No Evidence of Discrimation Accessing COVID-19 Immunization Clinics Across the City of Toronto*

Christina Wei

January 23, 2023

COVID-19 immunization clinic and ward profile datasets were used in the analysis of vaccine access across the City of Toronto. Based on analysis of number of clinics and household income, no systematic discrimation was discovered in this analysis. There is a slightly positive correlation between number of clinics and population, supporting the hypothesis that there is equitable access to COVID-19 vaccines across the city. Further analysis should be performed to analyze utilization data alongside COVID-19 clinic access.

Table of contents

---

*Code and data supporting this analysis is available at: https://github.com/christina-wei/INF3014-1-Covid-Clinics.git

1 Introduction

The new coronavirus (COVID-19) outbreak was deemed a global pandemic by the World Health Organization (WHO) on March 11, 2020 ("WHO Director-General's Opening Remarks at the Media Briefing on COVID-19 - 11 March 2020" 2020). The world has gone into a global lockdown in order to control the spread of the disease. As the vaccine became available in early 2021, Canada began preparations to roll out COVID-19 vaccines. Specifically in the City of Toronto, the government opened various immunization clinics across the city, as well as partnered with hospitals and pharmacies to provide access to vaccines. However, some criticized the government for inequitable access to vaccines, discriminating against vulnerable populations ("Ontario Court Tosses Lawsuit Claiming Province's COVID-19 Vaccination Rollout Was Discriminatory" 2021). As the vaccine rollout progresses, there are concerns that high-risk individuals in Ontario are lagging in vaccinations (Ogilvie 2022).

One of the factors that could contribute to discrimination of COVID-19 vaccines is access to immunization sites to receive the vaccination. Therefore, it is important to analyze if there is equitable distribution of immunization clinics across the city, as well as potential bias that may exist in the system. In this paper, I will analyze a snapshot of COVID-19 immunization clinic data as of January 22, 2023 alongside 2016 Census data to investigate whether household income or population have impacts on the availability of COVID-19 clinic sites.

In the Data section, I will discuss the different data sources used for analysis, as well as the data cleaning process applied to these datasets. The following Results section will present my data analysis in detail, presenting trends and correlations discovered along the way. The Discussion section shares more insights into the data findings, as well as discussing measurement errors and limitations of this analysis. Finally the paper wraps up with Conclusion summarizing the important discoveries from this paper.

2 Data

Data used in this paper are retrieved from Open Data Toronto Portal through the library `opendatatoronto` (Gelfand 2022). Three different data sources were leveraged to analyze COVID-19 immunization clinics across Toronto: `COVID-19 Immunization Clinics` (Data 2023), `Ward Profiles, 2018 (25-Ward Model)` (Data 2021), and `City Wards` (Data 2022). Data was cleaned and analyzed using the open source statistically programming language R

Table 1: Sample of cleaned Covid-19 clinic data

location_id	location_name	location_type	geometry	ward_name
153	Metro Toronto Convention Centre	City	c(-79.386829859627, 43.643971128415)	Spadina-Fort York
157	Scarborough Town Centre	City	c(-79.258930615442, 43.776990231769)	Scarborough Centre
158	Cloverdale Mall	City	c(-79.556501959627, 43.632603612174)	Etobicoke-Lakeshore
650	Woodbine Mall	City	c(-79.599279647844, 43.719178000313)	Etobicoke North
150	Mitchell Field Community Centre	City	c(-79.408229001951, 43.774637558851)	Willowdale

(R Core Team 2022), using functionalities from `tidyverse` (Wickham et al. 2019), `ggplot2` (Wickham 2016), `dplyr` (Wickham et al. 2022), `readr` (Wickham, Hester, and Bryan 2022), `tibble` (Müller and Wickham 2022), `janitor` (Firke 2021), `kableExtra` (Zhu 2021) and `knitr` (Xie 2014). Details of the data extraction and cleaning processes are discussed in the subsections below.

2.1 COVID-19 Immunization Clinics

This dataset published by Toronto Public Health (Data 2023) captures information about COVID-19 immunization clinics across the City of Toronto and is refreshed on a daily basis. The data used for analysis in this paper is captured as of January 22, 2023. Based on the data features described on the portal, the data should include information on clinic location, date added, contact information, eligibility criteria, availability and appointment booking details. However, through preliminary exploration of data that many data elements are not recorded properly for appointment, availability and eligibility related information. The columns designed to capture these data are not populated. Instead, information are typed in as free form comments in the `info` column, making it difficult to systematically extract details from this field. As a result, only columns with reasonable data are kept in the data cleaning process (see Table 1). Basic data cleaning was applied to the column values to shorten the description of different types of clinics for readabiliy.

