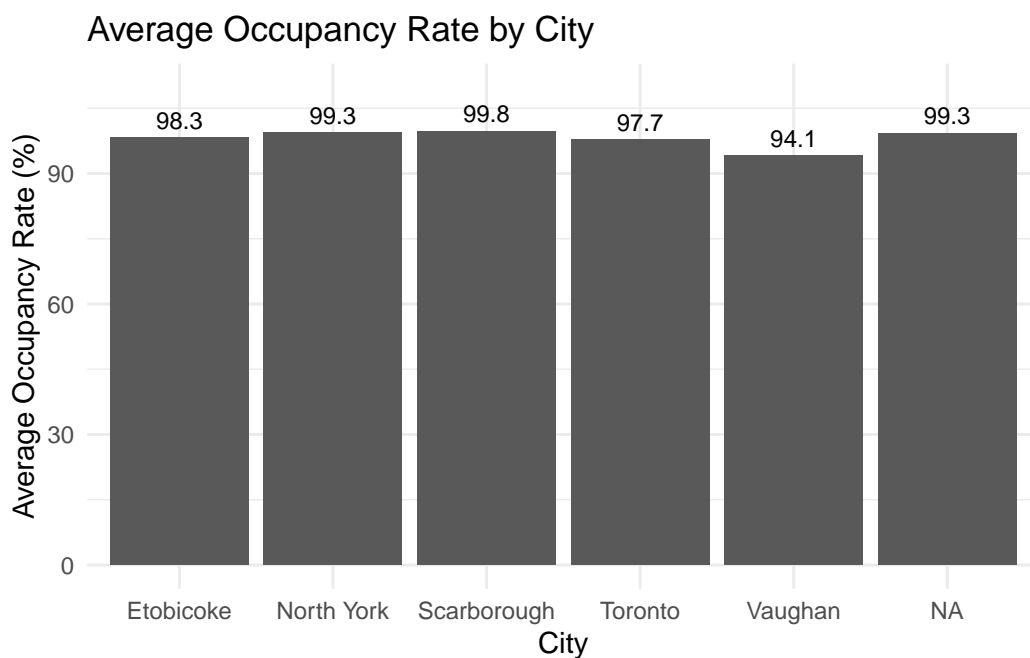## 2.4 Geographic Patterns of Homelessness

Figure 2: Average Occupancy Rate by City

Here are occupancy rates, rounded up at the first decimal place, across six categories: five cities—Etobicoke, North York, Scarborough, Toronto, Vaughan—and an unspecified category labeled "NA" (see Figure 2).

The narrow range between the highest and lowest occupancy rates and such high rates could imply limited vacancies and potential competitiveness in occupancy for homeless people seeking

accommodation. For policymakers, shelter operators, and advocacy groups, these figures might signal the need for increased shelter supply.

3 Discussion

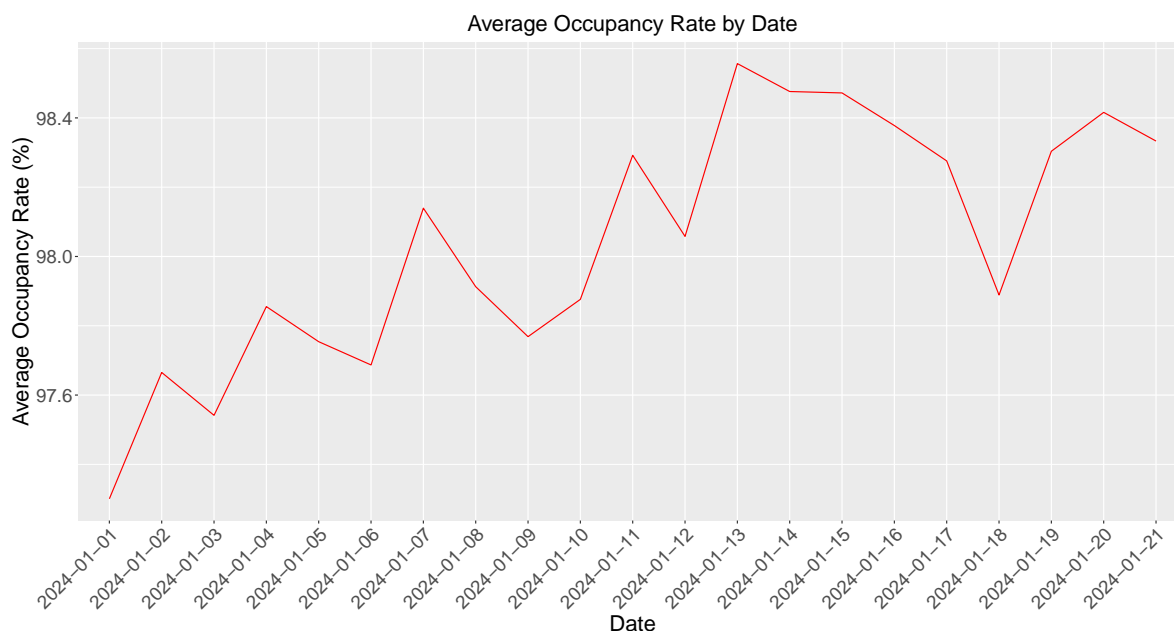## 3.1 Trends in Shelter Occupancy Rates by Dates

