New Indian restaurant in toronto

IBM Applied Data Science Capstone from Coursera

# Background

I have been thinking of moving to Toronto next year and is thinking about opening an Indian restaurant. The Indian cuisine is one of the best in the world, and opening a restaurant has always been a big dream. However, I have never been to Toronto, so a location has to be chosen from information available via internet. I would like to find a location without any Indian restaurants at this point, in order to avoid competition.

Data from various sources will be used for this project. Wikipedia will be used to extract geological details of the Toronto neighbourhoods. Foursquare location data will be used in order to come up with the best possible location for this restaurant.

# Problem

Opening a restaurant in a new place is challenging, especially if you have never been in such a place. This project will look into the best neighbourhood in Toronto to open an Indian restaurant in. The client wants to find a neighbourhood without any other Indian restaurants.

# Data Collection

The following sources will be used to collect data from:

**Wikipedia**

Wikipedia will be used to extract the neighbourhood data. The website <https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M> will be the main source for neighbourhoods in Toronto.

**Geospatial data**

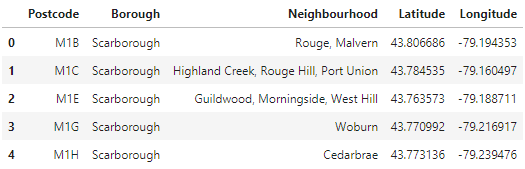
The geospatial data from <https://cocl.us/Geospatial_data/Geospatial_Coordinates.csv> will be used to collect the geospatial coordinates of the different neighbourhoods.

**Foursquare**

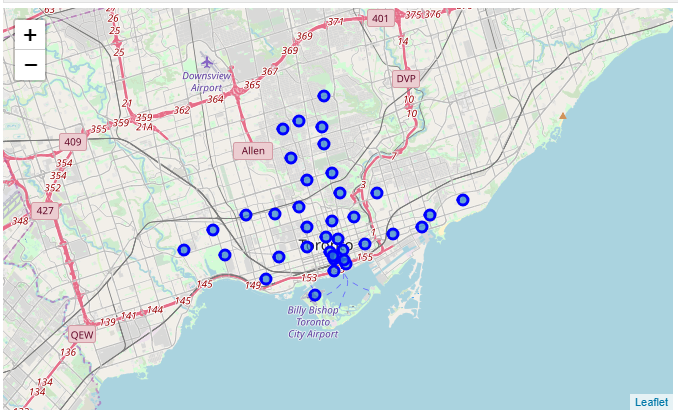
The Foursquare API will be used to find venues on different locations and find top venues per neighbourhood.

# Methodology

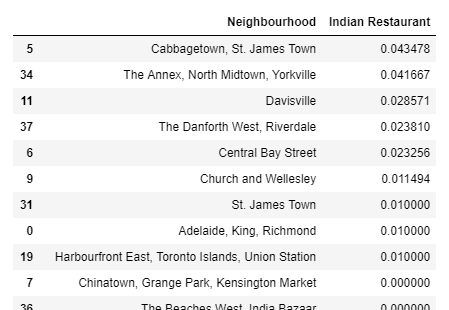
After loading all necessary Python packages, a list of Canadian neighbourhoods were extracted from <https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M>. The data was filtered, an all neighbourhoods of Toronto were listed. The geospatial data was extracted from <https://cocl.us/Geospatial_data/Geospatial_Coordinates.csv> and the postcode, latitude and longitude were added to the neighbourhood data.



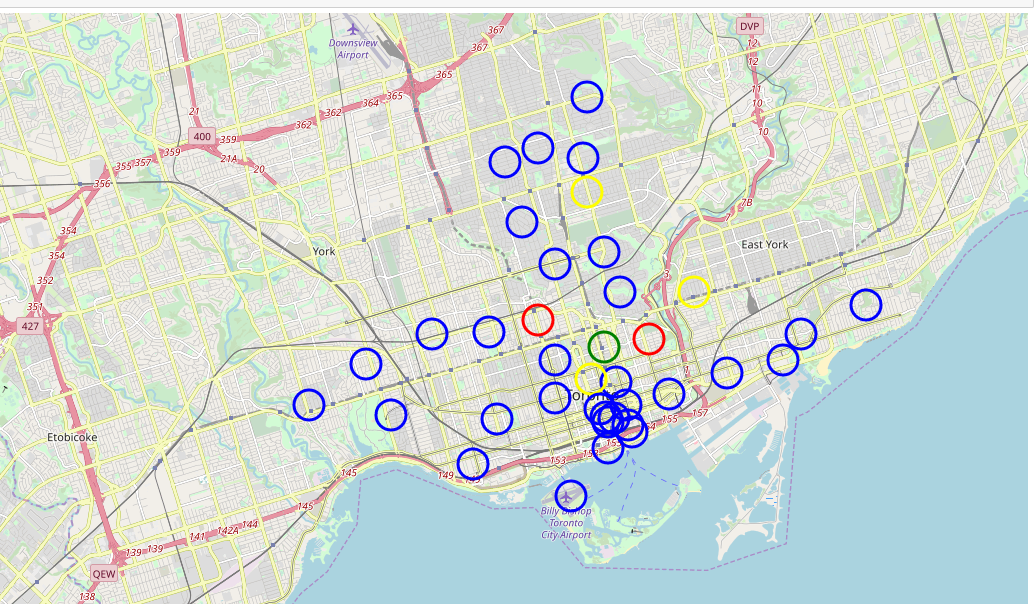
A map was now created of the different neighbourhoods in Toronto.



Now the Foursquare API was used to extract all the Indian Restaurants in Toronto grouped by neighbourhood and sorted by frequency. From this data, it follows that most Indian restaurants are currently located in Cabbagetown, St James town, The Annex, North Midtown and Yorkville as shown in the following table.



This data was merged with the latitude and longitude data and a heat map was made. Red and yellow indicate that there are already several restaurants in the area. Blue and green indicate that there is no or very little competition in the area.



Extracting all green and blue locations, the green locations indicate little competition by other Indian restaurants, the blue ones have almost none to no Indian restaurants present. The recommendation is to look at these neighbourhoods for future investments.

# Results

Surprisingly, there are 33 neighbourhoods in Toronto with little or no Indian restaurants. There is therefore a big market that could be reached by moving into one of those areas.

# Discussion

This provides initial assessment of suburbs for Indian restaurant. Final decision can be taken based on additional requirements such as population, mall nearby , spending power etc.

# Conclusion

Data science can be used to make a decision based on only data. Above example shows selection of neighbourhoods without any knowledge of the city. Data science is very powerful.