2.2 Ward Profiles, 2018 (25-Ward Model)

Ward profiles such as income and population would be interesting factors to analyze alongside COVID-19 immunization clinics. Therefore, the dataset for ward profiles (Data 2021) based on 2016 census data has been included in analysis as well. This dataset is published by City Planning, and was last updated on February 11, 2021. The 25-ward model was chosen over the 47-ward model because it is matching the ward classification in the COVID-19 immunization clinic data. This dataset contains demographic, social and economic information for each ward, such as income and population.

Ward profile data is stored in an Excel file with multiple tabs. The relevant data to be used for this paper's analysis is included the first tab, `2016 Census One Variable`. As such, only data

Table 2: Sample of cleaned ward profile data

ward_code	population	income
1	118040	73374
2	118020	128448
3	129080	105994
4	108805	99784
5	116685	67954

for this tab was downloaded for analysis. Further data cleaning was performed to transpose the data and only keeping information relevant to income and population level for each ward (see Table 2).

However, there is no ward name in the dataset that was downloaded. Based on the Open Data Toronto Portal, there is a CSV file in the package, 25 Ward Names and Numbers, that contains the mapping between ward codes and ward names. Unfortunately this particular resource does not download properly through the opentorontodata library (Gelfand 2022). For the sake of reproducibility, another dataset, City Wards (Data 2022), has been identified to fill in this gap.

2.3 City Wards

The city wards dataset (Data 2022) is published by City Clerk's Office, and was last updated on November 29, 2022. It contains information regarding the codes, names, and boundaries of Toronto's 25 wards. Data elements of ward code and ward name are selected from this dataset. Also the columns have been renamed to ward_code and ward_name to mach the naming conventions in the other datasets.

2.4 Combining Ward Data

Using the merge function, ward profile data now contains information on ward code, ward name, population, and average income. COVID-19 immunization clinic data can be summarized by each ward name and the number of clinics per ward. This data is then injected into the ward profile dataset, with Table 3 as the final result.

Table 3: Sample of ward profile details

ward_name	ward_code	population	income	num_clinics
Etobicoke North	1	118040	73374	11
Etobicoke Centre	2	118020	128448	14
Etobicoke-Lakeshore	3	129080	105994	28
Parkdale-High Park	4	108805	99784	18
York South-Weston	5	116685	67954	19
York Centre	6	104320	86717	19
Humber River-Black Creek	7	108035	65458	9
Eglinton-Lawrence	8	114395	162674	15
Davenport	9	108470	80807	12
Spadina-Fort York	10	115510	103047	27
University-Rosedale	11	104310	170832	28
Toronto-St. Paul's	12	107900	155470	14
Toronto Centre	13	103805	75382	29
Toronto-Danforth	14	106875	101323	17
Don Valley West	15	102510	216158	13
Don Valley East	16	94580	80648	14
Don Valley North	17	110080	87491	15
Willowdale	18	118800	87416	17
Beaches-East York	19	109465	104123	22
Scarborough Southwest	20	110280	78561	12
Scarborough Centre	21	112605	70624	27
Scarborough-Agincourt	22	105540	75768	12
Scarborough North	23	98800	78984	15
Scarborough-Guildwood	24	102390	72289	15
Scarborough-Rouge Park	25	102275	99715	12

