Investment in Popular Venues of Neighborhoods

Manoj Banisetty

May 20, 2020

1. Introduction

1.1 Background

Rich people are all across the globe. They earn very much. The main aim of everybody is to have money and should be able to double what we have. So , that's where the concept of investment came into play. Everybody likes to invest what they have in something they like to earn more from that investment. So, investment is important in everyone's life. Investment can be done in anything now. Be it a Corporate Company , Multi National Company , or it could even be in restaurants, shops,etc. Anything a person believes they can gain from, they will invest in it.

1.2 Problem

Data that might help investors who are looking to invest in anything which is popular around a particular city/neighborhoods – in our case, we will be checking the neighborhoods of Toronto- to understand the people living in that neighborhood. We can simply understand it by the derived data by estimating what business is more prevailing in that city/neighborhood so that the investors can invest in that business.

1.3 Target Audience

The advantage in this problem is that we don't need to search for a particular person who is looking to invest in a particular stream. Instead, any person who wants to invest in any business can look up at this data to understand and decide where they will invest their investments. Any person who is looking to invest in anything can refer to these conclusive datasets and come up with a plan to setup a business at a particular neighborhood where there is more demand for a particular product.

2. Data acquisition and cleaning

2.1 Data sources

I have obtained the required data from the following web pages:

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

	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

This link consists of all the Neighborhoods, Boroughs along with their postal codes.

https://cocl.us/Geospatial_data

	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

This CSV file consists of all the geographical coordinates of the Neighborhoods, Boroughs, etc.

2.2 Data cleaning

1) Firstly, we will remove all the rows which contain the ‘Not assigned’ or ‘NaN’ values from the postal codes dataset.

2) Secondly, we will do the same with the dataset containing the geographical coordinates of the Neighborhoods and all the venues present in their particular neighborhoods.

3) Lastly, we can merge the 2 data sets into one by using the common column in both the data sets which is the Postal Code column which in future will help us locate the coordinates on the map.

	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
5	M9A	Etobicoke	Islington Avenue
6	M1B	Scarborough	Malvern, Rouge
7	M3B	North York	Don Mills
8	M4B	East York	Parkview Hill, Woodbine Gardens
9	M5B	Downtown Toronto	Garden District, Ryerson
10	M6B	North York	Glencairn

2.3 Feature selection

The features required in the selection from the data is we need all the coordinates of the different venues present in different neighborhoods.

We have to particularly select a neighborhood and go through all of the various venue categories in order to arrive to a particular decision and mapping those venues.

3. Methodolgy

3.1 Exploratory Data Analysis

Exploring the data present is essential as we have to know what we are dealing with first before dealing with it. So , we have to divide the data we have into data sets before arriving to the dataset which we require to complete the task.

So, the first data set contains the names of various neighborhoods present in Toronto along with their respective Boroughs and Postal Codes.

	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

After cleaning the dataset :

	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
5	M9A	Etobicoke	Islington Avenue
6	M1B	Scarborough	Malvern, Rouge
7	M3B	North York	Don Mills
8	M4B	East York	Parkview Hill, Woodbine Gardens
9	M5B	Downtown Toronto	Garden District, Ryerson
10	M6B	North York	Glencairn

Now, the second data set contains the geographical coordinates which are linked to the Neighborhoods , Boroughs from the first dataset.

	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

So, in order to combine these 2 datasets and link the respective neighborhoods with their respective geographical coordinates , we will be using the ‘merge’ function from the pandas library.
Now as you notice that we have the ‘Postal Code’ column common in both the datasets, we can use it to merge these 2 datasets and arrive to the dataset which we require to map the 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
5	M9A	Etobicoke	Islington Avenue	43.667856	-79.532242
6	M1B	Scarborough	Malvern, Rouge	43.806686	-79.194353
7	M3B	North York	Don Mills	43.745906	-79.352188
8	M4B	East York	Parkview Hill, Woodbine Gardens	43.706397	-79.309937
9	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937

As we now have the complete dataset, we can now map them using ‘folium’ library.

