

# Shelter System Flow Data in Toronto Needs Improvement\*

Krupali Bhavsar

February 8 2022

## Abstract

Homelessness is prominent in the city of Toronto with 30 in every 10 000 people experiencing homelessness, with this comes the importance of tracking shelter usage to ensure an adequate means of measuring the progress towards reducing homelessness across the city. Data was obtained from the Toronto Open Data portal and was studied using tables and graphs. The city of Toronto has a long ways to go in their methods of both researching and aiding the homeless community. With better data and sufficient research, adequate plans can be put in place to aid the situation and reduce the population of homeless people in the city of Toronto.

## Contents

<b>Introduction</b>	<b>2</b>
<b>Data</b>	<b>2</b>
Data Source . . . . .	2
Data Characteristics . . . . .	2
Data Cleaning . . . . .	3
<b>Discussion</b>	<b>4</b>
Success of Moving to Permanent Housing . . . . .	4
Shelter Flow Based on Month of the Year . . . . .	5
Shelter Flow Based on Demographic Breakdown . . . . .	6
<b>Conclusion</b>	<b>7</b>
<b>References</b>	<b>9</b>

---

\*Code and data are available at: <https://github.com/krupalibhavsar/sta304-paper-1>

# Introduction

Homelessness has been a concern in the city of Toronto for many years. Recent statistics tell us that for every 10,000 people in the city of Toronto, 30 are homeless.(About Toronto homelessness, 2021). However, with something as prominent as homelessness there is little need for statistics to explain to the larger population how grave the problem is. A quick stroll through central downtown tells us enough to know that the problem of homelessness in the City of Toronto should be researched, studied and resolved. While the process will inevitably be long and hard it would be a disservice to the community not to take the step to aid the underprivileged. Using data about homelessness, it is the obligation of the privileged to analyse the data and put the results to use to find a solution for the problem at hand.

Data about Toronto shelter system flow was obtained. As homelessness in the city of Toronto is such a broad topic I found it valuable to break down my research purpose into three different aspects ensuring that an overload of information and interpretation did not hinder the adequacy of the exploratory analysis. The data was analysed thoroughly and once I understood the content of the data set I began cleaning. I then created multiple tables and graphs comparing different aspects of the data to come to a conclusion about my three focus points, shelter flow based on demographic breakdown, shelter flow based on month of the year and overall success of moving to permanent housing having used the shelter system. Results were obtained for all three topics and additional research served as further evidence to support the given results. However, as the analysis progressed it became clearer that limitations in the data set restricted the validity of the results.

Overall, the results obtained from the analysis were an adequate starting point for measuring the success and flow of Toronto homeless shelters. Looking into the shelter system flow based on the month of the year, we were able to identify some effects that the global pandemic had on the homeless population. With study of flow based on demographic group we were able to identify the lack of detail in the data set as concrete comparisons were not able to be made based on the information from the data set alone. Additionally, measuring the success of moving to permanent housing lead us to the discovery of the Toronto Homelessness Solutions Service Plan, allowing us to delve deeper into the topic of non-recurring homelessness. Again, due to the many existing limitations of the data identified through the analysis, one major suggestion points for more detailed research needing to be pursued. With a more detailed data set about the homeless population in the city a better analysis can be performed. As a result adequate plans can be put in place to aid the situation and reduce the population of homeless people in the city of Toronto.

## Data

### Data Source

Data from the Toronto Open Data portal was obtained to gain a better understanding of shelter flow based on demographic breakdown, shelter flow based on month of the year and overall success of moving to permanent housing having used the shelter system (Gelfand 2020). The data, published by Shelter, Support & Housing Administration, was concerned with the Toronto Shelter System Flow. Information about homeless people entering and leaving the shelter system is collected monthly from the Shelter System Flow data.

### Data Characteristics

The raw data consists of 180 observations and 18 variables. Monthly data is broken down into seven different groups, all population, chronic, refugees, families, youth, single adult and non-refugees. For each month in 2020 and 2021, from January to December, data based on Shelter flow was record for each of these groups. Using R (R Core Team 2020), tidyverse (Wickham et al. 2019), knitr (Xie, 2021), dplyr (Wickham et al. 2021), tidyr (Wickham 2021) and patchwork (Pedersen 2020), I cleaned and extracted the necessary data to begin my exploratory analysis. Reviewing the data and keeping my research purposes in mind I was able to reduce the variable numbers from 18 to 7.

