



CAPSTONE PROJECT

The Battle of Neighbourhoods Week 2

Capstone Project

The Battle of Neighbourhoods Week 2

List of Content

1. Requirements:	1
2. Problem definition:	2
3. Methodology followed	4
Initial location definition	4
Radius of the area assessed	5
Python necessary libraries	5
Foursquare API	5
Data processing and cleaning	6
Visualize the location	7
4. Results	8
5. Conclusions	9

1. Requirements:

A full report consisting of all of the following components:

1. Introduction where you discuss the business problem and who would be interested in this project.
2. Data where you describe the data that will be used to solve the problem and the source of the data.
3. Methodology section which represents the main component of the report where you discuss and describe any exploratory data analysis that you did, any inferential statistical testing that you performed, if any, and what machine learnings were used and why.
4. Results section where you discuss the results.
5. Discussion section where you discuss any observations you noted and any recommendations you can make based on the results.
6. Conclusion section where you conclude the report.

2. Problem definition:

The problem

The problem to be solved consists of opening of a gym business in the city of Valencia, Spain. More precisely, I look for a good location nearby the city centre of the city.

People usually go to the closest gym of their neighbourhoods so I will try to know if opening a gym in the city centre is a good business opportunity.

In reference to the data, I am going to use Foursquare location data to know where are the gyms located nearby the city centre of Valencia.

This will allow me to know if the city centre is a good place to start my gym business.

The context

Valencia is a city in Spain, capital of one of the 17 regions into which Spain is divided, the *Comunidad Valenciana*. It has a population of more than 800,000 inhabitants (2020), which rises to more than 1,500,000 inhabitants (2020) if its metropolitan area is included. This makes it the third most populated city and metropolitan area in Spain, only behind Madrid and Barcelona.



Figure 1. Geographical location of the City of Valencia

Geographically, the city is located on the banks of the Turia River, on the east coast of the Iberian Peninsula, right in the centre of the Gulf of Valencia. When the city was founded by the Romans, it was located on an island in the Turia River, about four kilometres from the sea. However, the city now extends to the sea facade, being the seafront one of the busiest areas of the city.

Its historical centre is one of the largest in Spain, with approximately 169 hectares and thanks to its historical and monumental heritage and its diverse scenic and cultural spaces, it is one of the cities with the greatest influx of national and international tourism in the whole country. Among its most representative monuments are the Miguelete, the Cathedral, the Torres de Serranos and de Quart, the Lonja de la Seda, declared a World Heritage Site by UNESCO in 1996.

