Data Science Capstone Final Project

PROJECT NAME : MY INDIAN RESTAURANT

Introduction

- This is a capstone project for IBM Data Science Professional Certificate. In this project, I am creating a hypothetical scenario for a concept that there may not be enough Indian Restaurants in Toronto Area.

- Therefore it might be a great opportunity for an entrepreneur who is based in Canada.

- Indian food is popular among the Asian community, so The entrepreneur might think of opening its business in areas where the Asian community resides. With the purpose in mind, finding the location to open such a restaurant is one of the most important decisions for this entrepreneur and I am designing this project to help him find the most suitable location.

Data & its Acquisition

To solve this problem, we will need below data:

- List of neighborhoods in Toronto, Canada
- Latitude and Longitude of these neighborhoods
- Venue data related to Indian restaurants. This will help us find neighborhoods that are more suitable to open an Indian Restaurant.
- the list of neighborhoods from Wikipedia: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

EXTRACTING THE DATA

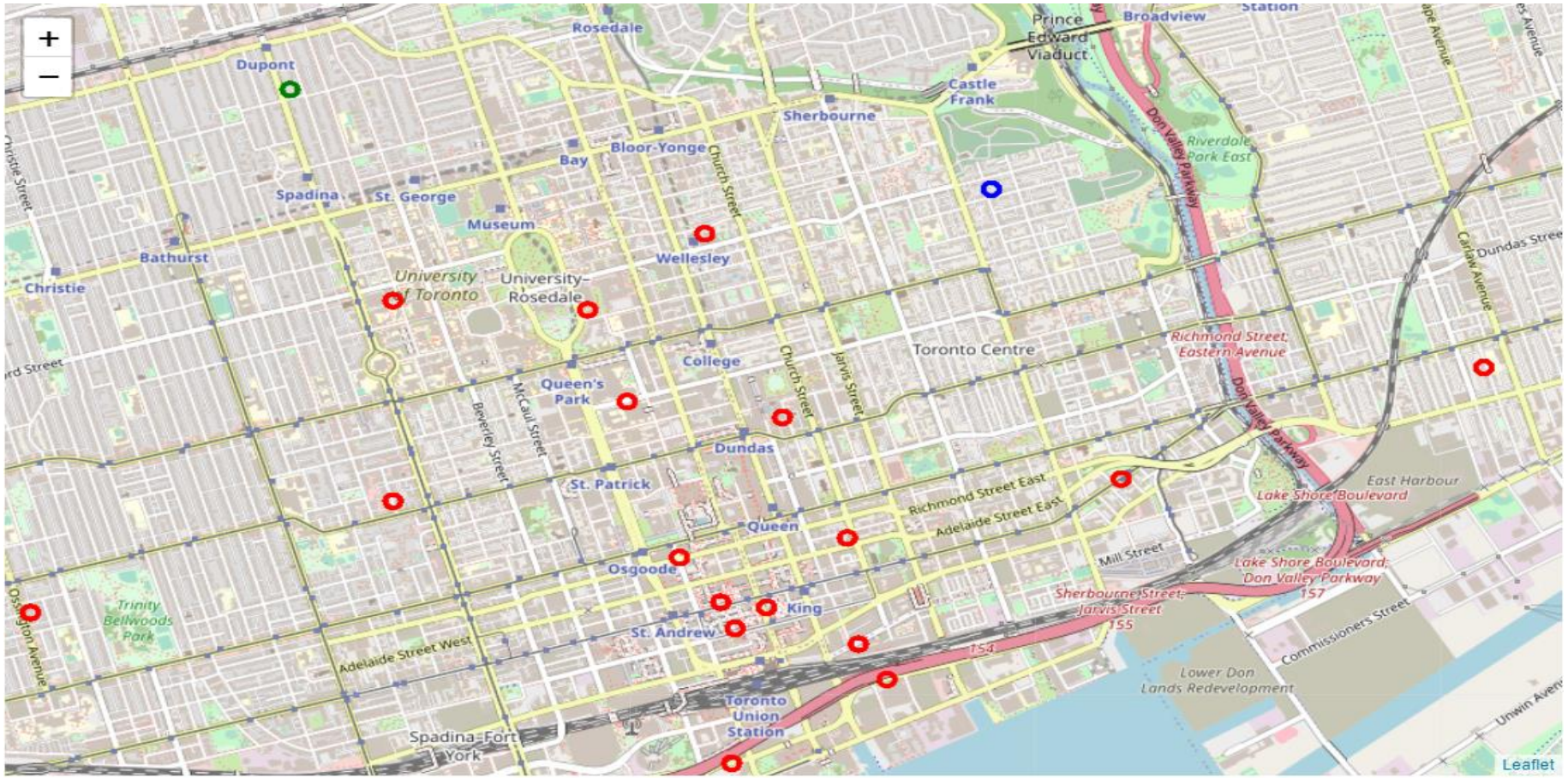
- The scrapping of Toronto neighborhoods via Wikipedia
- Getting Latitude and Longitude data of these neighborhoods via Geocoder package
- Using Foursquare API to get venue data related to these neighborhoods

METHODOLOGY

- I did the web scraping by utilizing pandas HTML table scraping method as it is easier and more convenient to pull tabular data directly from a web page into the data frame
- I have created a Foursquare developer account in order to obtain account ID and API key to pull the data. From Foursquare, I am able to pull the names, categories, latitude, and longitude of the venues.
- I made a justification to specifically look for "Indian restaurants". Lastly, I performed the clustering method by using k-means clustering.
- Based on the results (the concentration of clusters), I will be able to recommend the ideal location to open the restaurant.

RESULT

CLUSTERS



Continue...

- Cluster 0: Neighborhoods with almost no Indian restaurants.
- Cluster 1: Neighborhoods with a more number of Indian restaurants.
- Cluster 2: Neighborhoods with less number of Indian restaurants.

The results are visualized in the above map with Cluster 0 in green, Cluster 1 in blue, Cluster 2 in red.

RECOMMENDATIONS

- Looking at nearby venues it seems cluster 0 might be a good location as there are not a lot of Indian restaurants in these areas. Therefore, this project recommends the entrepreneur to open an authentic Indian restaurant in these locations.
- For Cluster 3 also, The entrepreneur has a golden opportunity to start Indian restaurant.

Conclusion

- This project recommends the entrepreneur to open an authentic Indian restaurant in such locations in Toronto where he can .grow his/her business by targeting especially the Indian Customers.

Thank You…