Battle of the Neighbourhoods: Recommending neighbourhoods in Toronto to open a Cannabis retail store

1. Introduction

Cannabis has been legal in Canada for medical purposes since 2001. However, on 17th October 2018, the Federal Government of Canada legalized the use of cannabis for recreational use. Since then, a number of retail dispensaries have come up in and around Toronto.
Initially, the Ontario government allotted limited retail licenses through a lottery system which slowed the rollout as compared to other provinces. Only the ones willing to shed a considerable amount were able to get licenses. This has created a shortage of legal stores in Toronto as currently, there are less than 15 legal stores in the entire city of Toronto. The government has since announced and taken efforts to create a more open market and aims to issue 20 licenses every month. (Freeman & Aguilar, 2019)
Even though there are legal retail stores operating in Toronto, 40% consumers still depend on the black market to acquire their goods (Dandapani, 2020). The reason for this as quoted by many is 'due to a lack of store in their neighborhood' as most cannabis stores in Toronto are packed around the Downtown Toronto area.
So, for this Capstone Project, given the government's initiative to make stores more accessible, I thought it would be interesting to identify neighborhoods to open a cannabis dispensary.

1.1 Background

For the purpose of this project, I have formulated a hypothetical scenario to serve as a case study–

'We are a new Cannabis retail company, named XYZ co., and have recently acquired a license to open a retail store in Toronto. We sell a range of products from flowers, terpenes, vapes to glassware and edibles.
As a response to our product offering we have two customer groups according to age. First group consists of customers between the ages of 21 and 40 and the second group is between the ages of 40 and 65. Majority of our sales comes from the first group. Therefore, we are looking for neighborhoods that have a high percentage of population within this age group. Within these age groups, a high number of our customers are single and married couples with no children. In terms of income, we have observed that our customers earn at least $ 30,000 annually.

We want to explore the neighborhoods in Toronto to find the best suited neighborhood to open our new store'

Based on the information stated above and characteristics of a Cannabis store the following criteria was decided in order to identify an 'ideal neighborhood':

Table 1: Selected Criteria and Assigned Weights

	Criteria
1.	Above average population
2.	Above average population density
3.	High percentage of population in the age group of 20-39
4.	Average-High percentage of population in the age group of 40-64
5.	Average income above 30,000$
6.	High percentage of population without children
7.	High percentage of never married population
8.	Low crime rate

1.2 Problem Statement

'To identify the top 10 neighborhoods in Toronto to open a new Cannabis Retail Store for XYZ co. using neighborhood census data and Foursquare API location data.'

1.3 Aim

The aim of this project is to identify and recommend neighborhoods to XYZ co., for opening a new cannabis retail store based on the information provided.
With this in mind, the project seeks to answer the following questions:

a.) How many cannabis retail stores are there in Toronto? What are the demographics of the neighbourhoods wherer these stores are located?
b.) Based on the decided criterion, which neighborhoods are ideal for opening a new retail store?

1.4 Target Audience

This project is an exploratory and prescriptive analytics study in the retail cannabis business in Toronto.
This project is intended for retailers looking to open a new cannabis store as well as for government and planning authorities to identify neighborhoods that are currently under served by existing stores.

2. Data Description

2.1. Data Requirements:

To accomplish this project, the following data will be required:

a.) **Census Data on Neighborhood level:** To map the demographics of neighborhoods in Toronto, Census data on population, population density, number of people in the age groups of 21-39 and 40-64 and average income.
b.) **Crime Rate on Neighborhood level:** Data on different crimes related to property and person to calculate the crime rate in every neighborhood
c.) **Venues in a Neighborhood:** To identify neighborhood clusters, location data on venues.
d.) **Toronto Neighborhood Boundaries:** To visualize geospatial data, neighborhood shapes and boundaries will be required

2.2. Data Collection:

The data required for this project was collected from:

a.) **Neighborhood**

This dataset is available on the Toronto Open Data Portal and is described as - *'Boundaries of Toronto Neighbourhoods.'*
Data Format required: .geojson (for boundaries)

b.) **Neighborhood Profiles**

This dataset is available on the Toronto Open Data Portal and is described as –
'The Neighbourhood Profiles provide a portrait of the demographic, social and economic characteristics of the people and households in each City of Toronto neighbourhood. The data is based on tabulations of 2016 Census of Population data from Statistics Canada.'

c.) **Neighbourhood Crime Rates (Boundary File)**

This dataset is available on the Toronto Police Service - Public Safety Data Portal and is described as -
'Toronto Neighbourhoods Boundary File includes 2014-2018 Crime Data by Neighbourhood. Counts are available for Assault, Auto Theft, Break and Enter, Robbery, Theft Over and Homicide. Data also includes four-year averages and crime rates per 100,000 people by neighbourhood based on 2016 Census Population.'

d.) **Foursquare API**

Foursquare is a technology company that has built a huge location dataset through crowd-sourcing. Fousquare's data is currently used by companies like Uber, Snapchat, Twitter and Apple Maps.
For this project, Foursquare API will be used to collect data on:

i.) Cannabis Retail Stores in Toronto

ii.) Neighborhood Venues (100 locations within a radius of 1km from the coordinates of every neighborhood)

