New Indian Restaurant

(The Battle of the Neighborhoods)

Capstone Project (Week 2)

# Table of contents

1. [Introduction: Business Problem](#_bookmark0)
2. [Data](#_bookmark1)
3. [Methodology](#_bookmark2)
4. [Results](#_bookmark3)
5. [Discussion](#_bookmark4)
6. [Conclusion](#_bookmark5)

# Introduction: Business Problem

## Background:

For many people in the world it is always a dream to have a business in America or Europe. When people look towards these destinations they have many questions in their mind.

## Problem:

In this project we will try to find a one of the best location in the world’s top two business centers (Toronto Vs San Francisco) for opening a new Indian restaurant. Choice of business city will depend upon many factors but here for opening a new Indian restaurant we will consider Only two of them 1. Size of Indian Community and 2. Immigration and Business Policies. Based on these two factors we will have a city of choice.

Once a business city is decided, we will try to find a best location for Indian restaurant in that city. In this process we will try to find pre existing Indian Restaurants in that city and look for a location where there is little ore no Indian restaurant restaurants available and it is close to the city center.

## Key Question:

So our key question is - How to find a best location for opening a Indian Restaurant in selected city?

## Target -

Target of this project will be Indian origin people who interested in opening a Indian restaurant in abroad like Toronto or San Francisco

We will use data science tools and Foursquare API to generate and process the neighborhoods information and location. Advantages of each area will then be clearly expressed for best possible final location for stakeholders.

# Data acquisition, cleaning & Analysis

Based on definition of our problem, factors that will influence our decision are:

* + Number of Indian community in Toronto and San Francisco
  + Immigration and Business Policies of USA and Canada
  + Number of any type existing restaurants in the neighborhood
  + Number of Indian Restaurant in the Neighborhood
  + Distance to Indian Restaurants in the neighborhood, if any
  + distance of neighborhood from city center

We decided to use circular area search using foursquare API with radius 5 KM Following data sources will be needed to extract/generate the required information:

***Data Source –***

1. Wikipedia Page - Indo Canadian Data<https://en.wikipedia.org/wiki/Indo-Canadians>
2. Indian Immigration to Canada - 1980 - 2013 data<https://ibm.box.com/shared/static/lw190pt9zpy5bd1ptyg2aw15awomz9pu.xlsx>
3. Canada M Postal Codes, borough, neighborhood https://en.wikipedia.org/wiki/List\_of\_postal\_codes\_of\_Canada:\_M
4. Canada M Postal Codes, latitude , longitude<http://cocl.us/Geospatial_data>
5. Number of restaurants, type, location and neighborhood - **Foursquare API**

[https://foursquare.com](https://foursquare.com/)

1. Foursquare categories of Restaurant (Indian)<https://developer.foursquare.com/docs/resources/categories>

***Data cleaning–***

1. Data downloaded or scraped from multiple sources were combined into dataframe(Table).
2. Find missing value
3. Try to replace missing value
4. Remove missing values rows(If can not be repaired or replaced)
5. Check for duplicate values
6. Remove duplicate values

***Data cleaning–***

1. Data returned by Foursquare contains large number of features but we have used only selected features –

name categories address distance

formattedAddress lat

lng id

1. We have also added some by splitting ‘categories’ features into like 'categoriesID', and 'categoriesName',
2. We have also added some features of multiple dat frame 'neighborhoodName', 'neighborhoodLatitude', 'neighborhoodLongitude'

***We have also try to save web scrapped and cleaned data into csv file for offline load and use***

***Toronto Vs San Francisco(City of Choice )***

Toronto and San Francisco are well developed and multi ethnic and multi culture business centers but Due to resent changes in USA business policies and H1-B visa and reforms, it has become a bit risky and fificult for outsider to open a new business.

So Toronto, Canada will be the first city of choice for opening a new Indian restaurant. It has very Open and favorable business policies and environment for outsiders.

