

IBM Applied Data Science Capstone

Capstone Project

The Battle of Neighborhoods – Part II

Opening an Indian Restaurant in Toronto

By: Vaishali Ujlayan



Introduction:

Canada contains the world's eighth largest Indian diaspora. The highest concentrations of Indian Canadians are found in the provinces of Ontario and British Columbia, followed by growing communities in Alberta and Quebec as well, with the majority of them being foreign-born.

The demographics of Toronto, Ontario, Canada make Toronto one of the most multicultural and multiracial cities in the world. In 2016, 51.5% of the residents of the city proper belonged to a visible minority group, compared with 49.1% in 2011, and 13.6% in 1981. Toronto also has established ethnic neighbourhoods such as the multiple Chinatowns, Corso Italia, Little Italy, Little India, Greektown, Koreatown, Little Jamaica, Little Portugal and Roncesvalles, which celebrate the city's multiculturalism.

As people all over the world are liking Indian cuisine, so why not explore some of the best locations for Indian restaurants throughout Toronto. Potentially, the owner of the new Indian restaurant can have great success and huge profit. However, as for starting off a new business, one must be aware of the consequences if things are not implemented strategically or whether or not the business will be profitable. So, a few points should be considered for opening a new business. In our case, an Indian restaurant. Thus, the main idea behind for the attempt of this project is to answer a few questions such as -" Where should the investor invest to open an Indian restaurant?" and "Where should I go if I want to experience great Indian cuisines?"

Objective:

The main objective of this project is to analyse and select the best locations in Toronto to open a new Indian restaurant. Using Data Science methodology and tools such as data analysis and visualization, this project aims to provide solutions to answer the above-mentioned business requirements.

This project can be useful for the investors looking to open an Indian restaurant, targeting the multicultural population in Toronto.

Target Audience:

This project is particularly beneficial for the investors & constructors looking forward to invest in opening an Indian restaurant in Toronto. It is a great place to open an Indian restaurant providing ethnic Indian cuisines, because the city has a significant Indian population. And people from around the world travel to India to experience Indian flavours. So, why not provide the citizens of Toronto with such facilities i.e., by opening an Indian restaurant in foreign place, in their home country?

🚦 Data:

To start off with the above-mentioned problem statement, the following data will be required:

- Toronto city data having features as postal code, neighbourhoods, boroughs.
- Geo coordinates for the cities present in the above-mentioned data.
- Venue data for restaurants for further analysis and visualization.

🚦 Data description:

- Toronto city data will be fetched using web scraping from an open data source.
- Furthermore, feature engineering will be performed to keep the data in a particular format for easy data analysis and visualization.
- Geocoding of city data will be performed using geocoder library in Python.
- Venue data will be fetched using Foursquare API, as it provides the data for many categories. In our case, it is for Indian restaurants.

The project will require many data science skills from web scraping, working with API, data cleaning, data wrangling, data engineering to data visualisation.

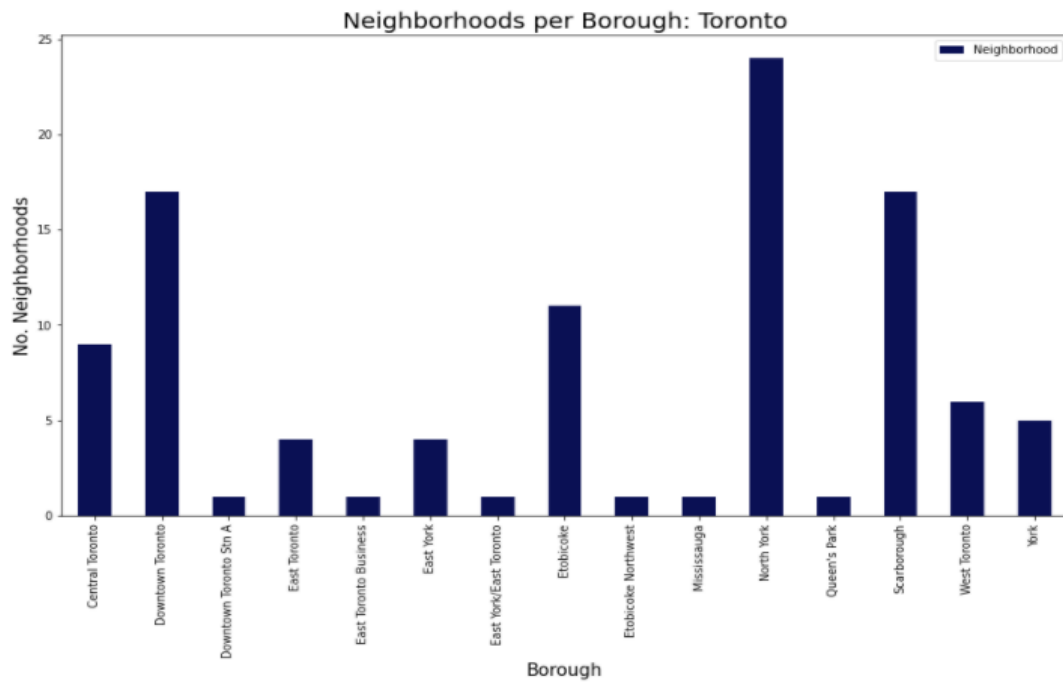
The next part of the project will elaborate other data science, statistical and machine learning techniques used.