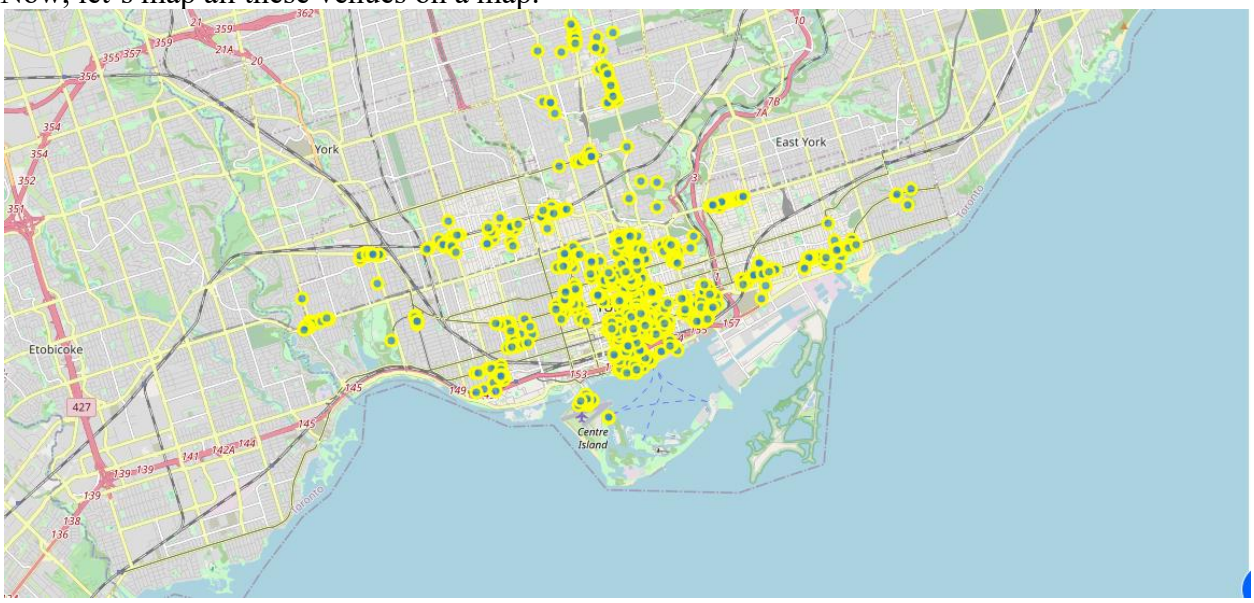
Map of Totonto Neighborhood :

Now, we are going to use “FourSquare API” to retrieve all the required venues locations and form another dataset.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Regent Park, Harbourfront	43.65426	-79.360636	Roselle Desserts	43.653447	-79.362017	Bakery
1	Regent Park, Harbourfront	43.65426	-79.360636	Tandem Coffee	43.653559	-79.361809	Coffee Shop
2	Regent Park, Harbourfront	43.65426	-79.360636	Morning Glory Cafe	43.653947	-79.361149	Breakfast Spot
3	Regent Park, Harbourfront	43.65426	-79.360636	Cooper Koo Family YMCA	43.653249	-79.358008	Distribution Center
4	Regent Park, Harbourfront	43.65426	-79.360636	Body Blitz Spa East	43.654735	-79.359874	Spa
5	Regent Park, Harbourfront	43.65426	-79.360636	Impact Kitchen	43.656369	-79.356980	Restaurant
6	Regent Park, Harbourfront	43.65426	-79.360636	Corktown Common	43.655618	-79.356211	Park
7	Regent Park, Harbourfront	43.65426	-79.360636	The Extension Room	43.653313	-79.359725	Gym / Fitness Center
8	Regent Park, Harbourfront	43.65426	-79.360636	The Distillery Historic District	43.650244	-79.359323	Historic Site
9	Regent Park, Harbourfront	43.65426	-79.360636	Figs Breakfast & Lunch	43.655675	-79.364503	Breakfast Spot
10	Regent Park, Harbourfront	43.65426	-79.360636	Distillery Sunday Market	43.650075	-79.361832	Farmers Market
11	Regent Park, Harbourfront	43.65426	-79.360636	SOMA chocolatemaker	43.650622	-79.358127	Chocolate Shop
12	Regent Park, Harbourfront	43.65426	-79.360636	Rooster Coffee	43.651900	-79.365609	Coffee Shop
13	Regent Park, Harbourfront	43.65426	-79.360636	Starbucks	43.651613	-79.364917	Coffee Shop
14	Regent Park, Harbourfront	43.65426	-79.360636	Dominion Pub and Kitchen	43.656919	-79.358967	Pub

