

Opening a New Pharmacy in town



Introduction

Opening a successful independent pharmacy is a challenging undertaking for any entrepreneur, especially with the competitiveness of today's retail pharmacy market. Building a new pharmacy from the ground up doesn't guarantee foot traffic, and leasing or owning a property doesn't mean you're ready to open for business.

It takes a lot of hard work, determination, and patience in today's competitive market. One of the most important questions is to choose the physical location, the chosen location should sustain another pharmacy.

Business problem

The pharmacy industry is governed by state and federal regulations that are in place to provide safety and accountability for all involved. One must be compliant with all regulations on top of getting your business up and running before your grand opening.

During a pandemic or any health crisis, this will help business owners to make efficient decisions on selecting the best place in Toronto to open a pharmacy.

Location - Are there other pharmacies/chains/independents in the market?

Traffic - Do pharmacy customers regularly travel through this area?

Opportunity - Are there nearby businesses, such as medical offices, that will provide a constant flow (medical refills) of potential customers?

Access - Can people easily enter and exit? Is there plenty of parking?

Size - Can your pharmacy grow in this location?

Data Requirements

This project uses neighborhood data in Toronto.

Foursquare, Put the most trusted, independent location data, and technology platform to work for any business.

The foursquare location platform will be used as the sole data source since all the stated required information can be obtained through the API. Using API collect various features like latitude and longitude, zip codes, and venue data.

After finding the list of neighborhoods, I accessed the Foursquare API to gather information about venues inside each and every neighborhood on a 500 meters radius.

Data source :

- Wikipedia:
https://en.wikipedia.org/wiki/Special_wards_of_Tokyo#List_of_postal_codes_of_Canada:_M
- Geospatial data: http://cocl.us/Geospatial_data
- Foursquare APIs

Methodology

Data preparation and Exploratory Data Analysis

The First step I did is to scrape data from Wikipedia related to the list of postal codes in the Toronto region. Using *requests*, I got “Postal code”, “Neighborhood” and “Borough” features in a table. This table is used to create a data-frame. For this, I’ve used *pandas* to transform the table data into a data

frame. The next step is to gather “Latitude” and “Longitude” information across all the rows. This data is crucial in order to find the best location for opening a pharmacy. I retrieved geographical coordinates from this URL: 'http://cocl.us/Geospatial_data'. The table was matched against the Postal code. I merged both the table into a new data frame that consists of all the necessary data as below:

	Postal Code	Borough	Neighborhood	Latitude	Longitude
0	M3A	North York	Parkwoods	43.753259	-79.329656
1	M4A	North York	Victoria Village	43.725882	-79.315572
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494
...
98	M8X	Etobicoke	The Kingsway, Montgomery Road, Old Mill North	43.653654	-79.506944
99	M4Y	Downtown Toronto	Church and Wellesley	43.665860	-79.383160
100	M7Y	East Toronto	Business reply mail Processing Centre, South C...	43.662744	-79.321558
101	M8Y	Etobicoke	Old Mill South, King's Mill Park, Sunnylea, Hu...	43.636258	-79.498509
102	M8Z	Etobicoke	Mimico NW, The Queensway West, South of Bloor,...	43.628841	-79.520999

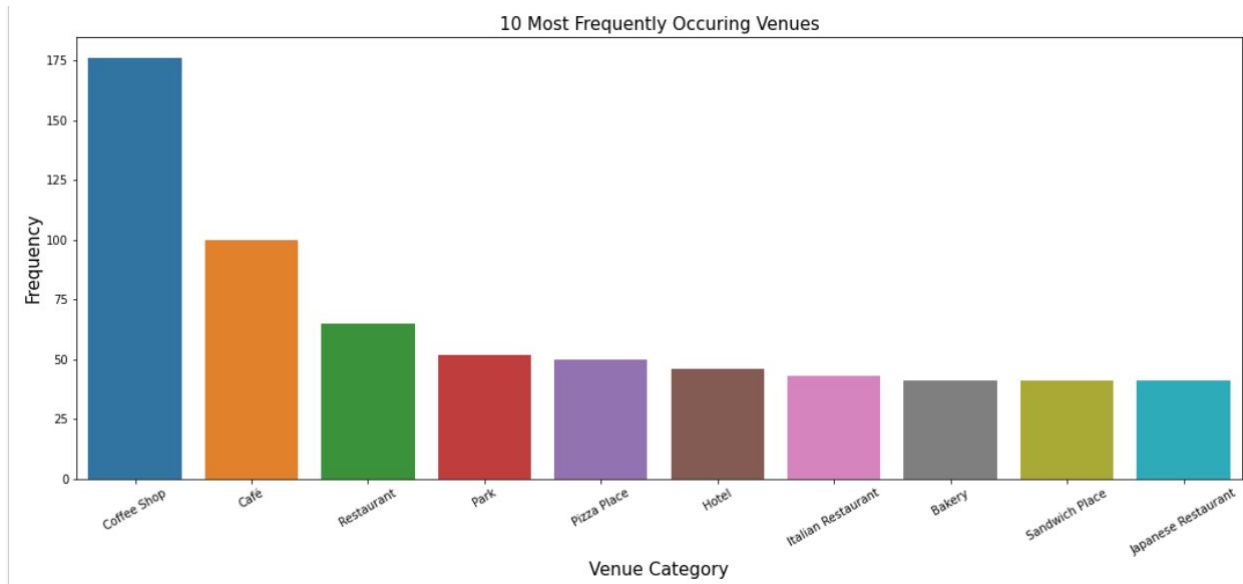
The Foursquare APIs allowed me to gather neighborhood information based on Venue category for all the locations listed in the data frame

