CAPSTONE PROJECT - THE BATTLE OF NEIGHBORHOODS

Prepared by:

Olga Hnat

Table of Contest

Where Should I Live?

As a part of the final course of IBM Data Science Professional Certificate - we have worked on a final Capstone Project. Main objectives of the project were as follows:

- Define a problem for your capstone project
- Discuss the data that you will use to solve the problem
- Carry out the work to complete the capstone project using available means such as Python libraries and Foursquare location data

In this project, we will go through all the steps from problem identification, data gathering to final analysis and finish up with a conclusion.

1. Introduction

For this project I have decided to select not necessarily a true business problem, but, more of a life one.

Imagine the simplest task each of us have to face when you are looking to move to another city - where should I live? Which area should i pick? What neighborhood? I was faced with this task about 3 years ago when I was moving to Toronto for work. Unfortunately, at that point of time I didn't have enough of data science knowledge to automate the search process, thus, decision was made based on lots and lots of manual research. However, having all this valuable Python knowledge, let's try to see how I could have tackled this problem.

Problem Statement

To reiterate - the goal is to identify the most suitable neighborhood in Toronto to move into based on the following criteria:

- proximity to Toronto Downtown
- availability of various venues around. Venue list should include:
    - Movie Theater
    - Dance Studio

- o Food
    - ▪ Gluten-free Restaurant
    - ▪ Vegetarian / Vegan Restaurant
- o Coffee Shop
- o Grocery Store
- o Pharmacy
- o Shopping Mall
- o Nightlife Spot
- o Gym / Fitness Center
- o Park

Target Audience

Who may be interested in this project?

- Most certainly, people who are looking to relocate into a new city due to any reason.
- Most likely, people who are looking to settle down and purchase a property, however, they are confused which area is the best. With some little tweaks to the process - this can become quite valuable and, most importantly, time saving way to arrive to final decision.
- And, of course, my fellow Data Analysts and Data Scientists, who are looking to analyze the neighborhoods of any city, in my case - Toronto, using a variety of techniques to reach desired results and draw subsequent conclusions.

2. Data

The data used for this project is as follows:

1. Wikipedia page https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M, containing the list of Toronto postal codes, Boroughs and Neighborhoods. We will scrape this page, get the data we need and create a pandas data frame. A bit of assumptions:

   a. The data frame will consist of three columns: Postal Code, Borough, and Neighborhood.

   b. We will only process the cells that have an assigned borough. We shall ignore cells with a borough that is 'Not assigned'.

2. A CSV file that has the geographical coordinates of each of Toronto's postal codes: http://cocl.us/Geospatial_data. We will use this link to obtain needed data and create new data frame containing following columns:

   - Postal Code
   - Borough
   - Neighborhood
   - Latitude of the Neighborhood
   - Longitude of the Neighborhood

3. We also will be using Foursquare API to obtain needed information on the venues in the neighborhoods of Toronto. The API returns a JSON file and we will be transforming it into a data frame for our analysis.

3. Methodology and Analysis

Data Preparation

To start our analysis, we had obtained following data frames.

Data frame containing the list of Boroughs, Neighborhoods and their postal codes. This data frame was created by using Beautiful Soup package and scraping Wikipedia page.

	Postal Code	Borough	Neighborhood
0	M1A	Not assigned	NaN
1	M2A	Not assigned	NaN
2	M3A	North York	Parkwoods
3	M4A	North York	Victoria Village
4	M5A	Downtown Toronto	Regent Park, Harbourfront

Data Frame Containing Information Scraped from Wikipedia Page

Obviously, some data cleaning was required – as discussed prior. We have dropped rows with 'Not assigned' values or 'NaN' values. We have also replaced '/ ' to ',' for better analysis down the road. Final results looked like:

	Postal code	Borough	Neighborhood
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

Data Frame Containing Cleaned Data

As a next step – we have created a data frame with coordinates assigned for each of the postal codes.

	Postal Code	Latitude	Longitude
0	M1B	43.806686	-79.194353
1	M1C	43.784535	-79.160497
2	M1E	43.763573	-79.188711
3	M1G	43.770992	-79.216917
4	M1H	43.773136	-79.239476

Data Frame Containing Coordinates of Toronto Postal Codes

And to create our final Toronto Boroughs and Neighborhoods data frame – we simply merged two data frames.

```
# merging both data frames based on Postal code column

df_final=pd.merge(df, df1, on='Postal code', how='left')
print(df_final.shape)
df_final.head()
```

Code Snippet

The result looked like below. This data frame contains 10 unique boroughs and 103 neighborhoods.

