**Contents**

[**IBM FINAL CAPSTONE PROJECT** 2](#_Toc41722936)

[INTRODUCTION 2](#_Toc41722937)

[TARGET AUDIENCE 3](#_Toc41722938)

[DATA 3](#_Toc41722939)

[EXTRACTING THE DATA 3](#_Toc41722940)

[METHODOLOGY 3](#_Toc41722941)

[Step 1: Demographics Data Extraction from Wikipedia: 3](#_Toc41722942)

[Step 2: Get the Geo Coordinates for the neighborhoods, from Step 1: 3](#_Toc41722943)

[STEP 3: Use the neighborhood Geo Coordinates from step 2, to extract venues around these neighborhoods: 4](#_Toc41722944)

[STEP 4: Use of venue details from Step 3 for the analysis: 4](#_Toc41722945)

[STEP 5: Find the best Neighborhood to open a new restaurant: 4](#_Toc41722946)

[RECOMMENDATIONS 5](#_Toc41722947)

**IBM FINAL CAPSTONE PROJECT**

# INTRODUCTION

This is a capstone project for IBM Data Science Professional Certificate. This project is about, smart data analysis using advanced Analytics and Geographical data points to find the best location/Neighborhood in Toronto Downtown for a Restaurant start-up.

Since the project is to identify the best suitable location in Toronto for the Restaurant business, let’s start with little background on Toronto and understand the New business opportunity and challenges.

Toronto is the provincial capital of [Ontario](https://en.wikipedia.org/wiki/Ontario) and the [most populous city in Canada](https://en.wikipedia.org/wiki/List_of_the_100_largest_municipalities_in_Canada_by_population). The city is the anchor of the [Golden Horseshoe](https://en.wikipedia.org/wiki/Golden_Horseshoe), an urban agglomeration of 12 million people surrounding the western end of [Lake Ontario](https://en.wikipedia.org/wiki/Lake_Ontario). Toronto is an international centre of business, finance, arts, and culture, and is recognized as one of the most multicultural and cosmopolitan cities in the world.

B**USINESS PROBLEM**

Because of it’s multicultural environment and being the financial hub of Canada, the opportunity is enormous and the same time market is highly competitive. Hence, any new business venture needs to be thought through and analysed carefully supported by data facts.

Through this project we are going to help find the best strategical location for a new Restaurant business setup in Toronto, Canada.

By using data science methods and tools along with machine learning algorithms such as clustering, this project aims to provide solutions to answer the business question: What part of Toronto, is the best suitable location to start new Restaurant Business

# TARGET AUDIENCE

The entrepreneur who wants to find the location to open a new restaurant business in Toronto Downtown.

# DATA

To solve this problem, we will need below data:

* Geo Coordinates of Neighbourhoods in Toronto Downtown, Canada
* Venue data around Toronto Neighborhood to understand two things:
  + - Happening places around downtown
    - Neighborhood with a smaller number of Restaurants.

## EXTRACTING THE DATA

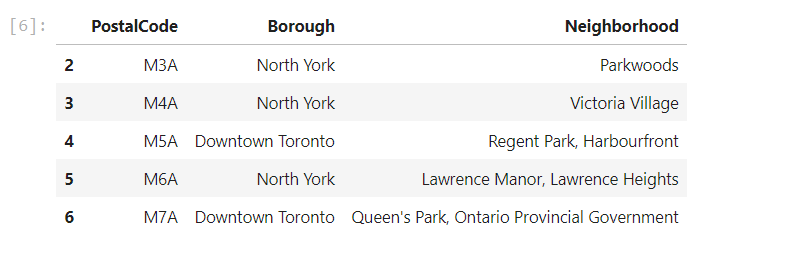
* The scrapping of Toronto Neighborhoods via Wikipedia
* Get Latitude and Longitude of these Neighbourhoods via Geocoder package
* Using Foursquare API to get venue data related to these Neighbourhoods

# METHODOLOGY

## Step 1: Demographics Data Extraction from Wikipedia:

First, we need to get the list of neighbourhoods in Toronto, Canada. This is possible by extracting the list of neighbourhoods from Wikipedia: ​<https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M>​

Did the web scraping on Wikipedia page, by utilizing BeautifulSoup4 and got the data of interest extracted into a DataFrame for further processing



## Step 2: Get the Geo Coordinates for the neighborhoods, from Step 1:

The data extracted from Wikipedia is only a list of neighborhoods along with it’s postal codes, which as it is does not mean anything. So we need to get the Geo coordinates of neighborhood utilising Geocoder package.

Geocoder Package was not working as expected, hence changed my strategy and used CSV file provided by IBM team to match the coordinates of Toronto neighborhoods.