I used Foursquare API to get the top 100 venues that are within a radius of 500 meters. I registered for a Foursquare Developer Account in order to obtain the Foursquare ID and Foursquare secret key. Foursquare will return the venue data in JSON format, this allowed me to extract the venue name, venue category, venue latitude, and longitude as shown below:

	Neighborhood	Latitude	Longitude	VenueName	VenueLatitude	VenueLongitude	VenueCategory
0	Parkwoods	43.753259	-79.329656	Brookbanks Park	43.751976	-79.332140	Park
1	Parkwoods	43.753259	-79.329656	Variety Store	43.751974	-79.333114	Food & Drink Shop
2	Victoria Village	43.725882	-79.315572	Victoria Village Arena	43.723481	-79.315635	Hockey Arena
3	Victoria Village	43.725882	-79.315572	Portugril	43.725819	-79.312785	Portuguese Restaurant
4	Victoria Village	43.725882	-79.315572	Tim Hortons	43.725517	-79.313103	Coffee Shop

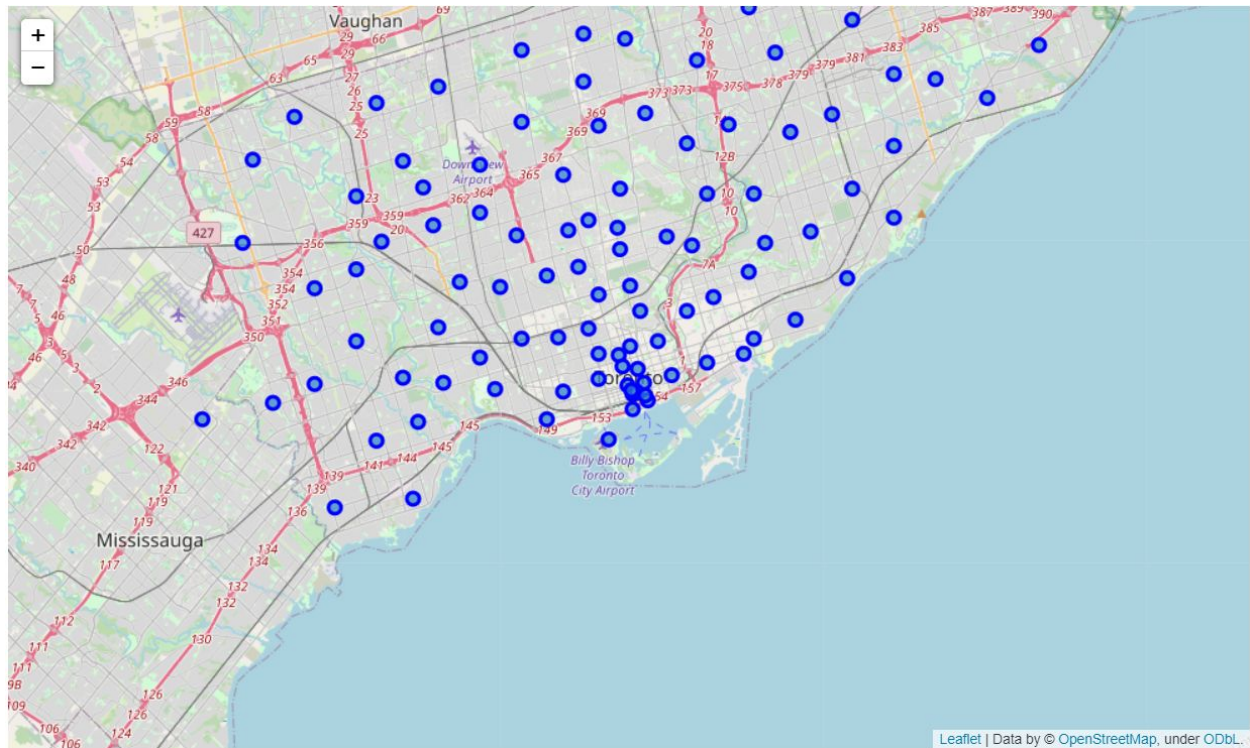
With this data, I explored venues that were returned for each neighborhood and examined how many unique categories can be curated from all the returned