Figure 3: Average Occupancy Rate by Date

The latest analysis of shelter occupancy data reveals a marked upward trend in average occupancy rates, rounded up at the first decimal place, over 21 days from January 1st to January 21st in 2024 (see Figure 3). While daily fluctuations were observed, the general direction was a steady increase in the need for shelter services. However, to provide a more thorough understanding, several critical aspects need to be considered in the discussion:

1. **Overall Rising Trend and Demographic Patterns**: The data shows an overall increase in occupancy rates. This trend aligns with the demographic patterns of homelessness, suggesting a balanced need for both bed-based and room-based capacities in shelters. Notably, there's an obvious requirement for bed-based capacity, particularly for family user groups. This insight is crucial for shelter resources to meet the specific needs of different homeless populations effectively.

2. **Geographic Patterns of Homelessness and Occupancy Rate**: The occupancy rates across all listed cities are consistently high, indicating a widespread demand for

shelter services (see Figure 2). This requires effective management of existing resources in these areas. Additionally, providing safe travel options to the vacant shelters in other cities and access to education to go back to society for the homeless might be significant considerations. These services can play an important role in enhancing the quality of life for homeless individuals and families, and in some cases, may aid in reducing the incidence of homelessness.

3. **Possible Issues**: Despite the high occupancy rates, none of the shelters reached or exceeded 100% capacity. This observation is contrasted with the visible presence of many homeless individuals on the streets. This contradiction indicates potential issues with accessibility or awareness of shelter services. To address this gap, shelter operators and advocacy groups need to intensify their efforts. This could involve improving outreach programs, enhancing the accessibility of shelters, and removing barriers that prevent homeless individuals from seeking shelter. By doing so, more individuals can be brought under the protection of shelter programs, thereby reducing the number of homeless people without adequate shelter.

3.2 Limitations and Next Steps

3.2.1 Limitations

1. **Data Timeframe and Recency**: One of the primary limitations of this study is the timeframe of the data analyzed. The 'Daily Shelter & Overnight Service Occupancy & Capacity' (Data 2024a) data contains the period from January 1st to January 21st, 2024. While the data has been updated daily since 2024, this analysis does not account for any updates or changes that may have occurred beyond January 21st, 2024. This restriction could potentially lead to minor inaccuracies or the omission of recent developments that could impact the findings and their relevance to the current situation. Additionally, the absence of data before 2024 limits the ability to observe significant trends and changes over a longer historical period.

2. **Scope of Data Variables**: The analysis currently focuses on general occupancy rates without considering specific variables that could offer more detailed insights. For instance, the inclusion of variables such as the program model of shelters, the type of overnight services provided, or the program area could potentially yield more detailed and informative results. These factors can significantly influence shelter occupancy patterns, and their exclusion from the analysis may limit the depth and applicability of the conclusions drawn.

3.2.2 Next Steps

1. **Incorporation of Recent Data**: To address the limitation of the study's timeframe, future analyses should follow the most recent data update containing a wider timeframe available. This approach will ensure that the findings are up-to-date and reflect the current state of shelter occupancy. Regular updates to the analysis will also help in monitoring trends over time, providing a dynamic view of the homelessness situation.

2. **Expanding Data Parameters**: Future studies should aim to include additional variables such as the program model, service type, and program area of shelters. This expansion will allow for a more thorough analysis, potentially uncovering patterns and trends that are not evident from occupancy rates alone. For example, understanding how different program models affect occupancy rates can inform more targeted and effective policy decisions and resource allocation.

In conclusion, while this study provides important insights into shelter occupancy trends, its scope and the recency of data are somewhat limited. Future research efforts should aim to overcome these limitations by incorporating more thorough and up-to-date data, thereby enriching the understanding of homelessness and informing more effective responses to this pressing social issue.

References

Barghiel, Naomi. 2024. "'Frustrating and Deadly': How the Extreme Cold Is Hitting Canada's Homeless." https://globalnews.ca/news/10235191/homeless-protections-extreme-cold-canada/.

Data, Toronto Open. 2022. "Daily Shelter Occupancy." https://open.toronto.ca/dataset/daily-shelter-occupancy/.

———. 2024a. "Daily Shelter & Overnight Service Occupancy & Capacity." https://open.toronto.ca/dataset/daily-shelter-overnight-service-occupancy-capacity/.

———. 2024b. "Oronto Shelter System Flow." https://open.toronto.ca/dataset/toronto-shelter-system-flow/.

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.

Müller, Kirill, and Jennifer Bryan. 2020. *Here: A Simpler Way to Find Your Files.* https://cran.r-project.org/web/packages/here/index.html.

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, and Kirill Müller. 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. 2014. *Knitr: A Comprehensive Tool for Reproducible Research in R.* Edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC. http://www.crcpress.com/product/isbn/9781466561595.