## Data Cleaning

Table 1: Shelter Usage for January 2020

date (month, year)	population group	returned from housing	returned to shelter	moved to housing	no recent shelter use	percentage
Jan-20	All Population	64	487	567	845	100.0
Jan-20	Chronic	12	47	225	133	35.0
Jan-20	Refugees	15	25	392	107	29.7
Jan-20	Families	15	10	380	75	27.3
Jan-20	Youth	8	46	27	96	10.0
Jan-20	Single Adult	41	431	160	674	62.8
Jan-20	Non-refugees	49	462	175	738	70.3

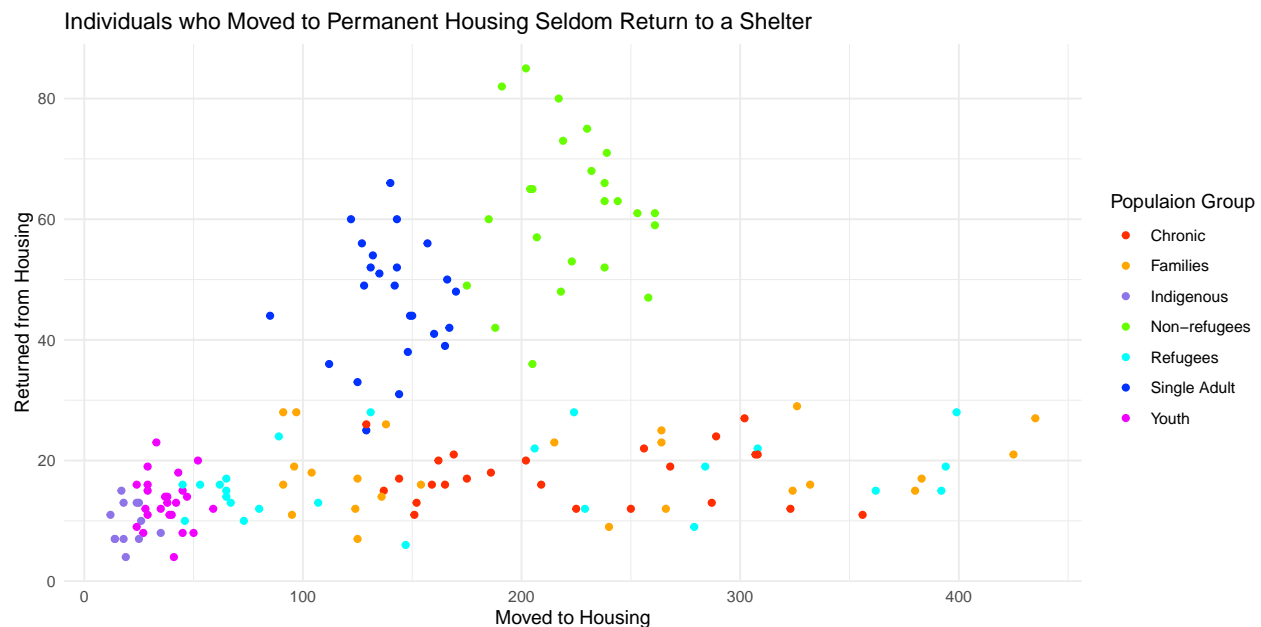
Table 1 illustrates the first 7 rows of the dataset with the columns being the only including the necessary variables, more specifically it depicts the data collected for the month of January in the year 2020. The table gives me an idea of how the data is organized and what variables I can use for the purpose of the investigation. The first column of the data set is the month and year for which the data was collected. Each row in the data set corresponded to a certain month and year and this date represented the time period for which the data in each row was being published. The second variable is population group and each row in this column corresponds with a different group; all population, chronic, refugees, families, youth, single adult and non-refugees. It is important to note that one demographic group is missing from the table, Indigenous, as this group was added in 2021. Although missing from the table, I will be taking this group into consideration. So far the first two columns tell us two important features of the observations. When it was collected for and what demographic group it was concerning.

