The Battle of Neighborhoods | Business Proposal

Introduction

Business Problem: The problem that I am approaching with this project is trying to determine the best neighborhood in the region of North York, Ontario to introduce a new bakery business. This project aims to analyze various neighborhoods and compare them to determine the best location for the prospective business owner. This includes factors such as similar businesses nearby (cafe's bakeries, grocery stores), as well as information about the neighborhood as a whole. This analysis will help the future business owner make the best decision for where to set up their bakery in order to be the most successful.

Interest: The parties interested in this data will be the hypothetical business owner who has hired me for this analysis, as well as future clients who may be looking for similar information. My report will allow the business owners to make the best decision for their future, as well as use definitive data to apply for loans if needed.

Data:

Data Sources:

Neighborhood listings for the North York area can be found on Wikipedia.

Foursquare: Foursquare API is a valuable source that provides information about millions of places, including location, pictures, tips, and information about the businesses. I will be able to use the foursquare data to see what neighborhoods have similar businesses (such as cafes and other bakeries).

Data Cleaning:

I will webscrape neighborhood data from Wikipedia and filter the results to include only postal codes in the North York area. I can then use the foursquare data to cluster neighborhoods using K-means clustering and see what neighborhoods have the fewest bakeries and cafes, as well as seeing what else in in the local area. This should allow me to narrow down the proper place to propose the new business. This data should help me accurately suggest a neighborhood for the bakery business. Once I analyze the foursquare data, I can use it to differentiate multiple neighborhoods that may fit both of the prerequisites stated above. I can further sort by other popular venues in the area that may benefit from a bakery nearby.

Libraries that will be used for this project are; Pandas, Beautiful Soup, Requests, Folium, JSON, XML, Matplotlib, Geocoder, and Scikit Learn.

Methodology:

Exploring the data:

I started by cleaning up the data that was scraped from Wikipedia. The data provided gave me much too broad of an area, so I filtered the data to include only those neighborhoods in North York for my analysis.

After filtering the data, I was able to start my exploration into the neighborhoods in the North York Borough. I used JSON and the foursquare API to gather information about the venues in the area, and sorted them by the most popular venues so that I could rule out any neighborhoods with an abundance of similar businesses.

```
----Don Mills North----
            venue   freq
0          Bakery    0.2
1        Bus Line    0.2
2   Metro Station    0.1
3     Bus Station    0.1
4            Park    0.1


----Glencairn----
                  venue   freq
0      Indian Restaurant    0.4
1  Vietnamese Restaurant    0.2
2              Pet Store    0.2
3     Chinese Restaurant    0.2
4            Yoga Studio    0.0


----Parkwoods----
                     venue   freq
0       Fast Food Restaurant    1.0
1                Yoga Studio    0.0
2                Music Venue    0.0
3   Middle Eastern Restaurant    0.0
4          Miscellaneous Shop    0.0


----Victoria Village----
                        venue   freq
0   Construction & Landscaping   0.67
1                          Bar   0.33
2        New American Restaurant   0.00
3    Middle Eastern Restaurant   0.00
4           Miscellaneous Shop   0.00
```

As you can see in Figure 1, the "Don Mills North" neighborhood's most popular venue type is "bakery". This means that it would be unwise to pursue another bakery business in this area as it would be drowned out quickly by well established businesses.

A neighborhood like Victoria Village for example doesn't seem to have many food options and may provide a better location for a new bakery business.

(Figure 1)

I wanted a little bit more information about the remaining neighborhoods, so I gathered the top 10 most common venues from each one in a dataframe.