3. Methodology:

3.1. Data Preprocessing:

Once the data was collected, it was filtered, formatted and transformed to suit the needs of this project.

a.) **Neighborhood:** Data on Neighborhood boundaries was in .csv format which was converted into a geopandas dataframe. Geopandas is a python library that makes working with geospatial data easy by converting shape files into pandas dataframe like structure.

	neighborhood	neighborhood_num	latitude	longitude	geometry
0	Wychwood	94	43.676919	-79.425515	POLYGON ((-79.43592 43.68015, -79.43492 43.680...
1	Yonge-Eglinton	100	43.704689	-79.403590	POLYGON ((-79.41096 43.70408, -79.40962 43.704...
2	Yonge-St.Clair	97	43.687859	-79.397871	POLYGON ((-79.39119 43.68108, -79.39141 43.680...
3	York University Heights	27	43.765736	-79.488883	POLYGON ((-79.50529 43.75987, -79.50488 43.759...
4	Yorkdale-Glen Park	31	43.714672	-79.457108	POLYGON ((-79.43969 43.70561, -79.44011 43.705...

Figure 1: Processed dataframe containing neighbourhood coordinates and geometry

As seen above, after cleaning, only neighborhood name, neighborhood number, latitude, longitude and geometry columns were kept from the original data.

b.) **Neighborhood Profiles:**

The original dataset provided by the Toronto Open Data Portal had 2383 rows.

```
There are 2383 rows and 146 columns in the dataframe.
```

	_id	Category	Topic	Data Source	Characteristic	City of Toronto	Agincourt North	Agincourt South-Malvern West	Alderwood	Annex	...	Willowdale West	Willowrid Martingro Richvi
0	1	Neighbourhood Information	Neighbourhood Information	City of Toronto	Neighbourhood Number	NaN	129	128	20	95	...	37	
1	2	Neighbourhood Information	Neighbourhood Information	City of Toronto	TSNS2020 Designation	NaN	No Designation	No Designation	No Designation	No Designation	...	No Designation	Designat
2	3	Population	Population and dwellings	Census Profile 98-316-X2016001	Population, 2016	2,731,571	29,113	23,757	12,054	30,526	...	16,936	22,
3	4	Population	Population and dwellings	Census Profile 98-316-X2016001	Population, 2011	2,615,060	30,279	21,988	11,904	29,177	...	15,004	21,
4	5	Population	Population and dwellings	Census Profile 98-316-X2016001	Population Change 2011-2016	4.50%	-3.90%	8.00%	1.30%	4.60%	...	12.90%	3.8

Figure 2: Neighbourhoods Profiles Dataset from Toronto Data Portal

However, as identified in the data requirements section only the required records were kept and the resulting dataset had 140 rows and 10 columns. The table below summarizes the preprocessing steps taken.

Table 2: Census data processing and transformation summary

Dataframe	**Characteristic**	**Column Name in original dataframe**	**Transformation**	**Column Name in new dataframe**
age_df	Population and Age census records	Neighborhood Name	Transposed entire dataframe to get the neighborhood names as a column	neighborhood_name
		Neighborhood Number	Converted from type string to type int	neighborhood_num
		Population, 2016	Converted to type float	population
		Density per square kilometre	Converted to type float	population_density
		Male and Female (age 0 to 100 and over)	• Aggregated columns to get totals for two age groups: 1. 20-39yrs 2. 40-64yrs • Calculated percentages	Two columns: 1. % 20-34yrs 2. % 40-64yrs

			for each age group	
mar_df	Marital status census records	Maritial status for population age 15 and over	• Converted to type float • Used in calculating percentages	Column dropped
		Never Married	• Converted to type float • Changed values to percentages	% never_married
		Couples without children	• Converted to type float • Changed values to percentages	% couples_without_children
		Couple census families in private households	• Converted to type float • Used in calculating percentages	Column dropped
age_df	Income census records	Total income, Average amount	Converted to type float	average_income

	neighborhood	neighborhood_num	population	population_density	% 20-39yrs	% 40-64yrs	% never_married	% couples_without_children	average_income
0	Agincourt North	129	29113.0	3929.0	0.2485	0.3535	0.2910	0.3720	30414.0
1	Agincourt South-Malvern West	128	23757.0	3034.0	0.2909	0.3458	0.3192	0.3968	31825.0
2	Alderwood	20	12054.0	2435.0	0.2584	0.3800	0.2922	0.4167	47709.0
3	Annex	95	30526.0	10863.0	0.4082	0.2875	0.4283	0.6592	112766.0
4	Banbury-Don Mills	42	27695.0	2775.0	0.2154	0.3522	0.2666	0.4744	67757.0

Figure 3: Transformed Dataframe containing census data

c.) **Neighborhood Crime rates:** The original dataset was filtered to get 'Assault Rate', 'Auto Theft Rate', 'Break and Enter Rate', 'Homicide Rate', 'Robbery Rate', 'Theft Over Rate'. The below table summarizes the columns and weights applied.

