

# COURSERA CAPSTONE

IBM APPLIED DATA SCIENCE CAPSTONE

## **OPENING A NEW CHINESE RESTAURANT IN TORONTO, CANADA**

BY: ZUFAR HASEEB  
JULY 2019

# BUSINESS PROBLEM

- Location of the Chinese restaurant is one of the most important decisions that will determine if the restaurant will be a success or a failure.
- Objective: To analyze and select the best locations in Toronto, Canada to open a new Chinese restaurant.
- This project is timely as there is an increasing demand for Chinese cuisine due to the increasing diversity of population in Canada.
- Business Question
  - In the city of Toronto, if an investor/ franchise is planning to open a new Chinese restaurant, where would you suggest them to open it ?

# DATA

- Data Required

- List of neighborhoods in Toronto
- Latitude and Longitude coordinates of neighborhoods
- Venue data, particularly data related to Chinese restaurants

- Sources of Data

- Wikipedia page for neighborhoods

[https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada: M](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)

- Geocoder package for latitude and longitude coordinates
- Foursquare API for venue data

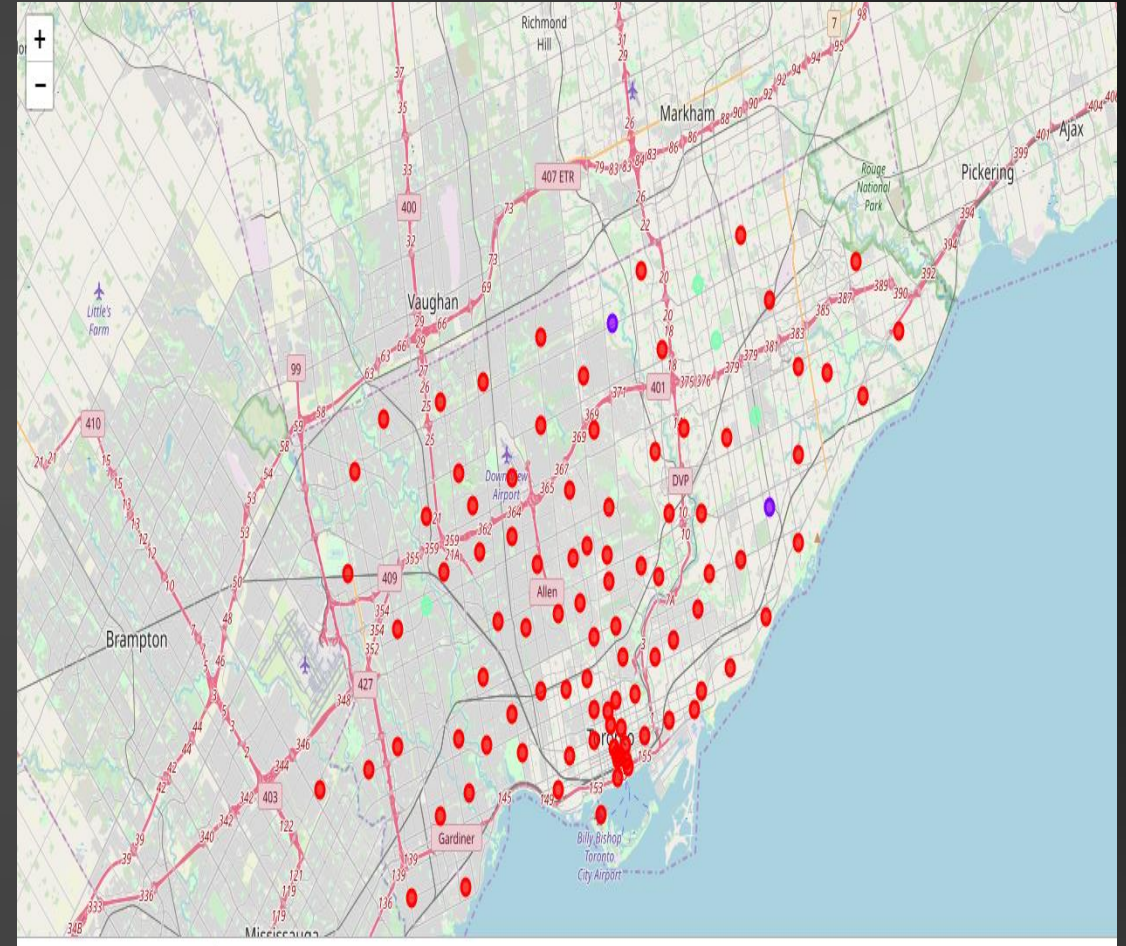
# METHODOLOGY

- Web scraping Wikipedia page for neighborhood list
- Get latitude and longitude coordinates using Geocoder
- Use Foursquare API to get Venue data
- Group data by postcode and taking the mean of the frequency of occurrence of each venue category
- Filter venue category by Chinese restaurants
- Perform clustering on the data using k-means clustering
- Visualize the clusters in a map using folium



# RESULTS

- Categorized the neighborhood into 3 clusters
  - Cluster 0: Neighborhoods with low number to no existence of Chinese restaurants (Red dots)
  - Cluster 1: Neighborhoods with moderate number of Chinese restaurants (Purple dots)
  - Cluster 2: Neighborhoods with high concentration of Chinese restaurant (Mint green dots)



# DISCUSSION

- Most of the Chinese restaurants are concentrated on the east side of Toronto
- Highest number in cluster 2 and moderate number in cluster 1
- Cluster 0 has very low number to no Chinese restaurants in its neighborhoods
- Neighborhoods in cluster 2 are likely to be suffering from over supply of Chinese restaurants

# RECOMMENDATIONS

- Open Chinese restaurants in cluster 0 with little to no competition
- Can also open Chinese restaurants in cluster 1 with moderate level of competition
- Avoid neighborhoods in cluster 2 , already high concentration of Chinese restaurants and intense competition

# CONCLUSION

- Answer to the Business Question:

“ The neighborhoods in cluster 0 are the most preferred locations to open a new Chinese restaurant”

- Findings of this project will help the relevant stakeholders to capitalize on the opportunities on high potential locations while avoiding overcrowded areas in their decisions to open a new Chinese restaurant.



**Thank You !**