

[https://github.com/sjderrick0127/Coursera_IBM_capstone/blob/master/Battle_of_Neighbourhoods_code_notebook%20\(1\).ipynb](https://github.com/sjderrick0127/Coursera_IBM_capstone/blob/master/Battle_of_Neighbourhoods_code_notebook%20(1).ipynb) Capstone Project -- Battle of Neighbourhoods

1.0: Background

Gloulou investment&Property consulting is a Toronto-based start-up focusing on providing specialized and customized consulting and agent services for people who intends to buy or rent a property. Due to intense market competitions, Gloulou decides to adopt data science techniques including machine learning to offer a brand-new recommendation system for its clients.

1.1: Business Problem

Recently, an old couple approached Gloulou. They wish to move from downtown Toronto to Scarborough and they have the following requirements.

1. They want to keep their lifestyle as before, we should recommend neighbourhoods which are similar to their former one. (Central Bay Street)
2. They want their neighbourhood has variety selection of Chinese restaurant.
3. They want their neighbourhood has grocery stores.

2.0: Data

We will utilize neighbourhood's information of Toronto and Scarborough, Foursquare information on venue categories, top tips, location data, ratings to tackle this task. Neighbourhood information can be accessed

through: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

3.0: Methodology

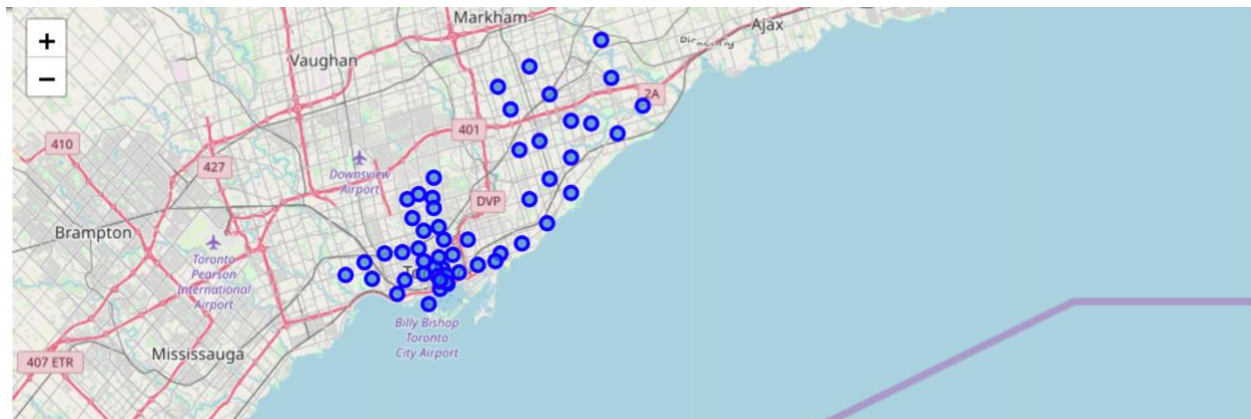
3.1: Summary

First requirement is solved by applying K-means clustering to perform city segmentation. With the neighbourhoods being clustered into clusters, this task becomes finding the neighbourhoods within the same cluster as 'Central bay street'.

The Second and Third requirement can be solved by providing venue and ratings information from Foursquare API. The neighbourhoods with high ratings of Chinese restaurants and grocery store will become our target neighbourhoods.

