



The battle of Neighbourhoods

KP Realtor Inc.



Introduction/Business Problem

- Identify the best location(Neighbourhood) of Toronto for client (Retired Couple) to buy their first home that fits their requirements

Neighbourhood should have:

- Parks
- Pet friendly places
- Drug Store Restaurants
- Coffee shops
- Trails
- Music friendly places

Data

- **Data 1:** list of Borough and neighbourhood associated with the postal codes of Toronto city
https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

Table 1:

	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

Data

- **Data 2:** Toronto city geographical coordinates from the Foursquare API for neighborhood venues in Toronto
- geographical data were also for Toronto from http://cocl.us/Geospatial_data

Out[46]:

	PostalCode	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

Methodology

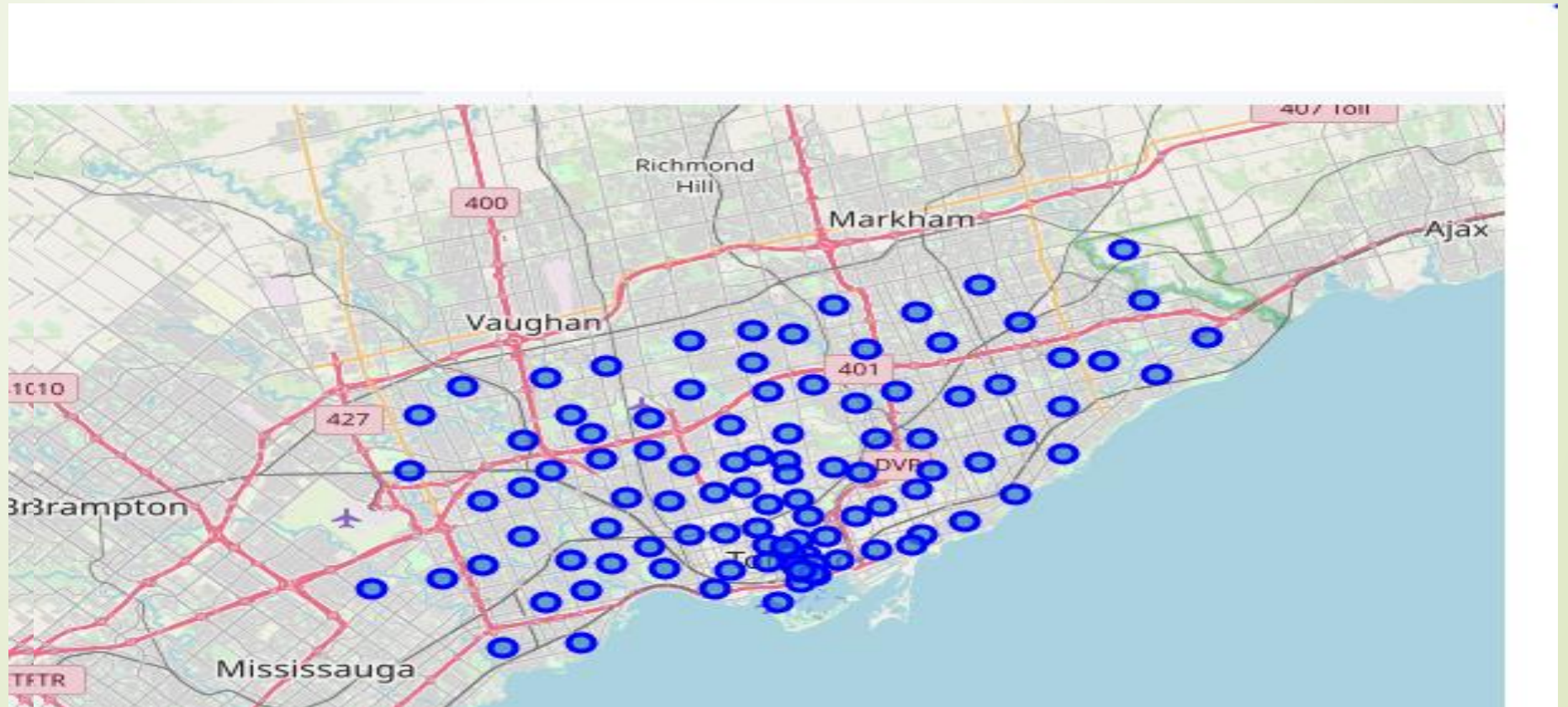
- Imported neighbourhood data with postal codes and merged with the geographical data for Toronto to assign respective longitude and latitude

Out[138]:

	PostalCode	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

Methodology

- Foursquare API - get venues surrounding neighbourhoods of Toronto



Methodology

- Data wrangling/editing for analysis and identify Top 5 most venues for each neighbourhood and 1st to 10th most common venues for neighbourhoods

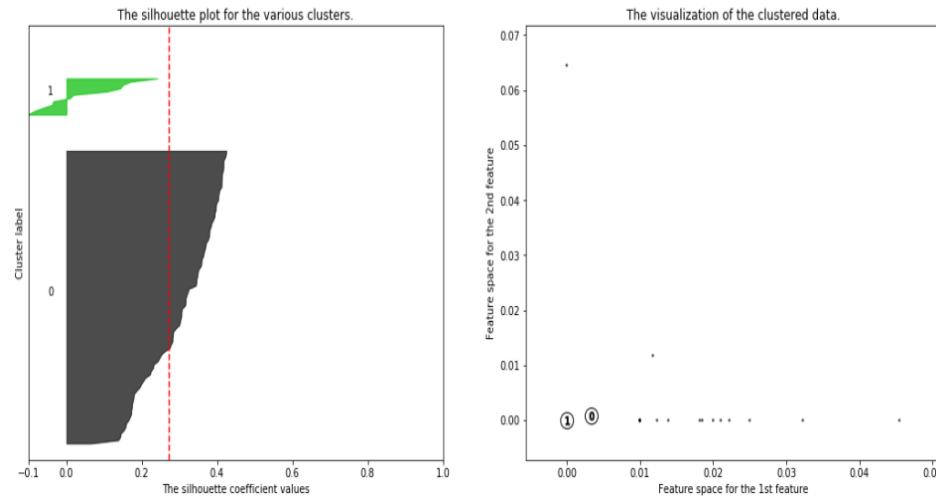
Out[115]:

	Yoga Studio	Accessories Store	Afghan Restaurant	Airport	Airport Food Court	Airport Gate	Airport Lounge	Airport Service	Airport Terminal	American Restaurant	Amphitheater	Animal Shelter	Antique Shop	Aquarium	Arcade	Argentinian Restaurant	A Galle
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Neighborhood	Yoga Studio	Accessories Store	Afghan Restaurant	Airport	Airport Food Court	Airport Gate	Airport Lounge	Airport Service	Airport Terminal	American Restaurant	Amphitheater	Animal Shelter	Antique Shop	Aquarium	Arcade	Air
0	Adelaide,King,Richmond	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.02	0.0	0.0	0.0	0.0	0.0
1	Agincourt	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.0	0.0	0.0	0.0	0.0
2	Agincourt North,L'Amoreaux East,Milliken,Steel...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.0	0.0	0.0	0.0	0.0
3	Albion Gardens,Beaumont Heights,Humbergate,Jam...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.0	0.0	0.0	0.0	0.0
4	Alderwood,Long Branch	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.0	0.0	0.0	0.0	0.0