As we can see , we have all the data we need to locate and retrieve and map all the venues present on a map.

Now, let’s map all these venues on a map.



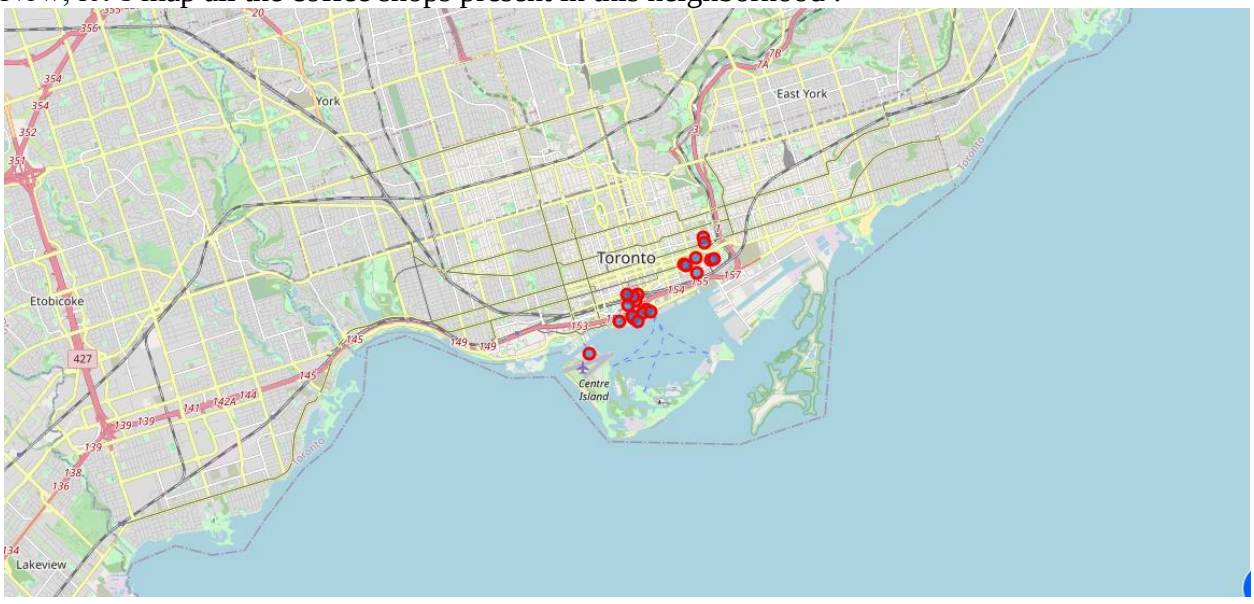
Now, we can divide the data we have into various components and check out the neighborhoods. For example :

Let us take all the neighborhoods which are at ‘Harbourfront’ and check which business is more active in those neighborhoods.

	Venue Category
Coffee Shop	22
Café	6
Aquarium	5
Park	5
Hotel	4
Bakery	4
Restaurant	4
Scenic Lookout	3
Brewery	3
Italian Restaurant	3
Pub	3

As we can see coffee shops are more active and more in number in these areas . So, now the investor can decide whether or not it is safe to invest or setup a business which he wants in this neighborhood or not.

