Navigating Ontario's Shelter and Overnight Services: A Analysis of the Change in Distribution, Program Types, Composition of Users and Occupancy Situation between 2021 and 2024*

Yiyi Feng

September 27, 2024

In this study, we investigated how the number of reported COVID-19 cases in Toronto between 2020 and 2024 relate to the homeless population across different age groups in the city's shelter system. While it's known that pandemics often coincide with increased homelessness, their specific impact on different age groups among the homeless remains less explored. Using linear models, we found a negative correlation between the overall homeless population and reported COVID-19 cases, including a more positive correlation with elder homeless population, and a more negative correlation with younger homelss population. Our models predict a decrease in the population of homeless individuals aged 45 and above as COVID-19 cases decrease, while those below 45 are expected to increase. This suggests a potential rise in the percentage of younger individuals within the homeless community in the future.

Table of contents

*Code and data are available at: https://github.com/kqlqkqlqF/Analyzation-of-shelter-overnight-occupancy.git.

1 Introduction [1]

In recent years, due to the impact of COVID-19 and the influx of refugees, the number of homeless population in Toronto has continued to rise. These individuals are facing increasingly difficult challenges when seeking overnight shelter. Even for those who living in downtown Toronto, where shelters are most concentrated, encountering homeless individuals on the streets has become a common sight—shelters simply cannot accommodate more people. Official reports confirm this reality. As early as 2020, the General Manager of the city's Shelter, Support, and Housing Administration publicly stated that an average of 72 people were turned away daily due to overcrowding, and this number has been steadily rising (Gibson 2023). By the winter of 2022, the number of people denied shelter had surged to 168 per day (National 2022).

This situation sparked our curiosity. We wanted to understand how the shelter system operates. For example, for those shelters outside of downtown Toronto, where are they located? Has the overall distribution of shelters changed significantly in recent years? What categories do shelters and overnight services fall into, and what are the proportions of each? Are Toronto's shelters truly as overcrowded as described earlier? If so, have shelters added more beds to cope with the increasing demand? To address these questions, we turned to the "About Daily Shelter & Overnight Service Occupancy & Capacity" dataset available on OpendataToronto, hoping that by analyzing this dataset, we could find some answers.

Using this dataset, we first analyzed the distribution of shelters and overnight service programs in Toronto over the past few years, comparing how their locations have changed each year.

[1] Please check https://github.com/kqlqkqlqF/Analyzation-of-shelter-overnight-occupancy.git for more information.

Next, we examined the different Program Models, Service Types, and Program Areas of these shelters and services to better understand their classifications and proportions. We also looked into the composition of the individuals using these shelters and services to explore changes in the makeup of Toronto's homeless population. Lastly, we investigated shifts in the total number of available beds and rooms, as well as their occupancy rates.

From these analyses, we drew several conclusions. First, between 2021 and 2024, aside from a few locations being removed or relocated, the distribution of shelters and overnight service programs in Toronto remained largely unchanged. The majority of these shelters are classified as emergency programs, which do not require a referral for access. We also observed a clear response from the shelter system to meet current demands, such as the COVID-19 response program from 2021 to 2023 and the refugee program introduced in 2023 and 2024. Among those using these services, mixed adults make up the largest group, while youth represent the smallest. However, the overall number of users has been on the rise. Lastly, although the number of available beds has increased in recent years, shelters are indeed very crowded, with the number of available beds and rooms closely matching the number of occupied ones, confirming the earlier observations.

To summarize, in this study, we analyzed the change in distribution, program types, composition of users and occupancy situation of Toronto shelter and overnight service program system between 2021 and 2024 to better understand their structure and proportions. It is important and valuable to understand the dynamics of shelter and overnight service program in Toronto. It highlights the challenges faced by Toronto's social welfare system over different periods and the broader societal responses to these events, such as the proactive measures taken in response to COVID-19 and the influx of refugees.