3 Results

3.1 COVID-19 Immunization Clinic Statistics

As of January 22, 2023, there are 434 COVID-19 immunization clinics across the City of Toronto. Further breakdown shows that there are 5 clinics that are city operated, 16 clinics that are run by hospitals, and 415 are offered in local pharmacies. On average, there are 17 clinics per ward in Toronto, with standard deviation of 6. The wards with the highest number of clinics are: Toronto Center (Ward 13) with 29 clinics, Etobicoke-Lakeshore (Ward 3) with 28 clinics, and University-Rosedale (Ward 11) with 28 clinics. The wards with the lowest number of clinics are: Humber River-Black Creek (Ward 7) with 9 clinics, Davenport (Ward 9) with 12 clinics, Scarborough Southwest (Ward 20) with 12 clinics, Scarborough-Agincourt (Ward 22) with 12 clinics, and Scarborough-Rouge Park (Ward 25) with 12 clinics.
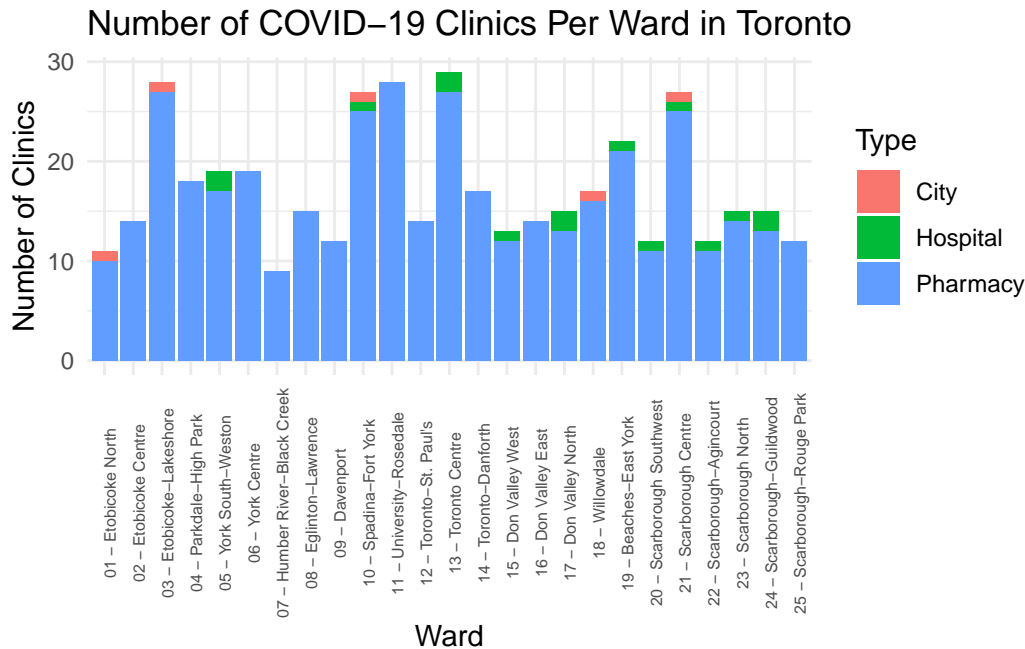

Figure 1: Number of COVID-19 Clinics Per Ward in Toronto

3.2 Ward Profile Based on 2016 Census Data

There are 25 wards in the City of Toronto. Based 2016 census data, the average population per ward is 109,263, with standard deviation of 7,472. The wards with the highest population are: Etobicoke-Lakeshore (Ward 3) at 129,080, Willowdale (Ward 18) at 118,800, and Etobicoke North (Ward 1) at 118,040. The wards with the lowest population are: Don Valley East (Ward

16) at 94,580, Scarborough North (Ward 23) at 98,800, and Scarborough-Rouge Park (Ward 25) at 102,275.

Average household income level per ward is $101,161, with standard deviation of $37,772. The wards with the highest household income are: Don Valley West (Ward 15) at $216,158, University-Rosedale (Ward 11) at $104,310, and Eglinton-Lawrence (Ward 8) at $162,674. The wards with the lowest household income are: Humber River-Black Creek (Ward 7) at $65,458, York South-Weston (Ward 5) at $67,964, and Scarborough Centre (Ward 21) at $70,624.

