Where to open your business office ?

(Canada- Toronto)

Bola Ashraf

August 6, 2020

1. Introduction

1.1 Background

How will you go about where to build your business office in Toronto? Well it is hard to determine the factors that this choice depend on.
So, I do my best to help you to choose the right place.

1.2 Problem

So, in **Canada - Toronto**, you must carefully choose the right **Borough** and **Neighborhood**:-

- The **Borough** must have a lot of **Hotels** in it for your clients.
- The **Neighborhoods** in the **Borough** must be as **safe** as possible.
- Then you must take the **best mixture** of these factors above.

2. Data acquisition and cleaning

2.1 Data sources

in this project I will use 2 datasets (CSV Files) and 1 Geojson file, also I will use Foursquare API**:-**

- **Kaggle:** I will use this Dataset that describes the crimes in Toronto from 2014 to 2019, with deleting unnecessary columns.
- **Wikipedia:** I will use this Dataset because the neighborhoods names in Kaggle dataset is unclear, I will use this for fixing it.
- **GitHub:** I will use this GeoJson File for get the geographical boundaries of each neighborhood for coming maps.
- **Foursquare:** It will be used for exploring Toronto and venues within each neighborhood.

2.2 Data cleaning

There is a lot of unnecessary columns in the above datasets, so I will just delete these columns**.**

3. Exploratory Data Analysis

3.1 After Cleaning datasets

Kaggle data frame: which give us the details about crimes in Toronto from 2014 to 2019.

Hood_ID	Lat	Long	Neighbourhood	offence	occurrencehour	occurrencedayofweek	occurrencedate	reportedhour	reporteddayofweek
1	43.748962	-79.602997	West Humber-Clairville (1)	Assault Bodily Harm	19	Sunday	2016-05-08T19:00:00.000Z	0	Monday
1	43.723656	-79.619942	West Humber-Clairville (1)	Robbery - Swarming	23	Saturday	2014-05-24T23:35:00.000Z	0	Sunday
1	43.719158	-79.603462	West Humber-Clairville (1)	Theft Of Motor Vehicle	23	Wednesday	2018-09-26T23:00:00.000Z	4	Friday
1	43.733555	-79.619995	West Humber-Clairville (1)	B&E	0	Tuesday	2015-09-29T00:25:00.000Z	0	Tuesday
1	43.698036	-79.586571	West Humber-Clairville (1)	Assault	14	Saturday	2018-03-03T14:00:00.000Z	14	Saturday
1	43.699749	-79.601471	West Humber-Clairville (1)	Assault	20	Tuesday	2014-03-11T20:00:00.000Z	14	Wednesday
1	43.686760	-79.595566	West Humber-Clairville (1)	Theft Of Motor Vehicle	18	Friday	2019-09-06T18:08:00.000Z	18	Friday
1	43.715561	-79.594795	West Humber-Clairville (1)	Theft Of Motor Vehicle	0	Saturday	2019-01-26T00:00:00.000Z	20	Saturday
1	43.737843	-79.610771	West Humber-Clairville (1)	Theft Over	11	Saturday	2018-09-01T11:17:00.000Z	13	Thursday
1	43.743919	-79.605499	West Humber-Clairville (1)	Theft Of Motor Vehicle	11	Saturday	2019-01-26T11:30:00.000Z	12	Saturday

Wikipedia data frame: which give us the details about each **borough** and **neiborhood** in Toronto.

BoroughID	Borough	Neighborhoods
1	West Humber-Clairville	West Humber-Clairville
2	Mount Olive-Silverstone-Jamestown	Smithfield
3	Thistletown-Beaumond Heights	Thistletown-Beaumond Heights
4	Rexdale-Kipling	Rexdale
5	Elms-Old Rexdale	The Elms,Rexdale
6	Kingsview Village-The Westway	Kingsview Village,Richview
7	Willowridge-Martingrove-Richview	Willowridge-Martingrove-Richview
8	Humber Heights-Westmount	Humber Heights-Westmount
9	Edenbridge-Humber Valley	Humber Valley
10	Princess-Rosethorn	Princess Anne Manor, Thorncrest Village,,Princ...

3.2 Work toward our goal

1- First I get **the number crimes committed** in each neighborhood by grouping the Kaggle data frame by Hood_ID which is unique number for every Borough in Toronto, then I check every group and get number of rows in it which describes the number of crimes committed in this borough. Finally, I add this number or crimes committed column to the previous data frame.

