2020

Final capstone project – Coursera – IBM – Data science - Report

THE BATTLE OF NEIGHBOURHOODS – ASIAN RESTAURANTS IN TORONTO CANADA

SINGARAM, KRISHNAVENI

Contents

Good Asian Restaurants in Toronto

Introduction

The project is intended to submit to Coursera-IBM as a Final Capstone project.

The problem take to analyse is to explore the neighbourhoods of Toronto Canada, cluster the neighbourhoods and short list the good Asian restaurants for the Indians or any traveller who visits Canada Toronto for a short time.

Problem Statement

Many Indian people go to Canada on business trips who usually search for Good Asian restaurants where they get good Asian food.

Now-a-days Indian people prefer also the other Asian countries food than Indian food. Some of the Chinese, Japanese, Thai food items became regular in-takes whenever they visit hotels both within India or other countries

And people check the ratings of the hotels before they choose to visit

The project is to help the people who go on business trips to Canada - Toronto to find better staying place which has good Asian restaurants

Target Audience

The Indian business people or anyone who is visiting Canada - Toronto for a short term stay

The audience is not restricted to Indian people but all the visitors of Toronto who are interested in Asian food items

The visitors who usually refer the ratings of the restaurants before they choose to visit them

Data

Source

Wikipedia lists the postal codes of Toronto in the follwing link https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

The data source will be scraped to obtain the postal codes

Assumption

This project makes an assumption that the most people visits downtowns for the business purposes. And hence after scraping the data in the link above, the filtered data would have only downtowns

Geographical coordinates

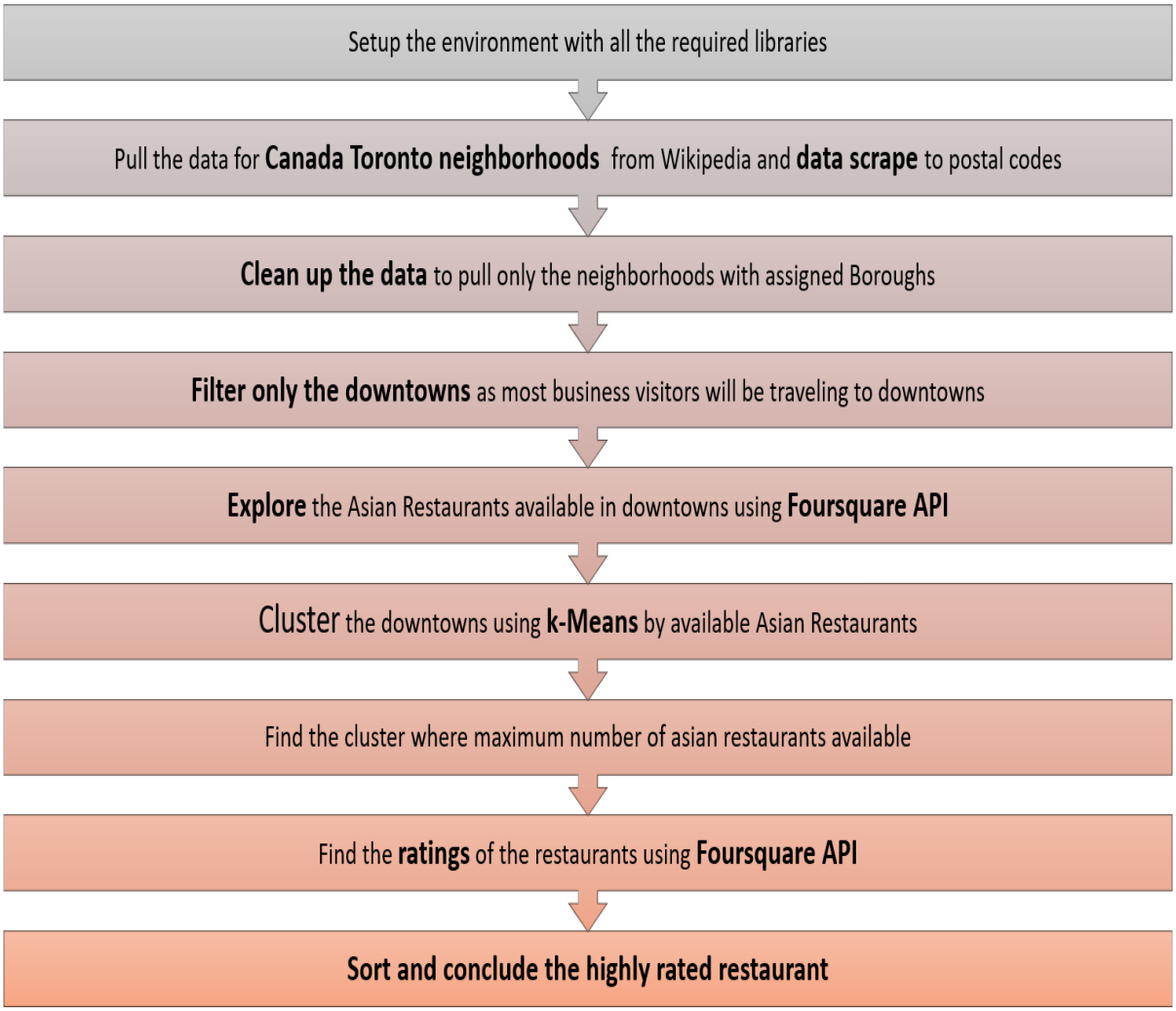
To get the geographical coordinates of the neighbourhoods using the Geocoder package, the following link is used to pull a csv file that has the geographical coordinates of each postal code: http://cocl.us/Geospatial_data

