

How homeless shelter systems affect people's lives and work

QIU WANTONG

02/07/2022

Abstract

An eye on the data about the supply relationship between overnight service and daily shelter can reveal the development of a city or even a country. Because homeless people reflect the local economic development, the executive power of government management, the gap between the rich and poor, and a perfect and healthy shelter system is to reduce the numbers of homelessness, which plays a vital role in social development. Compared with the past, this report can see whether the homeless shelters are improved, and more intuitively see whether the social problems are attached importance by the government and the country.

Introduction

With the Omicron outbreak occurring from the second half of 2021 to February 2022, people's lives and work have been greatly affected. According to Open Data Toronto's Data Base on SMIS, occupancy of daily shelters is close to 100%, which means that the homeless shelter community is still very large. In this situation, the government is also very concerned about homelessness and the improvement of the shelter system, and the government also increases the budget on shelter systems and epidemic prevention.(Manager 2021) Toronto is considering setting up more warming centers to protect homeless people from the pandemic.(Draaisma 2022)

Balancing the supply relationship of the daily shelter and data analysis of the overnight program are very important for the development and planning of the whole city. Because from the application demands of this homelessness, we can see whether the current social environment is a sustainable and healthy development. To put it simply, in this data from Open Data Toronto, we find that the occupancy rate of the daily shelter is so high that it is worth exploring some of the reasons and what impact it will have on society. Especially because of the outbreak of Omicron, we need to pay more attention to whether the supply of this homeless shelter is in place because without these data, it is difficult for relevant departments to help make protective measures of homelessness in each area according to corresponding data, rising numbers of the Omicron are also likely to be hard to control.

In order to have an in-depth understanding of the data set about occupancy and capacity of daily shelter, overnight service. I will select some variables in the data set before starting the analysis. To state analyzes the purpose and importance of Overnight service, daily shelter occupancy, and capacity. I will focus on analyzing some variables that contribute a lot to data analysis and try to find the connection between variables. Therefore, I will use the proportion of room occupancy in the program area variable to distinguish room occupancy for overnight service or winter program or temporary response program, or COVID-19 program. After that, I try to choose to observe the different number of occupied rooms in each location city to reflect the regional development of each city. Subsequently, we can see that different groups of people have different demands for services, according to the different data in the variable of the sector and the number of overnight service users. Finally, I performed more calculations for data of different types of overnight services in the table, so as to provide more data types for analysis. I will analysis and process my data in R.(R Core Team 2021) At the beginning I install the tidyverse(Wickham et al. 2019) package, dplyr(Wickham et al. 2021) package and knitrXie (2014) package to be used in my following R markdown

report. After that, I can make the table and figures by `ggplot2`(Wickham 2016). And the data set which I analyze are from the Open Data Toronto(Gelfand 2020).After creating the table and figures, I will give my paper a general introduction to help my audience have a pre-understanding of my topic.And then I write a abstract for leading the audience to see what happens in this story, and I try to present.Of course I need to give a glowing and straightforward title for the paper. And reference is really important for the paper in this way to respect the works of others. Finally, after I finishing writing the paper I need to install `tinytex`Xie (2019)to change my R markdown report to the form of PDF.

Data

Data Source

The data in the Open Data Toronto Portal is collected from the Shelter Support and Shelter Management Information System (SMIS) which analyzed the overnight service, demand, and supply of the daily shelter in Ontario in 2022. SMIS is a centralized information management system that provides information on shelter occupancy rates in Toronto. It can provide real-time data on occupancy rates. ((SMIS) 2019) In this data, it includes information about the shelter's occupancy rate, the shelter's location in Toronto, and a detailed breakdown of the shelter's capacity, including the number of rooms, beds, and so on.