BoroughID	Borough	Neighborhoods	crimes_from_2014_to_2019	lat	lng
77	Waterfront Communities-The Island	Distillery District, Old Town, St. Lawrence	7747	43.705953	-79.396487
76	Bay Street Corridor	Bay Street, Financial District	6817	43.705087	-79.398437
75	Church-Yonge Corridor	Church,Wellesley,Ryerson University	6232	43.705884	-79.393832
1	West Humber-Clairville	West Humber-Clairville	5702	43.712859	-79.368867
73	Moss Park	Moss Park, Corktown,Garden District	4786	43.706760	-79.394414
27	York University Heights	York University Heights	3989	43.706610	-79.392474
26	Downsview-Roding-CFB	Downsview	3974	43.709237	-79.388584
78	Kensington-Chinatown	Alexandra Park, Chinatown, Grange Park, Kensin...	3823	43.707127	-79.395671
137	Woburn	Woburn	3798	43.706375	-79.401986
136	West Hill	West Hill	3497	43.704420	-79.401680

2- Second I search for **Hotels** with Foursquare API and get all hotels details such as their names and locations, then I sort them in a data frame.

	name	lat	lng
0	Hotel X	43.632886	-79.411770
1	Hotel Gelato	43.703478	-79.414311
2	Toronto Plaza Hotel	43.719001	-79.514489
3	The Drake Hotel	43.643147	-79.424597
4	Four Seasons Hotel Toronto	43.671796	-79.389457
5	Gladstone Hotel	43.642616	-79.426936
6	Thompson Hotel	43.642753	-79.401558
7	Le Germain Hotel Toronto Mercer	43.645756	-79.390904
8	SoHo Metropolitan Hotel	43.644633	-79.392145
9	Radisson Suite Hotel Toronto Airport	43.692173	-79.576619
10	Sandman Signature Toronto Airport Hotel	43.689918	-79.590427
11	Sheraton Gateway Hotel	43.686456	-79.620002
12	Sheraton Centre Toronto Hotel	43.650594	-79.384530
13	One King West Hotel & Residence	43.649139	-79.377876
14	The Omni King Edward Hotel	43.649191	-79.376006

3- Then the hard part comes. It is getting the **Borough** that each hotels lie in using its coordinates only. So to solve this problem I used my GeoJson File which contains each **Borough** and their Geographical Polygons (Geographical Boundaries), then I define a polygon for each borough and store them in Python Dictionary. Then I use **Shapely Python Library** to define polygons and loop over all hotels coordinates and check which hotels lie on any borough then is add final results to a new data frame.

	BoroughID	Borough	Neighborhoods	crimes_from_2014_to_2019	hotels_num	lat	lng
0	77	Waterfront Communities-The Island	Distillery District, Old Town, St. Lawrence	7747	0	43.705953	-79.396487
1	76	Bay Street Corridor	Bay Street, Financial District	6817	6	43.705087	-79.398437
2	75	Church-Yonge Corridor	Church,Wellesley,Ryerson University	6232	5	43.705884	-79.393832
3	1	West Humber-Clairville	West Humber-Clairville	5702	6	43.712859	-79.368867
4	73	Moss Park	Moss Park, Corktown,Garden District	4786	1	43.706760	-79.394414
5	27	York University Heights	York University Heights	3989	0	43.706610	-79.392474
6	26	Downsview-Roding-CFB	Downsview	3974	1	43.709237	-79.388584
7	78	Kensington-Chinatown	Alexandra Park, Chinatown, Grange Park, Kensin...	3823	2	43.707127	-79.395671
8	137	Woburn	Woburn	3798	0	43.706375	-79.401986
9	136	West Hill	West Hill	3497	0	43.704420	-79.401680

4- Then I **normalize** these values**.**

	BoroughID	Borough	Neighborhoods	crimes_from_2014_to_2019	hotels_num	lat	lng
0	77	Waterfront Communities-The Island	Distillery District, Old Town, St. Lawrence	1.000000	0.000000	43.705953	-79.396487
1	76	Bay Street Corridor	Bay Street, Financial District	0.874222	1.000000	43.705087	-79.398437
2	75	Church-Yonge Corridor	Church,Wellesley,Ryerson University	0.795104	0.833333	43.705884	-79.393832
3	1	West Humber-Clairville	West Humber-Clairville	0.723424	1.000000	43.712859	-79.368867
4	73	Moss Park	Moss Park, Corktown,Garden District	0.599540	0.166667	43.706760	-79.394414
5	27	York University Heights	York University Heights	0.491750	0.000000	43.706610	-79.392474
6	26	Downsview-Roding-CFB	Downsview	0.489721	0.166667	43.709237	-79.388584
7	78	Kensington-Chinatown	Alexandra Park, Chinatown, Grange Park, Kensin...	0.469299	0.333333	43.707127	-79.395671
8	137	Woburn	Woburn	0.465918	0.000000	43.706375	-79.401986
9	136	West Hill	West Hill	0.425210	0.000000	43.704420	-79.401680

4. Results and Visualizations

4.1 Results

With the previous data frame we almost near to finish. Which left is to find **Best Mixture** between

The **Safety** and **Means of Housing (Hotels)** for the convenience of your customers.

