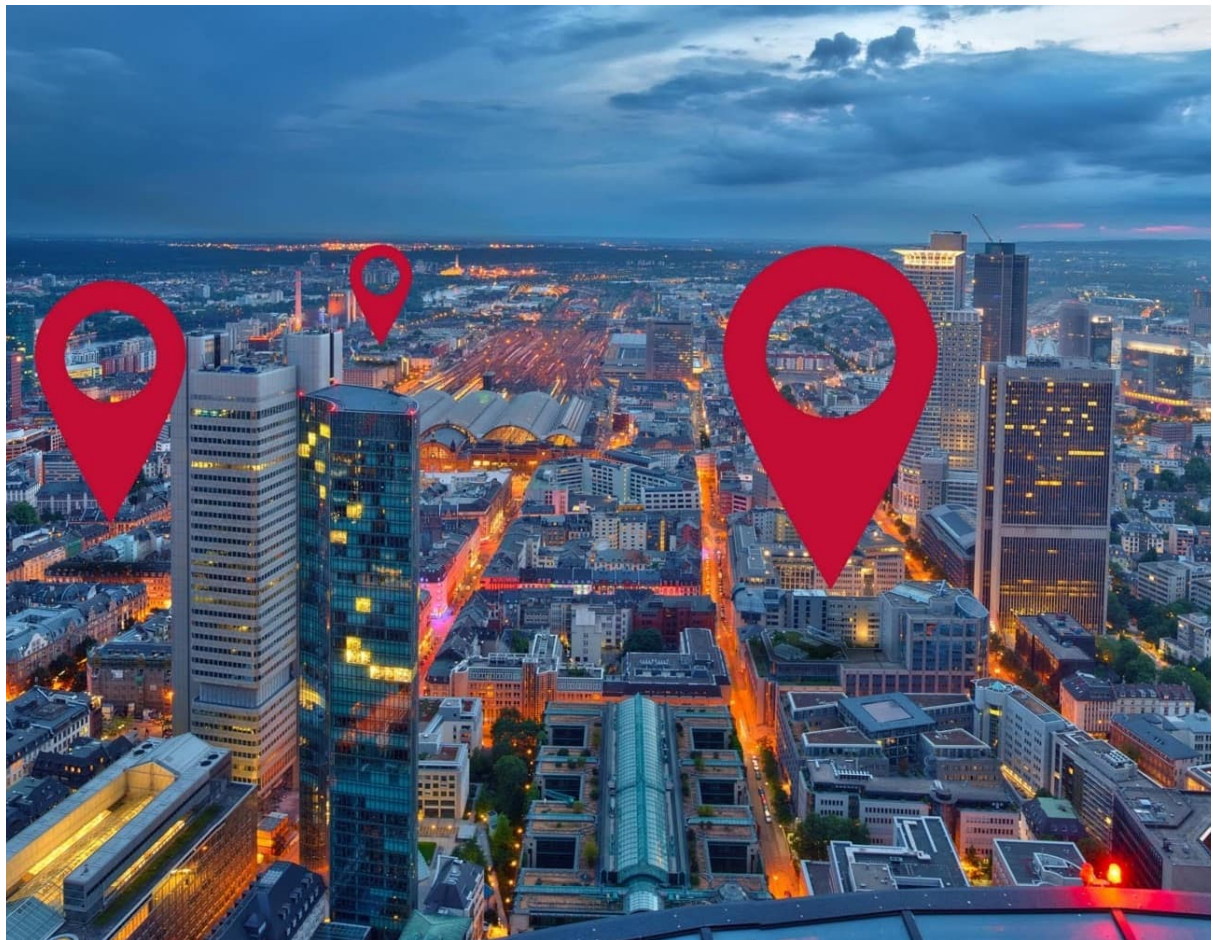# Finding the right location for your business

In this project we will try to find an optimal location for a restaurant but this method can be

used for any business which needs a physical location. Specifically, this report will be targeted to stakeholders interested in opening a **restaurant** in **Toronto**, Canada.

Since there are lots of restaurants in Toronto, 7500 to be precise. We will try to find locations that are not already saturated with restaurants. We are also particularly interested in areas with no Italian restaurants in vicinity. We would also prefer locations as close to city centre as possible, assuming that first two conditions are met.

Data

	Postal Code	Borough	Neighbourhood	Latitude	Longitude
0	M3A	North York	Parkwoods	43.75188	-79.33036
1	M4A	North York	Victoria Village	43.73042	-79.31282
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.65514	-79.36265
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.72321	-79.45141
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.66449	-79.39302
5	M9A	Etobicoke	Islington Avenue, Humber Valley Village	43.66277	-79.52831
6	M1B	Scarborough	Malvern, Rouge	43.81153	-79.19552
7	M3B	North York	Don Mills	43.74929	-79.36169
8	M4B	East York	Parkview Hill, Woodbine Gardens	43.70794	-79.31160
9	M5B	Downtown Toronto	Garden District, Ryerson	43.65736	-79.37818
10	M6B	North York	Glencairn	43.70799	-79.44838
11	M9B	Etobicoke	West Deane Park, Princess Gardens, Martin Grov...	43.65279	-79.55406

Toronto dataset will be used which we scrapped from wikipedia on third week. Dataset consisting of latitude and longitude, zip codes.