Methodology and Data Collection

This data is collected from shelter's occupancy data collected by date record node. Once you've checked into Shelter, your check-in data will be collected at 4am the next day, meaning that all of that day's data will be counted for the next day. Due to the impact of COVID-19, overnight service is also affected, because the epidemic requires more physical distance between people, so the city is ready to add two emergency shelters to confirm the physical distance. (Aguilar 2022)

On the other hand, This data set is relatively comprehensive in the classification of shelter occupancy rate, including more detailed classifications related to the capacity of actual rooms and the capacity of funding rooms. In this way, analysts can accurately understand the supply relationship of these shelters and overnight services when they receive the data so that they can make some realistic plans for the future. In addition, the data set also provides the corresponding ID numbers of program, shelter, and organization, so that all data can be made more concise during data analysis, and no matter what changes are made to these names, the ID numbers will not be changed. In this way, the corresponding data will not be found due to different names in data analysis. However, although we have detailed information about overnight service and shelter occupancy in this data set, there are a lot of missing values in these useful occupancy observations. This missing information will lead to deviations in data statistics, leading to unreliable data.

Data Characteristics

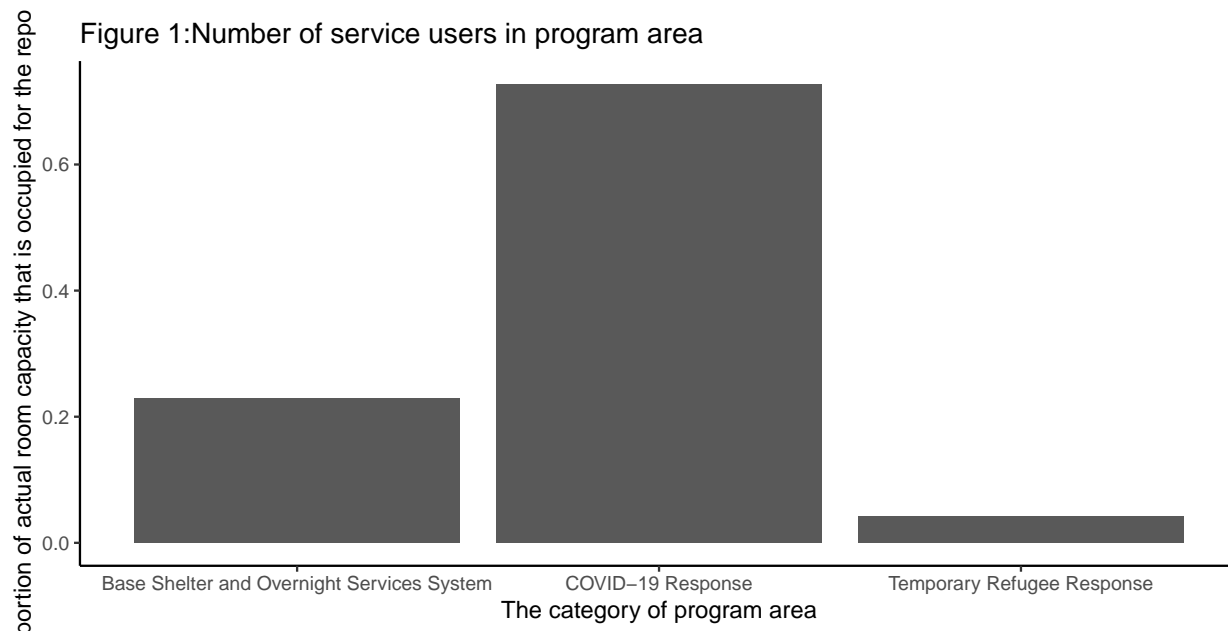
There are 5007 observations of 32 variables in this data set, including occupied date, the number or percentage of beds and rooms of occupancy. Type of Capacity, number of actual and funding beds and rooms, organizations providing services, locations information, types of Overnight service, etc. This data is classified in detail. All the data related to occupancy and capacity of overnight service and daily shelter that can be analyzed are collected in this data set.

Some of the variables contribute significantly to the analysis of the whole data set, for example, counting the overnight service types and users, which indicates that people make different choices for different overnight services. Then the occupancy of the room, this variable can let us see the data response of shelter's occupancy in different aspects. The program area indicates whether the program is more about the base shelter and overnight services system, or is about a temporary response program. Sector refers to the data classified by gender, age, and household size for those who need shelter. Next, I will describe the relationship between these variables in charts and explain them.

