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

The Battle of Neighbourhoods – Asian Restaurants in Toronto Canada

Singaram, Krishnaveni

2020

Contents

[Good Asian Restaurants in Toronto 2](#_Toc46133513)

[Introduction 2](#_Toc46133514)

[Problem Statement 2](#_Toc46133515)

[Target Audience 2](#_Toc46133516)

[Data 3](#_Toc46133517)

[Source 3](#_Toc46133518)

[Assumption 3](#_Toc46133519)

[Geographical coordinates 3](#_Toc46133520)

[Explore the neighbourhoods in Toronto using Foursquare API 3](#_Toc46133521)

[Methodology 4](#_Toc46133522)

[Statistics 4](#_Toc46133523)

[Exploring the neighbourhoods for filtering only downtowns 4](#_Toc46133524)

[Explore the Asian Restaurants in Downtowns 6](#_Toc46133525)

[ML Algorithm – k-Means 6](#_Toc46133526)

[Asian Restaurants East Toronto 7](#_Toc46133527)

[Conclusion 9](#_Toc46133528)

# 

# 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

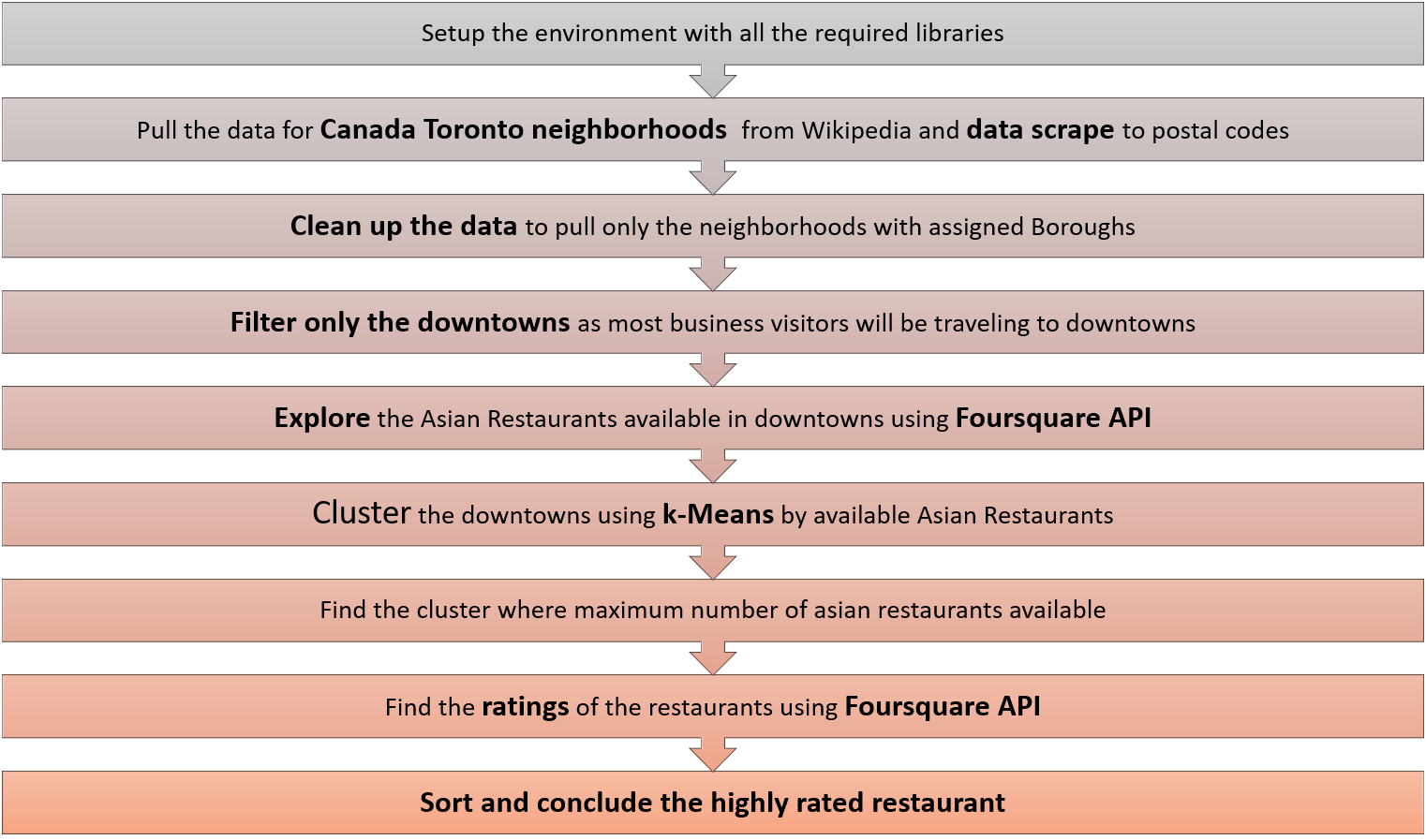
To get the geographical coordinates of the neighbourhoods using the Geocoder package, the following link is used to pull a csv file that has the geographical coordinates of each postal code: <http://cocl.us/Geospatial_data>

