

IBM Capstone – Toronto

Project's Overview

- This project aims to monitor and analyze data related to Toronto city's characteristics.
- Data describing Toronto's wider area were extracted from Wikipedia and latitudes –longitudes from Foursquare through API calls.

What is the Problem?

- A major problem about establishing a new shop, restaurant or office, is the location.
- The location can affect dramatically the productivity and development of a business.
- Which areas are most appropriate for establishing a new restaurant?

Data Acquisition and Cleaning

- The data required for this analysis were extracted from the following link:
https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M.
- Many libraries were used in this project. (Pandas, BeautifulSoup, geocode, folium etc.)
- After the page was downloaded as an html file, the first table (containing the related data) was accessed (through an iteration) and converted to a pandas data frame.

Methodology

- After the main data manipulation was completed, I had to use Foursquare API credentials in order to monitor the statistics and venues for my area of study.
- The data frame was processed again, in a way that reveals the venues in each neighborhood and specifically the most popular venue.
- The frequency of each venue is maybe one of the major factors which have to be related with the location choice. Also, the population is a crucial factor and the neighborhood's location in comparison with other neighborhoods.

Results

- I took five (5) neighborhoods which in my opinion represent a good establishing location for a new restaurant in the city of Toronto.
- The top 5 trending venues frequency were taken into consideration for the following results.
- The results for these neighborhoods are the following:

Alderwood, Long Branch

- In those neighborhoods, the trending venues are closely related to entertainment. There is a decent frequency in restaurants (0.14 Pizza place). That fact may indicate that the specific area is a place where people visit for entertaining activities.

Bayview Village

- In this area there are no restaurants in trending venues but also the park and dog run indicates an area with enough free space for sports and hobbies. Also, the high frequency of gas stations may be an indicator about high traffic which means that a restaurant with takeaway and drive-through services, could be a considerable option.

Birch Cliff, Cliffside West

- This is a neighborhood similar to the first group. The frequencies indicate the presence of teenagers and young people. This could be an appropriate place for a fast-food restaurant.

Caderbrae

- In this example, a restaurant could work as a supplementary service to gaming cafes. Many young people who spend hours in gaming cafes could be also potential customers in a lowbudget restaurant.

Cental Bay Street

- In this last example, we have a central street which obviously has many health regulated establishments. Although, the frequency in each one except coffee shops, is low. That means that the main activities in this area are related to entertainment and food & beverages. So, having a unique idea of a restaurant may fit in this area.

Conclusion

- Having the appropriate data set could be a major advantage in setting a new business.
- The location of a business, especially in customer service sector, directly affects its productivity and development.
- Using all those methods and tools could be a game-changing factor for any business, company or entrepreneur.

Thank you for your attention.