	Postal code	Borough	Neighborhood	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	Regent Park, Harbourfront	43.654260	-79.360636
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494

Merged Data Frame

To get more visual picture of the neighborhoods, we have used Folium package and plotted all coordinates on Toronto map:
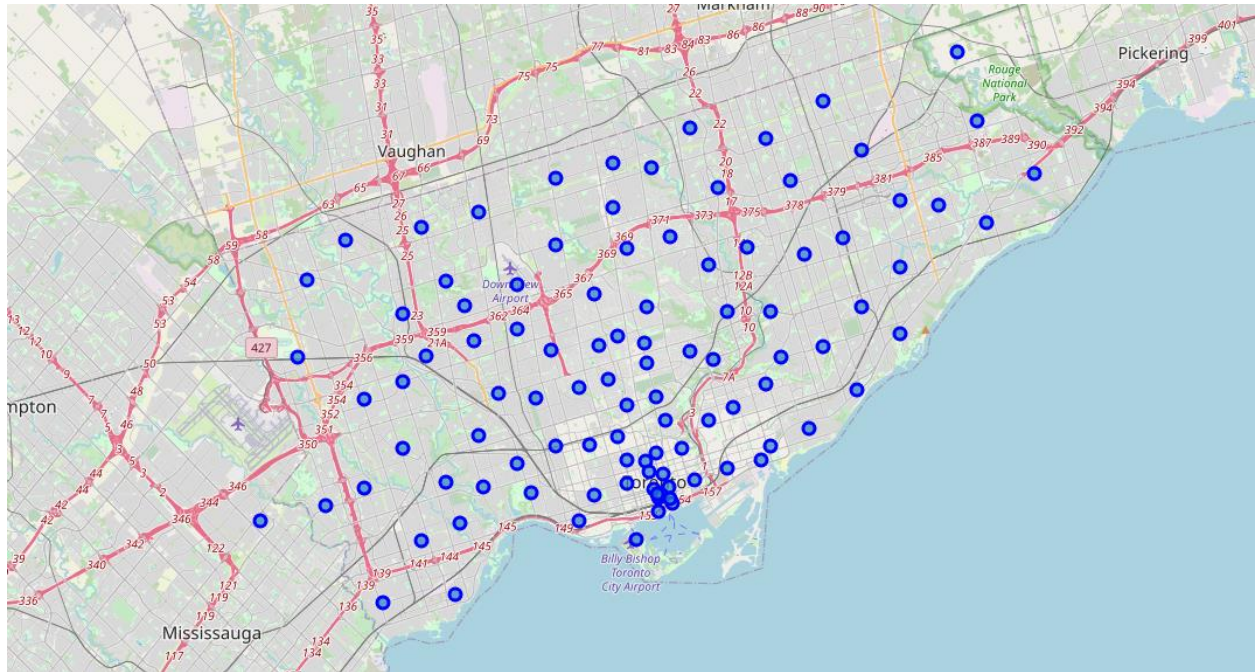
Toronto Map – 103 Neighborhoods

As a final step in the process of obtaining out full data, we have used Foursquare API to get a list of venues for each of listed above neighborhoods. We have set the radius of 1,000 meters and limit of 200 venues per neighborhood. Upon reading JSON file - we have compiled new data frame with the following columns:

- Neighborhood – contains the name of the neighborhood
- NeighbLat – neighborhood latitude
- NeighbLong – neighborhood longitude
- Venue – contains the name of the venue
- VenLat – venue latitude
- VenLong – venue longitude
- VenCat – venue category

	Neighborhood	NeighbLat	NeighbLong	Venue	VenLat	VenLong	VenCat
0	Parkwoods	43.753259	-79.329656	Allwyn's Bakery	43.759840	-79.324719	Caribbean Restaurant
1	Parkwoods	43.753259	-79.329656	Brookbanks Park	43.751976	-79.332140	Park
2	Parkwoods	43.753259	-79.329656	Tim Hortons	43.760668	-79.326368	Café
3	Parkwoods	43.753259	-79.329656	A&W	43.760643	-79.326865	Fast Food Restaurant
4	Parkwoods	43.753259	-79.329656	Bruno's valu-mart	43.746143	-79.324630	Grocery Store
5	Parkwoods	43.753259	-79.329656	Food Basics	43.760549	-79.326045	Supermarket
6	Parkwoods	43.753259	-79.329656	Shoppers Drug Mart	43.745315	-79.325800	Pharmacy
7	Parkwoods	43.753259	-79.329656	High Street Fish & Chips	43.745260	-79.324949	Fish & Chips Shop
8	Parkwoods	43.753259	-79.329656	Shoppers Drug Mart	43.760857	-79.324961	Pharmacy
9	Parkwoods	43.753259	-79.329656	Pizza Pizza	43.760231	-79.325666	Pizza Place

Data on Venues

This data frame has 4909 rows, which means it has the information on 4909 venues in various Toronto neighborhoods. Now – this should be sufficient data for us to perform analysis and draw some conclusions.

Exploratory Analysis

Let's understand our data a bit more and answer some exploratory questions.

1. <u>How many unique venue categories do we have in our data frame</u>? This is a simple question that can be answered with one line of code.

```python
print('There are {} uniques categories.'.format(len(toronto_venues['VenCat'].unique())))
```
```
There are 330 uniques categories.
```

<p align="center">Code Snippet</p>

We have 303 unique categories. Definitely – Toronto boroughs has a lot to offer.

2. <u>What is the most popular venue in all of the data set</u>? In other words, what venue category has the most number of venues?

