The Battle Of The Neighborhoods – Get Baked

1. Introduction/Business Problem:

A client is looking to open an Bakery in Toronto called "Get Baked". They are looking to open it in a location where there are no other Bakeries around, if possible. They are looking to you for your expertise as you are a Data Scientist and are hoping that your report for them makes sense and meets the criteria that they have set.

2. Data

In order to acheive what the client is wanting, I will be connecting Postal Codes data with Geographical Coordinates data and then using Foursquare to help me in planning out where to place the Bakery. I will also be using libraries imported in such as Folium in order to plot the data out on a map. These libraries will have comments on them to describe what they are for when imported. After getting all of this data using Python, I will be cleaning it up, connecting it together, and creating graphical charts and KMeans Clustering to make sure the data is all set up to find out how to answer the Business Problem at hand.

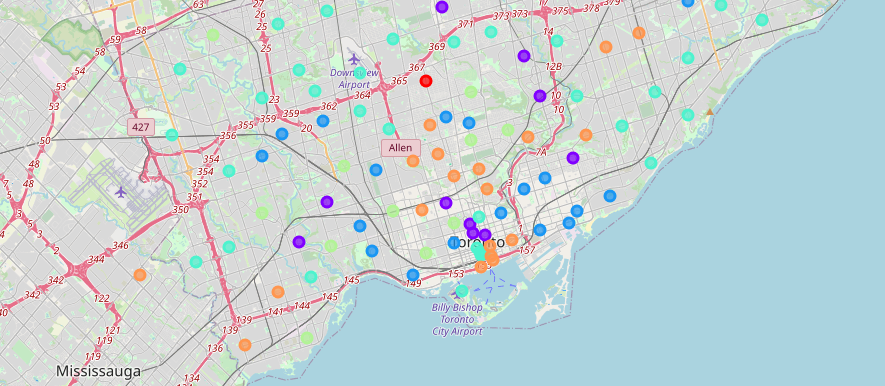
Toronto Postal Code Data: https://en.wikipedia.org/wiki/List\_of\_postal\_codes\_of\_Canada:\_M

Toronto Geographical Coordinates Data: https://cocl.us/Geospatial\_data

Foursquare

3. Methodology

In order to manipulate the raw data, I utilized the skills learned throughout all 9 of the courses for the IBM Data Science Professional Certificate to pull the raw data from the websites and load them into a data frame using the Pandas repository. After doing so I utilized Folium to create multiple graphs to show the data visually to help me get a better understanding of what I was working with in the city of Toronto. I then created algorithms to help with getting the Venues from the data and merged all of the tables together to use KMeans Clustering in finding which cluster would be the best place to put a Bakery.



4. Results

Comparing the results and different clusters, it appears that one of the 6 different clusters stands out among the rest as far as Bakery Frequency. The other clusters vary in size and all seem to have at least one Bakery around.



5. Discussion

When comparing the clusters, is it really a good thing to place a Bakery where there are none around? Is there a reason for this? Should we maybe change our criteria and find out what is causing there to be no bakeries around? Will we make money here with a new Bakery or will it fail?

6. Conclusion

Out of the 6 different clusters, only one of the clusters (Cluster 3) has results where there are no Bakeries at all. Judging by this it appears that anywhere in this cluster it would be the best area to put a Bakery as our client is wanting to place the bakery in an area where no other bakeries are located if possible.