Your new neighbourhood - recommender system (The Battle of Neighbourhoods)

Data

The system described in the introduction section will use following data sources.

1. Neighbourhood information

In this pilot study we'll be using neighbourhood data for Toronto, Canada which postal codes starts with letter M and can be obtained here:

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

Which can be converted to following data frame

	Postal Code	Borough	Neighbourhood
0	M3A	North York	Parkwoods
1	M4A	North York	Victoria Village
2	M5A	Downtown Toronto	Regent Park, Harbourfront
3	M6A	North York	Lawrence Manor, Lawrence Heights
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government

2. Geolocation information

The latitude and longitude data for each postal code extracted by "pgeocode" python package which resulted in data frame bellow

	Postal Code	Borough	Neighbourhood	Latitude	Longitude
0	M3A	North York	Parkwoods	43.7545	-79.33
1	M4A	North York	Victoria Village	43.7276	-79.3148
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.6555	-79.3626
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.7223	-79.4504
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.6641	-79.3889

3. Points of interests

The study uses two endpoints of the Foursquare API.

https://developer.foursquare.com/docs/places-api/endpoints/

The the full list of points of interest which are available from Foursquare extracted from "venue categories" and resulted in following data frame

	Main Category ID	Main Category Name	Category ID	Category Name
0	4d4b7104d754a06370d81259	Arts & Entertainment	56aa371be4b08b9a8d5734db	Amphitheater
1	4d4b7104d754a06370d81259	Arts & Entertainment	4fceea171983d5d06c3e9823	Aquarium
2	4d4b7104d754a06370d81259	Arts & Entertainment	4bf58dd8d48988d1e1931735	Arcade
3	4d4b7104d754a06370d81259	Arts & Entertainment	4bf58dd8d48988d1e2931735	Art Gallery
4	4d4b7104d754a06370d81259	Arts & Entertainment	4bf58dd8d48988d1e4931735	Bowling Alley

The Foursquare API "search" endpoint was used to obtain extensive points of interest data for each neighbourhood resulted in summarised data frame bellow

	Postal Code	Borough	Neighbourhood	Latitude	Longitude	Amphitheater	Aquarium	Arcade	Art Gallery	Bowling Alley	...	Road	Taxi Stand	Taxi	Tourist Information Center
0	M3A	North York	Parkwoods	43.7545	-79.33	0	0	0	0	0	...	0	0	0	0
1	M4A	North York	Victoria Village	43.7276	-79.3148	0	0	0	1	0	...	0	0	0	0
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.6555	-79.3626	0	0	4	47	1	...	0	0	0	2
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.7223	-79.4504	0	0	2	1	1	...	0	0	0	0
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.6641	-79.3889	1	1	3	45	1	...	0	0	1	0