

What's the best place to open a CrossFit gym?

Joana Chong

November 4, 2019

1. Business introduction

1.1. Background

Cardiovascular diseases, type 2 diabetes, high blood pressure and high cholesterol rates have been increasing over the past decades and are some of the most frequent causes of death. As it's known, our daily life has become more and more sedentary, stressful, everything is getting easier to buy (i.e. home delivery services), the need to use our body to do what was meant to do – move – is decreasing and the development of tasty processed food is not encouraging nor helping us maintaining our physical health. However, there is a solution to reverse this situation: diet and exercise – and for this to be successful in the long run, is absolute key to do them both in a balanced way.

1.2. Problem

The reality previously described has spread out over the world and gyms/health centers are getting very popular and frequently attended. However, the monotonous conventional gym doesn't make lots of people happy to exercise and neither do the group dance classes ... And there is where CrossFit appears, which is a constantly varied training methodology with high-intensity functional movements.

1.3. Stakeholders

So, given the problem of people getting bored with conventional gyms, CrossFit gyms have been increasingly sought and this high demand constitutes a really good business opportunity with a promising potential. This data science project will be targeted to stakeholders interested in opening a CrossFit gym in Lisbon.

2. Data Sources and cleaning

2.1. Data Acquisition

Based on the business introduction, the factors that will influence the goal of finding the best place to open a CrossFit gym are: **(1)** Finding a place with a low number of CrossFit gyms/health centers, **(2)** determine the venues nearby the current gyms, **(3)** find a location with the highest number of venues and **(4)** be close to the city center.

To do so, the following data sources will be required: **Foursquare API** to find the number of CrossFit gyms/health centers, the most common venues nearby, the location with more venues nearby, and **geopy.geocoders** to convert the addresses into latitude and longitude values and vice-versa.

2.2. Data Cleaning

Using **geopy.geocoders**, Lisbon's latitude and longitude coordinates were retrieved and then the gyms nearby were found using **Foursquare API**. In order to visualize all those locations, a map was used (**Figure 1**).

To clean the retrieved data of the gyms' locations, only the 'Postalcode', 'Name', 'Latitude', 'Longitude' and 'Distance' columns were taken into account, since no missing values or heterogeneities were found (**Table 1**).

	Postalcode	Name	Latitude	Longitude	Distance
0	1250-145	Holmes Place Avenida da Liberdade	38.718227	-9.142977	1291
1	1000	Fitness Hut	38.743323	-9.139145	3966
2	1000-139	Fitness Park	38.734569	-9.143383	3043
3	1950-242	Crossfit Alvalade Oriente	38.742091	-9.102260	4848
4	1600	Solince Laranjeiras	38.754550	-9.166829	5833

Table 1: Head of the table of the retrieved gyms.

3. Methodology and Results

- A. All the gyms in a 6000m radius of Lisbon's city center mapped (**Figure 1**), the first step was to retrieve all the venues in a radius of 100m of each gym and count them. Fifty gyms were found.

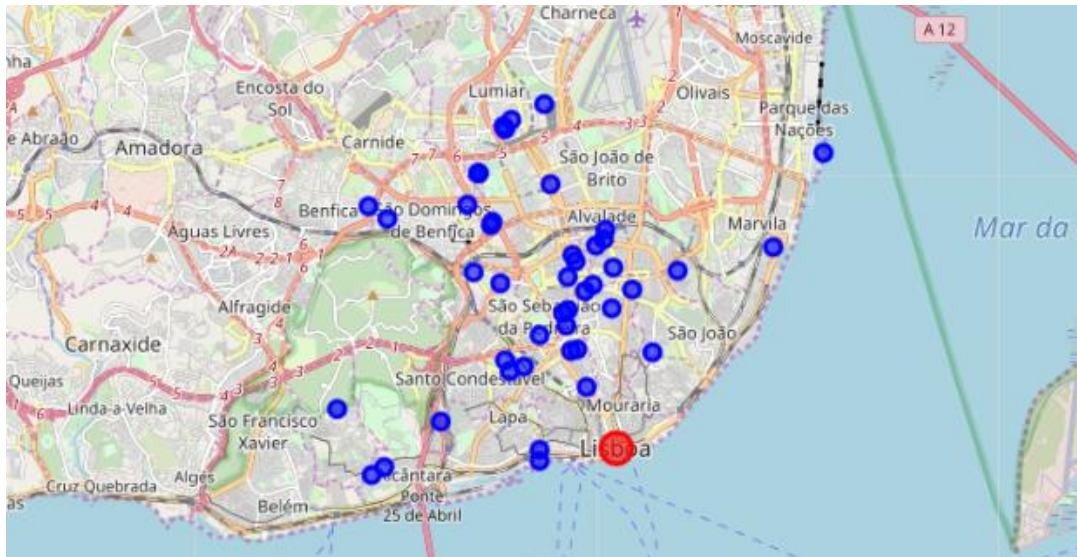


Figure 1: Representation of city centre and gyms nearby. The city centre is in red, while the gyms are in blue.