\*\*\*Data Source – Resent News and Media Report

**Toronto**

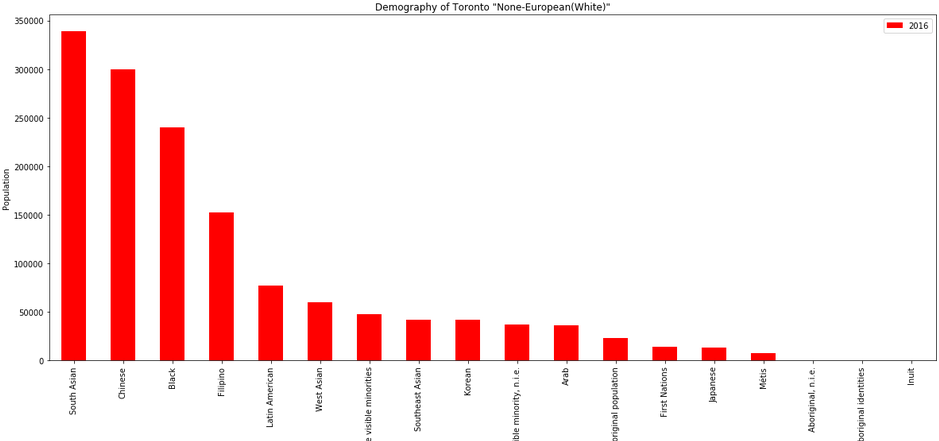
## Data Sample

1. **Percentage of Indian population in Canada**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **City** | **Province** | **Indian** | **Percentage** |
| **0** | Toronto | Ontario | 643370 | 10.40% |
| **1** | Greater Vancouver | British Columbia | 243135 | 10.00% |
| **2** | Montréal | Quebec | 48485 | 1.20% |
| **3** | Calgary | Alberta | 90625 | 6.50% |
| **4** | Edmonton | Alberta | 72245 | 5.50% |
| **5** | Ottawa | Ontario | 28945 | 2.20% |
| **6** | Winnipeg | Manitoba | 30800 | 4.00% |
| **7** | Hamilton | Ontario | 23390 | 3.10% |
| **8** | Victoria | British Columbia | 9180 | 2.50% |
| **9** | Kitchener | Ontario | 19295 | 3.70% |

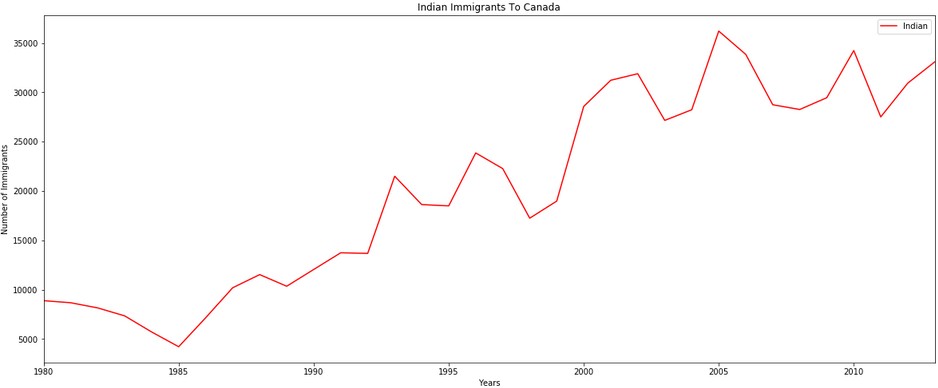
1. **Demographics of Toronto 2006 to 2016**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **2016** | **2011** | **2006** |
| **European (White)** | 1282750 | 1292365 | 1300330 |
| **South Asian** | 338965 | 317100 | 298370 |
| **Chinese** | 299460 | 278390 | 283075 |
| **Black** | 239850 | 218160 | 208555 |
| **Filipino** | 152715 | 132445 | 102555 |

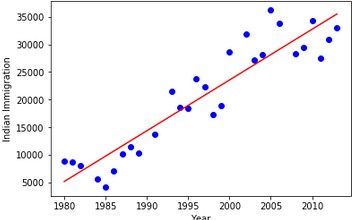


1. **Indian immigration to Canada (1980 to 2013)**

|  |  |
| --- | --- |
|  | **Indian** |
| **Year** |  |
| **1980** | 8880 |
| **1981** | 8670 |
| **1982** | 8147 |
| **1983** | 7338 |
| **1984** | 5704 |