Moving on we have the numeric variables describing number of individuals associated with a specific action. In the third column we have “returned from housing” representing individuals who were previously recorded as moved to permanent housing and returned to an overnight service in the reporting month. Next there is “Returned to Shelter,” number of individuals who returned to a shelter after three months or longer of not using one. “Moved to housing” are number of people who were previously using the shelter and have now moved to permanent housing. “No recent shelter use” is the amount of people who were using the shelter system previously but have not used it in the past three months. Lastly, “Percentage” refers to the percentage of each group calculated from all the population.

Other variables included in the data are pertaining to age and gender however in the spirit of answering the proposed research questions and insuring that an overload of information and interpretation did not hinder the adequacy of the exploratory analysis, these variables were removed.

# Discussion

## Success of Moving to Permanent Housing



To understand the success of moving to permanent housing I decided to create a scatter plot with “moved to housing” on the y-axis and “returned to housing” on the x-axis. Before delving into an analysis it is important to note that “All population” was removed from the graph as I believe it was more valuable to view the information grouped in their different demographic groups. Therefore, “all population” which represents individuals without considering demographic breakdown, is not needed in Figure 1.

Each dot on the graph represents the number of individuals in a specific demographic group who have Moved to Housing and Returned from Housing in a particular month. Of course as the number of people who have moved to permanent housing increases the number of people who return to the shelter is also bound to increase however I am interested in the difference between the two numbers. Overall, this figure illustrates that for every group, the number of people who have moved into permanent housing is significantly greater than the number of people who have returned to a shelter after having moved to permanent housing. From this graph alone I am lead to believe that in general the homeless population in the city of Toronto has great success transitioning from shelter usage to moving into permanent housing. It seems that a very few portion of people return to a shelter once they have moved to permanent housing.

Although from the graph alone my analysis deems that the homeless population in the city of Toronto has great success transitioning from shelter usage to moving into permanent housing, our data has limitations that could influence the results. The data only reflect those who have returned to overnight service leaving out those who use other homelessness services and those who sleep exclusively outdoors. This implies that permanent housing could have failed for people, however because they did not use an over night service they would not be accounted for in the returned from housing column.

