Capstone Project - The Battle of the Neighborhoods:

Introduction: Business Problem:

In this project we are going to find a location in Toronto to open a restaurant. This report will be targeted to stakeholders interested in opening a **Chinese restaurant** in **Toronto**, Canada.

Since Toronto is a well-diversified city, you can find cuisines from all over the world. Toronto has hundreds of thousands Chinese people. There are lots of authentic Chinese food in Greater Toronto Area. We are looking for locations where there are few Chinese food and locations near city centre

We are going to use data science to observe neighborhoods that meet our criteria's will deliver the advantages of each locations so the stakeholders can make best possible decisions

Data

We are going to investigate the following points to make decisions:

- the number of restaurants in each neighborhood
- number of and distance to Chinese restaurants
- the distance from Chinese restaurants to the city centre

We find the latitude and longitude of each neighborhood in Toronto, transferred the latitude and longitude into X/Y co-ordinates to find out the distance of each neighborhood to Toronto city centre and visualized each neighborhood in folium map. We also leverage the Foursquare (Foursquare API) location data to explore the Chinese restaurants in each neighborhood

Methodology

In this project we are going to implement two selection process and ensemble them together to deliver a result

The first process is to find out how many Chinese restaurants in each neighborhood and the distance of each neighborhood to Toronto city centre (Yonge-Dundas Square). we find the neighborhood(s) that will few Chinese restaurants but with moderate distance

The second process is to find out the distance of each restaurant to Toronto centre(Yonge-Dundas Square), grouped by neighborhood ,find the mean value of the distance to Toronto centre within each neighborhood.

Finally find the intersection of the two results and narrow it down

Method 1

We first perform some basic data analysis techniques to explore the data, first we are going to observe the number of Chinese restaurants in each Neighborhood. And We can observe that the top 5 neighborhoods are in Toronto downtown area.

Out[22]:

	Neighborhood	id	distance to neighborhood/KM
0	Central Bay Street	50	0.820184
1	University of Toronto, Harbord	50	2.383689
2	Queen's Park, Ontario Provincial Government	50	1.552538
3	Richmond, Adelaide, King	49	0.471690
4	Kensington Market, Chinatown, Grange Park	49	1.873503
5	First Canadian Place, Underground city	49	0.831382
6	Toronto Dominion Centre, Design Exchange	49	1.045900
7	Garden District, Ryerson	47	0.827280
8	St. James Town	46	1.039831
9	Commerce Court, Victoria Hotel	46	0.971820
10	Willowdale, Willowdale East	45	18.867945
11	Agincourt	39	26.571952

And added the number of Chinese restaurants to the data frame:

In [22]:
```
grouped_rest_by_neigh[["Neighborhood","id","distance to neighborhood/KM"]]
```

6	Toronto Dominion Centre, Design Exchange	49	1.045900
7	Garden District, Ryerson	47	0.827280
8	St. James Town	46	1.039831
9	Commerce Court, Victoria Hotel	46	0.971820
10	Willowdale, Willowdale East	45	18.867945
11	Agincourt	39	26.571952
12	Church and Wellesley	38	1.983309
13	Milliken, Agincourt North, Steeles East, L'Amo...	33	28.309320
14	Stn A PO Boxes	31	1.545560
15	Harbourfront East, Union Station, Toronto Islands	26	2.043592
16	Berczy Park	21	1.863131
17	The Annex, North Midtown, Yorkville	15	3.981595
18	Steeles West, L'Amoreaux West	13	24.554913
19	Clarks Corners, Tam O'Shanter, Sullivan	13	22.477924

We have already finished the first method and demonstrated the data frame indicates the number of Chinese restaurants and the distance to city centre, we find that most neighborhoods with high number of Chinese restaurant are either located in Downtown Toronto or pretty far away from Toronto centre and that makes sense. These neighborhoods are out of considerations because it could be too competitive. but there are a few communities near Toronto Centre but with fewer Chinese restaurants. they are **Berczy Park, The Annex, North Midtown, Yorkville and East Toronto, Harbourfront East, Union Station, Toronto Islands, Broadview North (Old East York)**. these neighborhoods have around 20 Chinese restaurants means that Chinese food are still popular but not too competitive.

