

Capstone Project - The Battle of the Neighborhoods (Week 1)

Applied Data Science Capstone by IBM/Coursera

Table of contents

- [Introduction: Business Problem](#)
- [Data](#)

Introduction: Business Problem

In this project we will try to find an optimal location for a gym. Specifically, this report will be aimed at those interested in opening a gym specializing in Pilates in Paris, France.

Since there are many gyms in Paris, we will try to detect locations that are not already crowded with gyms.

We are also particularly interested in areas without Pilates gyms nearby.

We also prefer locations as close to the city center as possible, assuming that the first two conditions are met.

We will use our data science powers to generate some of the most promising neighborhoods based on this criterion.

The advantages of each area will then be clearly expressed so that the best possible final location can be chosen by those interested.

Data

Based on the definition of our problem, the factors that will influence our decision are

- number of gyms in the neighborhood (any type of gym)
- number and distance of Pilates gyms in the neighborhood, if any
- distance from the neighbourhood to the city centre

We decided to use a grid of regularly spaced places, centered in the center of the city, to define our neighborhoods.

The following data sources will be needed to extract/generate the required information:

- the centers of the candidate areas will be algorithmically generated and the approximate addresses of the centers of those areas will be obtained using the OpenCage Reverse geocoding API
- The number of gyms and their type and location in each neighborhood will be obtained using the API Foursquare.ç
- The coordinate of the center of Paris will be obtained using OpenCage Geocoding API from the known location of Paris (Champs-Élysées)