

# **Choosing the best location to open a bar in Toronto**

Crapha

*November, 2019*

## 1. Introduction

This project aims at identifying the best place to open a bar within the city of Toronto. This analysis is meant for entrepreneurs, investors or private persons willing to open a bar in Toronto or financing such a bar.

We make the assumption that the best places to open a bar are those where we can currently find the most popular (existing) bars : for various reasons that we will not analyze (but which are probably related to population density, distance to offices and other venues, attractiveness of the neighborhood, etc.) a bar in such a location has better chances of being popular/successful than in other places where les bars can be found.

The idea is therefore to detect the areas within Toronto with the highest density of popular bars.

## 2. Methodology

First, we will construct a dataframe listing all neighborhoods in Toronto, along with their latitude and longitude values.

Then, we will use Foursquare data to retrieve the most popular venues in each neighborhood.

We will then able to analyse this informations and list the neighborhoods which have « bars » as most popular venues. These will be the neighborhoods which will be kept as potential candidates for opening a new bar.

Initially, we also wanted to run a clustering algorithm on some detailed data about bars, available from Foursquare. However the standard « explore » request returns only very few useful information about bars (no rating, no price range, etc.). It seems that the only way to obtain this kind of information would be to use « Get Details of a venue » kind of requests. However, these are premium requests, which are limited to 500 daily. We will therefore not able to use them / obtain the information we needed. Instead, we will use the « section » parameter in our « explore » requests, where we chose « topPicks ». This is a cheap way of keeping only bars that are qualified as « topPicks » (instead of having a proper rating and other useful information). Also, we will run a clustering algorithm on all « topPicks » venues in Toronto in order gain some additional insight.

### 3. Data acquisition and cleaning

The following Wikipedia page is used to obtain the list of neighborhoods in Toronto :

[https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada:\\_M](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)

This list is transformed into a pandas data frame as follows :

	Postcode	Borough	Neighbourhood
0	M1B	Scarborough	Rouge, Malvern
1	M1C	Scarborough	Highland Creek, Rouge Hill, Port Union
2	M1E	Scarborough	Guildwood, Morningside, West Hill
3	M1G	Scarborough	Woburn
4	M1H	Scarborough	Cedarbrae

It is then « cleaned » in order to :

- Remove rows with unassigned borough
- Fill the missing neighborhood values with the corresponding borough value
- Concatenate rows with the same post code

Which yields a data frame which looks as follows :

	Postcode	Borough	Neighbourhood
0	M1A	Not assigned	Not assigned
1	M2A	Not assigned	Not assigned
2	M3A	North York	Parkwoods
3	M4A	North York	Victoria Village
4	M5A	Downtown Toronto	Harbourfront

We then add the longitudes and latitudes for each row in the data frame. To do so, we first import some geospatial data available at [http://cocl.us/Geospatial\\_data](http://cocl.us/Geospatial_data) :

	Postal Code	Latitude	Longitude
0	M1B	43.806686	-79.194353
1	M1C	43.784535	-79.160497
2	M1E	43.763573	-79.188711
3	M1G	43.770992	-79.216917
4	M1H	43.773136	-79.239476