VenCat	Neighborhood	NeighbLat	NeighbLong	Venue	VenLat	VenLong	Borough
Coffee Shop	414	414	414	414	414	414	414
Café	233	233	233	233	233	233	233
Park	166	166	166	166	166	166	166
Pizza Place	165	165	165	165	165	165	165
Restaurant	150	150	150	150	150	150	150
Italian Restaurant	114	114	114	114	114	114	114
Grocery Store	111	111	111	111	111	111	111
Bakery	107	107	107	107	107	107	107
Japanese Restaurant	99	99	99	99	99	99	99
Sandwich Place	98	98	98	98	98	98	98

<p align="center">Data Frame of the Most Popular Venue Categories</p>

Seems like residents of Toronto really like their coffee – there are 414 venues in Coffee Shop category. The next in line – all sorts of food venues – cafes, pizza places, restaurants, Italian, Japanese, sandwiches. This makes me think that Toronto is a foodie city!

3. <u>Which borough has the highest number of venues</u>?

Borough	Neighborhood	NeighbLat	NeighbLong	Venue	VenLat	VenLong	VenCat
Downtown Toronto	1674	1674	1674	1674	1674	1674	1674
North York	1017	1017	1017	1017	1017	1017	1017
Central Toronto	567	567	567	567	567	567	567
West Toronto	540	540	540	540	540	540	540
East Toronto	407	407	407	407	407	407	407

Data Frame of Boroughs with the Highest Number of Venues

Downtown Toronto is in the lead. This borough has over 1600 venues of various categories. Definitely, you won't be bored in Toronto Downtown.

4. <u>Which borough has the highest number of unique venue categories</u>?

Borough	VenCat
Downtown Toronto	196
West Toronto	157
North York	151
East Toronto	132
Central Toronto	126

Data Frame of the Boroughs with the Highest Number of Unique Venue Categories

Downtown Toronto is in the lead again! Not only this Borough has the highest number of venues in total, it also has the highest number of unique venue categories. However, take a note of the first 5 boroughs for this data frame and the one above – boroughs are matching. It means – these 5 boroughs have the highest number of venues overall and the highest number of unique venues. We are adding these 5 boroughs along with all corresponding neighborhoods and venues into a separate data frame called **'Most Desired Boroughs'** for further analysis.

Let's move from borough plane to the level of neighborhoods and do some exploratory analysis there.

5. <u>Which neighborhood has the **maximum** number of venues?</u>

Neighborhood	NeighbLat	NeighbLong	Venue	VenLat	VenLong	VenCat
Willowdale	112	112	112	112	112	112
Little Portugal, Trinity	100	100	100	100	100	100
Church and Wellesley	100	100	100	100	100	100
Garden District, Ryerson	100	100	100	100	100	100
Regent Park, Harbourfront	100	100	100	100	100	100

Data Frame of the Neighborhoods with the Highest Number of Venues

Seems like Willowdale is the leader as it has 112 venues.

6. <u>What would be the **minimum** number of venues returned per neighborhood?</u>

Neighborhood	NeighbLat	NeighbLong	Venue	VenLat	VenLong	VenCat
Woburn	8	8	8	8	8	8
Old Mill South, King's Mill Park, Sunnylea, Humber Bay, Mimico NE, The Queensway East, Royal York South East, Kingsway Park South East	8	8	8	8	8	8
Rouge Hill, Port Union, Highland Creek	5	5	5	5	5	5
York Mills, Silver Hills	4	4	4	4	4	4
Northwest	3	3	3	3	3	3

Data Frame of the Neighborhoods with the Lowest Number of Venues

Well, you probably will not want to live in any of those – as that would be a definite challenge.

7. <u>Which neighborhood has the highest number of unique venue categories?</u> In other words – what neighborhood can offer the most diverse multitude of venues:

Neighborhood	VenCat
Church and Wellesley	67
Queen's Park, Ontario Provincial Government	65
The Annex, North Midtown, Yorkville	63
Harbourfront East, Union Station, Toronto Islands	62
Richmond, Adelaide, King	61

Data Frame of the Neighborhoods with the Highest Number of Unique Venue Categories

Seems like Church and Wellesley are in the lead.

One-hot Encoding and Neighborhood Clustering

We have applied one-hot encoding technique to the list of the venues in the **most desired boroughs** which would be the boroughs with the highest number of unique venue categories. That would be: Downtown Toronto, West Toronto, North York, East York and Central Toronto. As the result, we have obtained a data frame containing the list of the first 10 most common venues new neighborhood that are a part of most desired boroughs.

	Neighborhood	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	Bathurst Manor, Wilson Heights, Downsview North	Coffee Shop	Bank	Convenience Store	Supermarket	Frozen Yogurt Shop	Fried Chicken Joint	Dog Run	Sandwich Place	Sushi Restaurant	Shopping Mall
1	Bayview Village	Bank	Grocery Store	Gas Station	Japanese Restaurant	Chinese Restaurant	Restaurant	Park	Intersection	Café	Trail
2	Bedford Park, Lawrence Manor East	Italian Restaurant	Coffee Shop	Bank	Sandwich Place	Bridal Shop	Fast Food Restaurant	Intersection	Sushi Restaurant	Juice Bar	Sports Club
3	Berczy Park	Coffee Shop	Café	Hotel	Japanese Restaurant	Restaurant	Beer Bar	Park	Seafood Restaurant	Cocktail Bar	Creperie
4	Brockton, Parkdale Village, Exhibition Place	Café	Coffee Shop	Restaurant	Bar	Furniture / Home Store	Bakery	Tibetan Restaurant	Gift Shop	Supermarket	Indian Restaurant

