

Substance Treatment Center Frequencies in Toronto Neighbourhoods*

Sevnur Kulak

07 February 2022

Abstract

While substance usage has becoming a significant problem, this report analyzes the frequency and distribution of treatment centers in different neighbourhoods in the city of Toronto. Then it compares treatment center locations with specific neighbourhood characteristics using two datasets from Toronto Open Data Portal. Though the dataset used in this report might have a small sample size, therefore leading to have a statistical bias, further investigation on updated and different neighbourhood characteristics will be needed to explain any potential pattern on the substance treatment center distributions.

1 Introduction

Substance usage has becoming a dangerous problem among the nations. According to the National Survey on Drug Use and Health, around 40.3 million Americans were struggling on substance usage in 2020. Similarly, approximately 6 million Canadians struggle with substance usage and addiction according to Canadian Mental Health Association. When the problem controls over a big portion of the population, the negative externalities on other people due to substance users' unconscious behaviours as well as the impact on the country's economy, health, crime rates, and legal issues -causing approximately 600billion every year - are unavoidable. Therefore, in order to prevent an increase on the substance usage or even to try decreasing it among the people, treatment centers and treatment availability are extremely important. According to Addiction Center statistics, only 10% of addicted Americans receive treatment. So, it is important to further investigate and assess the availability of substance treatment.

In this paper, I will use substance use treatment locations in Toronto to further analyze whether the availability of treatment centers correlate with any other factors or not. For this analysis, I will pull treatment centers' and neighbour characteristics datasets from Toronto Open Data portal to look at their frequency and distribution depending on the neighborhoods. I will then look into neighbourhood characteristics and investigate more on potential factors such as population, immigration percentage, average income, education degree, and unemployment rate, which might affect the availability of treatment centers within the region. I will process my analysis in R, using mainly tidyverse and dplyr packages and will generate the output on R Markdown. The paper will proceed Data and References sections that I will be talking about data sources, data characteristics, data analysis, and further analysis on other potential contributing factors.