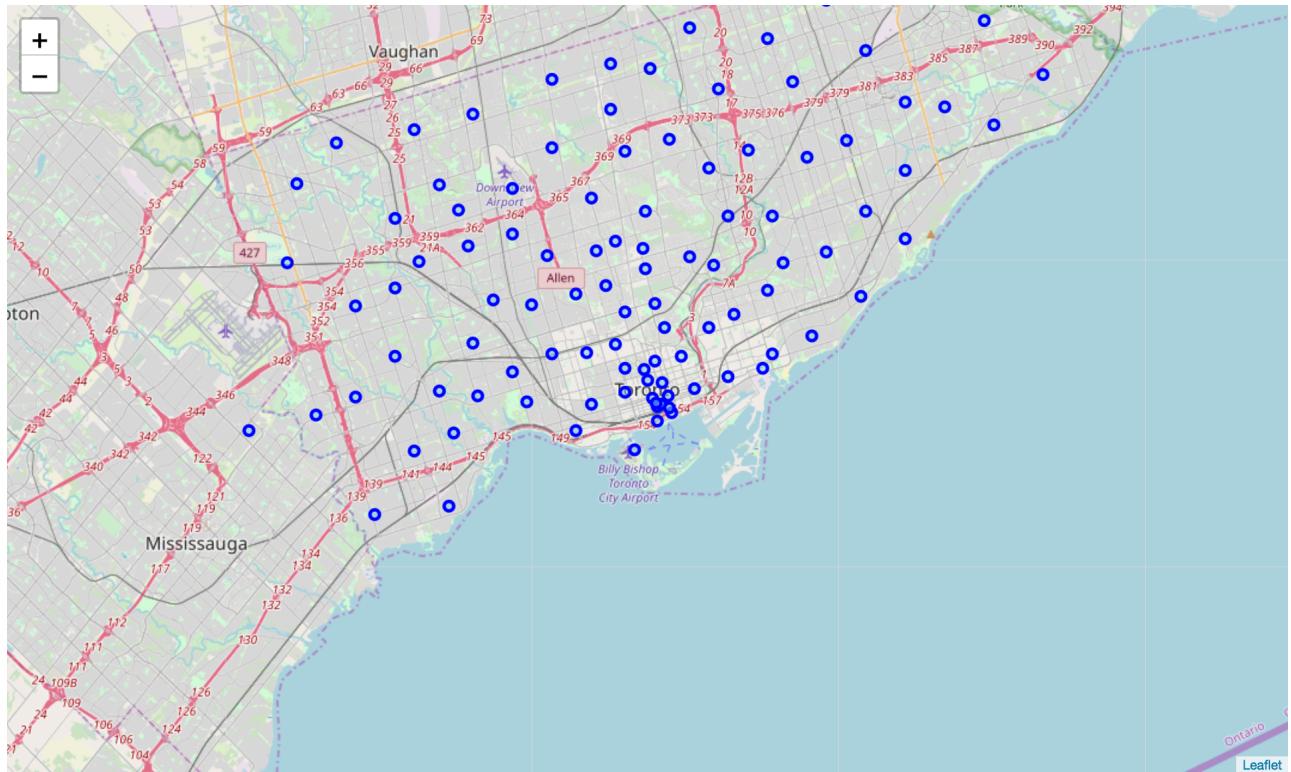
We then merge this data frame with the one constructed previously, using the post codes as common key :

	Postcode	Borough	Neighbourhood	Latitude	Longitude
0	M1B	Scarborough	Rouge, Malvern	43.806686	-79.194353
1	M1C	Scarborough	Highland Creek, Rouge Hill, Port Union	43.784535	-79.160497
2	M1E	Scarborough	Guildwood, Morningside, West Hill	43.763573	-79.188711
3	M1G	Scarborough	Woburn	43.770992	-79.216917
4	M1H	Scarborough	Cedarbrae	43.773136	-79.239476

We will also make Foursquare queries in order to retrieve information about the various venues that can be found in the different neighborhoods listed above. More details are given in the following and also in the Jupyter notebook.

## 4. Analysis and Results

We first plot a map of Toronto and its neighborhoods (blue circles) :



We then make Foursquare requests in order to locate the most popular (« topPicks ») venues located in a 500 m radius around each of these neighborhoods :

	Neighbourhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Rouge, Malvern	43.806686	-79.194353	Wendy's	43.807448	-79.199056	Fast Food Restaurant
1	Highland Creek, Rouge Hill, Port Union	43.784535	-79.160497	Royal Canadian Legion	43.782533	-79.163085	Bar
2	Guildwood, Morningside, West Hill	43.763573	-79.188711	G & G Electronics	43.765309	-79.191537	Electronics Store
3	Guildwood, Morningside, West Hill	43.763573	-79.188711	Big Bite Burrito	43.766299	-79.190720	Mexican Restaurant
4	Guildwood, Morningside, West Hill	43.763573	-79.188711	Swiss Chalet Rotisserie & Grill	43.767697	-79.189914	Pizza Place
5	Guildwood, Morningside, West Hill	43.763573	-79.188711	Eggsmart	43.767800	-79.190466	Breakfast Spot
6	Woburn	43.770992	-79.216917	Toronto Ali Sami Yen Spor Kompleksi (Lionhill)	43.767670	-79.218210	Soccer Field
7	Woburn	43.770992	-79.216917	Leslie & Giles Insurance Limited	43.773116	-79.219559	Insurance Office
8	Woburn	43.770992	-79.216917	Korean Grill House	43.770812	-79.214502	Korean Restaurant
9	Woburn	43.770992	-79.216917	Starbucks	43.770037	-79.221156	Coffee Shop
10	Woburn	43.770992	-79.216917	Tim Hortons	43.770827	-79.223078	Coffee Shop

We can then analyse these data in order to identify the most common venue types for each neighborhood :

	Neighbourhood	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	Adelaide, King, Richmond	Bar	Thai Restaurant	Café	Coffee Shop	Steakhouse	Clothing Store	Restaurant	Breakfast Spot	Cosmetics Shop	Bakery
1	Agincourt	Lounge	Breakfast Spot	Latin American Restaurant	Chinese Restaurant	Filipino Restaurant	Fast Food Restaurant	Farmers Market	Falafel Restaurant	Empanada Restaurant	Electronics Store
2	Agincourt North, L'Amoreaux East, Milliken, St...	Playground	Park	Yoga Studio	Dog Run	Farmers Market	Falafel Restaurant	Empanada Restaurant	Electronics Store	Eastern European Restaurant	Dumpling Restaurant
3	Albion Gardens, Beaumont Heights, Humbergate, ...	Pharmacy	Pizza Place	Beer Store	Sandwich Place	Fried Chicken Joint	Fast Food Restaurant	Donut Shop	Falafel Restaurant	Empanada Restaurant	Electronics Store
4	Alderwood, Long Branch	Skating Rink	Gym	Pizza Place	Pool	Pub	Dog Run	Falafel Restaurant	Empanada Restaurant	Electronics Store	Eastern European Restaurant

Since we are mostly interested by bars, we can see that 3 groups of neighborhoods have bars as most common venue type :

