Capstone Project – The Battle of Neighborhoods

Introduction/Business Problem

The city of Toronto is the financial capital of Canada. It has diverse cultures and enormous potential for new businesses. As a global hub of commerce, a lot of people come to the city to look for new opportunities in the market every year. However, great opportunity also means cruel competition. To reduce the risk of failure and explore the benefit as much as we can, careful analysis needs to be done before investing in the real market.

Meanwhile, bubble tea, a popular Taiwanese tea-based drink has become famous throughout the world. Its perfect combination of tea, milk and sugar, together with its different toppings satisfies people's cravings from different continents. A growing number of bubble tea shops have popped up in Toronto in the last couple of years. More and more people enjoy this asian taste during their weekday breaks or weekend free time. We are interested in opening a new bubble tea shop in the city of Toronto and would like to enter the market with a smooth start. Therefore, it is important for us to know about the potential competitors around different boroughs and find the best location for our new shop.

In this project, we will investigate the distribution of bubble tea shops in Toronto and pick the best location to open a new bubble tea shop.

Data

The first dataset we are using is based on the Wikipedia page: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M, which contains postal codes of different boroughs and neighborhoods of Toronto. We are going to extract the table and save the information into a pandas dataframe. The table in the webpage looks like below,

Postcode	Borough	Neighbourhood
M1A	Not assigned	Not assigned
M2A	Not assigned	Not assigned
M3A	North York	Parkwoods
M4A	North York	Victoria Village
M5A	Downtown Toronto	Harbourfront
M5A	Downtown Toronto	Regent Park
M6A	North York	Lawrence Heights
M6A	North York	Lawrence Manor
M7A	Queen's Park	Not assigned
M8A	Not assigned	Not assigned
M9A	Etobicoke	Islington Avenue

The second dataset has the geographical coordinates of each postal code and we are using the dataset provided in the capstone project week 3: http://cocl.us/Geospatial_data. Just like what we did in project week 3, we combined the geographical data with the postal codes, boroughs and neighborhoods table. The snippet of the final table is shown below,

	Postcode	Borough	Neighbourhood	Latitude	Longitude
0	M3A	North York	Parkwoods	43.753259	-79.329656
1	M4A	North York	Victoria Village	43.725882	-79.315572
2	M5A	Downtown Toronto	Harbourfront,Regent Park	43.654260	-79.360636
3	M6A	North York	Lawrence Heights,Lawrence Manor	43.718518	-79.464763
4	M7A	Queen's Park	Queen's Park	43.662301	-79.389494

In addition to the above two datasets, we are utilizing the FourSquare api to retrieve the bubble tea shops information around each interested neighborhood. Based on all these information, we will pick the best location for our new bubble tea shop.