Exploring the Neighborhoods in Madrid: Data Science for finding a place of a new hindu restaurant

Introduction

As a part of the final IBM Capstone Project, we get to know of what data scientists go through in real life. Objectives of the final assignments were to define a business problem, look for data in the web and use Foursquare location data to compare different districts within the city of Madrid to figure out which neighborhood is suitable for starting an hindu restaurant.

In this report, we discussed about where is the best place to start an hindu restaurant in Madrid.

Business Problem

The city of Madrid has a population of 6 millions of people approximately. During last decades, it has received thousands of immigrants from all places in the world. This fact, linked to the growing number of fans who loved this cuisine aimed us to provide solutions about where it would be the best places to set up a new hindu restaurant.

Target Audience

This project is particularly useful for entrepreneurs who are decided to start up a restaurant in the capital of Spain, Madrid. This project is timely as the number of international restaurants has been constantly growing during the last years.

Data Selection

To realize that, data from 2 different sources were used. Data about hindu restaurants in Madrid was taken from Foursquare tool and eltenedor.es website. It was collected information related to postal code, name of the restaurant, address, latitude, longitude and average price per night and it was integrated into a database which contains 90 observations about hindu restaurants from Madrid. For a better analysis, some data about the population of each neighborhood was collected from institutional Madrid website https://www.madrid.es/.

The final database which combines Foursquare location data and Madrid restaurants data will be used to develop our machine learning models and to cluster Madrid areas neighborhoods in order to provide the best places to start our restaurant based on several factors.

Results

We have obtained all the hindu restaurants that are situated in Madrid. We have applied a k-means technique in order to select the best neighborhoods. The results from the k-means clustering show that we can categorize the neighborhoods into 3 clusters based on the amount of hindu restaurants that already exist in Madrid:

- Cluster 0: Neighbourhoods with high number of hindu restaurants.

- Cluster 1: Neighbourhoods with moderate number of hindu restaurants.

- Cluster 2: Neighbourhoods with low number of hindu restaurants.

Conclusions

In this study, I analyzed the number of hindu restaurants in Madrid in order to find the best places to open one. Most restaurants are concentrated in the centre and in the business districts. Therefore they already have a wide offer of international restaurants. Ciudad Lineal and Hortaleza represent a good opportunity for any investor because there are little competition and have a lively nightlife and good communications.

Another aspect that should be taken into account is the average income of citizens who live in these neighborhoods. We have checked that the selected neighborhoods are middle-class.