Table 3: Weights applied in calculate crime score

Columns selected from original dataframe	Weight Applied	Reason
Assault Rate 2019	0.1	Crimes affecting store owners directly have been given more weight than crimes affecting store visitors
Auto Theft Rate 2019	0.15	
Break and Enter Rate 2019	0.2	
Homicide Rate 2019	0.1	
Robbery Rate 2019	0.2	
Theft Over Rate 2019	0.25	

All features values were multiplied by their assigned weights and crime rate for each neighborhood was calculated by adding the weighted feature values.

	neighborhood	crime_rate
0	Yonge-St.Clair	99.770
1	York University Heights	373.290
2	Lansing-Westgate	153.145
3	Yorkdale-Glen Park	422.870
4	Stonegate-Queensway	125.745
...	...	...
135	Milliken	229.400
136	Pleasant View	91.040
137	Wychwood	137.630
138	Leaside-Bennington	112.030
139	Briar Hill-Belgravia	196.400

Figure 4: Dataframe with calculated crime rate

d.) **Foursquare API:** First, using Foursquare API location data of cannabis stores in Toronto was returned. The results returned had missing stores, duplicate stores and stores that are no longer operational. Using blogTO's article that listed all the legal stores in Toronto and OCS, the result was filtered and then each store's neighborhood was noted by visualizing the locations on the neighborhood map of Toronto.

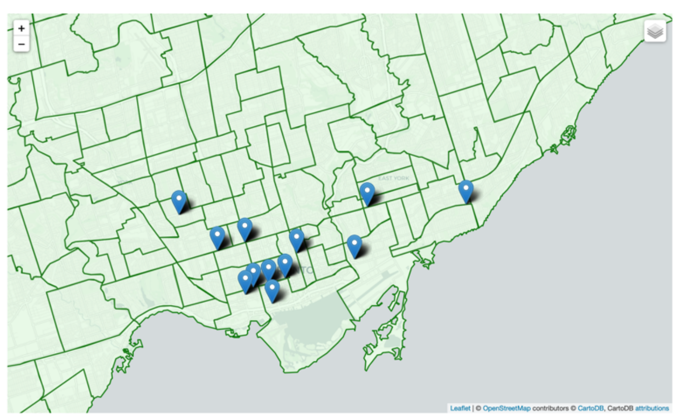


Figure 5: Visualisation to find neighbourhoods of cannabis stores

This resulted in a dataframe with 13 stores and their locations.

	neighborhood	neighborhood_lat	neighborhood_lng	store_name	store_lat	store_lng	store_location
0	Niagara	43.636681	-79.412420	Tokyo Smoke	43.643663	-79.411682	850B Adelaide Street West
1	South Riverdale	43.649292	-79.335651	Tokyo Smoke	43.658130	-79.350150	100 Broadview Avenue
2	The Beaches	43.671050	-79.299601	Session	43.680614	-79.287420	964 Kingston Rd
3	Trinity-Bellwoods	43.650176	-79.415342	Tokyo Smoke	43.643663	-79.411682	850B Adelaide Street West
4	Waterfront Communities-The Island	43.633880	-79.377202	CAFE - Cannabis and Fine Edibles	43.640033	-79.396399	66 Fort York Blvd
5	Waterfront Communities-The Island	43.633880	-79.377202	Nova Cannabis	43.648228	-79.398421	499 queen st w
6	Annex	43.671585	-79.404001	Tokyo Smoke 570 Bloor St W	43.665200	-79.411840	570 Bloor Street West
7	Bay Street Corridor	43.657511	-79.385721	Canna Cabana	43.660725	-79.382891	435B Yonge St
8	Corso Italia-Davenport	43.677661	-79.447469	CAFE St. Clair	43.676424	-79.448885	1321 St Clair Ave W
9	Danforth	43.684025	-79.329819	Canvas Cannabis	43.679447	-79.342964	435B Yonge St
10	Dovercourt-Wallace Emerson-Junction	43.665677	-79.438541	CAFE	43.661699	-79.427517	932 Bloor Street West
11	Kensington-Chinatown	43.653554	-79.397240	Tokyo Smoke 570 Bloor St W	43.665200	-79.411840	570 Bloor Street West
12	Church-Yonge Corridor	43.657511	-79.385721	Tokyo Smoke	43.657108	-79.380483	333 Yonge St.