Explore the neighbourhoods in Toronto using Foursquare API

Once we have the filtered data and the geographical coordinates, The neighbourhoods of Toronto will be explored using Foursquare places API https://developer.foursquare.com/places

The places will be shortlisted for Asian restaurants and would be checked for ratings using Four square API

Methodology



Statistics

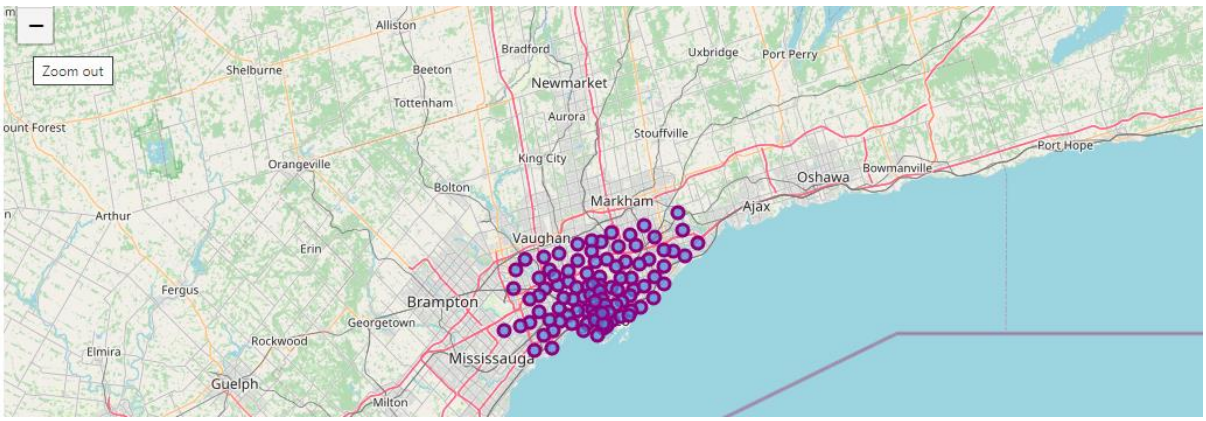
Exploring the neighbourhoods for filtering only downtowns

There are 103 Neighbourhoods are found when we data scrape the Wikipedia page and clean them with ONLY BOROUGHS ASSIGNED

Out[10]:

	Postal Code	Borough	Neighborhood
0	M3A	North York	Parkwoods
1	M4A	North York	Victoria Village
2	M5A	Downtown Toronto	Regent Park, Harbourfront
3	M6A	North York	Lawrence Manor, Lawrence Heights
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government
...	...	...	...
98	M8X	Etobicoke	The Kingsway, Montgomery Road, Old Mill North
99	M4Y	Downtown Toronto	Church and Wellesley
100	M7Y	East Toronto	Business reply mail Processing Centre, South C...
101	M8Y	Etobicoke	Old Mill South, King's Mill Park, Sunnylea, Hu...
102	M8Z	Etobicoke	Mimico NW, The Queensway West, South of Bloor,...

103 rows × 3 columns



The neighbourhoods are filtered only for Downtowns as per the assumption that most of the business visitors would be traveling to Downtowns (It may not be the real case but this project is intended to the visitors who would be traveling to Downtowns.

