Report – Battle of Neighborhoods

19th June 2020

Author: Kanishk Singh

Business Problem

Problem

Where to open Indian restaurants (either new restaurant or a new branch of existing restaurant) to attract the maximum business and demand?

Context

Today, there is an influx of immigrants to Canada due to the encouraging migration policy adopted by the Canadian government. The highest number of immigrants have historically been from India, basis statistics explored in previous assignments of the IBM - Coursera program on Data Science. Further, Toronto being the financial and commercial capital of Canada attracts most of the Indian immigrants. Thus, there has been a huge a boost in the demand of Indian food in Toronto in the recent past.

Owners of Indian restaurants do not know where to open new branches or restaurants to attract the maximum customer business in Toronto.

Proposed Solution

This exercise will help in identifying areas in Toronto which are not saturated with Indian restaurants, and also have high population density, so that the new restaurants/branches can enjoy adequate demand for their offering.

Assumptions

1. If an area has over ~~10~~ 3 restaurants (number will be modified basis Indian restaurants in a radius of 1 sq. km), it is considered saturated.
2. Those areas with higher population density and lower saturation will be the best regions to open Indian restaurants.
3. Population distribution from 2016 data of Toronto from official website, indicates average of 27k people living in each area. Thus, targeting areas with >27k would be viable and help secure business for the restaurant from the higher population driving higher demand.
4. Given the increased demand of Indian food, mainly driven by Indian immigrants and partly by non-Indian customers, opening a restaurant away from saturated areas might provide access to white spaces among the Indian cuisine customers.
5. The location data provided by Foursquare is accurate and reliable.
6. The majority of the Indian restaurants have the below key-words in the title of the restaurant:

- India
- Punjab
- Biryani

Data

1. Foursquare location data will be leveraged to complete this assignment through consuming the API service available on the website.
2. Population data will be leveraged from the official source https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/hlt-fst/pd-pl/comprehensive.cfm to determine areas with lower population density which might not be suitable for opening Indian restaurants.
3. The API call will be made with below Python sample code:

```
import json, requests
url = 'https://api.foursquare.com/v2/venues/explore'
params = dict(
client_id='CLIENT_ID',
client_secret='CLIENT_SECRET',
v='20180323',
ll='40.7243,-74.0018',
query='coffee',
limit=1
)
resp = requests.get(url=url, params=params)
data = json.loads(resp.text)
```

4. Parameter details:
   - *client_id*: Obtained from my Foursquare account
   - *client_secret*: Obtained from my Foursquare account
   - *ll*: Latlongs would be mapped against postal codes for areas in Toronto. This data will be obtained from Wikipedia page and geocoder module of python
   - *query*: This would include various phrases commonly used in the titles of Indian restaurants listed below-
     - India
     - Punjab
     - Biryani
5. The sample response of the API:

```
{
 "meta": {
  "code": 200,
  "requestId": "5ac51d7e6a607143d811cecb"
 },
 "response": {
  "venues": [
   {
    "id": "5642aef9498e51025cf4a7a5",
```

```
      "name": "Mr. Purple",
      "location": {
        "address": "180 Orchard St",
        "crossStreet": "btwn Houston & Stanton St",
        "lat": 40.72173744277209,
        "lng": -73.98800687282996,
        "labeledLatLngs": [
          {
            "label": "display",
            "lat": 40.72173744277209,
            "lng": -73.98800687282996
          }
        ],
        "distance": 8,
        "postalCode": "10002",
        "cc": "US",
        "city": "New York",
        "state": "NY",
        "country": "United States",
        "formattedAddress": [
          "180 Orchard St (btwn Houston & Stanton St)",
          "New York, NY 10002",
          "United States"
        ]
      },
      "categories": [
        {
          "id": "4bf58dd8d48988d1d5941735",
          "name": "Hotel Bar",
          "pluralName": "Hotel Bars",
          "shortName": "Hotel Bar",
          "icon": {
            "prefix": "https://ss3.4sqi.net/img/categories_v2/travel/hotel_bar_",
            "suffix": ".png"
          },
          "primary": true
        }
      ],
      "venuePage": {
        "id": "150747252"
      }
    }
  ]
 }
}
```