Figure 6: Dataframe with Cannabis retail store locations

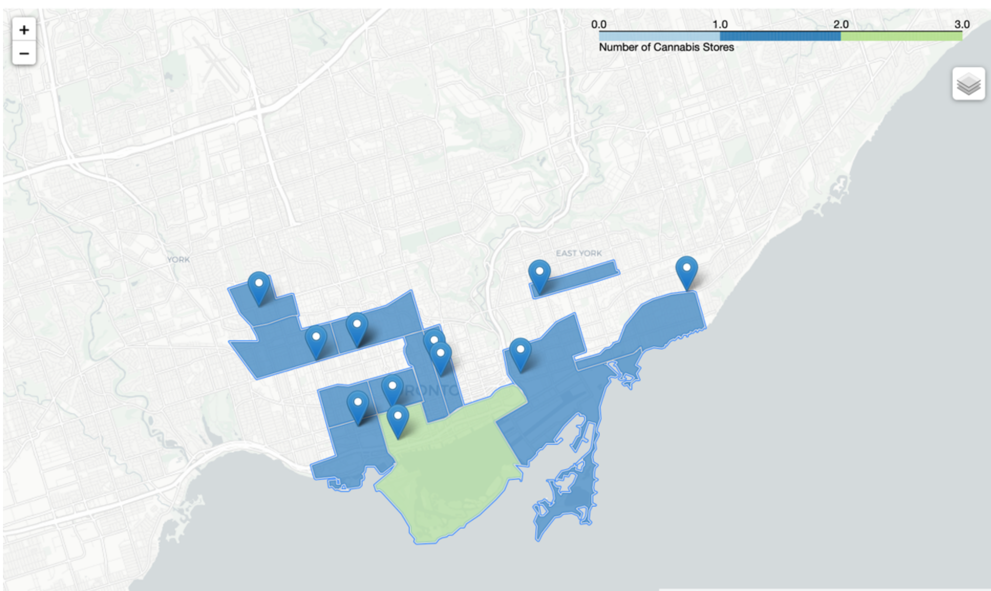


Figure 7: Map showing Cannabis store locations and number of stores per neighbourhood

Next, Foursquare API was once again called upon to get the location data on venues in each neighborhood. For this, the limit on results was set to 100 locations and the radius was set as 1Km. This resulted in a dataframe with 5698 venues.

Using One-hot encoding, categorical data on venues was transformed into numerical data and the mean frequency of each venue category was calculated

	neighborhood	ATM	Accessories Store	Adult Boutique	Afghan Restaurant	African Restaurant	American Restaurant	Amphitheater	Animal Shelter	Antique Shop	...	Vietnamese Restaurant	Warehouse Store	Whis I
0	Wychwood	0	0	0	0	0	0	0	0	0	...	0	0	
1	Wychwood	0	0	0	0	0	0	0	0	0	...	0	0	
2	Wychwood	0	0	0	0	0	0	0	0	0	...	0	0	
3	Wychwood	0	0	0	0	0	0	0	0	0	...	0	0	
4	Wychwood	0	0	0	0	0	0	0	0	0	...	0	0	

Figure 8: Turning categorical values into numerical values using One-hot encoding

Using the mean frequencies of each venue category, a new dataframe was created to show the top 10 most common venues in every neighborhood

	neighborhood	neighborhood_num	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Agincourt North	129	Chinese Restaurant	Sandwich Place	Indian Restaurant	Bakery	Vietnamese Restaurant	Coffee Shop	Bank	Discount Store	Pharmacy	Pizza Place
1	Agincourt South-Malvern West	128	Chinese Restaurant	Cantonese Restaurant	Shopping Mall	Restaurant	Boutique	Korean Restaurant	Motorcycle Shop	Noodle House	Clothing Store	Sushi Restaurant
2	Alderwood	20	Park	Gas Station	Convenience Store	Pharmacy	Pizza Place	Moroccan Restaurant	Discount Store	Grocery Store	Gym	Sandwich Place
3	Annex	95	Café	Coffee Shop	Restaurant	Italian Restaurant	Bakery	French Restaurant	Vegetarian / Vegan Restaurant	Museum	Tea Room	Gastropub
4	Banbury-Don Mills	42	Coffee Shop	Restaurant	Pizza Place	Bank	Japanese Restaurant	Café	Burger Joint	Mexican Restaurant	Supermarket	Botanical Garden

Figure 9: Dataframe showing 10 most common venues in every neighbourhood

A Word Cloud was also created to visualise the most common venues across all neighbourhoods in Toronto. By looking at the Word Cloud, it is clear that Coffee shop, Café, Restaurant, Pizza Place and Park are the most common venues across all neighbourhoods.

Figure 10: Word cloud showing most common venues across all neighbourhoods

3. Exploratory analysis

Once the data had been collected, processed and transformed into the required specific format, exploratory analysis was done to visualize which neighborhoods have the highest percentage of both target groups.

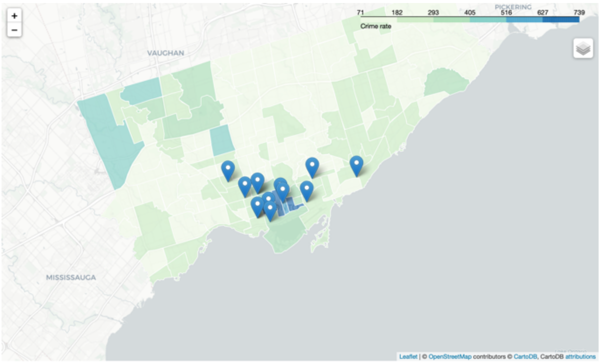


Figure 11: Choropleth Map showing % of Population within the age group of 20-39 and Cannabis retail store locations

