The Effects of Homeless Shelter Occupancy Rates on Non-Fatal Opioid Overdoses*

Justin Klip

December 3, 2024

1 Introduction

Overview Paragraph: Opioid overdoses have been a prominent issue in North America for some time. Opioid-related deaths have doubled in Canada from 2019 to the end of 2021 (Bains 2024). Homeless opioid overdoses in shelters have climbed as well during the pandemic, with opioid deaths in Ontario shelters more than tripling during the pandemic (Alhmidi 2024). While many avenues could be explored for this climb, one area to potentially look at is the effects of the shelter conditions themselves.

The remainder of this paper is structured as follows. Section 2 describes how the data was acquired, how it was measured, and provides visualizations on the overdoses and occupancy in shelters.

2 Data

2.1 Data Source

I make use of two data sources in order to run my analysis, both from Open Data Toronto ("Daily Shelter Overnight Service Occupancy and Capacity Data" 2024). Firstly I make use of Open Data Toronto's Daily Shelter Overnight Capacity Data (Shelter and Services 2024). Secondly I use Open Data Toronto's Fatal and Non-Fatal Suspected Opioid Overdoses in the Shelter System data set (Services 2024).

Using the statistical programming language R (R Core Team 2023), the data was then cleaned, tested, analyzed, and merged using the R-packages arrow, (Richardson et al. 2024), dplyr

*Code and data are available at: [https://github.com/justinklip/toronto_shelter_overdose_study).

(Wickham et al. 2024a), here (Müller et al. 2024), gtools (Warnes et al. 2024), MASS (Ripley and Venables 2024), tidyr (Wickham et al. 2024d), opendatatoronto (Burt 2024), tidyverse (Wickham et al. 2024e), ggplot2 (Wickham et al. 2024b), and testthat (Wickham et al. 2024c). The daily shelter attendance data was aggregated from the daily data in the shelter data base to a quarterly record in order to merge with the overdose data set. Details are in the appendix on how exactly observations were dropped and merged are available in the appendix. Some dummy variables were added into the data set in order to account for selection induced upon dropping missing values.

2.2 Measurement

The Daily Shelter Overnight Capacity Data from Open Data Toronto is measured directly from administrative data (Shelter and Services 2024). According to Open Data Toronto, this administrative data comes from the Toronto Shelter and Support Services division's Shelter Management Information (SMIS) database. The data is generated as follows: a shelter records metrics in their shelter at exactly 4:00am every day (likely through a computer or check-in system) such as attendance numbers, capacity numbers, and respective capacity utilization rates, then uploads that data daily in compliance with the SMIS data requirements. Generally shelters measured occupancy and capacity in one of two ways: beds or rooms. Bed-based occupancy was used in shelters that had communal sleeping areas, whereas room-based occupancy was used for more family-oriented programs. An "individual" in this data set would counts as someone who was served in a service area or program. Interestingly data is only recorded for the first quarter of 2023, this and other limitations of the data are discussed further in the appendix.

The Fatal and Non-Fatal Suspected Opioid Overdoses in the Shelter System Data Set comes from paramedic rather than shelter level data (Services 2024). This data only counts for specific kinds of shelters: shelter-hotels, emergency shelters, and shelter-hotels, meaning that when the data is merged, only shelters coming from this data set will be used. An entry in the data set is formed as follows: a shelter member or employee calls 911 for an emergency, when Paramedics arrive and if they determine on the scene that this is a suspected overdose, then that location gets an observation added to the data set. It's important to note also that if a particular address has less than 5 non-fatal overdoses in a particular quarter, then the true amount is not published for anonymity purposes. This data set also includes fatal overdoses, and a fatal overdose is only added as an entry if the Coroner's office determines the cause of death to be an overdose. This data is also not matched at the shelter level and is only aggregate statistics, so it will be ignored for the purpose of analysis since I am more interested in the relationship between capacity and overdoses.

