Tracking Homelessness: A Statistical Analysis of Toronto's Shelter System Flow*

Jiaxuan Song

October 9, 2025

This report is based on the Toronto Shelter System Flow dataset from Toronto Open Data, providing an in-depth analysis of how the city's homelessness services are functioning. The data offers valuable insights into the movement of people within the shelter system, helping to measure Toronto's progress in reducing homelessness. The goal of this analysis is to better understand the persistence and patterns of homelessness, ultimately providing evidence to support efforts to reduce homelessness to rare, brief, and non-recurring instances.

Table of contents

*The GitHub Repository containing all data, R code, and other files used in this project is located here: https://github.com/Jiaxuan-Song/STA304-Paper1-Initial-Submission.git

1 Introduction

As society continues to evolve, homelessness has become a major social challenge faced by cities around the world, including Toronto. The causes behind it are complex and multifaceted, typically involving personal, economic, social, and structural factors. Economic difficulties, such as low income, unemployment, and high housing costs, are among the primary reasons for homelessness. Additionally, personal factors such as mental health issues, substance abuse or addiction, serious illnesses, and family breakdowns also increase the risk of individuals becoming homeless (Nishio et al. 2017). Chronic homelessness not only has a severe impact on an individual's physical and mental health but also places a significant burden on healthcare systems, social welfare services, and public resources. People experiencing homelessness are typically at higher risk for illness and have a lower quality of life, which further complicates their ability to escape poverty and secure stable housing (Goodman, Saxe, and Harvey 1991). Given these challenges, addressing homelessness has become increasingly urgent.

In response to this issue, the City of Toronto has implemented a series of measures specifically designed to support individuals experiencing homelessness. These measures include emergency shelters, respite services, and additional programs such as hotel/motel initiatives and warming centers, aimed at providing essential shelter and resources for those in need (Hwang 2000). However, offering shelter alone is not enough. To better address homelessness and develop effective policies, the city tries to create an statistical system to track and manage individuals who access these services. Through this system, Toronto can not only monitor trends in homelessness in real-time but also allocate resources more accurately, helping vulnerable populations move toward long-term stability. This system provides crucial data that enables the government to tackle homelessness in a more targeted and effective manner. In the future, expanding the scope of support services, improving service quality, and collaborating with community organizations will be crucial steps in reducing homelessness (Gaetz 2020).

2 Data

2.1 Overview

In this report, the dataset used is the "Toronto Shelter System Flow" dataset from the City of Toronto Open Data platform Gelfand (2022) . This dataset provides information on individuals entering and exiting the shelter system each month, including the number of unique individuals who have used the shelter system at least once in the past three months and are still considered homeless (i.e., they have not been placed in permanent housing). Additionally, this dataset is open for copying, redistribution, and use for any purpose, as long as proper credit is given, a link to the license is provided, and any modifications are indicated.

The dataset analyzed in this report contains 604 entries and 20 variables, covering homelessness-related information in Toronto. It includes details such as the date, population groups (e.g.,

All Population, Chronic, Refugees, Families, Youth), and various metrics like the number of individuals returning from housing, newly identified as homeless, moved to housing, or became inactive. Age categories are broken down into groups ranging from under 16 to 65 and over, along with gender distribution for males, females, and those identifying as transgender, non-binary, or Two-Spirit.