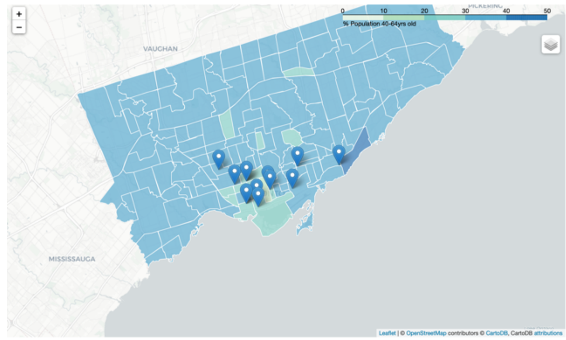


Figure 12: Choropleth Map showing % of Population within the age group od 40-64 and Cannabis retail store locations

Observations:

- From the above maps, it is clear that Target group 1(Ages 20-39) mainly resides in Downtown Toronto, whereas Target Group 2 (Ages 40-69) is concentrated outside of Downtown Toronto.
- Outside Downtown Toronto, the percentage of Target group 1 are in the range of 20%-40% with most neighborhoods reflecting this. As for our Target group 2, they make 30%-40% of population of most neighborhoods in Toronto.
- Waterfront Communities Neighborhood has the highest percentage of people in the age group 20-39 whereas Birchcliffe-Cliffeside has the highest percentage of people in the age group of 40-64.
- Cannabis retail locations are mainly located in areas that have a high percentage of population within the age group of 20-39. In the decision matrix, this feature will be given higher weightage.

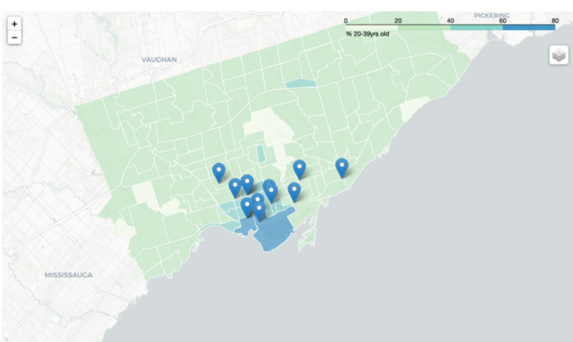


Figure 13: Choropleth map of crime rates in Toronto and Cannabis retail store locations

Observations:

- From the choropleth map, it can be seen that cannabis retailers do not consider low neighbourhood crime rate to be a very important factor in selecting a retail location as most current locations are located in medium to high crime rate areas Therefore, in the final analysis, crime rate should have comparatively low weightage

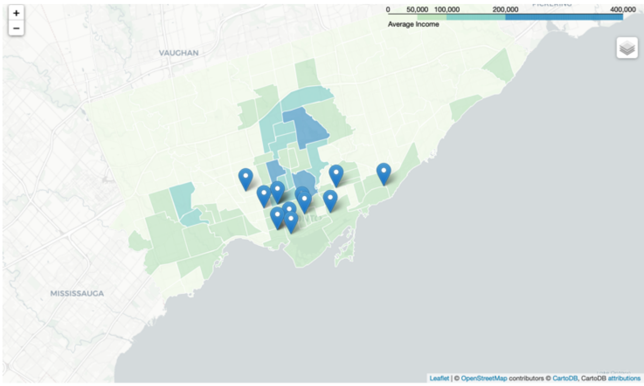


Figure 14: Choropleth map of average income in neighbourhoods and Cannabis store locations

Observations:

- Average neighbourhood income can be seen as a factor in choosing store location as all stores are concentrated in and around areas with incomes higher than 50000

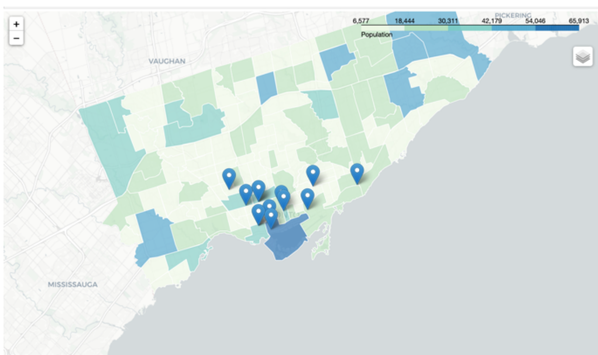


Figure 16: Choropleth map of population and Cannabis store locations

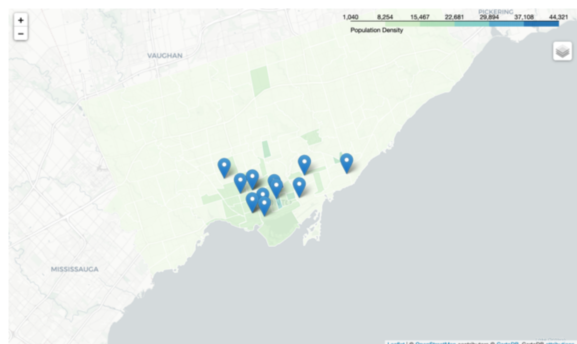


Figure 15: Choropleth map of population density and Cannabis store locations

Observations:

- The choropleth map of Population shows that Cannabis stores are usually found within medium to high population areas. Thus, implying that neighbourhood population is a considerable feature in final analysis. This feature will carry a higher score.
- The choropleth map of Population density shows that majority of neighborhoods have a population density of 8240 people per sq.km. Cannabis locations are present in neighbourhoods that have population density of higher than 8240 people per sq.km.