2 Data

2.1 Data Source

The data for this analysis was collected from the shared Toronto Open Database (Shelter-Support-&-Housing-Administration 2018). This study utilizes and analyzes the dataset titled "About Daily Shelter & Overnight Service Occupancy & Capacity". This data set provides a daily list of active overnight shelter and allied services in the Toronto Shelter and Support Services division's Shelter Management Information System (SMIS) database. It provides information on shelter occupancy and service provision in Toronto from January 2021 to August 2024. It includes details such as the unique identifiers for organizations and shelters, the dates of occupancy records, and the specific types of overnight services offered. The dataset categorizes shelter programs based on their operational models, including emergency and transitional shelters, and identifies the sectors they serve, such as adult men, women, mixed adults, youth, and families. Furthermore, it tracks the number of service users, bed and room capacities, and occupancy rates, enabling analysis of trends in shelter usage and

responses to evolving needs, including those arising from the COVID-19 pandemic and the influx of refugees.

The dataset, published by Toronto Shelter & Support Services, has been updated daily since January 1, 2021. As of the submission of this paper, the dataset includes data up to September 26, 2024. This dataset has received full marks for Freshness, Metadata, Accessibility, Completeness, and Usability on the Opendatatoronto website, leading us to consider it highly credible and use it as the primary data source for this study. Since the dataset does not yet cover all of September 2024, and in our analysis of "Composition of Users" and "Occupancy and Capacity," we divided the data into four-month intervals, the data included only extends up to August 31, 2024. However, for the "Distribution" and "Program Types" analyses, where we compare and segment data on a yearly basis, the updated data for September 2024 has been included.

2.2 Features

This dataset records the program model, overnight service type, and program area categories for each shelter, which include specific classifications relevant to the shelter system. First, the program model is divided into two types: Emergency, which can be accessed by any individual, and Transitional, which is only available to individuals with a referral. Next, the overnight service type is categorized based on the services provided, including Shelter, 24-Hour Respite, Motel/Hotel, Interim Housing, Warming Centre, 24-Hour Women's Drop-in, and Isolation/Recovery Site. It is important to note that the Isolation/Recovery Site was established specifically for homeless individuals infected with COVID-19 who require urgent isolation and medical services. Additionally, the 24-Hour Women's Drop-in not only accommodates homeless women but also supports transgender or gender-non-binary individuals. Lastly, the Program Area distinguishes whether a program belongs to a shelter, overnight service program, or temporary response program, with categories such as Base Shelter and Overnight Services System, Temporary Refugee Response, COVID-19 Response, and Winter Response.

The dataset also introduces a classification of Bed-Based or Room-Based Capacity Types. Bed-based capacity typically applies to programs with shared sleeping areas, while room-based capacity is more common for family programs and hotel-based programs, where bedrooms are not shared by individuals from different households. This classification is designed to prevent over-reporting of capacity in room-based programs. For each capacity type, the dataset records five key figures: Capacity Actual, Capacity Funding, Occupied, Unoccupied, and Occupancy Rate. Capacity Actual refers to the number of spaces available as shown in the management system, while Capacity Funding represents the number of spaces that have been approved for funding in the program. Occupied indicates the number of spaces in use on a given day. Unoccupied is the difference between the capacity actual and the number of occupied spaces. Occupancy Rate is the result of dividing the number of occupied spaces by the capacity actual.

2.3 Methodology

The data analysis was conducted using R (R Core Team 2022), a versatile statistical programming language. We utilized a range of packages to enhance our analysis. The tidyverse (Wickham et al. 2019) suite of packages provided a comprehensive toolkit for efficient data manipulation and visualization. Package ggplot2 (Kassambara 2023) allowed us to create compelling visualization. The here (Müller 2020) package simplified file management within our project directory structure. We also used ggmap (Kahle and Wickham 2013) in combination with the Google API to generate a map showing the distribution of shelters in Toronto. Additionally, kableExtra (Zhu 2021) was employed to generate visually appealing and customizable tables, enhancing the presentation of our findings. For Bayesian analysis, we utilized the rstanarm (Goodrich et al. 2020) package, which provided an elegant interface to Stan, a cutting-edge platform for statistical modeling and computation. This allowed us to estimate relationships within our data using a Bayesian framework, providing valuable insights into our research questions. Report generation was seamlessly managed using knitr (Xie 2023), enabling the integration of R code within our document. Other essential packages included tibble (Müller and Wickham 2022), stringr (Wickham 2020), lubridate (Grolemund and Wickham 2020), janitor (Firke 2023), and testthat (Wickham and RStudio 2020), each contributing to various aspects of our data analysis process, from data manipulation to quality assurance.

