An Analysis of daily shelter and overnight service usage in Toronto focusing on shelter occupancy and capacity*

Angel Xu

September 27, 2024

This paper analyzes the overnight service occupancy and capacity of shelters in Toronto, utilizing data from OpenDataToronto. Employing the R programming language along with various packages, the analysis focuses on shelter usage trends, specifically examining changes in occupancy rates and actual capacity over the years. The results indicate an overall upward trend in shelter occupancy rates, while a significant decline in actual capacity was observed in 2024. These findings highlight the need for strategic resource allocation and improved response strategies concerning shelter occupancy in the future.

Table of contents

*Code and data are available at: https://github.com/Anjojoo/Toronto-Overnight-Servive-Occupancy-Capacity.

1 Introduction

The growing issue of homelessness in urban centers like Toronto has created an urgent need to analyze shelter occupancy and capacity, as shelters are critical for providing temporary housing to those in need. To effectively manage and improve these services, understanding usage trends is essential. This paper focuses on analyzing the "Daily Shelter & Overnight Service Occupancy & Capacity" dataset from OpenDataToronto(Gelfand (2022)), covering data from 2021 to 2024. This dataset includes daily records of shelter occupancy, capacity, and types of overnight services provided, offering a comprehensive view of the situation in Toronto. The motivation behind this study is to address the fluctuating demand for shelter services and inform strategies for more efficient allocation of resources to support vulnerable populations.

The study involves a detailed analysis of shelter usage trends, employing R programming(R Core Team (2023)) and Alexander (2023) mainly. Besides, associated packages of R are applied to clean, manipulate, and visualize the data. Tools like tidyverse(Wickham et al. (2019)), ggplot2(Wickham (2016)), dplyr(Wickham et al. (2023)), and others facilitated data processing and visualization, enabling a deeper examination of occupancy rates and capacity changes over the years. The findings reveal an overall upward trend in shelter occupancy rates, suggesting increased demand for these services. A significant decline in actual capacity was observed in 2024, which underscores a mismatch between available resources and the needs of shelter users. Summary statistics were computed to highlight trends in service user counts and occupancy rates, providing insights into how shelter usage has evolved.

This paper's findings can be helpful for policy makers, social service organizations, and urban planners. By shedding light on the trends in shelter occupancy and capacity, the study identifies gaps in service and potential areas for improvement. The analysis has implications for future resource allocation, as understanding the patterns of shelter usage can help design better support strategies, optimize capacity, and improve the overall response to homelessness in urban areas. Additionally, it can serve as a model for analyzing shelter systems in other cities, offering a framework for global cases.

The remainder of this paper is structured as follows. Section 2 provides details on the data source, including the specific variables used, and describes the data cleaning process to prepare for analysis. Data tools and methodologies applied in this paper are also stated in this section. Measurement and ethics are considered as well. Section 3 presents the key results, focusing on occupancy rates, service user counts, and actual capacity over time; Visualizations are used to intuitively convey these findings. Section 4 offers a discussion of the results and their implications,also emphasizing the importance of balancing shelter occupancy and capacity for effective service provision while limitations point out some weakness in this paper. The comprehensive structure ensures that the reader gains a full understanding of what was analyzed, the findings, and their significance.

2 Data

2.1 Data Source and Overview

This paper analyzes data from the "Daily Shelter & Overnight Service Occupancy & Capacity" dataset available on the OpenDataToronto portal. The dataset contains a daily record of overnight shelters and service programs collected through Toronto's Shelter Management Information System (SMIS). Data spans from 2021 to 2024, offering current insights, and is updated daily, except for weekends and holidays, with uploads resuming on the next business day.