2.2 Results

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

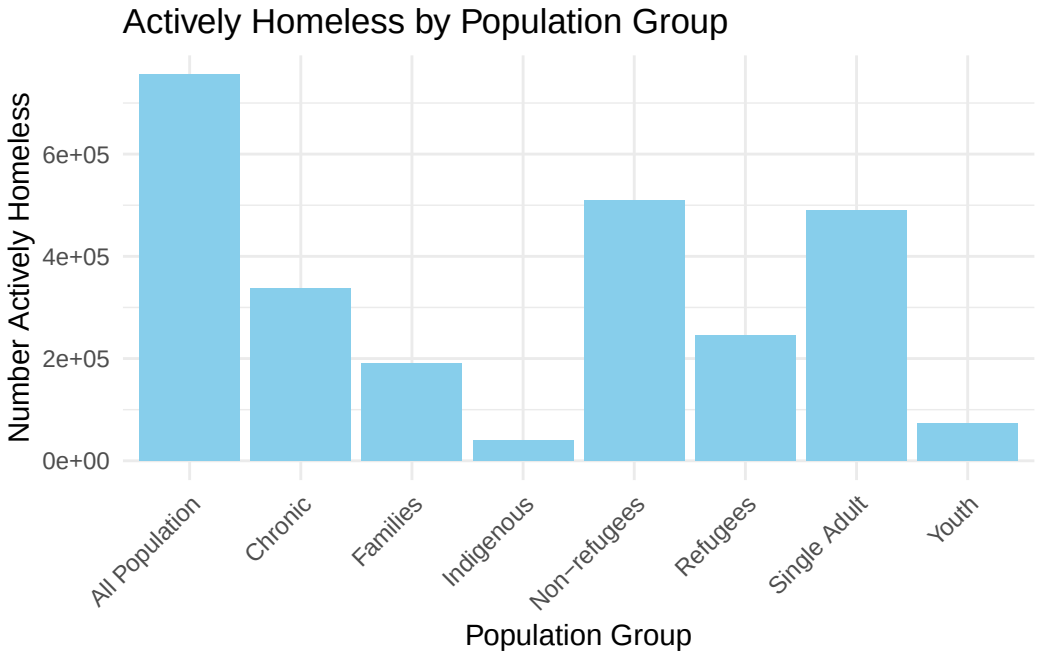

Figure 1: Actively homeless individuals across different population groups in Toronto 2018-2024

The bar chart Figure 1 illustrates the number of actively homeless individuals across different population groups in Toronto. The x-axis represents various population groups, while the y-axis shows the number of actively homeless individuals within each group. The chart clearly shows that homelessness is primarily concentrated in a few key groups: chronically homeless individuals, single adults, and non-refugees. Although vulnerable groups such as refugees, youth, and Indigenous people are also represented, their proportion is relatively small. The family group has the lowest number of actively homeless individuals.

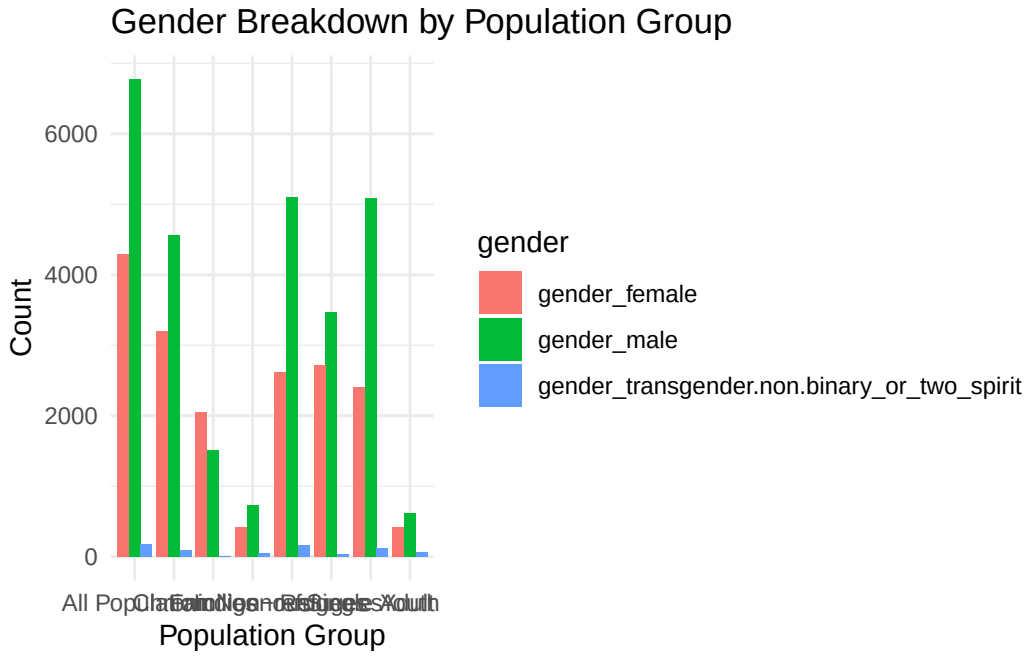

Figure 2: gender breakdown of actively homeless individuals across different population groups in Toronto 2018-2024

