

Battle of Neighborhoods

Description of the Problem and Discussion of the Background

Exploring the neighborhoods of Toronto ,India and categorising the top ten venues.

Toronto is the capital of Ontario and the most populous city in Canada, with a population of 2,731,571 in 2016. Current to 2016, the Toronto CMA . Toronto is the fastest growing city in North America, and is the anchor of an urban agglomeration, known as the Golden Horseshoe in Southern , located on the northwestern shore of [Lake Ontario](#). Toronto is an international centre of business, finance, arts, and culture, and is recognized as one of the most [multicultural](#) and [cosmopolitan](#) cities in the world.

Target Audience

1. New person to the city or any tourist, to find reasonable rest place close to his location.
2. Budding Data Scientists, who wants to implement some of the most used Exploratory Data Analysis techniques to obtain necessary data, analyze it and, finally be able to tell a story out of it.



Initial Data Preparation

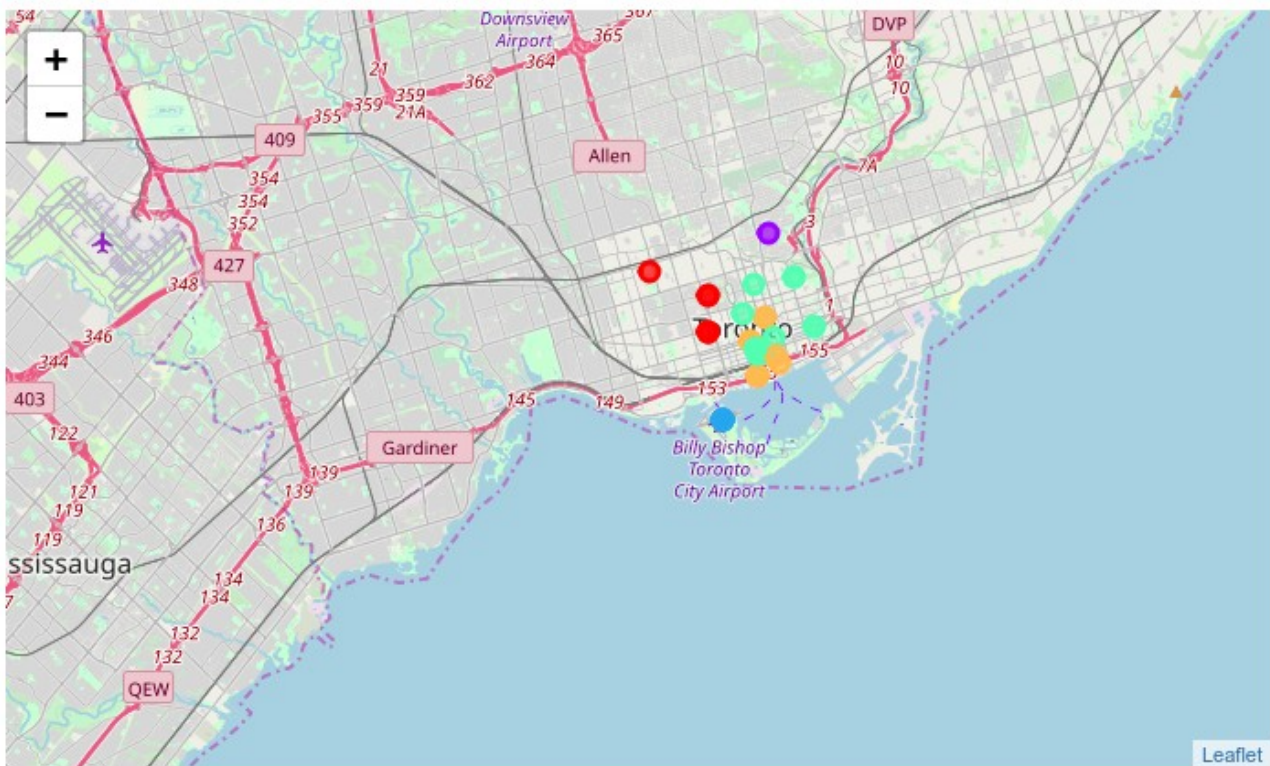
1. use the pandas to obtain the data frame from wikipedia
2. The data frame consist of three columns: PostalCode, Borough, and Neighbourhood
2. ignored the cells with a borough that is not assigned.
3. combinig two rows with neighbourhood seperated by comma.
4. assigning neighbourhood as tha of Borough it it is not assigned.
5. Then i merge up the data set with the locations
6. now the data set is ready..

	Postcode	Borough	Neighbourhood	Latitude	Longitude
0	M3A	North York	Parkwoods	43.753259	-79.329656
1	M4A	North York	Victoria Village	43.725882	-79.315572
2	M5A	Downtown Toronto	Harbourfront	43.654260	-79.360636
3	M5A	Downtown Toronto	Regent Park	43.654260	-79.360636
4	M6A	North York	Lawrence Heights	43.718518	-79.464763
5	M6A	North York	Lawrence Manor	43.718518	-79.464763
6	M7A	Queen's Park	Not assigned	43.662301	-79.389494
7	M9A	Etobicoke	Islington Avenue	43.667856	-79.532242
8	M1B	Scarborough	Rouge	43.806686	-79.194353
9	M1B	Scarborough	Malvern	43.806686	-79.194353

7. Then i create a data frame by doing exploratory Analysis that contains top ten venues of the niegbourhood of the Toronto.

	Postcode	Borough	Neighbourhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
0	M5A	Downtown Toronto	Harbourfront	43.654260	-79.360636	3	Coffee Shop	Bakery	Park	Brea
1	M5A	Downtown Toronto	Regent Park	43.654260	-79.360636	3	Coffee Shop	Bakery	Park	Brea
2	M5B	Downtown Toronto	Ryerson	43.657162	-79.378937	4	Café	Clothing Store	Coffee Shop	Bu
3	M5B	Downtown Toronto	Garden District	43.657162	-79.378937	4	Café	Clothing Store	Burger Joint	
4	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418	3	Coffee Shop	Gastropub	Restaurant	It Restau

8.clustering the neighbourhoods and visualising it using Folium.



Discussion

the most common venues are mexican resturants and cafe
so the people of toronto are very fon of mexican cuisine and
offcourse coffee..

Conclusion

Finally to conclude this project, We have got a small glimpse of
how real life data-science projects look like. I
have made use of some frequently used python libraries to scrap
web-data, use Foursquare API to explore the
major neighbourhood of Toronto and saw the results of
segmentation of districts using Folium leaflet map.