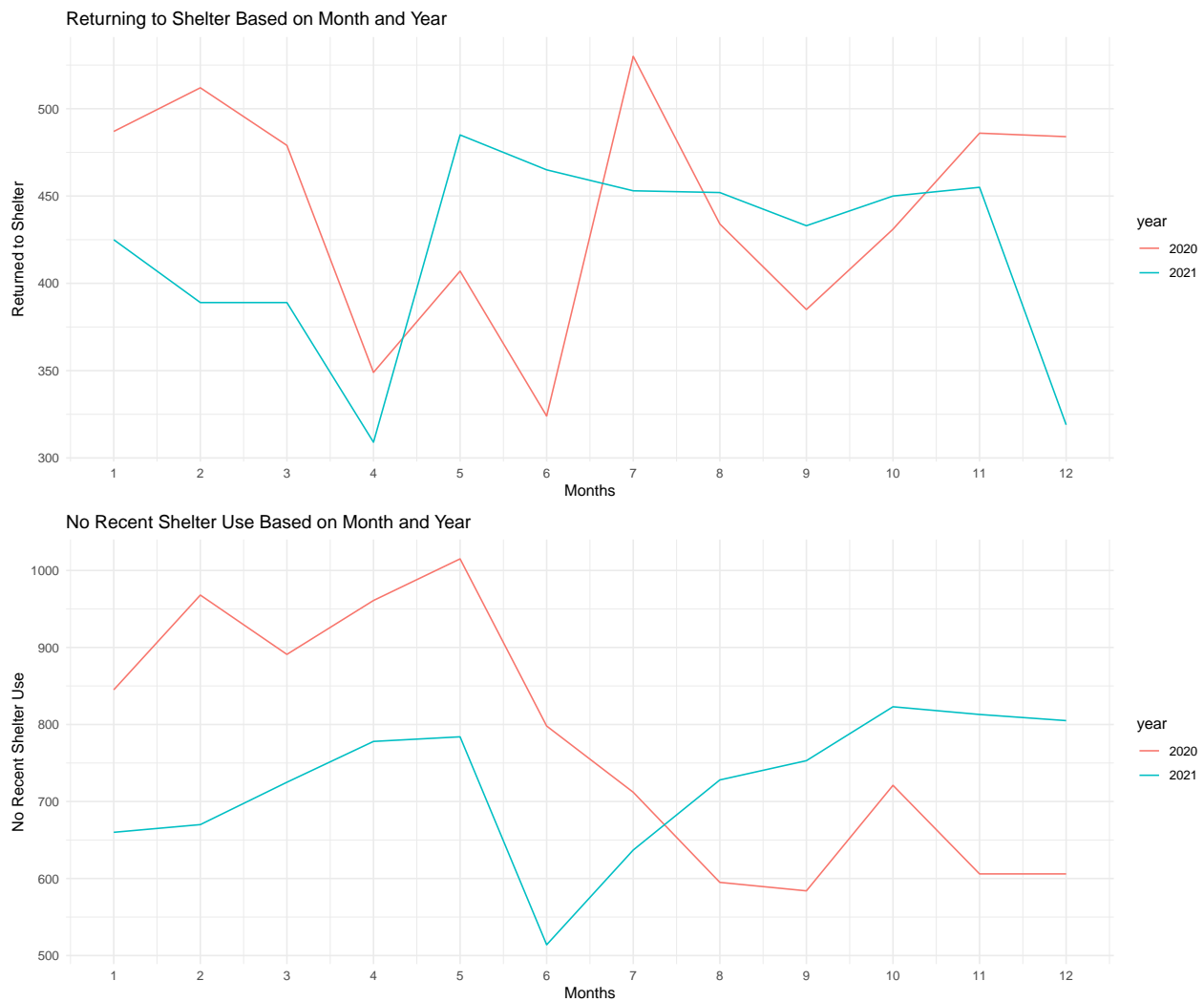
With the existing limitations I decided to research more about the success of permanent housing in Toronto. According to the Shelter, Support and Housing Administration, the city has implemented a Homelessness Solutions Service Plan, (Bédard, 2021). The plan was implemented in 2019 and the measurements of progress indicate that the success of individuals moving to permanent housing was depicted in the three month with an average of only 96 individuals returning to homelessness. The outcome of this success solidified the non-recurring nature of homelessness in the city, (Bédard, 2021). This study solidified my initial analysis and although our data came with limitations the report from the city itself shows that Figure 1 is reflecting the truth. People who move to permanent housing after having used an overnight service are less likely to return

to a shelter.

I see discrepancies in the different population groups however considering that the proportion of the population being a part of a certain group varies I do not believe that any core assumptions would be valid when considering the differences or similarities of returning to a shelter from permanent housing based on the different demographic groups. I am still interested in shelter flow based on demographic breakdown thus it will be explored later in the paper.

## Shelter Flow Based on Month of the Year

To analyze shelter flow based on the month of the year I had to add and remove certain columns and observations. To begin all observations under population group that was not equal to “All Population” was removed. I was not concerned with demographic breakdown therefore keeping the demographic breakdown would have been counter productive. Additionally, I added a column that numerically represents the months from 1 to 12 representing January to December. The second column I added was “year” representing the year in which the data was collected for.



I expected to see more consistency with the two years in Figure 2 where “return to shelter” and “no recent shelter use” increase and decrease at around the same month each year, staying consistent throughout the year. However, that is not always reflected in the graphs above. In April of both years the returned to shelter numbers dropped significantly, this could be due to weather as April is when it get warmer, although there is

no way to tell using the data alone. However, other than that there are not many instances where the two lines on the graph mirrored each other.