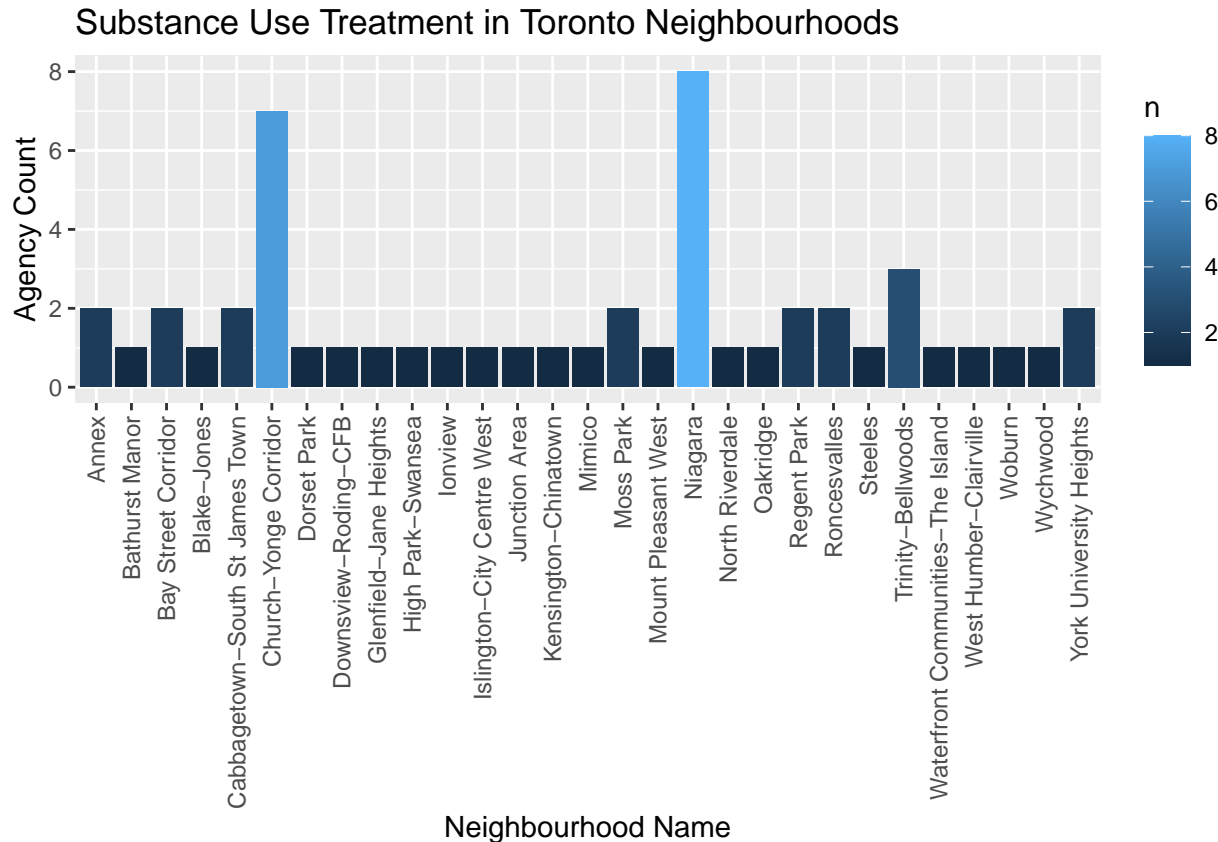
2 Data

To investigate the accessibility and frequency of treatment centers in Toronto, I used Wellbeing Youth – Substance Use Treatment data from the Toronto Open Data Portal. The data was last updated on March 16, 2020 and published by Social Development, Finance & Administration. The raw data includes 56 different agent and agent-related information out of 29 neighbourhoods in Toronto. Agent-related information consists of agency name and address, neighbourhood name, phone number, e-mail, website, eligibility, description and application for the treatment centers, languages used, accessibility, working hours, legal status, geographical

*Code and data are available at: https://github.com/sevnurkulak/Paper_1

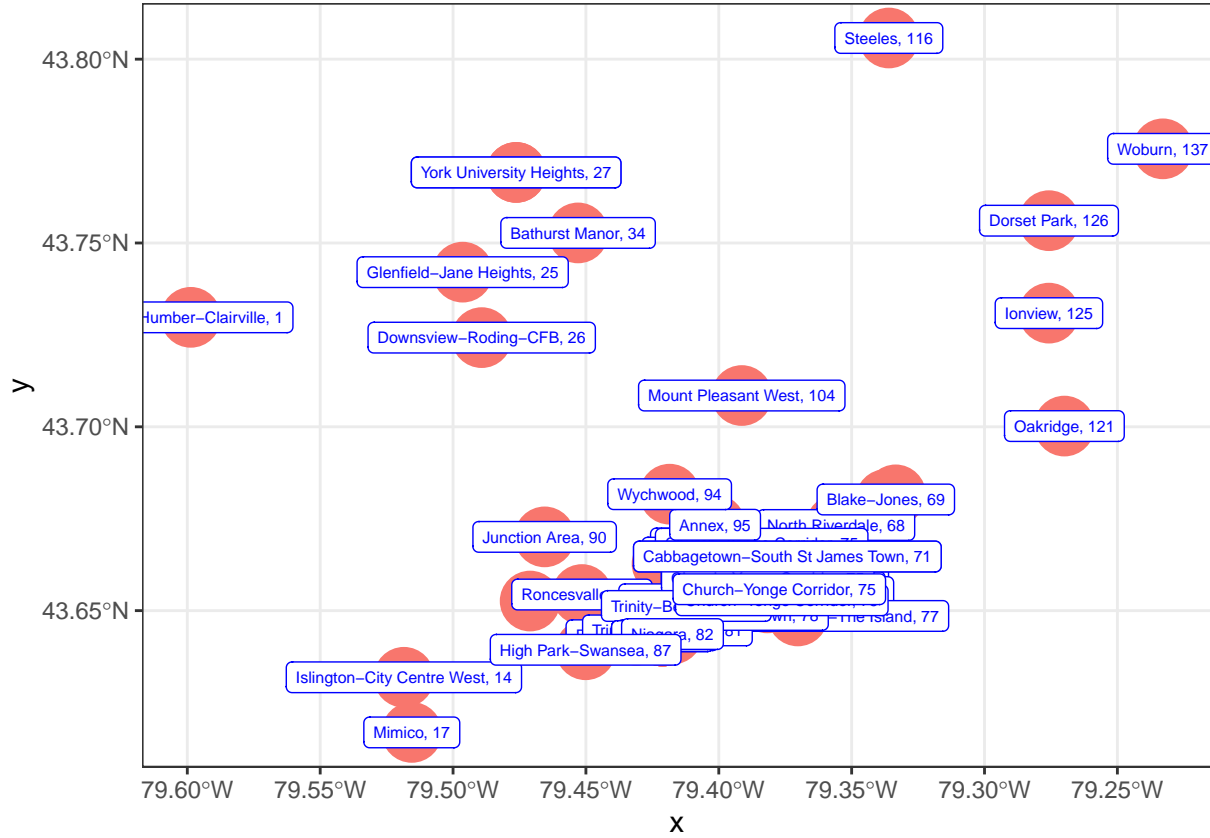
location, postal code, and municipality. I extracted, cleaned, analyzed, and graphed my data using R, tidyverse, dplyr, janitor, tidyr, knitr, ggplot2, sf.