1. **Regression Model of Indian Immigration to Canada**



Mean absolute error: 2698.25

Residual sum of squares (MSE): 11865109.78 R2-score: 0.80

## Indian Immigration prediction by 2025 (Canada)

|  |  |
| --- | --- |
| **Year** | **Indian** |
| 2020 | 42305 |
| 2021 | 43239 |
| 2022 | 44173 |
| 2023 | 45108 |
| 2024 | 46042 |
| 2025 | 46976 |

1. **Toronto Neighborhood and It’s Postal Code**

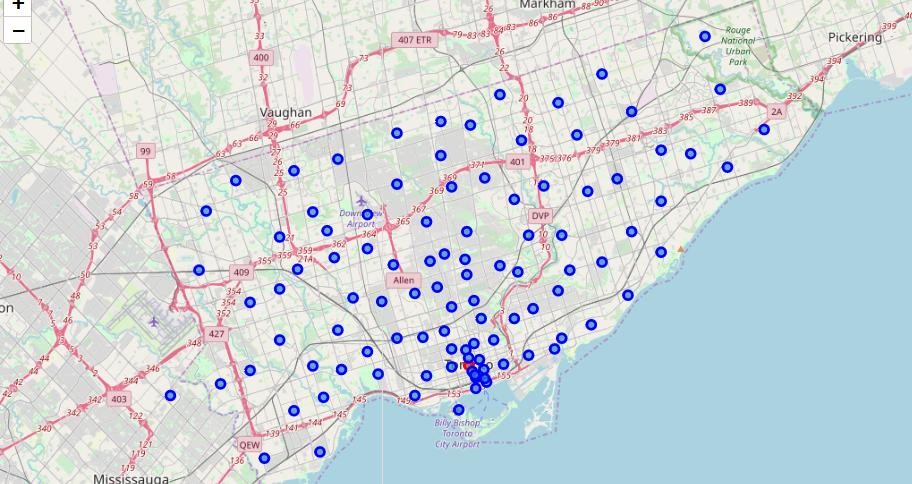
|  |  |  |
| --- | --- | --- |
| **Postal Code** | **Borough** | **Neighborhood** |
| M3A | North York | Parkwoods |
| M4A | North York | Victoria Village |
| M5A | Downtown Toronto | Harbourfront |
| M5A | Downtown Toronto | Regent Park |
| M6A | North York | Lawrence Heights |

1. **Toronto Postal code latitude and longitude**

|  |  |  |
| --- | --- | --- |
| **Postal Code** | **Latitude** | **Longitude** |
| M1B | 43.806686 | -79.194353 |
| M1C | 43.784535 | -79.160497 |
| M1E | 43.763573 | -79.188711 |
| M1G | 43.770992 | -79.216917 |
| M1H | 43.773136 | -79.239476 |

1. **Toronto Neighborhood with latitude and longitude**

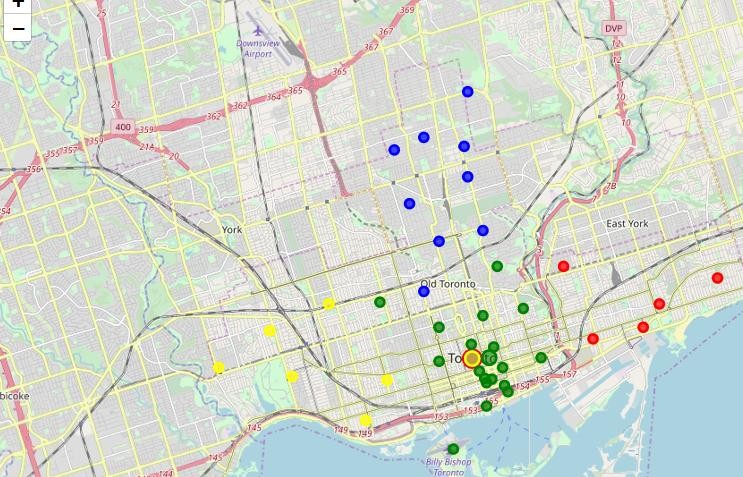
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Postal Code** | **Borough** | **Neighborhood** | **Latitude** | **Longitude** |
| M1B | Scarborough | Rouge, Malvern | 43.806686 | -79.19435 |
| M1C | Scarborough | Highland Creek, Rouge Hill, Port  Union | 43.784535 | -79.1605 |
| M1E | Scarborough | Guildwood, Morningside, West Hill | 43.763573 | -79.18871 |
| M1G | Scarborough | Woburn | 43.770992 | -79.21692 |
| M1H | Scarborough | Cedarbrae | 43.773136 | -79.23948 |