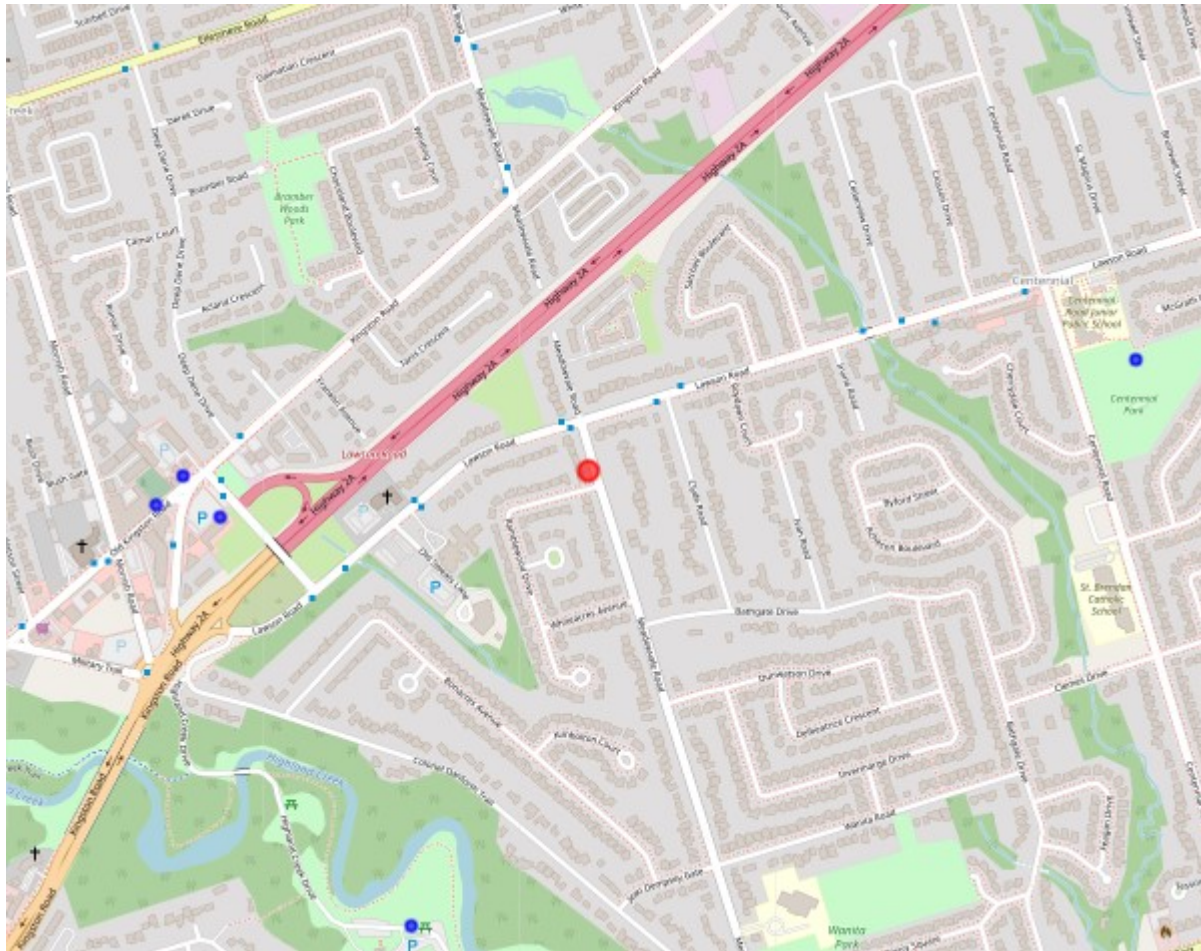|	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
3	Don Mills North	Bakery	Bus Line	Metro Station	Bus Station	Park	Soccer Field	Ice Cream Shop	Intersection	Pharmacy	Moroccan Restaurant
10	Glencairn	Indian Restaurant	Vietnamese Restaurant	Pet Store	Chinese Restaurant	Yoga Studio	Music Venue	Middle Eastern Restaurant	Miscellaneous Shop	Mobile Phone Shop	Modern European Restaurant
17	Parkwoods	Fast Food Restaurant	Yoga Studio	Music Venue	Middle Eastern Restaurant	Miscellaneous Shop	Mobile Phone Shop	Modern European Restaurant	Monument / Landmark	Moroccan Restaurant	Museum
18	Victoria Village	Construction & Landscaping	Bar	New American Restaurant	Middle Eastern Restaurant	Miscellaneous Shop	Mobile Phone Shop	Modern European Restaurant	Monument / Landmark	Moroccan Restaurant	Museum

(Figure 2)

In Figure 2, you can see that the Parkwoods and Glencairn neighborhoods also contains many food options such as Fast Food Restaurants, Middle Eastern Restaurants, Modern European restaurants, Vietnamese Restaurants, Chinese Restaurants, Indian Restaurants and Moroccan Restaurants. This is likely not the best area to be starting a new food business. At this point, it seems safe to assume that Victoria Village Neighborhood has the least competition for an incoming bakery business.

I decided to dive a little bit deeper into the Victoria Village neighborhood to see what competition arose. I used the foursquare API to make a call for the venues nearest to the Victoria Village Neighborhood. Figure 3 shows the results.

	name	categories	lat	lng
0	Fratelli Village Pizzeria	Italian Restaurant	43.784008	-79.169787
1	Shamrock Burgers	Burger Joint	43.783823	-79.168406
2	Ted's Restaurant	Breakfast Spot	43.784468	-79.169200
3	Colonel Danforth Park	Playground	43.777507	-79.164303
4	Centennial Park	Park	43.786257	-79.148776

(Figure 3)

I also created a map of the area to show where these venues were located as seen in Figure 4.

(Figure 4)

It looks like venues in this area are quite spread out, and do not include anything related to a bakery that would pose competition to the new incoming business. Victoria Village seems like an excellent neighborhood choice.

Results and Discussion:

Based on the data above, Victoria Village is the best neighborhood in the North York area to open a new bakery business. This considers the neighborhood, proximity to other venues, and overall type of venue in the area. Using the foursquare API and folium, we could see what the Victoria Village neighborhood had an advantage. It is the only neighborhood listed that had little to no food options, compared to the other neighborhoods, one of which had "bakery" listed as their most common venue.

Conclusion:

In this study, I discovered the best neighborhood for a new bakery business in the North York area. This should help the future business owner decide where to set up their bakery, as well as give them data to use when applying for loans or other planning processes. The Victoria park area was the best neighborhood due to the fact that it had no other bakeries, and was relatively spaced out when it came to venues in general which allows for plenty of neighborhood coverage for their future clients.