Due to the extensive variety of data classifications within this dataset, with a total of 32 features, the data cleaning process was quite challenging. To maintain an organized structure in the final QMD file that generates the PDF, we first cleaned the overall dataset and saved it as a CSV file. From this cleaned overall dataset, we then extracted and processed the necessary data for each figure separately, storing the results in individual CSV files.

3 Result

3.1 Distribution of Active Shelters and Overnight Service Programs in Toronto

The data provided in **?@tbl-one** and Figure 1 offers insight into the dynamic changes in the homeless population within the Toronto shelter system from 2018 to 2023.

A notable trend observed is the fluctuation in the annual count of newly identified individuals entering the shelter system. Peaking in 2018 and gradually declining over the following years, the trend suggests a persistent need for shelter services, with a gradual increase observed from 2020 onwards.

Another significant aspect of the data is the return of individuals from permanent housing to the shelter system. This phenomenon is concerning as it indicates potential housing instability for the population. Fortunately, the numbers have remained relatively stable from 2018 to 2023, and they are lower than those transitioning to permanent housing each year. However, the declining trend in individuals moving to stable housing from 2018 to 2021 is worrisome.

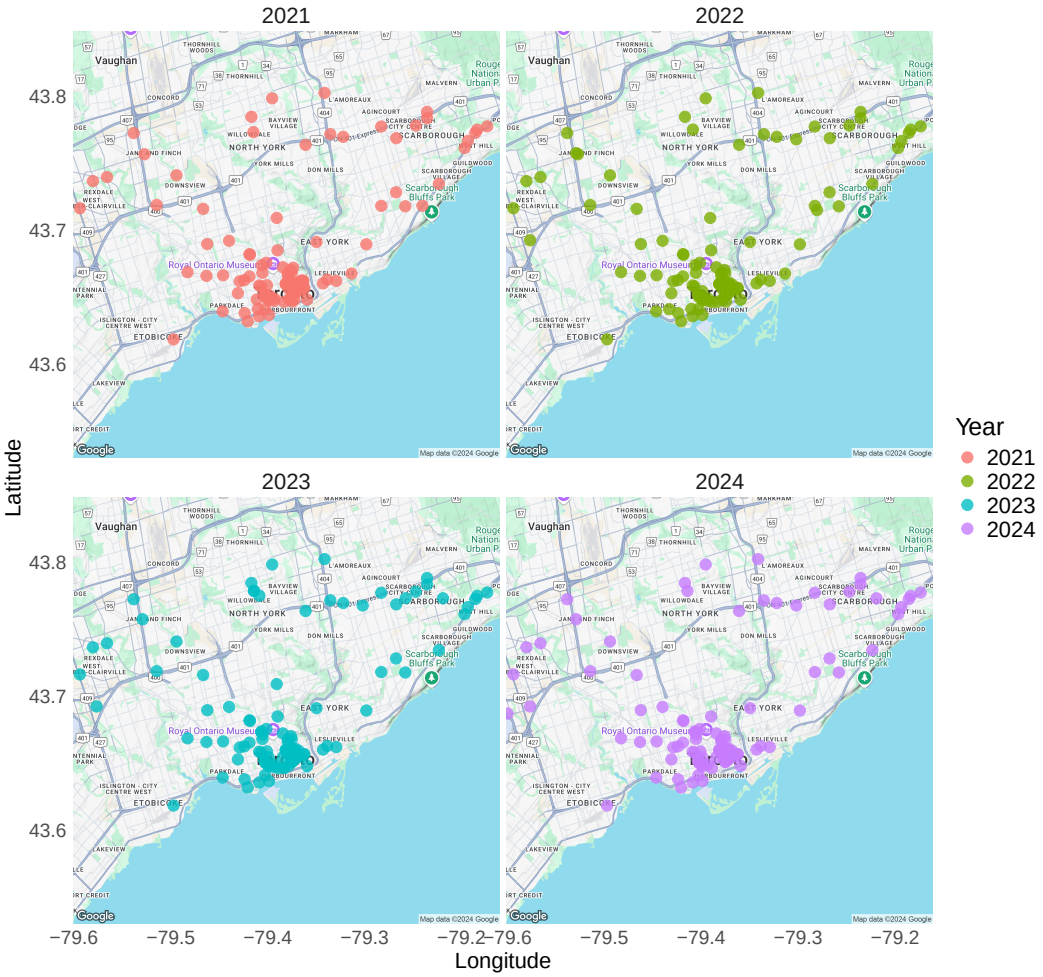

Figure 1: Distribution of active shelters and overnight service programs in Toronto from 2021 to 2024

Although there is a slight increase in 2022 and 2023, it remains below the 2018 levels, suggesting a decrease in individuals transitioning to permanent housing.