Data Frame of the 10 Most Common Venues per Neighborhood

We have performed K-Means machine learning (code below) to split the list of available neighborhoods into 4 clusters. But, first, we needed to find the best K – meaning the most accurate number of clusters to use for our splitting exercise.
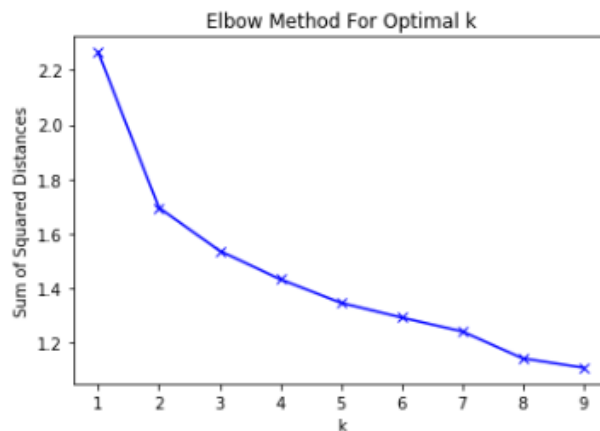

The Best Value of K

We have split the list of available neighborhoods into 4 clusters. And for better visual picture –
we have plotted all of them on Toronto map.

Neighborhood	NeighbLat	NeighbLong	Borough	Cluster Labels	1st Most Common Venue	2nd Common V
Parkwoods	43.753259	-79.329656	North York	3	Park	Convenience :
Victoria Village	43.725882	-79.315572	North York	0	Coffee Shop	Hockey A
Regent Park, Harbourfront	43.654260	-79.360636	Downtown Toronto	0	Coffee Shop	Th
Lawrence Manor, Lawrence Heights	43.718518	-79.464763	North York	0	Clothing Store	Fast Resta
Queen's Park, Ontario Provincial Government	43.662301	-79.389494	Downtown Toronto	0	Coffee Shop	Sushi Resta

Data Frame of the Neighborhoods with Cluster Value Assigned
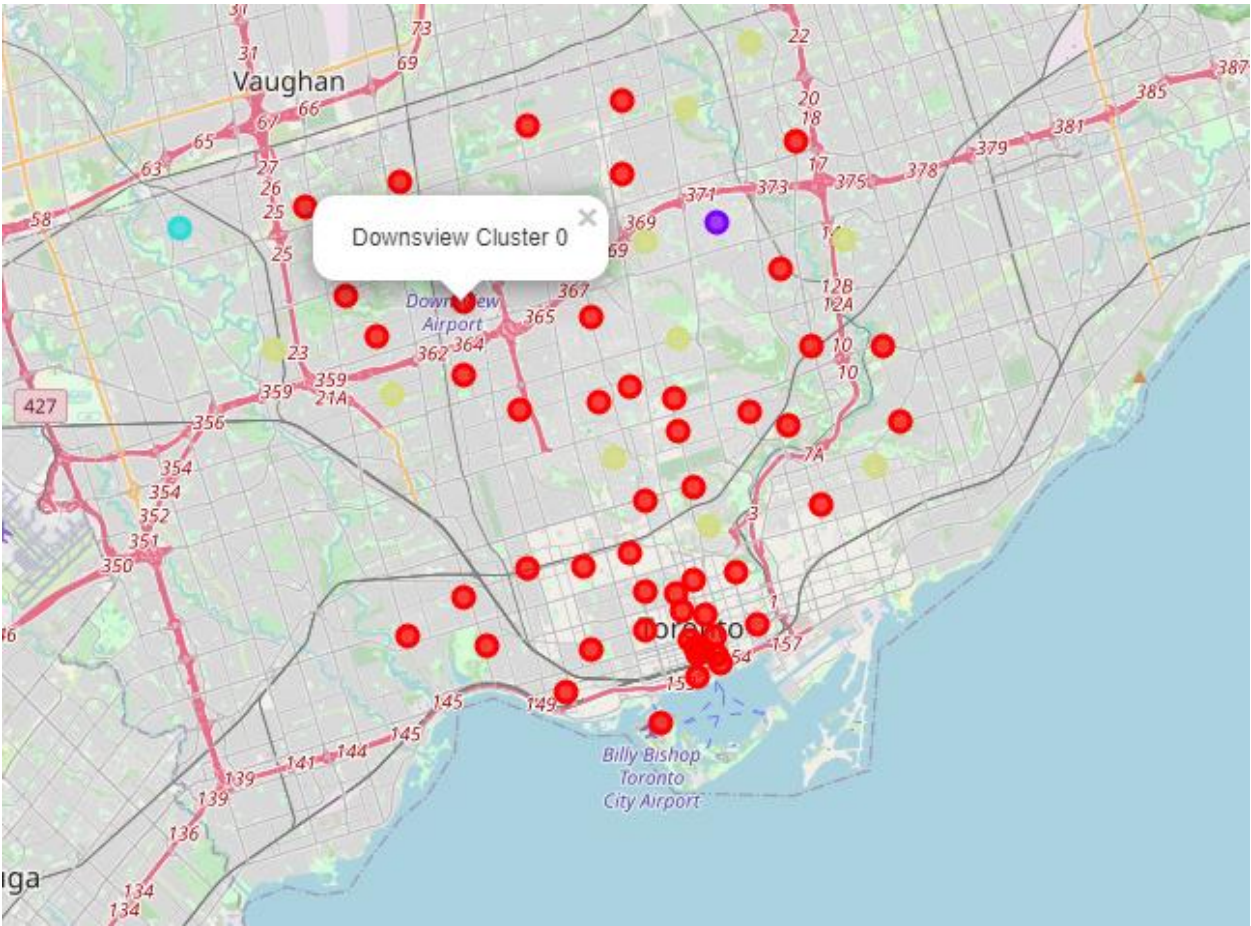
Toronto Map with Mapped Neighborhoods

