

# The Battle of the Neighborhoods

## Introduction:

Toronto is a vibrant, big-time city abuzz with activity. Some of the world's finest restaurants are found here, alongside happening bars and clubs and eclectic festivals. A rich, first world city known for its low crime rates and unique nightlife, Toronto stands out as a prime candidate as a place to open a new bar. However, a key piece of criteria when making a decision about where to open a new bar is getting the location right. Get it right and you could make a fortune, get it wrong and it could be a financial disaster. This project aims to help businesspeople to choose an area that is not saturated with bars and in a small way help them find a gap in the market.

## Business Problem:

What are the best Neighborhoods in downtown Toronto for setting up a new bar? A question any budding entrepreneur worth his or her salt would be dying to know! Armed with data, this project uses Machine Learn and Data Science methods to provide answers to this business problem. Q. What are the best Neighborhoods in downtown Toronto for setting up a new bar?

## Data Acquisition and Cleaning:

Data was pulled from the following areas -

- Canadian Postal Code Wikipedia page for lists of data about Postal Codes and their relevant Borough and Neighbourhood.
- Geocoder Package to get coordinates of Neighbourhoods in Downtown Toronto.
- Data about nearby bars from Foursquare crowdsourcing data.

Data was scraped from the Canadian Postal Code Wikipedia page using the Beautiful Soup Library. The column needed to be re-named from 'Neighbourhood' to 'Neighborhood' so that processing issues later down the line could be avoided because 'Neighborhood' came up in multiple instances.

Cells with a borough were processed, cells without were ignored. I also combined postal codes which are listed twice and cells which had a borough but a Not assigned neighborhood by replacing 'Not assigned' with an entry which is same as the borough.

I used the CSV file to the GPS coordinates and Postal Codes of Toronto boroughs, before merging that data with data about the Named boroughs and neighborhoods - and create the final dataframe.

## Methodology:

The model we use to solve this problem is K-means clustering. We use the K-means clustering algorithm and label all the neighborhoods in downtown Toronto into 7 clusters. Here are some sample data from some of the clusters generated by our model:

	Neighborhood	Bar	Cluster Labels	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
9	Design Exchange, Toronto Dominion Centre	0.02	0	43.647177	-79.381576	Cafe Landwer	43.648753	-79.385367	Café
10	First Canadian Place, Underground city	0.02	0	43.648429	-79.382280	Boxcar Social Temperance	43.650557	-79.381956	Bar
10	First Canadian Place, Underground city	0.02	0	43.648429	-79.382280	Cafe Landwer	43.648753	-79.385367	Café
10	First Canadian Place, Underground city	0.02	0	43.648429	-79.382280	Pumpnickel's Deli	43.648832	-79.381970	Deli / Bodega
10	First Canadian Place, Underground city	0.02	0	43.648429	-79.382280	Mos Mos Coffee	43.648159	-79.378745	Café
10	First Canadian Place, Underground city	0.02	0	43.648429	-79.382280	Walrus Pub & Beer Hall	43.647375	-79.379515	Pub
10	First Canadian Place, Underground city	0.02	0	43.648429	-79.382280	John & Sons Oyster House	43.650656	-79.381613	Seafood Restaurant
10	First Canadian Place, Underground city	0.02	0	43.648429	-79.382280	Starbucks	43.646936	-79.384428	Coffee Shop
10	First Canadian Place, Underground city	0.02	0	43.648429	-79.382280	Estiatorio Volos	43.650329	-79.384533	Greek Restaurant
10	First Canadian Place, Underground city	0.02	0	43.648429	-79.382280	Hy's Steakhouse	43.649505	-79.382919	Steakhouse
10	First Canadian Place, Underground city	0.02	0	43.648429	-79.382280	Design Exchange	43.647972	-79.380104	Art Gallery
10	First Canadian Place, Underground city	0.02	0	43.648429	-79.382280	King Taps	43.648476	-79.382058	Gastropub
10	First Canadian Place, Underground city	0.02	0	43.648429	-79.382280	Brick Street Bakery	43.648815	-79.380605	Bakery
10	First Canadian Place, Underground city	0.02	0	43.648429	-79.382280	Pilot Coffee Roasters	43.648835	-79.380936	Coffee Shop
10	First Canadian Place, Underground city	0.02	0	43.648429	-79.382280	Sam James Coffee Bar (S.J.C.B.)	43.647881	-79.384332	Café

```
In [62]: #Cluster 2
cluster_2=bar_cpy.loc[bar_cpy["Cluster Labels"]==2]
cluster_2
```