The data also captures changes in individuals returning to shelters and those becoming inactive within the shelter system. While there is a gradual decline in the number of individuals becoming inactive homeless over the five years, the decrease is minimal, similar to the trend in individuals returning to shelters. Without further explanation, it's challenging to draw reliable conclusions on the decrease in the number of inactive homeless individuals or the reasons behind individuals returning to shelters since their experience after leaving the shelter system is unknown. They could be moving to other cities, finding housing, or dying. Therefore, if there's no further detail for the reason they left the shelter system, it is hard to understand the implications of the data fully, and no conclusion can be drawn.

Lastly, individuals from the actively homeless category highlight the ongoing demand for shelter services and support. Although there was a decline in the number of actively homeless individuals in 2020 and 2021, it surged to new highs in 2022 and 2023. This suggests a significant increase in the number of individuals utilizing overnight shelter services in the past two years, indirectly indicating a potential increase in Toronto's homeless population from 2022 to 2023.

In conclusion, these data paint a complex picture of homelessness within the Toronto shelter system, with a trend of increasing actively homeless individuals.

3.2 Analysis of Program Models, Service Types and Program Areas for Active Shelters and Overnight Services

Before beginning the analysis of Figure 2, several points need to be explained. Firstly, within the population groups depicted in this graph, apart from refugees and non-refugees, the other options are not mutually exclusive. This means that a person can simultaneously belong to the chronic, families, and refugee groups, but cannot be a refugee and a non-refugee at the same time. This explains why the sum of individuals in these population groups exceeds the number in all population categories. Secondly, as mentioned earlier, the number of individuals in the indigenous group was only recorded starting in 2020, so the actual number of indigenous individuals who have used the shelter system should be higher than what is shown in the graph.

Individuals aged between 25 and 44 constitute the largest portion of the homeless population, while those aged between 45 and 64 represent the second largest group, and those aged 65 and above represent the smallest group. From the perspective of population groups, non-refugees constitute the largest group, followed by refugees and then chronic individuals, with indigenous individuals being the smallest group. From this distribution trend, we can infer the following information:

Figure 2: Analysis of Program Models, Service Types and Program Areas for Active Shelters and Overnight Services in Toronto from 2021 to 2024

Firstly, except for the age group under 16, chronic individuals constitute a significant portion of the homeless population in other age groups, indicating that a considerable number of homeless individuals rely on overnight shelter services for an extended period and find it difficult to secure permanent housing, thus perpetuating their homelessness.

Secondly, although refugees constitute a smaller proportion of the homeless population compared to non-refugees overall, in the age group under 16, the number of refugees far exceeds the proportion of non-refugees. This suggests that there are a significant number of parenting teenagers or even younger children among the refugee population who are homeless. Additionally, we observe that the proportion of families in the age group under 16 also significantly increases, slightly surpassing the proportion of refugees and almost equaling the "all population" category. This evidence confirms our previous conclusions and supports the notion that these young refugee homeless individuals are likely to be wandering alongside their refugee family members.

Thirdly, unlike refugees, homeless individuals categorized as non-refugees are mainly distributed in the age group of 25 and above. Particularly in the age group of 65 and above, the number of non-refugees is almost equivalent to the "all population" category.

In summary, chronic individuals are prominent across age groups, highlighting challenges in securing permanent housing. Refugees tend to enter the shelter system as a whole family with their young children, indicating a concerning trend of homelessness among refugee families with children, while the non-refugees dominate the older age groups.

3.3 Composition of Users for Active Shelters and Overnight Services

The data presented in Figure 3 provides insights into the gender dynamics within Toronto's homeless population across different population groups from 2018 to 2023. A consistent trend observed throughout the years is the higher representation of males compared to females and individuals identifying as transgender, non-binary, or two-spirit. This imbalance suggests a prevalent gender disparity among those experiencing homelessness, with males comprising a larger proportion of the homeless population.