1. **Filtered borough that contains word “Toronto”**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Postal Code** | **Borough** | **Neighborhood** | **Latitude** | **Longitude** |
| M4E | East Toronto | The Beaches | 43.676357 | -79.29303 |
| M4K | East Toronto | The Danforth West, Riverdale | 43.679557 | -79.35219 |
| M4L | East Toronto | The Beaches West, India Bazaar | 43.668999 | -79.31557 |
| M4M | East Toronto | Studio District | 43.659526 | -79.34092 |
| M4N | Central Toronto | Lawrence Park | 43.72802 | -79.38879 |

1. **Clustering of Each Borough (Toronto)**



1. **Neighborhoods of borough Central Toronto**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Postal Code** | **Borough** | **Neighborhood** | **Latitude** | **Longitude** |
| M4N | Central Toronto | Lawrence Park | 43.72802 | -79.38879 |
| M4P | Central Toronto | Davisville North | 43.712751 | -79.3902 |
| M4R | Central Toronto | North Toronto West | 43.715383 | -79.40568 |
| M4S | Central Toronto | Davisville | 43.704324 | -79.38879 |
| M4T | Central Toronto | Moore Park, Summerhill East | 43.689574 | -79.38316 |



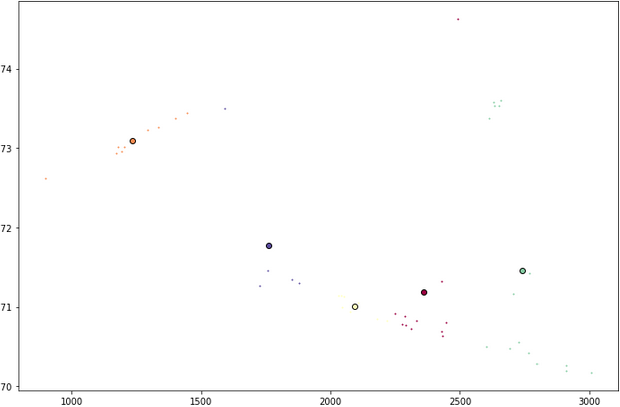
1. **Top 100 Restaurants that are in Central Toronto - Lawrence Park within a radius of 5 km meters**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **name** | **categories** | **location.addre ss** | **location.distan ce** | **id** |
| Indian Bread Bar | [{'id':  '54135bf5e4b08f3d2429dfd d', 'name': 'N... | 3305 Yonge St | 1179 | 54dd1704498e6bef053d2 cdf |
| Banjara Indian Cuisine | [{'id':  '4bf58dd8d48988d10f9417 35', 'name': 'I... | 164 Eglinton Ave E | 2278 | 4b7ccc72f964a520e3a52f e3 |
| Indian Street Food  Co. | [{'id':  '4bf58dd8d48988d10f9417 35', 'name': 'I... | 1701 Bayview | 2448 | 5650eed9498e376aac0c2 478 |
| Alleycat z Live Jazz Bar | [{'id':  '4bf58dd8d48988d1e79317 35', 'name': 'J... | 2409 Yonge St. | 2106 | 4ad4c05df964a52032f620 e3 |