2.3 Summary Statistics

Table 1: Summary Statistics for Non-Fatal Overdoses, Average Occupancy Rate, and Daily Average Users

Variable	Observations	Mean	Median	Standard Deviation	Minimum	Maximum
Non-Fatal Overdoses	271	7.206642	5.00000	10.058059	0.0000000	51.0000
Average Occupancy Rate	492	96.436265	99.39367	7.338487	21.8904348	100.0000
Daily Average Users	492	71.313895	44.08242	72.927412	0.0777778	383.0989

Table 1 gives the summary statistics of our variables of interest. We can see that most of the occupancy rate is extremely high and centered near 96 percent for the mean and 99 percent for the median, although there is some standard deviation. As for daily average users, the median amount daily is about 44, suggesting most shelters are about medium size. The really low minimum amount here is likely seen by a shelter that was temporarily open, as averages are calculated by dividing by the number of days in the quarter. The mean is much higher than the median at 71 meaning that some shelters are a lot larger, dragging the average up. This is indicated by the maximum daily average users of one shelter being and the large standard deviation. The non-fatal overdoses (excluding entries labelled "<5") are centered around 5, suggesting the true median is likely lower and within that less than 5 category.

2.4 Outcome variables

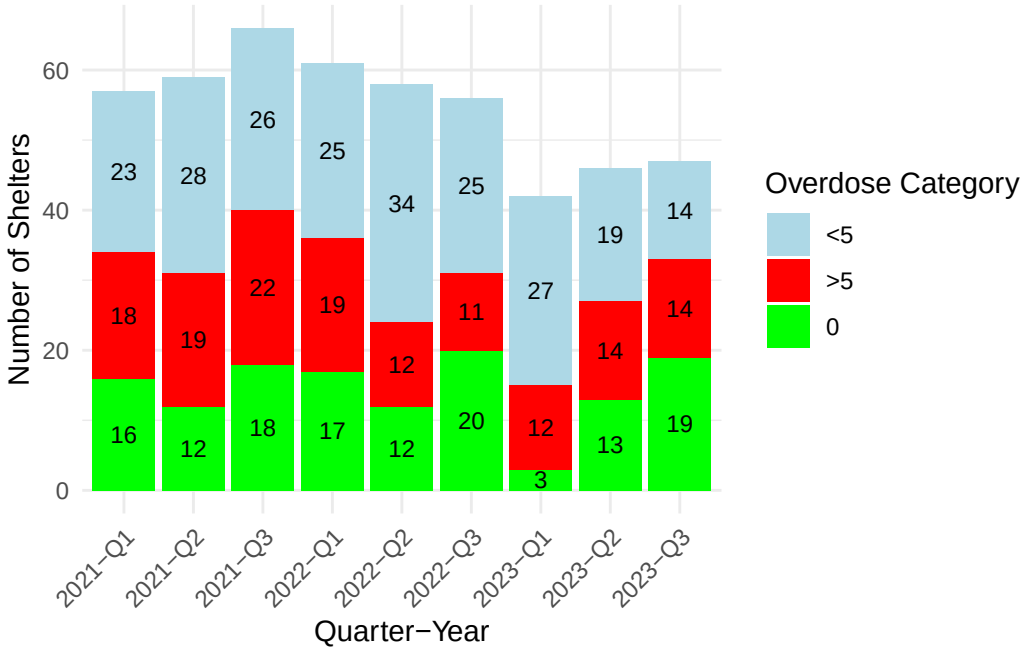

Figure 1: Shelters Overdoses by Quarter-Year By Data Category

Figure 1 provides information on the number of overdoses for each shelter in the data set from quarter to quarter. It is important to note that the data set does not include Q4 data. The first thing to note is that the number of shelters seems to decrease in 2023, this is likely consistent with COVID-19 specific shelters beginning to close. Another important thing to note is that the number of overdoses greater than 5 and the amount less than 5 are generally split evenly between the shelters. This suggests we have some "high overdose" level shelters, likely because they are bigger, and others which are smaller and less drug prone.