A similar dataset, "Daily Shelter Occupancy," is also available on OpenDataToronto but only provides data up to March 10, 2022, and includes fewer variables. In contrast, "Daily Shelter & Overnight Service Occupancy & Capacity" features the most recent data, refreshed on 25 September, 2024. It builds upon and enhances the earlier dataset by incorporating additional key variables, including "overnight service type" and "actual capacity." The "overnight service type" variable categorizes the types of overnight services offered in Toronto's shelters, helping analyze variations in user counts by service type. The "actual capacity" variable represents the true available capacity of shelters, as opposed to the funding-based capacity in the previous dataset, offering a more accurate measure of occupancy rates for further analysis.

2.2 Data Tools

The data was simulated, extracted, and cleaned by R(R Core Team 2023) from the opendatatoronto package(Gelfand 2022). Tidyverse(Wickham et al. 2019), here(Müller 2020), ggplot2(Wickham 2016), dplyr(Wickham et al. 2023), lubridate(Grolemund and Wickham 2011), janitor(Firke 2023) and knitr(Xie 2023) are all applied in the analysis. Styler(Müller and Walthert 2024) is used to correct code styles.

2.3 Data Measurement and Ethics

For the data measurement of "Daily Shelter & Overnight Service Occupancy & Capacity" from opendatatoronto(Gelfand (2022)), variables are mainly supported by Shelter Management Information System(SMIS). The occupancy data is recorded at 4:00 a.m. each morning, reflecting the previous evening's overnight service. Service user counts measures the number of service users at each program referring to the occupancy date. Actual capacity of beds and rooms are the data available on SMIS and used to calculate the occupancy rate with another available variable, occupied capacity which is also available on SMIS.

The ethical consideration included ensuring that the data used reflects accurate and timely reporting from the open-source platform and respecting the privacy of the individuals indirectly represented in the data. The data are collected anonymized and does not contain personally identifiable information, ensuring privacy and confidentiality of shelter users.

2.4 Clean Data

Prior to analysis, data cleaning is essential to remove unnecessary columns and create new ones for further examination. The occupancy date variable is converted to a date format, allowing the extraction of the year for subsequent analysis. After this conversion, a new 'year' column is mutated to show trends in service user counts and occupancy rates from 2021 to 2024. Once these steps are complete, relevant variables are selected from the raw dataset to simplify and clarify the analysis. The chosen variables include year, occupancy date, overnight service type, service user count, capacity type, actual capacity and occupancy rates. NAs exist in actual capacity and occupancy rates, but there are no needs to eliminate them since due to different capacity type of bed and room, NA appears in the opposite actual capacity and occupancy rates.

2.5 Summary Statistics of Variables

To gain a better overview of the dataset, a table including key variables is provided. Table 1 presents year, overnight service type, service user count and occupancy rates. Year represents the year where the observation is collected. Overnight service type represents what service has been provided. Service user count records number of users taking service during that night. Occupancy rates record the proportion of actual capacity that is occupied.

Table 1: Table 1: Sample of the key variables in Occupancy & Capacity Data

year	overnight service type	service user count	occupancy rate beds	occupancy rate rooms
2021	Motel/Hotel Shelter	74	NA	89.66
2021	Motel/Hotel Shelter	3	NA	100.00

Table 1: Table 1: Sample of the key variables in Occupancy & Capacity Data

year	overnight service type	service user count	occupancy rate beds	occupancy rate rooms
2021	Motel/Hotel Shelter	24	NA	82.14
2021	Motel/Hotel Shelter	25	NA	100.00
2021	Motel/Hotel Shelter	13	NA	92.86
2021	Shelter	6	75	NA

Beside the sample table of ley variables, a summary statistics table of key variables can be provided to clarify some information of the data. Table 2 presents the mean and standard deviation of service user counts and two occupancy rates. Service user counts has a mean at 60.58 and a standard deviation at about 76.09% which is pretty high and shows the great difference in user picks between different services. Occupancy rate of beds has a high mean at 96.24% and a standard deviation at 9.28%. This illustrates that the occupency rate of beds are approximately near 100 and are full occupied for most of the situations. Occupancy rate of rooms also has a high mean at 97.03% which is slightly higher than occupancy rate of beds and a standard deviation of 10.82%, suggesting that capacity of shelter rooms reaches nearly full and fuller than shelter rooms. However, the actual capacity of beds and rooms aren't included in this table; thus, the usage of shelter beds and rooms need to be considered more thoroughly by observing their actual capacity numbers.