Methodology

Three main data sources would be leveraged:

(1) Foursquare – for location of Indian restaurants in neighborhoods of Toronto
(2) Geocoder – python module for determining lat-longs
(3) StatsCanada – for population distribution in neighborhoods of Toronto (2016)

Through data wrangling we arrive at a dataset which would include:

(1) Postal codes
(2) Boroughs
(3) Neighborhoods
(4) Latitude
(5) Longitude
(6) Count of Indian Restaurants (in each postal code area)
(7) Saturation (y or n)
(8) Population (in each postal code, 2016)

Through data visualization and descriptions, we would define:

(1) Saturation - Areas with 4 or more Indian restaurants considered saturated
(2) Population density – Areas with >27000 population considered adequately dense regions

Results

Two main findings:

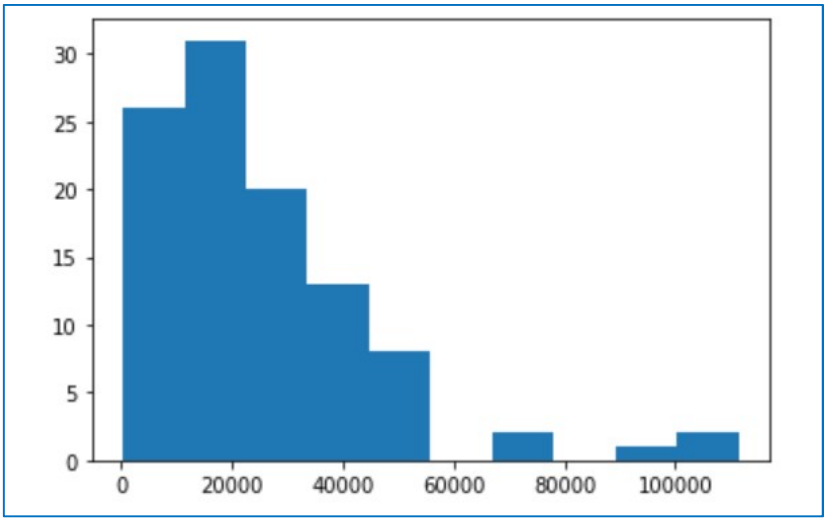
(1) Areas with >3 Indian restaurants within a 1km radius are saturated, so those with less than 4 restaurants should be considered for opening restaurants
Mean = 3.3 Indian restaurants per area

	Latitude	Longitude	Count_Indian_Restaurants	Population, 2016
count	102.000000	102.000000	102.000000	102.000000
mean	43.705271	-79.395009	3.343137	26785.676471
std	0.052286	0.095146	5.374113	15160.057881
min	43.602414	-79.594054	0.000000	0.000000
25%	43.661782	-79.462654	0.000000	16717.000000
50%	43.700636	-79.388790	1.000000	24866.500000
75%	43.745613	-79.340923	3.750000	36933.000000
max	43.836125	-79.160497	23.000000	75897.000000