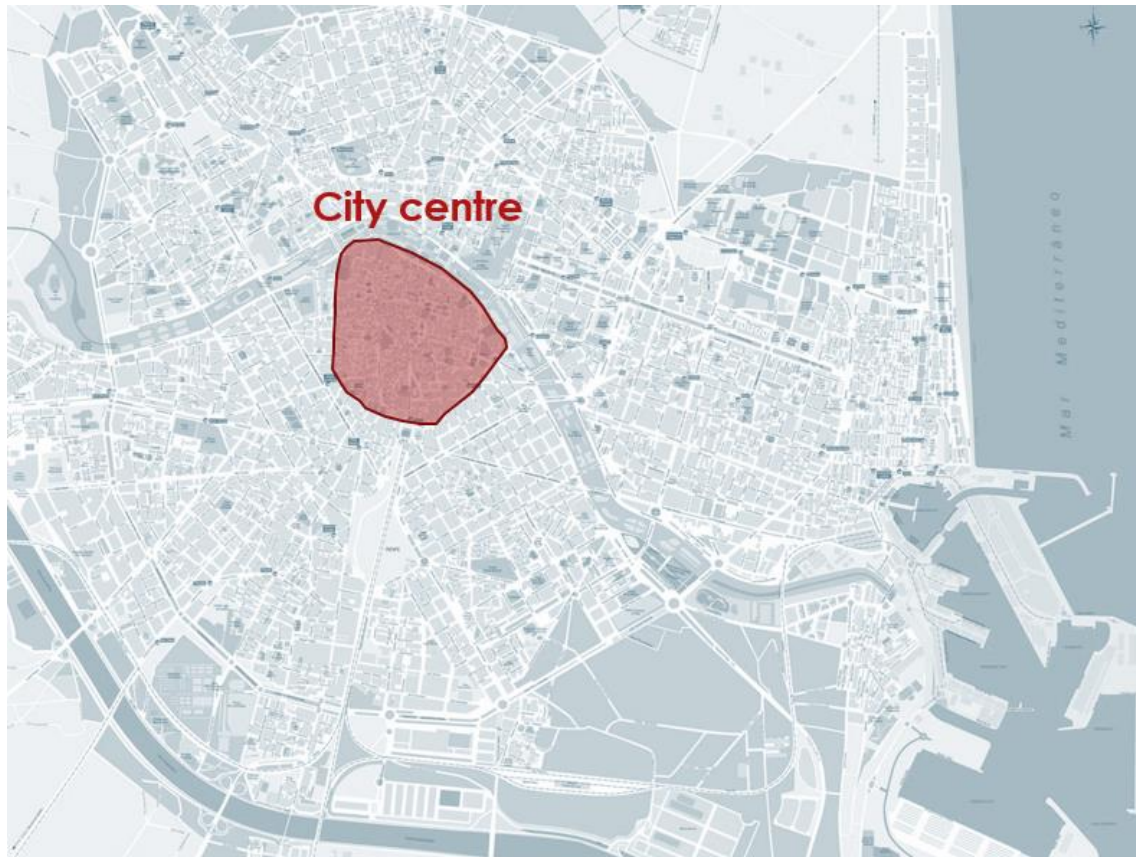


Figure 2. Geographical location of the City centre of Valencia

At present, the economy of the city is centred on services, as nearly 84% of the active population is employed in the service sector. However, the city maintains an industrial base, with a percentage of employed population of 5.5%. On the other hand, agricultural activities, although relatively minor with only 1.9% of the active population employed, remain in the municipality with a total of 3973 hectares, which are mostly occupied by orchards and citrus fruits.

Potential location for a gym business

The city of Valencia is divided 19 districts, being the city centre one of them. Despite of being one of districts with lower population inside the city, the city centre concentrates a lot of business and, therefore, **workers** that can be **potential clients** of the new gym business.

The street map of Valencia has a radial structure, with several concentric axes, being the first concentric axis is the “Ronda Interior”. The “Ronda Interior” surrounds the Valencia City centre and presents several axes to access the city's streets. This city structure facilitates the traffic to and from the city, which could allow people from **surrounding districts** (2,3,4,5 & 6) to access the city centre to be **potential clients in our gym business**.

