

IBM Applied Data Science Capstone Project

“Analysis of a potetial neighbourhood where to open an spanish food place in Guadalajara, Mexico”

April 11, 2019

Hernández Jara Jorge

1. Introduction/Business Problem

The purpouse of this Project is to use the Foursquare API along with some machine learning techniques to retrieve location data and regional clustering of venue information to decide which might be the best neighbourhood in my home city, Guadalajara located at the state of Jalisco in Mexico, to open an authentic spanish food place.

Guadalajara is considered one of the largest cities in Mexico in territorial extension, with 803 km² and also constitutes third economic nucleus of the country. Therefore, this suggests that the city is very likely to be a very highly populated multicultural place in a short term.

Another argument that supports the idea of this project is that I, as a resident, can assure that there are very few spanish food places along city, the most common present cuesines are italian, chinise and of course mexican.

Due to the previously mentioned fact that Guadalajara is constantly growing, migration from foreign countries are proportionaly growing. Several spanish commuties are moving on to the city so they are naturally considered the target audience because they can provide the authentic flavors of their home country cuisine

The hypothesis and success criteria for this project is that base on the non ocurrence of spanish places, the population and the economics indicators of an specific neighbourhoods the machine learning algorithm can suggest potential places where to open a restaurant of this kind.

2. Data description

The data that is going to be used come from a governamental open source of data sets that are provided by the Institute of Statistical and Geographical Information of the State of Jalisco which is the public organization responsible for understanding the state situation and dynamic reality, with the aim of offering a guide for the decision-making of economic and social actors; is a recognized frame of reference for objective evaluation of public interventions and changes that interest society.

The data set contains sociodemographic information by neighbourhoods that belong to the Metropolitan Area of Guadalajara, Jalisco [1]. Although the data is not up-to-date, since they date from 2010, are still valuable due the fact that only the neighbourhoods,

population, and features related with the economical status are of interest such as the population that is economically active and unactive and also the quantity of people that are and are not employed.

References

[1][https://datos.jalisco.gob.mx/dataset/informacion-sociodemografica-por-colonia-area-metropolitana-de-guadalajara-jalisco/resource#{}\]](https://datos.jalisco.gob.mx/dataset/informacion-sociodemografica-por-colonia-area-metropolitana-de-guadalajara-jalisco/resource#{})