Out[62]:

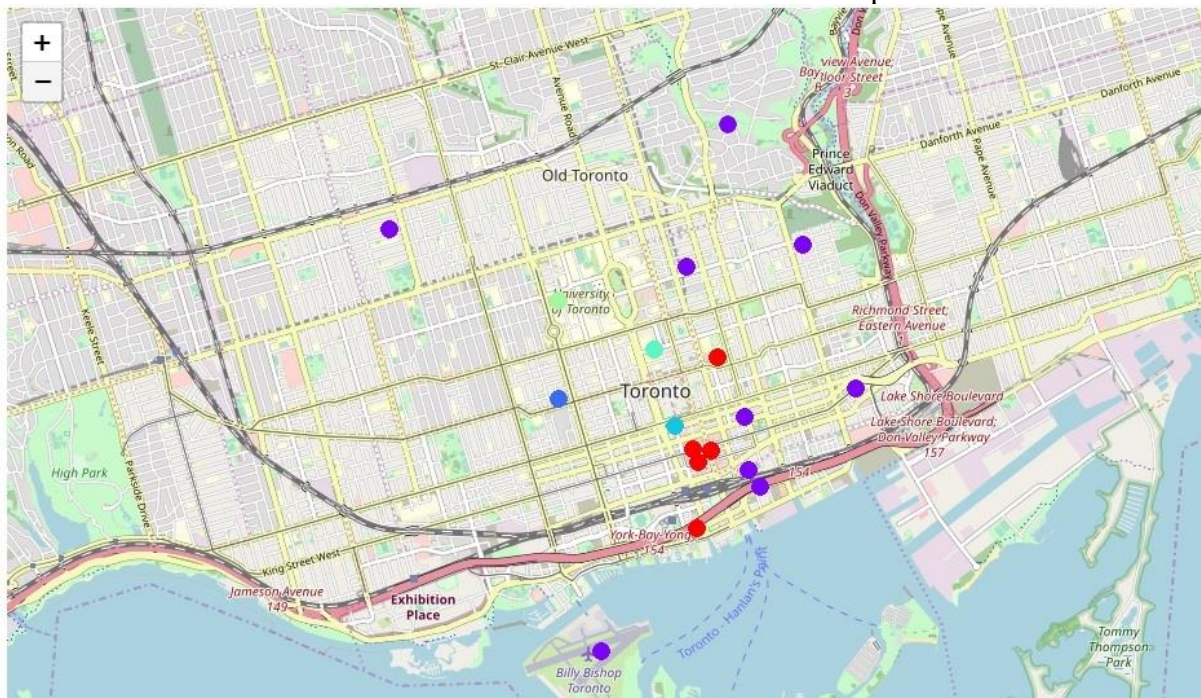
	Neighborhood	Bar	Cluster Labels	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
11	Harbord, University of Toronto	0.058824	2	43.662696	-79.400049	Gyubee	43.667088	-79.400571	Japanese Restaurant
11	Harbord, University of Toronto	0.058824	2	43.662696	-79.400049	The Beer Store	43.665385	-79.403477	Beer Store
11	Harbord, University of Toronto	0.058824	2	43.662696	-79.400049	Yasu	43.662837	-79.403217	Japanese Restaurant
11	Harbord, University of Toronto	0.058824	2	43.662696	-79.400049	Piano Piano	43.662949	-79.402898	Italian Restaurant
11	Harbord, University of Toronto	0.058824	2	43.662696	-79.400049	Rasa	43.662757	-79.403988	Restaurant
11	Harbord, University of Toronto	0.058824	2	43.662696	-79.400049	The Dessert Kitchen	43.662823	-79.402746	Dessert Shop
11	Harbord, University of Toronto	0.058824	2	43.662696	-79.400049	Almond Butterfly	43.662836	-79.403365	Bakery
11	Harbord, University of Toronto	0.058824	2	43.662696	-79.400049	Bakka Phoenix Books	43.662959	-79.402601	Bookstore
11	Harbord, University of Toronto	0.058824	2	43.662696	-79.400049	Her Father's Cider Bar + Kitchen	43.662448	-79.404703	Beer Bar
11	Harbord, University of Toronto	0.058824	2	43.662696	-79.400049	Cafe Cancan	43.662735	-79.403447	French Restaurant
11	Harbord, University of Toronto	0.058824	2	43.662696	-79.400049	Harbord House	43.662466	-79.405410	Bar

```
In [66]: #Cluster 6
cluster_6 = bar_cpy.loc[bar_cpy["Cluster Labels"]==6]
cluster_6
```