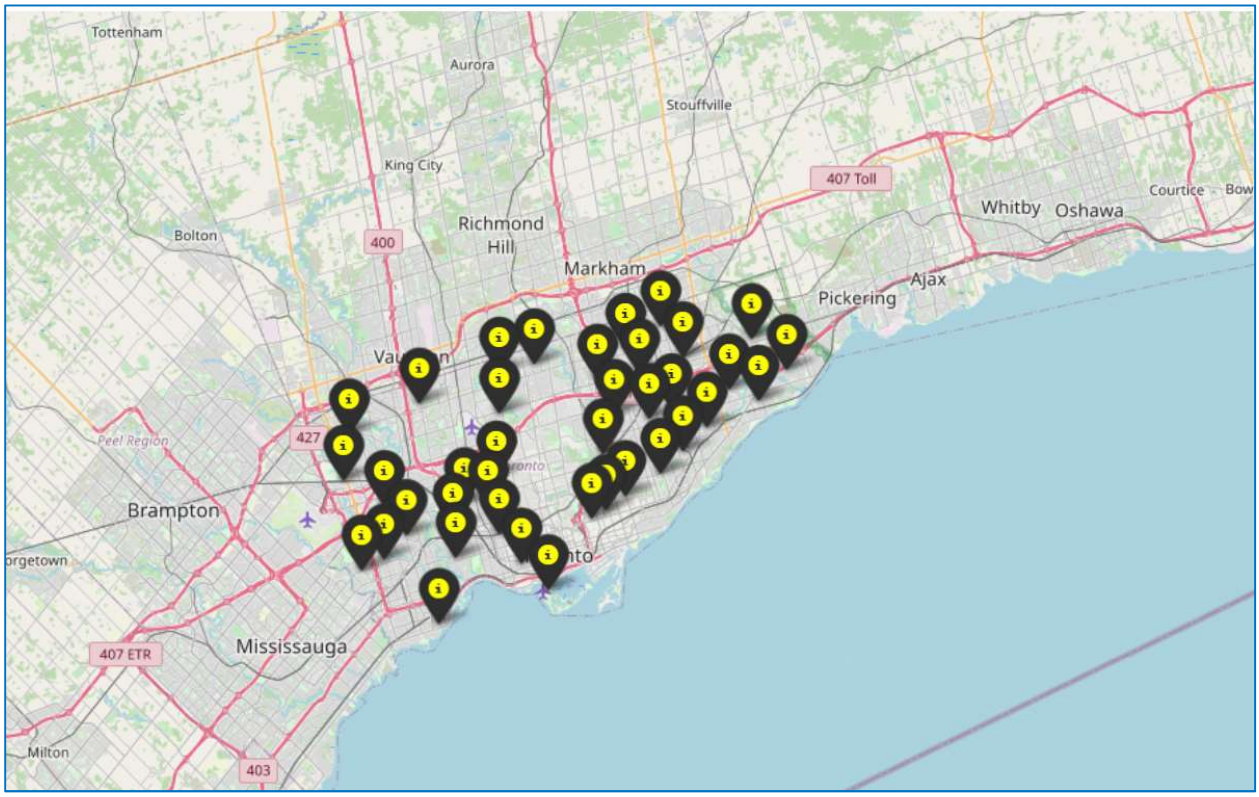
Postal Code	Borough	Neighborhood	Latitude	Longitude	Count_Indian_Restaurants	Saturation
M3A	North York	Parkwoods	43.753259	-79.329656	0	n
M4A	North York	Victoria Village	43.725882	-79.315572	3	n
M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636	5	y
M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763	0	n
M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494	12	y
...	...	...	...	...	...	...
M8X	Etobicoke	The Kingsway, Montgomery Road, Old Mill North	43.653654	-79.506944	2	n
M4Y	Downtown Toronto	Church and Wellesley	43.665860	-79.383160	9	y
M7Y	East Toronto	Business reply mail Processing Centre, South C...	43.662744	-79.321558	5	y
M8Y	Etobicoke	Old Mill South, King's Mill Park, Sunnylea, Hu...	43.636258	-79.498509	0	n
M8Z	Etobicoke	Mimico NW, The Queensway West, South of Bloor,...	43.628841	-79.520999	1	n

(2) Areas with <27000 people are below average population density, so those areas with more than 27,000 people should be considered to drive high demand for the Indian restaurant

	Latitude	Longitude	Count_Indian_Restaurants	Population, 2016
count	102.000000	102.000000	102.000000	102.000000
mean	43.705271	-79.395009	3.343137	26785.676471
std	0.052286	0.095146	5.374113	15160.057881
min	43.602414	-79.594054	0.000000	0.000000
25%	43.661782	-79.462654	0.000000	16717.000000
50%	43.700636	-79.388790	1.000000	24866.500000
75%	43.745613	-79.340923	3.750000	36933.000000
max	43.836125	-79.160497	23.000000	75897.000000



38 neighborhoods emerged as suitable for opening Indian restaurants due to lower competition from other Indian restaurants and higher demand from greater population density in their proximity. The areas are depicted on the map of Toronto, followed by their names below:



Discussion

For finding Indian restaurants in an area, initially 39 query phrases were chosen and there were 103 postal codes of Toronto in the dataset. Thus, total number of API calls would be ~4000, but the limit for the free developer profile per day is 950. So, the number of query phrases was reduced to 3. Due to unavailability of population data for 1 neighborhood, it was dropped from consideration, and 102 areas were finally analyzed.

