

Coursera

Opening a new Bakery in the Toronto city

IBM DATA SCIENCE CAPSTONE

by Khyati Jha

Feb 2020

Business Problem

The objective of this capstone project is to analyse and select the best locations in the city of Toronto, Canada to open a new Bakery. Using data science methodology and machine learning techniques like clustering, this project aims to provide solutions to answer the business question: In the city of Toronto, Canada, if someone is looking to open a new Bakery, where would you recommend that they open it?

To solve the problem, we will need the following data:

1. List of neighborhoods in Toronto. This defines the scope of this project which is confined to the city.
2. Latitude and longitude coordinates of those neighborhoods. This is required in order to plot the map and also to get the venue data.
3. Venue data, particularly data related to Bakery. We will

Method

- Web scraping using BeautifulSoup
- Used Geocoder module for latitude and longitude
- Used Foursquare api to get venue data
- Calculated frequency of venues
- Filtered venue by Bakery
- Clustering using k-means clustering
- Visualization using folium

Results

The results from the k-means clustering show that we can categorize the neighbourhoods into 3 clusters based on the frequency of occurrence for “Bakery”:

Cluster 0: Neighborhoods with almost no bakeries.

Cluster 1: Neighborhoods with good number of Bakeries

Cluster 2: Neighborhoods with highest concentration of Bakeries

Cluster 3: Neighborhoods with Low concentration of Bakeries

Cluster 0

	Neighborhoods	Bakery	clusters	PostalCode	Borough	Latitude	Longitude
29	Roselawn	0.0	0	M5N	Central Toronto	43.711695	-79.416936
17	Forest Hill North, Forest Hill West	0.0	0	M5P	Central Toronto	43.696948	-79.411307
22	Lawrence Park	0.0	0	M4N	Central Toronto	43.728020	-79.388790
23	Little Portugal, Trinity	0.0	0	M6J	West Toronto	43.647927	-79.419750
28	Rosedale	0.0	0	M4W	Downtown Toronto	43.679563	-79.377529
13	Deer Park, Forest Hill SE, Rathnelly, South Hi...	0.0	0	M4V	Central Toronto	43.686412	-79.400049
12	Davisville North	0.0	0	M4P	Central Toronto	43.712751	-79.390197
11	Davisville	0.0	0	M4S	Central Toronto	43.704324	-79.388790
37	The Beaches West, India Bazaar	0.0	0	M4L	East Toronto	43.668999	-79.315572
24	Moore Park, Summerhill East	0.0	0	M4T	Central Toronto	43.689574	-79.383160
8	Christie	0.0	0	M6G	Downtown Toronto	43.669542	-79.422564
25	North Toronto West	0.0	0	M4R	Central Toronto	43.715383	-79.405678
35	The Annex, North Midtown, Yorkville	0.0	0	M5R	Central Toronto	43.672710	-79.405678
26	Parkdale, Roncesvalles	0.0	0	M6R	West Toronto	43.648960	-79.456325
4	CN Tower, Bathurst Quay, Island airport, Harbo...	0.0	0	M5V	Downtown Toronto	43.628947	-79.394420
3	Business Reply Mail Processing Centre 969 Eastern	0.0	0	M7Y	East Toronto	43.662744	-79.321558
27	Queen's Park	0.0	0	M7A	Downtown Toronto	43.662301	-79.389494
36	The Beaches	0.0	0	M4E	East Toronto	43.676357	-79.293031
9	Church and Wellesley	0.0	0	M4Y	Downtown Toronto	43.665860	-79.383160
30	Runnymede, Swansea	0.0	0	M6S	West Toronto	43.651571	-79.484450

Cluster 1,2

	Neighborhoods	Bakery	clusters	PostalCode	Borough	Latitude	Longitude
21	High Park, The Junction South	0.043478	1	M6P	West Toronto	43.661608	-79.464763
19	Harbourfront	0.065217	1	M5A	Downtown Toronto	43.654260	-79.360636
34	Studio District	0.048780	1	M4M	East Toronto	43.659526	-79.340923
5	Cabbagetown, St. James Town	0.044444	1	M4X	Downtown Toronto	43.667967	-79.367675
2	Brockton, Exhibition Place, Parkdale Village	0.047619	1	M6K	West Toronto	43.636847	-79.428191
18	Harbord, University of Toronto	0.057143	1	M5S	Downtown Toronto	43.662696	-79.400049

	Neighborhoods	Bakery	clusters	PostalCode	Borough	Latitude	Longitude
15	Dovercourt Village, Dufferin	0.142857	2	M6H	West Toronto	43.669005	-79.442259

Cluster 3

	Neighborhoods	Bakery	clusters	PostalCode	Borough	Latitude	Longitude
31	Ryerson, Garden District	0.020000	3	M5B	Downtown Toronto	43.657162	-79.378937
33	Stn A PO Boxes 25 The Esplanade	0.021053	3	M5W	Downtown Toronto	43.646435	-79.374846
32	St. James Town	0.030000	3	M5C	Downtown Toronto	43.651494	-79.375418
0	Adelaide, King, Richmond	0.030000	3	M5H	Downtown Toronto	43.650571	-79.384568
16	First Canadian Place, Underground city	0.020000	3	M5X	Downtown Toronto	43.648429	-79.382280
14	Design Exchange, Toronto Dominion Centre	0.020000	3	M5K	Downtown Toronto	43.647177	-79.381576
10	Commerce Court, Victoria Hotel	0.020000	3	M5L	Downtown Toronto	43.648198	-79.379817
7	Chinatown, Grange Park, Kensington Market	0.024691	3	M5T	Downtown Toronto	43.653206	-79.400049
6	Central Bay Street	0.024390	3	M5G	Downtown Toronto	43.657952	-79.387383
1	Berczy Park	0.036364	3	M5E	Downtown Toronto	43.644771	-79.373306
20	Harbourfront East, Toronto Islands, Union Station	0.020000	3	M5J	Downtown Toronto	43.640816	-79.381752
38	The Danforth West, Riverdale	0.024390	3	M4K	East Toronto	43.679557	-79.352188

Refrence

Category:Suburbs in Toronto. *Wikipedia*. Retrieved from
(https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)

Foursquare Developers Documentation. *Foursquare*. Retrieved
from <https://developer.foursquare.com/docs>