

Opening a new restaurant in Barcelona

Sala, Albert
Coursera Capstone
6/25/20

Introduction

In Barcelona, food is a lifestyle. Traditional and avant-garde delicacies are put together in an outstanding expression of Mediterranean cuisine and eating out is both a social need and a custom deeply rooted in local culture. From tiny and crowded tapas bars to stunning fine dining venues are spread all over the Condal city territory.

However trendy spots, streets or even neighbourhoods usually change all over the time, and the public for these areas is also different. Barcelona counts with locals, expats and with uncountable tourists all year round, and the preferences for each group are significantly different. This makes the decision on opening a restaurant, a complex project where the type of cuisine, attention, price and targeted public could define the likely location where the business would flourish.

Business Problem

The objective for this project is to analyse with Data Science and Machine Learning the current market of restaurants in Barcelona by small areas within the city, to provide a quantifiable detail rather than a recommendation from a friend telling that one zone is popping. Besides the zonal business analysis, the society will also be considered in the study. Barcelona is an amazing walking city, so people usually wander around their living zone and end up in a restaurant, so that leads to a question: Is the background and income of the society influencing the restaurant industry? How?

Target Audience of this project

Any entrepreneur who is planning to open a restaurant would be the first ones interested in the findings of this project, but they are not the only ones. The insights can be used for opening any business, from ice cream shops to jewellery businesses. The behaviour of the society, together with their income, nationality and age can define the prosperity for the zone studied, and this is universally applicable.

Data

To solve this project, I will count with different sources of data:

Data sets from the Open Data BCN portal, the Ajuntament de Barcelona's open data service

Open Data BCN, a project that was born in 2010, implementing the portal in 2011, has evolved and is now part of the Barcelona Ciutat Digital strategy, fostering a pluralistic digital economy and developing a new model of urban innovation based on the transformation and digital innovation of the public sector and the implication among companies, administrations, the academic world, organizations, communities and people, with a clear public and citizen leadership.

From this source I will use:

- population.csv containing

Year	District.Code	District.Name	Neighborhood.Code	Neighborhood.Name	Gender	Age	Number
------	---------------	---------------	-------------------	-------------------	--------	-----	--------

- nationalitynsex.csv containing

Year	District.Code	District.Name	Neighborhood.Code	Neighborhood.Name	Gender	Nationality	Age	Number
------	---------------	---------------	-------------------	-------------------	--------	-------------	-----	--------

Foursquare API

After that, I will use Foursquare API to get the venue data for those neighbourhoods. Foursquare has one of the largest database of 105+ million places and is used by over 125,000 developers. Foursquare API will provide many categories of the venue data, I am particularly interested in the Restaurant categories.

Geopy

Geopy's geocoding web services will be used to obtain the longitude and latitude of the different neighbourhoods analysed.

Methodology

This is a project that will make use of many data science skills working with API (Foursquare), data cleaning, data wrangling, to machine learning (K-means clustering) and map visualization (Folium). In the next week, I will update this Methodology section where we will discuss the steps taken in this project, the data analysis that I did and the machine learning technique that was used in more detail.