

Capstone Project - The Battle of Neighborhoods (Week 1)

1 A description of the problem and a discussion of the background

We are planning to build a luxury hotel in Barcelona, and we want to know which neighbourhood is the best.

Barcelona was the 20th-most-visited city in the world by international visitors and the fifth most visited city in Europe after London, Paris, Istanbul and Rome, with 5.5 million international visitors in 2011. By 2015, both Prague and Milan had more international visitors. With its Rambles, Barcelona is ranked the most popular city to visit in Spain.

Barcelona as internationally renowned a tourist destination, with numerous recreational areas, one of the best beaches in the world, mild and warm climate, historical monuments, including eight UNESCO World Heritage Sites, and developed tourist infrastructure.

Barcelona is divided into 10 districts. These are administrated by a councillor designated by the main city council, and each of them have some powers relating to issues such as urbanism or infrastructure in their area. The current division of the city into different districts was approved in 1984. In 2009, in Barcelona started using a new division of 73 neighbourhoods (the 10 districts are still in use), a division that was done for a better service from the City Council.

Some of these districts have a previous history as independent municipalities which were integrated into the city of Barcelona during the late 19th century and the first half of the 20th century, such as Sarrià, Les Corts, Sant Andreu de Palomar, Gràcia or Sant Martí de Provençals. However, other municipalities which are contiguous to Barcelona (such as L'Hospitalet de Llobregat, Badalona, Sant Adrià de Besòs or Montcada i Reixac) have remained separate towns to this day and are part of the much larger metropolitan area of Barcelona.

We want to use an index of the family income (RFD) and the quantity of the hotels by neighbourhood to create clusters and see which one fits better with our goal.

2 A description of the data and how it will be used to solve the problem.

For this project we are going to use different datasets:

- Barcelona's last published income index by Neighbourhood (2017)
 - Source: https://opendata-ajuntament.barcelona.cat/data/dataset/79bdf758-dae1-485b-800c-be9f8cfa9360/resource/e7206797-e57b-4ded-8c6c-62e9b4cb54f7/download/2017_distribucio_territorial_renda_familiar.csv
 - Description: RFD (Renta Familiar Disponible per càpita) is the amount of income available to resident families for consumption and savings, once the depreciation or consumption of fixed capital in family economic operations and direct taxes and contributions paid to Social Security have been deducted. Eurostat recommends the use of the RFD as the main regional economic aggregate.

- Barcelona Hotels Foursquare
 - Source: Foursquare API
 - Description: Using the Foursquare API we'll get all the hotels in Barcelona.
- Barcelona Hotels Open Data
 - Source: <https://opendata-ajuntament.barcelona.cat/data/dataset/88efe464-2bcd-4794-85b0-8b0bbfd9e4c0/resource/eced0fe8-9892-4926-b035-4fdc7328e31d/download>
 - Description: We've looked for another hotel's source because we felt Foursquare API wasn't complete enough. This data set includes an accurate description of every hotel in the city