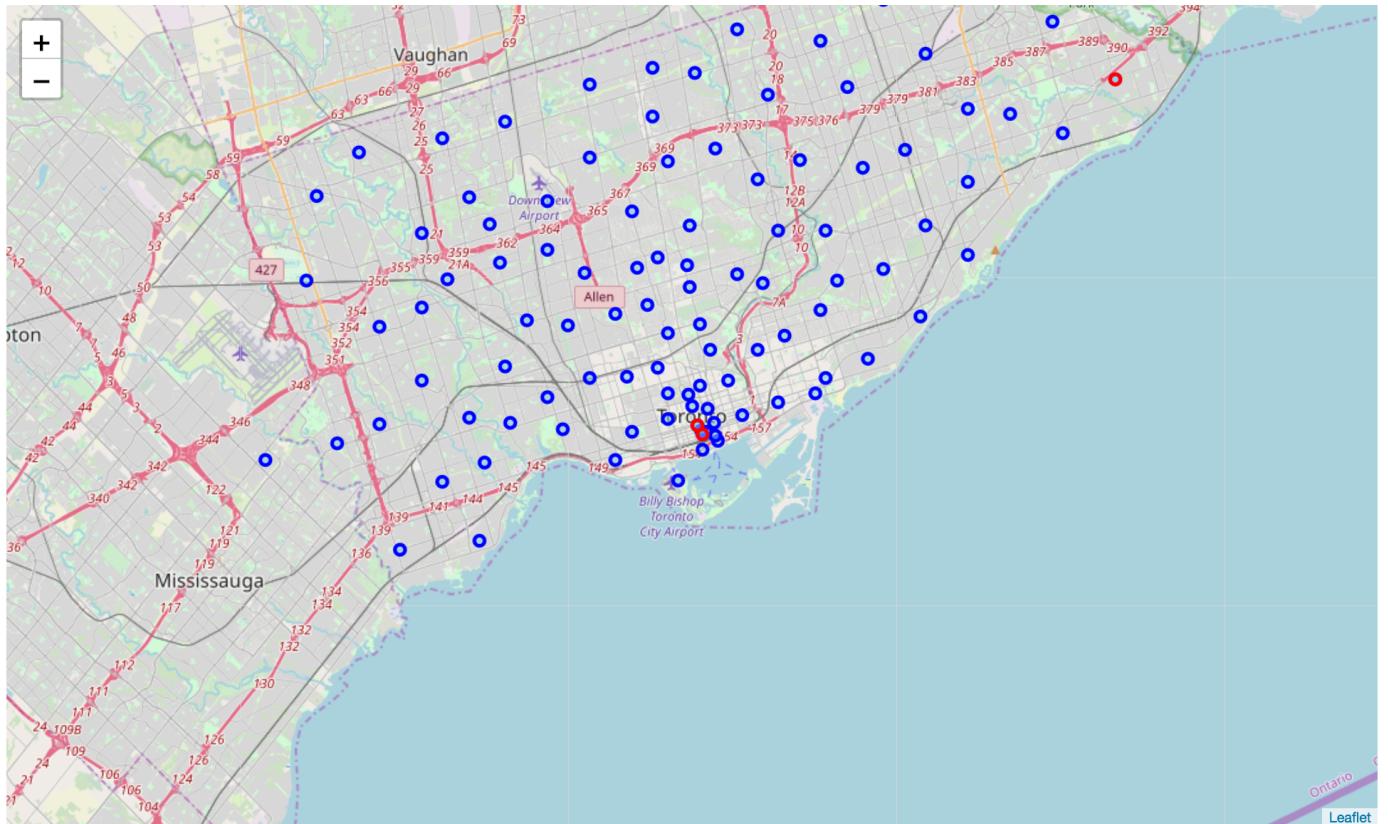
	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Adelaide, King, Richmond	Bar	Thai Restaurant	Café	Coffee Shop	Steakhouse	Clothing Store	Restaurant	Breakfast Spot	Cosmetics Shop	Bakery
32	Design Exchange, Toronto Dominion Centre	Bar	Japanese Restaurant	Coffee Shop	Roof Deck	Tailor Shop	Yoga Studio	Farmers Market	Falafel Restaurant	Empanada Restaurant	Electronics Store
53	Highland Creek, Rouge Hill, Port Union	Bar	Yoga Studio	Fish Market	Filipino Restaurant	Fast Food Restaurant	Farmers Market	Falafel Restaurant	Empanada Restaurant	Electronics Store	Eastern European Restaurant

This is a great insight for our task. It means that we should focus on these neighborhoods in order to decide where to open a bar. Let's list them again fully :