Program area and the percentage of the occupied room

According to the information provided by the SMIS in open data Toronto, in the first figure, I analyzed the relationship between the program area variable and the percentage of the occupied room variable. In the

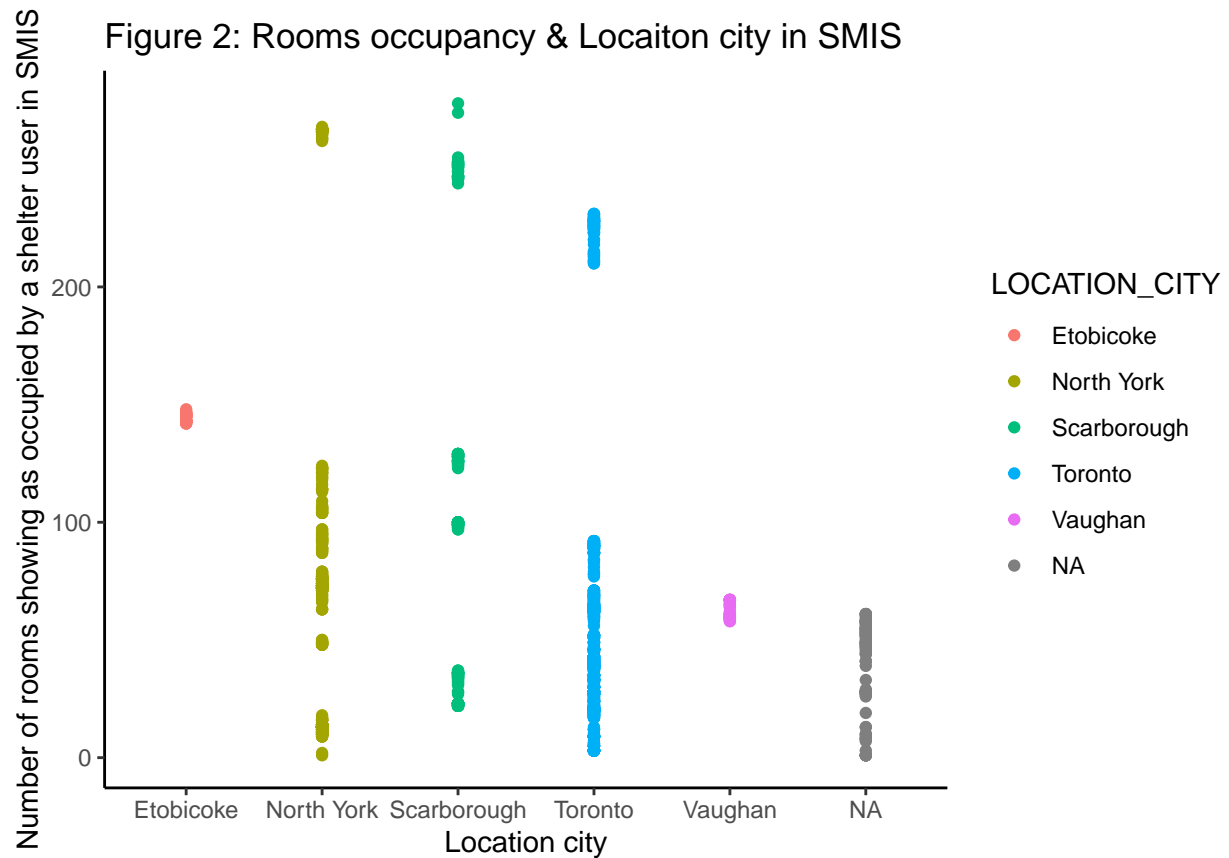
program area variable of data set, we can see that it is divided into three areas: Base Shelter and Overnight Services System, COVID-19, and Temporary Refugee. We can see the occupancy ratio of rooms in these three sections in this bar chart .



In this bar chart, we can see that the response of COVID-19 accounts for the largest proportion, which is over 60%, while temporary refugee accounts for the least proportion, which is less than 20%. Then, base shelter and overnight service systems accounted for more than 20%. As we can see from this ratio, the shelter's occupancy rate was greatly affected by Omicron's explosion.

