# Vizualizing Toronto Homess Shelter data Across a Global Pandemic*

## A Look Into How The Pandemic disrupted the Flow Of Homeless Shelters And How The Situation Has Improved Since

Adam Labas

06 February 2022

**Abstract**

In addition to a global pandemic, homlessness and housing is a self apparent issue in the city of Toronto. Taking a stroll through certain areas in the downtown core, it's almost impossible to miss homeless tent cities and homeless people lined outside of shelters awaiting food or entry in the afternoon. Upon further examination of data collected from the Open Data Portal of the City of Toronto, it was easy to find interesting trends in homeless shelter data in the years 2020 and 2021. We found trends in the diminishing use of shelters by various demographic of homeless people and age groups. This paper attempts to visualize these trends as well as touch on some problems that one might experience in trying to extract real meaning from the data.

## Contents

*Code and data are available at: https://github.com/adam-labas/adam_labas_paper_1_opendatatoronto

# 1　Introduction

The year 2020 was no doubt a year of unspoken discomfort and uncharted territories for everyone. For myself, the year 2020 started of like all the others. Winter commutes to school and winter festivities like ice skating with friends. As the semester started to develop and I started to get into a routine with my new classes, we started hearing of the corona virus which had supposedly arrived from China and was rapidly making its way around the globe. Although this seemed as nothing but a farce and a scene from a movie, the reality and gravity of the situation rapidly started to be apparent. As school was brought to an abrupt stop after the World Health Organization had declared COVID-19 as a pandemic on March 11th 2020 (https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020), the magnitude on the situation started to manifest itself in everyone daily lives. However, many lives were affected much more than others.

While I was mostly confined to the four corners of my home for the vast part of 2020. Being indoor during the cold Canadian winter has never been more of a challenge. However, it does require a brief explanation to claim that some people under no confinement at all, would be in a much more grave situation. For long, homelessness has been a problem in the city of Toronto. Going for a stroll through certain areas in the downtown core, it's almost impossible to miss homeless tent cities and homeless people lined outside of shelters awaiting food or entry in the afternoon. The privilege that people who live in homes have is immense and the difficulties that homeless people face daily is a reality that no one ever wants to live.

In this analysis, I had a look at Toronto Homeless shelter data, which is described with great extent in the Data Section 2. In my analysis, I decided to focus on how the global pandemic has impacted the homeless community of Toronto by creating visuals on the progression of the shelter usage thought the 24 months from January 2020 to December 2021. To further break down the analysis, I decided to look into the distribution of the aforementioned data but segregating the data not only by the various age group bins provided in the data sent but also the different demographic population groups provided. In general, it was found that shelter usage dropped drastically, immediately after the start of the pandemic and has recently been slowly trending back to its original usage flow. In this paper, I will attempt to take you on a journey to explore the data that was used to produce the findings, I will visually present to you the findings and I we will speak about the implications of the findings and later on in the Limitations Section 6, we will discuss the limitations of our data and problems that this may have when wanting to extract true, ethical, unbiased information from the implications of the findings.

You can and should cross-reference sections and sub-sections. For instance, Section 2. R Markdown automatically makes the sections lower case and adds a dash to spaces to generate labels, for instance, Section 6.1.

abstract: "In addition to a global pandemic, homelessness and housing is a self apparent issue in the city of Toronto. Taking a stroll through certain areas in the downtown core, it's almost impossible to miss homeless tent cities and homeless people lined outside of shelters awaiting food or entry in the afternoon. Upon further examination of data collected from the Open Data Portal of the City of Toronto, it was easy to find interesting trends in homeless shelter data in the years 2020 and 2021. We found trends in the diminishing use of shelters by various demographic of homeless people and age groups. This paper attempts to visualize these trends as well as touch on some problems that one might experience in trying to extract real meaning from the data."

# 2  Data

In this Data Section 2, I will provide a look into the data acquisition and processing methodology as well as a deep dive into the contents of the data.