Finally, to validate the observations, a correlational matrix was built to see the relation between the selected features and store count. The correlation matrix shows that % never married, % of couples without children and % 20-39yrs have high correlational with number of stores. Therefore, based on this matrix and above choropleth maps, weights will be decided for the final decision matrix.

```
population                    0.321834
population_density            0.240975
% 20-39yrs                    0.565383
% 40-64yrs                   -0.407314
% never_married               0.439782
% couples_without_children    0.481944
average_income                0.051673
crime_rate                    0.375631
store_count                   1.000000
Name: store_count, dtype: float64
```

Clustering:

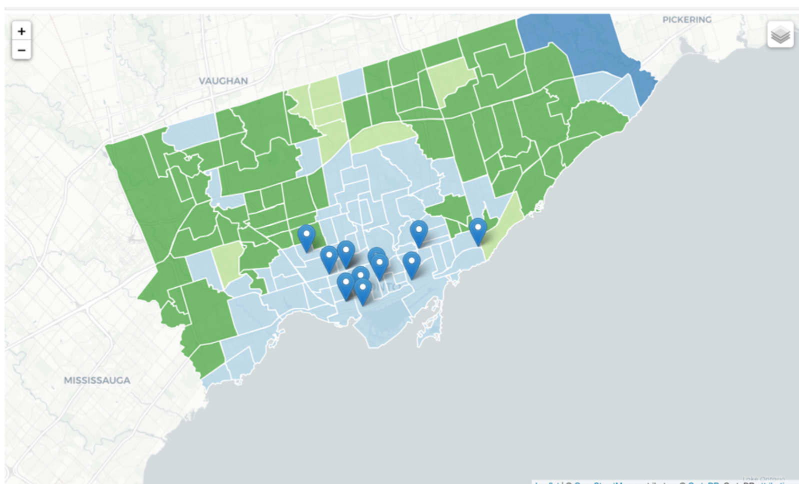
Now, using scikit-learn library and K-Means clustering algorithm, neighbourhoods were clustered into 4 clusters based on the mean frequency of venue categories.
The machine learning algorithm produced the following results:

The resulting clusters were visualized and existing store locations were plotted.

```
print ('There are {} neighborhoods in Cluster 0'.format(clus0.shape[0]))
print ('There are {} neighborhoods in Cluster 1'.format(clus1.shape[0]))
print ('There are {} neighborhoods in Cluster 2'.format(clus2.shape[0]))
print ('There are {} neighborhoods in Cluster 3'.format(clus3.shape[0]))
```

```
There are 68 neighborhoods in Cluster 0
There are 1 neighborhoods in Cluster 1
There are 8 neighborhoods in Cluster 2
There are 63 neighborhoods in Cluster 3
```

Figure 17: Code showing number of neighbourhoods in each cluster



Cluster 0 and Cluster 3 have 131 neighbourhoods together and neighbourhoods in these clusters are filled with commercial venues such as stores, restaurants, coffee shops, bars. As you can see, neighbourhoods in Cluster 0 are business areas with lots of restaurants, shops, cafes, etc. whereas neighbourhoods in Cluster 3 have a mix of residential and commercial venues such as parks, gas stations, coffee shops, fast food restaurants, etc. Cluster 1 is a cluster of residential neighbourhoods with very few commercial venues and Cluster 2 is a neighbourhood with a zoo and a park. All licensed Cannabis stores are in neighbourhoods within Cluster 0 which tells us that Cannabis retailers look for locations in commercial and business areas. So, overall, neighbourhoods in these two clusters are suitable for a Cannabis store.

Figure 20: Word Cloud Cluster 0

Figure 19: Word Cloud Cluster 3

Neighbourhoods with existing store locations were removed from Cluster 1 and the remaining 120 neighbourhoods were chosen for further analysis.

