NEW INDIAN RESTAURANT IN TORONTO,CANADA

1. BUSINESS PROBLEM

In this project I am going to discuss what are the aspects that we should consider when we want to open a new restaurant . Here I would like discuss a particular case and discuss the factors to consider while opening a new Indian restaurant in Toronto, Canada. The methodology is applicable for opening of any restaurant in any location in the world provided we have adequate data to support our findings. Our problem here is the find the best location(neighbourhood) in Toronto where we can invest in a new Indian restaurant and get maximum benefit from our investment. Our target audience for this project is anyone with some interest in restaurant investment, Indian food and consider investment in Toronto, CA.

2. DATA

The most important and crucial section of the project is data collection. We need to determine the data we need and then try to find the data from different available sources. The data needed to solve given business problem is very specific to the particular problem.
In our case we need the name of the neighborhoods in Toronto and their respective geographic location to explore the nearby venues of the neighborhoods. We also need the population , annual income and demographics of the people living in each neighborhoods.

- **Data Sources:**
  The names of the neighborhoods are extracted via webscaping this wikipedia page(**https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M**).
  The geographic location of the neighborhoods are available here (http://cocl.us/Geospatial_data).
  The demographics (Total Population, Annual Family Income and Indian population) are obtained from the Toronto Open Data Portal (http://map.toronto.ca/wellbeing).
- **Data Cleaning:**
  The table on the wikipedia page has the useful data.

Postalcode	Borough	Neighborhood
M1A	Not assigned	Not assigned
M2A	Not assigned	Not assigned
M3A	North York	Parkwoods
M4A	North York	Victoria Village
M5A	Downtown Toronto	Regent Park, Harbourfront

The table is scrapped from the web and loaded into a pandas data frame for further processing. The table has 3 columns: 'Postal Code' , 'Borough' and 'Neighborhood'.

First in our data cleaning process, we ignore all rows with a 'Borough' that is 'Not Assigned'. If a row has a borough but a 'Not assigned' neighborhood, then the neighborhood has been assigned the same as the borough**.** Also more than one neighborhood exists for one Borough. These rows are combined into one row with the neighborhoods separated with a comma.

Postalcode	Borough	Neighborhood
M3A	North York	Parkwoods
M4A	North York	Victoria Village
M5A	Downtown Toronto	Regent Park, Harbourfront
M6A	North York	Lawrence Manor, Lawrence Heights
M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government

The geospatial data as shown below

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

and the demographics data are downloaded and saved from their respective website.

Neighborhood	Total Population	Average Family Income	Healthy Food Index	Indian
West Humber-Clairville	33312.0	72820.0	23.82	13920.0
Mount Olive-Silverstone-Jamestown	32954.0	57411.0	37.57	11095.0
Thistletown-Beaumond Heights	10360.0	70838.0	42.26	2555.0
Rexdale-Kipling	10529.0	69367.0	23.31	1420.0
Elms-Old Rexdale	9456.0	61196.0	24.71	840.0

Finally all the data are cleaned, merged and loaded for data analysis.

Postalcode	Borough	Neighborhood	Latitude	Longitude	Total Population	Average Family Income	Healthy Food Index	Indian
M2K	North York	Bayview Village	43.786947	-79.385975	21396.0	92800.0	32.12	1360.0
M2H	North York	Hillcrest Village	43.803762	-79.363452	16934.0	81484.0	48.46	1525.0
M9L	North York	Humber Summit	43.756303	-79.565963	12416.0	65757.0	32.33	3590.0
M6C	York	Humewood-Cedarvale	43.693781	-79.428191	14365.0	105770.0	32.85	320.0
M1J	Scarborough	Scarborough Village	43.744734	-79.239476	16724.0	61631.0	35.48	5520.0
M4E	East Toronto	The Beaches	43.676357	-79.293031	21567.0	139757.0	31.53	645.0