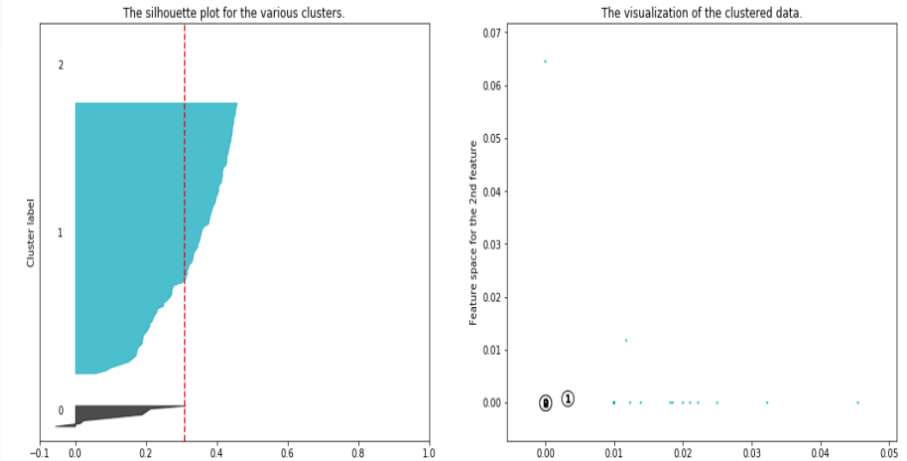
Methodology

➤ KMeans clustering

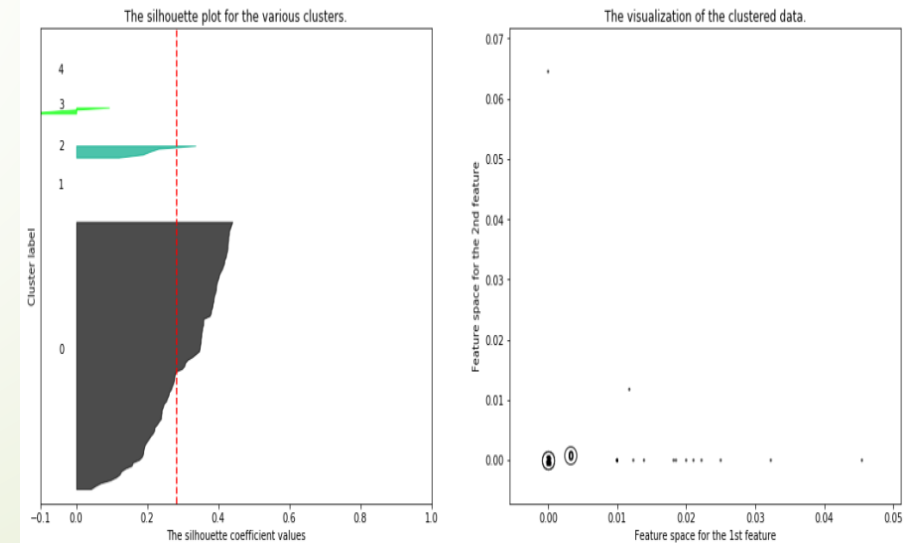
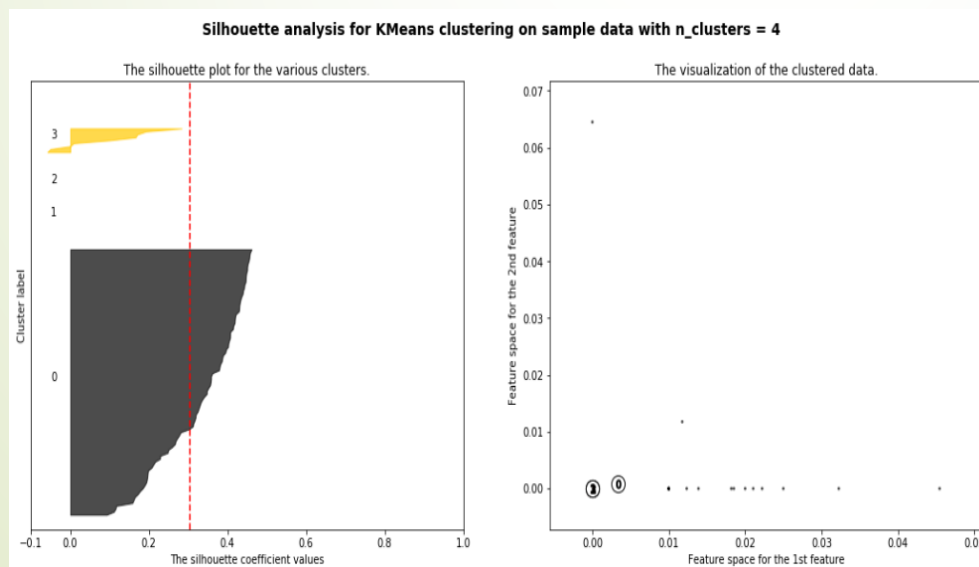
Silhouette analysis for KMeans clustering on sample data with $n_clusters = 2$