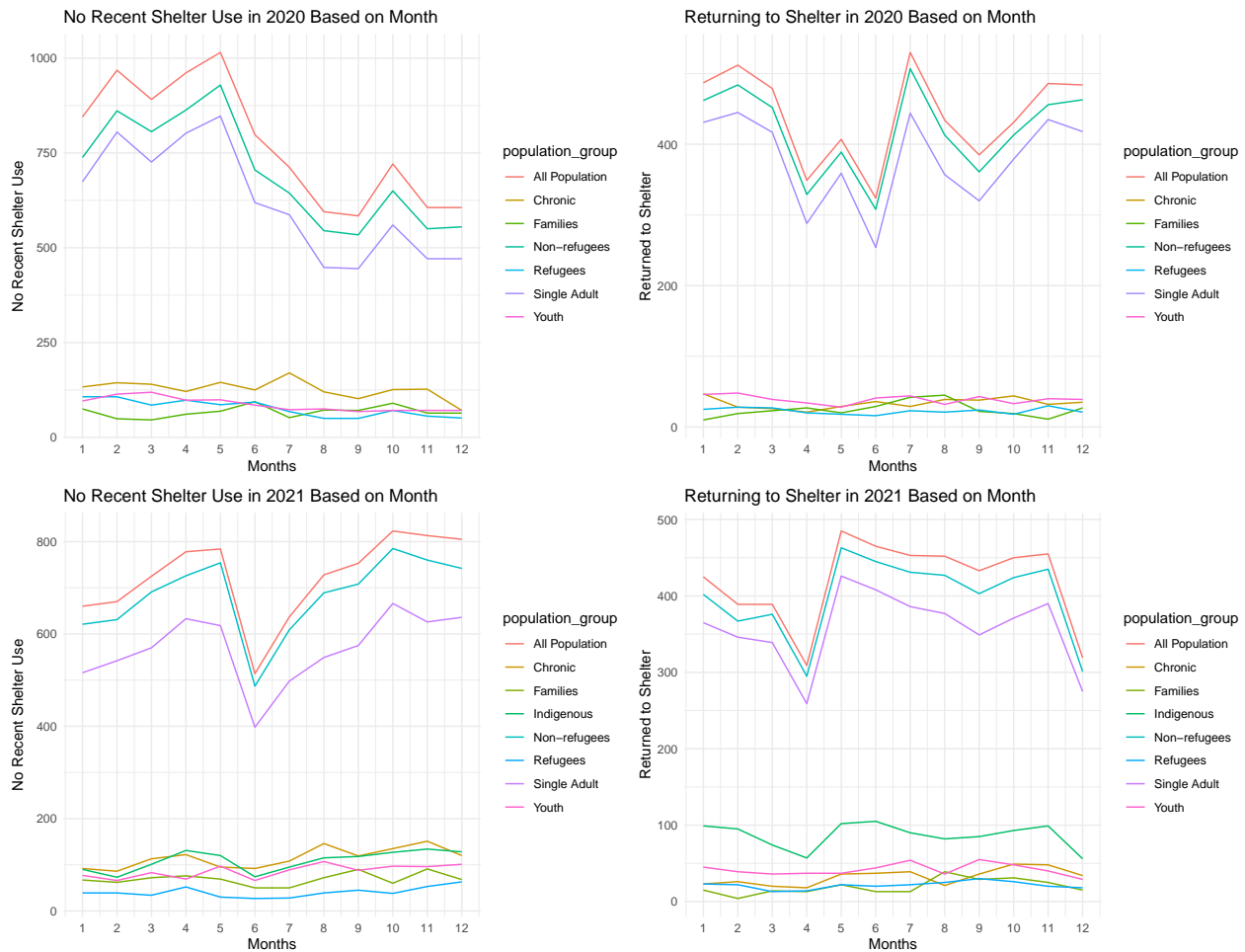
Looking at only the lines for 2020 I noticed something fairly interesting. “Return to Shelter” was the lowest in April May and June meaning that few individuals returned to the overnight service after not using it for a couple months. Additionally, in those same three months in 2020 “no recent shelter use” was high compared to the rest of the months meaning many individuals who were using the shelter system previously were no longer using it. This could be due to the Covid-19 pandemic. In March 2020 the city of Toronto went into strict lock down and it only seems logical to assume that this immensely affected the flow in the shelter system. This sentiment is also reflected in the graphs shown above with shelter use decreasing during months of strict lock down and spiking right after.

The results being influenced by the pandemic can be confirmed upon reading an article written by Kate McGillivray that followed the affects of the pandemic on individuals suffering from homelessness. It was shown that many homeless individuals who had previously used shelters where opting to sleeping outside in tents due to the high risk of contracting Covid-19 in homeless shelters at the peak of the pandemic, (McGillivray, 2021). As was mentioned earlier in this paper, the data Toronto Shelter System Flow data does not include information about individuals sleeping outdoors and thus it further explains the trends we see in the graph for “return to shelter” and “no recent shelter use” for 2020.