We have also calculated the number of neighborhoods per cluster. Upon examining all clusters, we have determined the following:

- Cluster 1 – very busy cluster of neighborhoods, filled with coffee shops and restaurants.
- Cluster 2 – unique cluster that offer park, pool and zoo.
- Cluster 3 – unique cluster with arts and crafts, bakery and a park.
- Cluster 4 – more quieter cluster with neighborhoods filled with park, book stores, parks, but, as well offering the benefit of coffee shops and restaurants.

```python
# finding out how many neighborhoods are in each cluster

for cluster_num in range(kclusters):
    num_of_nbh = toronto_merged[toronto_merged['Cluster Labels'] == cluster_num].shape[0]
    print('The number of neighbourhoods in cluster {} is {}'.format(cluster_num+1, num_of_nbh))
```

```
The number of neighbourhoods in cluster 1 is 50
The number of neighbourhoods in cluster 2 is 1
The number of neighbourhoods in cluster 3 is 1
The number of neighbourhoods in cluster 4 is 11
```

Code Snippet

Neighborhood	NeighbLat	NeighbLong	Borough	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
Victoria Village	43.725882	-79.315572	North York	0	Coffee Shop	Hockey Arena	Boxing Gym	Lounge	Café	Park	Golf Course	Men's Store	Grocery Store	Gym / Fitness Center
Regent Park, Harbourfront	43.654260	-79.360636	Downtown Toronto	0	Coffee Shop	Theater	Restaurant	Park	Café	Diner	Bakery	Italian Restaurant	Pub	Breakfast Spot
Lawrence Manor, Lawrence Heights	43.718518	-79.464763	North York	0	Clothing Store	Fast Food Restaurant	Coffee Shop	Restaurant	Accessories Store	Dessert Shop	Sushi Restaurant	Furniture / Home Store	Fried Chicken Joint	Vietnamese Restaurant
Queen's Park, Ontario Provincial Government	43.662301	-79.389494	Downtown Toronto	0	Coffee Shop	Sushi Restaurant	Park	Japanese Restaurant	Ramen Restaurant	Gastropub	Café	Thai Restaurant	Bookstore	Italian Restaurant
Don Mills	43.745906	-79.352188	North York	0	Coffee Shop	Restaurant	Japanese Restaurant	Burger Joint	Bank	Supermarket	Asian Restaurant	Gym	Beer Store	Pizza Place
Don Mills	43.725900	-79.340923	North York	0	Coffee Shop	Restaurant	Japanese Restaurant	Burger Joint	Bank	Supermarket	Asian Restaurant	Gym	Beer Store	Pizza Place
Parkview Hill, Woodbine Gardens	43.706397	-79.309937	East York	0	Fast Food Restaurant	Pizza Place	Brewery	Gym / Fitness Center	Athletics & Sports	Pet Store	Bakery	Coffee Shop	Café	Rock Climbing Spot
Garden District, Ryerson	43.657162	-79.378937	Downtown Toronto	0	Coffee Shop	Gastropub	Japanese Restaurant	Café	Hotel	Italian Restaurant	Restaurant	Diner	Theater	Park

Cluster 1

Neighborhood	NeighbLat	NeighbLong	Borough	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
York Mills, Silver Hills	43.75749	-79.374714	North York	1	Park	Pool	Zoo	Ethiopian Restaurant	Donut Shop	Dry Cleaner	Dumpling Restaurant	Eastern European Restaurant	Electronics Store	Elementary School

Cluster 2

Neighborhood	NeighbLat	NeighbLong	Borough	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
Humber Summit	43.756303	-79.565963	North York	2	Pizza Place	Bank	Arts & Crafts Store	Bakery	Park	Shopping Mall	Electronics Store	Pharmacy	Italian Restaurant	Falafel Restaurant

