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	3	3	5	6	6	8	6	6	7	7
top	Downtown Toronto	NaN	Coffee Shop	Hotel	Café	Restaurant	Japanese Restaurant	Asian Restaurant	Seafood Restaurant	Gym	Theater	American Restaurant
freq	8	NaN	5	4	4	3	3	1	2	3	2	2

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	Trail	Dog Run	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Event Space	Ethiopian Restaurant
freq	1	NaN	1	1	1	1	1	1	1	1	1	1

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	Rental Car Location	Airport Service	Airport Terminal	Boat or Ferry	Sculpture Garden	Coffee Shop	Harbor / Marina	Airport Lounge	Airport Gate	Airport Food Court
freq	1	NaN	1	1	1	1	1	1	1	1	1	1

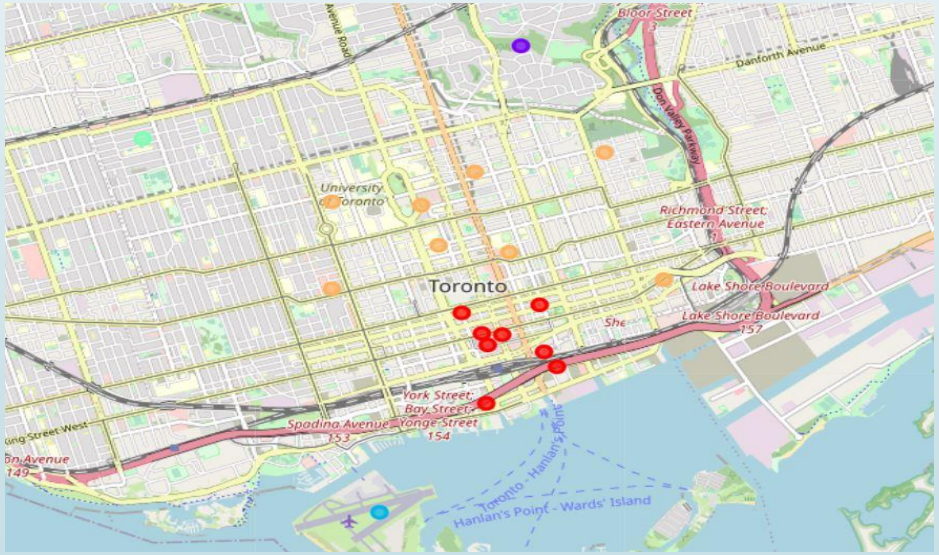
Cluster 3

	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	Restaurant	Bakery	Nightclub	Baby Store	Candy Store	Athletics & Sports
freq	1	NaN	1	1	1	1	1	1	1	1	1	1

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	8	8	7	8	7	8	7	7
top	Downtown Toronto	NaN	Coffee Shop	Japanese Restaurant	Café	Coffee Shop	Sandwich Place	French Restaurant	Pizza Place	Gastropub	Gym	Italian Restaurant
freq	8	NaN	6	1	1	1	2	1	2	1	2	2



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	23	23	23	23	23	23	23	23	23	23	23
unique	23	13	19	15	20	19	18	18	20	20	21
top	South Ozone Park	Pizza Place	Pizza Place	Deli / Bodega	Park	Deli / Bodega	Sandwich Place	Donut Shop	Supermarket	Metro Station	Supermarket
freq	1	6	3	4	2	3	3	4	3	2	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	1	1	1	1	1	1	1	1	1	1	1
unique	1	1	1	1	1	1	1	1	1	1	1
top	Neponsit	Beach	Park	Pizza Place	Zoo	Event Space	Falafel Restaurant	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant
freq	1	1	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	Bayswater	Playground	Park	Men's Store	Athletics & Sports	Event Space	Falafel Restaurant	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant
freq	1	1	1	1	1	1	1	1	1	1	1

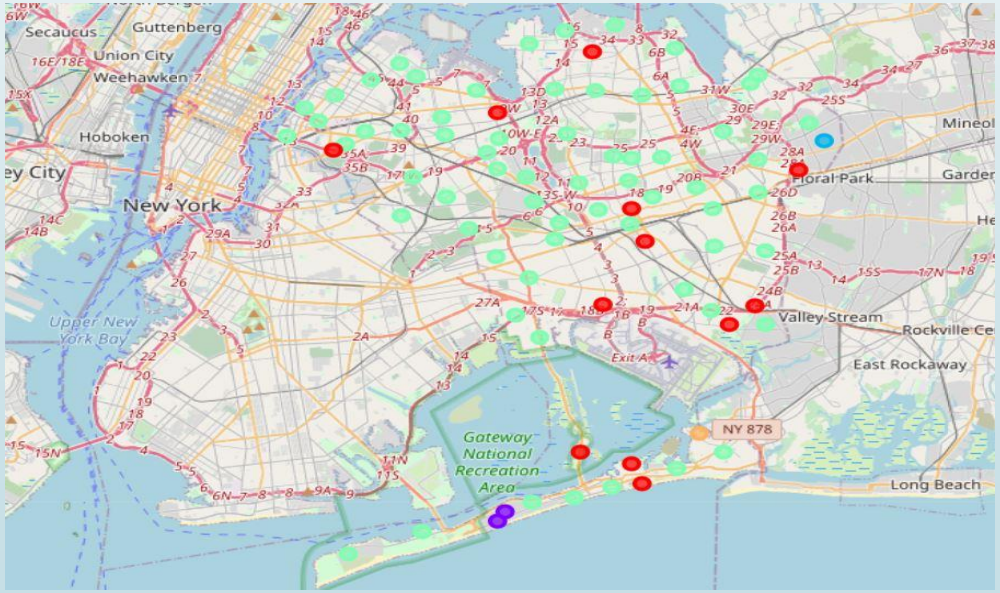
Cluster 3

	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	52	52	52	52	52	52	52	52	52	52	52
unique	52	29	33	35	37	40	37	38	37	41	43
top	Pomonok	Chinese Restaurant	Pizza Place	Bakery	Donut Shop	Bakery	Sandwich Place	Bar	Fast Food Restaurant	Chinese Restaurant	Sandwich Place
freq	1	5	5	5	4	3	3	5	3	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	4	4	4	4	4	4	4	4	4	4	4
unique	4	1	4	3	4	4	4	4	4	3	4
top	Hammels	Beach	Supermarket	Donut Shop	Dog Run	Bar	Beach Bar	Mexican Restaurant	Bagel Shop	Fast Food Restaurant	Seafood Restaurant
freq	1	4	1	2	1	1	1	1	1	2	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 224 venues. The most common venues are:

- Queens borough has 81 neighborhoods and 302 venues. The most common venues are:

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.