

Clustering Toronto Venues with K-means

(The Battle of Neighborhoods Project)

Sittisak Tantiprasongchai

March 7,2020

1. Introduction

1.1 Background

Toronto is the provincial capital of Ontario and the most populous city in Canada, with a population of 2,954,024 as of July 2018.

Toronto is an international center of business, finance, arts, and culture, and is recognized as one of the most multicultural and cosmopolitan cities in the world.

1.2 Problem

This project is to help people who interest in moving their home to a Toronto in Canada.

Some people might like to travel. Some people might like to eat food. Or resting in the park. What they do is Search in internet or just ask the people around.

And use those suggestions To decide.

1.3 Interest

Obviously, people who trade house very interested in area

That have advantage like park, garden, restaurant and other venues. Others who are interested in place around Toronto such as who decide to move their home in town and Tour guide may also be interested.


2. Data acquisition and cleaning

2.1 Data Sources

Postal, borough, neighborhoods data can be found in Wikipedia Website.

https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

dataset, however, lack data for neighborhoods. For example, the postal code stats dataset in M1A, and the postal code stats dataset in M2A,. To complement these, I scraped Wikipedia for neighborhood stats.



WIKIPEDIA
The Free Encyclopedia

- Main page
- Contents
- Featured content
- Current events
- Random article
- Donate to Wikipedia
- Wikipedia store

Interaction

- Help
- About Wikipedia
- Community portal
- Recent changes
- Contact page

Tools

- What links here
- Related changes
- Upload file
- Special pages

Article [Talk](#)

List of postal codes of Canada: M

From Wikipedia, the free encyclopedia

This is a list of [postal codes in Canada](#) where the first letter is M. Postal codes beginning with characters are listed, corresponding to the Forward Sortation Area.

[Canada Post](#) provides a free postal code look-up tool on its website,^[1] via its [applications](#) for s CD-ROMs. Many vendors also sell validation tools, which allow customers to properly match at offices, and some libraries.

Toronto - FSAs [\[edit \]](#)

Note: There are no rural FSAs in Toronto, hence no postal codes should start with M0, however Mississauga, suggesting that Canada Post may be allocating the M0 FSA for high volume addre

Postcode ↕	Borough ↕	Neighbourhood ↕
M1A	Not assigned	Not assigned
M2A	Not assigned	Not assigned
M3A	North York	Parkwoods
M4A	North York	Victoria Village
M5A	Downtown Toronto	Harbourfront

And latitude, longitude can be found on http://cocl.us/Geospatial_data

	PostalCode	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

For venues and location JSON data can be found on Foursquare API call.
<https://foursquare.com/>

```
results = requests.get(url).json()
results

{'meta': {'code': 200, 'requestId': '5e6271c5f7706a001b9dba32'},
 'response': {'headerLocation': 'The Beaches',
 'headerFullLocation': 'The Beaches, Toronto',
 'headerLocationGranularity': 'neighborhood',
 'totalResults': 5,
 'suggestedBounds': {'ne': {'lat': 43.680857404499996,
 'lng': -79.28682091449052},
 'sw': {'lat': 43.67185739549999, 'lng': -79.29924148550948}},
 'groups': [{ 'type': 'Recommended Places',
 'name': 'recommended',
 'items': [{ 'reasons': { 'count': 0,
 'items': [{ 'summary': 'This spot is popular',
 'type': 'general',
 'reasonName': 'globalInteractionReason' } ] },
 'venue': { 'id': '4bd461bc77b29c74a07d9282',
```