## 2.1  Data Collection

All the data that is used in this paper was retrieved from the City of Toronto Open Data Portal and is titled "Toronto Shelter System Flow." Using the R package opendatatoronto (Gelfand 2020), we loaded the data in the R-script titled *Data Acquisition and Processing*. The data is uploaded and funded by the City of Toronto and is refreshed monthly, last being updated on January 7th 2022 as of February 6th 2022. As the webpage in the Open Data Toronto Portal explains, "the City of Torontofunds and operates services dedicated to people experiencing homelessness in Toronto." They go on to explain that the services which they provide like emergency shelters, hotels/motels, warming centers etc. all use the Shelter Management Information System (SMIS) to track and log people entering the municipally funded homeless shelter system. When a new person enters the SMIS for the first time, many data points are collected. These variables are presented in our data. They include 14 variables like the current month and year for which the count is being recorded, the number of people who are returning/moving to/from housing, returning to shelter and more importantly for our analysis, the population group and the age group which the homeless person fits in.

## 2.2  Variables of interest

Population group is divided into seven categories, All Population, Chronic, Families, Indigenous, Non-refugees, Refugees, Single Adult and Youth, with Indigenous only being a population group which the SMIS began tracking in January 2021. Certain group definitions are important. First off, the population group Youth is defined by the Open Data Toronto Portal as "unaccompanied youth and includes people who are between 16 and 24 years old by the last date of the reporting month and are not members of a family." Next, the age groups in the data set are divided as such: peoples aged under 16, aged 16 to 24, aged 25 to 44, aged 45-64 and aged 65 and above. Importantly, we also use the date variable to

## 2.3  Data Processing

To process our data, we began by loading all the necessary libraries like opendatatoronto @citeodt whichwas used to load the data into R @citeR, a package called lubridate @citelubridate which will be used to process date Data

Table 1: First ten rows of a dataset of shelture usage

| Population Group | Aged less than 16 | Aged 16-24 | Aged 25-44 | Aged 45-64 | Aged 65 |
|---|---|---|---|---|---|
| All Population | 1440 | 1227 | 3893 | 2852 | 504 |
| Chronic | 429 | 432 | 1096 | 1205 | 307 |
| Refugees | 969 | 341 | 1177 | 427 | 27 |
| Families | 1438 | 238 | 790 | 231 | 9 |
| Youth | 0 | 987 | 0 | 0 | 0 |
| Single Adult | 0 | 0 | 3103 | 2621 | 495 |
| Non-refugees | 471 | 886 | 2716 | 2425 | 477 |
| All Population | 1434 | 1185 | 3909 | 2795 | 511 |
| Chronic | 405 | 435 | 1029 | 1251 | 309 |
| Refugees | 961 | 334 | 1174 | 407 | 26 |



Use of shelter users under age 16 by population group.
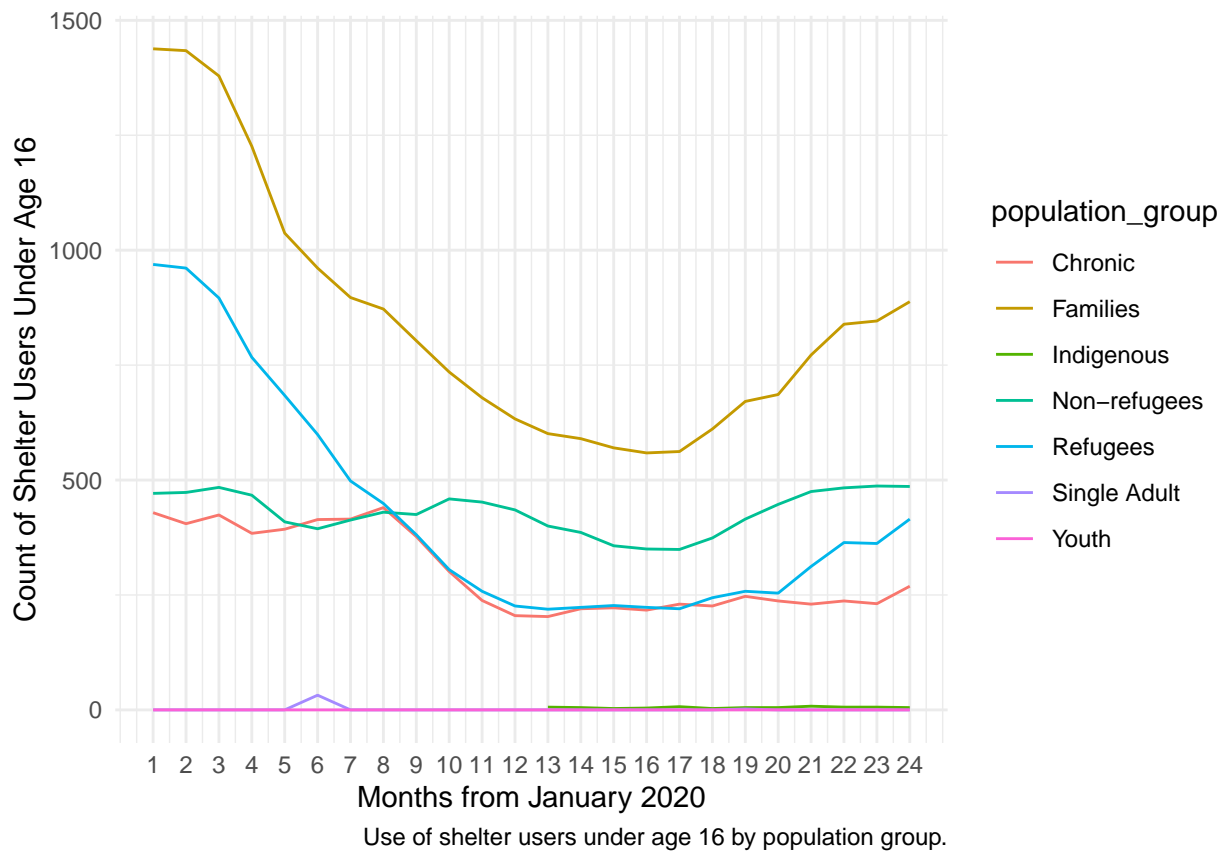
Talk more about it.