3.3 Comparing Number of Clinics with Ward Population

As seen in Figure 3, there is a slightly positive correlation between the number of clinics and the average household income in a ward, with correlation coefficient of 0.26. The top three wards with the highest population levels, Etobicoke-Lakeshore (Ward 3), Willowdale (Ward 18) and Etobicoke North (Ward 1) have a wide spread of number of clinics, with 28, 17, 11 clinics respectively. The bottom three wards with lowest population, Don Valley East (Ward 16), Scarborough North (Ward 23) and Scarborough-Rouge Park (Ward 25), have average number of clinics per ward, with 14, 15, and 12 clinics respectively.
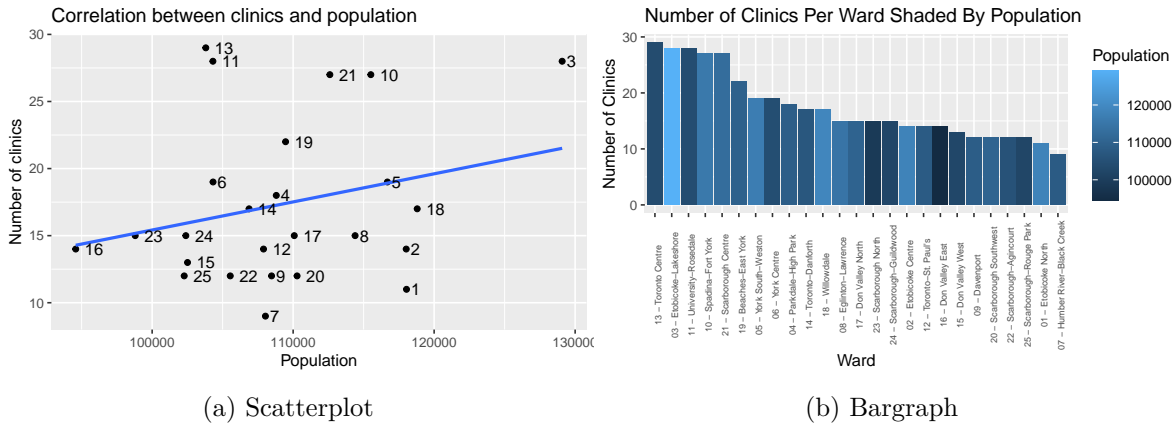

(a) Scatterplot (b) Bargraph

Figure 2: Two visualizations for number of clinics vs. population

3.4 Comparing COVID-19 Clinics with Ward Income

As seen in Figure 3, there is no observable correlation between the number of clinics and the average household income in a ward, with correlation coefficient of 0.06.
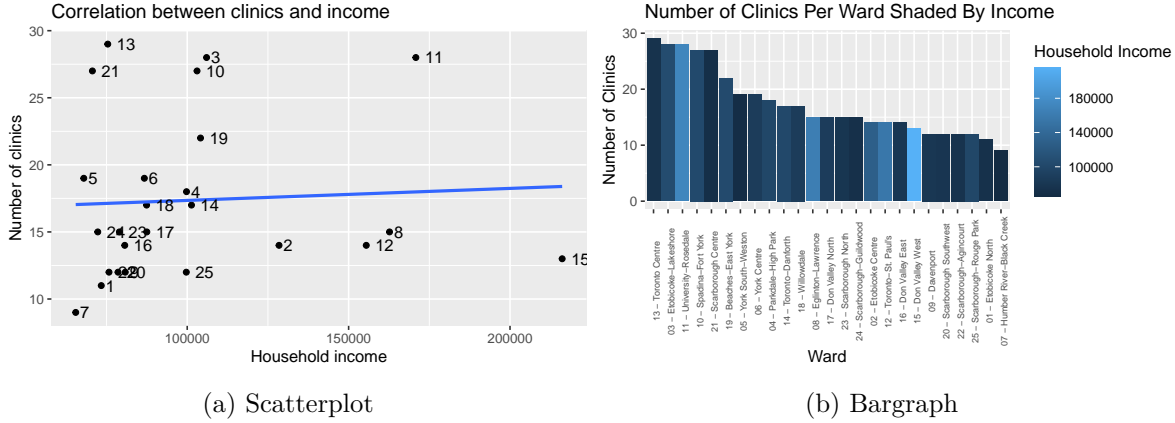
(a) Scatterplot   (b) Bargraph

Figure 3: Two visualizations for number of clinics vs. income

4 Discussion

Overall there is a slightly positive relationship between the number of COVID-19 immunization clinics and population within each ward, which is matching to the expectation that clinic sites are established based on demographic data. There is no correlation found between number of clinics and household income levels, which demonstrated there is no evidence of systematic discrimination against vaccine access based on income alone.

Based on Figure 1, most of the currently active COVID-19 clinics (96%) are operated through local pharmacies like Shoppers Drug Mart or Rexall. As such, this analysis may be a reflection of the existing pharmacy infrastructure across Toronto instead. The results above demonstrated that existing pharmacy network is reasonable equitable across Toronto, therefore it is a good infrastructure to provide access to COVID-19 to the general public.

