

# CLUSTERING TORONTO NEIGHBORHOOD AREAS BY SURROUNDING RESTAURANT TYPE

---

Advay Janardhan

June 24, 2020

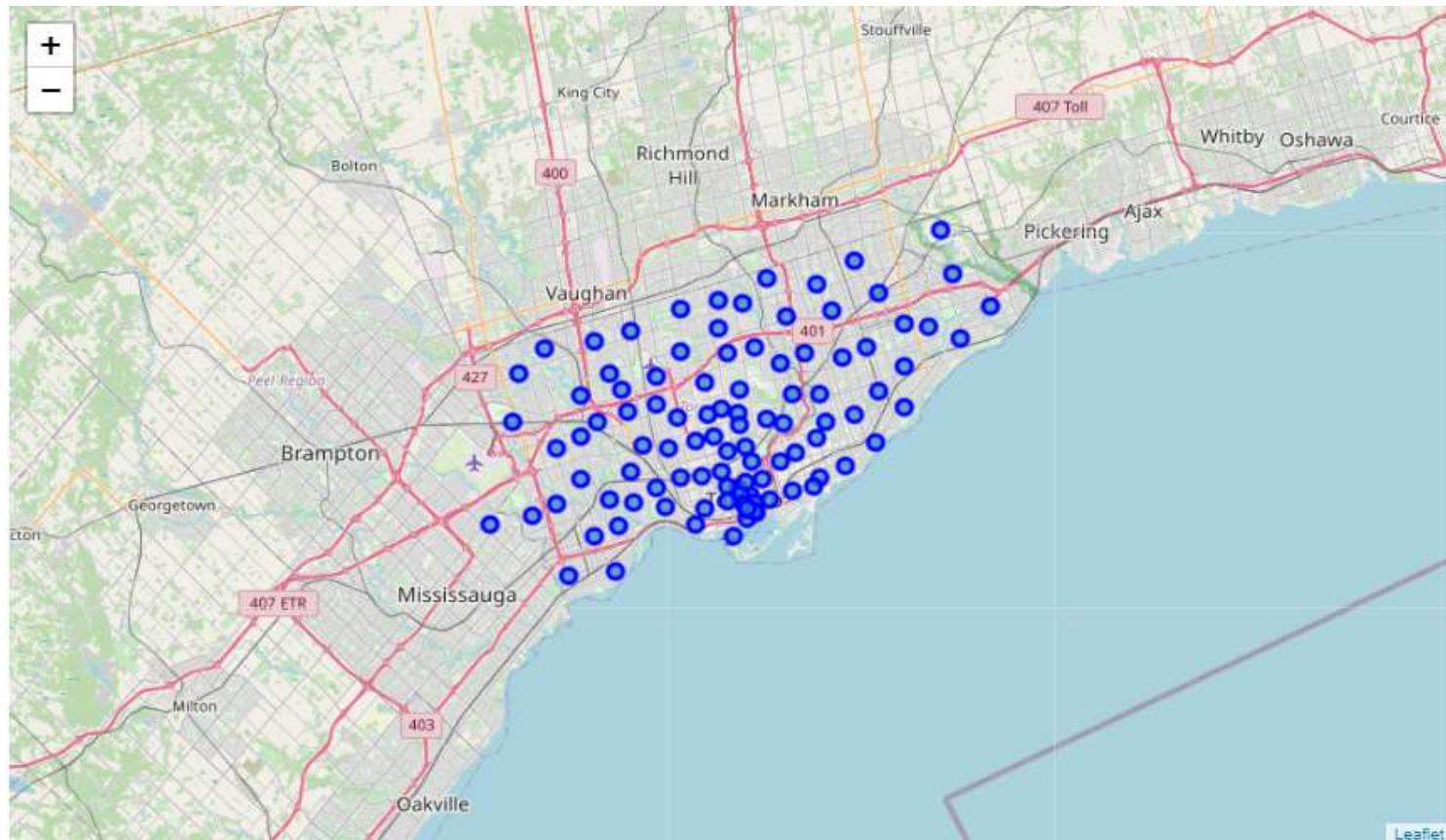
# Useful for Tourists and Restaurants

- Toronto has a booming Tourism and Restaurant industry
- Tourists will benefit from information on neighborhoods
  - They know which neighborhoods have what restaurant types
  - They can choose lodging about the neighborhood of their choice
- Restaurants can choose where to build locations
  - They have information on competitor locations

# Data Extraction

- Postal code and Neighborhood data scraped from [Wikipedia page on Toronto postal codes](#).
- Latitude and Longitude data for postal codes imported from .csv file.
- Information on venues nearby each location found from FourSquare API
- Missing information was dropped