Now, let's map all the coffee shops present in this neighborhood :



Now,let us take again all the neighborhoods which are now at 'Riverdale' and check which business is more active in those neighborhoods.

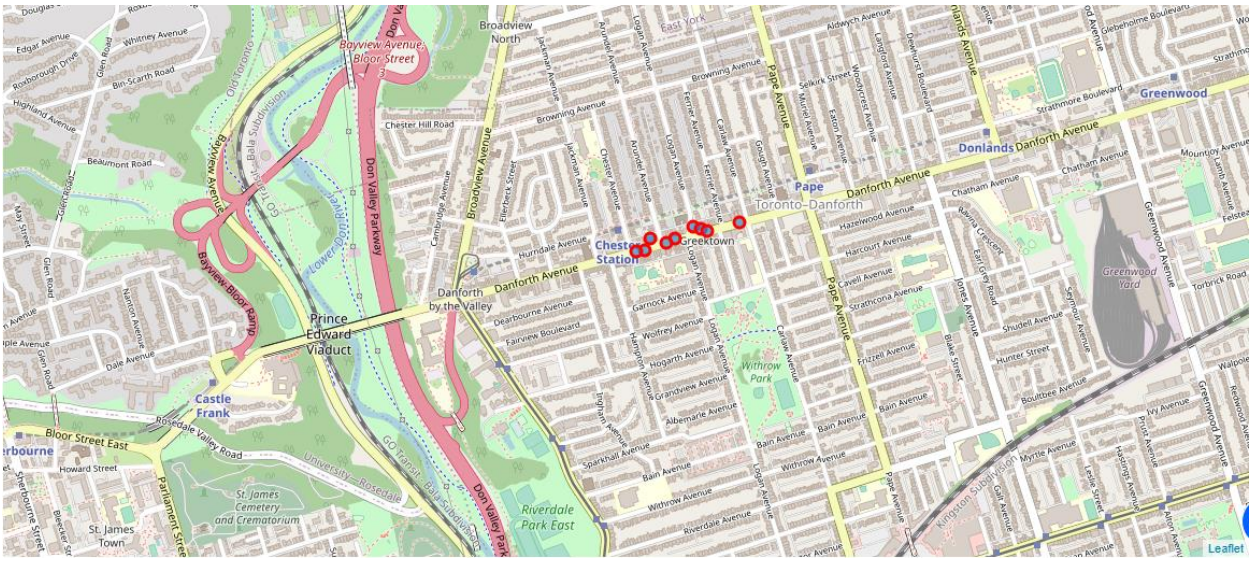
	index	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	645	The Danforth West, Riverdale	43.679557	-79.352188	MenEssentials	43.677820	-79.351265	Cosmetics Shop
1	646	The Danforth West, Riverdale	43.679557	-79.352188	Pantheon	43.677621	-79.351434	Greek Restaurant
2	647	The Danforth West, Riverdale	43.679557	-79.352188	Cafe Fiorentina	43.677743	-79.350115	Italian Restaurant
3	648	The Danforth West, Riverdale	43.679557	-79.352188	Dolce Gelato	43.677773	-79.351187	Ice Cream Shop
4	649	The Danforth West, Riverdale	43.679557	-79.352188	La Diperie	43.677530	-79.352295	Ice Cream Shop
5	650	The Danforth West, Riverdale	43.679557	-79.352188	Moksha Yoga Danforth	43.677622	-79.352116	Yoga Studio
6	651	The Danforth West, Riverdale	43.679557	-79.352188	Mezes	43.677962	-79.350196	Greek Restaurant
7	652	The Danforth West, Riverdale	43.679557	-79.352188	Louis Cifer Brew Works	43.677663	-79.351313	Brewery
8	653	The Danforth West, Riverdale	43.679557	-79.352188	Alexandros	43.678304	-79.349486	Greek Restaurant
9	654	The Danforth West, Riverdale	43.679557	-79.352188	Valley Farm Produce	43.677999	-79.349969	Fruit & Vegetable Store
10	655	The Danforth West, Riverdale	43.679557	-79.352188	7 Numbers	43.677062	-79.353934	Italian Restaurant
11	656	The Danforth West, Riverdale	43.679557	-79.352188	The Auld Spot Pub	43.677335	-79.353130	Pub
12	657	The Danforth West, Riverdale	43.679557	-79.352188	Caffé Demetre	43.677683	-79.351608	Dessert Shop
13	658	The Danforth West, Riverdale	43.679557	-79.352188	Rikkochez	43.677267	-79.353274	Restaurant
14	659	The Danforth West, Riverdale	43.679557	-79.352188	Pizzeria Libretto	43.678489	-79.347576	Pizza Place
15	660	The Danforth West, Riverdale	43.679557	-79.352188	The Big Carrot Organic Juice Bar	43.677438	-79.352683	Juice Bar
16	661	The Danforth West, Riverdale	43.679557	-79.352188	Re: Reading	43.678507	-79.347678	Bookstore
17	662	The Danforth West, Riverdale	43.679557	-79.352188	Messini Authentic Gyros	43.677827	-79.350569	Greek Restaurant
18	663	The Danforth West, Riverdale	43.679557	-79.352188	Don Valley Trail	43.676331	-79.353923	Trail