Location city and occupied rooms

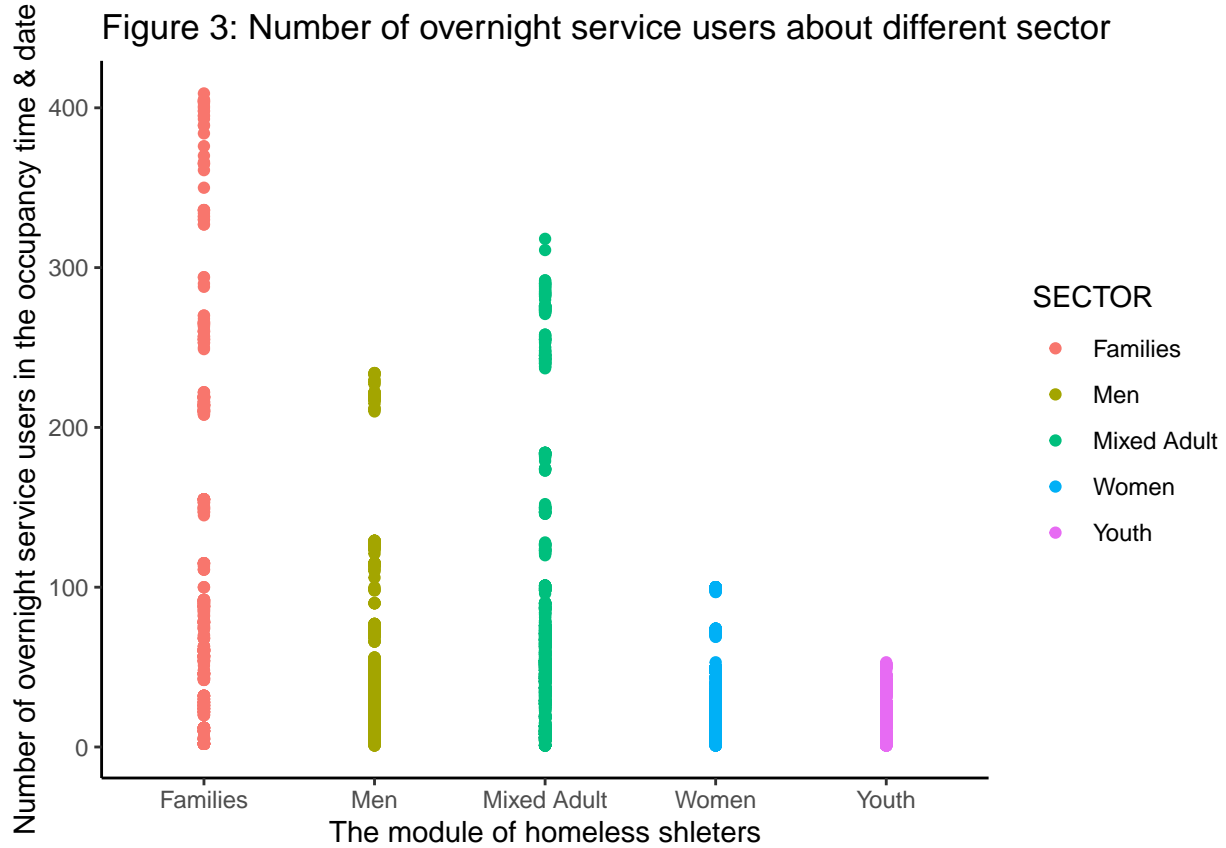
For the second figure, I use a geometric point plot to show the relationship between location city and rooms occupancy variables. Location city contains 5 regions and data of missing value. In order to ensure data integrity, I did not filter out the missing value here. The occupied rooms variable is about the number of occupied rooms by the SMIS database in open data Toronto on a specific date and time. Here, we can observe the distribution relationship between these two variables through the plot.