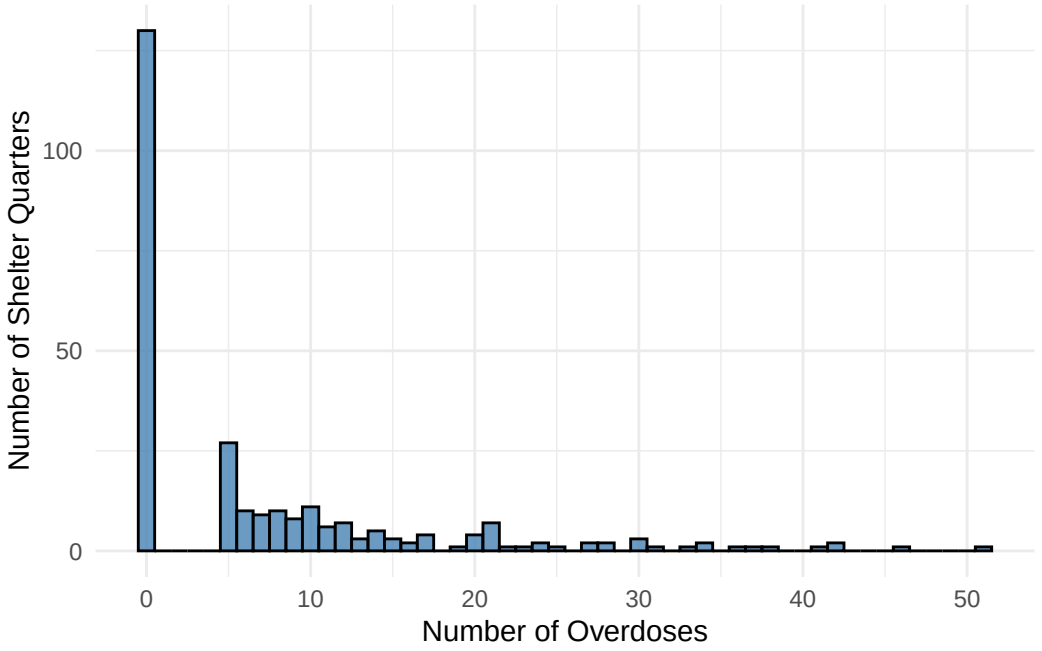

Figure 2: Distribution of Overdoses Counts, Shelter-By-Quarter, excluding nonzero less than 5 counts.

Figure 2 plots the distribution of the number of overdoses for each quarter in each shelter, except those that are marked "<5". The data seems to indicate a clear mass around 0, with about 150 out of approximately 500 total observations being located at this point. Considering the shape of the data, it would be reasonable to assume that there are less and less observations as the number of overdoses per quarter increases at a certain address.