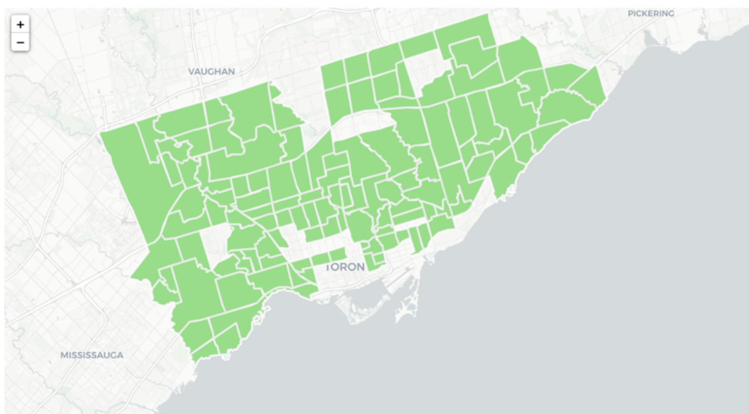


Figure 21: Map showing neighbourhoods selected for analysis

4. Analysis:

neighborhood	neighborhood_num	population	population_density	% 20-39yrs	% 40-64yrs	% never_married	% couples_without_children	average_income	crime_rate
Agincourt North	129	29113.0	3929.0	0.2485	0.3535	0.2910	0.3720	30414.0	113.030
Agincourt South-Malvern West	128	23757.0	3034.0	0.2909	0.3458	0.3192	0.3968	31825.0	215.295
Alderwood	20	12054.0	2435.0	0.2584	0.3800	0.2922	0.4167	47709.0	113.250
Annex	95	30526.0	10863.0	0.4082	0.2875	0.4283	0.6592	112766.0	303.045
Banbury-Don Mills	42	27695.0	2775.0	0.2154	0.3522	0.2666	0.4744	67757.0	127.845

Figure 22: Dataframe used for analysis

In finding the best suited neighborhoods, each neighborhood had to be scored for meeting or exceeding the criteria. Scores were calculated by subtracting the minimum criteria from

individual values and the dividing it by the minimum criteria. By this formula, if a neighborhood attribute meets the minimum criteria defined in Table 3 then it should have a positive score and if it does not meet the criteria it should have a negative score. Neighborhoods that perfectly match the criteria will have a score of 0 across all attributes.
Individual feature scores were multiplied by their assigned weights and the summed up to compute a total score.
The table below summarizes the criteria selection and weights used in the decision matrix

Table 3: Selected Criteria and Assigned Weights

	Criteria (description)	Minimum criteria	Weight
1.	Above average Population	Mean	0.075
2.	Above average Population density	Mean	0.025
3.	High percentage of population in the age group of 20-40	Atleast 30% (mean = 27.7%)	0.25
4.	Average-High percentage of population in the age group of 40-64	Atleast 35% (mean = 34.5)	0.15
5.	Average income above 30,000$	30000$	0.1
6.	High percentage of population without children	Mean	0.15
7.	High percentage of never married population	Mean	0.15
8.	Low crime rate	Mean	0.1

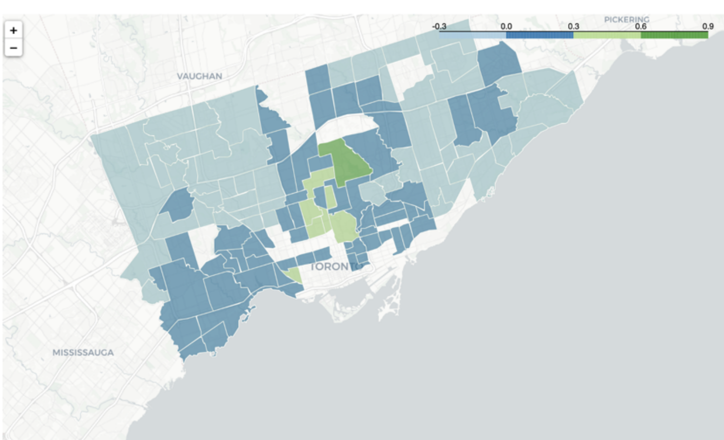


Figure 21: Choropleth Map showing scores by neighbourhood

	neighborhood	neighborhood_num	score
0	Bridle Path-Sunnybrook-York Mills	41	0.7381
1	Rosedale-Moore Park	98	0.5576
2	Forest Hill South	101	0.5533
3	Casa Loma	96	0.4702
4	Mount Pleasant West	104	0.4130
...	...	...	...
115	Oakridge	121	-0.1162
116	West Humber-Clairville	1	-0.1278
117	Humbermede	22	-0.1362
118	Yorkdale-Glen Park	31	-0.1800
119	Humber Summit	21	-0.2582

Figure 22: Final Dataframe with computed scores

5. Results:

Out of the 120 neighborhoods in the study area, 68 neighborhoods had a positive score implying that they met the minimum required criteria. The result of the analysis show that the following 10 neighborhoods exceed the minimum criteria and will be suitable locations for opening a retail cannabis store.

	neighborhood	neighborhood_num	score
0	Bridle Path-Sunnybrook-York Mills	41	0.7381
1	Rosedale-Moore Park	98	0.5576
2	Forest Hill South	101	0.5533
3	Casa Loma	96	0.4702
4	Mount Pleasant West	104	0.4130
5	Yonge-St.Clair	97	0.4087
6	Lawrence Park South	103	0.3578
7	Little Portugal	84	0.3160
8	Willowdale East	51	0.2986
9	North St. James Town	74	0.2900