In this point plot, if the missing value is ignored, Scarborough district occupies the largest number of rooms for daily shelter, and Vaughan district occupies the least. At the same time, shelters in North York and Toronto each occupied more than 200 rooms. From this figure, we can see that the three areas with relatively dense population distribution occupy more rooms. Meanwhile, due to some missing values, the accuracy of data analysis needs to be improved.

Sector and Service user count

Based on the data from open Data Toronto, We can make a geometric point plot to observe the relationship between sector variable and service user count variable. The Sector variable was divided into 5 categories, namely families, men, mixed adults, women, and youth. According to this plot, we can see the number of people selected overnight service in 5 different homeless shelters, respectively.



In this point plot, we can see that the number of families with the highest demand for overnight service exceeds 400 users. Then the group with the least demand is the youth group, with less than 100 users. Mixed adult has the second-largest demand, with more than 200 users. Men have overnight services users of over 200, and women have over 100 fewer users than men. From this figure, we can intuitively see that families need to receive high attention from government departments, and it is also worth thinking about how to help them.

Data about overnight service types

Table 1: More Data about Overnight Service Type

OVERNIGHT_SERVICE_TYPE	min	Q1	median	Q3	max	IQR	mean	sd
24-Hour Respite Site	12	27	40	47.00	184	20.00	51.58559	46.30259
24-Hour Women's Drop-in	16	18	24	37.00	44	19.00	27.40278	10.19619
Isolation/Recovery Site	1	13	48	54.75	62	41.75	36.14000	21.42287
Motel/Hotel Shelter	2	38	65	106.00	409	68.00	94.53007	89.70374
Shelter	1	15	25	47.00	234	32.00	38.13192	42.31483

OVERNIGHT_SERVICE_TYPE	min	Q1	median	Q3	max	IQR	mean	sd
Warming Centre	2	15	58	68.00	71	53.00	43.65979	25.37259

According to the database from SMIS, this table is intended to provide more data about overnight Service type. It includes minimum and maximum values, the values of the median, first and third quantile, Inter Quartile Range, and Standard Deviation. The mean and median of overnight service types are similar at 52, 27, 36, 95, 38, 44, which supports that the distribution of the number of overnight service types is not symmetric. The sample standard deviation (46.3025934, 10.1961878, 21.4228659, 89.703737, 42.3148342, 25.3725929) is relatively big, meaning the uncertainty of the data is relatively high.

Reference

- Aguilar, Bryann. 2022. “Toronto to Open Two Emergency Shelters, Distribute More Than 310k N95 Masks to Homeless,” 1.
- Draaisma, Muriel. 2022. “City to Consider Opening More Warming Centres for Unhoused People When Extreme Cold Hits Again,” 1.
- Gelfand, Sharla. 2020. *Opendatatoronto: Access the City of Toronto Open Data Portal*. <https://CRAN.R-project.org/package=opendatatoronto>.
- Manager, City. 2021. “COVID-19 Response Update: Protecting People Experiencing Homelessness and Ensuring the Safety of the Shelter System,” 1.
- R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- (SMIS), Shelter Management Information System. 2019. “TORONTO,” 1.
- 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 Golemund, 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. 2021. *Dplyr: A Grammar of Data Manipulation*. <https://CRAN.R-project.org/package=dplyr>.
- Xie, Yihui. 2014. “Knitr: A Comprehensive Tool for Reproducible Research in R.” In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC. <http://www.crcpress.com/product/isbn/9781466561595>.
- . 2015. *Dynamic Documents with R and Knitr*. 2nd ed. Boca Raton, Florida: Chapman; Hall/CRC. <https://yihui.org/knitr/>.
- . 2019. “TinyTeX: A Lightweight, Cross-Platform, and Easy-to-Maintain LaTeX Distribution Based on TeX Live.” *TUGboat*, no. 1: 30–32. <https://tug.org/TUGboat/Contents/contents40-1.html>.
- . 2021a. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. <https://yihui.org/knitr/>.
- . 2021b. *Tinytex: Helper Functions to Install and Maintain TeX Live, and Compile LaTeX Documents*. <https://github.com/yihui/tinytex>.