

# Site Selection for Japanese Restaurants in Toronto

IBM Data Science Capstone

Liang Nie

# Introduction

- Japanese cuisine is extremely popular throughout North America, primarily in suburban and urban metros like Toronto.
- A good location is crucial to the new restaurant's performance.
- This study explores all the neighborhoods in Toronto in terms of the frequency of occurrence of Japanese Restaurants.
- Cluster analysis is applied to group similar neighborhoods together.
- Business problem: Where's suitable to open a Japanese restaurant in Toronto?
- This study targets entrepreneurs who are thinking about starting a Japanese restaurant in Toronto.

# Data

- Scraping Wikipedia with python  
([https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada:\\_M](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)).
- Since we are only interested in urban area, the final list contains 38 neighborhoods in 4 boroughs.
- Latitudes and longitudes of these boroughs was obtained using Python Geocoder package.
- Venue data for each neighborhood were gathered using Foursquare API.

# Methodology

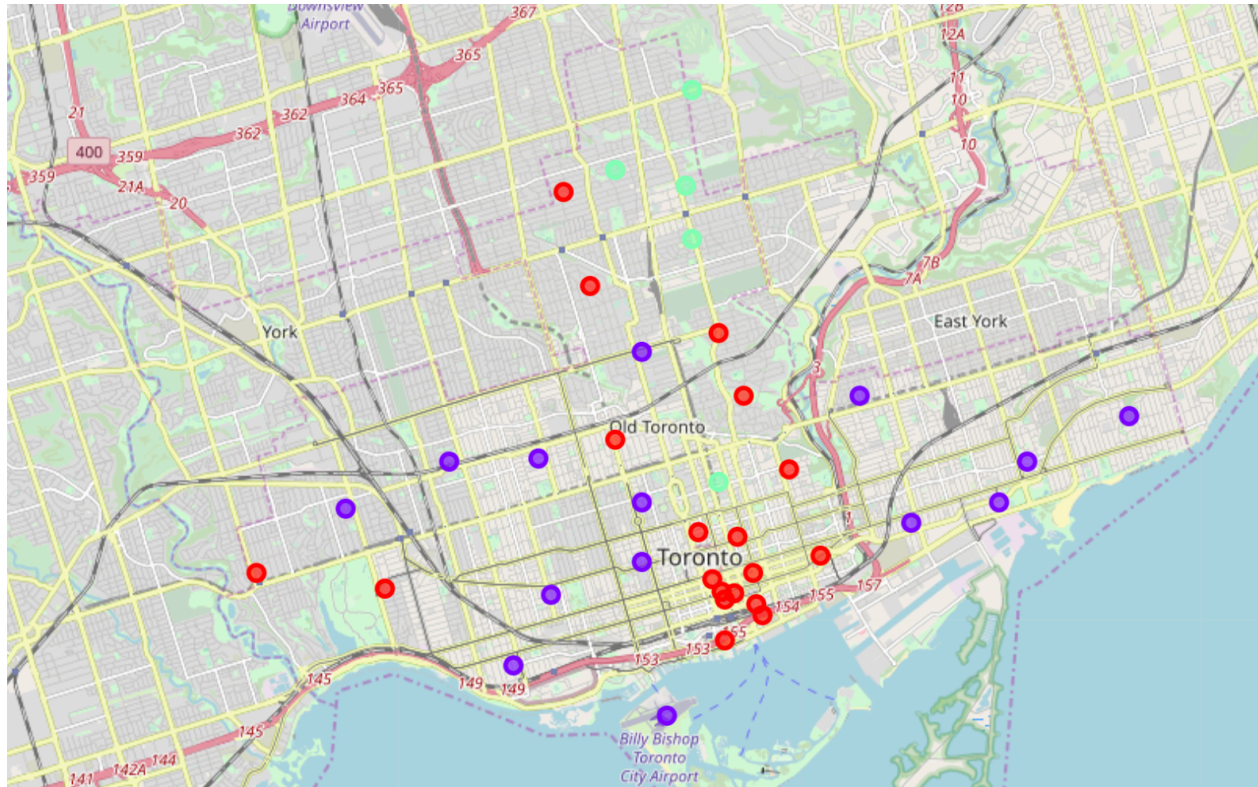
- k-means clustering to cluster the 38 neighborhoods into 3 clusters based on the frequency of occurrence of Japanese restaurant.
- k-means is a simple and popular unsupervised machine learning algorithm. It aims to group similar observations together and thus discover underlying patterns.
- The result allowed us to identify neighborhoods with high, medium and low concentration of Japanese restaurants and would help us to locate the suitable site for the new restaurant.

# Results

- Cluster 0: Neighborhoods with moderate concentration of Japanese restaurants
- Cluster 1: Neighborhoods with relatively low concentration of Japanese restaurants
- Cluster 2: Neighborhoods with relatively high concentration of Japanese restaurants

# Visualization

- We visualized the resulting clusters on the map below with Cluster 0 marked as red, Cluster 1 purple and Cluster 2 green.



# Discussion

- Four out of five Cluster 2 neighborhoods with the most concentration of Japanese restaurants are in Central Toronto area.
- Most of the Cluster 0 neighborhoods with moderate saturation of Japanese restaurants are in Downtown Toronto. Downtown is always crowded with large population density with plenty of business opportunities.
- Cluster 1 neighborhoods, mostly locate in East Toronto and West Toronto, have relatively low saturation of Japanese restaurants.

# Conclusion

- In this project, we gathered location data of Toronto by web scraping and by utilizing Foursquare API.
- We mainly applied k-means clustering, a popular machine learning technique, to group similar neighborhoods in terms of the frequency of occurrence of nearby Japanese restaurants.
- Based on the clustering results, we recommended a few ideal spots for opening a Japanese restaurant. The findings will help the stakeholders to capitalize on the locations with the most business potentials.