Also bills and their average (Figure **??**). (Notice how you can change the height and width so they don't take the whole page?)

```
ageunder16_overtime <- all_population %>%
  ggplot(aes(x = date,
             y = ageunder16)) +
  labs(x = "Date",
       y = "Under 16",
       caption = "Shelter use by those aged under 16 in 2020 and 2021") +
  geom_line() +
```
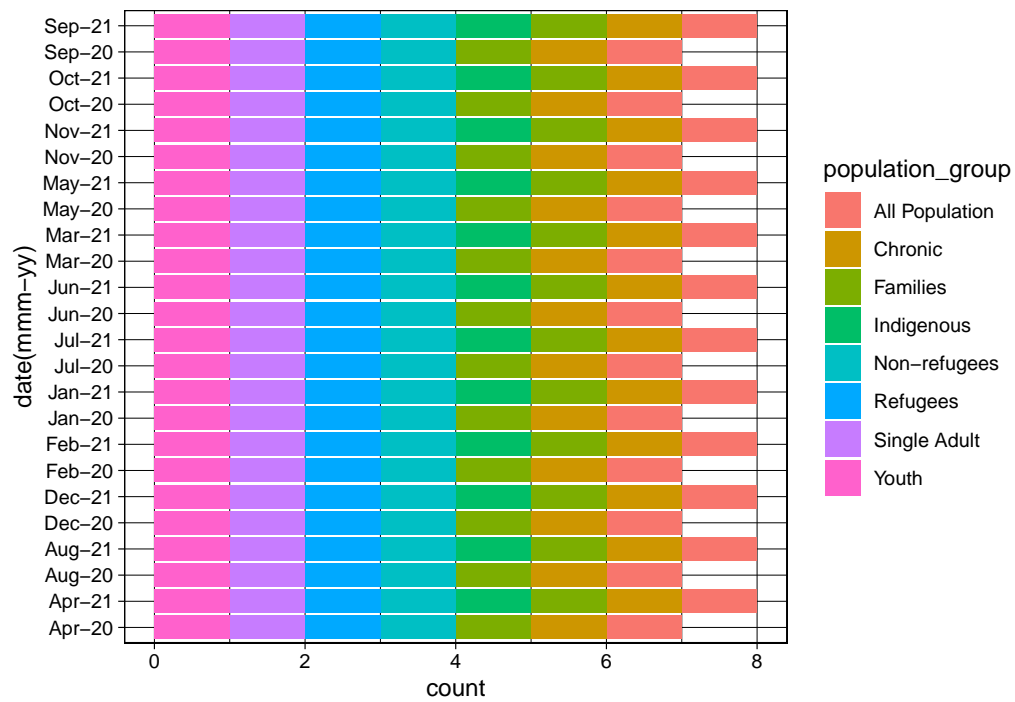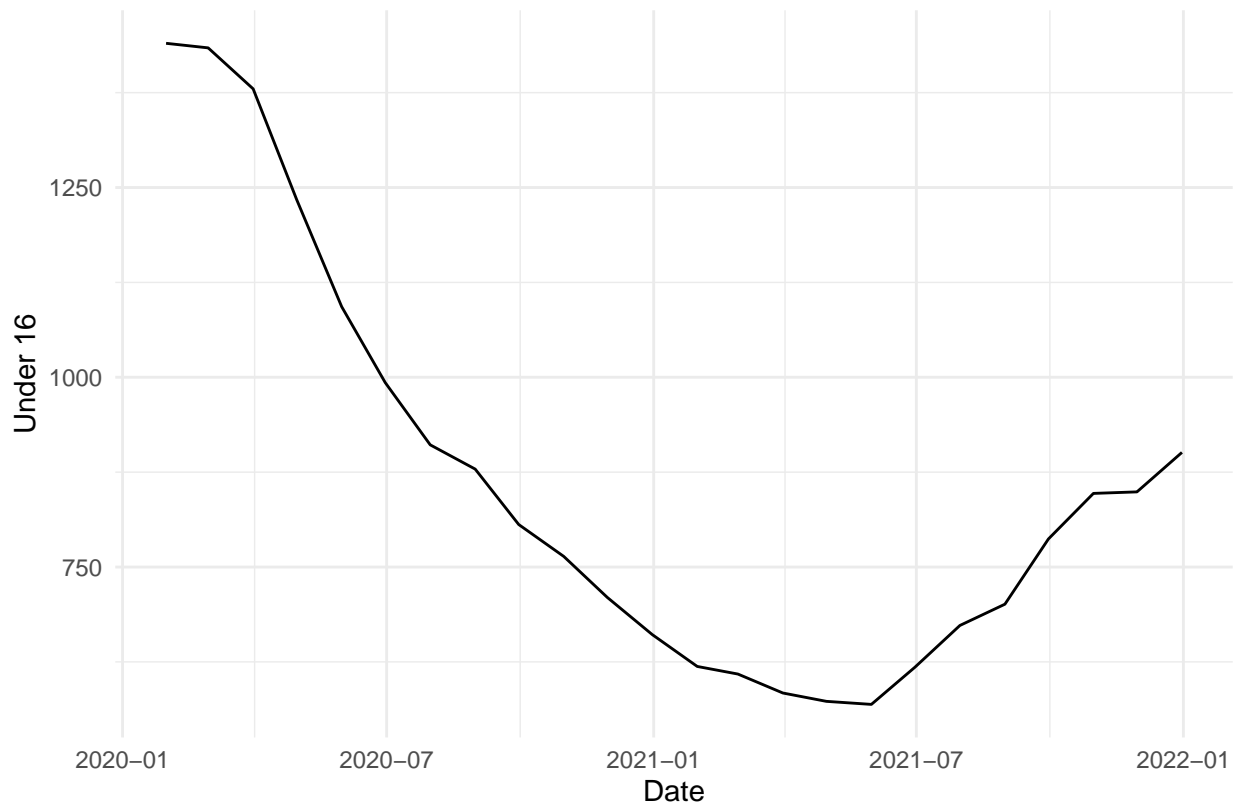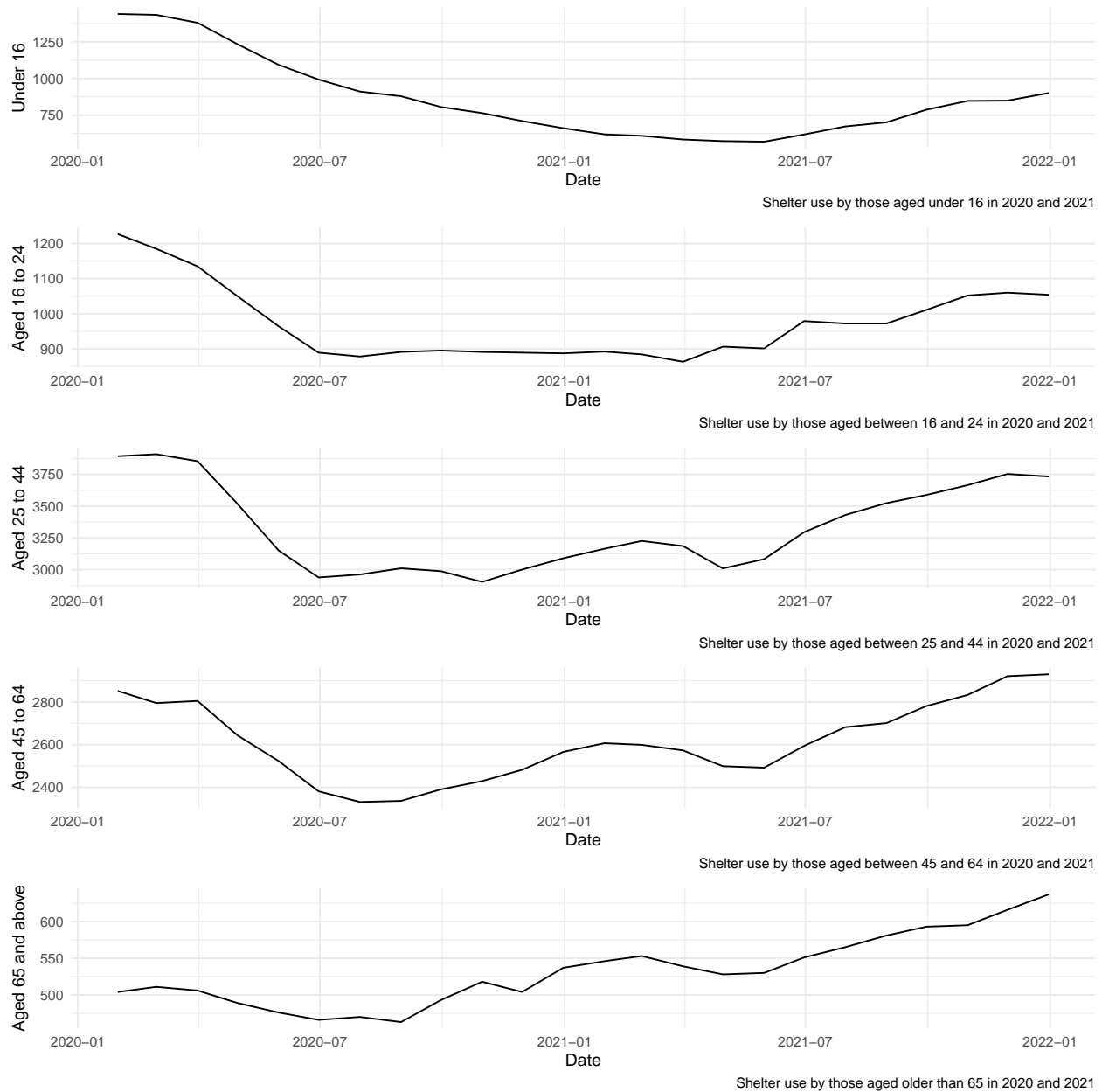
Figure 1: There are no Indeginous data for 2020

```
  theme_minimal()
ageunder16_overtime
```