2.5 Predictor variables

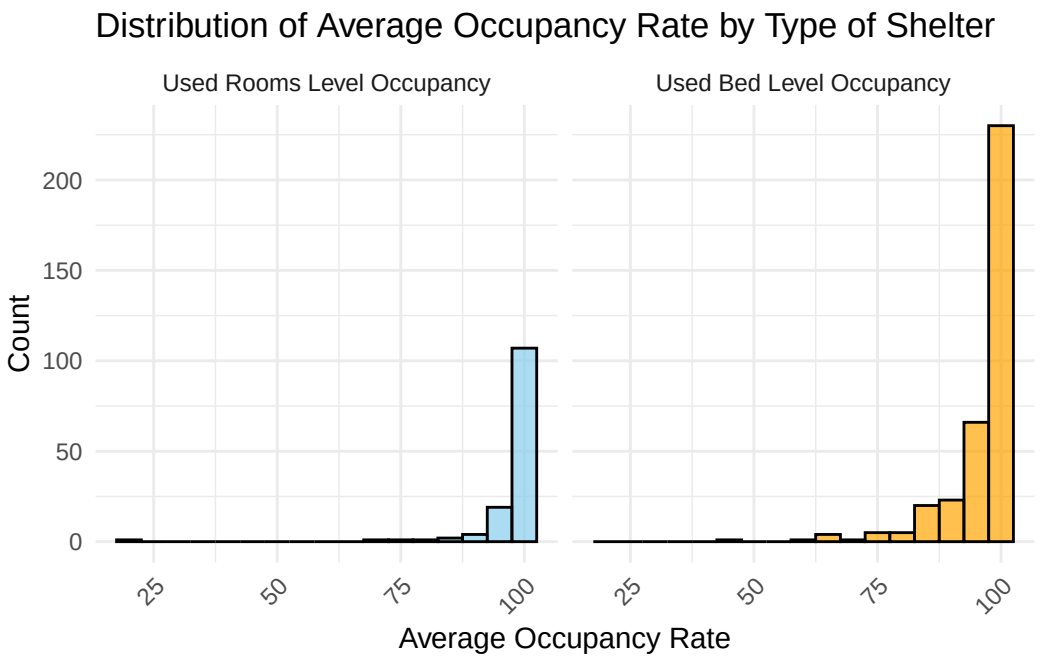

Figure 3: Distribution of Average Quarterly Occupancy Rates in Toronto Shelters by Type of Shelter

Figure 3 plots the distribution of the average occupancy rate by quarter-year, as seen in some quarters, some shelters do have less than 100 percent occupancy, although a lot spend their time at the 95-100 percent occupancy rate. For the bed data there are a lot more observations with less than 95 to 100 percent occupancy. This makes sense as it is a lot easier for a shelter to fill up all their rooms than it is to fill all their beds.

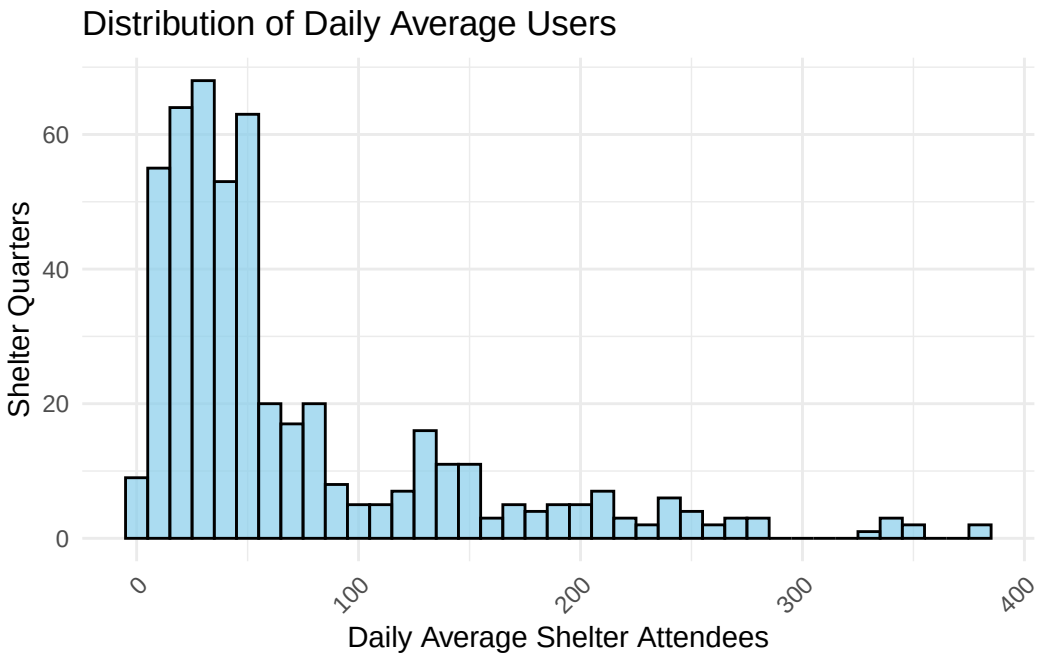

Figure 4: Distribution of Daily Average Number of People of Toronto Homeless Shelters

Figure 4 plots the distribution of the daily average number of people in Toronto Shelters each quarter, we see that most shelters are relatively small-medium scale, with less than 100 average people staying a shelter per night on average each quarter. There are some larger shelters too, explaining why the tail of the distribution moves right, with a bunch of shelters having more than 200 daily average users each quarter.