- Number of restaurants and their type and location in every neighbourhood will be obtained using Foursquare API.

- Coordinates of Toronto centre will be obtained using Google Maps API geocoding.

Foursquare api will be used to retrieve data of different neighbourhoods. Foursquare will give us venues around a neighbourhood their type and ratings which will be used to segment neighbourhoods and decide which will be the best option. The following data will be obtained from foursquare

- Neighbourhood
- Neighbourhood Latitude
- Neighbourhood Longitude
- Venue
- Name of the venue e.g. the name of a store or restaurant
- Venue Latitude
- Venue Longitude
- Venue Category

Methodology

After collecting the required data of every restaurant within 5km from the centre of Toronto. We will start our exploration of 'restaurant density' across different areas of Toronto - we will use heatmaps to identify a few promising areas close to centre with low number of restaurants in general and focus our attention on those areas. After finding the most suitable areas we will create clusters of locations that meet our requirements we will take into consideration locations with no more than three restaurants in radius of 250 meters. We will present map of all such locations but also create clusters (using k-means clustering) of those locations to identify general zones / neighbourhoods / addresses which should be a starting point for final 'street level' exploration and search for optimal venue location by stakeholders.

Results

TOP 5 Neighborhood of Toronto having max Japanese Restaurant
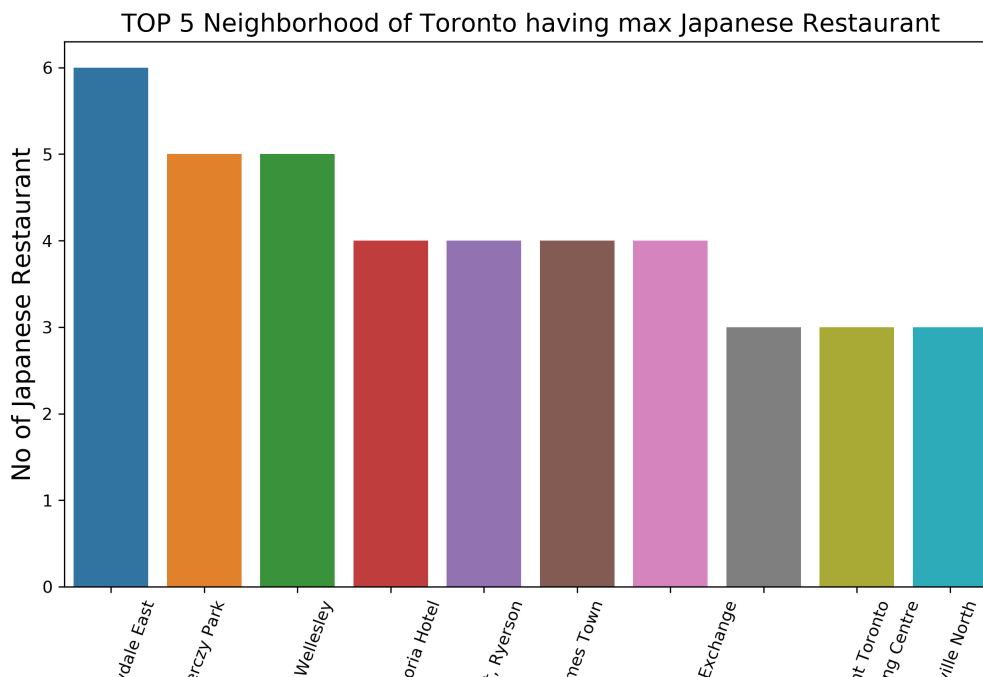
Our analysis shows that although there is a great number of restaurants in Toronto, there are pockets of low restaurant density fairly close to city center. Highest concentration of restaurants was detected downtown Toronto, so we focused our attention to areas around it. Those location candidates were then clustered to create zones of interest which contain greatest number of location candidates. Addresses of centre of those zones were also generated using reverse geocoding to be used as markers/starting points for more detailed local analysis based on other factors.

Result of all this is 15 zones containing largest number of potential new restaurant locations based on number of and distance to existing venues - both restaurants in general This, of course, does not imply that those zones are actually optimal locations for a new restaurant! Purpose of this analysis was to only provide info on areas close to Toronto center but not crowded with existing restaurants (particularly Japanese) - it is entirely possible that there is a very good reason for small number of restaurants in any of those areas, reasons which would make them unsuitable for a new restaurant regardless of lack of competition.

Discussion

Recommended zones therefore be considered only as a starting point for more detailed analysis which could eventually result in location which has not only no nearby competition but also other factors taken into account and all other relevant conditions met.

Conclusion

Finally to conclude this capstone project, I have got a small glimpse of how real life data-science projects look like. In this project I have imported different types of python libraries such as panda, numpy, matplotlib I have also used BeautifulSoup package to web scrape data.

I have also used Foursquare api to get the latitude and longitude data of Toronto City by Geopy Client. I have explored the different Borough, Neighborhood of Toronto city and analyse the data to get different outcome for Indian Restaurants of different parts of the city.

I have also used Machine learning technique K-Means clustering to cluster the neighborhoods and predicted a result which may help many business enthusiasts for opening Indian Restaurant in Toronto city where profit will be maximum and the demand is high.