The Battle of Neighborhoods

Determine Next best locations for the Client to Expand Operations

Outdoor Recreation and Professional Services Inc

Business Problem

- The Outdoor Recreation and Professional Services Inc is located in Rosedale, Downtown Toronto. The client had a successful run in the present neighborhood of the Toronto City.
- They offer outdoor activity and recreation products and services to their customer
- .Now with its existing business maturing in Rosendale, they would like to expand their operations in another neighborhood.
- The bulk of the clients existing business comes from the venues in the existing neighborhood. The client's business operating model is therefore proven and as such the client wants to reduce the risk and accentuate the possibility of success by investing in neighborhoods with similar venue profile.
- The objective therefore is to facilitate the client’s selection of the neighborhoods by proposing neighborhoods that are similar to the present one.

Data : Requirements and Sources

The following data set was ascertained to be used for the initial iteration for the project:

A. List of all the boroughs and neighborhoods associated with Toronoto

- The following link was found to contain all the relevant borough's alongwith the postal codes https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M'

B. List of all the boroughs associated with Toronoto

- In order to explore the venues for the boroughs, the geo coordinates for the borough's were also required. The same was preprocessed and made available postal code wise. The csv file is to be downloaded from http://cocl.us/Geospatial_data'

C. FOURSQUARE api with the Explore endpoint is planned to be used to fetch the details for the venues near the latitude and longitude coordinates for each borough.

- Inorder to call the api, the developer access to Foursquare Developers was registered and the client_id and client secret keys were generated as a pre-requisite. The api call returns json with a multitude of keys and value pairs. The relevant data sets are to be extract by writing suitable functions. The json data can then be tranformed into data frame using some of the existing python packages.
- All the above data sets were progressively used, combined and transformed for the effective analysis, exploration, clustering and final recommendations to meet the project objective.

Data : Features

Sr. No.	Feature	Source
1	Neighborhood	Wikipedia page
2	Latitude	Enriched using the Postal Code Geo Coding file
3	Longitude	Enriched using the Postal Code Geo Coding file
4	Venue	Derived using the FourSquare API with Explore endpoint
5	Venue Latitude	Derived using the FourSquare API with Explore endpoint
6	Venue Longitude	Derived using the FourSquare API with Explore endpoint
7	Venue Category	Derived using the FourSquare API with Explore endpoint

Methodology

- The clustering algorithm used in K-Means. Based on the venue profile for each of the neighborhood, K-means shall help ascertain the similarity amongst the neighborhoods by grouping them into the clusters.

- Once the clusters are assigned to the neighborhoods, it will be easy to locate other neighborhoods that are similar to the client's present business location, Rosendale. Thereby the other neighborhoods belonging to the same cluster as Rosendale shall be the recommended locations for the next expansion location.