Now, we will see what business is more active :

	Venue Category
Greek Restaurant	9
Coffee Shop	3
Italian Restaurant	3
Furniture / Home Store	2
Restaurant	2
Ice Cream Shop	2
Fruit & Vegetable Store	1
Pizza Place	1
Liquor Store	1
American Restaurant	1
Trail	1
Yoga Studio	1
Bakery	1
Bookstore	1
Frozen Yogurt Shop	1
Cosmetics Shop	1
Café	1
Bubble Tea Shop	1
Brewery	1

There are more ‘Greek Restaurants’ in this area.

Let us map those restaurants now :



In this way we can check out for each neighborhood and present them to your investor , which will help them very much in understanding the type of businesses which are more active in those areas and where to invest to gain more returns.

4. Results :

From the above obtained data, we can safely say that any person who is looking to invest in various businesses in the city of Toronto can rely on outputs of the above data.
This will give them a clear understanding of how those neighborhoods are reacting to different businesses and whether are not the business which is the investor is willing to open will be compatible enough to survive in those neighborhoods for a very long time.

Thus, a picture can be framed as to where would the businesses of a certain kind can survive for a long time if they are opened in various neighborhoods.

5. Discussions :

As I said , this problem has a very good advantage where not only a certain kind of people can look into it, anyone who is willing to take up any ideas on any businesses can even experiment in investing in various neighborhoods and still earn what they want to.
Many investors who have a similar idea of opening a business can even come together and share their opinions if they are willing to partnership. In this way we can check out for each neighborhood and present them to your investor , which will help them very much in understanding the type of businesses which are more active in those areas and where to invest.

6. Conclusion:

In this study I have analyzed the data which contains various neighborhoods and boroughs and their respective coordinates. I've used FourSquare API to retrieve those coordinates and plot them on a map using the folium library. This data can help any person to help with their doubt of investments and gives a particular direction to invest the money on a particular venue.
For example, if you want to invest in any kind of business in a particular neighborhood , say, Riverdale .From the data we have obtained we can suggest that investor to open a 'Greek Restaurant' as they have a high demand in that place. If that person doesn't have enough money to open up a restaurant we can go to the second choice which is the 'Coffee Shop'.

The models of maps in this study mainly focused on neighborhoods and venues. You can even perform the above operations with appropriate data on various major cities and use the data to attract various investors in investing into your firm.