# Map of Neighborhoods



# Data Processing

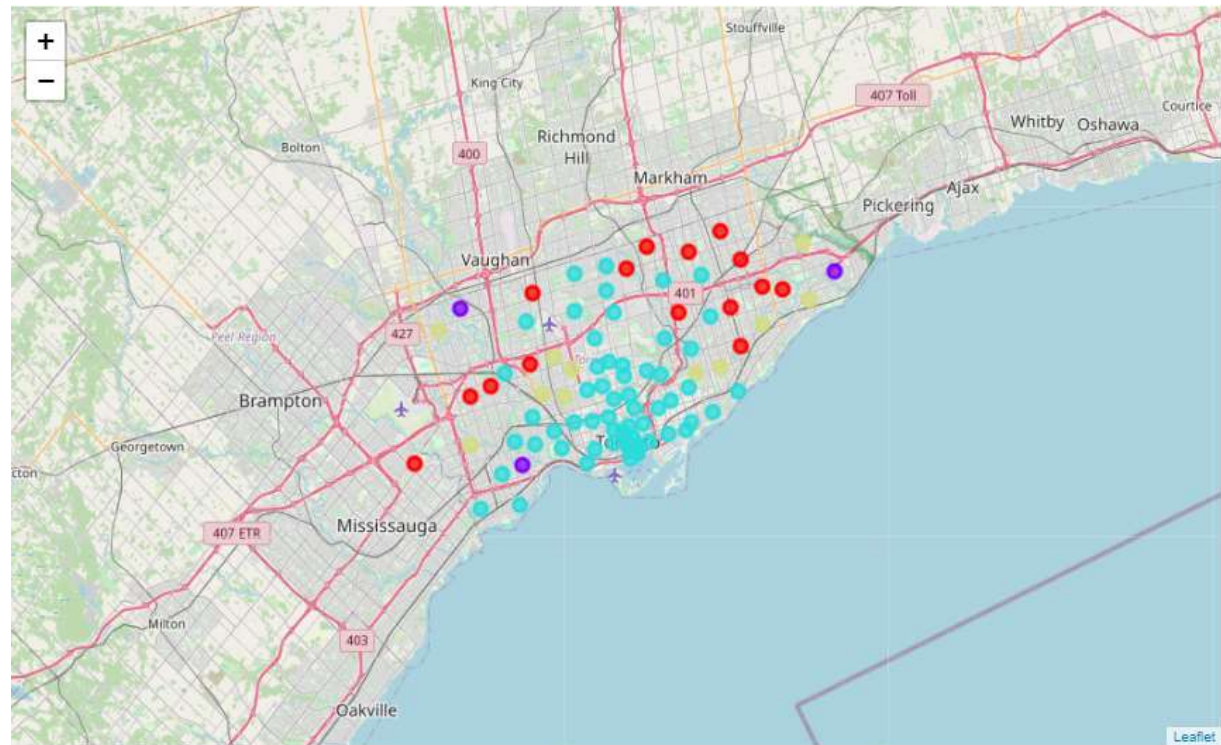
- Venue category data from FourSquare filtered for restaurants
- Restaurant type created from filtered Venue category column
  - The category was in the form “<type> Restaurant”
  - The type was extracted from that string
  - Categories just saying “Restaurant” were dropped

# K-means clustering

- K-means clustering was run with 4 clusters
- Data was encoded by transforming the category variable into many binary variables, transforming categorical variable into a quantitative variable
- Data was grouped by neighborhood to yield one entry for each neighborhood that had frequency of each restaurant type in the row.
- Clusters were marked 0, 1, 2, and 3.

# Map with clusters

- Cluster 0 is red
- Cluster 1 is purple
- Cluster 2 is blue
- Cluster 3 is beige



# Final Dictionary Output

- A final dictionary output was generated
  - Neighborhoods as keys
  - List of surrounding restaurant types as the values
- Cluster 0 characterized by Asian cuisine
- Cluster 1 characterized by Italian cuisine
- Cluster 2 characterized by an equal variety of restaurant types
- Cluster 3 characterized by fast food cuisine



# Conclusions

- The neighborhoods with a variety of options are clustered in central and downtown Toronto
- The neighborhoods characterized by Asian cuisine lie on the outskirts of the area
- The neighborhoods characterized by Italian cuisine also lie on the outskirts of the area
- The neighborhoods characterized by fast food cuisine lie between central and the outskirts of Toronto (both east and west)
- With this information tourists can choose where to lodge and businesses can choose where to build based on their preferences.