Method 2

We first transferred the latitude and longitude of Chinese restaurants into X/Y co-ordinates to calculate the distance of each restaurants to Toronto centre then aggregate them by neighborhood and find the mean distance of each neighborhood.

```
grouped_2_neighborhood
```

Out[43]:

	Neighborhood	location.lat	location.lng	Latitude	Longitude	distance to Toronto centre/KM
0	Garden District, Ryerson	43.655793	-79.382815	43.657162	-79.378937	0.891494
1	Commerce Court, Victoria Hotel	43.650770	-79.382495	43.648199	-79.379817	0.915078
2	St. James Town	43.653109	-79.380153	43.651494	-79.375418	0.931716
3	Richmond, Adelaide, King	43.653658	-79.387705	43.650571	-79.384568	0.968876
4	Toronto Dominion Centre, Design Exchange	43.651893	-79.386897	43.647177	-79.381576	1.036200
5	First Canadian Place, Underground city	43.651893	-79.386897	43.648429	-79.382280	1.036200
6	Stn A PO Boxes	43.649058	-79.379251	43.646435	-79.374846	1.184906
7	Berczy Park	43.647321	-79.379142	43.644771	-79.373306	1.212020
8	Central Bay Street	43.656292	-79.392754	43.657952	-79.387383	1.384786
9	Harbourfront East, Union Station, Toronto Islands	43.645527	-79.383315	43.640816	-79.381752	1.408815
10	Queen's Park, Ontario Provincial Government	43.658249	-79.391901	43.662302	-79.389494	1.505390
11	Kensington Market, Chinatown, Grange Park	43.654458	-79.396894	43.653206	-79.400049	1.594382

And then we added the number of restaurants of each neighborhood.

Out[46]:

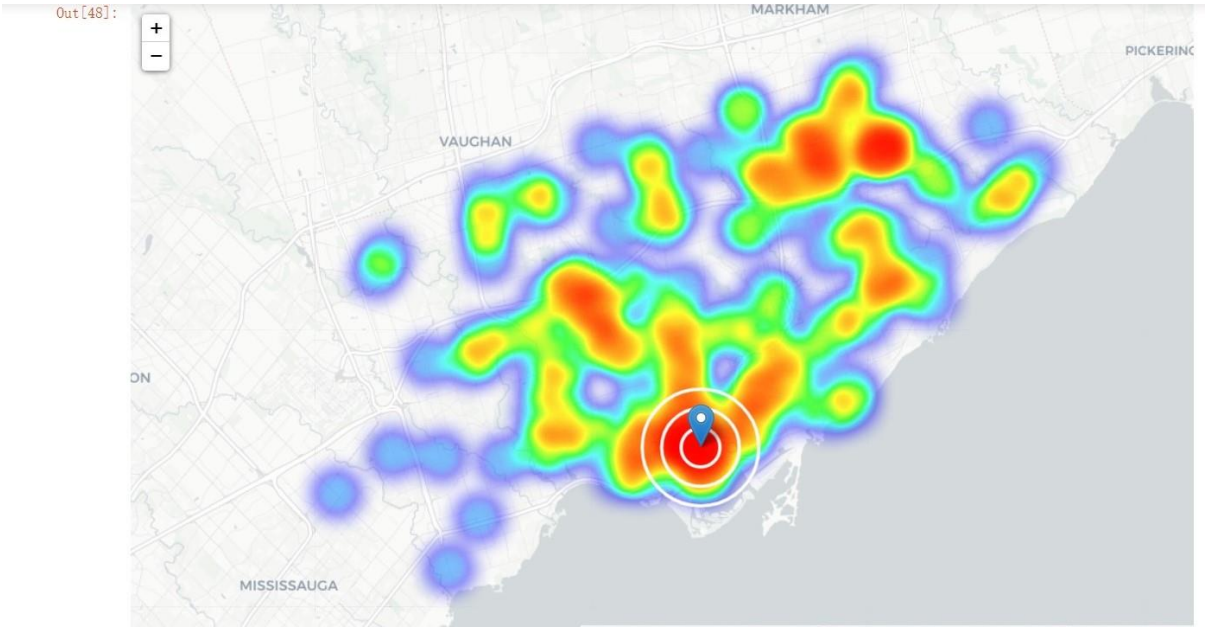
	Neighborhood	location.lat	location.lng	Latitude	Longitude	distance to Toronto centre/KM	num of Chinese restaurant
0	Garden District, Ryerson	43.655793	-79.382815	43.657162	-79.378937	0.891494	47
1	Commerce Court, Victoria Hotel	43.650770	-79.382495	43.648199	-79.379817	0.915078	46
2	St. James Town	43.653109	-79.380153	43.651494	-79.375418	0.931716	46
3	Richmond, Adelaide, King	43.653658	-79.387705	43.650571	-79.384568	0.968876	49
4	Toronto Dominion Centre, Design Exchange	43.651893	-79.386897	43.647177	-79.381576	1.036200	49
5	First Canadian Place, Underground city	43.651893	-79.386897	43.648429	-79.382280	1.036200	49
6	Stn A PO Boxes	43.649058	-79.379251	43.646435	-79.374846	1.184906	31
7	Berczy Park	43.647321	-79.379142	43.644771	-79.373306	1.212020	21
8	Central Bay Street	43.656292	-79.392754	43.657952	-79.387383	1.384786	50
9	Harbourfront East, Union Station, Toronto Islands	43.645527	-79.383315	43.640816	-79.381752	1.408815	26
10	Queen's Park, Ontario Provincial Government	43.658249	-79.391901	43.662302	-79.389494	1.505390	50
11	Kensington Market, Chinatown, Grange Park	43.654458	-79.396894	43.653206	-79.400049	1.594382	49

