

IBM APPLIED DATA SCIENCE CAPSTONE

Opening Bubble Tea Shop in Toronto: Report

By Asmar Aliyeva

October 2019

Introduction

Since Green Grotto opened at Yonge and Sheppard in 1993, the Toronto's thirst for bubble tea—that sweet, tea-based drink studded with tapioca balls, or *boba*—has really popped. Today, bubble tea shops are proliferating in the downtown core. Bubble tea in Toronto is so popular, you'll find international chains and local faves expanding to all corners of the city at a ferocious pace.

One of the most significant factors when thinking about opening a bubble tea franchise is the location of the shop. It is crucial to make sure that the café will be in an area with high traffic that will come in and peruse your store for food and drink. To open a successful bubble tea business, area that has less competition will be the best choice.

Business Problem

This capstone project will analyse neighbourhoods of Toronto city to choose the best location for opening a new Bubble Tea Shop. Data Science and Machine Learning techniques such as K-Means Clustering will be used to analyse this business problem and to find the neighbourhood with lowest competition and less concentration of Bubble Tea Shops. So, the main question of this research will be to find the location where is it most beneficial to open Bubble Tea Shop.

Target Audience

The result of this capstone project will be particularly useful for investors looking to invest in cafes and restaurants, particularly, in Bubble Tea Shops in Toronto, ON. Bubble Tea Shops are getting more popular and new Shops are opening every year. Yet, there are still available and undiscovered area where investors can take advantage of.

Data:

Required Data

1. The data regarding neighbourhoods of Toronto.
2. Latitude and longitude of neighbourhoods
3. Venue data for each neighbourhood

Data collection

1. Data regarding neighbourhoods of Toronto will be collected from Wikipedia Page (https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M), which contains the list of all neighbourhoods with their corresponding postal codes and boroughs. I will use web scrapping techniques, Python Request and BeautifulSoup, to extract the data.
2. To collect corresponding latitude and longitude coordinates of neighbourhoods I will use Python Geocoder package.
3. To collect venue data of each neighbourhood I will use Foursquare API as it has the largest database of locations with over 105 million places. Along with venue name, it provides information about the category of the venue and its latitude and longitude. Among venue categories I am interested in Bubble Tea Shops.