There are potential measurement errors in the dataset used for this analysis. First, with regards to the measurement of COVID-19 immunization clinics, the data is not inclusive of all vaccine site, such as doctor's office or pop-up sites. There is also the potential for human error for data entry. For example, two immunization sites were wrongfully categorized to "NA". Either the sites did not belong in Toronto or they were not matched to the correct wards. Third, even though the dataset itself is listed as being refreshed daily, the reliability of data sources and frequency of data updates is unknown.

Also, with the usage of ward profile data based on 2016 Census, the data may be outdated it is collected 7 years ago. There are also various errors that may exist in census data ("Guide to the Census of Population, 2021 / Chapter 9 – Data Quality Evaluation" 2021): *coverage errors* when data is incorrectly processed like omissions or misclassifications; *non-response errors* when respondent did not not provide full information; *response errors* when a problem was not understood properly by the respondent; *processing errors* as information is entered

incorrectly; and *sampling errors* where the sample selected for census does not represent the general population.

There are several limitations to this analysis. As the COVID-19 immunization clinic dataset is a snapshot as of a point in time, it is not possible to analyze the changes in clinic availability over time. Also, the dataset only records which locations offer COVID-19 vaccines, but does not properly contain information with regards to opening hours or eligibility, which are also important factors to assess availability. Lastly, it would have been useful to include the utilization rates for each clinic for better assessment of public needs. Unfortunately this dataset is not available on the Open Data Toronto Portal (Gelfand 2022).

5 Conclusion

This paper analyzed COVID-19 immunization clinic data to assess whether there is equitable access to vaccines across City of Toronto. Our data analysis showed that there is a slightly positive correlation between the number of clinics and the population in wards. There is no evidence of systematic discrimination in the distribution of clinic sites as there is no correlation between the number of clinics and household incomes in wards. As most COVID-19 vaccines are offered through major pharmacy chains, this analysis can also be extended to indicate that there is a reasonably equitable access to pharmacies across the city. Future analysis with newer census data, clinic distribution over time, as well as including clinic utilization rate in analysis will be useful to continue investigate into the access to COVID-19 vaccines within the City of Toronto.

Reference

Data, Toronto Open. 2021. "Ward Profiles, 2018 (25-Ward Model)." https://open.toronto.ca/dataset/ward-profiles-2018-25-ward-model/.
———. 2022. "CityWards." https://open.toronto.ca/dataset/city-wards/.
———. 2023. "COVID-19 Immunization Clinics." https://open.toronto.ca/dataset/covid-19-immunization-clinics/.
Firke, Sam. 2021. *Janitor: Simple Tools for Examining and Cleaning Dirty Data.* https://CRAN.R-project.org/package=janitor.
Gelfand, Sharla. 2022. *Opendatatoronto: Access the City of Toronto Open Data Portal.* https://CRAN.R-project.org/package=opendatatoronto.
"Guide to the Census of Population, 2021 / Chapter 9 – Data Quality Evaluation." 2021. https://www12.statcan.gc.ca/census-recensement/2021/ref/98-304/2021001/chap9-eng.cfm.
Müller, Kirill, and Hadley Wickham. 2022. *Tibble: Simple Data Frames.* https://CRAN.R-project.org/package=tibble.

Ogilvie, Megan. 2022. "High-Risk Ontarians Lagging on Third Doses of COVID Vaccine." https://www.thestar.com/news/gta/2022/01/21/high-risk-ontarians-lagging-on-third-doses-of-covid-vaccine.html.

"Ontario Court Tosses Lawsuit Claiming Province's COVID-19 Vaccination Rollout Was Discriminatory." 2021. https://www.cbc.ca/news/canada/toronto/covid-vaccine-discrimination-court-challenge-1.6014892.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

"WHO Director-General's Opening Remarks at the Media Briefing on COVID-19 - 11 March 2020." 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.

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 Grolemund, 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. 2022. *Dplyr: A Grammar of Data Manipulation.* https://CRAN.R-project.org/package=dplyr.

Wickham, Hadley, Jim Hester, and Jennifer Bryan. 2022. *Readr: Read Rectangular Text Data.* https://CRAN.R-project.org/package=readr.

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.

Zhu, Hao. 2021. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax.* https://CRAN.R-project.org/package=kableExtra.