When the data frame is fileted for only downtowns, we get the following list of downtowns

Filtering only the places who have Toronto in the name

```
In [43]: filtered_data = TorontoPostalData[TorontoPostalData['Borough'].str.contains('Toront
         filtered_data.head()
```

Out[43]:

	Borough	Neighborhood	Latitude	Longitude
0	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
1	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494
2	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937
3	Downtown Toronto	St. James Town	43.651494	-79.375418
4	East Toronto	The Beaches	43.676357	-79.293031

Explore the Asian Restaurants in Downtowns

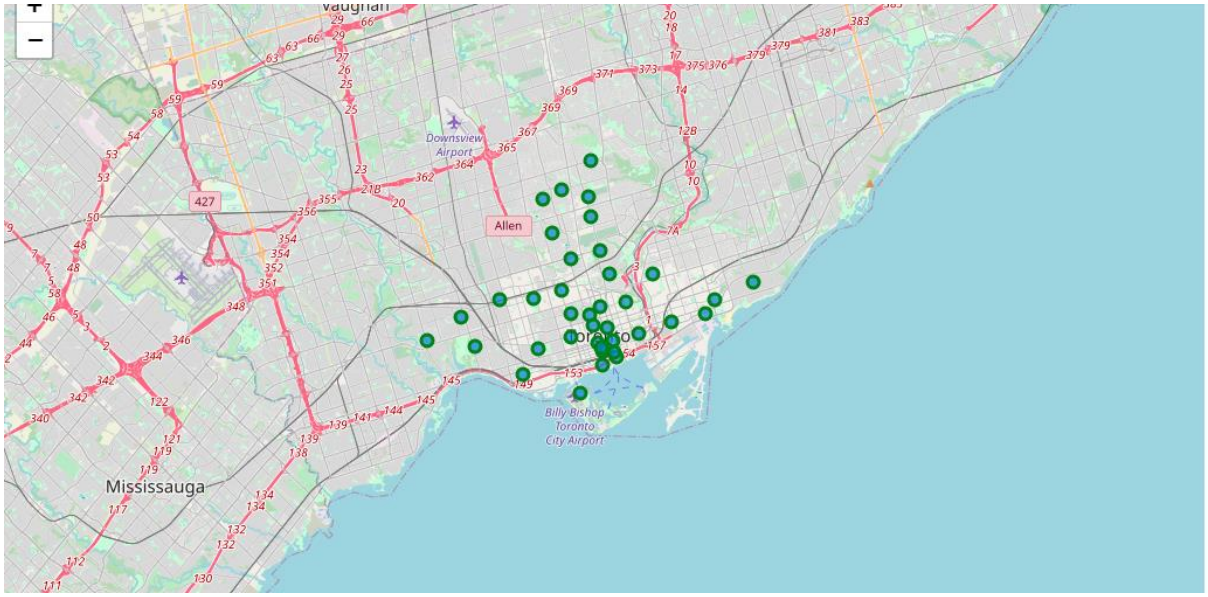
ML Algorithm – k-Means

k-Means is used to cluster the neighbourhoods. Three clusters are defined and others are outliers.

Four square API is used to explore the venues in the downtowns which have Asian Restaurants.

Once the list Asian restaurants is pulled, the downtowns are clustered using k-Means by Neighbourhoods

The below map shows the clustering of Neighbourhoods by Asian Restaurants.



Asian Restaurants East Toronto

While analysing the clusters, it is deduced that East Toronto Downtown has maximum number of Asian Restaurants (4) out of total 8 restaurants available in all the Downtowns

[21]:
```
<bound method NDFrame.head of             Borough                     Neighborhood
0  Downtown Toronto       Regent Park, Harbourfront  50e88fc6e4b007fcbb57aae8
1      East Toronto                     The Beaches  4c535e4afd2ea593856b4f28
2      East Toronto                     The Beaches  4dbc53541e72b351caafda72
3      West Toronto        Little Portugal, Trinity  4af369d0f964a52060ed21e3
4      West Toronto        Little Portugal, Trinity  54c46166498edeaa806cabc8
5      West Toronto        Little Portugal, Trinity  58b0e82d5d6ec60662cbda46
6      East Toronto    The Danforth West, Riverdale  4adcb1a9f964a520d52e21e3
7      East Toronto  India Bazaar, The Beaches West  5859d15c1d21ba53f2f9b4ef

                         Name
0                       Izumi
1                   The Goof
2         Seaspray Restaurant
3               Foxley Bistro
4                     Hanmoto
5              Pinky's Ca Phe
6  Danforth Dragon Restaurant
7                   Lake Inez  >
```