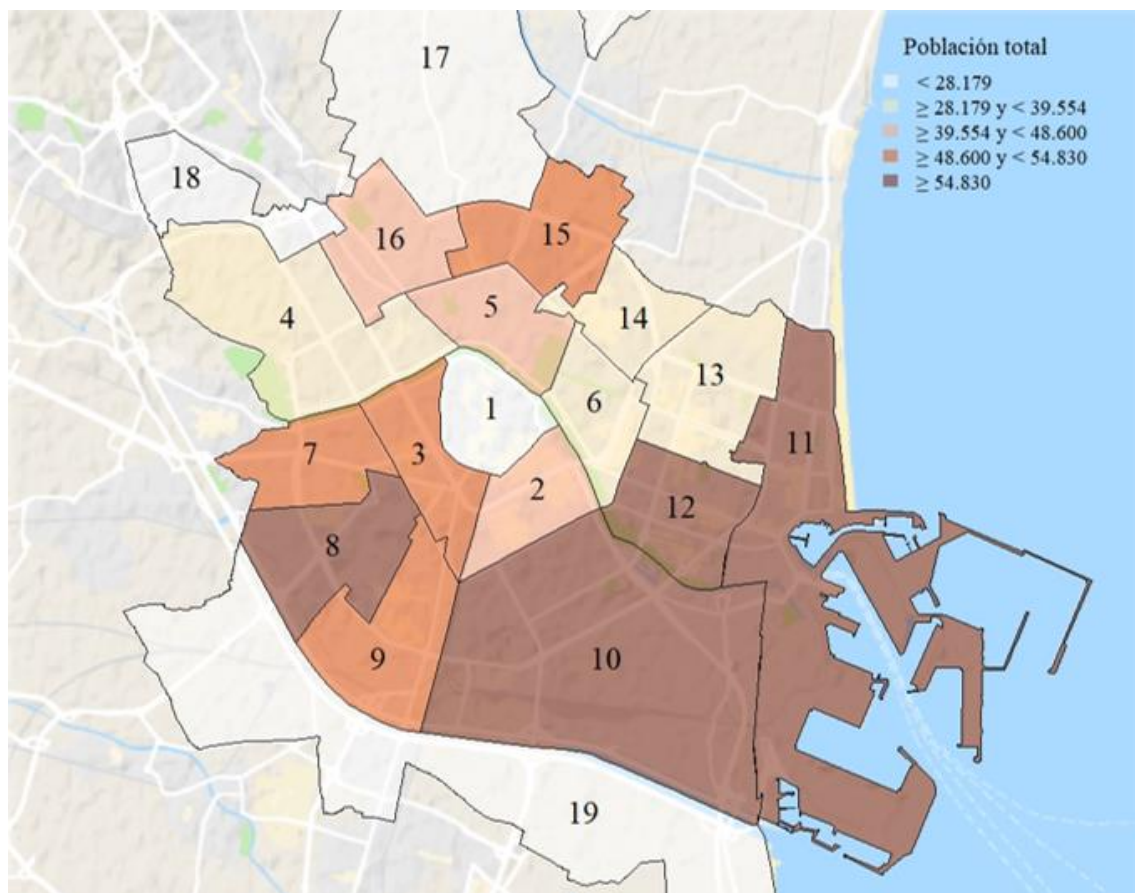


Figure 3. Geographical distribution of the districts of the City of Valencia and its population
Source: City Hall of Valencia Website

3. Methodology followed

1. Define the **initial location** of our potential business area → City centre of Valencia
2. Define the **radius** to identify other gyms inside the area
3. Import the **necessary python libraries** to solve the problem (i.e. requests, pandas, numpy, etc.)
4. Use **Foursquare API** and credentials to identify gyms nearby the area assessed
5. Send the Get request and **examine the results** of the Foursquare API
6. **Transform the data** into pandas and clean the dataset
7. **Visualize** the location of our potential competitors in the selected area
8. **Define an area** to establish our gym business

Initial location definition

For our gym business we set up the centre of our search nearby the "Mercado Central de Valencia", the main market that is located in the city centre:

Coordinates of "Mercado Central de Valencia": 39°28'25.5"N 0°22'46.9"W

- Latitude: 39.473757
- Longitude: -0.379694

Radius of the area assessed

We also set up a radius for our analysis of 2 kilometers.

We define the initial location of the neighborhood to start the gym business (center and radius)

For our gym business we set up the center of our search nearby the "Mercado Central de Valencia", the main market that is located in the city center:

Coordinates of "Mercado Central de Valencia": 39°28'25.5"N 0°22'46.9"W

- Latitude: 39.473757
- Longitude: -0.379694

Besides, we also set up a radius for our analysis of 2kilometers:

- Radius: 2km

```
## Variable definition:
```

```
latitude = 39.473757  
longitude = -0.379694  
radius = 2000  
search_query = 'Gym'
```

Python necessary libraries

To solve the problem, we also need to import the required libraries. Following figure illustrates the libraries used in the analysis:

Importing the required libraries to solve the problem

```
import requests # library to handle requests  
import pandas as pd # library for data analysis  
import numpy as np # library to handle data in a vectorized manner  
import random # library for random number generation  
  
!conda install -c conda-forge geopy --yes  
from geopy.geocoders import Nominatim # module to convert an address into Latitude and Longitude values  
  
# Libraries for displaying images  
from IPython.display import Image  
from IPython.core.display import HTML  
  
# transforming json file into a pandas dataframe library  
from pandas.io.json import json_normalize  
  
!conda install -c conda-forge folium=0.5.0 --yes  
import folium # plotting library  
  
print('Folium installed')  
print('Libraries imported.')
```