It is important to note that with the limitations of the data, the flow in out of shelters based on month cannot be shown accurately. Although the irregularities of the lines on the graph are surprising at first, when we delve into the limitations it is easier to understand why there is not a visible pattern. The data only reflects individuals who have used an overnight service and those who sleep outside are not accounted for. There also exists over night shelters that do not use Shelter Management Information System and others that are funded by other levels of government as a result these shelters are not accounted for in the data. Therefore it seems that with the information I have right now it is not possible to completely understand and make assumptions about the shelter usage flow in the city of Toronto in the different months.

## **Shelter Flow Based on Demographic Breakdown**

To evaluate the shelter flow based on demographic breakdown it would be useful to study both “no recent shelter use” and “returned to shelter” separated by demographic breakdown.



The graphs in Figure 3 depict an interesting pattern concerning “all population,” “non-refugees” and “single adults. I see that all three of these population groups follow a very similar pattern. Out of all the demographic groups non-refugees and single adults is higher up on all of the graphs in comparison to the other groups. This tells me a few things about shelter flow based on demographic breakdown. Considering these characteristics it is clear that non-refugees and single adults make up a larger portion of the total population being considered meaning that more individuals using the overnight shelter systems are a part of a family and/or a single adult compared to any other demographic group being considered in the data. This also means that most individuals considered in this data who identify as a single adult also identify as being a part of a family.

Seeing that the line on the graph representing individuals who suffer from chronic homelessness is not closer to the pattern followed by all population was surprising. This tells me that people who suffer from chronic homelessness are not a large population of those using the overnight shelter services included in this data. The limitations of the data restrict me from finding out more about why chronic homelessness is not a common trait of the people represented in the data. Because we are only able to see the flow of individuals using certain overnight services it is not possible to find where a larger population of those suffering from chronic homelessness reside. Thus, I can make no further assumptions about the result.

## Conclusion

Overall, the results obtained from the analysis were an adequate starting point for measuring the success and flow of Toronto homeless shelters. We were able to identify some of the adverse effects of the global

pandemic on the homeless population based on a study of shelter system flow by month. By examining the flows according to demographic groups, we were able to identify the lack of detail in the data set as concrete comparisons could not be made based on the information in the data. Additionally, measuring the success of moving to permanent housing led us to the discovery of the Toronto Homelessness Solutions Service Plan, allowing us to delve deeper into the topic of non-recurring homelessness. One of the major suggestions points to a need for further investigation due to the many limitations inherent with the data identified through analysis. With a more detailed data set about the homeless population in the city a better analysis can be performed. As a result adequate plans can be put in place to aid the situation and reduce the population of homeless people in the city of Toronto.



## References

- “About Toronto Homelessness.” Homes First, February 9, 2021. <https://homesfirst.on.ca/about-toronto-homelessness>.
- Bédard, Mary-Anne. “Homelessness Solutions Service Plan - Toronto.ca.” Homelessness Solutions Service Plan . Accessed February 6, 2022. <https://www.toronto.ca/legdocs/mmis/2021/ec/bgrd/backgroundfile-171730.pdf>.
- Kate McGillivray. “Ontario’s Homeless 5 Times More Likely to Die of Covid-19, Study Finds | CBC News.” CBCnews. CBC/Radio Canada, January 12, 2021. <https://www.cbc.ca/news/canada/toronto/ontario-s-homeless-5-times-more-likely-to-die-of-covid-19-study-finds-1.5869024>.
- Gelfand, Sharla. 2020. *Opendatatoronto: Access the City of Toronto Open Data Portal*. <https://CRAN.R-project.org/package=opendatatoronto>.
- Pedersen, Thomas Lin. 2020. *Patchwork: The Composer of Plots*. <https://CRAN.R-project.org/package=patchwork>.
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Wickham, Hadley. 2021. *Tidyr: Tidy Messy Data*. <https://CRAN.R-project.org/package=tidyr>.
- 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>.