### 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

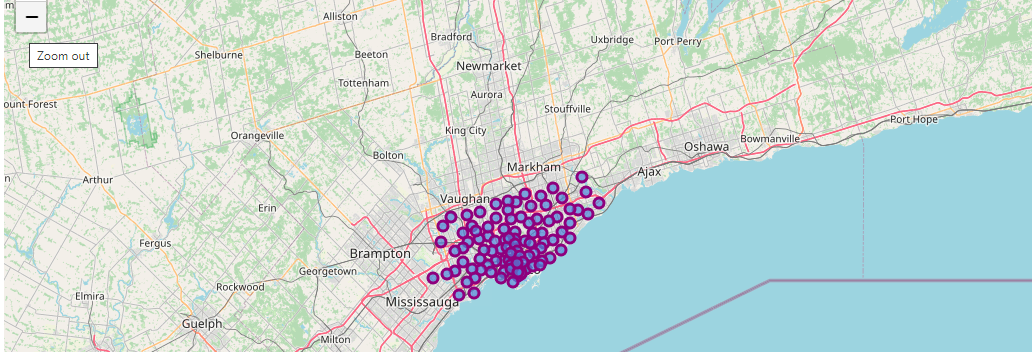


## Statistics

### Exploring the neighbourhoods for filtering only downtowns

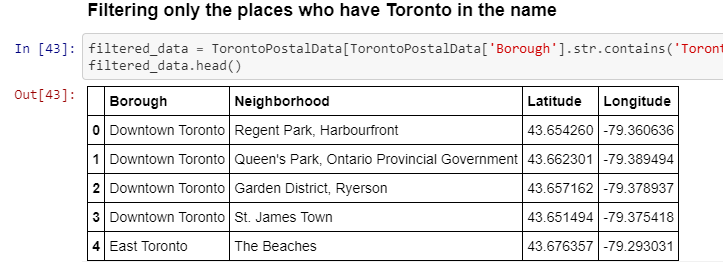
There are 103 Neighbourhoods are found when we data scrape the Wikipedia page and clean them with ONLY BOROUGHS ASSIGNED





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



## 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.

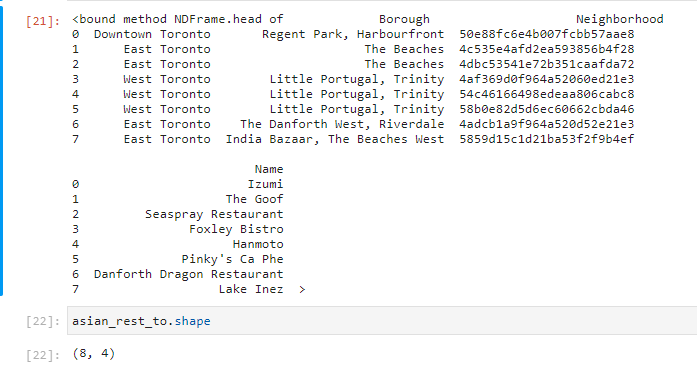
Once the list Asian restaurants is pulled, the downtowns are clustered using k-Means by Neighbourhoods

