

The Battle of Neighborhoods (Week 2)

1. Introduction


The travel agency would like to open a travel package to bring tourist into Toronto City so the travel agency have to understand in this city for creating travel package and finding area where able to be response to tourist by that area should have restaurants, cafe, museum and shopping mall.

For traveler that want to visit, travel or person who have to be in Toronto.

2. Data

To consider the objective, I listed the below data sources used for the analysis.

a) Toronto Neighborhood Data: The following Wikipedia page was scraped to pull out the necessary information: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M



The screenshot shows the Wikipedia page for 'List of postal codes of Canada: M'. The page includes a sidebar with navigation links, a search bar, and a table of postal codes and neighborhoods. The table has three columns: Postcode, Borough, and Neighbourhood. The data is as follows:

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
M6A	North York	Lawrence Heights
M6A	North York	Lawrence Manor

The information obtained i.e. the table of postal codes was transformed into a pandas data frame for further analysis.

b) Coordinate data for each Neighborhood in Toronto: The following csv file gave us the geographical coordinates of each postal code: http://cocl.us/Geospatial_data

3. Methodology

3.1 Load Toronto Neighborhood Data and transformed into a pandas data frame for further analysis.

	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

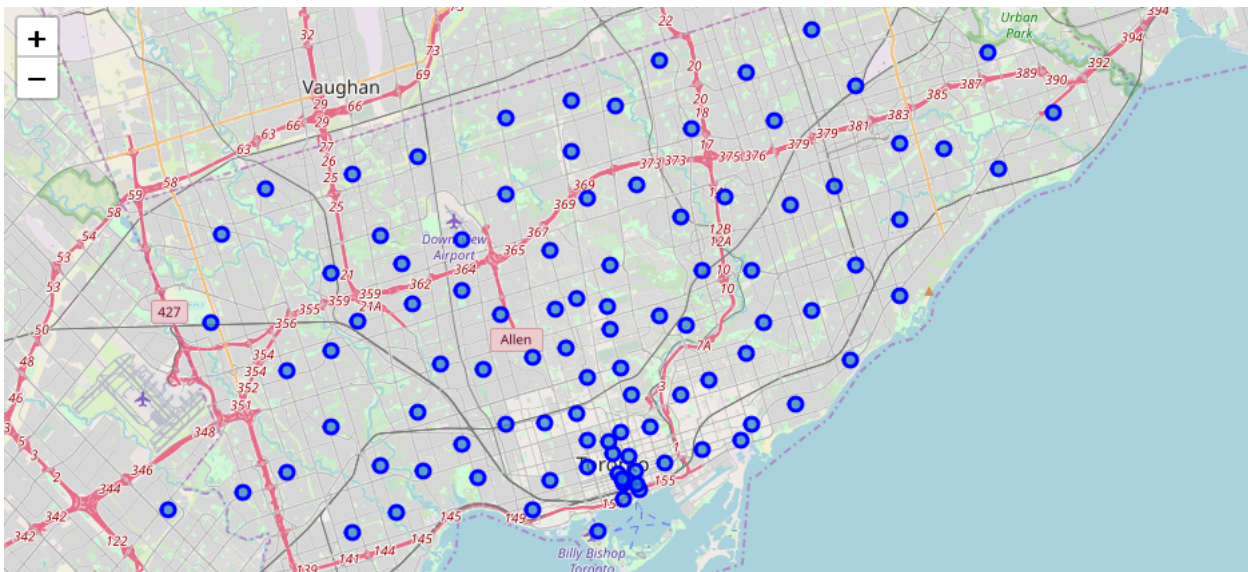
3.2 Load Coordinate data for each Neighborhood in Toronto and put them to data frame.