[22]:
```
asian_rest_to.shape
```

[22]:
```
(8, 4)
```

```
plt.xlabel('Borough', fontsize = 15)
#On y-axis
plt.ylabel('No.of Asian Restaurant', fontsize=15)
#giving a bar plot
asian_rest_to.groupby('Borough')['ID'].count().plot(kind='bar', color='purple')
#Legend
plt.legend()
#displays the plot
plt.show()
```

As we have maximum number of Asian restaurants in East downtown as per the graph above, the Foursquare venue details API is used to pull the details about restaurants for data like Likes, Ratings, visits etc.,

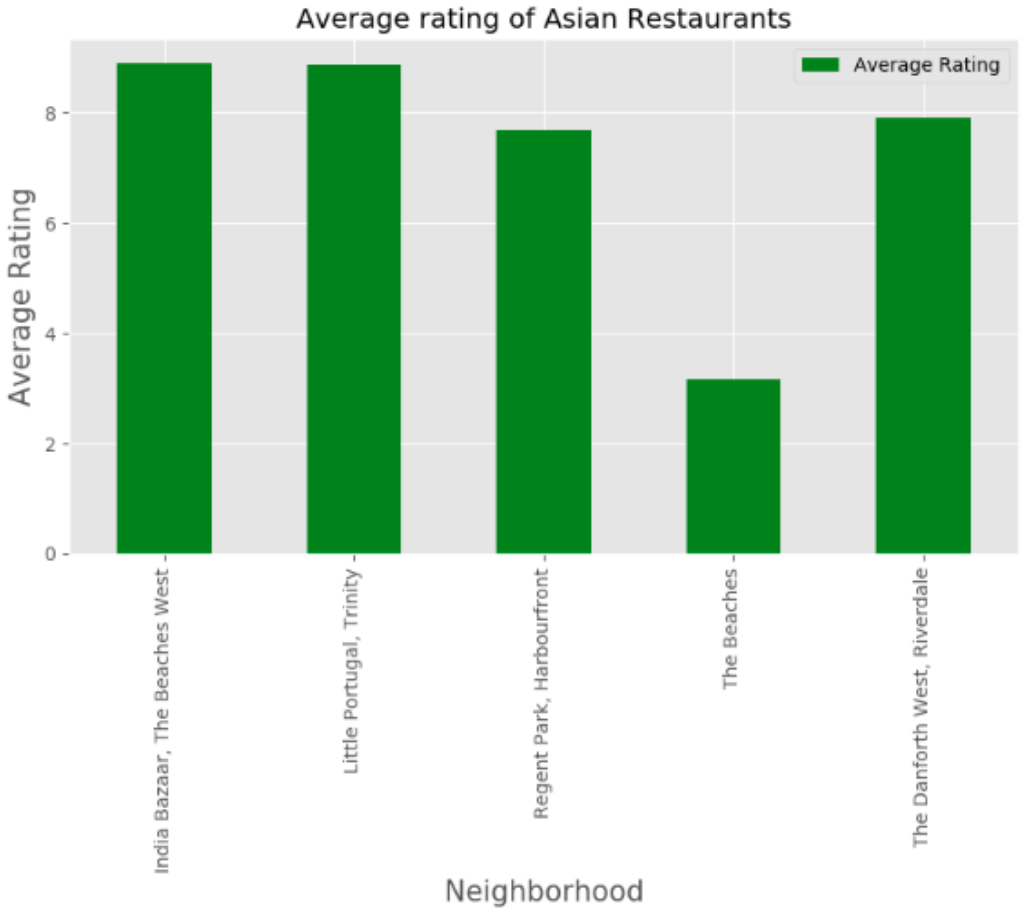
It will be wiser to check the average rating to select the top restaurant and hence the average rating is calculated and the details are as per the below graph

```
rest_ratings.sort_values(['Average Rating'],ascending=False).head(10)
```

[75]:

	Neighborhood	Average Rating
0	India Bazaar, The Beaches West	8.900000
1	Little Portugal, Trinity	8.866667
4	The Danforth West, Riverdale	7.900000
2	Regent Park, Harbourfront	7.700000
3	The Beaches	3.150000

Conclusion



k-Means algorithm has been used to cluster and explore the neighbourhoods. **Three clusters** are created and the cluster with higher dense of Asian Restaurants was explored to find the best rated Asian Restaurant. The chosen downtown was **East Downtown.**

As per the statistics explained above and as per the graph show here, **Inda Bazaar, The Beaches West is the best rated in East Toronto and the East Toronto has the maximum number of Asian Restaurants.**