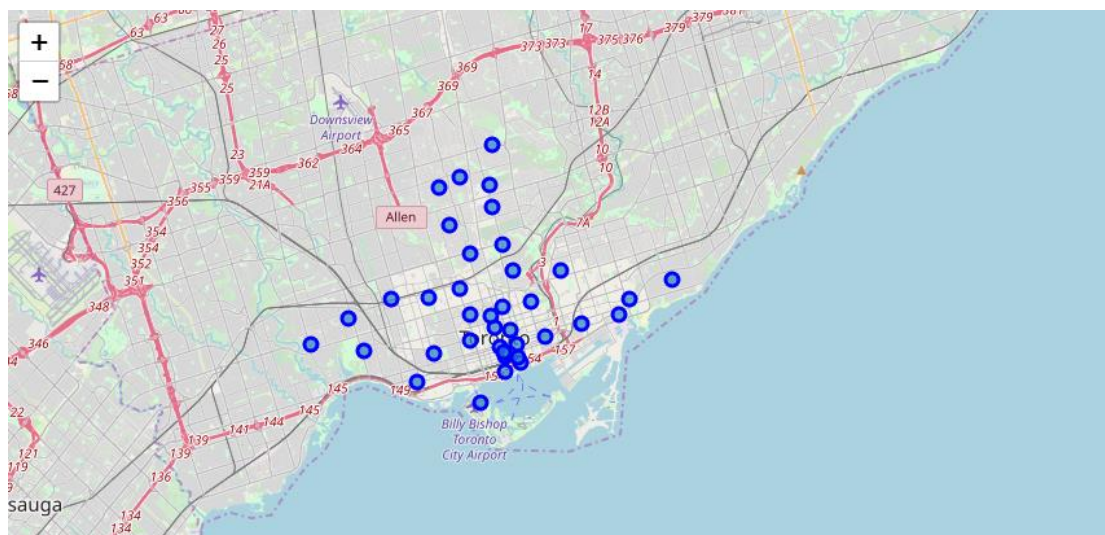
3. Methodology

- Data is scraping from Wikipedia page by Beautifulsoup Package
- Then, Convert to Data frame

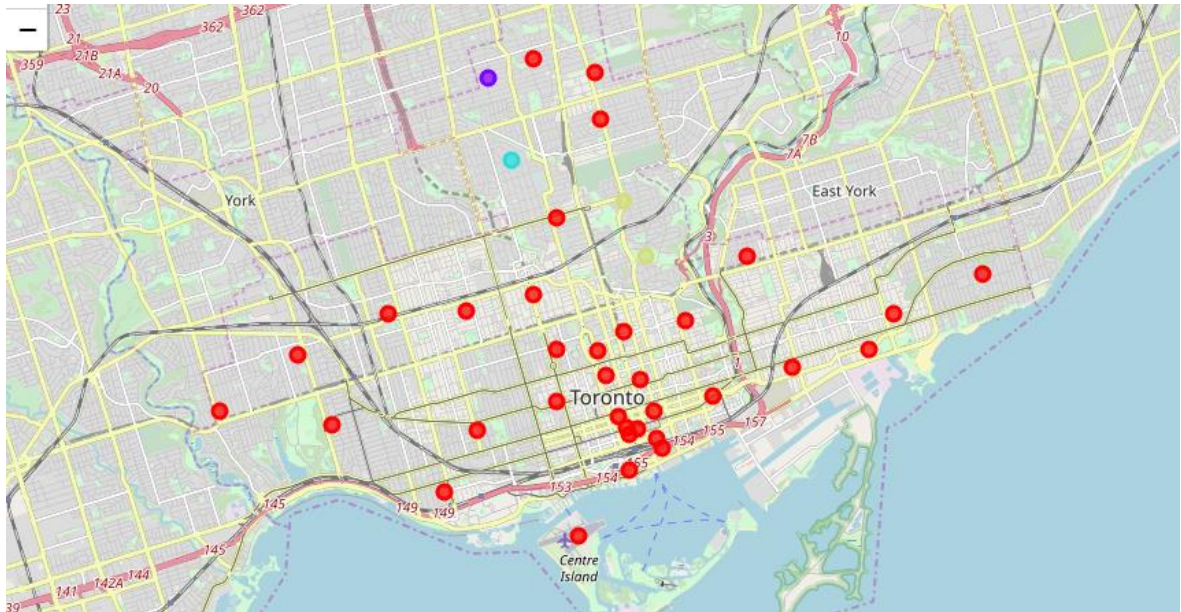
	PostalCode	Borough	Neighborhood
0	M1B	Scarborough	Rouge, Malvern
1	M1C	Scarborough	Highland Creek, Rouge Hill, Port Union
2	M1E	Scarborough	Guildwood, Morningside, West Hill
3	M1G	Scarborough	Woburn
4	M1H	Scarborough	Cedarbrae

- Merge latitude and longitude data to the data frame and show map by folium.