Cluster 3

Neighborhood	NeighbLat	NeighbLong	Borough	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
Parkwoods	43.753259	-79.329656	North York	3	Park	Convenience Store	Shopping Mall	Pharmacy	Bus Stop	Laundry Service	Café	Caribbean Restaurant	Tennis Court	Discount Store
Woodbine Heights	43.695344	-79.318389	East York	3	Coffee Shop	Park	Pizza Place	Skating Rink	Sandwich Place	Pub	Athletics & Sports	Farmers Market	Plaza	Café
Hillcrest Village	43.803762	-79.363452	North York	3	Park	Coffee Shop	Pharmacy	Restaurant	Shopping Mall	Chinese Restaurant	Sandwich Place	Korean Restaurant	Bank	Bakery
Bathurst Manor, Wilson Heights, Downsview North	43.754328	-79.442259	North York	3	Coffee Shop	Bank	Convenience Store	Supermarket	Frozen Yogurt Shop	Fried Chicken Joint	Dog Run	Sandwich Place	Sushi Restaurant	Shopping Mall
Bayview Village	43.786947	-79.385975	North York	3	Bank	Grocery Store	Gas Station	Japanese Restaurant	Chinese Restaurant	Restaurant	Park	Intersection	Café	Trail
North Park, Maple Leaf Park, Upwood Park	43.713756	-79.490074	North York	3	Coffee Shop	Chinese Restaurant	Pizza Place	Park	Gas Station	Bakery	Athletics & Sports	Mediterranean Restaurant	Dim Sum Restaurant	Convenience Store
Humberlea, Emery	43.724766	-79.532242	North York	3	Golf Course	Convenience Store	Storage Facility	Bakery	Park	Intersection	Gas Station	Discount Store	Zoo	Ethiopian Restaurant
Lawrence Park	43.728020	-79.388790	Central Toronto	3	Park	Bookstore	Café	Trail	Gym / Fitness Center	Coffee Shop	College Quad	College Gym	Bus Line	Farm
York Mills West	43.752758	-79.400049	North York	3	Coffee Shop	Park	Restaurant	Convenience Store	Business Service	Bowling Alley	French Restaurant	Gas Station	Bank	Grocery Store
Forest Hill North & West	43.696948	-79.411307	Central Toronto	3	Park	Bank	Coffee Shop	Café	Gym / Fitness Center	Sushi Restaurant	Skating Rink	Burger Joint	Pharmacy	Japanese Restaurant
Rosedale	43.679563	-79.377529	Downtown Toronto	3	Coffee Shop	Park	Grocery Store	Sandwich Place	Bank	Mattress Store	BBQ Joint	Candy Store	Japanese Restaurant	Breakfast Spot

Cluster 4

Analysis of data based on certain venue criteria

As you remember, earlier we have created data frame consisting of most desired boroughs, their corresponding neighborhoods and venues data.

We also have a specific list of venue categories that are detrimental in figuring out which neighborhood is the best to settle in. Those categories are listed in the very beginning of this report and are comprised of:

- o Movie Theater
- o Dance Studio
- o Food
  - ▪ Gluten-free Restaurant
  - ▪ Vegetarian / Vegan Restaurant
- o Coffee Shop
- o Grocery Store
- o Pharmacy
- o Shopping Mall
- o Nightlife Spot
- o Gym / Fitness Center
- o Park

Let's dive into our data frame and solve the mystery! Which neighborhood if the best to settle in?

To start with – we are filtering data frame of most desired boroughs and getting all the venues matching our detrimental venue categories.

```
# setting search criteria as a list of venue categories important for a customer
# generating a DF with the list of these venue categories

search_values = ['Movie Theater', 'Dance Studio', 'Gluten-free Restaurant', 'Vegetarian / Vegan Restaurant', 'Coffee Shop', 'Grocery Store', 'Pharmacy', 'Shopping Mall', 'Nightlife Spot', 'Gym / Fitness Center', 'Park']
specific_boroughs=df_most_desired_boroughs[df_most_desired_boroughs.VenCat.str.contains('|'.join(search_values))]
specific_boroughs.head(10)
```

Code Snippet

	Neighborhood	NeighbLat	NeighbLong	Venue	VenLat	VenLong	VenCat	Borough
1	Parkwoods	43.753259	-79.329656	Brookbanks Park	43.751976	-79.332140	Park	North York
4	Parkwoods	43.753259	-79.329656	Bruno's valu-mart	43.746143	-79.324630	Grocery Store	North York
6	Parkwoods	43.753259	-79.329656	Shoppers Drug Mart	43.745315	-79.325800	Pharmacy	North York
8	Parkwoods	43.753259	-79.329656	Shoppers Drug Mart	43.760857	-79.324961	Pharmacy	North York
14	Parkwoods	43.753259	-79.329656	Sandover Park	43.760277	-79.333305	Park	North York
19	Parkwoods	43.753259	-79.329656	La Notre	43.760704	-79.325396	Coffee Shop	North York
21	Parkwoods	43.753259	-79.329656	Parkwoods Village Centre	43.760735	-79.324873	Shopping Mall	North York
25	Parkwoods	43.753259	-79.329656	Broadlands Park	43.746746	-79.322502	Park	North York
28	Parkwoods	43.753259	-79.329656	Donwood Plaza	43.745369	-79.325435	Shopping Mall	North York
30	Victoria Village	43.725882	-79.315572	Tim Hortons	43.725517	-79.313103	Coffee Shop	North York

Data Frame of the List of Required Venues within the Most Desired Boroughs

We have 712 venues total in 5 boroughs that are matching the categories of venues.

```
In [72]: specific_boroughs.shape
Out[72]: (712, 8)
```

Code Snippet

Among all 5 boroughs – what are the most popular venue category? It is coffee shop, of course. This is not surprising, especially knowing that coffee shops are the most popular venue category among all available boroughs.

