Coursera Capstone Project
The Battle of the Neighborhoods (Week 2)

Table of contents

Anna Fabiola Cedeno Maselli
Data Science Professional Certificate

03/02/2021

1. Introduction: Business Problem

An Italian firm located in Texas City, United States, decides to move its headquarters to New York City or Toronto, Canada. They don't know which city is the best for them. The firm wants to know local businesses and neighborhoods to locate the company. The project will analyze the neighborhoods between New York City and Toronto, understand the differences and similarities, group the neighborhoods, visualize these groups on a map, and provide the best decision.

The target audience are the investors interested in moving their headquarters to the best city and may need an objective advice to choose the location for the company and its employees.

2. Data and Methodology

a. The websites that collect the information about Toronto and New York borough and their locations. The pages are: https://geo.nyu.edu/catalog/nyu_2451_34572 and https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M.

b. The Foursquare API, that collect venues and their categories for each location within a radius of 700 meters.

c. The Geopy and Folium libraries to get the coordinates of every location of Toronto and New York City. The page is https://cocl.us/Geospatial_data.

d. Cluster venues of each neighborhood using k-means algorithm and analyze the top 10 most common venue in each cluste.

2. Data and Methodology

e. Visualize clusters on the map, thus showing the best locations.

f. The pandas library used for data manipulation and analysis.

g. The Numpy library used to work with arrays.

h. The Requests used to send HTTP/1.1 requests.

i. Matplotlib library used to create static, animated, and interactive visualizations in Python.

j. The json used to transfer data as text that can be sent over a network.

k. The Urllib used to fetch URLs (Uniform Resource Locators).

3. Results

Downtown Toronto

Cluster 0

	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
count	8	8.0	8	8	8	8	8	8	8	8	8	8
unique	1	NaN	2	3	6	8	7	6	6	6	7	6
top	Downtown Toronto	NaN	Coffee Shop	Café	Café	Gastropub	Restaurant	Japanese Restaurant	Gym	Seafood Restaurant	American Restaurant	Pizza Place
freq	8	NaN	6	3	3	1	2	2	3	3	2	2
mean	NaN	0.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Cluster 1

	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
count	1	1.0	1	1	1	1	1	1	1	1	1	1
unique	1	NaN	1	1	1	1	1	1	1	1	1	1
top	Downtown Toronto	NaN	Park	Playground	Gym / Fitness Center	Trail	Yoga Studio	Discount Store	Falafel Restaurant	Event Space	Ethiopian Restaurant	Electronics Store
freq	1	NaN	1	1	1	1	1	1	1	1	1	1
mean	NaN	1.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

3. Results

Cluster 2

	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
count	1	1.0	1	1	1	1	1	1	1	1	1	1
unique	1	NaN	1	1	1	1	1	1	1	1	1	1
top	Downtown Toronto	NaN	Airport Service	Airport Terminal	Rental Car Location	Coffee Shop	Harbor / Marina	Airport Lounge	Sculpture Garden	Bar	Music Venue	Pier
freq	1	NaN	1	1	1	1	1	1	1	1	1	1
mean	NaN	2.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

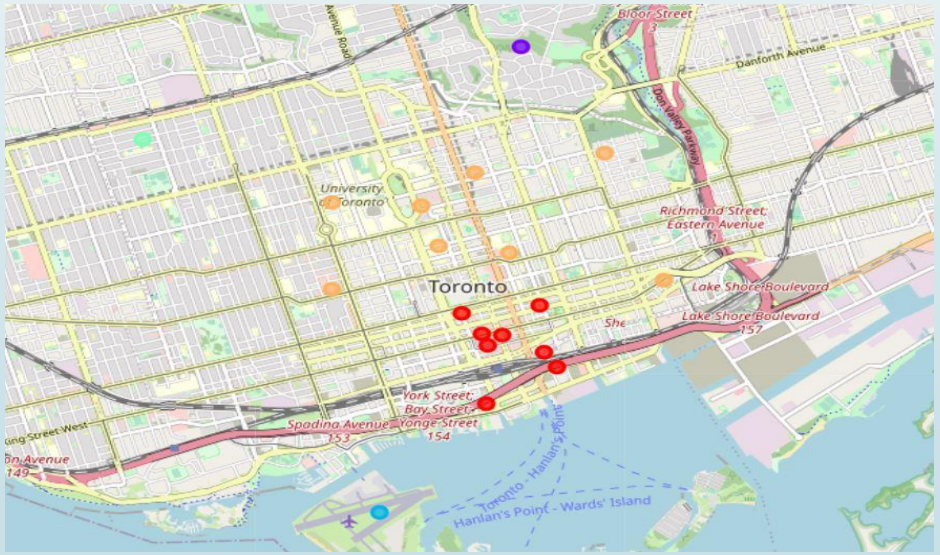
Cluster 3

	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
count	1	1.0	1	1	1	1	1	1	1	1	1	1
unique	1	NaN	1	1	1	1	1	1	1	1	1	1
top	Downtown Toronto	NaN	Grocery Store	Café	Park	Coffee Shop	Candy Store	Bakery	Playground	Beer Store	Italian Restaurant	Nightclub
freq	1	NaN	1	1	1	1	1	1	1	1	1	1
mean	NaN	3.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

