Gary Swider
Coursera Capstone Project

Introduction/Business Problems

A description of the problem and a discussion of the background.

People moving to different locations are constantly searching for information about their surrounding area. When buying a new home, people always look for what is in proximity to their neighborhood, how much homes in that area cost, nearby schools, crime rates, restaurants, and fun things to do in the area. This project is designed to inform people of the 'important information' when determining whether they will be happy with a move to a new location. In this project, I will continue to use the Scarborough data from the previous assignments to analyze and explore the data to make a well-rounded opinion on a particular neighborhood in the Scarborough area, based on key location information from FourSquare.

We can consider a 'good neighborhood' as one that has a considerable amount of commercial presence, including dining, entertainment, as well as house pricing, schools, and general safety.

Data Description and how it will be used

Scarborough is a popular destination in Canada. It is said to be one of the most diverse and multicultural areas Toronto. This diverse culture is comparable to New York, due to the amounts of immigration. This trend seems to be continuing, especially since immigration in both the United States and Canada have been increasing.

Where am I getting the data?

This project will use FourSquare API as its prime data gathering source. FourSquare has a database of millions of places acquired through crowdsourcing. As a free user of FourSquare's API, certain features are limited and therefore, certain calls to the API are limited throughout the project.

Mainly, to compare the similarities of two cities, I am going to explore the neighborhoods, segment them using KMeans, and then group them into clusters to find similar neighborhoods in New York and Toronto.

This will be done in a Jupyter Notebook, using Python, which requires several libraries such as Pandas, Folium, SciKit Learn, Geopy, and Requests. Pandas will be used to create and manipulate dataframes. Folium will be used to create maps to visualize the neighborhoods. Scikit Learn is the machine learning tool for importing k-means clustering. Geopy will be used to retrieve the location data. Finally, Requests will be used to handle http requests.

Methodology

Foursquare API:

For this project, I used FourSquare API as the main data-gathering source. FourSquare crowdsourced data on a large amount of places and organized them into databases. The data is very easy to access as the API allowed me to Search/Explore and Request data on the business or location. I had to create a developer account with assigned me with a unique Client ID and Private Key. Using these credentials, I was able to mine the particular pieces of data that were applicable to my queries.

To complete this task, I needed to:
1. Acquire and Clean the Data such as postcodes and venue names
2. Organize the data into a dataframe
3. Create the map using the coordinates/Data that was mined
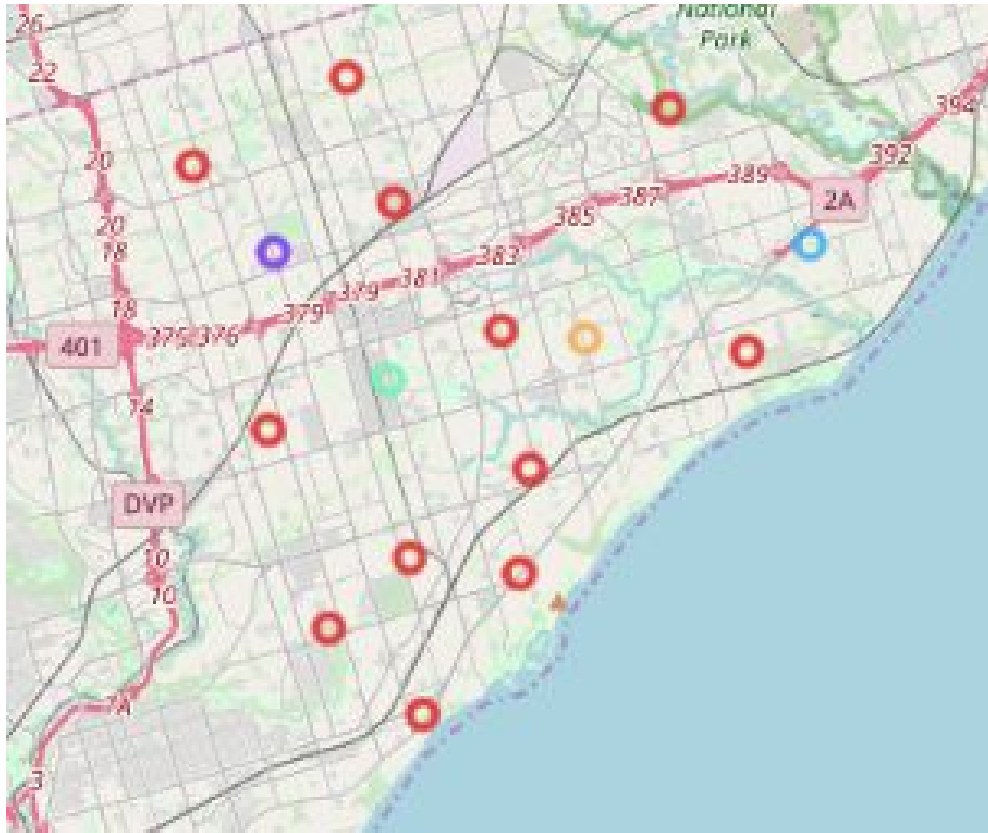4. Create Clusters using machine learning.

