Coursera Capstone Project: Where to open the next Paris CrossFit center?

By Nicolas de Wergifosse

Introduction

Paris is well known to be an active city where a lot of workers are looking for sport activities and one of the last trend is the CrossFit. Promoted as both physical and competitive fitness sport elements this sport made more than 15.000 affiliated recently just in the US. The trend arrived in Europe these last 5 years and lot of new gyms are proposing the infrastructures for their members. While there is a recent increase in classic gyms all around in Paris, only few propose CrossFit trainings to their members and as someone who's truly believed that it's a full body working that might still attract a lot of new affiliated I am looking for the best location in Paris where to open a new center.

Data Required

For this project we need the following data :

- Paris neighborhoods with their postal code, the population density, the price per square meter and their latitude and longitude.

  Data source : Official government data

  Description : This data set contains the required information that will be used to explore various locality of Paris.

- Nearby places in each locality of Paris.

  Data source : Foursquare API

  Description : By using this API we will get all the CrossFit and fitness centers in each neighborhood called "Arrondissements".

Methodology

Import libraries

As done in previous Coursera labs, the first step was to import all the necessary libraries like numpy, panda, matplotlib and others that was required to further manipulate the data.

Cleaning up the data

The second step was to club all the prerequisite information into a single data frame. Therefore I collected information from official government data regarding commercial rental prices, population density in arrondissements, and geographical coordinates, and collated that into the Jupiter Notebook. To finally reformatted in a way that would allow further analysis as well as better understanding.

Collect data from Foursquare API

I called the Foursquare database using the dedicated API. Therefore I had to define a radius from the city center geographical coordinates and ran a query to extract information related to CrossFit and Fitness centers already opened in Paris Arrondissements. These numbers were normalized and based on the neighborhood where these centers were located I incremented the initial data frame.

Visualize on Folium

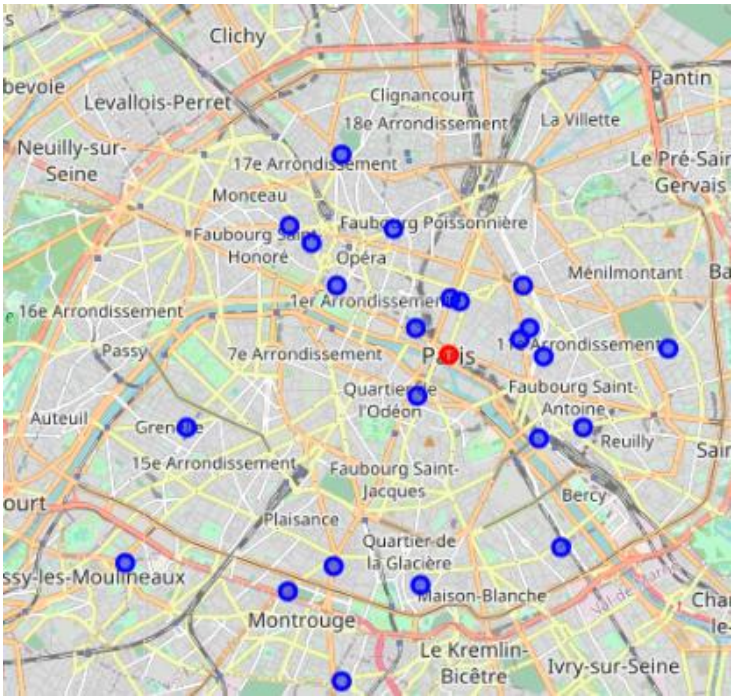
Based on incremented data from data frame I used Folium library to visualize the concentration of centers on the map being centered based on Paris latitude and longitude.

Other quantitative factors

On top of location information, we also ran histogram charts to easily compare the population density (in population per squared km) and the property prices (in EUR per squared meter). The conclusion from these graphs will be used in conjunction with the location information derived above in order to conclude which arrondissement will present the best business opportunity in terms of opening a new CrossFit center.

Results and discussion

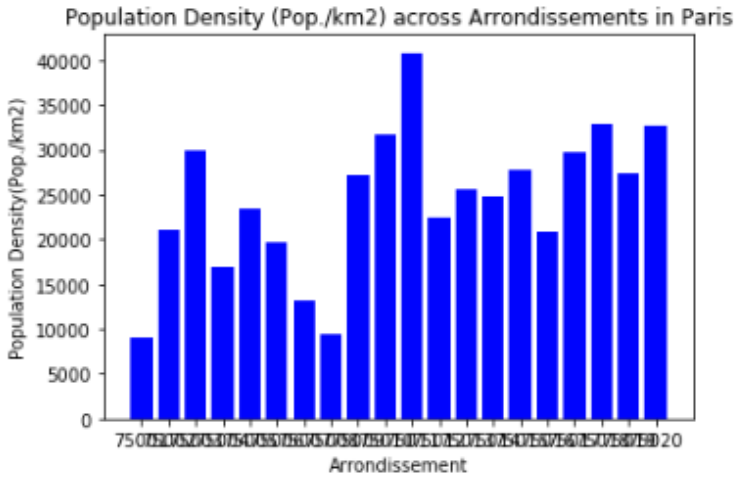
Displaying Foursquare results onto a Folium map and further by clustering the data we can observe that a first cluster of CrossFit centers is located in the city center along the Seine and a second one in the South of Paris:



It’s obvious that CrossFit center has open yet in the North of Paris and while extracting gym/fitness centers information from Foursquare I could see that there are well some gyms opened there in the North and the target clients being the same I see there an opportunity to open a new center in the 10,11,18,19,20th Arrondissements.

Taking a look at the population density of each arrondissement in Paris, we see that the 1st, 7th, and 8th arrondissements see the lower population densities while the 10th, 11th and 18th arrondissements have the top three highest population density. However, while it may be good to have a high population density, since it potentially represents higher customer flow, it is also important to evaluate the demographics of the population represented. For example, is the population more dense because the area is inhabited by more low-income population? This is where qualitative observations based on personal observations also come into picture.

Paired with my personal experience, it’s clear that even if the wealthiest people are located in the city center, a wave of wealthy people and young workers moved to the East and North of Paris these last 10 years because of more affordable property price (that will be analyzed later) and these are definitely target customers for our new CrossFit center.



In running a business, it is also important to keep costs low. One of the highest cost drivers in a city such as Paris is real estate. Taking a look at the commercial property price, we see that the first seven arrondissements see the highest prices, as they are located in the most central locations in Paris, while the 10th, 12th, 13th, 19th and 20th arrondissements see the lowest prices in EUR per meter squared.

Conclusion

By investigating location data, property prices, and population densities, it is possible to come to the conclusion that the 18th or 19th arrondissements would pose as the best location to open a CrossFit center. In particular, the 19th arrondissement serves as a very interesting opportunity because of the following reasons:

1) No other CrossFit center in the neighborhood.
2) Other gyms/fitness centers are successful in the area, meaning the target customers are there. Same principle as opening a Burger King next to a Mc Donald.
3) The population density is as high as in the South of Paris were the concept of CrossFit centers has already been a success.
4) Lower property price compared to surrounding Arrondissements. The migration wave to the North of Paris has started and better to be there in time before prices reach Paris regular property prices.
5) Proximity from multiple working area where big companies opened office centers for similar economic reasons (like BNP Real Estate who heavily invested in this area).