Silhouette analysis for KMeans clustering on sample data with $n_clusters = 3$



Silhouette analysis for KMeans clustering on sample data with $n_clusters = 5$



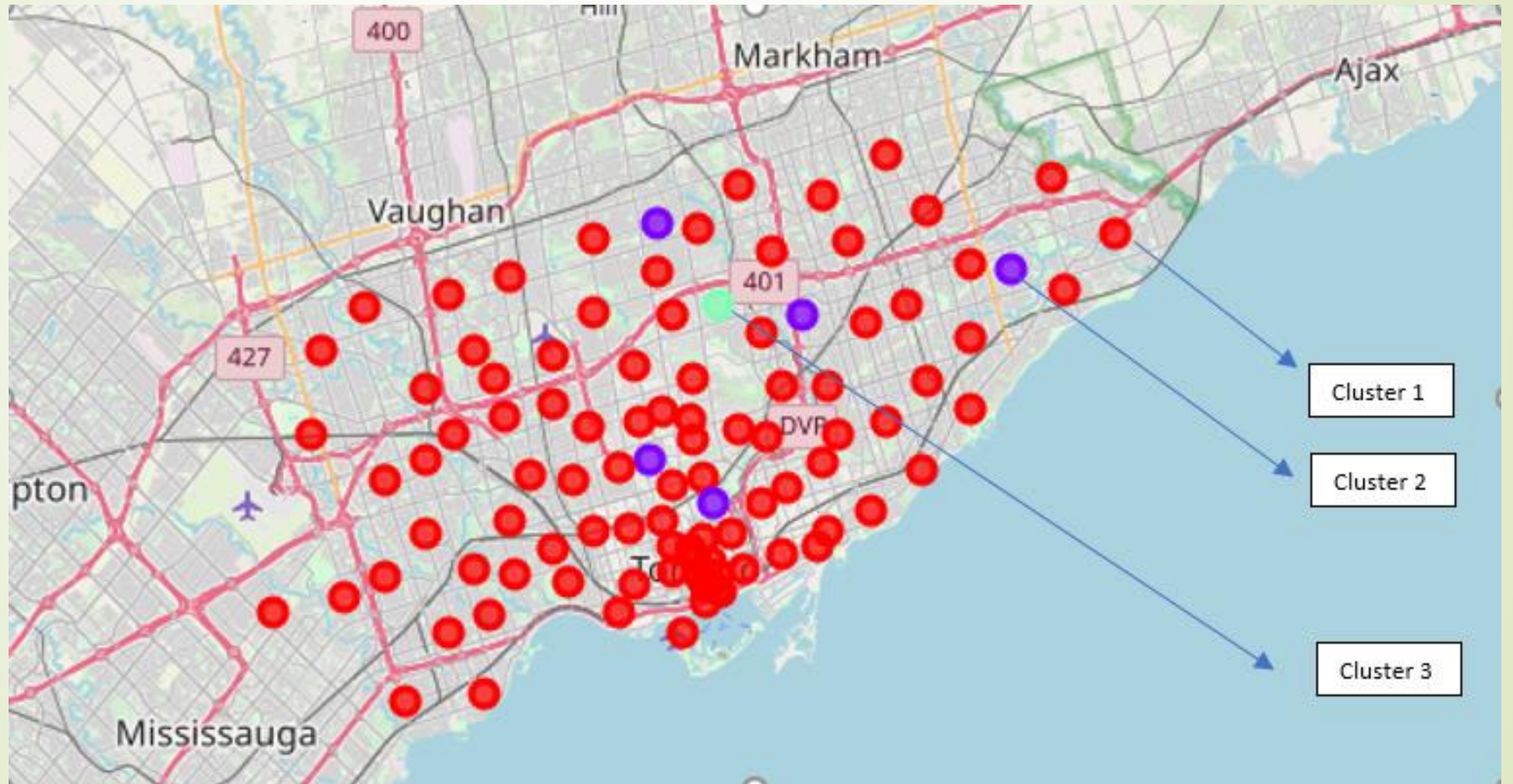
Methodology

- Dataset (Neighbourhoods) divided in 3 clusters

	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
0	M1B	Scarborough	Rouge,Malvern	43.806686	-79.194353	0.0	Fast Food Restaurant	Hobby Shop	Business Service	Coffee Shop	Spa	Bus Station	Construction & Landscaping	Drugstore
1	M1C	Scarborough	Highland Creek,Rouge Hill,Port Union	43.784535	-79.160497	0.0	Breakfast Spot	Burger Joint	Bar	Women's Store	Dive Bar	Dog Run	Doner Restaurant	Donut Shop
2	M1E	Scarborough	Guildwood,Morningside,West Hill	43.763573	-79.188711	0.0	Pizza Place	Fast Food Restaurant	Thrift / Vintage Store	Park	Rental Car Location	Moving Target	Bus Line	Electronics Store
3	M1G	Scarborough	Woburn	43.770992	-79.216917	1.0	Coffee Shop	Park	Business Service	Women's Store	Drugstore	Diner	Discount Store	Dive Bar
4	M1H	Scarborough	Cedarbrae	43.773136	-79.239476	0.0	Bakery	Coffee Shop	Indian Restaurant	Thai Restaurant	Caribbean Restaurant	Gas Station	Fried Chicken Joint	Athletics & Sports

Methodology

- Dataset (Neighbourhoods) divided in 3 clusters based on most common venues



Methodology

- Dataset (Neighbourhoods) divided in 3 clusters based on most common venues

Out[149]:

	Cluster1	Cluster2	Cluster3
0	Rouge, Malvern		
1	Highland Creek, Rouge Hill, Port Union		
2	Guildwood, Morningside, West Hill		
3		Woburn	
4	Cedarbrae		
5	Scarborough Village		
6	East Birchmount Park, Ionview, Kennedy Park		
7	Clairlea, Golden Mile, Oakridge		
8	Cliffcrest, Cliffside, Scarborough Village West		
9	Birch Cliff, Cliffside West		

Results and Discussions

- **Cluster 1** - Many fast food and speciality restaurants; coffee shops.
- Pretty much clustered through out the whole city of Toronto.
- very less parks and natural places in this neighbourhood. These places seem to be well populated and business friendly.

