Battle of Neighborhoods

Description of problem:

The problem we are dealing with is that if we were to open up a store in an area one would have to decide upon a number of factors.
One of those factors is the local competition. If one was to open up a store for example a food store in an area where there are already 20+ stores present then the opening up of a new store might not have a great impact. There is another way to look at this which is that from the size of current market you can judge that how big a particular industry is in that market. For example if you have a different idea or approach for your idea you can reap larger profits if you enter a region where food interests a larger number of people.

Description of data:

I will be using the New York data provided by Capstone in an a=earlier assignment and the Toronto data obtained from foursquare in order reach upon a particular conclusion.

Methodology:

I will use the foursqaure API database in order to find the areas of interest in Manhattan
and Toronto and after obtaining those areas of interest I will be performing cluster analysis in order to segment the areas of interest. After the clustering we can study in detail about the interests and market size for individual cities and are their any markets which are absent in one and which could be exploited. Is any market already saturated such that exploiting that any further might not give out a reasonable profit.

Results:

We performed a clustering on the locations of Manhattan and Toronto and based on the clustering we performed some analysis on the results we got. To keep the analysis simple we have limited the number of cluster to 5 and at first have printed out the size of individual clusters. Which can be seen below:

```
The size of cluster 0 of Toronto is 27
The size of cluster 0 of Manhattan is 19
The size of cluster 1 of Toronto is 8
The size of cluster 1 of Manhattan is 9
The size of cluster 2 of Toronto is 2
The size of cluster 2 of Manhattan is 1
The size of cluster 3 of Toronto is 1
The size of cluster 3 of Manhattan is 10
The size of cluster 4 of Toronto is 1
The size of cluster 4 of Manhattan is 1
```

We can see that the cluster 0 of both Manhattan and Toronto are the dominant ones. So we further explore these dominant clusters.

```
[92] clstr0_tr = tr_clstr.loc[tr_clstr["Cluster Labels"]==0].reset_index(drop=True)
     clstr0_tr.head(27)
```

	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
0	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636	0	Coffee Shop	Bakery	Park	Pub	Restaurant
1	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494	0	Coffee Shop	Sushi Restaurant	Yoga Studio	Bank	Beer Bar
2	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937	0	Clothing Store	Coffee Shop	Bubble Tea Shop	Café	Japanese Restaurant
3	Downtown Toronto	St. James Town	43.651494	-79.375418	0	Café	Coffee Shop	Cocktail Bar	American Restaurant	Gastropub
4	Downtown Toronto	Berczy Park	43.644771	-79.373306	0	Coffee Shop	Cocktail Bar	Bakery	Café	Cheese Shop
5	Downtown Toronto	Central Bay Street	43.657952	-79.387383	0	Coffee Shop	Café	Italian Restaurant	Japanese Restaurant	Sandwich Place

(Cluster 0 of Toronto)

```
clstr0_ny = ny_clstr.loc[ny_clstr["Cluster Labels"]==0].reset_index(drop=True)
clstr0_ny.head(19)
```

	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
0	Manhattan	Chinatown	40.715618	-73.994279	0	Chinese Restaurant	Bakery	Cocktail Bar	Bubble Tea Shop	Coffee Shop
1	Manhattan	Central Harlem	40.815976	-73.943211	0	African Restaurant	Seafood Restaurant	Gym / Fitness Center	American Restaurant	Bar
2	Manhattan	Upper East Side	40.775639	-73.960508	0	Italian Restaurant	Coffee Shop	Gym / Fitness Center	Bakery	Yoga Studio
3	Manhattan	Yorkville	40.775930	-73.947118	0	Coffee Shop	Gym	Italian Restaurant	Bar	Sushi Restaurant
4	Manhattan	Lenox Hill	40.768113	-73.958860	0	Coffee Shop	Pizza Place	Italian Restaurant	Sushi Restaurant	Café
5	Manhattan	Upper West Side	40.787658	-73.977059	0	Bar	Italian Restaurant	Dessert Shop	Wine Bar	Indian Restaurant
6	Manhattan	Lincoln Square	40.773529	-73.985338	0	Plaza	Italian Restaurant	Café	Gym / Fitness Center	Theater

(Cluster 0 of Manhattan)

We notice that both the clusters are actually pretty similar in the fact that at both places the most common venue is generally a food outlet but the general difference we can see is that the most visited places in Toronto for cluster 0 has majority food outlets and in that to the most common one is either a Cafe or a Coffee shop. But in case of Manhattan it also has Gyms, Spas and Yoga studious within the cluster and Coffee Shops although still present are not as dominant as in case of Toronto.

Therefore, we can assume that cluster 0 of Manhattan might be near a residential area whereas the cluster 0 of Toronto might be near the workplace.

After exploring the clusters we come across the observation that cluster 0 in Toronto and cluster 3 in Manhattan generally contain areas related to food outlets. So, if one was to decide on where the competition is already high or where one could find a more choice we could use this information.

So we compare the size of the clusters and come to the conclusion that Toronto has a much more tighter competition if one were to open up a store there as compared to Manhattan which only has around 10 areas where food is the main item of attraction.

Conclusion:

In the end I would like to conclude that while the analysis of both the neighborhoods is not that accurate we can get a general idea of the different cultures of both the cities for example Manhattan has 331 categories of venues whereas Toronto has only 237. Further more we have limited the total venues to a 100 and have defined a radius of 500, so we can conclude that the market place of Manhattan is much more diverse than that of Toronto.
We saw that the clusters of Manhattan had a little bit more uniform distribution when compared with Toronto as Toronto was top heavy.