VenCat	Neighborhood	NeighbLat	NeighbLong	Venue	VenLat	VenLong	Borough
Coffee Shop	333	333	333	333	333	333	333
Dance Studio	12	12	12	12	12	12	12
Grocery Store	79	79	79	79	79	79	79
Gym / Fitness Center	34	34	34	34	34	34	34
Indie Movie Theater	3	3	3	3	3	3	3
Movie Theater	12	12	12	12	12	12	12
Park	118	118	118	118	118	118	118
Pharmacy	50	50	50	50	50	50	50
Shopping Mall	28	28	28	28	28	28	28
Theme Park Ride / Attraction	1	1	1	1	1	1	1
Vegetarian / Vegan Restaurant	42	42	42	42	42	42	42

Data Frame of the Most Popular Required Venue Categories within Most Desired Boroughs

Let's list all the boroughs and see what they have to offer. This data frame depicts how many venues of each category exist in each borough. Look at Coffee Shops in Downtown Toronto! 151 total of Coffee shops only!

Borough	VenCat	Neighborhood	NeighbLat	NeighbLong	Venue	VenLat	VenLong
Central Toronto	Coffee Shop	44	44	44	44	44	44
	Grocery Store	11	11	11	11	11	11
	Gym / Fitness Center	7	7	7	7	7	7
	Indie Movie Theater	1	1	1	1	1	1
	Movie Theater	4	4	4	4	4	4
	Park	20	20	20	20	20	20
	Pharmacy	11	11	11	11	11	11
	Vegetarian / Vegan Restaurant	9	9	9	9	9	9
Downtown Toronto	Coffee Shop	151	151	151	151	151	151
	Dance Studio	10	10	10	10	10	10
	Grocery Store	20	20	20	20	20	20
	Gym / Fitness Center	13	13	13	13	13	13
	Indie Movie Theater	1	1	1	1	1	1
	Movie Theater	4	4	4	4	4	4
	Park	44	44	44	44	44	44
	Pharmacy	1	1	1	1	1	1
	Shopping Mall	7	7	7	7	7	7
	Vegetarian / Vegan Restaurant	25	25	25	25	25	25
East York	Coffee Shop	21	21	21	21	21	21
	Grocery Store	7	7	7	7	7	7
	Gym / Fitness Center	2	2	2	2	2	2
	Park	5	5	5	5	5	5
	Pharmacy	5	5	5	5	5	5
	Shopping Mall	4	4	4	4	4	4
North York	Coffee Shop	82	82	82	82	82	82
	Grocery Store	36	36	36	36	36	36
	Gym / Fitness Center	10	10	10	10	10	10
	Movie Theater	3	3	3	3	3	3
	Park	34	34	34	34	34	34
	Pharmacy	29	29	29	29	29	29
	Shopping Mall	17	17	17	17	17	17
West Toronto	Coffee Shop	35	35	35	35	35	35
	Dance Studio	2	2	2	2	2	2
	Grocery Store	5	5	5	5	5	5
	Gym / Fitness Center	2	2	2	2	2	2
	Indie Movie Theater	1	1	1	1	1	1
	Movie Theater	1	1	1	1	1	1
	Park	15	15	15	15	15	15
	Pharmacy	4	4	4	4	4	4
	Theme Park Ride / Attraction	1	1	1	1	1	1
	Vegetarian / Vegan Restaurant	8	8	8	8	8	8

Data Frame of the Required Venue Categories per Borough

Next question we will answer would be – there are 11 unique venue categories detrimental to our main goal. Is there any borough that has all 11 categories? Apparently, no. The closest boroughs to satisfy the requirement are Downtown Toronto and West Toronto. This is in line with our previous observation - these 2 boroughs have the highest number of unique venues and overall venues in total. The third on the list – Central Toronto with 8 unique categories present. These 3 boroughs will be our main focus further on.

```
Borough
Central Toronto     8
Downtown Toronto    10
East York           6
North York          7
West Toronto        10
Name: VenCat, dtype: int64
```

List of Unique Venue Categories per Borough

We will come down to the level of neighborhoods – is there any neighborhood is all 5 boroughs that satisfies the requirement of all 11 unique venue categories? And the answer is – no. There are 3 neighborhoods with maximum 7 unique venue categories.

Neighborhood	VenCat
Queen's Park, Ontario Provincial Government	7
Willowdale	7
Garden District, Ryerson	7
Harbourfront East, Union Station, Toronto Islands	6
Davisville North	6
Regent Park, Harbourfront	6
Summerhill West, Rathnelly, South Hill, Forest Hill SE, Deer Park	6
Runnymede, Swansea	6
First Canadian Place, Underground city	6
Downsview	6

Data Frame of the Number of Unique Venue Categories per Neighborhood

For the next step of our research – we shall zero in 3 boroughs that have the highest number of unique categories on our list – Downtown Toronto, West Toronto and Central Toronto. We

have filtered a list of all neighborhoods to get only those within these 3 boroughs. Let's map those on Toronto map to get better visual picture.

