



Battle of Neighborhood Restaurant Business in Toronto

IBM Data Science Professional Certificate – Capstone Project

Date : 11 May 2021
SHROWTHI Bharadwaja

Introduction

- ❖ Toronto is the capital city of the Canadian province of Ontario. It is the most populous city in Canada and the fourth most populous city in North America.
- ❖ The cuisine of Toronto reflects Toronto's size and multicultural diversity.
- ❖ Canadian cuisine varies widely depending on the regions of the nation. The four earliest cuisines of Canada have indigenous, English, Scottish and French roots. The traditional cuisine of English Canada is closely related to British cuisine.
- ❖ Overtime, with subsequent waves of immigration in the 19th and 20th centuries, Canadian food has been shaped and impacted by those of indigenous people, settlers, and immigrants.
- ❖ Different ethnic neighborhoods throughout the city focus on variety of cuisines. Examples: Chinese, Indian, Italian, Japanese, Caribbean, Jewish, Vegetarian/Vegan, American, Mediterranean, Fast Food Centers etc.
- ❖ A number of culinary festivals take place in Toronto each year. Any trip to Toronto is incomplete without checking out its food tours. Going on any one of these, will without a doubt, leave you satiated.

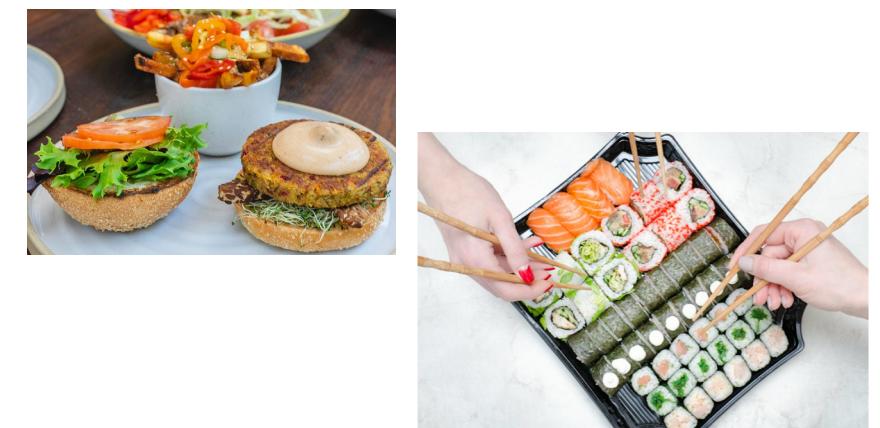
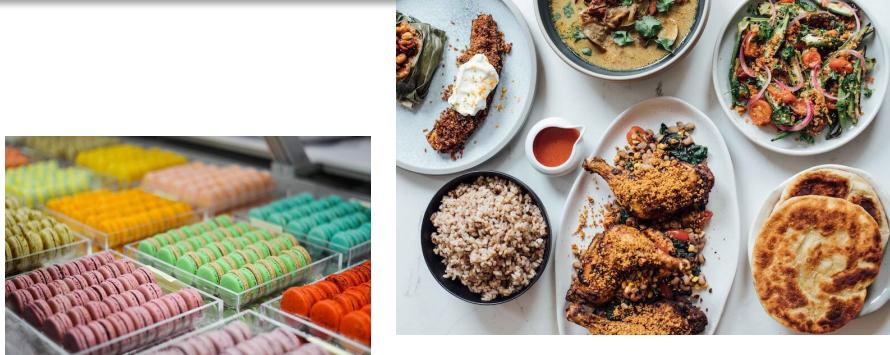


References:

https://en.wikipedia.org/wiki/Cuisine_in_Toronto
<https://tastytourstoronto.com>

Business Problem

- ❖ Toronto is one of the most densely populated areas in Canada with an estimated population of ~ 6 million.
- ❖ The city is home of the finest and most diverse haute cuisine restaurants in the world.
- ❖ Toronto is well known for its great food. Canadian culinary includes an array of international cuisines influenced by multiculturalism of the town.
 - Example: Chinese, Indian, Italian, Japanese, Caribbean, Jewish, Vegetarian/Vegan, American, Mediterranean, Fast Food Centers etc.
- ❖ The objective of this project is:
 - to use Foursquare location data and regional clustering of venue information to determine the ‘best’ neighborhood in Toronto to open a new restaurant with a specific cuisine style.
 - to evaluate the most suitable neighborhood in Toronto for an entrepreneur to invest in a restaurant business. This analysis will provide vital information that can be used by the target audience (Entrepreneurs).



❖ Objective:

1. To recommend the best neighborhood for a restaurant business in Toronto
1. To understand the similarities and differences between the neighborhoods using Unsupervised K-Mean Clustering Algorithm.

❖ Data Sources:

1. Neighborhoods of the city of Toronto for it's Boroughs and Venues.
[https://en.wikipedia.org/w/index.php?title=List_of_postal_codes_of_Canada:_M&oldid=1008658627'\).text](https://en.wikipedia.org/w/index.php?title=List_of_postal_codes_of_Canada:_M&oldid=1008658627').text)
2. Geo-coordinates data for each Neighborhood in Toronto is collected from the following csv file:
https://cocl.us/Geospatial_data
3. Foursquare Venues data for Restaurants and their geographical location using Client credentials.

❖ Data Cleaning (or preprocessing):

- Foursquare API gives latitude and longitude values for specified venues covered within a given circle of radius.
- The foursquare API data will be merged with Toronto's Borough's and Neighborhoods data frame.
- All the NaN for venues are not considered in this analysis
- Boroughs with NaN are replaced with Neighborhoods names.

❖ Feature Selection:

- Toronto has 10 unique Boroughs and has 103 Neighborhoods.
List of Boroughs: Scarborough', 'North York', 'East York', 'East Toronto', 'Central Toronto', 'Downtown Toronto', 'York', 'West Toronto', 'Mississauga', 'Etobicoke
- Out of these, only Boroughs that contains “Toronto” word are chosen for this analysis
- The top venues are retrieved using “group by neighborhoods” and Foursquare API . This list is further filtered for restaurants with the cuisine style.

❖ Analytical Method:

- Unsupervised K-Mean Clustering Algorithm is implemented to segregate the neighborhoods and the respective restaurants distribution per neighborhood.
- Total 5 clusters are considered for this analysis.
- Once the data frame is clustered, only top 10 most common restaurant venues are considered for their geographic distribution per cluster
- Based on the restaurant and their cuisine style distribution per neighborhood, the entrepreneur can consider the suitable location to start a restaurant. The new restaurant cuisine style can be decided based on the existing restaurant distribution. The choice of cuisine style is an option for the entrepreneur investment strategy.

Python packages and Dependencies

❖ The following python libraries (but not limited) are mainly used for the analysis:

- | | |
|----------------|---|
| Pandas | - Library for Data Analysis |
| NumPy | - Library to handle data in a vectorized manner |
| JSON | - Library to handle JSON files |
| Geopy | - To retrieve Location Data |
| Requests | - Library to handle http requests |
| Matplotlib | - Python Plotting Module |
| Sklearn | - Python machine learning Library |
| Folium | - Map rendering Library |
| Beautiful Soup | - Web Scraping and Data Wrangling |
| Foursquare API | - To collect the top trending venues |

Results and Analysis (to be continued in week#5)