Coursera Capstone

IBM Applied Data Science Capstone

Opening a New Restaurant in Toronto, Canada

By:

Pranav Viswanathan

BY:
PRANAV
VISWANATHAN

Business Problem :

The objective of this capstone project is to analyse and select the best locations in the city of Toronto,Canada to open a new Restaurant. Using data science methodology and machine learning techniques like clustering, this project aims to provide solutions to answer the business question:
In the city of Toronto, Canada, if a property developer is looking to open a new restaurant, where would you recommend that they open it?

DATA:

To solve the problem, we will need the following data:

- List of neighbourhoods in Toronto. This defines the scope of this project which is confined to the city of Toronto, the city of the country Canada.
- Latitude and longitude coordinates of those neighbourhoods. This is required in order to plot the map and also to get the venue data.
- Venue data, particularly data related to public places,tourist spot . We will use this data to perform clustering on the neighbourhoods.

SOURCES OF DATA

- This Wikipedia page (https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M) contains a list of neighbourhoods in canada.
- we will get the geographical coordinates of the neighbourhoods using (https://cocl.us/Geospatial_data)which will give us the latitude and longitude coordinates of the neighbourhoods
- we will use Foursquare API to get the venue data for those neighbourhoods

METHODOLOGY

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

Results

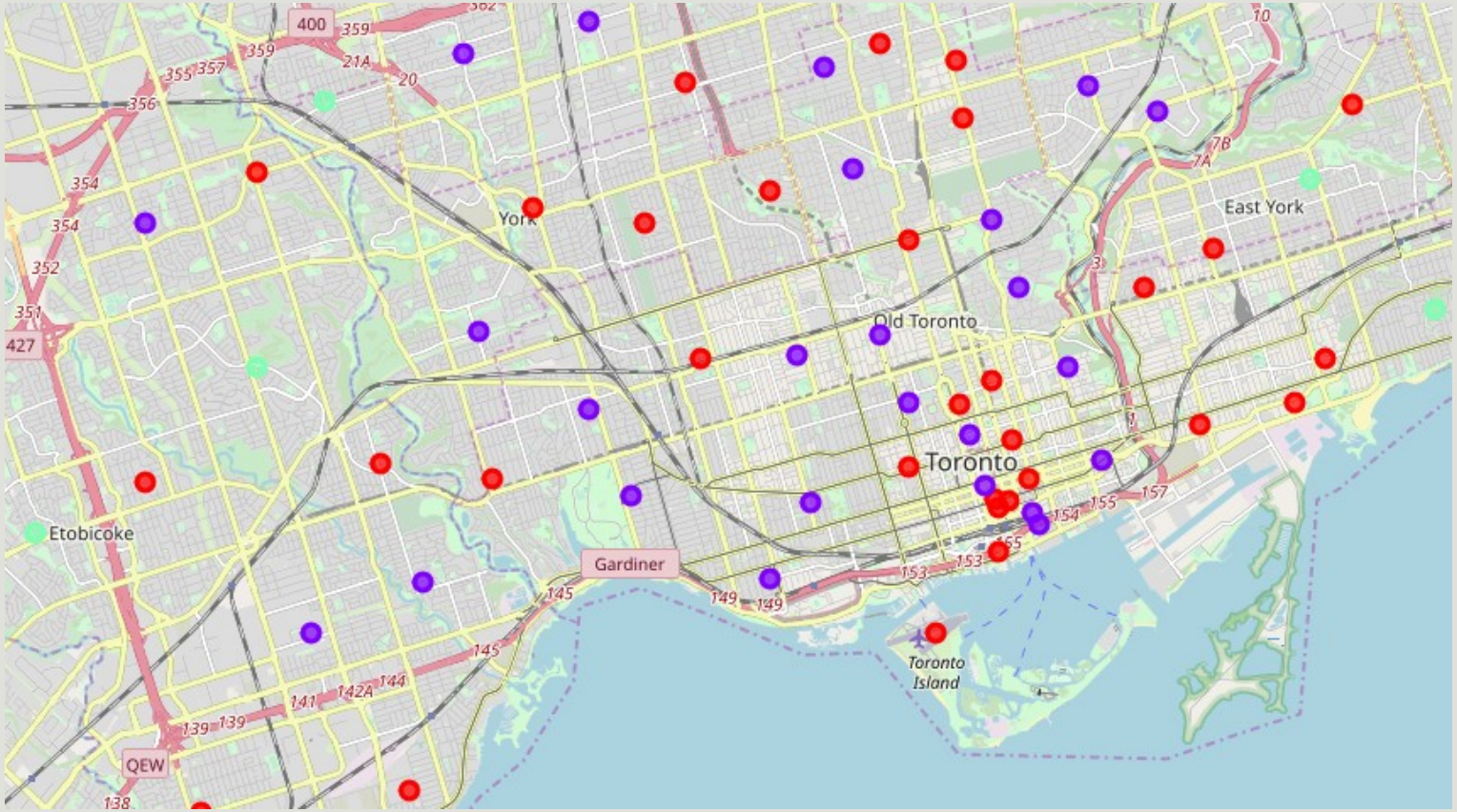
- Categorized the neighbourhoods into 3 clusters :

➢Cluster 0: Neighbourhoods with moderate number of restaurant’s

➢Cluster 1: Neighbourhoods with low number to no existence of restaurant’s

➢Cluster 2: Neighbourhoods with high concentration of restaurant’s

FOLIUM MAP


400
York
East York
Old Toronto
Toronto
Etobicoke
Gardiner
QEW
Toronto Island

Discussion

- Most of the Restaurants are concentrated in the central area of the city
- Highest number in cluster 2 and moderate number in cluster 0
- Cluster 1 has very low number to no restaurant’s in the neighbourhoods
- Oversupply of restaurant’s mostly happened in the central area of the city, with the suburb area still have very few Restaurant’s.

Recommendations

- Open new Restaurant in neighbourhoods in cluster 1 with little to no competition
- Can also open in neighbourhoods in cluster 0 with moderate competition if have unique selling propositions to stand out from the competition
- Avoid neighbourhoods in cluster 2, already high concentration of Restaurant’s and intense competition

Conclusion

- Answer to business question: The neighbourhoods in cluster 1 are the most preferred locations to open a new 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 restaurant’s.