🚦 Methodology:

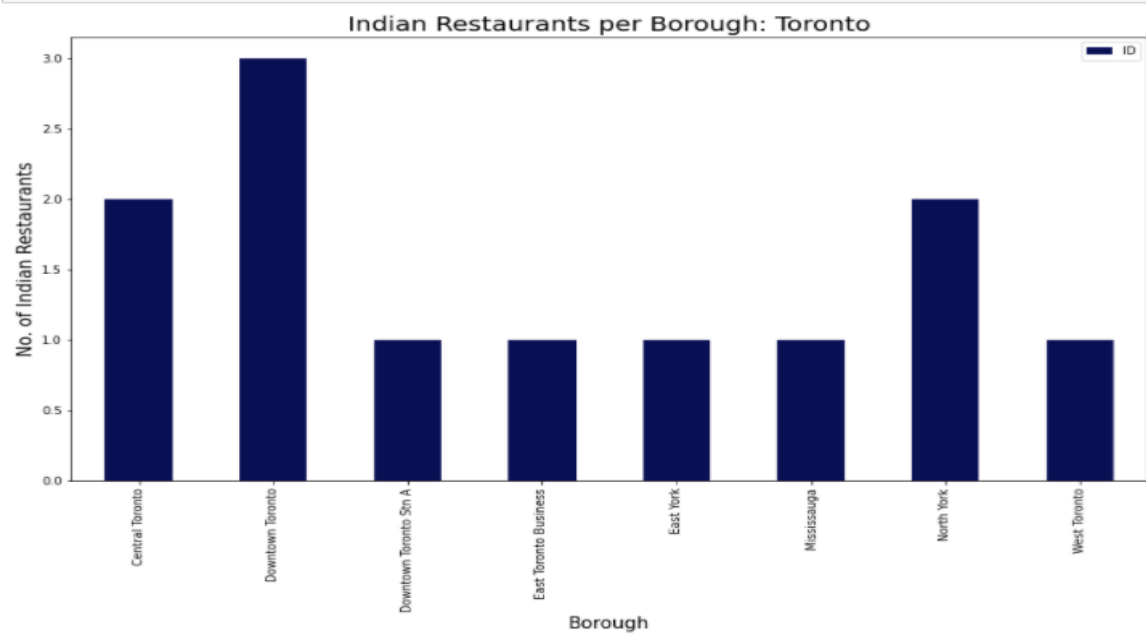
- Data will be collected from:
https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M and cleaned and processed into a dataframe.
- Foursquare be used to locate all venues and then filtered by Indian restaurants. Ratings, tips, and likes by users will be counted and added to the dataframe.
- Data will be sorted based on rankings.
- Finally, the data be will be visually assessed using graphs.

Results:

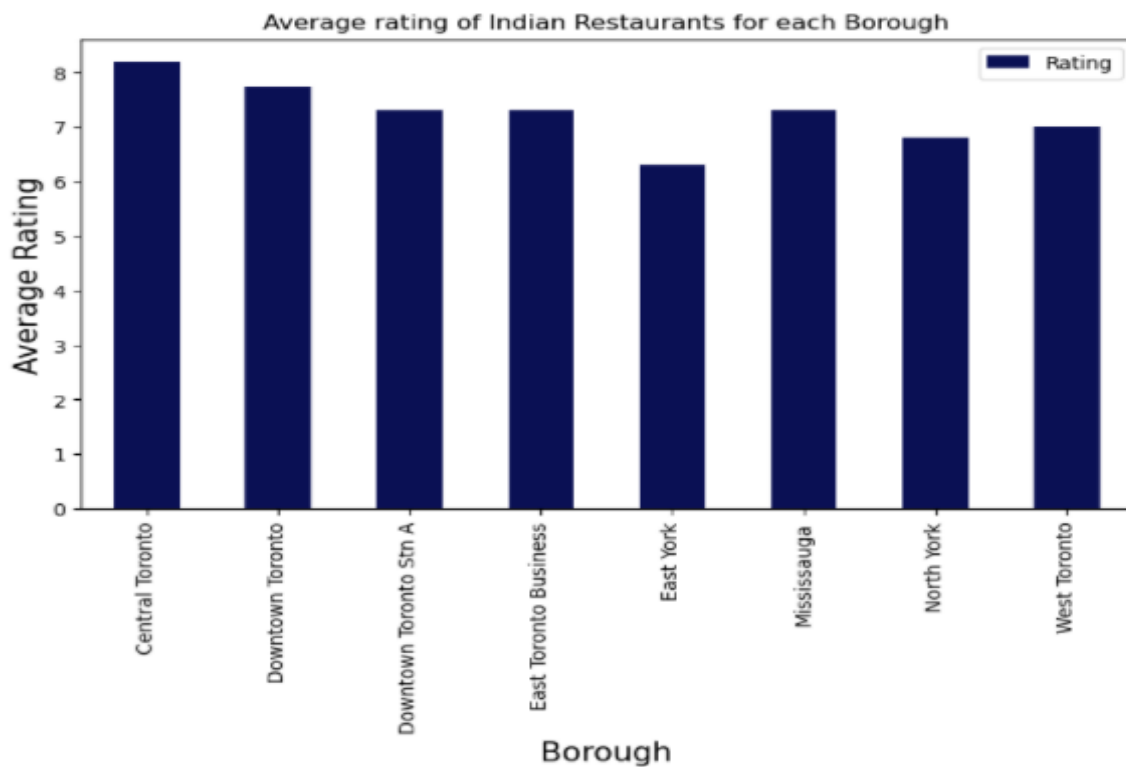
1. North York has the highest count of neighbourhood compared to others.