Before selecting the columns I want to use, I cleaned my data and removed any “n/a” parts. In order to see the frequency of treatment centers in different neighbourhoods, I only used neighbourhood column to see the total number of treatment centers counting them with the count function according to the neighbourhood they belong to. In order to continue my analysis neatly, I also extracted the numerical part attached to each neighbourhood name through select function. That way I only included the name of each neighbourhood. In order to better see the distribution, I graphed my counted neighbourhood data using ggplot.



Although most of the neighbourhoods have 1-2 treatment centers on average, the neighbourhoods Church - Yonge Corridor, Niagara, and Trinity-Bellwoods have significantly higher number of treatment centers. In order to visualize this difference I also included a data table of the counts in my analysis.

Before moving to further investigation on the neighbourhood characteristics which might be explanatory on such difference, I also used ggplot2 to draw a map of Torontonians neighbourhoods with treatment centers and see if there are any clusters of treatment centers. To do that, I first extracted the geographical location column of each neighbourhoods in the raw data. I then labeled the locations of each treatment center in Toronto on a map including the neighbourhood names to see the distribution of the centers in the city.



As it can be seen from the map, most of the treatment centers are clustered in downtown Toronto and gets significantly less frequent when moved out of the city center. Although it might be because of factors such as higher population, therefore higher demand, I also wanted to compare these findings with other neighbourhood characteristics.

To do that, I used Neighbourhood Profiles data from the Toronto Open Data Portal. The data was last updated on October 7, 2019 and published by Social Development, Finance & Administration. The raw data includes detailed information on all 140 neighbourhoods of Toronto. Among over 100 variables, in this paper I focus mainly on population and components of population such as working-age percentage and senior percentage, immigration percentage, percentage of people who obtained bachelor's degree education, average income of a household, and unemployment rate within the neighbourhood. The main reason for this focus within the dataset is to see if different components of a population will be correlated or explanatory of the cluster of treatment centers in the city. Additionally, because income and level of education might be other potential factors on willing to receive treatment on substance usage, I include these variables in my analysis as well. I extracted, cleaned, and analyzed the data again using R, tidyverse, dplyr, janitor, tidyr, knitr, and ggplot2. After cleaning my data, I only included the neighbourhoods that has treatment centers and neighbourhood specific variables that I talked above. In order to make it visually more appealing, I changed rows to columns and columns to rows. So that the neighbourhood names will be rows and neighbourhood specific variables will be columns. Then finally, I composed my data table for the characteristics of each neighbourhood.