Neighborhood	VenCat	Postal code	Borough	Latitude	Longitude
Queen's Park, Ontario Provincial Government	7	M7A	Downtown Toronto	43.662301	-79.389494
Garden District, Ryerson	7	M5B	Downtown Toronto	43.657162	-79.378937
Davisville North	6	M4P	Central Toronto	43.712751	-79.390197
Regent Park, Harbourfront	6	M5A	Downtown Toronto	43.654260	-79.360636
Summerhill West, Rathnelly, South Hill, Forest Hill SE, Deer Park	6	M4V	Central Toronto	43.686412	-79.400049
Church and Wellesley	6	M4Y	Downtown Toronto	43.665860	-79.383160
Harbourfront East, Union Station, Toronto Islands	6	M5J	Downtown Toronto	43.640816	-79.381752
St. James Town	6	M5C	Downtown Toronto	43.651494	-79.375418
Runnymede, Swansea	6	M6S	West Toronto	43.651571	-79.484450
First Canadian Place, Underground city	6	M5X	Downtown Toronto	43.648429	-79.382280

Data Frame of the Unique Venue Categories in Downtown Toronto, West Toronto and Central Toronto
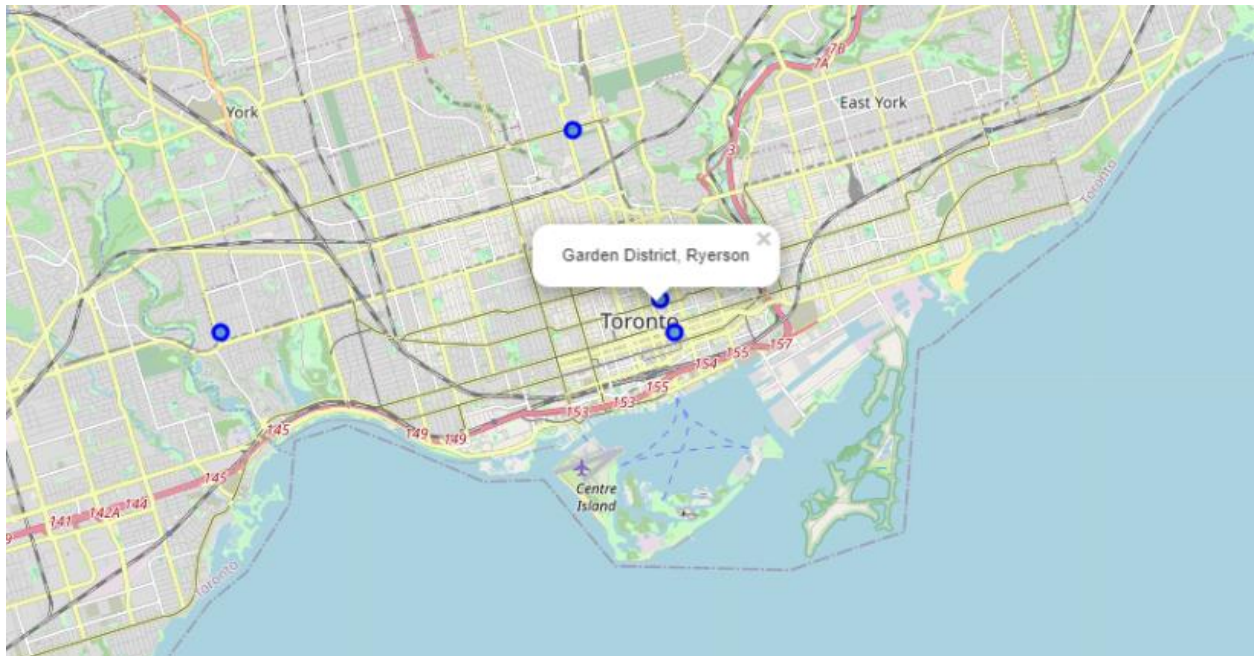

Mapped Neighborhoods

Now, let's try to answer the main question – where would the person want to settle based on the list of required venues being available around? Out of all 103 neighborhoods with over 4,000 venues available – we have short listed 5 most probably neighborhoods to settle. They offer availability of 6-7 required venue categories. But, which one is closer to Toronto Downtown? Well – we are in luck – 2 neighborhoods with the highest number of unique venue categories are actually in Toronto Downtow. Let's map those.

	Neighborhood	VenCat	Postal code	Borough	Latitude	Longitude
0	Queen's Park, Ontario Provincial Government	7	M7A	Downtown Toronto	43.662301	-79.389494
1	Garden District, Ryerson	7	M5B	Downtown Toronto	43.657162	-79.378937
4	Summerhill West, Rathnelly, South Hill, Forest...	6	M4V	Central Toronto	43.686412	-79.400049
7	St. James Town	6	M5C	Downtown Toronto	43.651494	-79.375418
8	Runnymede, Swansea	6	M6S	West Toronto	43.651571	-79.484450

Data Frame of the Neighborhoods with the Highest Number of Unique Venue Categories

Mapped Neighborhoods

4. Discussion

We have performed various analysis to answer the following question – based on the list of required venue categories – which neighborhood will be the best to settle in?

Ultimately – there are options. Based on thorough and detailed analysis, we have determined that 3 boroughs are populated with unique venue categories and overall number od venues the most. Those are:

- Central Toronto
- Downtown Toronto
- West Toronto

However, unfortunately, none of boroughs and none of neighborhoods within those boroughs offer access to all required venue categories. **Our recommendation would be to pick Downtown Toronto as a borough and Garden District as a neighborhood**. Even though Garden Distric has only 7 unique venue categories, it is in the close proximity from other neighborhoods providing the exposure to other venues and venue types.