Capstone Project- The Battle of Neighborhoods (Week 1)

Background

The Coffee Shops market worldwide is projected to grow by US$58.7 Billion. The shifting dynamics supporting this growth makes it critical for businesses in this space to keep abreast of the changing pulse of the market. Poised to reach over US$143.4 Billion by the year 2025, Coffee will bring in healthy gains adding significant momentum to global growth.

More people in urban settings both residential and commercial, expect coffee shops within walking distance. This has given rise to local community coffee shops along with big brand names like Starbucks. The age of coffee drinkers has also got younger with the youngest being 14 years. The combination of millennial influence, a casual social setting, free wi-fi and a warm atmosphere is constantly driving growth of approx. 7% across the coffee shop culture.

A coffee shop stop has now become ritualistic, both on the way to and back from work for most working Americans. Additionally, as we suffer from a sleep disorder epidemic, more Americans are drinking more coffee every day.

Problem

The city of Toronto supports one of the large urban settings in the world. With the likes of Financial District - Toronto Stock Exchange (3rd largest in North America), Canadian Film & TV Industry, The Digital Corridor and more, the need for coffee is boundless around the city. We will look to explore which neighborhoods in Toronto are closest to coffee shops to residents. This can influence the decision to stay/rent/buy housing in those neighborhoods for coffee lovers.

1. Find neighborhoods cluster(s) with the most coffee shops near them
2. Identify a few desirable neighborhoods to stay that have access to most coffee shops
3. Find the neighborhoods cluster with the lowest coffee shops near them
4. This presents a business opportunity for entrepreneurs to invest in or open new coffee shops near those neighborhoods.

Data Description & Usage

1. We will source Toronto neighborhood data from Wikipedia
2. Cleanup empty and NaN cells
3. Get geolocation data from cocl.us (used in previous labs)
4. Merge the two data to obtain latitude and longitude for the neighborhoods
5. Plot the data on a map
6. Using FourSquare API, get the venues data for the neighborhood latitude/longitudes
7. Filter venues to use data for Coffee Shops
8. Run KMeans clustering
9. Plot the cluster data on a map
10. Analyze and deduce best clusters/neighborhoods for coffee lovers
11. Analyze and deduce neighborhoods where there are low or now coffee shops