Table 2: Table 2: Mean and Standard Deviation of Key Variables

Variable	Mean	Standard Deviation
Service User Count	60.58689	76.105111
Occupancy Rate of Beds	96.23710	9.276826
Occupancy Rate of Rooms	97.03618	10.812602

For deeper explore in service user count and capacity in different years, the sum of these variables from 2021 to 2024 needs to be examined. Occupancy rate of beds and rooms cannot reflect the total difference usage of all shelters, as well as the mean of service user counts. They can only represent the common usage situation of every shelter in total. While Table 3 provides a table of total service user counts and actual capacity for both beds and rooms. Within these data, a more complete result can be figured out and will be used for data visualization later.

Table 3: Table 3: Total Service User Count and Actual Capacity between years

Year	Total Service User Counts	Total Actual Bed Capacity	Total Actual Room Capacity
2021	2329502	1024657	1030161
2022	2924216	1128856	1211399
2023	3288505	1457984	1078307
2024	2633817	1232028	777790

3 Result

Data visualizations can be used to present results more intuitively, and the data used to visualize is from opendatatoronto@citeopendatatoronto.

3.1 Overnight Service Usage

The graph below displays the user counts for various overnight service types, offering a clear comparison. This visualization helps identify the most frequently used service across all types.

Figure 1 shows the distribution of overnight service usage. It is evident that Shelter and Motel/Hotel Shelter services have the highest user counts, with values nearly identical. The 24-Hour Respite Site ranks third, while the remaining services have significantly fewer users compared to these top three. This highlights that Shelter and Motel/Hotel Shelter are the most commonly chosen services in Toronto, likely due to the more comprehensive support they offer compared to other service types.

3.2 Service Count Change

Figure 2 shows the number of service users across different years from 2021 to 2024, as also detailed in Table 3. The graph highlights a clear upward trend from 2021 to 2023 and reached to a peak of 3288505 in 2023, followed by a decline back to 2633817 in 2024. However, the total number of service users in 2024 remains higher than in 2021.

3.3 Average Occupancy Rate Change

Figure 3 illustrates the trend in the average occupancy rates of beds and rooms from 2021 to 2024. Both categories exhibit an upward trend, although the increase in the occupancy rates of beds and rooms from year to year is quite modest. Overall, the average occupancy rate