The method I use for achieving that is just supposing that **these two factors are 2d points** and I will calculate the **distance** between it and the best choice ever which is **0 Crimes and Max number of hotels**. Then **the borough with minimum distance result will be our choice to build our office in,** because that means that **this distance is the nearest to our best ideal choice.**

So after doing this method the result is **The Final Report Data Frame**s:-

1- The Top 5 Choices

	BoroughID	Borough	chance2d	crimes_from_2014_to_2019	hotels_num
78	84	Little Portugal	0.345148	1015	4
11	95	Annex	0.480998	2917	4
3	1	West Humber-Clairville	0.723424	5702	6
2	75	Church-Yonge Corridor	0.812384	6232	5
7	78	Kensington-Chinatown	0.815283	3823	2

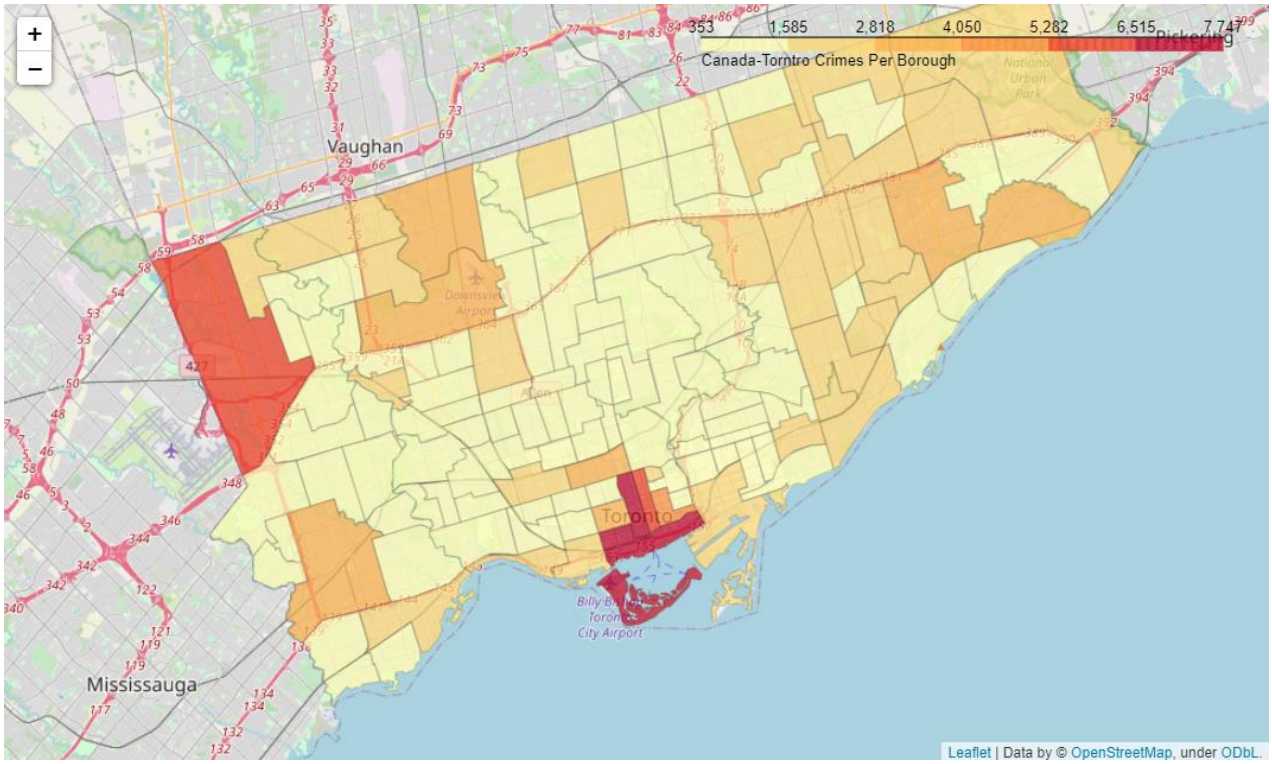
2- The Top Worst 5 Choices

	BoroughID	Borough	chance2d	crimes_from_2014_to_2019	hotels_num
0	77	Waterfront Communities-The Island	1.414214	7747	0
5	27	York University Heights	1.114369	3989	0
8	137	Woburn	1.103213	3798	0
9	136	West Hill	1.086648	3497	0
10	14	Islington-City Centre West	1.075852	3287	0

- I want to say that these results **make sense**, because obviously the worst choices we could ever choose is these Boroughs in Toronto that have **no Hotels** in it and have **the highest crimes-committed-rate**.

