Applied Data Science Capstone

Analyzing Neighborhoods of Toronto to Starting A New Restaurant

Capstone Project (Week 4)

1. Introduction: Business Problem

Toronto is the capital city of the Canadian province of Ontario. With a recorded population of 2,731,571 in 2016, it is the most populous city in Canada and the fourth most populous city in North America. Toronto's growing population coincides with increased development and investment in the city and surrounding region. This is why the Toronto area an excellent place to set up a new restaurant.

Opening a succesful restaurant is not an easy task. One of the things that need to be done is to determine the right location to open a new restaurant. In this project, we will study the neighborhoods and find the most suitable location to starting a new restaurant. Target audience of this project is someone who want to open new restaurant in Toronto and have no idea where to build it.

2. Data

In this project, we need the following data :

- Neighborhoods of Toronto
  - Source
    https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
  - Description
    This data was scrapped from Wikipedia page using BeautifulSoup (python library) and will be store in a dataframe. The dataframe will consist of three columns: PostalCode, Borough, and Neighborhood.
  - Example

	PostalCode	Borough	Neighborhood
0	M3A	North York	Parkwoods
1	M4A	North York	Victoria Village
2	M5A	Downtown Toronto	Regent Park, Harbourfront
3	M6A	North York	Lawrence Manor, Lawrence Heights
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government
5	M9A	Etobicoke	Islington Avenue, Humber Valley Village

- Geographical coordinates of the neighborhoods (latitude and longitude)
    - Source
      http://cocl.us/Geospatial_data
    - Description
      In order to utilize the Foursquare location data, we need to get the latitude and the longitude coordinates of each neighborhood. . The dataframe will consist of three columns: PostalCode, Latitude, and Longitude.
    - Example

	PostalCode	Latitude	Longitude
0	M1B	43.806686	-79.194353
1	M1C	43.784535	-79.160497
2	M1E	43.763573	-79.188711
3	M1G	43.770992	-79.216917
4	M1H	43.773136	-79.239476

- Venue data related to restaurants in Toronto
    - Source : Foursquare
    - Description
      This data will consist of four columns: venue name, categories, latitude, and longitude.
    - Example

	name	categories	lat	lng
0	Roselle Desserts	Bakery	43.653447	-79.362017
1	Tandem Coffee	Coffee Shop	43.653559	-79.361809
2	Cooper Koo Family YMCA	Distribution Center	43.653249	-79.358008
3	Morning Glory Cafe	Breakfast Spot	43.653947	-79.361149
4	Body Blitz Spa East	Spa	43.654735	-79.359874