



Applied Data Science Capstone

The Battle of Neighborhoods - Restaurants in Barcelona

Author: Adolfo López-Cerdán

Contact: adlpecer@gmail.com

Table of Contents

Introduction	2
Data	3

Introduction

Have you ever imagined setting up your own restaurant?

Opening a new and successful restaurant is probably a tough but exciting challenge. There are several points to consider to achieve it. One of the most important is where the new restaurant will be placed. Indeed, the physical location of the food-service business is a key point to get known, popular and to set the best possible price range.

In this project, we will work with restaurant data from Barcelona (Spain) to determine which neighborhoods are more convenient to start this kind of business.

Why Barcelona?

Barcelona is a coastal city in the northeast of Spain. It is the second most populated Spanish city, with a population of 1.6 million, and it heads one of the most populous metropolitan areas in the European Union. Besides, Barcelona is considered one of the world's main tourist, economic and cultural centers. Therefore, it is a trendy and cosmopolitan city, plenty of people from all over the world.

Such a variety of people and cultures makes Barcelona an inviting pool of potential customers but also a competitive market to start a business in the food-service sector. In that sense, this project could be helpful in the first steps of setting up a restaurant.

Business Problem

The objective of this project is to identify the most suitable neighborhoods in Barcelona to open a restaurant. To achieve this goal, data science methodology will be applied to Barcelona neighborhoods and venues data. Exploratory analysis of data and neighborhood clustering will be used to answer the business question: Which neighborhoods of Barcelona would be more advisable to consider if you want to maximize the popularity of your restaurant?

Target Audience

This project could be particularly interesting to:

- New entrepreneurs who want to start their own restaurant.
- Restaurant owners who are planning to move the location of their business.
- Investors looking for opportunities in the real estate sector.
- Food franchises looking for new locations.

Data

The following data has been used in this project:

Neighborhoods list

The list of neighborhoods in Barcelona has been extracted from the *Open Data BCN* service (<https://opendata-ajuntament.barcelona.cat/>), a platform managed by the Municipal Data Office. This data source contains a wide and open catalogue of datasets from several public administrations in Barcelona.

In this case, the dataset *Administrative units of the city of Barcelona* has been used. This dataset include a table with a total of 73 records that correspond to each neighborhood of the city with the name of the neighborhood, its ID and the borough to which it belongs (table 1).

Table 1: 5 first rows of *Administrative units of the city of Barcelona* dataframe.

	BoroughID	Borough	NeighborhoodID	Neighborhood
0	1	Ciutat Vella	1	el Raval
1	1	Ciutat Vella	2	el Barri Gòtic
2	1	Ciutat Vella	3	la Barceloneta
3	1	Ciutat Vella	4	Sant Pere, Santa Caterina i la Ribera
4	2	Eixample	5	el Fort Pienc

Neighborhoods boundaries

One of the visualization tools most used in this project is the choropleth map. This tool allows shading areas of a map in proportion to a numerical variable. However, the geospatial data of the neighborhood's limits is required to plot them on a map.

To achieve this goal, a *GeoJSON* file containing those limits has been extracted from the GitHub repository *bcn-geodata* (<https://github.com/martgnz/bcn-geodata>), created by Martín González.

Foursquare data

All the restaurant data have been extracted from *Foursquare Database* by making the right calls to *Foursquare Places API*. Two kinds of data have been obtained in this way:

- Restaurant names, IDs and categories by making the regular “*GET explore*” call for each neighborhood.
- The number of likes, Rating and Price tier by making the premium “*GET details*” call for each restaurant.

All the data has been concatenated to get a data frame containing all restaurant data.