

REPORT

Capstone Project - The Battle of Neighborhoods - Relocation Helper

Github Link:

https://github.com/victoriturbe2005/Coursera_Capstone/blob/master/The%20Battle%20of%20Neighborhoods.ipynb

1. INTRODUCTION:

a. Background

It is often that people decide to move to another city due to a new job or maybe other reasons. Going through this process presents a list of challenges. One of these challenges is locating an area that we like. Even though the distance from the workplace would be a criterion, people would also consider culture to be an important criteria to reflect upon. Location culture could translate into local gatherings where people meet to eat, drink, etc. Therefore, this code helps people find an area cluster) in which they could be interested in based on "most common venues". In this case we have chosen two cities that we have worked with before: New York city and Toronto.

b. Problem

Most online sites that provide information of what to do in certain cities do not break the search down my sections of the city nor offer different options to what the user is looking for. For example: if you are looking for Chinese restaurants around you, a simple google search may give you a list of results. However, it may not give you other possible options such as Italian or Peruvian restaurants around you.

2. Data

a. Data Sources

New York City and Toronto datasets were modified to include Boroughs, Neighborhoods, Latitude, and Longitude.

Toronto data source: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

New York data source: https://cocl.us/new_york_dataset

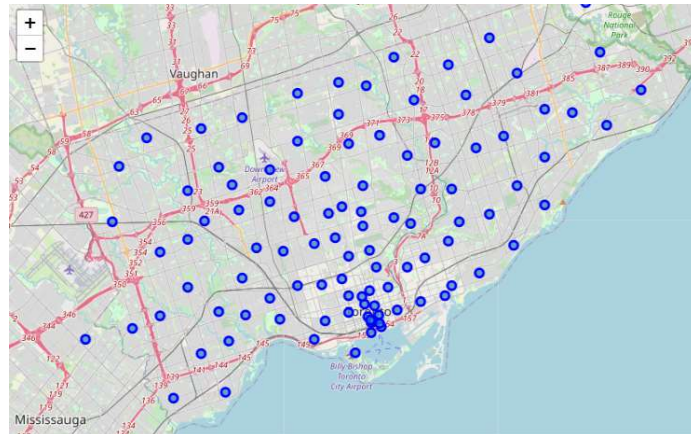
b. Methodology

Toronto Data

The Toronto data source provided a list of postal codes, Borough, and Neighborhoods; a Geospatial data set provided Latitudes and Longitudes for these Postal Codes, the both data frames were matched and merged.

	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	M1B	Scarborough	Malvern, Rouge	43.806686	-79.194353
1	M1C	Scarborough	Rouge Hill, Port Union, Highland Creek	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

Python folium library was used to create a map of Toronto with boroughs superimposed on top.



Foursquare API was used to get the most common venues based on Latitude and Longitude with a radius of 750 meters.

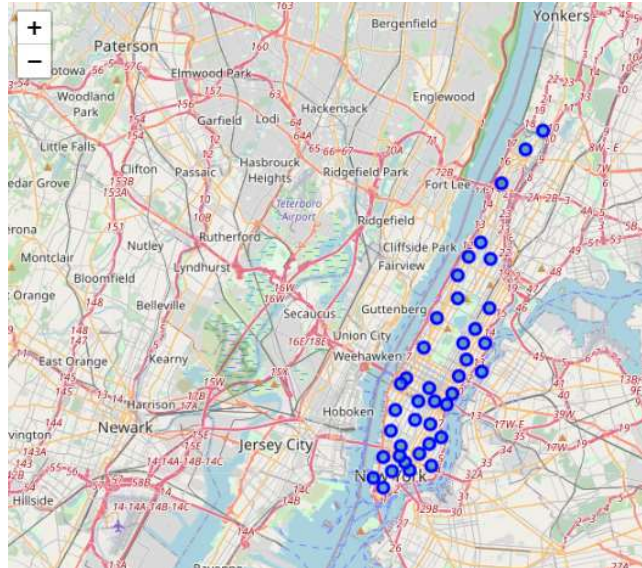
	Postal_Code	Postal_Latitude	Postal_Longitude	Venue	Venue_Latitude	Venue_Longitude	Venue_Category
0	M1B	43.806686	-79.194353	Wendy's	43.807448	-79.199056	Fast Food Restaurant
1	M1B	43.806686	-79.194353	Wendy's	43.802008	-79.198080	Fast Food Restaurant
2	M1B	43.806686	-79.194353	Staples Morningside	43.800285	-79.196607	Paper / Office Supplies Store
3	M1B	43.806686	-79.194353	Tim Hortons	43.802000	-79.198169	Coffee Shop
4	M1B	43.806686	-79.194353	Lee Valley	43.803161	-79.199681	Hobby Shop