1. Adelaide
2. King
3. Richmond

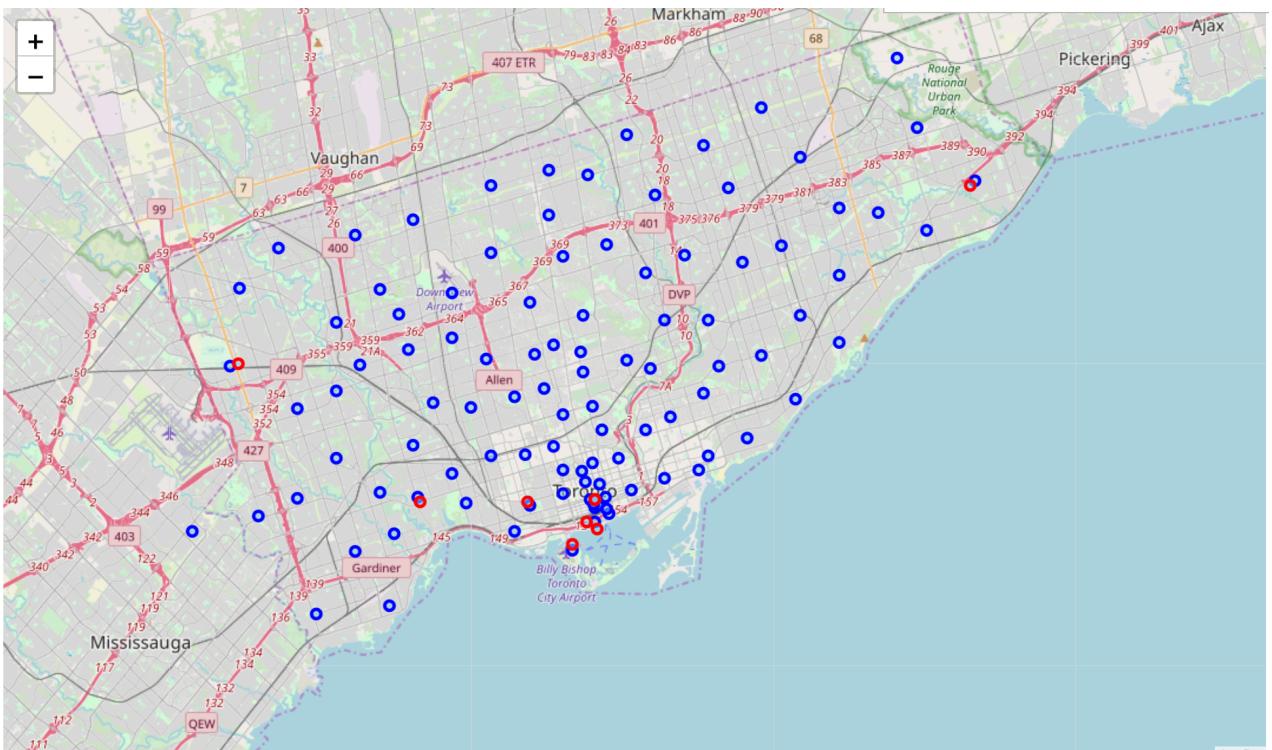
4. Design Exchange
5. Toronto Dominion Centre
6. Highland Creek
7. Rouge Hill
8. Port Union

These are distributed as follows (red circles on the map) :



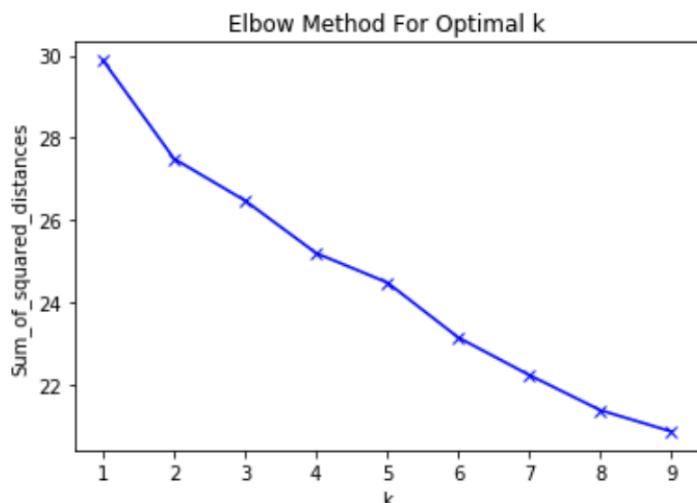
Apparently, the neighborhoods with bars as most common venue are located in the city centre, except for an outlier in the north east.

We also plotted the most popular bars in Toronto, just to compare :