2. Downtown Toronto has the highest numbers of Indian restaurants i.e., 3 then Central Toronto and North York having 2 Indian restaurants respectively.



- Central Toronto has the Indian restaurant with high ratings than all the restaurants in other Boroughs.
Thus, I would recommend to visit Indian restaurants present in Central Toronto, then Downtown Toronto.



- Top 10 locations for Indian restaurants based on their ratings.

	Borough	Neighborhood	Latitude	Longitude	Average Rating
0	North York	Bedford Park, Lawrence Manor East	43.73545	-79.41916	7.3
1	Downtown Toronto	Berczy Park	43.64536	-79.37306	7.3
2	Downtown Toronto	Church and Wellesley	43.66659	-79.38133	8.0
3	Central Toronto	Davisville	43.70340	-79.38659	8.0
4	Mississauga	Enclave of L4W	43.64869	-79.38544	7.3
5	East Toronto Business	Enclave of M4L	43.64869	-79.38544	7.3
6	Downtown Toronto Stn A	Enclave of M5E	43.64869	-79.38544	7.3
7	West Toronto	Runnymede, Swansea	43.64982	-79.47548	7.0
8	Downtown Toronto	St. James Town, Cabbagetown	43.66788	-79.36649	7.9
9	Central Toronto	The Annex, North Midtown, Yorkville	43.67484	-79.40452	8.4

📌 Discussion Section:

Central Toronto and Downtown Toronto have the best rated Indian restaurants on average. West Toronto has the least number of Indian restaurants per borough. However, of note, the Annex, North Midtown, Yorkville in Central Toronto has the highest number of Indian Restaurants in all of Toronto. Despite Downtown Toronto having the least number of neighbourhoods in all the boroughs, it has the greatest number of Indian restaurants. Based on the above information, I would state that Central Toronto and Downtown Toronto are the best locations for Indian cuisine in Toronto. To have the best shot of success, I would open an Indian restaurant in Central Toronto. Central Toronto has multiple neighbourhoods with average ratings exceeding 7.0 on a scale of 1.0 to 10.0 and has a smaller number of Indian restaurants than Downtown Toronto, making competition easier. Also, we should keep in mind, that real estate prices in Brooklyn are much cheaper than in Manhattan. Finally, I would go to Little India Restaurant in Mississauga, Downtown Toronto Stn A or East Toronto Business for the best Indian food based on 127 likes. As a final note, all of the above analysis is depended on the accuracy of Foursquare data. A more comprehensive analysis and future work would need to incorporate data from other external databases.

📌 Limitations & future scope:

The major limitation of this project is that the data is fetched from open source so there might be chances of discrepancy and authenticity of the data. Also, third party API has been used for fetching venue details i.e., ratings, reviews, likes, venues etc. The major limitation here for using API is the number of calls made to get the desired results as free Sandbox Tier Account was used and furthermore, this limitation can be overcome by using paid account and incorporating external authentic databases for more features to be added in the data.

📌 Conclusions:

In this project, various data science methodologies were used such as:

- Identifying business problem
- Specifying the data required
- Extracting data from various sources and keeping data in a standard format
- Performing data analysis
- Providing answers to the pre-mentioned questions based on the analysis and techniques used for getting the answers.

Thus, the findings of the project will help the relevant investor better understand the advantages and disadvantages of different neighbourhoods/boroughs for opening the Indian restaurant.