Then the top 5 most common venues were ranked for each Borough. Also, clusters were assigned to later help with the relocation of the user.

	PostalCode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	M1B	Scarborough	Malvern, Rouge	43.806686	-79.194353	4	Fast Food Restaurant	Coffee Shop	Hobby Shop	Bus Station	Construction & Landscaping
1	M1C	Scarborough	Rouge Hill, Port Union, Highland Creek	43.784535	-79.160497	4	Breakfast Spot	Bar	Italian Restaurant	Burger Joint	Yoga Studio
2	M1E	Scarborough	Guildwood, Morningside, West Hill	43.763573	-79.188711	4	Fast Food Restaurant	Pizza Place	Sports Bar	Beer Store	Bank
3	M1G	Scarborough	Woburn	43.770992	-79.216917	0	Coffee Shop	Park	Business Service	Construction & Landscaping	Electronics Store
4	M1H	Scarborough	Cedarbrae	43.773136	-79.239476	4	Indian Restaurant	Coffee Shop	Yoga Studio	Bank	Gas Station

New York Data

The New York data source was a repository which already included a list of Boroughs and Neighborhoods along with Latitude and Longitude. Then, Manhattan was segmented as the neighborhood of choice of relocation.



Foursquare API was used to get the top 5 most common venues were ranked for each Neighborhood.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Battery Park City	Park	Hotel	Coffee Shop	Wine Shop	Memorial Site
1	Carnegie Hill	Coffee Shop	Pizza Place	Yoga Studio	Gym	Wine Shop
2	Central Harlem	African Restaurant	Seafood Restaurant	Bar	French Restaurant	Art Gallery
3	Chelsea	Art Gallery	Coffee Shop	Café	American Restaurant	Juice Bar
4	Chinatown	Chinese Restaurant	Cocktail Bar	Bakery	Optical Shop	Salon / Barbershop

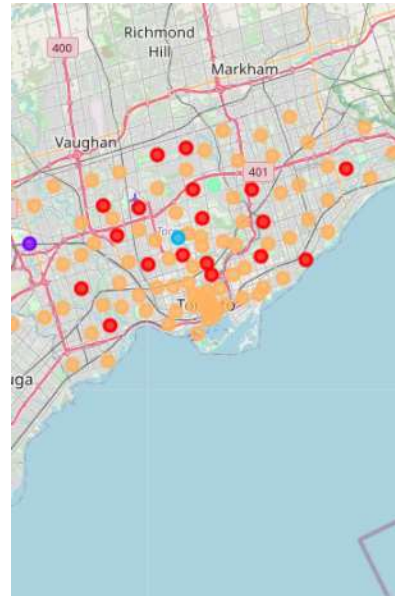
Clustering

5 clusters were created running k-means for Neighborhoods in New York city (Manhattan) and boroughs in Toronto. Then maps were created for both.