Thus, there might be more Indian restaurants in the areas determined to be undersaturated, however, this bias is somewhat offset by the higher population density in these areas. The higher population would generate more demand and perhaps drive the business of any new Indian restaurants in the area.

While we have indicative data about where to open restaurants, we could refine this model further by identifying how ethnic groups are distributed across Toronto. Typically, there are areas like China Town, Little India etc. in most metropolitans in the world where people from populous countries have settled down for generations. It would be interesting to see if these areas are saturated or if there is scope for opening up new businesses catering to the populations.

Conclusion

Opening new Indian restaurants or expanding existing ones in the determined areas should be the starting point to be considered by business owners. This study/approach can be extended to almost all other popular cuisines like Chinese, Japanese, Mediterranean etc. This is especially important for the food and beverages industry which deals with perishables and where customer's preference is directly related to quickness of service and quality (partly constituted by freshness/recency of preparation). So, it is safe to deduce that proximity to the customer directly improves the quickness of service as well as freshness of meals, thereby improving the business demand.

By using this approach, and refining it for specific business objectives, depending upon richness and accuracy of the dataset, specific regions can be selected for best business outcomes for almost any new business or expansion.

Appendix

List of 38 areas suitable for opening Indian restaurants in Toronto.

Postal Code	Borough	Neighborhood
M3A	North York	Parkwoods
M9A	Etobicoke	Islington Avenue, Humber Valley Village
M1B	Scarborough	Malvern, Rouge
M6B	North York	Glencairn
M9B	Etobicoke	West Deane Park, Princess Gardens, Martin Grov...
M1C	Scarborough	Rouge Hill, Port Union, Highland Creek
M3C	North York	Don Mills
M4C	East York	Woodbine Heights
M9C	Etobicoke	Eringate, Bloordale Gardens, Old Burnhamthorpe...
M1E	Scarborough	Guildwood, Morningside, West Hill
M6E	York	Caledonia-Fairbanks
M1G	Scarborough	Woburn
M3H	North York	Bathurst Manor, Wilson Heights, Downsview North
M6H	West Toronto	Dufferin, Dovercourt Village
M1J	Scarborough	Scarborough Village

M2J	North York	Fairview, Henry Farm, Oriole
M4J	East York	East Toronto, Broadview North (Old East York)
M6J	West Toronto	Little Portugal, Trinity
M1K	Scarborough	Kennedy Park, Ionview, East Birchmount Park
M4K	East Toronto	The Danforth West, Riverdale
M1L	Scarborough	Golden Mile, Clairlea, Oakridge
M2M	North York	Willowdale, Newtonbrook
M6M	York	Del Ray, Mount Dennis, Keelsdale and Silverthorn
M3N	North York	Downsview
M6N	York	Runnymede, The Junction North
M1P	Scarborough	Dorset Park, Wexford Heights, Scarborough Town...
M1R	Scarborough	Wexford, Maryvale
M2R	North York	Willowdale, Willowdale West
M9R	Etobicoke	Kingsview Village, St. Phillips, Martin Grove ...
M1S	Scarborough	Agincourt
M6S	West Toronto	Runnymede, Swansea
M1T	Scarborough	Clarks Corners, Tam O'Shanter, Sullivan
M1V	Scarborough	Milliken, Agincourt North, Steeles East, L'Amo...
M5V	Downtown Toronto	CN Tower, King and Spadina, Railway Lands, Har...
M8V	Etobicoke	New Toronto, Mimico South, Humber Bay Shores
M9V	Etobicoke	South Steeles, Silverstone, Humbergate, Jamest...
M1W	Scarborough	Steeles West, L'Amoreaux West
M9W	Etobicoke	Northwest, West Humber - Clairville