Lack of Indian Restaurants

Abijith Pradeep

Introduction

Knowing the lack of Indian Foods and Cuisine in the now known Toronto region, I would like to propose a project to help resolve this issue and help Indian Cuisine grow. This can be done by opening new Indian restaurants in the area.

Business Problem

Now arises a business problem which can or cannot be tackled with data science. But what is that problem?
The problem is what location in Toronto region to choose that can be best for the restaurant to be opened. Now an entrepreneur would definitely like to take up the task to set up this new restaurant as it could boom his business. We can give a small help to the entrepreneur by using data science tools such as machine learning's famous clustering algorithm. The where part of the restaurant is the question to be answered.

Data and its Extraction and Target Audience

To solve the problem, we will need the following data:

1. List of neighbourhoods in and around Toronto
2. Latitude and Longitude of these Neighbourhoods
3. Surrounding data regarding location that favours the setting up of an Indian Restaurant.

Extraction:

1. Wikipedia
2. PyPostalcode
3. Foursquare API

The target audience are the entrepreneurs and business experts willing to set up an Indian Restaurant.

Methodology

First, we need to get a list of neighbourhoods in Toronto, Canada. This is possible if we extract the data from https://en.wikipidea.org/wiki/List _of_postal_codes_of_Canada:_M

Using pypostal code and the Wikipedia packages we can accomplish the task of creating a data frame of postal codes of the entirety of Toronto region. Also, the csv file provided by the IBM team or Coursera didn't have all the postal codes required to be present. Next, using folium we can visualize maps and verify whether these coordinates are true or not. Next, we can make use of the Foursquare API to pull a limit of 100 venues within a 500 meters radius. Following the tutorials to creating a developer personal Foursquare account one can always create that account and generate their CLIENT-ID and CLIENT-SECRET. Then we can create a venues list containing all the venues along with which neighbourhood they belong to, latitudes, longitudes etc.

Finally, we can go on with our main approach of classification of the Indian restaurants present in Toronto region within 500 meters. K-means clustering is what I found apt to be

used and have clustered them into 5 different clusters. Although the size will be less due to the low presence of Indian restaurants throughout Toronto.

Results

The result shows us a certain order up to which an entrepreneur can decide where to set up an Indian restaurant.

First cluster 0, then cluster 3, 4, then cluster 1 and finally cluster 2, where all Indian restaurants are present.

Recommendations

It is best to set up the restaurants as said in the order above to face less business challenges.