We can't show the entire view of the data frame because it is too long. Only the first ten rows are displayed for simplicity.

We have already completed the second selection process and demonstrated the above data frame. We can observe that most neighborhood near Toronto centre have many Chinese restaurant (around 50) we observe **Berczy Park, Harbourfront East, Union Station, Toronto Islands, Harbourfront, St.Regent Park, Harbourfront, James Town, Cabbagetown, The Annex, North Midtown, Yorkville** are good place to open Chinese restaurants

Then we added a heat map of the density of Chinese restaurants in Toronto

Out[48]:



The heat map below shows the density of Chinese restaurants in the Toronto. It is clear that area in red represents high density, yellow and green represents median density and blue represents low density. We can see there are lots of in **downtown Toronto** and **North York Centre Station** (the centre of North York)

We recall the results we generated from the first selection process: **Berczy Park, The Annex, North Midtown, Yorkville and East Toronto, Harbourfront East, Union Station, Toronto Islands, Broadview North (Old East York)** as well as the results we generated by the second process: **Berczy Park, Harbourfront East, Union Station, Toronto Islands, Harbourfront, St.,Regent Park, Harbourfront, James Town, Cabbagetown,The Annex, North Midtown, Yorkville**

We take the intersection of those two results: **Berczy Park,(The Annex, North Midtown, Yorkville),(Harbourfront East, Union Station, Toronto Islands).**