Shelter use by those aged under 16 in 2020 and 2021

Shelter use by those aged under 16 in 2020 and 2021



Shelter use by those aged between 16 and 24 in 2020 and 2021



Shelter use by those aged between 25 and 44 in 2020 and 2021



Shelter use by those aged between 45 and 64 in 2020 and 2021



Shelter use by those aged older than 65 in 2020 and 2021

Talk way more about it.

# 3 Model

$$Pr(\theta|y) = \frac{Pr(y|\theta)Pr(\theta)}{Pr(y)} \tag{1}$$

Equation (1) seems useful, eh?

Here's a dumb example of how to use some references: In paper we run our analysis in R (R Core Team 2020). We also use the `tidyverse` which was written by (**thereferencecanbewhatever?**) If we were interested in baseball data then (**citeLahman?**) could be useful.

We can use maths by including latex between dollar signs, for instance $\theta$.

# 4 Results

# 5 Discussion

# 6 Limitations

## 6.1 First discussion point

If my paper were 10 pages, then should be be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

## 6.2 Second discussion point

## 6.3 Third discussion point

## 6.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

# Appendix

# A    Additional details

# References

Gelfand, Sharla. 2020. *Opendatatoronto: Access the City of Toronto Open Data Portal.*

R Core Team. 2020. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.