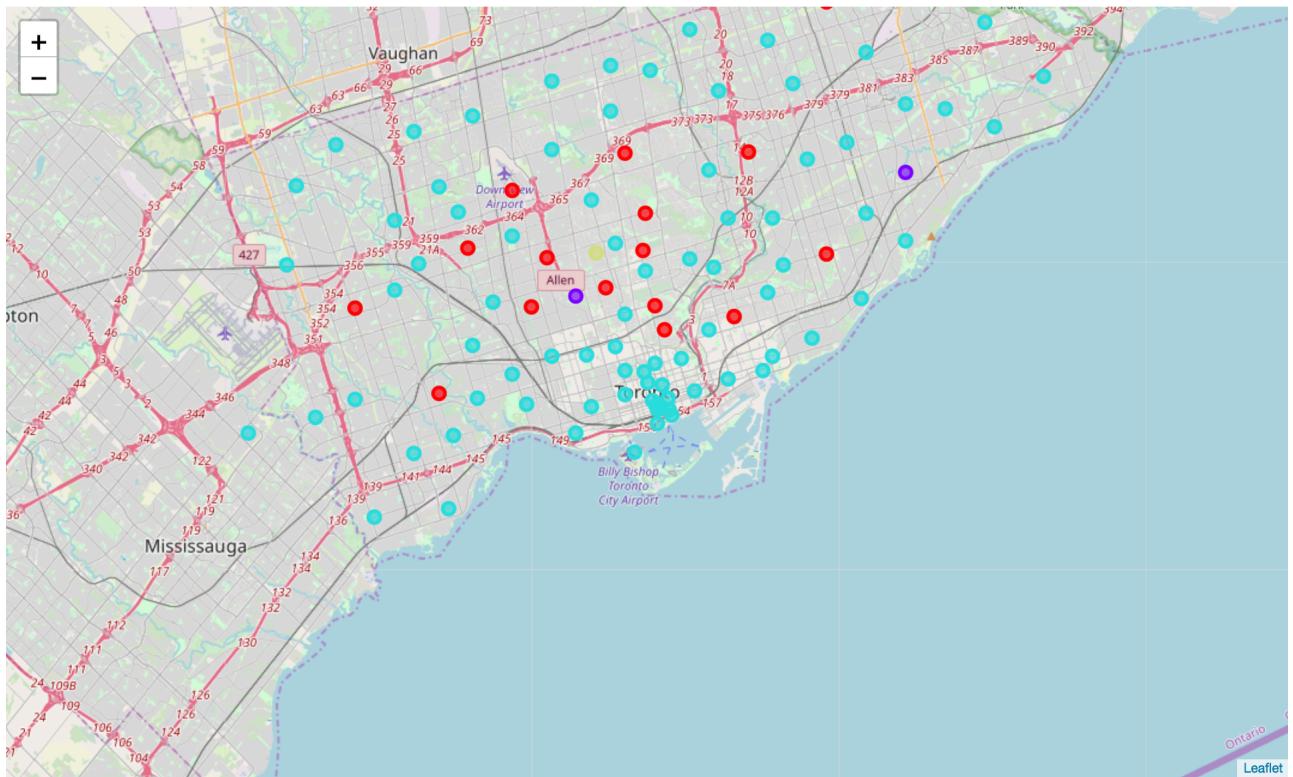


This map confirms that most of the popular bars are located in the city center.

We can also run a k-means clustering algorithm on the most popular (« toppicked ») venues in Toronto. Using the elbow method, we determine 4 as a good choice for the number of clusters (eventhough our graph does not show an ideal/neat elbow as we would expect) :



The result of a k-means clustering with 4 kernels is depicted below :



We can see that most neighborhoods are assigned to cluster no 2 (green circles). The neighborhoods we identified previously as having bars as most popular venues are included in this cluster too.

## 5. Discussion

Determining the location of popular bars is only one useful information among others when deciding where to open a new bar. Unfortunately, we did not have other information at hand to perform deeper analysis (clustering according to various other factors, as planned initially).

We were still able to see that the most popular bars are located mostly in the city center. This makes sense, as it is the densest part of the city, where people gather - not only to work, but also to have drinks afterwards...

Opening a bar in the city center would be a good option since it is « the place to be », however it is also where there is the most competition and probably where the rents are the most expensive. Also, opening a bar in this region means following the trend.

This statistical analysis should therefore be balanced by some human exploration of Toronto in person, in order to possibly identify other neighborhoods where a bar would be welcome, even if no such bar exists at the moment yet. We suggest to start an exploration in the 8 neighborhoods identified as having bars as most common venue.

## **6. Conclusion**

This final project was interesting as it was a good opportunity to apply data analysis techniques to a real world business case. We were able to gather various information about Toronto and its popular venues and bars. Our analysis showed that most popular bars are located in the city center, as we would have expected.