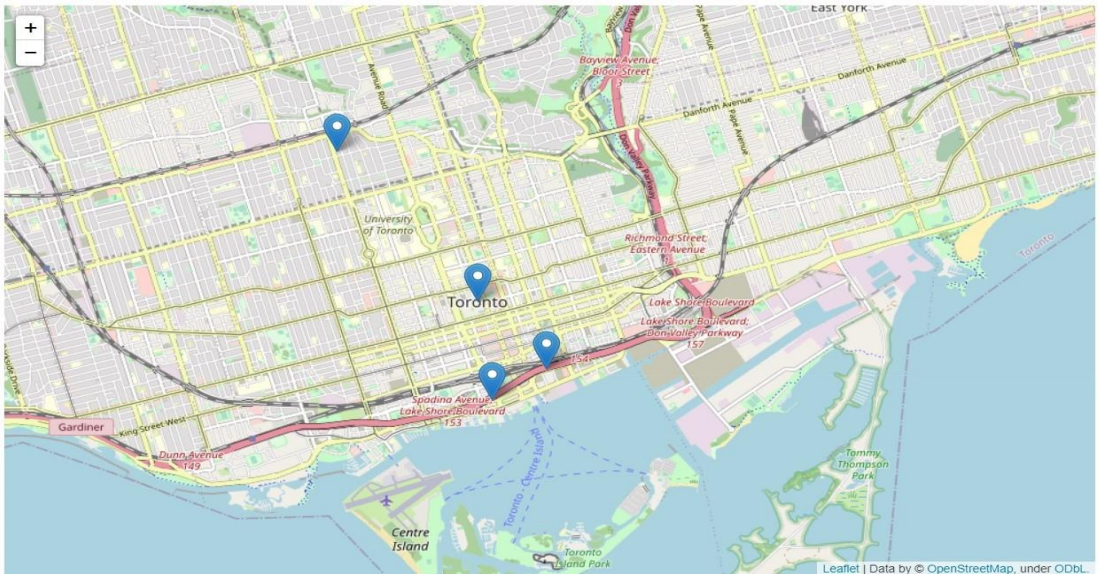
The result contains seven neighborhood which is still too many, we are going to deep compare these seven neighborhoods and narrow it down

```
In [58]: my_neighborhood=["Berczy Park","The Annex, North Midtown, Yorkville","Harbourfront East, Union Station, Toronto Islands"]
my_restaurant_df=pd.DataFrame(columns=grouped_2_neighborhood.columns)
for i in my_neighborhood:
    df=grouped_2_neighborhood[grouped_2_neighborhood["Neighborhood"]==i]
    my_restaurant_df=pd.concat([my_restaurant_df,df])
my_restaurant_df
```

Out[58]:

	Neighborhood	location.lat	location.lng	Latitude	Longitude	distance to Toronto centre/KM	num of Chinese restaurant
7	Berczy Park	43.647321	-79.379142	43.644771	-79.373306	1.212020	21
17	The Annex, North Midtown, Yorkville	43.666104	-79.401670	43.672710	-79.405678	2.969228	15
9	Harbourfront East, Union Station, Toronto Islands	43.645527	-79.383315	43.640816	-79.381752	1.408815	26

And showed the three neighborhoods on the map:



From the map as well as the data frame we observe that Harbourfront East, Union Station, Toronto Islands and Berczy Park neighborhoods have similar number of Chinese restaurants and similar distance to city centre but Berczy Park is much better because it is much near and the number is fewer than Harbourfront East, Union Station, Toronto Islands neighborhood. The Annex, North Midtown, Yorkville is much further (more than 2 times) but it is less competitive.

Results and Discussion

Our analysis shows there are lots of Chinese restaurants in the city of Toronto (around one thousand not included cities in York region and Peel region). However, most of them are located in downtown. To meet our criteria, we have to find a neighborhood in Toronto downtown or somewhere near downtown. After ensembled two selection method we have three neighborhoods: **Berczy Park, (The Annex, North Midtown, Yorkville),(Harbourfront East, Union Station, Toronto Islands)**. But we want to narrow down the result. We compared the three neighborhoods and labelled them on the map and we finally selected The Annex, North Midtown, Yorkville neighborhood because it is pretty near Toronto centre (3km) and market is not as competitive as the neighborhoods it near by.

The neighborhood we found is the most optimal based on our selection criteria, but it doesn't imply it is the best place. Selecting a community to start business does not simply consider these criteria. Rent, population densities, target customers are also factoring that should be considered. Take an example, Yorkville is one of the regions with highest rent and expense. But these factors are not included in the

project. We can only say The Annex, North Midtown, Yorkville neighborhood is the best choice based on our requirements initially set above.

Conclusion

The purpose of this project is to find the best neighborhood to start business based on the number of Chinese restaurants nearby and the distance to city centre. By calculating the distance with two methods and the number of Chinese restaurants in the neighborhood and narrow down analysis, we choose The Annex, North Midtown, Yorkville neighborhood. But we have to take considerations of other factors such as rent, population densities and target customers whether it nears bus stations or subway stations.