Foursquare API

We use the Foursquare API to identify the gyms that are located inside the area assessed:

Foursquare Credentials

```
CLIENT_ID = 'YOUR_CLIENT_ID'  
CLIENT_SECRET = 'YOUR_CLIENT_SECRET'  
ACCESS_TOKEN = 'YOUR_ACCESS_TOKEN'  
VERSION = '20180604'  
LIMIT = 30  
print('Your credentials:')  
print('CLIENT_ID: ' + CLIENT_ID)  
print('CLIENT_SECRET: ' + CLIENT_SECRET)
```



```
## Url definition:
url = 'https://api.foursquare.com/v2/venues/search?client_id={}&client_secret={}&ll={}&oauth_token={}&v={}&query={}&radius={}&limit={}'.format(CLIENT_ID, CLIENT_SECRET, LAT, LONG, OAUTH_TOKEN, VERSION, QUERY, RADIUS, LIMIT)
url
```

Then we use the GET Request and examine the results obtained from Foursquare API

```
results = requests.get(url).json()
results

{'meta': {'code': 200, 'requestId': '6007fb91e1a6d0750d468bbe'},
 'notifications': [{'type': 'notificationTray', 'item': {'unreadCount': 0}}],
 'response': {'venues': [{'id': '4fbbfd7ce4b0d314f1767484',
  'name': 'gym24',
  'location': {'lat': 39.48004615438323,
  'lng': -0.3928476145314537,
  'labeledLatLngs': [{'label': 'display',
  'lat': 39.48004615438323,
  'lng': -0.3928476145314537}],
  'distance': 1329,
  'cc': 'ES',
  'city': 'Valencia',
  'state': 'Comunidad Valenciana',
  'country': 'España',
  'formattedAddress': ['Valencia Comunidad Valenciana']},
  'categories': [{'id': '4bf58dd8d48988d176941735',
  'name': 'Gym',
  'pluralName': 'Gyms',
  'shortName': 'Gym',
  'icon': {'prefix': 'https://ss3.4sqi.net/img/categories_v2/building/gym_',
  'suffix': '.png'},
  'primary': True}]}]}
```

Data processing and cleaning

Once we have retrieved the data, we process it and clean it to have a proper format

```
# assign relevant part of JSON to venues
venues = results['response']['venues']

# transform venues into a dataframe
dataframe = json_normalize(venues)
dataframe.head()
```

/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages/ipykernel_launcher.py:5: FutureWarning: pandas.io.json.json_normalize is deprecated. Use pandas.json_normalize instead

	id	name	categories	referralId	hasPerk	location.lat	location.lng	location.labeledLatLngs	location.distance
0	4fbbfd7ce4b0d314f1767484	gym24	[{'id': '4bf58dd8d48988d176941735', 'name': 'Gym'}]	1611135889	False	39.480046	-0.392848	[{'label': 'display', 'lat': 39.48004615438323, 'lng': -0.3928476145314537}]	1329
1	5de8ca521016f8000844d266	Gym Boutique Alameda	[{'id': '4bf58dd8d48988d175941735', 'name': 'Gym'}]	1611135889	False	39.475635	-0.365503	[{'label': 'display', 'lat': 39.475635, 'lng': -0.365503}]	1237
2	4eb938d4f5b94bd85d61fdbb	Venice Gym	[{'id': '4bf58dd8d48988d176941735', 'name': 'Gym'}]	1611135889	False	39.481440	-0.372702	[{'label': 'display', 'lat': 39.48143993086462, 'lng': -0.372702}]	1045
3	51379782e4b0536567a71523	Gym & Tonic	[{'id': '4bf58dd8d48988d176941735', 'name': 'Gym'}]	1611135889	False	39.462570	-0.370740	[{'label': 'display', 'lat': 39.46257, 'lng': -0.37074}]	1463
4	5182b1b4e4b0ea1baed06438	Master Gym	[{'id': '4bf58dd8d48988d176941735', 'name': 'Gym'}]	1611135889	False	39.470801	-0.391775	[{'label': 'display', 'lat': 39.4708005498919, 'lng': -0.391775}]	1089