venues. I then analyzed each neighborhood by grouping the rows by neighborhood and taking the mean of the frequency of occurrence of each venue category. I found that the “Pharmacy” was not one of the top venue categories (refer below). This is very helpful for me to understand that there are not many pharmacies in the Toronto region



Visualization

To make sure all the coordinates are mapped accordingly, I used the python *folium* library to visualize geographic details of Toronto and its neighborhood and I created a map of Toronto with boroughs superimposed on top. I used latitude and longitude values to get the visual as below:

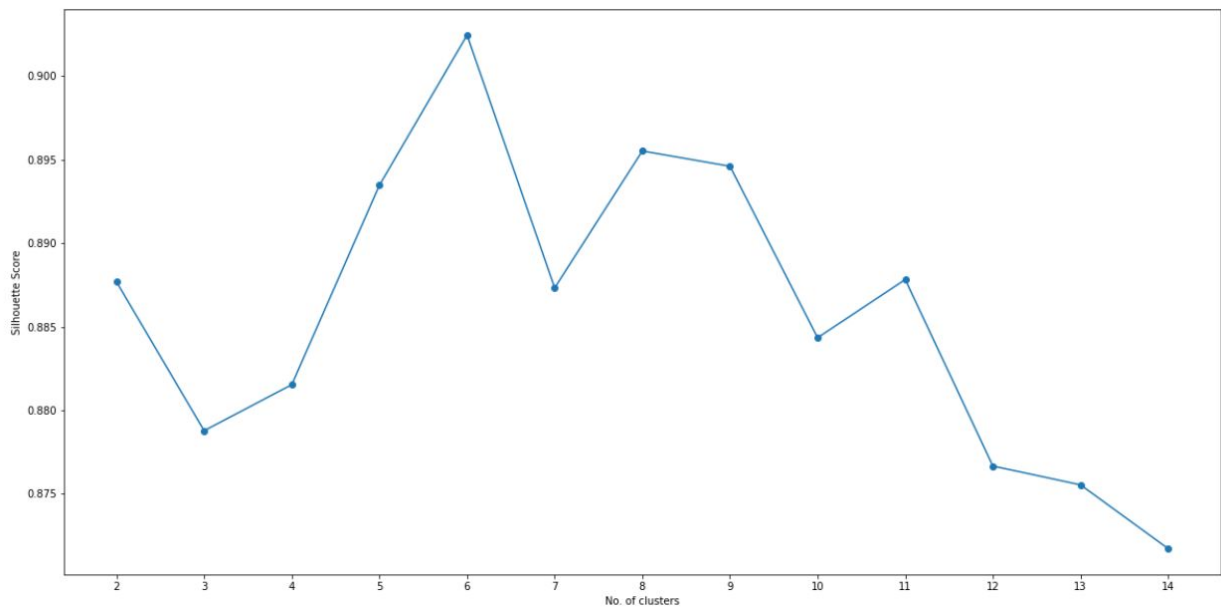


K means clustering

I used a simple unsupervised approach to cluster the locations.

k-means clustering is a method of vector quantization, originally from signal processing, that aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean (cluster centers or cluster centroid), serving as a prototype of the cluster.

The above data is now used to create cluster centroids and labels, allowing me to do further analysis.



I ran the algorithm with 6 clusters (figure above).

The results allowed me to identify which neighborhoods have a higher concentration of pharmacies while which neighborhoods have fewer pharmacies. Based on the occurrence/frequency of pharmacies in different neighborhoods, I was able to find the most suitable location for new pharmacies.