The clustered bar chart Figure 2 presents the gender breakdown of actively homeless individuals across various population groups in Toronto. The x-axis represents the population groups, while the y-axis shows the count of individuals. The colors in the bars indicate gender categories: dark blue for males, teal for females, and green for transgender/non-binary/two-spirit individuals. Males dominate the homeless population in most categories. Females are present in smaller numbers across all groups, but they represent a larger proportion in the Families. Transgender/non-binary/two-spirit individuals make up a very small portion of the homeless population across all groups.

3 Adding the labels to the first known cause data table

first_known_cause_data <- first_known_cause_data |> mutate(`Label` = first_known_cause_data_labels)
‘‘‘

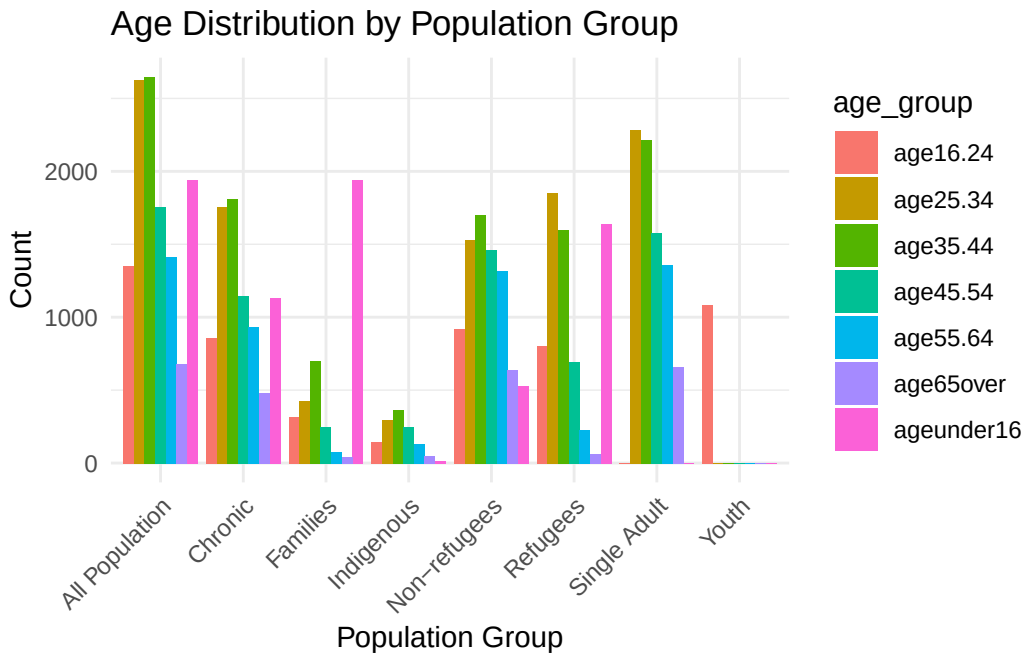

Figure 3: age-distribution of actively homepless individuals across different population groups in Toronto 2018-2024

The Dodged Bar Plot Figure 3 shows the age distribution across different population groups. The x-axis lists several population groups. The y-axis indicates the count or number of individuals within each age group.Families and Youth groups are predominantly represented by younger individuals, with the ageunder16 and age16.24 groups being the most prominent. Chronic and Single Adult groups exhibit a more balanced distribution across the middle age groups, with a noticeable presence in the age35.44 and age45.54 categories. Refugees and Indigenous groups display smaller counts overall but show a broader distribution across various age categories.