However, within specific population groups, variations in gender composition emerge. For instance, in the chronic population group, which likely encompasses individuals experiencing long-term homelessness, both males and females are prominently represented, indicating a diverse demographic within this subgroup. In contrast, the families population group, which likely consists of homeless families with children, demonstrates a higher count of females compared to males and individuals identifying as transgender, non-binary, or two-spirit. This gender distribution within homeless families may be influenced by factors such as caregiving responsibilities and access to support services, indicating the unique dynamics within this subgroup. The "Refugee" population group exhibits a relatively balanced distribution between males and females, with a smaller count of individuals identifying as transgender, non-binary, or two-spirit. This suggests a diverse composition within refugee populations experiencing

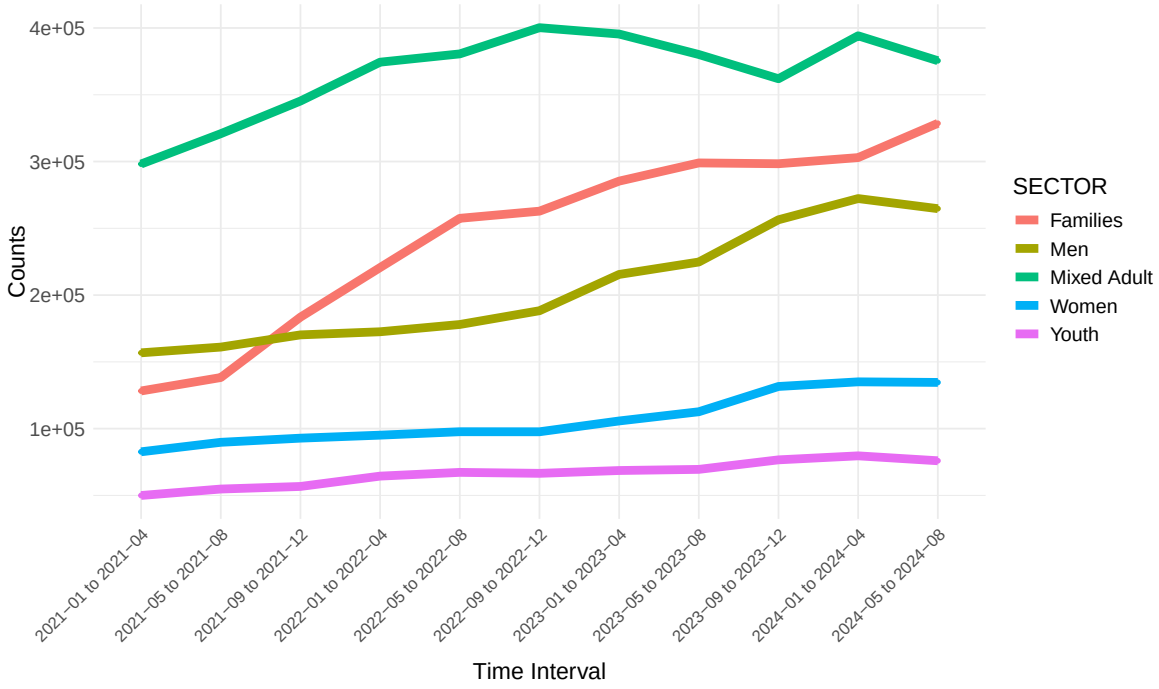

Figure 3: Changes in Composition of Users for Active Shelters and Overnight Services in Toronto between January 2021 to August 2024

homelessness, possibly influenced by factors such as migration patterns and resettlement challenges.

Furthermore, the data reveal temporal fluctuations in the chronic population group, which demonstrates a consistent increase in count from 2018 to 2023. This upward trend suggests a growing number of individuals experiencing long-term homelessness.

3.4 Analysis of Occupancy and Capacity for Active Shelters and Overnight Services

Trends of the Number of COVID-19 Cases demonstrated in Figure 4 reveal several peaks in the surge of COVID-19 infections, including from September 2020 to January 2021, February to June 2021, and November 2021 to February 2022. Although there were a few minor increases in infection numbers thereafter, overall, the number of infections has been consistently decreasing from February 2022 to February 2024.