Considering the cluster of the treatment center in downtown Toronto, population can be one of the explanations for it since the total number of population around the cluster-area is higher than the neighbourhoods far from downtown. However, I cannot conclude that population is the reason for individual neighbourhoods to have more treatment centers. Although Niagara, Church-Yonge Corridor and Trinity-Bellwoods have the highest number of treatment centers, other neighbourhoods that have 1-2 treatments centers on average have similar number of population. Therefore, population might not be the main cause for individual neighbourhoods to have more treatment centers but might be one of the reasons for the cluster around

Table 1: Toronto Neighbourhood Characteristics

	Population, 2016	Working Age (25-54 years)
niagara	31,180	23,320
church_yonge_corridor	31,340	18,780
trinity_bellwoods	16,556	9,210
annex	30,526	15,040
bay_street_corridor	25,797	13,065
cabbagetown_south_st_james_town	11,669	5,855
moss_park	20,506	12,530
regent_park	10,803	5,575
roncesvalles	14,974	7,990
york_university_heights	27,593	12,290
bathurst_manor	15,873	6,655
blake_jones	7,727	3,605
dorset_park	25,003	10,485
downsview_rodong_cfb	35,052	15,045
glenfield_jane_heights	30,491	11,590
high_park_swansea	23,925	11,570
ionview	13,641	6,050
islington_city_centre_west	43,965	20,640
junction_area	14,366	7,470
kensington_chinatown	17,945	8,760
mimico_includes_humber_bay_shores	33,964	17,695
mount_pleasant_west	29,658	17,100
north_riverdale	11,916	5,490
oakridge	13,845	6,050
steeles	24,623	9,500
waterfront_communities_the_island	65,913	45,105
west_humber_clairville	33,312	13,845
woburn	53,485	21,945
wychwood	14,349	6,420

Table 2: Toronto Neighbourhood Characteristics

	Pre-retirement (55-64 years)	Seniors (65+ years)
niagara	1,885	1,510
church_yonge_corridor	3,235	3,010
trinity_bellwoods	1,625	2,350
annex	3,480	5,910
bay_street_corridor	1,760	2,420
cabbagetown_south_st_james_town	1,715	2,250
moss_park	2,495	1,900
regent_park	1,100	730
roncesvalles	1,690	1,680
york_university_heights	2,965	3,530
bathurst_manor	2,030	2,940
blake_jones	940	895
dorset_park	3,330	3,780
downsview_rodin_cfb	4,380	5,535
glenfield_jane_heights	3,605	5,005
high_park_swansea	2,870	3,800
ionview	1,720	1,890
islington_city_centre_west	5,400	7,405
junction_area	1,750	1,500
kensington_chinatown	1,795	2,705
mimico_includes_humber_bay_shores	4,620	5,160
mount_pleasant_west	3,030	4,605
north_riverdale	1,645	1,660
oakridge	1,595	1,655
steeles	3,515	5,755
waterfront_communities_the_island	4,680	4,635
west_humber_clairville	3,990	4,980
woburn	6,245	8,010
wychwood	1,595	3,150

Table 3: Toronto Neighbourhood Characteristics

	Immigrants	Total immigrants
niagara	8,995	30,440
church_yonge_corridor	10,935	30,310
trinity_bellwoods	5,810	16,230
annex	8,270	28,640
bay_street_corridor	10,455	25,605
cabbagetown_south_st_james_town	3,770	10,990
moss_park	6,470	18,455
regent_park	5,015	10,675
roncesvalles	4,395	14,875
york_university_heights	15,595	27,565
bathurst_manor	8,020	15,575
blake_jones	2,420	7,685
dorset_park	14,685	24,375
downsview_rodong_cfb	19,515	35,000
glenfield_jane_heights	17,790	30,215
high_park_swansea	6,565	23,345
ionview	7,795	13,610
islington_city_centre_west	19,590	43,220
junction_area	4,550	14,075
kensington_chinatown	7,615	17,460
mimico_includes_humber_bay_shores	13,425	33,325
mount_pleasant_west	11,295	28,930
north_riverdale	2,925	11,680
oakridge	7,520	13,600
steeles	17,440	24,265
waterfront_communities_the_island	23,820	65,620
west_humber_clairville	19,230	32,890
woburn	30,185	53,000
wychwood	4,345	13,245