Out[66]:

	Neighborhood	Bar	Cluster Labels	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
5	Chinatown, Grange Park, Kensington Market	0.06	6	43.653206	-79.400049	Wanda's Pie in the Sky	43.656163	-79.402630	Bakery
5	Chinatown, Grange Park, Kensington Market	0.06	6	43.653206	-79.400049	Bunner's Bake Shop	43.655655	-79.402467	Bakery
5	Chinatown, Grange Park, Kensington Market	0.06	6	43.653206	-79.400049	Juicy Dumpling	43.652722	-79.398498	Dumpling Restaurant
5	Chinatown, Grange Park, Kensington Market	0.06	6	43.653206	-79.400049	Café Pamenar	43.656678	-79.402822	Café
5	Chinatown, Grange Park, Kensington Market	0.06	6	43.653206	-79.400049	I Deal Coffee	43.655058	-79.403254	Coffee Shop
5	Chinatown, Grange Park, Kensington Market	0.06	6	43.653206	-79.400049	Icha Tea	43.651336	-79.397370	Tea Room
5	Chinatown, Grange Park, Kensington Market	0.06	6	43.653206	-79.400049	Top Gun Steak	43.655039	-79.402167	Sandwich Place
5	Chinatown, Grange Park, Kensington Market	0.06	6	43.653206	-79.400049	Krispy Kreme	43.655834	-79.399417	Donut Shop
5	Chinatown, Grange Park, Kensington Market	0.06	6	43.653206	-79.400049	Buddha's Vegetarian	43.651904	-79.403312	Vegetarian / Vegan Restaurant
5	Chinatown, Grange Park, Kensington Market	0.06	6	43.653206	-79.400049	Pow Wow Cafe	43.654109	-79.401779	Taco Place
5	Chinatown, Grange Park, Kensington Market	0.06	6	43.653206	-79.400049	Otto's Berlin Döner	43.656387	-79.402788	Doner Restaurant
5	Chinatown, Grange Park, Kensington Market	0.06	6	43.653206	-79.400049	Nu Bügel	43.655547	-79.402528	Bagel Shop
5	Chinatown, Grange Park, Kensington Market	0.06	6	43.653206	-79.400049	Ronnie's Local 069	43.655104	-79.402675	Bar

We use folium to visualise the clusters and its distribution on the map.



## Results:

There are 7 clusters:

**Cluster (0)** – A decent variety of neighbourhoods with a moderate concentration of bars.

**Cluster (1)** : The most variety of neighbourhoods with a moderate concentration of bars.

**Cluster (2)** : Harbord and University of Toronto areas, contains a high amount of bars.

**Cluster (3)** : Covers 3 neighbourhoods (Adelaide, King, Richmond) and has a moderate level of bars.

**Cluster (4)** : Two neighbourhood (Ryerson, Garden District) with a low level of bars.

**Cluster (5)** : Single neighbourhood (Central Bay Street) with a moderate concentration of bars.

**Cluster (6)** : This cluster contains a high concentration of bars, which contains the Chinatown, Grange Park and Kensington.

## Observations/Conclusion:

It appears that the highest concentrations of bars are in Cluster 2 and 6. Cluster 2 covers the University of Toronto neighbourhood, this would make a lot of sense because it is a student neighbourhood. Cluster 6 has the highest level of bars so would be a very poor choice. Cluster 4, the Ryerson and Garden District areas have the lowest concentration of bars so would be a great place to consider investing in opening a new bar.