	PostalCode	Borough	Neighborhood	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
5	M1J	Scarborough	Scarborough Village	43.744734	-79.239476
6	M1K	Scarborough	East Birchmount Park, Ionview, Kennedy Park	43.727929	-79.262029
7	M1L	Scarborough	Clairlea, Golden Mile, Oakridge	43.711112	-79.284577
8	M1M	Scarborough	Cliffcrest, Cliffside, Scarborough Village West	43.716316	-79.239476
9	M1N	Scarborough	Birch Cliff, Cliffside West	43.692657	-79.264848
10	M1P	Scarborough	Dorset Park, Scarborough Town Centre, Wexford ...	43.757410	-79.273304



- Then, get 100 venues by Foursquare API and clustering the neighborhood data by K-mean Machine Learning Algorithm.
- I decide to divide cluster in to 4 groups (K = 4).
- Create map by Folium



- After that, the cluster are shown as 4 groups, I label them as most frequency of venues appears.
- Cluster are name as Restaurant, Garden, Jewelry Store, Park.

➤ Restaurant

	Borough	Cluster Labels	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	East Toronto	0	Health Food Store	Asian Restaurant	Pub	Trail	Discount Store	Department Store	Dessert Shop	Dim Sum Restaurant	Diner	Women's Store
1	East Toronto	0	Greek Restaurant	Coffee Shop	Italian Restaurant	Furniture / Home Store	Ice Cream Shop	Cosmetics Shop	Brewery	Bubble Tea Shop	Restaurant	Café
2	East Toronto	0	Park	Sushi Restaurant	Sandwich Place	Liquor Store	Burrito Place	Italian Restaurant	Fast Food Restaurant	Ice Cream Shop	Steakhouse	Fish & Chips Shop
3	East Toronto	0	Café	Coffee Shop	Gastropub	Bakery	Brewery	Italian Restaurant	American Restaurant	Bookstore	Sandwich Place	Cheese Shop
5	Central Toronto	0	Convenience Store	Gym	Department Store	Hotel	Sandwich Place	Food & Drink Shop	Breakfast Spot	Pizza Place	Park	Doner Restaurant
6	Central Toronto	0	Clothing Store	Coffee Shop	Yoga Studio	Sporting Goods Shop	Salon / Barbershop	Restaurant	Rental Car Location	Café	Chinese Restaurant	Mexican Restaurant
7	Central Toronto	0	Sandwich Place	Dessert Shop	Sushi Restaurant	Gym	Italian Restaurant	Café	Pizza Place	Coffee Shop	Pharmacy	Salon / Barbershop

➤ Garden

	Borough	Cluster Labels	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
22	Central Toronto	1	Garden	Women's Store	Deli / Bodega	Electronics Store	Eastern European Restaurant	Dumpling Restaurant	Donut Shop	Doner Restaurant	Dog Run	Distribution Center

➤ Jewelry Store

	Borough	Cluster Labels	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
23	Central Toronto	2	Jewelry Store	Trail	Bus Line	Sushi Restaurant	Women's Store	Department Store	Eastern European Restaurant	Dumpling Restaurant	Donut Shop	Doner Restaurant

➤ Park

	Borough	Cluster Labels	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
4	Central Toronto	3	Park	Lake	Bus Line	Swim School	Department Store	Eastern European Restaurant	Dumpling Restaurant	Donut Shop	Doner Restaurant	Dog Run
8	Central Toronto	3	Trail	Park	Playground	Summer Camp	Discount Store	Department Store	Dessert Shop	Dim Sum Restaurant	Diner	Women's Store
10	Downtown Toronto	3	Park	Playground	Trail	Cupcake Shop	Dumpling Restaurant	Donut Shop	Doner Restaurant	Dog Run	Distribution Center	Discount Store

4. Discussion and Conclusion

The clusters using K-means make similar venue type of neighborhood.

- Toronto is the best place for who want to enjoy with many restaurants
- Most of them are coffee shop.
- For Center Toronto have park and garden.