Then we define the information of interests and filter the data frame. Process to be followed:

1. We keep only columns that include venue name, and anything that is associated with location
2. We make a function that extracts the category of the venue
3. We filter the category for each row
4. We clean column names
5. We show the data frame filtered

The Battle of Neighbourhoods

	name	categories	lat	lng	labeledLatLngs	distance	cc	city	state	country	formattedAddress	address	postalCode	crossStreet
0	gym24	Gym	39.480046	-0.392848	[{"label": "display", "lat": 39.480046, "lng": -0.392848, "distance": 1329, "cc": "ES", "city": "Valencia", "state": "Comunidad Valenciana", "country": "España", "formattedAddress": "[Valencia Comunidad Valenciana]", "address": "NaN", "postalCode": "NaN", "crossStreet": "NaN"}]	1329	ES	Valencia	Comunidad Valenciana	España	[Valencia Comunidad Valenciana]	NaN	NaN	NaN
1	Gym Boutique Alameda	Gym / Fitness Center	39.475635	-0.365503	[{"label": "display", "lat": 39.475635, "lng": -0.365503, "distance": 1237, "cc": "ES", "city": "Valencia", "state": "Comunidad Valenciana", "country": "España", "formattedAddress": "[Paseo de La Alameda, 4, 46010 Valencia Comunidad Valenciana]", "address": "Paseo de La Alameda, 4, 46010 Valencia", "postalCode": "46010", "crossStreet": "NaN"}]	1237	ES	Valencia	Comunidad Valenciana	España	[Paseo de La Alameda, 4, 46010 Valencia Comunidad Valenciana]	Paseo de La Alameda, 4, 46010 Valencia	46010	NaN
2	Venice Gym	Gym	39.481440	-0.372702	[{"label": "display", "lat": 39.48143993086462, "lng": -0.372702, "distance": 1045, "cc": "ES", "city": "Valencia", "state": "Comunidad Valenciana", "country": "España", "formattedAddress": "[Calle del Poeta Bodria, 4, 46010 Valencia, 46010 Valencia]", "address": "Calle del Poeta Bodria, 4, 46010 Valencia", "postalCode": "46010", "crossStreet": "NaN"}]	1045	ES	Valencia	Comunidad Valenciana	España	[Calle del Poeta Bodria, 4, 46010 Valencia, 46010 Valencia]	Calle del Poeta Bodria, 4, 46010 Valencia	46010	NaN
3	Gym & Tonic	Gym	39.462570	-0.370740	[{"label": "display", "lat": 39.46257, "lng": -0.37074, "distance": 1463, "cc": "ES", "city": "Valencia", "state": "Comunidad Valenciana", "country": "España", "formattedAddress": "[Doctor Sumsi 13, Valencia Comunidad Valenciana]", "address": "Doctor Sumsi 13", "postalCode": "NaN", "crossStreet": "NaN"}]	1463	ES	Valencia	Comunidad Valenciana	España	[Doctor Sumsi 13, Valencia Comunidad Valenciana]	Doctor Sumsi 13	NaN	NaN
4	Master Gym	Gym	39.470801	-0.391775	[{"label": "display", "lat": 39.4708005498919, "lng": -0.391775, "distance": 1089, "cc": "ES", "city": "Valencia", "state": "Comunidad Valenciana", "country": "España", "formattedAddress": "[Martin El Humano 11, 46008 Valencia Comunidad Valenciana]", "address": "Martin El Humano 11", "postalCode": "46008", "crossStreet": "NaN"}]	1089	ES	Valencia	Comunidad Valenciana	España	[Martin El Humano 11, 46008 Valencia Comunidad Valenciana]	Martin El Humano 11	46008	NaN
5	Westin Gym	Gym / Fitness Center	39.473121	-0.361336	[{"label": "display", "lat": 39.473121, "lng": -0.361336, "distance": 1579, "cc": "ES", "city": "Valencia", "state": "Comunidad Valenciana", "country": "España", "formattedAddress": "[46010 Valencia Comunidad Valenciana]", "address": "NaN", "postalCode": "46010", "crossStreet": "NaN"}]	1579	ES	Valencia	Comunidad Valenciana	España	[46010 Valencia Comunidad Valenciana]	NaN	46010	NaN
6	campus gym	Gym	39.481848	-0.364016	[{"label": "display", "lat": 39.48184824501116, "lng": -0.364016, "distance": 1620, "cc": "ES", "city": "Valencia", "state": "Comunidad Valenciana", "country": "España", "formattedAddress": "NaN", "address": "NaN", "postalCode": "NaN", "crossStreet": "NaN"}]	1620	ES	Valencia	Comunidad Valenciana	España	NaN	NaN	NaN	NaN
7	Mô Gym Studio	Gym	39.482729	-0.363171	[{"label": "display", "lat": 39.482729, "lng": -0.363171, "distance": 1735, "cc": "ES", "city": "Valencia", "state": "Comunidad Valenciana", "country": "España", "formattedAddress": "[Calle Bachiller, 7, 46010 Valencia Comunidad Valenciana]", "address": "Calle Bachiller, 7", "postalCode": "46010", "crossStreet": "NaN"}]	1735	ES	Valencia	Comunidad Valenciana	España	[Calle Bachiller, 7, 46010 Valencia Comunidad Valenciana]	Calle Bachiller, 7	46010	NaN
8	Sala de Abdominales Club Metropolitan Gym	Gym	39.456852	-0.375063	[{"label": "display", "lat": 39.45685204525932, "lng": -0.375063, "distance": 1923, "cc": "ES", "city": "Valencia", "state": "Comunidad Valenciana", "country": "España", "formattedAddress": "[Calle Filipinas (Peris Y Valero), 46006 Valencia]", "address": "Calle Filipinas", "postalCode": "46006", "crossStreet": "Peris Y Valero"}]	1923	ES	Valencia	Comunidad Valenciana	España	[Calle Filipinas (Peris Y Valero), 46006 Valencia]	Calle Filipinas	46006	Peris Y Valero