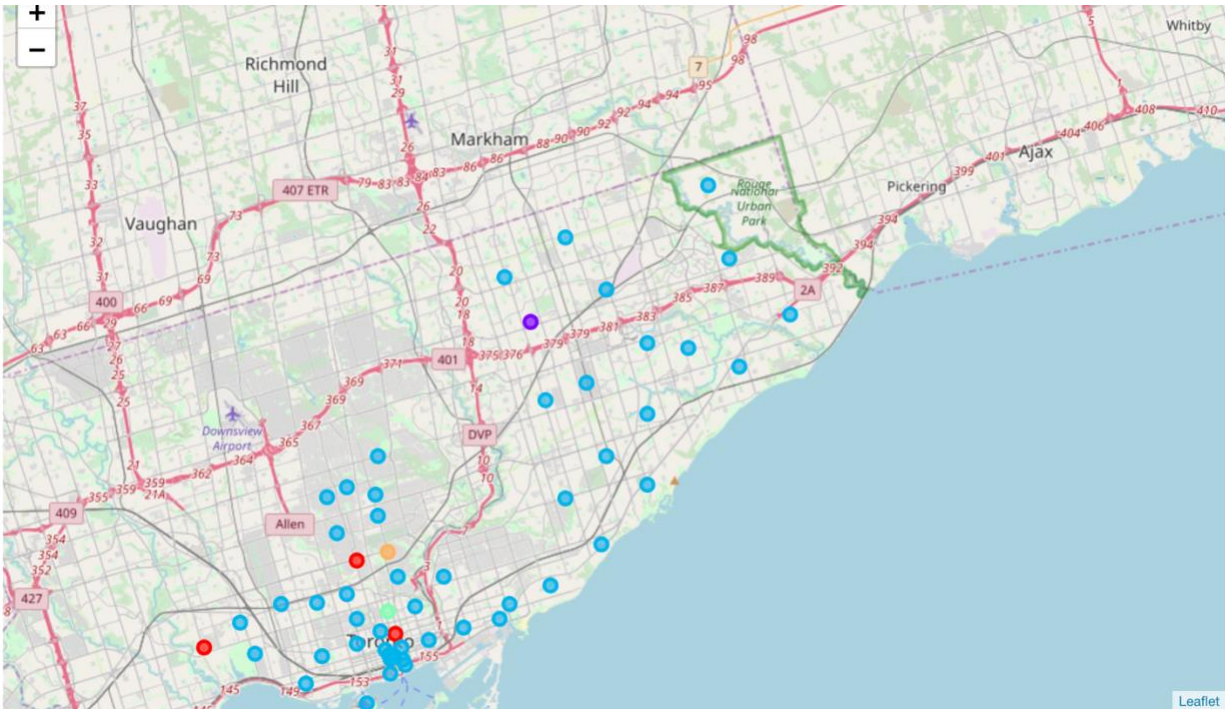
3.2: Procedures

I: In order to get an accurate city segmentation, the whole city of Toronto and Scarborough are our population. 55 neighbourhoods in total.

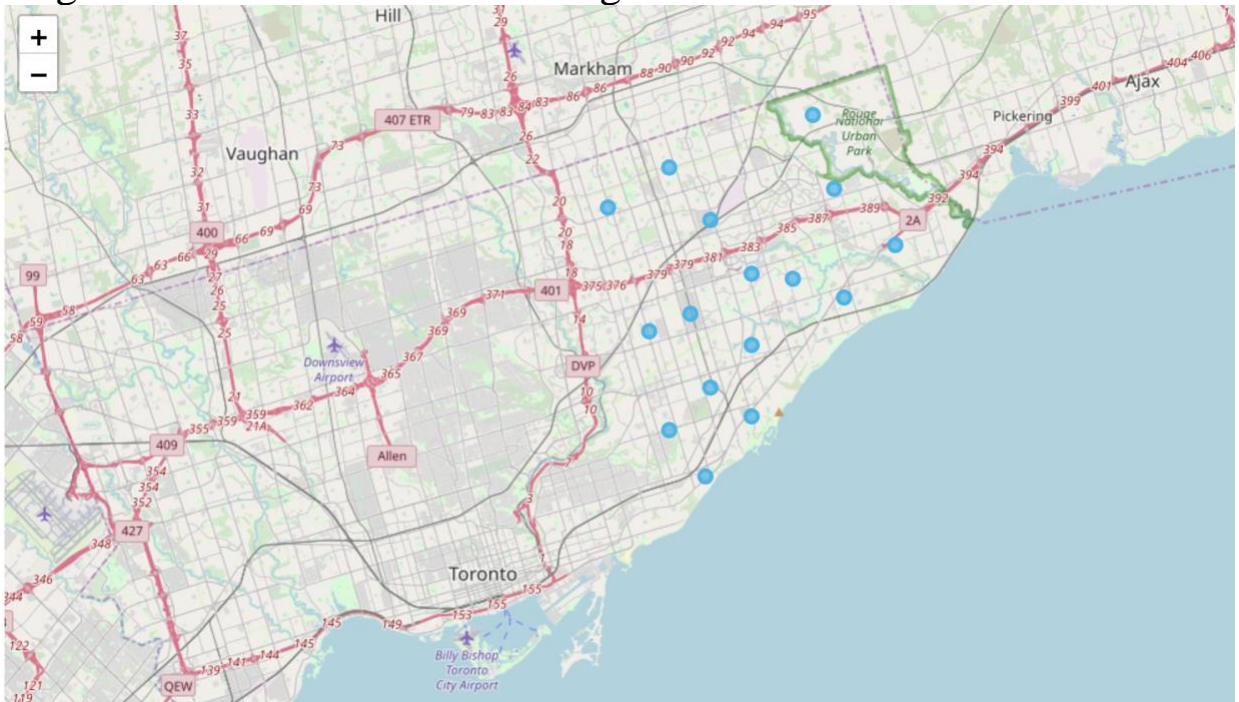


II: Access Foursquare location data of each neighbourhoods, using onehot technique to get top 10 venues of each neighbourhood.

III: Apply Kmeans Cluster to segment all 55 neighbourhoods into 5 clusters.



IV: The ‘Central Bay Street’ belongs to cluster 2, the other cluster 2 neighbourhoods within Scarborough:



V: Filter the data frame by searching for Chinese restaurants and grocery stores, turning out only 3 of all neighbourhoods has Chinese restaurants being one of their top venues. Narrow down our target neighbourhoods to these three neighbourhoods.

