IBM

Applied Data Science Capstone

Introduction

Considering the present scenario, many people around the globe love the flavor of the Indian spices. So, if I as a entrepreneur want to start a Indian restaurant in Toronto the best way out is to find a location with minimum Indian restaurants where most people will have Indian food. So, in general, Asian people are more likely to consume Indian food. Thus, In order to open a successful restaurant , I will have to locate the place which has maximum number of Asians.

Business Problem

The objective of this Capstone project is to find the best location for an entrepreneur to open an Indian restaurant in Toronto. So the main aim of this project is to provide the best possible localities to open an Indian Restaurant in Toronto using Data Science and Machine learning algorithms like k-means clustering.

Target Audience

New entrepreneurs who don't want to risk a chance of failure in order to open a Indian Restaurant.

Data Needed

List of nearby areas of the desired city here taking Toronto (as the data for this was easily available.)
Coordinates of theses areas.
Data related to the details of nearby Indian Restaurants.

Methodology

1. Extracting nearby places.

Link:https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

2. Web Scrapping using Pandas HTML table Scrapping to pull tables into data frams.
3. Coordinates using .csv file provided by IBM.

4. Folium package to visualize the map
5. Foursquare API to list neighbourhood.
6. K-means clustering to cluster the places accourding to the number of Indian Restaurants.

Data Extraction

Toronto neighbourhood- Wikipedia
Coordinates of these places - CSV by IBM
Foursquare API to get information about the venues.

Results and analysis

The areas are clustered according the number of Indian Restaurants in nearby areas. And these clusters are color coded as follows:
Cluster 0: Zero number of Indian restaurants
Cluster 1: Minimum number of Indian restaurants
Cluster 2: Maximum number of Indian restaurants

Conclusion

The clusters with the minimum number of Indian Restaurants were identified. So, the model recommends entrepreneurs some best localities to start an Indian restaurant.

Submission by- Niyati Sinha