Results

Neighborhoods' Top 10 Most Common Venues

	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
0	Agincourt	Shopping Mall	Supermarket	Park	Pool	Chinese Restaurant	Shanghai Restaurant	Skating Rink	Grocery Store	Sushi Restaurant	Bakery
1	Agincourt North, L'Amoreaux East, Milliken, St...	Pharmacy	Women's Store	Hobby Shop	Gym / Fitness Center	Gym	Grocery Store	Golf Course	Gift Shop	General Entertainment	Furniture / Home Store
2	Birch Cliff, Cliffside West	Gym Pool	Gym	General Entertainment	Skating Rink	College Stadium	Park	Electronics Store	Gym / Fitness Center	Grocery Store	Golf Course
3	Cedarbrae	Trail	Playground	Women's Store	Discount Store	Gym / Fitness Center	Gym	Grocery Store	Golf Course	Gift Shop	General Entertainment
4	Clairlea, Golden Mile, Oakridge	Bakery	Bus Line	Intersection	Coffee Shop	Soccer Field	Metro Station	Women's Store	Fast Food Restaurant	Gym	Grocery Store
5	Clarks Corners, Sullivan, Tam O'Shanter	Pizza Place	Pharmacy	Bus Stop	Fried Chicken Joint	Shopping Mall	Coffee Shop	Hobby Shop	Thai Restaurant	Golf Course	Chinese Restaurant
6	Cliffcrest, Cliffside, Scarborough Village West	Fast Food Restaurant	Discount Store	Liquor Store	Coffee Shop	Furniture / Home Store	Sandwich Place	Pharmacy	Pizza Place	Electronics Store	Gym
7	Dorset Park, Scarborough Town Centre, Wexford ...	Bakery	Gift Shop	Wine Shop	Women's Store	Electronics Store	Gym Pool	Gym / Fitness Center	Gym	Grocery Store	Golf Course
8	East Birchmount Park, Ionview, Kennedy Park	Discount Store	Coffee Shop	Department Store	Electronics Store	Gym Pool	Gym / Fitness Center	Gym	Grocery Store	Golf Course	Gift Shop
9	Guildwood, Morningside, West Hill	Gym / Fitness Center	Park	Construction & Landscaping	Women's Store	Electronics Store	Gym	Grocery Store	Golf Course	Gift Shop	General Entertainment
10	Highland Creek, Rouge Hill, Port Union	Bar	Women's Store	Electronics Store	Gym Pool	Gym / Fitness Center	Gym	Grocery Store	Golf Course	Gift Shop	General Entertainment
11	L'Amoreaux West	Grocery Store	Chinese Restaurant	Fast Food Restaurant	Pizza Place	Japanese Restaurant	Coffee Shop	Other Great Outdoors	Sandwich Place	Pharmacy	Electronics Store
12	Maryvale, Wexford	Auto Garage	Intersection	Convenience Store	Electronics Store	Gym / Fitness Center	Gym	Grocery Store	Golf Course	Gift Shop	General Entertainment
13	Rouge, Malvern	Home Service	Wine Shop	Gym Pool	Gym / Fitness Center	Gym	Grocery Store	Golf Course	Gift Shop	General Entertainment	Furniture / Home Store
14	Scarborough Village	Women's Store	Train Station	Indian Restaurant	Restaurant	Bar	Electronics Store	Bakery	Gym	Grocery Store	Golf Course
15	Woburn	Korean Restaurant	Business Service	Coffee Shop	Park	Women's Store	Fast Food Restaurant	Gym / Fitness Center	Gym	Grocery Store	Golf Course

Results (cont.)

Clusters of Venues for Scarborough, Canada

Discussion

Based on the information above, there are many diverse neighborhoods in Scarborough, Canada. If one prioritized Shopping, one could be interested in Agincourt as the Shopping Mall is Agincourt's most popular venue. If someone prioritized exercising, Ceaderbrae's most popular venue is its running trail and playground. As a man in my mid-twenties, I prioritize fast-food and bars, so if I were looking, I would start in Cliffcrest as it has a balance of restaurants, liquor stores, coffee shops, and a gym.

Conclusion

Through KMeans clustering,I separated Scarborough into 5 clusters. These 5 clusters show similar neighborhoods around them. By using the table above highlighting the top venues per neighborhood, one could make a decision on where to move to if venue proximity was the most important factor.