New York City clusters



Toronto clusters



NYC Cluster 2

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
17	Chelsea	Art Gallery	Coffee Shop	Café	American Restaurant	Juice Bar
27	Gramercy	Bagel Shop	Italian Restaurant	Bar	Coffee Shop	Playground
22	Little Italy	Chinese Restaurant	Italian Restaurant	Mediterranean Restaurant	Spa	Bakery
15	Midtown	Coffee Shop	Hotel	Theater	Clothing Store	Bakery
35	Turtle Bay	Coffee Shop	Italian Restaurant	Café	Wine Bar	Park
29	Financial District	Coffee Shop	Pizza Place	Hotel	American Restaurant	Sandwich Place
30	Carnegie Hill	Coffee Shop	Pizza Place	Yoga Studio	Gym	Wine Shop
25	Manhattan Valley	Coffee Shop	Spa	Pizza Place	Bar	Hawaiian Restaurant
38	Flatiron	Gym / Fitness Center	Italian Restaurant	American Restaurant	Wine Shop	Salon / Barbershop
39	Hudson Yards	Hotel	American Restaurant	Gym / Fitness Center	Café	Italian Restaurant
16	Murray Hill	Hotel	Sandwich Place	Coffee Shop	Gym / Fitness Center	Steakhouse
8	Upper East Side	Italian Restaurant	Bakery	Gym / Fitness Center	Juice Bar	Spa
9	Yorkville	Italian Restaurant	Coffee Shop	Gym	Deli / Bodega	Bar
18	Greenwich Village	Italian Restaurant	Coffee Shop	Gym	Ice Cream Shop	Bakery
34	Sutton Place	Italian Restaurant	Coffee Shop	Gym / Fitness Center	Park	Bagel Shop
31	Noho	Italian Restaurant	Coffee Shop	Japanese Restaurant	Grocery Store	Sandwich Place
10	Lenox Hill	Italian Restaurant	Coffee Shop	Pizza Place	Café	Cocktail Bar
23	Soho	Italian Restaurant	Mediterranean Restaurant	Gym	French Restaurant	Coffee Shop
21	Tribeca	Italian Restaurant	Park	Wine Bar	Café	Spa
13	Lincoln Square	Italian Restaurant	Plaza	Café	Gym / Fitness Center	Theater
24	West Village	Italian Restaurant	Wine Bar	American Restaurant	Coffee Shop	Park
12	Upper West Side	Italian Restaurant	Wine Bar	Bakery	Dessert Shop	Coffee Shop
33	Midtown South	Korean Restaurant	Hotel	Japanese Restaurant	Burger Joint	Lounge
32	Civic Center	Spa	American Restaurant	Hotel	French Restaurant	Coffee Shop
14	Clinton	Theater	Coffee Shop	Gym / Fitness Center	Gym	Wine Shop

NYC Cluster 3

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
6	Central Harlem	African Restaurant	Seafood Restaurant	Bar	French Restaurant	Art Gallery
2	Washington Heights	Café	Bakery	Mobile Phone Shop	Chinese Restaurant	Latin American Restaurant
1	Chinatown	Chinese Restaurant	Cocktail Bar	Bakery	Optical Shop	Salon / Barbershop
20	Lower East Side	Chinese Restaurant	Pharmacy	Art Gallery	Café	Cocktail Bar
5	Manhattanville	Coffee Shop	Chinese Restaurant	Seafood Restaurant	Italian Restaurant	Deli / Bodega
7	East Harlem	Mexican Restaurant	Bakery	Thai Restaurant	Latin American Restaurant	Deli / Bodega
3	Inwood	Mexican Restaurant	Café	Lounge	Pizza Place	Restaurant
36	Tudor City	Park	Café	Deli / Bodega	Mexican Restaurant	Pizza Place
28	Battery Park City	Park	Hotel	Coffee Shop	Wine Shop	Memorial Site
11	Roosevelt Island	Park	Pizza Place	Gym / Fitness Center	Greek Restaurant	Coffee Shop
19	East Village	Pizza Place	Cocktail Bar	Coffee Shop	Bar	Wine Bar

3. Results

The table above shows two out of the five clusters that were created. NYC cluster 2 shows “Italian Restaurants” as the most common venue often along with Coffee shops and bakeries.

For someone moving from Toronto to New York city and is looking for a *Terra Nostra* atmosphere on of the neighborhoods in this cluster could be a good choice.

On the other hand, someone looking for more *Salsa* may want to pick a neighborhood with cluster 3 given the number of Mexican and Latin venues.

4. Conclusion

It is much easier to pick a location once it is clustered including a few neighborhoods. Also, it is much easier to have a pretty good idea of a local culture based on venues. Even though, New York city would be considered a melting pot, some people may prefer to be close to his/her roots.

Your choice of a presentation or blogpost. **(10 marks)**