Discovery gastronomy diversity in Toronto's neighbourhood

I. Introduction: Business Problem

Toronto is the most populous city in Canada. Toronto also is an international centre of business, finance, arts and culture. It is recognised as one of the most multicultural and cosmopolitan cities in the world. Toronto's economy is highly diversified with strengths in technology, design, financial services, life sciences, education, fashion, food services and tourism. The diverse population of Toronto reflects its current and historical role as an important destination for immigrants to Canada. Torontonians speak English as their primary language, but over 160 languages are spoken in the city.

The idea of this project is to categorically segment the neighbourhoods of Toronto into major clusters of food habits and taste. By exploring venue data from Foursquare's 'Places API' and using unsupervised machine learning algorithm 'k-means clustering', it could help us to discover the diversity of a neighbourhood and its relationship with cuisines in the city.

This analysis can be used to understand the distribution of different cultures and cuisines over the city. It can also be utilized by a new food vendor who is willing to open his/her restaurant, or by a government who would like to examine and study their city's culture diversity.

II. Data

The following data sources will be used in the project:

1. Canada dataset:

We firstly use the following wikipedia website to obtain the data that is in the table of postal codes in Canada:
https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
Then, using the following link to a .csv file that has the geographical coordinates of each postal code in Canada: http://cocl.us/Geospatial_data
From which, we could extract the addresses of neighbourhoods of Toronto.

2. Foursquare API

Link: https://developer.foursquare.com/docs
Foursquare API, a location data provider, will be used to make RESTful API calls to retrieve data about venues in different neighbourhoods. This is the link to Foursquare Venue Category Hierarchy. Venues retrieved from all the neighbourhoods are categorised broadly into 'Arts & Entertainment', 'College & University', 'Event', 'Food', 'Nightlife Spot', 'Outdoors & Recreation', 'Professional and Other', 'Residence', Shop and Service', 'Travel and transport'. From which we could extract food dataset of Toronto's neighbourhoods.