4 Discussion

4.1 Summery of Findings

The analysis of homeless population dynamics in Toronto from 2018 to 2023 revealed fluctuating trends in shelter entry, with a peak in 2018 followed by a gradual decline and a slight increase from 2020 onwards. The data also demonstrated potential housing instability through individuals returning from permanent housing to shelters, although overall stability was observed. Chronic individuals were prevalent across age groups, indicating persistent challenges in securing permanent housing. Refugee families, often with young children, contributed to homelessness trends, while non-refugees dominated older age groups. Gender disparities were evident, with males comprising a larger proportion of the homeless population, especially among chronic individuals, while females were more represented in homeless families. These findings indicate an alarming rise in homelessness, shifts in demographic composition, and a gender imbalance, all demonstrated the multifaceted nature of homelessness.

From the generated models, we found that the trend in all homeless population is negatively correlated with the number of COVID-19 cases, while the age group above 45 exhibits a positive correlation and the age group below 45 shows a negative correlation. Moreover, overall, older age groups of homeless individuals tend to show stronger positive correlations, while younger age groups of homeless individuals tend to exhibit stronger negative correlations. These results were contradicting with our original expectations. Following this pattern, as the spread of the pandemic is further controlled, the occurrence of COVID-19 cases is expected to decrease, while the total homeless population is anticipated to increase. Additionally, the proportion of older individuals within the homeless population will gradually decrease, while the proportion of younger individuals will gradually increase.