Cluster 1

```
df1=toronto_cl1.loc[toronto_cl1['Cluster Labels'] == 0,toronto_cl1.columns[[2] + list(range(5, toronto_cl1.shape[1]))]]  
df1
```

[146]:

	Neighborhood	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
0	Rouge,Malvern	0	Fast Food Restaurant	Hobby Shop	Business Service	Coffee Shop	Spa	Bus Station	Construction & Landscaping	Drugstore	Discount Store	Dive Bar
1	Highland Creek,Rouge Hill,Port Union	0	Breakfast Spot	Burger Joint	Bar	Women's Store	Dive Bar	Dog Run	Doner Restaurant	Donut Shop	Drugstore	Dumpling Restaurant
2	Guildwood,Morningside,West Hill	0	Pizza Place	Fast Food Restaurant	Thrift / Vintage Store	Park	Rental Car Location	Moving Target	Bus Line	Electronics Store	Mexican Restaurant	Breakfast Spot
4	Cedarbrae	0	Bakery	Coffee Shop	Indian Restaurant	Thai Restaurant	Caribbean Restaurant	Gas Station	Fried Chicken Joint	Athletics & Sports	Asian Restaurant	Chinese Restaurant
5	Scarborough Village	0	Fast Food Restaurant	Coffee Shop	Pizza Place	Convenience Store	Farm	Dumpling Restaurant	Discount Store	Dive Bar	Dog Run	Doner Restaurant

Results and Discussions

- **Cluster 2** - Many parks and trails.
- There are only few assorted neighbourhoods in this cluster.
- These places seem to be less populated and quiet.

Cluster 2

```
df2=toronto_cl1.loc[toronto_cl1['Cluster Labels'] == 1,toronto_cl1.columns[[2] + list(range(5, toronto_cl1.shape[1]))]]  
df2
```

147]:

	Neighborhood	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
3	Woburn	1	Coffee Shop	Park	Business Service	Women's Store	Drugstore	Diner	Discount Store	Dive Bar	Dog Run	Doner Restaurant
20	Newtonbrook, Willowdale	1	Park	Coffee Shop	Music Venue	Trail	Drugstore	Diner	Discount Store	Dive Bar	Dog Run	Doner Restaurant
24	Parkwoods	1	Park	Food & Drink Shop	Pet Store	Women's Store	Dim Sum Restaurant	Diner	Discount Store	Dive Bar	Dog Run	Doner Restaurant
49	Rosedale	1	Park	Playground	Trail	Drugstore	Dim Sum Restaurant	Diner	Discount Store	Dive Bar	Dog Run	Doner Restaurant
63	Forest Hill North, Forest Hill West	1	Park	Jewelry Store	Trail	Sushi Restaurant	Women's Store	Drugstore	Diner	Discount Store	Dive Bar	Dog Run

Results and Discussions

- Cluster 3 - Unique neighbourhood with unique venues which are not common in cluster 1 and cluster 3.

Cluster 3

```
df3=toronto_cl1.loc[toronto_cl1['Cluster Labels'] == 2,toronto_cl1.columns[[2] + list(range(5, toronto_cl1.shape[1]))]]
df3
```

[148]:

	Neighborhood	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
19	Silver Hills, York Mills	2	Cafeteria	College Rec Center	Dim Sum Restaurant	Diner	Discount Store	Dive Bar	Dog Run	Doner Restaurant	Donut Shop	Drugstore

Conclusion:

- **Neighbourhoods listed in cluster 2 best suites for client** to chose a location that meets client's requirement since parks, coffee shops, drug stores and trails are the amongst the top 5 most common venues around this neighbourhood.
 - **Woburn**
 - **Newtonbrook**
 - **Willowdale**
 - **Parkwoods**
 - **Rosedale**
 - **Forest Hill North**
 - **Forest Hill West**
- Note: Hospital is not a most common venue in any of these neighbourhood.
- According to this analysis it is evident that the neighbourhood of Newtonbrook and Willowdale (North York/Postal code:M2) is the most fit location for this client since all venues that client would want for their new house location are common in this neighbourhood
- **The best suited location to buy new house for client:**
 - **Newtonbrook and Willowdale (North York/Postal code:M2)**