Table 4: Toronto Neighbourhood Characteristics

	Bachelor's Degree	Unemployment Rate
niagara	11835	4.5
church_yonge_corridor	10405	8.5
trinity_bellwoods	3515	6.1
annex	9135	6.7
bay_street_corridor	8370	10.2
cabbagetown_south_st_james_town	2925	8
moss_park	5480	6.9
regent_park	2060	9.6
roncesvalles	3435	6.7
york_university_heights	3910	10.7
bathurst_manor	3075	7.2
blake_jones	1550	8.5
dorset_park	3485	9.7
downsview_roding_cfb	3350	7.9
glenfield_jane_heights	1610	11.5
high_park_swansea	6390	6
ionview	1670	8.8
islington_city_centre_west	9970	7.3
junction_area	3250	5.6
kensington_chinatown	4070	8.2
mimico_includes_humber_bay_shores	7560	6.2
mount_pleasant_west	9275	6.7
north_riverdale	3030	5.3
oakridge	1660	14.6
steeles	4320	9.5
waterfront_communities_the_island	27280	5.2
west_humber_clairville	4245	9.6
woburn	7575	10.6
wychwood	2770	5.2

the downtown area. Due to increased population and potentially increased demand, there might be more treatment centers clustered. Additionally, although working-age population is the biggest portion for most of these 29 neighbourhoods, it is important to highlight the top-3 neighbourhoods with the largest number of treatment centers, the working-age population is significantly higher (80.9%, 71.6%, 65.6% respectively) than the average of the whole city of Toronto which is 51.2% according to the data source. Moving onto the immigration rates, although there is not clear relationship between neighbourhoods having more treatment centers and having less immigration rate, when the number of treatment center that each neighbourhood has decreased, the immigration rate increases compared to the ones with larger number of treatment centers. Similarly, neighbourhoods with less treatment centers have lower bachelor's degree percentage, whereas the ones with larger number of treatment centers have higher rates. There are no clear patterns on neither unemployment rates nor average income in neighbourhoods with them having more/fewer treatment centers. However, downtown Toronto includes most of the business centers. Therefore, different than any potential relationship between individual neighbourhoods and income, having lots of business centers might be an explanation for the cluster of treatment centers in the downtown Toronto. It might be because business centers want to protect their business from substance users so that have exclusive deals with either provincial government or treatment centers. However, further investigation and research on that will answer to that question. In addition, the question of willing to get the treatment might be another research topic for further investigation. The underlying reason for having a cluster within a city or individual neighbourhoods having more treatment centers than the average might be because people in those places are more willing to receive a substance treatment therefore have a high demand for them which will cause to have more treatment centers in specific parts of the city.

When looking both of the datasets, I extracted them from a reliable resource of Toronto Open Data Catalogue. Both of the datasets include very few missing values and have detailed components within, which can be used for other analysis. However, there are 2 major challenges with the datasets. First one is the sample size of the data is small which might lead to have a statistical bias in my analysis. For instance, we only have data for 29 neighbourhoods out of 140 neighbourhoods for the first dataset. Among 29 of these neighbourhoods there are 3 of them which has larger number of treatment centers. Such issue might cause sampling bias due to small sample size, therefore wrong assumptions and finding on the population, for instance treatment center distribution in Canada overall. Further investigation on substance treatment center frequencies as well as any potential correlation with neighbourhood characteristics might be needed from other provinces. Additionally, the first dataset is more recent-2020, whereas the second one is from-2016. Not only the difference might create a mismatch between the datasets and lead to a wrong conclusion, the true Neighbourhood characteristics might be different in 2020 since it is very likely those datapoints (percentages, rates, population, etc.) to change quickly. For further investigation, it will be better to use an updated dataset for neighbourhood characteristics to have a better comparison.

3 References