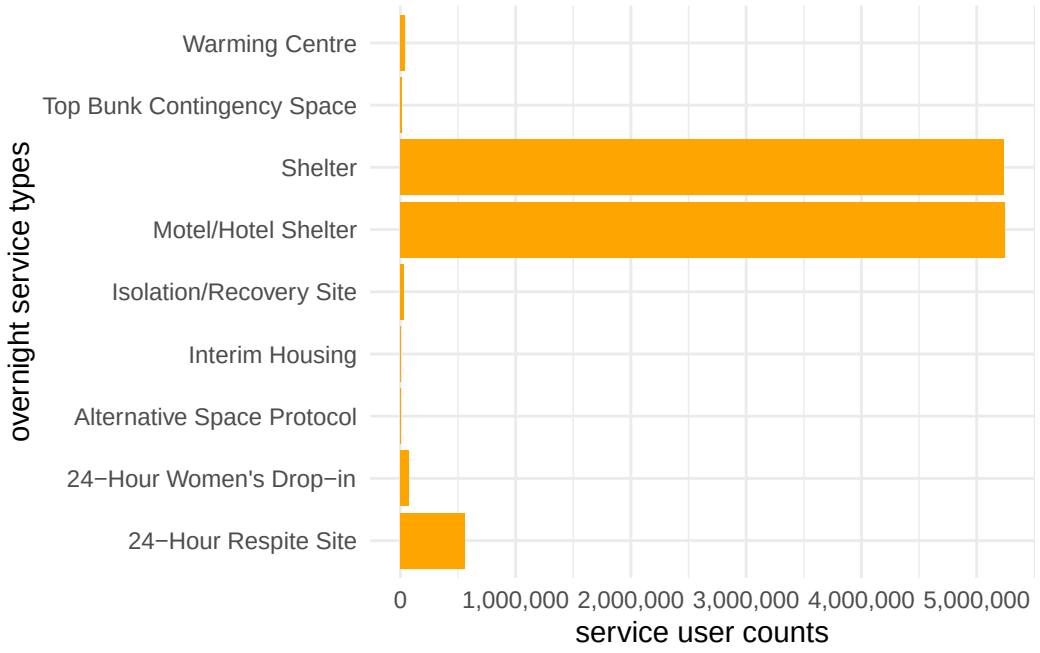

Figure 1: Total service user counts for different overnight service types

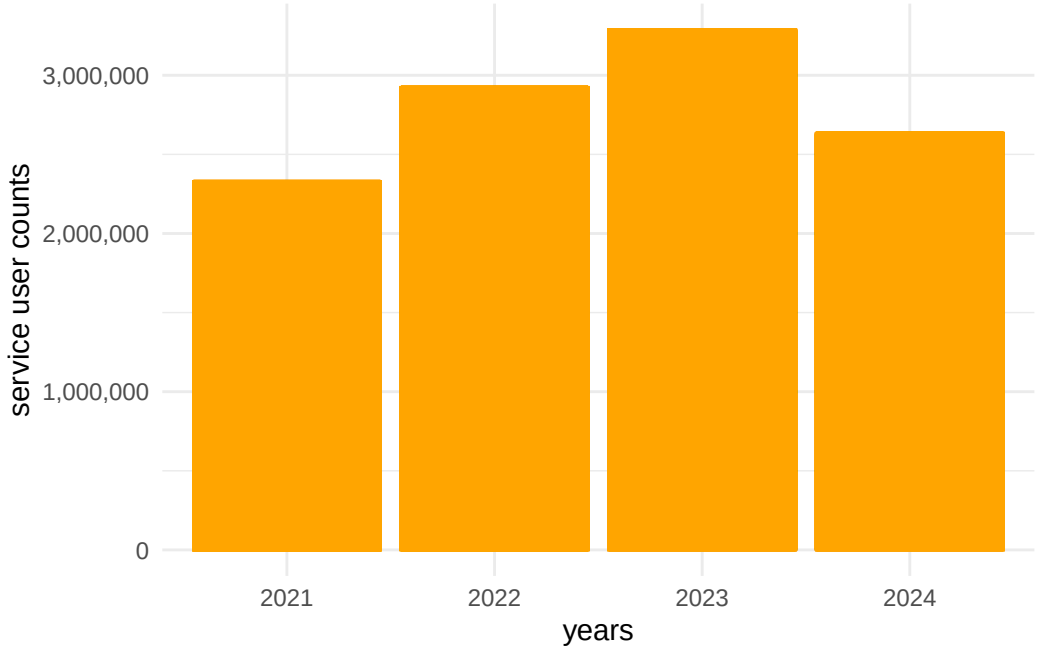

Figure 2: Service user count change from 2021 to 2024

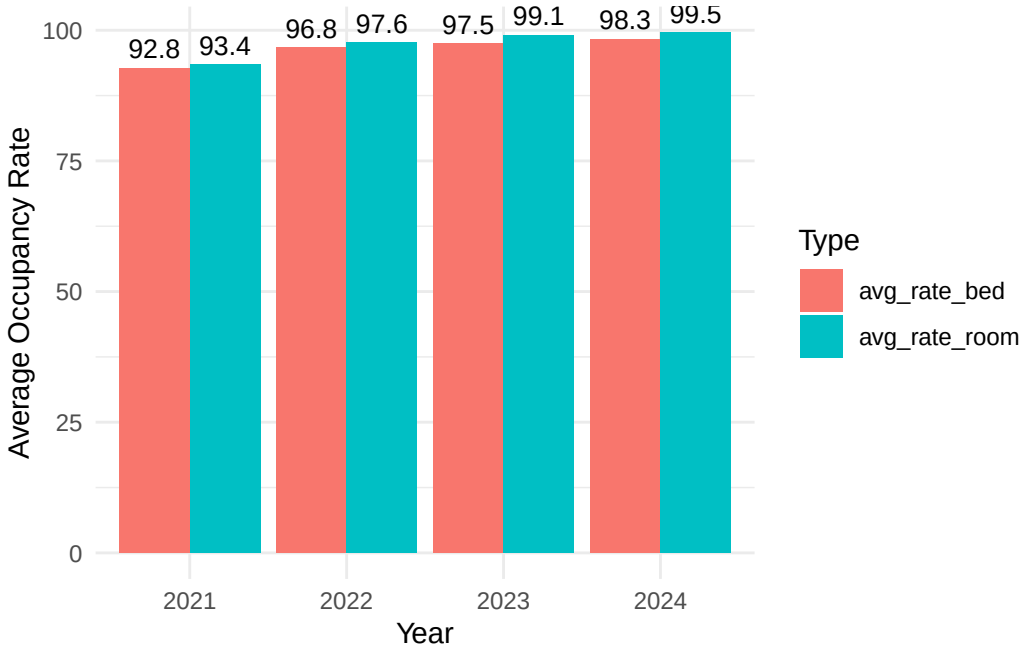

Figure 3: Average Occupancy Rate of Beds and Rooms from 2021 to 2024

for rooms consistently surpasses that of beds each year. In 2024, the average room occupancy rate reached 99.5%, indicating nearly full occupancy of shelter rooms, while the average bed occupancy rate achieved 98.3%, slightly lower than the average room occupancy rate.

3.4 Total Actual Capacity Change

Figure 4 illustrates the total actual capacity of beds and rooms from 2021 to 2024, allowing for a comparison of the capacities across two categories. Notably, as shown in Figure 3, rooms consistently exhibit a higher occupancy rate than beds over all four years. The plots indicate that the total actual capacity of rooms was only greater than that of beds in 2021 and 2022; however, in 2023 and 2024, the capacity of beds surpassed that of rooms significantly and raised to 1457984. Additionally, the total actual capacity of beds demonstrated an increasing trend from 2021 to 2023, followed by a decline to 1232028 in 2024. Conversely, the total actual capacity of rooms exhibited an upward trend from 2021 to 2022 which has its highest point at 1211399 of total actual capacity, after which it began to decrease until reaching its lowest point, 777790 in 2024.