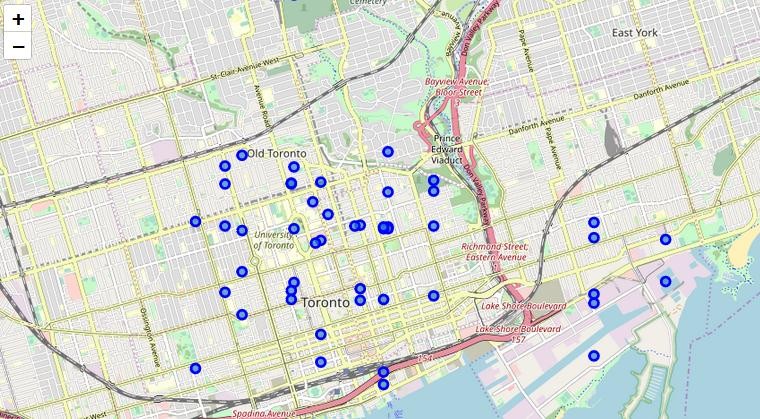
1. **K-Means Clustering**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **distance** | **lat** | **lng** |
| **Clus\_km** |  |  |  |
| **0** | 2093.75 | 43.710103 | -79.39558 |
| **1** | 2742.25 | 43.714588 | -79.396789 |
| **2** | 1235.666667 | 43.730912 | -79.403304 |
| **3** | 2361.454545 | 43.711813 | -79.39197 |
| **4** | 1760.8 | 43.71773 | -79.399427 |



|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **name** | **categories** | **address** | **distance** | **lat** | **lng** | **postalCode** | **state** | **id** |
| Aroma Fine Indian  Restaurant | [{'id':  '4bf58dd8d4898 8d10f941735',  'name': 'I... | 287 King St. W | 3194 | 43.64646 | -79.3896 | M5V 1J5 | ON | 4aef8854f96 |
| Aroma Fine Indian  Restaurant | [{'id':  '4bf58dd8d4898 8d10f941735',  'name': 'I... | 287 King St. W | 4525 | 43.64646 | -79.3896 | M5V 1J5 | ON | 4aef8854f96 |
| Aroma Fine Indian  Restaurant | [{'id':  '4bf58dd8d4898 8d10f941735',  'name': 'I... | 287 King St. W | 4827 | 43.64646 | -79.3896 | M5V 1J5 | ON | 4aef8854f96 |
| Aroma Fine Indian  Restaurant | [{'id':  '4bf58dd8d4898 8d10f941735',  'name': 'I... | 287 King St. W | 5884 | 43.64646 | -79.3896 | M5V 1J5 | ON | 4aef8854f96 |

1. **Indian Restaurant in Central Toronto Neighborhoods (Radius 5km)**



1. **Category ID for Indian Restaurant**

['54135bf5e4b08f3d2429dfe5','54135bf5e4b08f3d2429dff3','54135bf5e4b08f3d2429dff5', '54135bf5e4b08f3d2429dfe2','54135bf5e4b08f3d2429dff2','54135bf5e4b08f3d2429dfe1', '54135bf5e4b08f3d2429dfe3','54135bf5e4b08f3d2429dfe8','54135bf5e4b08f3d2429dfe9', '54135bf5e4b08f3d2429dfe6','54135bf5e4b08f3d2429dfdf','54135bf5e4b08f3d2429dfe4', '54135bf5e4b08f3d2429dfe7','54135bf5e4b08f3d2429dfea','54135bf5e4b08f3d2429dfeb', '54135bf5e4b08f3d2429dfed','54135bf5e4b08f3d2429dfee','54135bf5e4b08f3d2429dff4', '54135bf5e4b08f3d2429dfe0','54135bf5e4b08f3d2429dfdd','54135bf5e4b08f3d2429dff6', '54135bf5e4b08f3d2429dfef','54135bf5e4b08f3d2429dff0','54135bf5e4b08f3d2429dff1', '54135bf5e4b08f3d2429dfde','54135bf5e4b08f3d2429dfec','4bf58dd8d48988d10f941735']

# Methodology

In this project we have put our efforts to find a location in In Central Toronto where number of Indian restaurant are minimum. We will limit our analysis to area = 5km around each neighborhood based on postal code.

In first step we have collected the required information about Toronto demography and Indian immigration data to justify the selection of city for opening Indian restaurant.

In second step we have collected the required \*\*data: location and type (category) of postal codes of Toronto neighborhood.

In third step we collected the data related with Indian Restaurant of central Toronto neighborhoods within the radius of 5KM. We have identified and located Indian Restaurant in neighborhood as per Foursquare Categorization.