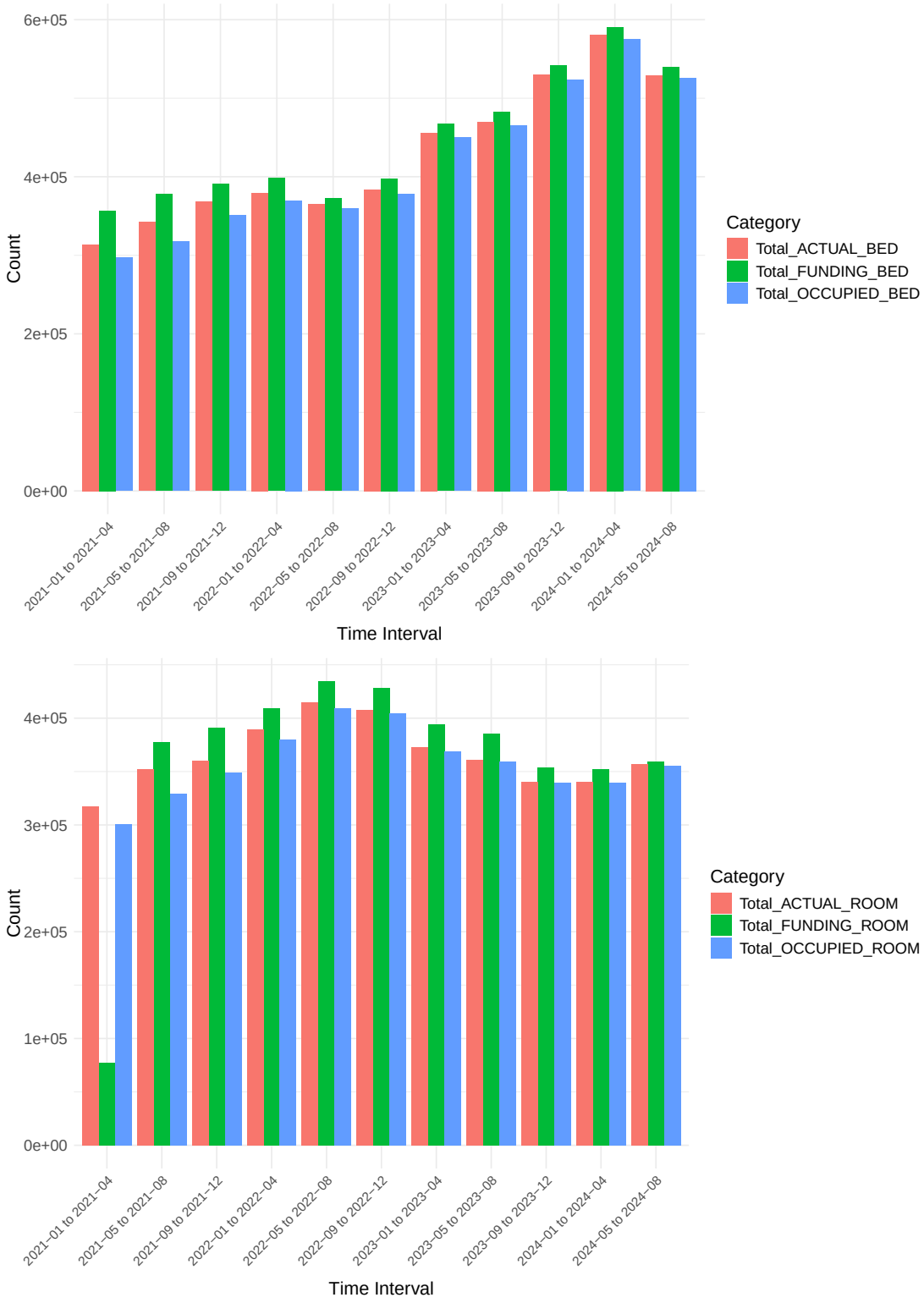

Figure 4: Trends of Occupancy and Capacity for Active Shelters and Overnight Services in Toronto from January 2021 to August 2024

4.2 Things learned from this study

Firstly, we found that the process of data collection, analysis, and research to reach conclusions is much more complex than we initially anticipated. In the Toronto shelter system Flow dataset we utilized, data is recorded in detail, including information on whether homeless individuals leave shelters to move into permanent housing. However, despite the comprehensive nature of the data, our analysis still falls short. As we raised in the results section, we have no way of knowing where homeless individuals go or what they do after becoming inactive, making it difficult to determine whether the increase or decrease in inactive status is beneficial or detrimental. We believe that in the future, there could be further sharing of data between shelter systems, such as sharing homeless individual information across cities or provinces, which would allow us to track their movements after leaving shelters in another jurisdiction.

Secondly, the proportion of child/youth homeless individuals is much higher than we anticipated, which came as a surprise. From the data, it appears that many of these young homeless individuals are refugees who are traveling with family members, with only a small fraction being residents. Additionally, around one-third of homeless children are classified as chronic, meaning they have been homeless for an extended period. These data indirectly highlight the difficulties refugees face in finding employment and obtaining legal status in a new country, which in turn affects the next generation. We hope that the government can develop relevant welfare policies for these refugee children to reduce the number of homeless children.

Thirdly, conducting overall homeless population counts poses significant challenges. Firstly, many homeless individuals do not stay overnight in shelters and spend both day and night on the streets or in parks. This means that we do not have an efficient and accurate way of collecting identity information for these homeless individuals. Secondly, as mentioned earlier regarding refugee issues, many homeless individuals lack legal status or identification documents, making it difficult for staff to determine the reliability of the information they provide. Additionally, the transient nature of homelessness complicates both the accurate collection of homeless individuals' numbers and their identity information. Some individuals may be homeless today, find temporary accommodation the next day, and then suddenly lose their accommodation and resume homelessness after a few days. This instability makes it challenging to accurately collect both the number and identity information of homeless individuals.

4.3 Limitations

Although the analysis provides insights into the dynamics of homelessness in Toronto, there are still limitations to consider. Firstly, these data primarily reflect individuals utilizing shelter services and may not capture the entire homeless population, including those living on the streets or in unstable housing conditions. Secondly, we need more detailed data for making more precise conclusions. For example, we cannot track the whereabouts of all homeless individuals who leave the shelter system; the data only record numbers without monitoring their activities. The reasons for individuals becoming inactive homeless may vary significantly,

such as death, disappearance, or relocation, leading to confusion and ambiguity in predicting future dynamics. Thirdly, our dataset is based on the trends of homelessness from 2018 to 2024, which were affected by the multifaceted impacts of the COVID-19 pandemic, including but not limited to: increased difficulty in data collection due to social distancing measures, the health, economic, and social impacts of COVID-19 on the homeless population, and the potential closure of shelters due to the spread of the disease. These factors may reduce the effectiveness and accuracy of data collection. Moreover, we can observe a significant decrease in shelter utilization in 2020 and 2021 during the pandemic, which contradicts our initial assumptions. Therefore, we have some reservations about the reliability of the data during this period. Furthermore, there exist significant differences in sample sizes among different age groups. As the number of homeless individuals aged 65 and above is inherently low, the credibility of model predictions for this group is naturally lower compared to the 25-44 age group. However, this issue is particularly concerning because the scarcity of homeless individuals aged 65 and above is a fact, and we cannot collect a similar volume of data as we can for the 25-44 age group due to the limited presence of elderly homeless individuals.