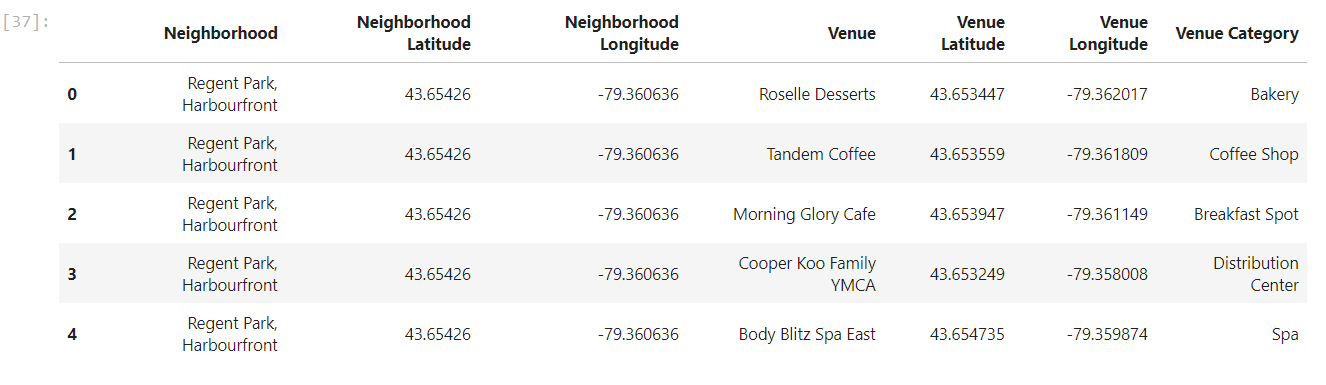
Then used the derived co-ordinates to visualize the map of Toronto neighborhood using Folium package to verify whether these are correct coordinates.



## 

## STEP 3: Use the neighborhood Geo Coordinates from step 2, to extract venues around these neighborhoods:

Next, used Foursquare API to pull the list of top 100 venues within 500 meters radius. Foursquare help us with very useful information such as names, categories, latitude, and longitude of the venues.



## STEP 4: Use of venue details from Step 3 for the analysis:

The venue details from step 3 are of great value for our analysis to find the best suitable place to open a new Restaurant in Toronto Downtown.

Used the Foursquare data:

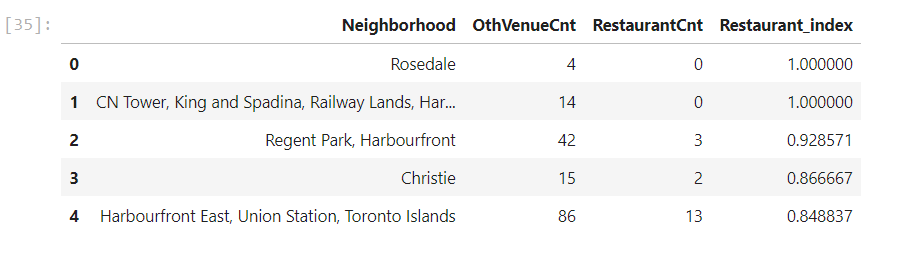
* + - To venue categorise into two further high-level categories. 1. Restaurants and 2. All other venues around each of the neighborhoods
    - All other venues (except Restaurant) was used to identify the most happening places around Toronto Downtown.
    - Restaurants data was used to find number of restaurants around neighborhoods, which will give us a feel of density of Restaurants that will help decide whether a specific neighborhood has a need for a new restaurant or there are already lot of them

## STEP 5: Find the best Neighborhood to open a new restaurant:

Based on the above analysis, we found out the most happening places (based on count of number of venues) around and it’s restaurant index (number of restaurants/number of venues) to finalise the best strategical location to open a new restaurant.

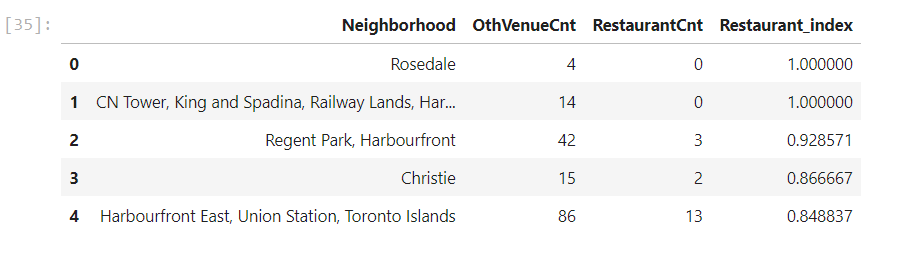
If there are may venues around, this means it’s a happening place in the Toronto downtown, so you can expect a lot of visitors and people around which is the first requirement for restaurant business. People normally go out to places with a lot options to spend time and when we are out, we love to enjoy good food with family and friends.

Once we identified the happening places, the next to thing to check is, how many restaurants around. If there are already many of them then does not make sense, so the best neighborhood to open a new restaurant is, neighborhood with most happening places around but smaller number of restaurants.



# RESULT DATA SET

**Given below the result data set (Top 5) for suitable neighborhood to open a New Restaurant**



# RECOMMENDATIONS

Based on the above Analysis, my recommendation for a New Restaurant Business would be "Regent Park" as there are many great happening places around but only 4 Restaurants in top 100 list. So strategically suitable neighborhood to start a new Restaurant Business