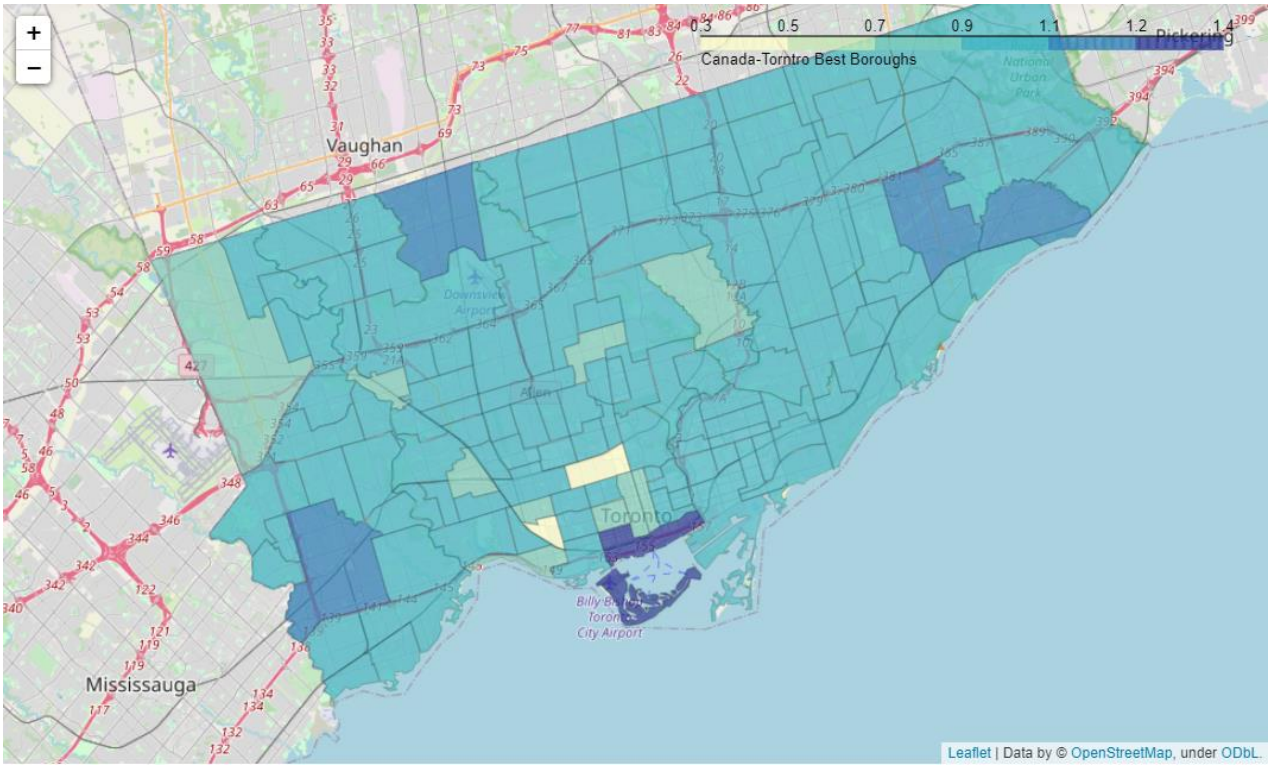
4.1 Visualization Maps

1- This map shows us the **Crimes Committed** in every **Borough** in Toronto



2- This map shows us the **Distribution of Hotels in Toronto** with markers on it**.**

3- Finally this map shows us the **Best Boroughs Choices** you can choose to build your office in



5. Observations

1- **Choosing any Boroughs from the east of Toronto will be as worst as possible**, because as shown in maps these boroughs have no Hotels, so you cannot meet your customers for business in these boroughs.

2- **For Safety choose the** boroughs in the center of Toronto,
For Means of Housing (Hotels) choose the boroughs in the west of Toronto

6. Conclusions

Finally, if you want to build your business office in Toronto, your best boroughs will be **Waterfront Communities-The Island** and **Bay Street Corridor** as shown above.