Fourth and final step includes analysis of Foursquare data '**No of Indian restaurant in neighborhoods**' for identifying locations with minimum number of Indian Restaurants

# Results

Let's perform some basic explanatory data analysis and derive some additional info from our raw data. First let's count the **number of Indian restaurants in every area neighborhood**:

**Due to selection of a bit larger radius in foursquare search, there are some overlaps in selected circular area and as result some restaurants selected in multiple neighborhoods of Central Toronto.**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **neighborhoodName** | **categoriesName** | **name** | **distance** | **lat** | **lng** | **id** |
| The Annex, North Midtown, Yorkville | Indian Restaurant | Aroma Fine Indian  Restaurant | 3194 | 43.64646 | - 79.39 | 4aef8854f964a5201cd921e3 |
| Deer Park, Forest Hill SE, Rathnelly,  South Hi... | Indian Restaurant | Aroma Fine Indian  Restaurant | 4525 | 43.64646 | - 79.39 | 4aef8854f964a5201cd921e3 |
| Moore Park, Summerhill East | Indian Restaurant | Aroma Fine Indian  Restaurant | 4827 | 43.64646 | - 79.39 | 4aef8854f964a5201cd921e3 |
| Forest Hill North, Forest Hill West | Indian Restaurant | Aroma Fine Indian  Restaurant | 5884 | 43.64646 | - 79.39 | 4aef8854f964a5201cd921e3 |

**We can overcome this problem by either reducing the search radius or remove restaurants reflected in multiple neighborhoods by using distance column**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **neighborhoodName** | **categoriesName** | **name** | **distance** | **lat** | **lng** | **id** |
| Davisville | Indian Restaurant | Banjara Indian Cuisine | 531 | 43.70781 | - 79.3933 | 4b7ccc72f964a52 |
| Davisville | Indian Restaurant | Eat Indian By Amaya | 1104 | 43.70559 | - 79.3752 | 4c51e0069d642d7 |
| Davisville | Indian Restaurant | Indian Street Food Co. | 1102 | 43.70803 | - 79.3761 | 5650eed9498e376 |
| Davisville | Indian Restaurant | Marigold Indian Bistro | 172 | 43.70288 | -79.388 | 5169d445e4b07d |
| Davisville North | Indian Restaurant | Mt Everest Restaurant | 1905 | 43.71327 | - 79.3665 | 504bcf32e4b0ef1 |
| Deer Park, Forest Hill SE, Rathnelly, South Hi... | Indian Restaurant | Chef of India | 525 | 43.68739 | - 79.3937 | 4b7c2a4ff964a520 |
| Forest Hill North, Forest Hill West | Indian Restaurant | Earth Indian Express | 969 | 43.70489 | - 79.4064 | 4ffb595be4b068d |
| Lawrence Park | North Indian Restaurant | Indian Bread Bar | 1179 | 43.73015 | - 79.4032 | 54dd1704498e6b |