Figure 23: Dataframe showing recommended neighbourhoods

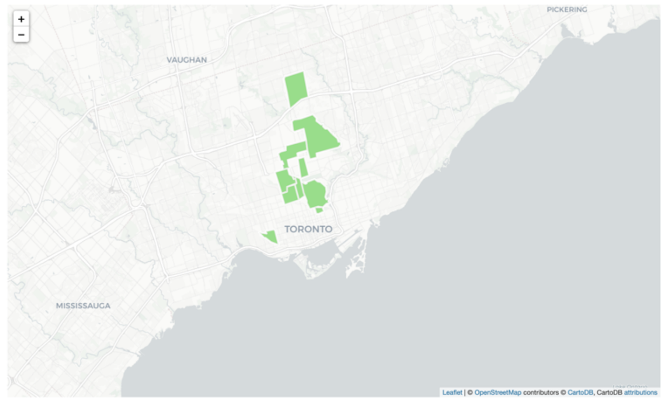


Figure 23: Top 10 recommended neighbourhoods

6. Discussion

Based on the calculated scores using the principles of a Decision Matrix, the following observations can be made:

- The most ideal neighbourhood to open a new Cannabis retail store based on the selected criteria would be **Bridle Path-Sunnybrook-York Mills (score: 0.7381).** This neighborhood currently has no store and overall scores high across all selected features.
- Though **Bridle Path-Sunnybrook-York** neighbourhood has a comparatively low percentage of Target group 1 (Ages 20 to 39), it has the highest average income across Toronto, low crime rate and almost 39% of its population is Target group 2 (Ages 40-64yrs). A cannabis store in this neighbourhood could benefit from adjusting its product mix with higher priced products such as pre-rolls, vapes, edibles, etc. Also, the company could focus its marketing on the medicinal uses of cannabis rather than recreational uses as customers in Target group 2 seek products that reflect the medicinal
- Locations of existing cannabis retail stores and the recommended neighbourhoods suggest that Downtown Toronto and Midtown Toronto house populations that shows characteristics of a cannabis customers.
- As the neighborhoods currently serviced by existing stores were left out of the final analysis, the results show neighbourhoods that are currently underserved.

7. Conclusion

The aim of the project was to identify ideal neighbourhoods for XYZ co. to open a new retail cannabis store. Data collected on cannabis store locations acquired from Foursquare validated the initial hypothesis that Cannabis stores in Toronto are mainly located in Downtown Toronto.
Further analysis on neighbourhood population census, income, crime rates and venues acquired from Toronto Data Portal and Public Safety Data Portal pointed to the fact that Cannabis retailers consider factors such as percentage of youth population, percentage of never married individuals, percentage of couples without children as well as neighbourhood income and crime rate to be important factors in choosing store locations.

The results recommended 10 neighbourhoods based on the selected criteria that would be ideal to open a new Cannabis store.

The methodology implemented and results of this project serve as a fundamental base for further analysis into retail cannabis market in Toronto. The project evaluated neighbourhoods based on a specific hypothesized scenario. Due to the limited scope of the project and resources available, the project lacks in considering additional features, that in real world would be important to a cannabis retailer. For example, commercial rent/lease, parking availability, store visibility, traffic, characteristics of the store, signage and zoning, regulations on store locations like proximity to schools, etc. The analysis considered only 119 neighbourhoods in the final study and the geographic scope of the project was limited to the city of Toronto. However, further analysis should look into regions within GTA and outside of the city of Toronto.

In addition to the data considered in this project, retailers can also leverage store level location data, in clustering stores based on individual store characteristics which would then be used for formulating a regional store location strategy.

Finally, as more retail locations open and more data is available on the retail store locations, building on the methodology and exploratory analysis implemented in this project, machine learning algorithms like Logistic Regression can be used in predicting neighbourhoods for opening Cannabis retail stores.

8. Next Steps…

This study helps to narrow down neighbourhoods for opening a new Cannabis retail store but before decision makers finalize a location, the next steps would be to ask additional questions. Some of the questions I would ask are:

1. Strategy: Has management decided on a high-level strategy and which neighbourhoods/locations fit this strategy?
2. Time frame: How long will it be before we open our new store?
3. Laws/Regulations: Will we have to comply with different laws and regulations for our new location?

4. Competition: Where are our competitors located and where will they be by the time we open a new store?
5. Products and Customers: What products we will be stocking? What is the detailed profile of our customer and what type of customers do we want to target with our new store? Which locations offer us the best coverage of our target customers?
6. Store Characteristics: How will the store look like and what elements of your brand will it have? How much space will be needed? Is there commercial space available to incorporate this vision? Is there labor available?
7. Do we need to consider other features specific to our store characteristic in evaluating neighbourhoods/locations? For example, parking availability, store visibility, traffic, characteristics of the store, signage and zoning, etc

Answers to these questions will allow decision makers to filter through neighbourhoods/locations and dive deeper into analyzing each neighbourhood/location. By the end of it, fewer neighbourhoods/locations will remain which are strategically relevant, have considerable population of our target group, are lawfully viable, not too competitive and have commercial real estate properties as per our store requirements and characteristics.

Bibliography

Dandapani, A. (2020, February 21). Challenges and Opportunities for Cannabis Retail in Canada Post-Legalization. *Retail Insider*.

Freeman, J., & Aguilar, B. (2019, December 13). Ontario scraps lottery system. *Ontario will scrap lottery system, move to expand retail cannabis market in 2020*.