M1K Scarborough East Birchmount Park, Ionview, Kennedy Park

M1P Scarborough Dorset Park, Scarborough Town Centre, Wexford Heights

M1W Scarborough L'Amoreaux West, Steeles West

VI: Utilize Foursquare to find grocery stores around each neighbourhood.

- For East Birchmount Park, Ionview, Kennedy Park, 5 grocery stores are found but only 1 of them has ratings: Ghadir Mid-Easton Grocery – 8.1, one of its tips is: “Best Halal butcher shop around”.
- For Dorset Park, Scarborough Town Centre, Wexford Heights, only one grocery store is found, which is Ghadir Min-Easton Grocery.
- For L'Amoreaux West, Steeles West, The Low Carb Grocery is found but it only has a rating of 7.2.

VII: Get rid of “L'Amoreaux West, Steeles West”, explore the Chinese restaurants around the other two neighbourhoods. It turns out they have almost same Chinese Restaurants around the two neighbourhood, with an average rating of 6.48.

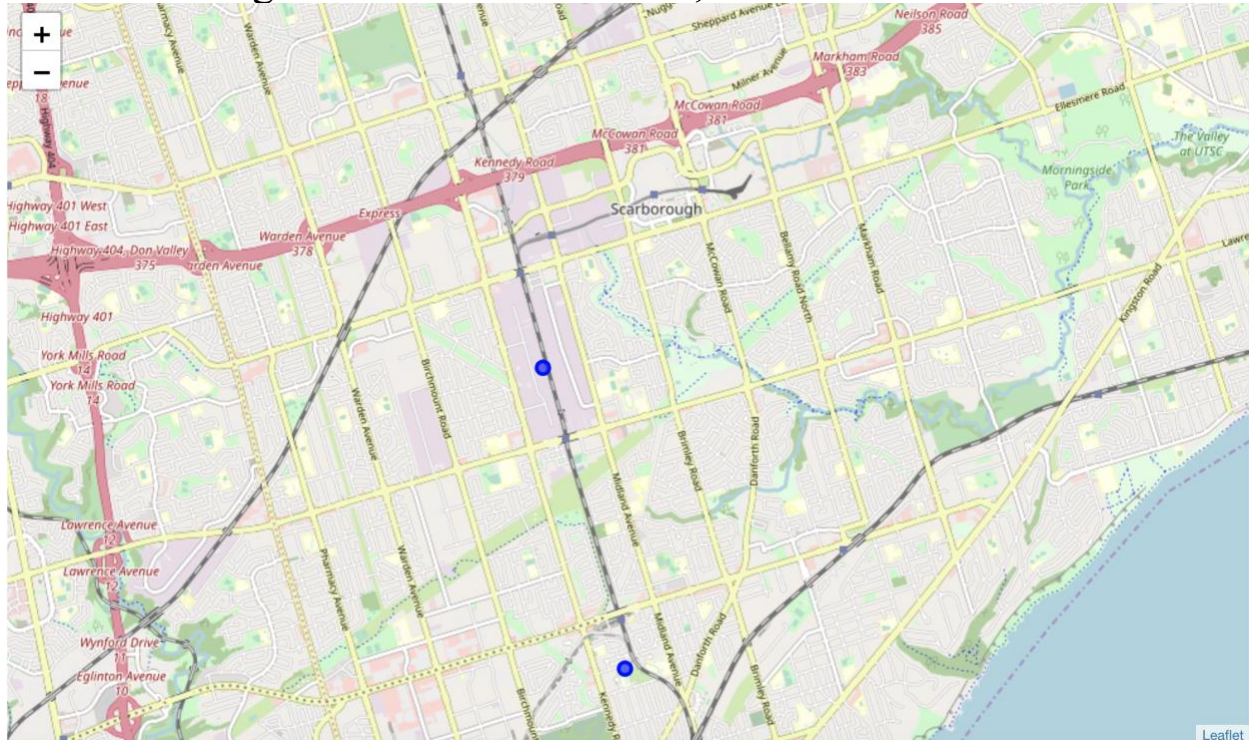
4.0: Results

By exploring the client's requirements of Chinese restaurants and grocery stores, the two neighbourhoods: ‘Dorset Park, Scarborough Town Centre, Wexford heights’, ‘L'Amoreaux West, Steeles West’

have the same Restaurants and Grocery Stores selection. Besides, they are within Scarborough and cluster 2, which fit all the requirements of our clients.

5.0: Discussion

These two neighbourhoods are so alike, let's see the location of them.



So they are on the same road and very close to each other, it is likely that those venues of requirements are just between them geographically. Thus, they should be recommended to our clients.

6.0 Conclusion

This result has limitations. The venues used in the project are from the top 10 venues of each neighbourhood, which might neglect neighbourhoods with more venues but less frequency of Chinese restaurants and grocery stores.

Secondly, we use 10km as our parameter of searching venues. This can be wrong if our clients don't have a car or there's no public transition to those venues.

In conclusion, "Dorset Park, Scarborough Town Centre, Wexford heights", "L'Amoreaux West, Steels West" meets the requirement of our client but better choices may exist.