4 Discussion

The study had conducted an in-depth analysis of the composition of Toronto's homeless population, focusing on homelessness across different demographic groups. Through the analysis of various population categories—such as chronically homeless individuals, refugees, families, youth, and single adults—we found that chronically homeless individuals and single adults

make up a significant portion of the homeless population, highlighting the systemic challenges these groups face (Aubry, Klodawsky, and Coulombe 2012). Additionally, the gender and age distribution revealed disparities within the homeless population. For example, males dominate across all groups, while females and youth are more concentrated in the refugee and family categories. These findings underscore the complexity of homelessness and not only enhance our understanding of the homeless population but also provide critical data to support future policy interventions (Spetter 1996).

Our findings align with previous research (Nishio et al. 2017), confirming that chronically homeless individuals, due to a lack of stable social support and housing opportunities, often remain in long-term homelessness. This trend is particularly evident in the single adult group, who frequently lack family support and face more severe economic and mental health challenges. Furthermore, the higher proportion of females in the refugee and family groups may be due to the fact that these groups are more likely to consist of family units, which face unique challenges in adapting to new environments and accessing resources.

At the same time, youth homelessness highlights the vulnerability of this group during the transition into adulthood. Without adequate employment and educational opportunities, they are at higher risk of falling into a cycle of homelessness. The underreporting of gender minorities, such as transgender and non-binary individuals, may be related to systemic recording deficiencies, reflecting the "invisibility" (Norris and Quilty 2021) of these groups within the homeless population and suggesting the need for greater attention and data collection in the future.

While this study provides valuable insights into Toronto's homeless population, it also has several limitations. First, the study only includes individuals who used overnight services, excluding those who sleep outdoors or use other non-municipally funded services, potentially underestimating the total number of homeless individuals. Second, the proportion of gender minorities in the data is low, indicating that these groups may be systematically overlooked or underreported within the homeless population. Additionally, the time range of the study limits our ability to analyze long-term trends, preventing us from fully understanding how homelessness evolves over time and across different spatial contexts. Finally, the study relied solely on municipal data, possibly missing key subgroups such as migrant or informal populations.

A Appendix

A.1 Dataset and Graph Sketches

Sketches depicting both the desired dataset and the graphs generated in this analysis are available in the GitHub Repository.

A.2 Data Cleaning

The data cleaning process involved filtering out some of the columns from the raw dataset and renaming some of the data entries for clarity and simplicity.

A.3 Attribution Statement

"Contains information licensed under the Open Data - City of Toronto (Gelfand 2022).

References

Alexander, Rohan. 2023. *Telling Stories with Data*. Boca Raton: CRC Press. https://tellingstorieswithdata.com/.

Aubry, Tim, Fran Klodawsky, and Daniel Coulombe. 2012. "Comparing the Housing Trajectories of Different Classes Within a Diverse Homeless Population." *American Journal of Community Psychology* 49: 142–55.

Gaetz, Stephen. 2020. "Making the Prevention of Homelessness a Priority: The Role of Social Innovation." *American Journal of Economics and Sociology* 79 (2): 353–81.

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

Goodman, Lisa A, Leonard Saxe, and Mary Harvey. 1991. "Homelessness as Psychological Trauma: Broadening Perspectives." *American Psychologist* 46 (11): 1219.

Hwang, Stephen W. 2000. "Mortality Among Men Using Homeless Shelters in Toronto, Ontario." *Jama* 283 (16): 2152–57.

Nishio, Akihiro, Ryo Horita, Tadahiro Sado, Seiko Mizutani, Takahiro Watanabe, Ryosuke Uehara, and Mayumi Yamamoto. 2017. "Causes of Homelessness Prevalence: R Elationship Between Homelessness and Disability." *Psychiatry and Clinical Neurosciences* 71 (3): 180–88.

Norris, Michelle, and Aideen Quilty. 2021. "Unreal, Unsheltered, Unseen, Unrecorded: The Multiple Invisibilities of LGBTQI+ Homeless Youth." *Critical Social Policy* 41 (3): 468–90.

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

Spetter, Victoria C. 1996. "As Government Assistance Decreases, Homelessness Increases: A Closer Look at Welfare, Housing and Homelessness." In *Hybrid*, 3:111. HeinOnline.

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.