Opening a new Restaurant in Canada

Introduction

With over 60% of Canadian adults eating out at least 1-2x per week, the restaurant / food service industry is one of the fastest growing industries in Canada today. While restaurant trends typically come and go,it is doubtful Canadians will stop visiting eating establishments anytime soon. It is VERY PROFITABLE!, because most of the Canadians don't know what REAL Indian Food is! Most of the BEST Indian Food is sold on Indian Streets and Dhaba's and Canadians are pretty much hung up like British and they want to have sanitized foods like a hospital food and play it safe!

Business Problem

The objective of this Capstone Project is to find a location for your restaurant in Toronto,canada. We can use Data Science techniques i.e. "Machine Learning" algorithms such as clustering and different "tools"

Target Audience

To recommend the correct location. The objective is to locate and recommend which location of Toronto city will be best choice to start a restaurant. This would interest anyone who wants to start a new restaurant in Toronto city.

Data section

To solve this Problem, we need following Data:

1. List of Neighborhoods in Toronto, Canada.

2. Laitute and longitute of these Neighborhoods.

3. Venue Data Related to Indian Restaurants.

	PostalCode	Borough	Neighborhood
3	M3A	North York	Parkwoods
4	M4A	North York	Victoria Village
5	M5A	Downtown Toronto	Regent Park, Harbourfront
6	M6A	North York	Lawrence Manor, Lawrence Heights
7	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government

:

	Postal Code	Latitude	Longitude
0	M1B	43.806686	-79.194353
1	M1C	43.784535	-79.160497
2	M1E	43.763573	-79.188711
3	M1G	43.770992	-79.216917
4	M1H	43.773136	-79.239476

Extracting the Data

1. We will use web scrabing techniques to extract data from the wikipedia, with the help of python requests and BeautifulSoup packages.

2. Getting Laitute and longitute of these Neighborhoods via Geocoder Package.

3. Using Foursquare API to get venue data related to these Neighborhoods.

Methodology

>Business understanding

Our main goal is to find optimum location for a new Restaurant in Toronto, Canada

>Analytic Approach

First we need to get the list of neighborhoods in Toronto, canada. Then web scraping has been applied by using pandas HTML table scraping method. In this Project I have used different Packages, Modules and API. For Clustering, K-means Clustring is used.

K-means clustering is a simple unsupervised learning algorithm that is used to solve clustering problems. It follows a simple procedure of classifying a given data set into a number of clusters, defined by the letter "k," which is fixed beforehand. The clusters are then positioned as points and all observations or

data points are associated with the nearest cluster, computed, adjusted and then the process starts over using the new adjustments until a desired result is reached.

I have clustered the neighborhoods in Toronto into 3 clusters based on their frequency of occurance for Indian food.

Result & Recommendations

Clusters:

Most of the Indian restaurants are in cluster 1 which is around Central Bay Street, Church and Wellesley, Berczy Park, Union Station, Richmond. And lowest in Cluster 0 areas which are in North Toronto West and Parkade areas Looking at nearby venues it seems cluster 2 might be a good location as there are not a lot of Indian restaurants in these areas. Therefore, this project recommends the entrepreneur to open an authentic Indian restaurant in these locations.