0

e

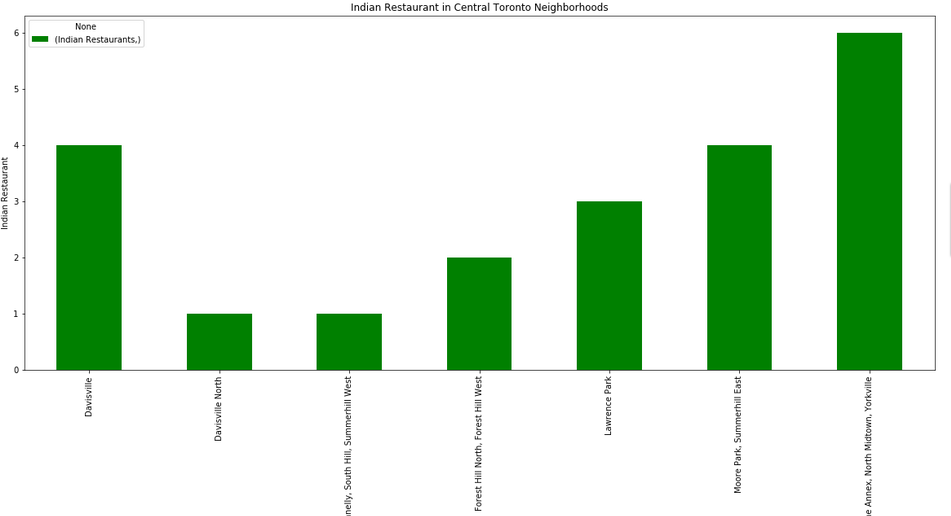
9

e

**\*Restaurant with minimum distance belongs to that neighborhood Indian Restaurant Group By Neighborhood(Grouping & Clustering)**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Lawrence Park | Indian Restaurant | Indian Crown | 4032 | 43.76008 | - 79.4122 | 5a3aa72566fc657 |
| Lawrence Park | Indian Restaurant | Patio Indian Restaurant | 4324 | 43.7221 | - 79.3357 | 59e94c0260255e6 |
| Moore Park, Summerhill East | Indian Restaurant | Indian Roti House | 2812 | 43.66438 | - 79.3805 | 547f94e9498e62f |
| Moore Park, Summerhill East | Indian Restaurant | Maja Indian Cuisine | 1984 | 43.67204 | - 79.3787 | 55f75919498ea27 |

|  |  |
| --- | --- |
|  | **Indian Restaurants** |
| **neighborhoodName** |  |
| **Davisville** | 4 |
| **Davisville North** | 1 |
| **Deer Park, Forest Hill SE, Rathnelly, South Hill, Summerhill West** | 1 |
| **Forest Hill North, Forest Hill West** | 1 |
| **Lawrence Park** | 3 |
| **Moore Park, Summerhill East** | 3 |
| **The Annex, North Midtown, Yorkville** | 5 |



# Discussion

In our analysis for opening a new Indian restaurant in world top business centers like Toronto in Canada or San Francisco in USA, we initially focused on deciding a city between Toronto and San Francisco for for Indian Restaurant.

Due to recent change in US polices related to visa and immigrants which might effect thousands of working people in US and there is an speculations about outflow of talent and business towards Canada. So choice was clear i.e Toronto.

initially after selection of Toronto we focused on analysis of its multi ethnicity. We looked and analyze the presence of Indian or South Asian community and try to predict the immigration trends of Indian community towards Canada by 2025. This gives us a picture of how a new Indian restaurant can grow by 2025. In this analysis we have considered South Indian or Indian community as our prime customers for new business. We found the data very encouraging because South Asian community was one of the largest community present in Toronto and immigration trends shows it will increase by time.

For location analysis we first try to find different neighborhoods of Toronto by using M Postal codes. We also grouped some of the neighborhoods having common postal codes and find their latitude and longitude for location. We found Toronto as big city with so many boroughs and their neighborhoods. So we again narrow our search areas by focusing on only boroughs that contains the word 'Toronto'. We further narrow our search area by selecting Central Toronto borough and its neighborhood as our prime search area.

Our analysis shows that large number of restaurants present in Central Toronto and its neighborhoods. Using Foursquare API search we looked for Indian Restaurant in each neighborhood within the radius of 5 KM. We have also used Foursquare Category for Indian Restaurant to filter the data. We find the density of Indian Restaurant by counting the number of Restaurants present in each neighborhood in Central Toronto borough.

During analysis we found that at radius of 5KM Foursquare API search area overlaps neighborhoods and this results in repeated names of Indian Restaurants but with different distance value. We used distance column with minimum value to fixed this problem and set the neighborhood of overlapping areas.

after analysis of Central Toronto Borough and its neighborhood for Indian Restaurant we found one of the neighborhood 'Davisville North' has only one Indian restaurant so this many be an ideal location for opening a new Indian Restaurant in Toronto. We also try to get ratings of its only Indian restaurant "Mt Everest Restaurant" which is not very good.

We found another candidate location for Indian Restaurant which has only one Indian restaurant "Deer Park..."

# Conclusion

Purpose of this project was to identify a location for opening a new Indian restaurants in order to aid stakeholders in narrowing down the search for optimal location. By analyzing restaurant distribution from Foursquare data we have first identified general boroughs that justify further analysis (Central Toronto), and then generated extensive collection of locations which satisfy some basic requirements regarding existing nearby Indian restaurants. Clustering of those locations was not requires because some of the neighborhoods has only few Indian restaurants present so nearby locations restaurants are good for opening a new Indian restaurant in Toronto

.This works as a starting point for final exploration by stakeholders.

Final decision on optimal restaurant location will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended zone, taking into consideration additional factors like attractiveness of each location (proximity to park or water), levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.