- B. The venues were grouped in a general group, i.e, all types of restaurants were counted as restaurants and all the hostels as hotels.

- C. The top 5 venues of each gym were found, as well as the top 5 most common venues among gyms (**Table 2**).

	Venue	Count
39	Restaurant	51.0
51	Train Station	23.0
20	Historic Site	20.0
14	Furniture / Home Store	19.0
9	Café	13.0

Table 2: Top 5 most common venues nearby gyms.

- D. A region of interest (ROI) of 18x28 equally spaced locations of potential addresses was calculated and mapped, corresponding to a total of 504 pairs of coordinates (**Figure 2**). The ROI was created with an attempt to avoid covering the ocean. It was not 100% doable, but in a further step it will be filtered and won't make a difference.

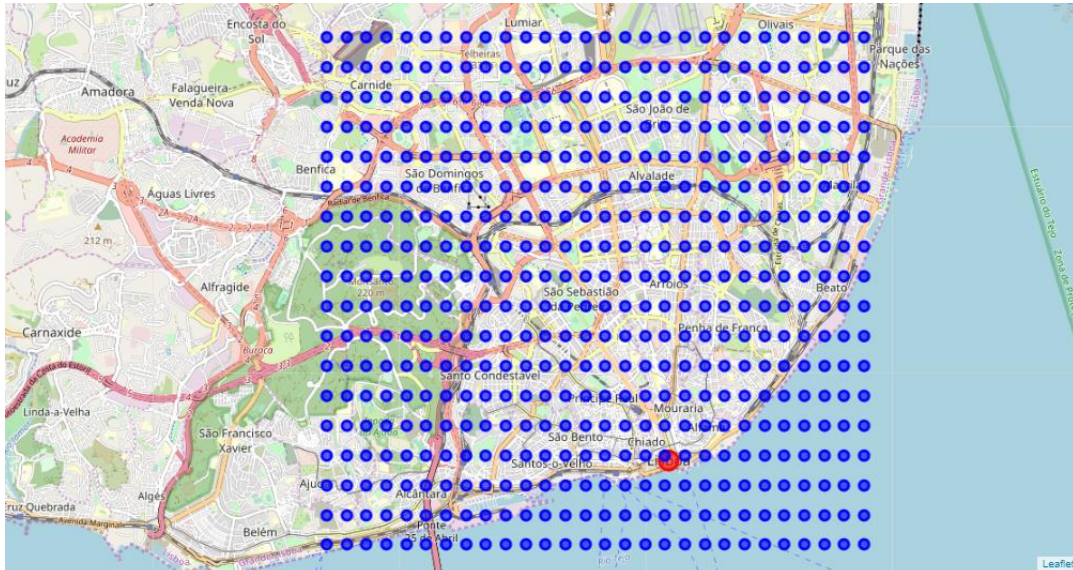


Figure 2: Representation of the ROI and the 504 pairs of coordinates of possible locations to open a gym.

- E. The most common venues around each location of the ROI (radius of 100m) were found and grouped/ counted by location ('Place') (**Table 3**).

	Place	Place Latitude	Place Longitude	Venues Count
0	127	38.712607	-9.144393	19
1	267	38.733327	-9.144393	18
2	129	38.712607	-9.137299	13
3	132	38.712607	-9.126658	13
4	183	38.720895	-9.144393	13

Table 3: Top 5 locations with more venues nearby.

- F. Only the ones with at least 1 venue were mapped in blue and the top 5 locations with more venues were mapped in yellow (Figure 3).