Appendix

3 Data Limitations and Improvements

Due to data issues, my modelling left a lot to be desired. In this segment I document segments of the observational data that I would include to improve data to allow for better modelling. I also lay out a potential way causal inference could be done given the data if we were to strengthen our data in a few ways.

3.1 Observational Data

Our approach to deal with censored values (where a shelter simply said their overdose count was less than 5 for a quarter) was simply to set the value to 1, as that would give me a lower bound estimate for the effect of shelter occupancy on overdoses. The idealized data would simply not have these censored observations, allowing the models to more accurately represent the data at hand. While privacy concerns were cited as the main reason for this, ideally I would sign some data confidentiality agreement in order to be able to use this data.

The next issue comes from the lack of information in the shelter overdose data set,

3.2 Idealized Survey

Another way that the data could be improved would be including surveys of shelter members. This would more adequately describe living conditions in a way that the current data could not. While I try to use average occupancy rates and total attendance amounts as a proxy for 'crowdedness', arguably a better way to do this would be to ask members of these shelters themselves. This is because most shelters try to stay near 100 percent occupancy at all times, so there is not much variation for analysis using this as a predictor. Surveys could be randomly assigned to a certain percentage of members of each shelter (e.g 10%) quarterly asking a question like "what degree of privacy do you feel like you have" with a scale. I would also assure members of their anonymity to avoid biases relating to going for middle answer. Combining both the survey, as well as the the average occupancy rate as a control, with the fixed overdose data would provide much more informative results.

References

Alhmidi, Maan. 2024. "Opioid Deaths in Ontario Shelters More Than Tripled During Pandemic, Study Finds | CBC News." *CBC News*. CBC/Radio Canada. https://www.cbc.ca/news/canada/toronto/ont-shelters-overdose-deaths-1.7238916.

Bains, Camille. 2024. "1 in 4 Deaths of Canadians Age 22 to 39 in 2021 Was Opioid Related in 2021, Researchers Find | CBC News." *CBC News*. CBC/Radio Canada. https://www.cbc.ca/news/health/opioid-young-people-1.7174098.

Burt, James. 2024. *Opendatatoronto: Access Open Data from the City of Toronto*. https://cran.r-project.org/package=opendatatoronto.

"Daily Shelter Overnight Service Occupancy and Capacity Data." 2024. *City of Toronto Open Data Portal*. https://open.toronto.ca/dataset/daily-shelter-overnight-service-occupancy-capacity/.

Müller, Kirill et al. 2024. *Here: A Simpler Way to Find Your Files*. https://cran.r-project.org/package=here.

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

Richardson, Neal et al. 2024. *Arrow: Integration to Apache Arrow*. https://cran.r-project.org/package=arrow.

Ripley, Brian D., and William N. Venables. 2024. *MASS: Support Functions and Datasets for Venables and Ripley's MASS*. https://cran.r-project.org/package=MASS.

Services, Toronto Paramedic. 2024. "Fatal and Non-Fatal Suspected Opioid Overdoses in the Shelter System." https://open.toronto.ca/dataset/fatal-and-non-fatal-suspected-opioid-overdoses-in-the-shelter-system/.

Shelter, Toronto, and Support Services. 2024. "Daily Shelter Overnight Service Occupancy and Capacity." https://open.toronto.ca/dataset/daily-shelter-overnight-service-occupancy-capacity/.

Warnes, Gregory R. et al. 2024. *Gtools: Various r Programming Tools*. https://cran.r-project.org/package=gtools.

Wickham, Hadley et al. 2024a. *Dplyr: A Grammar of Data Manipulation*. https://cran.r-project.org/package=dplyr.

——— et al. 2024b. *Ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics*. https://cran.r-project.org/package=ggplot2.

——— et al. 2024c. *Testthat: Unit Testing for r*. https://cran.r-project.org/package=testthat.

——— et al. 2024d. *Tidyr: Tidy Messy Data*. https://cran.r-project.org/package=tidyr.

——— et al. 2024e. *Tidyverse: Easily Install and Load the 'Tidyverse'*. https://cran.r-project.org/package=tidyverse.