Result

The results from the k-means clustering based on the frequency of occurrence for “Pharmacy”:

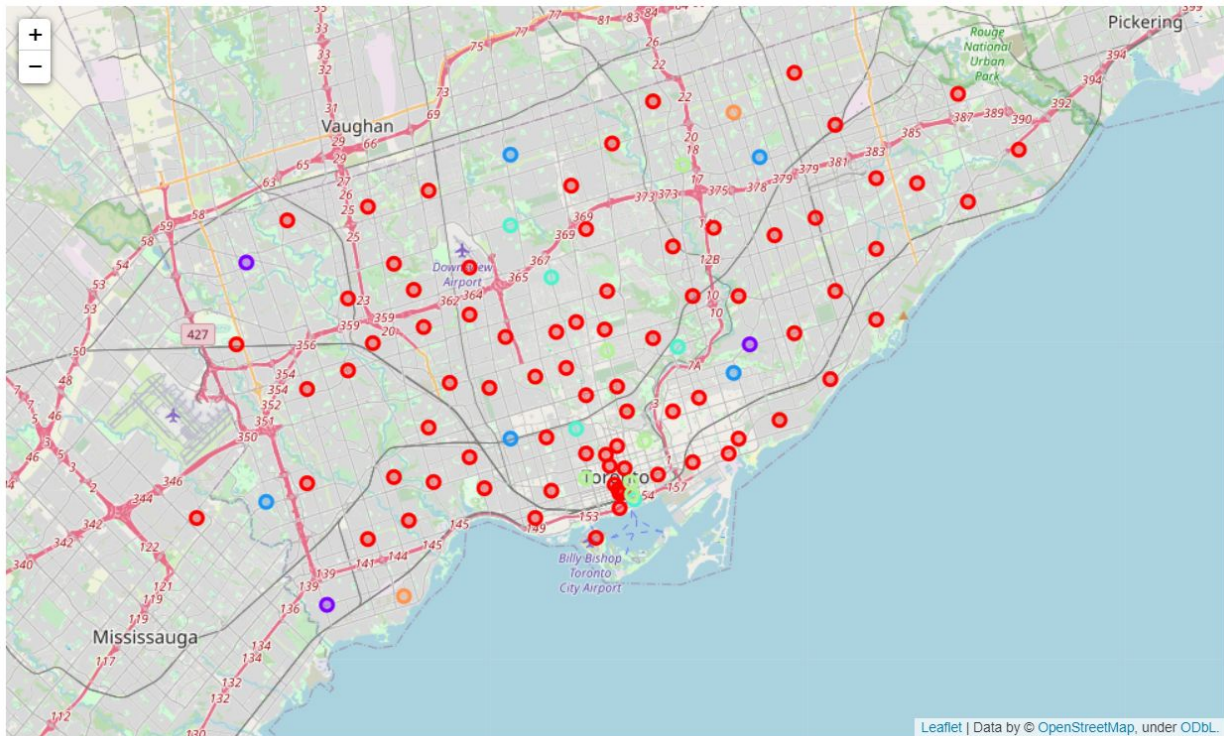
Cluster 0: Neighbourhoods shows with no sign of a pharmacy

Cluster 1: Neighbourhoods with a low number to no existence of pharmacy on the outskirts of the city

Cluster 2, Cluster 3: Neighbourhoods with a median concentration of pharmacy. There is not much traffic in these areas compared to the clusters 4 and 5

Cluster 4: Neighbourhoods with a high concentration of pharmacy with high traffic. Downtown Toronto comes with higher ratings.

Cluster 5: Neighbourhoods with a high concentration of pharmacy but median traffic. Etobicoke, and Scarborough



Discussion

Most of the Pharmacies are in Cluster 4 which is in the central area of Toronto and Downtown Toronto. Looking at nearby venues, it seems Cluster 4 and Cluster 5 might be a good location as there is a lot of traffic in these areas. People travel through these areas regularly.

By looking at the cluster data, we can see that cluster 0 (where there are no pharmacies) is the one that we are the most interested in order to find the best location for a new one. One of the main factors to look out while opening a pharmacy is refilling medical supplies and population density. There are nearby businesses, such as medical offices, that will provide a constant flow (medical refills) of potential customers.

Therefore, based on these findings it is recommended to open a pharmacy in locations closer to downtown Toronto, Etobicoke, and Scarborough with little competition.

Conclusion

Finding and investing in a location is possibly the most important thing to do before opening any pharmacy.

For an entrepreneur in Canada, the first stop will be to visit Ontario College of Pharmacist's web page on [Opening & Operating a Pharmacy](#). It will help with all the basics. Followed by extensive market research on the interest area.

The next logical step is to find a location.

This project is done with extensive feature analysis, using the best data science methods, and machine learning to provide solutions to the business problem. The results can be a powerful guide to show an entrepreneur to start a pharmacy the right way with the best location.