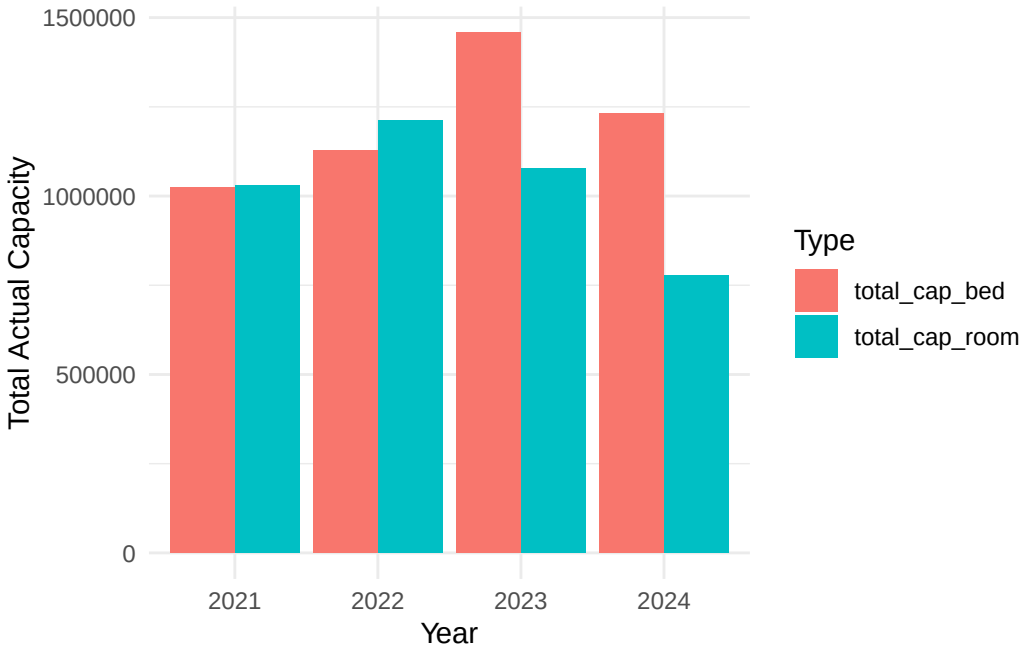

Figure 4: Total Actual Capacity of Beds and Rooms from 2021 to 2024

4 Discussion

4.1 Summary Conclusions and Implications

This paper analyzes trends in shelter occupancy and capacity in Toronto from 2021 to 2024, focusing on the increasing demand for shelter services and changes in actual capacity. Figures above highlight that traditional shelters and motel/hotel shelters are the most commonly used services, suggesting that these facilities are critical in addressing homelessness.

Furthermore, while the total service user count has generally increased, the average occupancy rates of beds and rooms have reached nearly full capacity, especially in 2024, where room occupancy approached 99.5%. This high occupancy rate underscores the strain on available resources and the urgent need for capacity expansion or better resource distribution.

Oppose to the upward trend in shelter occupancy rates, total actual capacity met a notable decline in 2024. This suggests that while demand for shelter services is rising, the availability of resources has not kept pace, particularly in the latest year. The analysis of service user counts, occupancy rates, and capacity provides insights into the operational challenges faced by shelters in meeting the needs of the city's homeless population.

4.2 Further Discussion

Further research should expand the time frame of analysis to understand long-term trends in shelter usage, occupancy and capacity. In addition, it would be beneficial to investigate the experiences and outcomes of shelter users to assess the quality and effectiveness of services provided. Comparative studies all over ther world facing similar homelessness challenges could provide valuable insights into better innovative solutions. Moreover, exploring the causes of capacity fluctuations during 2023 and 2024, such as funding changes or policy shifts, could help develop more sustainable strategies for managing shelter resources effectively.

Finally, developing predictive models to forecast shelter demand could assist policymakers and social service organizations in planning and allocating resources more efficiently as the demand model can help to predict the most appropriate resources that are needed by shelter users. These models can incorporate not only historical data but also external factors like economic conditions, housing market trends, and social policies which can help the analysis become more comprehensive.

4.3 Limitations

There are still some limitations in this paper. The data used from "Daily Shelter & Overnight Service Occupancy & Capacity" from opendatatoronto(Gelfand (2022)) spans only four years, which may limit the ability to generalize long-term trends in shelter usage. Within a longer terms of trend, prediction will be more accurate for Toronto's shelter usage situation. In turn, seasonal variations and short-term policy changes can have significant impacts on occupancy and capacity, which might not be fully captured in this analysis.

Additionally, while occupancy rates provide useful insights into shelter usage, they do not necessarily reflect the quality of services provided or the experiences of shelter users. Furthermore, the study relies on recorded data from the Shelter Management Information System(SMIS), which may not record all the actual situation in Toronto service programs .

The consideration of external factors that can affect average occupancy rates are not included in this paper. As the shelter demand can change at specific times, such as during extreme weather conditions or economic downturns.

References

Alexander, Rohan. 2023. *Telling Stories with Data: With Applications in r*. Chapman; Hall/CRC.

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/.

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

Müller, Kirill, and Lorenz Walthert. 2024. *Styler: Non-Invasive Pretty Printing of r Code*. https://CRAN.R-project.org/package=styler.

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.

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