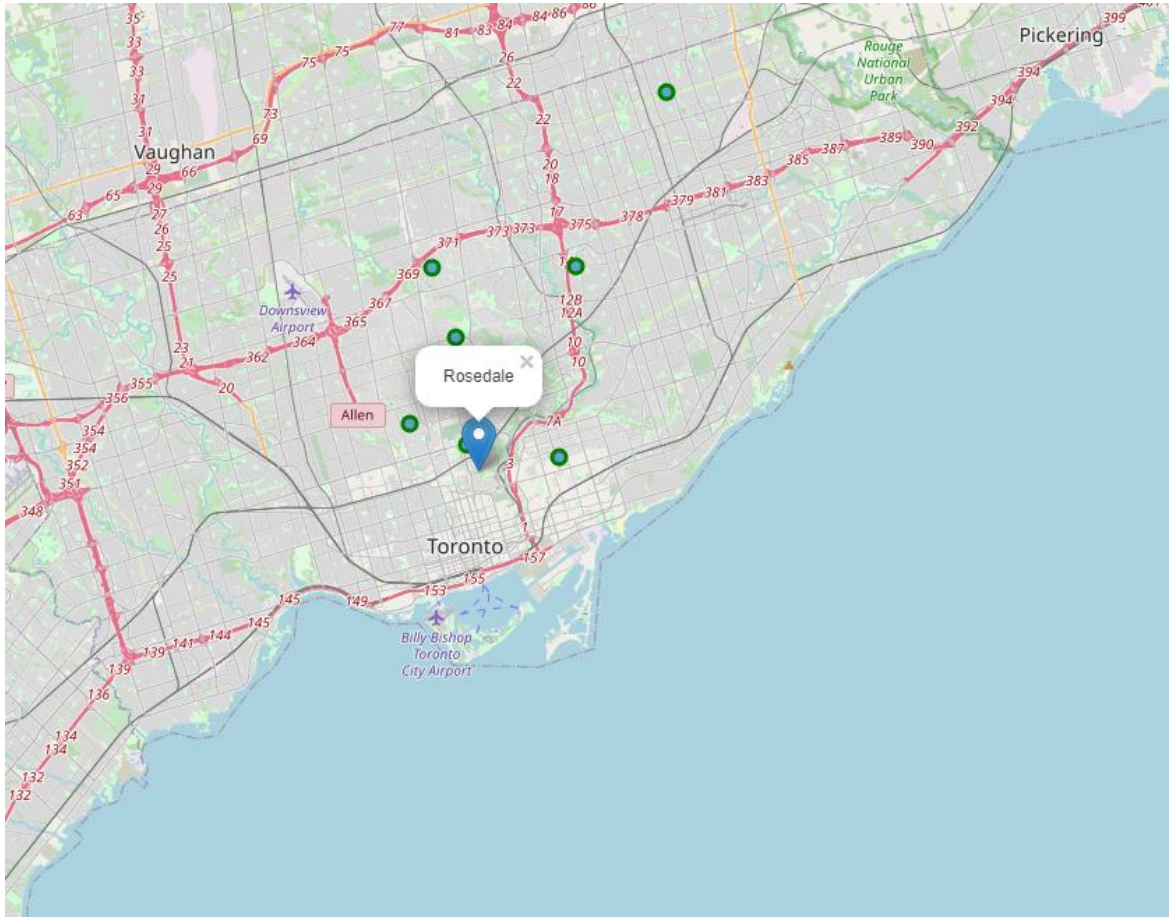
Results

Client’s current location, that is Rosendale has been grouped into Cluster 5. Therefore keeping in line with the projects objective of reduce the risk and maximizing return, all the locations grouped in cluster 5 are the recommended locations for the client, Outdoor Recreation and Professional Services Inc.

Neighborhoods recommended for the next business Location those are similar to Rosendale:

- Parkwoods
- East Toronto, Broadview North (Old East York)
- Lawrence Park
- York Mills West
- Forest Hill North & West, Forest Hill Road Park
- Moore Park, Summerhill East
- Milliken, Agincourt North, Steeles East, L'Amoreaux East

Results


Rosedale
Vaughan
Pickering
Toronto
Allen
Downsview Airport
Billy Bishop Toronto City Airport
Rouge National Urban Park

Results

- K-Means Cluster with similar Neighborhoods

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	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
M3A	North York	Parkwoods	43.753259	-79.329656	4.0	Park	Pool	Food & Drink Shop	Yoga Studio	Curling Ice	Drugstore	Donut Shop	Dog Run	Distribution Center	Discount Store
M4J	East York	East Toronto, Broadview North (Old East York)	43.685347	-79.338106	4.0	Park	Convenience Store	Rental Car Location	Yoga Studio	Curling Ice	Drugstore	Donut Shop	Dog Run	Distribution Center	Discount Store
M4N	Central Toronto	Lawrence Park	43.728020	-79.388790	4.0	Park	Swim School	Bus Line	Yoga Studio	Deli / Bodega	Eastern European Restaurant	Drugstore	Donut Shop	Dog Run	Distribution Center
M2P	North York	York Mills West	43.752758	-79.400049	4.0	Park	Convenience Store	Yoga Studio	Dance Studio	Eastern European Restaurant	Drugstore	Donut Shop	Dog Run	Distribution Center	Discount Store
M5P	Central Toronto	Forest Hill North & West, Forest Hill Road Park	43.696948	-79.411307	4.0	Park	Jewelry Store	Trail	Sushi Restaurant	Yoga Studio	Dance Studio	Drugstore	Donut Shop	Dog Run	Distribution Center
M4T	Central Toronto	Moore Park, Summerhill East	43.689574	-79.383160	4.0	Park	Trail	Yoga Studio	Curling Ice	Eastern European Restaurant	Drugstore	Donut Shop	Dog Run	Distribution Center	Discount Store
M1V	Scarborough	Milliken, Agincourt North, Steeles East, L'Amo...	43.815252	-79.284577	4.0	Playground	Park	Yoga Studio	Curling Ice	Drugstore	Donut Shop	Dog Run	Distribution Center	Discount Store	Diner
M4W	Downtown Toronto	Rosedale	43.679563	-79.377529	4.0	Park	Playground	Trail	Yoga Studio	Curling Ice	Drugstore	Donut Shop	Dog Run	Distribution Center	Discount Store

Discussion

Based on the result above, looks like that are seven neighborhoods that have similarity with the client’s existing business location.

These locations can there after be further subjected to diligence on factors such as proximity, resource availability etc to further shortlist on one that leverages the synergy and reduces cost to get the maximum possible returns on investment.