3. Results

Cluster 4

	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
count	8	8.0	8	8	8	8	8	8	8	8	8	8
unique	1	NaN	2	8	6	7	7	8	8	6	6	8
top	Downtown Toronto	NaN	Coffee Shop	Japanese Restaurant	Café	Bakery	Gastropub	Mexican Restaurant	Bubble Tea Shop	Ramen Restaurant	Diner	Korean Restaurant
freq	8	NaN	6	1	3	2	2	1	1	2	2	1
mean	NaN	4.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN



Cluster 0 (Red dots)

Cluster 1(Purple dots)

Cluster 2 (Blue dots)

Cluster 3 (Green dots)

Cluster 4 (Orange dots)

3. Results

Queens, New York City

Cluster 0

	Neighborhood	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
count	12	12	12	12	12	12	12	12	12	12	12
unique	12	7	11	9	10	11	10	11	12	9	10
top	South Jamaica	Deli / Bodega	Deli / Bodega	Deli / Bodega	Sandwich Place	Food Truck	Caribbean Restaurant	Bagel Shop	Mediterranean Restaurant	Chinese Restaurant	Sandwich Place
freq	1	5	2	2	2	2	2	2	1	3	2

Cluster 1

	Neighborhood	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
count	2	2	2	2	2	2	2	2	2	2	2
unique	2	1	2	2	2	2	2	2	2	2	2
top	Neponsit	Beach	Park	Pizza Place	Zoo	Pharmacy	Falafel Restaurant	Mexican Restaurant	Bagel Shop	Fast Food Restaurant	Donut Shop
freq	1	2	1	1	1	1	1	1	1	1	1

3. Results

Cluster 2

	Neighborhood	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
count	1	1	1	1	1	1	1	1	1	1	1
unique	1	1	1	1	1	1	1	1	1	1	1
top	Floral Park	Indian Restaurant	Ice Cream Shop	Grocery Store	Pizza Place	Fast Food Restaurant	Gift Shop	Bank	Bagel Shop	Dosa Place	Donut Shop
freq	1	1	1	1	1	1	1	1	1	1	1

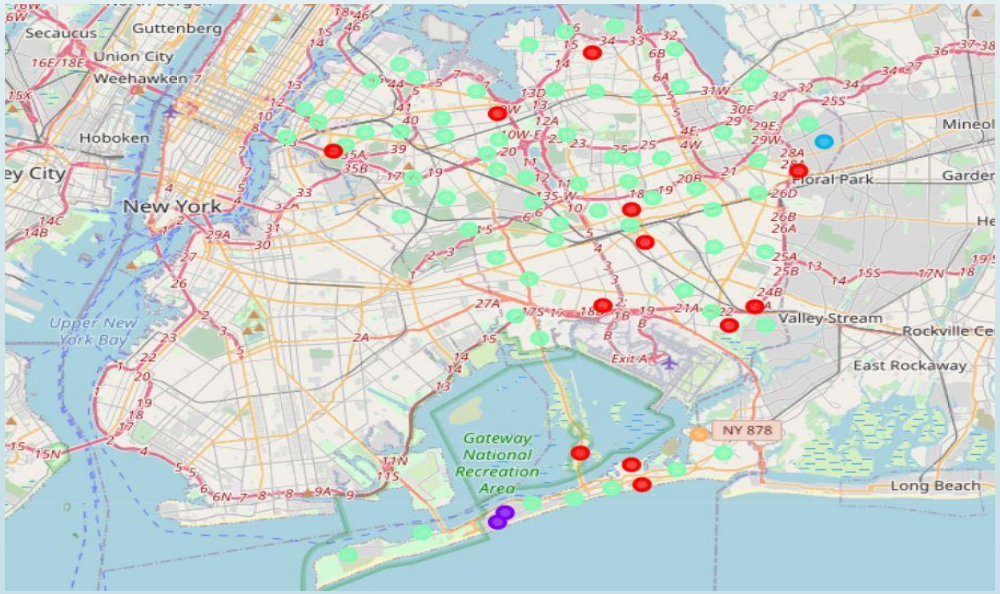
Cluster 3

	Neighborhood	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
count	65	65	65	65	65	65	65	65	65	65	65
unique	65	31	37	39	38	45	47	46	48	49	49
top	Pomonok	Pizza Place	Pizza Place	Pizza Place	Bakery	Grocery Store	Sandwich Place	Chinese Restaurant	Chinese Restaurant	Donut Shop	Bank
freq	1	11	8	7	6	5	4	4	4	4	3

3. Results

Cluster 4

	Neighborhood	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
count	1	1	1	1	1	1	1	1	1	1	1
unique	1	1	1	1	1	1	1	1	1	1	1
top	Bayswater	Playground	Indian Restaurant	Tennis Court	Park	Construction & Landscaping	Cycle Studio	Filipino Restaurant	Cosmetics Shop	Event Space	Falafel Restaurant
freq	1	1	1	1	1	1	1	1	1	1	1



Cluster 0 (Red dots)

Cluster 1(Purple dots)

Cluster 2 (Blue dots)

Cluster 3 (Green dots)

Cluster 4 (Orange dots)

4. Discussion

- Downtown Toronto has 19 neighborhoods and 222 venues. The most common venues are: Coffee shop, Grocery Store, Park and Airport service.

- Queens borough has 81 neighborhoods and 304 venues. The most common venues are: Deli/Bodega, Beach, Indian Restaurant, Pizza Place and playground.

5. Conclusion

Both cities are good but based on the quantity of venues and neighborhoods is the best for the Italian firm to choose Queens over Downtown Toronto to move its headquarters because offer more options for the company and its employees.