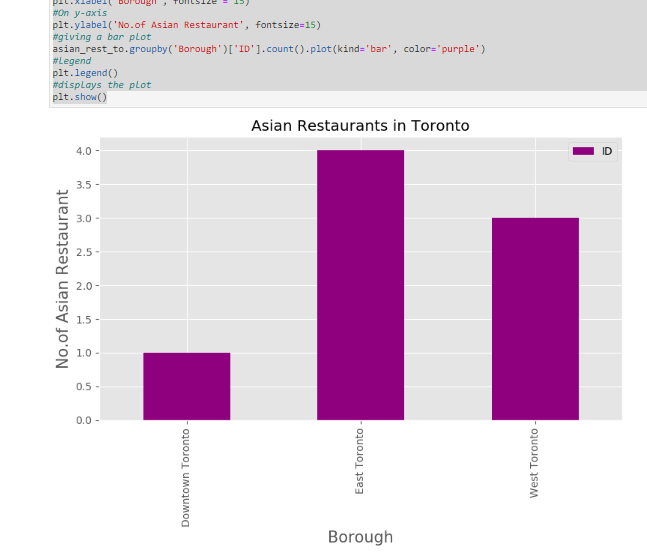
The below map shows the clustering of Neighbourhoods by Asian Restaurants.



## Asian Restaurants East Toronto

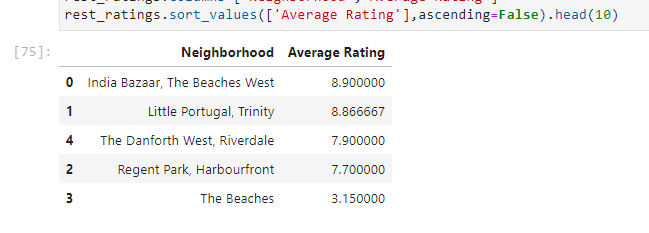
While analysing the clusters, it is deduced that East Toronto Downtown has maximum number of Asian Restaurants (4) out of total 8 restaurants available in all the Downtowns



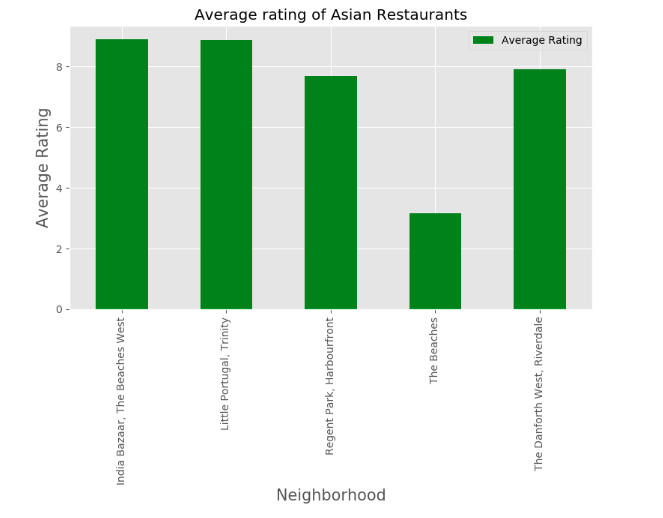


As we have maximum number of Asian restaurants in East downtown as per the graph above, the Foursquare venue details API is used to pull the details about restaurants for data like Likes, Ratings, visits etc.,

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



## 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.**