	Postal Code	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

3.3 Merge 2 data frame together.

	Postal Code	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

3.4 Explore and cluster the neighborhoods in Toronto.



3.5 Define Foursquare

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Rouge, Malvern	43.806686	-79.194353	Wendy's	43.807448	-79.199056	Fast Food Restaurant
1	Highland Creek, Rouge Hill, Port Union	43.784535	-79.160497	Royal Canadian Legion	43.782533	-79.163085	Bar
2	Highland Creek, Rouge Hill, Port Union	43.784535	-79.160497	Scarborough Historical Society	43.788755	-79.162438	History Museum
3	Guildwood, Morningside, West Hill	43.763573	-79.188711	Swiss Chalet Rotisserie & Grill	43.767697	-79.189914	Pizza Place
4	Guildwood, Morningside, West Hill	43.763573	-79.188711	G & G Electronics	43.765309	-79.191537	Electronics Store

3.6 Analysing Each Neighbourhood.

----Adelaide, King, Richmond----		
	venue	freq
0	Coffee Shop	0.06
1	Bar	0.04
2	Café	0.04
3	Steakhouse	0.04
4	Thai Restaurant	0.04
----Agincourt----		
	venue	freq
0	Skating Rink	0.2
1	Breakfast Spot	0.2
2	Lounge	0.2
3	Latin American Restaurant	0.2
4	Clothing Store	0.2

3.7 Take result to data frame.

	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
0	Adelaide, King, Richmond	Coffee Shop	Bar	Café	Thai Restaurant	Steakhouse	Restaurant	Bakery	Re
1	Agincourt	Clothing Store	Latin American Restaurant	Skating Rink	Lounge	Breakfast Spot	Falafel Restaurant	Event Space	
2	Agincourt North, L'Amoreaux East, Milliken, St...	Playground	Park	Yoga Studio	Dumpling Restaurant	Discount Store	Dog Run	Doner Restaurant	
3	Albion Gardens, Beaumont Heights, Humbergate, ...	Grocery Store	Pizza Place	Fried Chicken Joint	Coffee Shop	Beer Store	Pharmacy	Fast Food Restaurant	S
4	Alderwood, Long Branch	Pizza Place	Pub	Pharmacy	Gym	Sandwich Place	Coffee Shop	Skating Rink	

3.8 Answer the question of this project.

The travel agency can create the travel package from catogorie of 1st Most Common Venue

```
1 toronto_merged['1st Most Common Venue'].unique()
array(['Fast Food Restaurant', 'History Museum', 'Spa', 'Coffee Shop',
      'Caribbean Restaurant', 'Playground', 'Discount Store', 'Bakery',
      'American Restaurant', 'College Stadium', 'Indian Restaurant',
      'Clothing Store', 'Pizza Place', 'Dog Run', 'Café', 'Cafeteria',
      'Piano Bar', 'Ramen Restaurant', 'Park', 'Gym',
      'Miscellaneous Shop', 'Airport', 'Grocery Store', 'Baseball Field',
      'Liquor Store', 'Pub', 'Sporting Goods Shop', 'Greek Restaurant',
      'Pet Store', 'Dessert Shop', 'Summer Camp', 'Sandwich Place',
      'Garden', 'Mexican Restaurant', 'Bar', 'Airport Lounge',
      'Metro Station', 'Tennis Court', 'Gift Shop', 'Light Rail Station',
      'Furniture / Home Store', 'Rental Car Location'], dtype=object)
```

I assume that the travel package is travel Museum

```
1 toronto_merged.loc[toronto_merged['1st Most Common Venue'] == 'History Museum']
```

	Postal Code	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
1	M1C	Scarborough	Highland Creek, Rouge Hill, Port Union	43.784535	-79.160497	0	History Museum	Bar	Yoga Studio

4. Result

- If travel package is travel Museum, the travel agency bring tourist to Scarborough and walk around the Highland Creek, Rouge Hill and Port Union.
- This project show about if someone would like to use data to creating the travel package, you can do that!

5. Conclusion

- Have to set travel package first and then will able to plan path for tourist
- The accuracy of data depends purely depends on the data provided by FourSquare