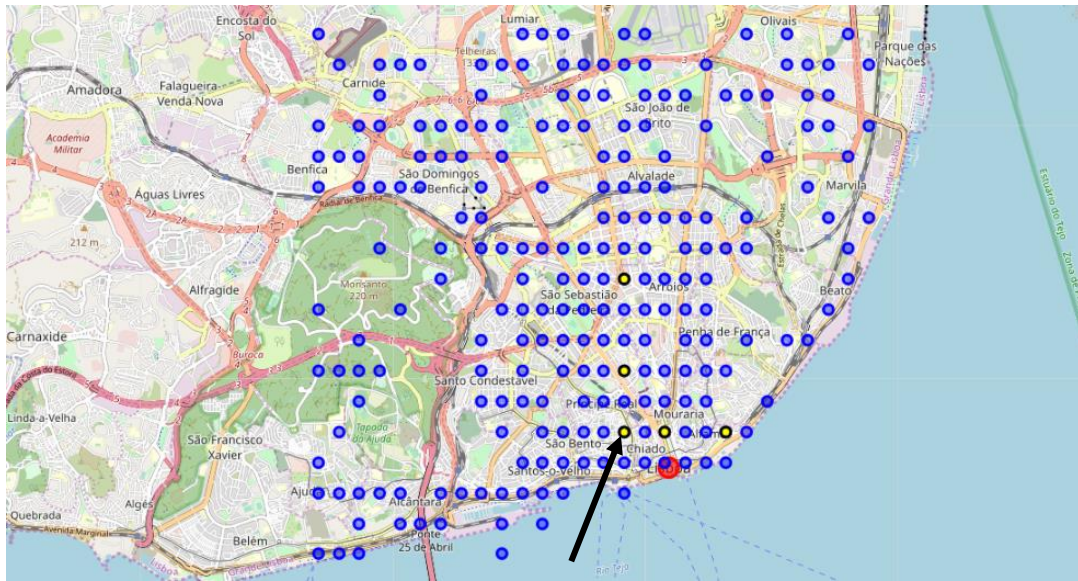


Figure 3: Representation of Lisbon centre in red, potential locations with at least 1 venue nearby in blue and the top 5 locations with more venues in a 100m radius in yellow.

- G. The addresses of the top 5 potential locations were retrieved using **geolocator.reverse** and the latitude and longitude coordinates. Using **Foursquares**, the number of gyms was retrieved (radius of 150m) and finally, the distance to the city centre was calculated. The output is represented down below:

- 1- **Address:** Bairro Alto, Misericórdia, Lisboa, 1200-050
Gyms nearby: 0
Distance to city centre: 866 m
- 2- **Address:** Atrium Saldanha, Saldanha, Lisboa, 1050-094
Gyms nearby: 2
Distance to city centre: 2924 m
- 3- **Address:** Rua de Santa Justa, Baixa, Lisboa, 1100-052
Gyms nearby: 1
Distance to city centre: 544 m
- 4- **Address:** Rua de Santo Estêvão, Alfama, Lisboa, 1100-615
Gyms nearby: 2
Distance to city centre: 1017 m
- 5- **Address:** Rua de Santa Marta, Santo António, Lisboa, 1150-255
Gyms nearby: 0
Distance to city centre: 1611 m

The first address is pointed with an arrow in **Figure 3**.

4. Discussion

Lisbon has a total of 50 gyms, that are mapped in **Figure 1**. Some of them are solo, others are close to each other, which is not ideal because they are business competitors of each other. The goal is to be as far as possible from another gym.

In **Table 2**, it's possible to see that restaurants are the most common venue nearby gyms, followed by train station, which makes sense. People often move around the city by public transports and have lunch/dinner out in a restaurant. So ideally, the new gym should be close to restaurants and public transport stations.

Table 3 shows the top 5 possible locations with more venues in a 150m radius. The first and second ones have almost the same number, which means that the number of venues is not enough to determine if it is the best place to open a gym. So, other factors have to be taken into account in order to separate them and have a clearer conclusion.

In **G.**, there is an output from the Jupyter Notebook that shows the gyms nearby each of the top 5 gyms, as well as their distance to city centre. As we can see, the first one (the one that has more venues nearby) is also the one without gyms in a 150m radius, when comparing to the second location with more venues nearby, and the second one closest to city centre (866m).

5. Conclusion

Given the problem of the increasing rates of cardiovascular diseases, one of the solutions is dieting and exercising, and people are more and more aware of it. With that said, people are looking for a gym to exercise. Therefore, opening a well-located CrossFit gym is key for a successful business.

In order to find the best location in Lisbon to open a CrossFit gym, this study was made and the best location is Bairro Alto because it has the highest number of venues nearby (venues = 19), it is the second closest to city center (866m) and it doesn't have any gym/CrossFit gym nearby. With that said, this location as the best factors to be successful.

Future developments may include finding the most relevant venues that people are looking for and creating a ROI more 'appropriate', perhaps considering cartesian coordinates and making each location 100m apart.