In terms of model analysis, as mentioned earlier, the COVID-19 pandemic is not the sole factor influencing the homeless population, and their correlation is unlikely to be entirely linear. The number of homeless individuals is closely tied to various factors such as economic conditions, housing prices, welfare policies, and more. These factors impose significant limitations on our models. Additionally, since the outbreak and containment of the pandemic occurred within approximately three years, our dataset and data reliability are relatively limited. This project itself is not conducive to long-term monitoring (e.g., over five to ten years), as determined by the specific form of viral epidemic. Moreover, due to the sharp increase in infection numbers during concentrated outbreaks of the virus, there is significant variability in monthly infection counts. This results in a plethora of low-infection count data points and scarce data during periods of high infection counts, affecting the overall accuracy of the model predictions.

4.4 Future Study

As the homeless population continues to grow steadily and faces increasingly dire challenges in obtaining permanent housing, we urge the government to intervene urgently to expand housing capacity, reduce housing costs, and strengthen support services. Future research should explore the fundamental causes of homelessness, identify effective interventions to prevent homelessness from recurring and assess the long-term impacts of housing policies and support programs. By understanding the root causes of homelessness and implementing targeted interventions, people can be protected from homelessness, thereby reducing its occurrence at a more fundamental level. As time passes, the impact of the COVID-19 pandemic on homelessness dynamics is expected to diminish. If similar studies were to be conducted in the future, We believe it would be beneficial to first incorporate more factors, such as the average rental prices and cost of living in Toronto. Additionally, using non-linear models for prediction may yield

better results. However, regardless of the approach, generating models for the homeless population poses significant challenges. Furthermore, it is essential to incorporate a wider range of comprehensive factors into our research when collecting data, which may include reasons for homelessness and dynamic trajectories, to provide deeper insights into which factors have a more profound impact on homelessness. Moreover, the government needs to establish a more comprehensive data collection and measurement system, including tracking the dynamics of inactive homeless individuals and conducting categorized surveys of homeless families. This will enable a more accurate assessment of the challenges faced by the homeless population and facilitate the development of targeted and effective measures to address these issues.

Bibliography

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://CRAN.R-project.org/package=janitor.

Gibson, Victoria. 2023. "Why the Homelessness Crisis Could Get Even Worse." 2023. https://www.thestar.com/news/gta/why-the-homelessness-crisis-could-get-even-worse/article_9054cdcc-f846-5a33-a9e4-f972860686a1.html.

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2020. *Rstanarm: Bayesian Applied Regression Modeling via Stan*. https://mc-stan.org/rstanarm.

Grolemund, Garrett, and Hadley Wickham. 2020. *Lubridate: Make Dealing with Dates a Little Easier*. https://CRAN.R-project.org/package=lubridate.

Kahle, David, and Hadley Wickham. 2013. *Ggmap: Spatial Visualization with Ggplot2*. https://cran.r-project.org/package=ggmap.

Kassambara, Alboukadel. 2023. *Ggpubr: 'Ggplot2' Based Publication Ready Plots*. https://rpkgs.datanovia.com/ggpubr/.

Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. https://CRAN.R-project.org/package=here.

Müller, Kirill, and Hadley Wickham. 2022. *Tibble: Simple Data Frames*. https://CRAN.R-project.org/package=tibble.

National, CBC News. 2022. "Toronto Shelters Turning Away Average of 168 People a Day." 2022. https://www.youtube.com/watch?v=vUZhZIS-XRo&t=2s.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Shelter-Support-&-Housing-Administration. 2018. *About Toronto Shelter System Flow*. https://open.toronto.ca/dataset/toronto-shelter-system-flow/.

Wickham, Hadley. 2020. *Stringr: Simple, Consistent Wrappers for Common String Operations*. https://CRAN.R-project.org/package=stringr.

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, and RStudio. 2020. *Testthat: Unit Testing for r*. https://CRAN.R-project.org/package=testthat.

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

Zhu, Hao. 2021. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=kableExtra.