Visualize the location

Once we have identified the locations of the gyms nearby the city centre of Valencia, we can visualize their locations nearby the "Mercado Central de Valencia". Process to be followed:

1. We generate the map in our centre location (Mercado Central de Valencia)
2. We add a marker in the "Mercado Central de Valencia"
3. We add blue spots in the gyms that are inside the radius
4. We display the map

```
# We generate map centred in the "Mercado Central de Valencia"
venues_map = folium.Map(location=[latitude, longitude], zoom_start=15)

# We add in the map the "Mercado Central de Valencia" as a red circle mark
folium.features.CircleMarker(
    [latitude, longitude],
    radius=10,
    popup='Mercado_Central',
    fill=True,
    color='red',
    fill_color='red',
    fill_opacity=0.6
).add_to(venues_map)

# add popular spots to the map as blue circle markers
for lat, lng, label in zip(dataframe_filtered.lat, dataframe_filtered.lng, dataframe_filtered.categories):
    folium.features.CircleMarker(
        [lat, lng],
        radius=5,
        popup=label,
        fill=True,
        color='blue',
        fill_color='blue',
        fill_opacity=0.6
    ).add_to(venues_map)

# display map
venues_map
```


4. Results

The results of our analysis shows all the locations of the gyms nearby the area assessed, the City Centre of Valencia. As can be seen in the map, there is a lack of gyms in the neighbourhood of the City Centre of Valencia, so that, geographically, it seems to be a good district to establish a gym business.

We do the map again including a green circular area without any gym inside. This area has a radius of 800 meters, which requires more or less a 10 to 15min walk to go to the gym.

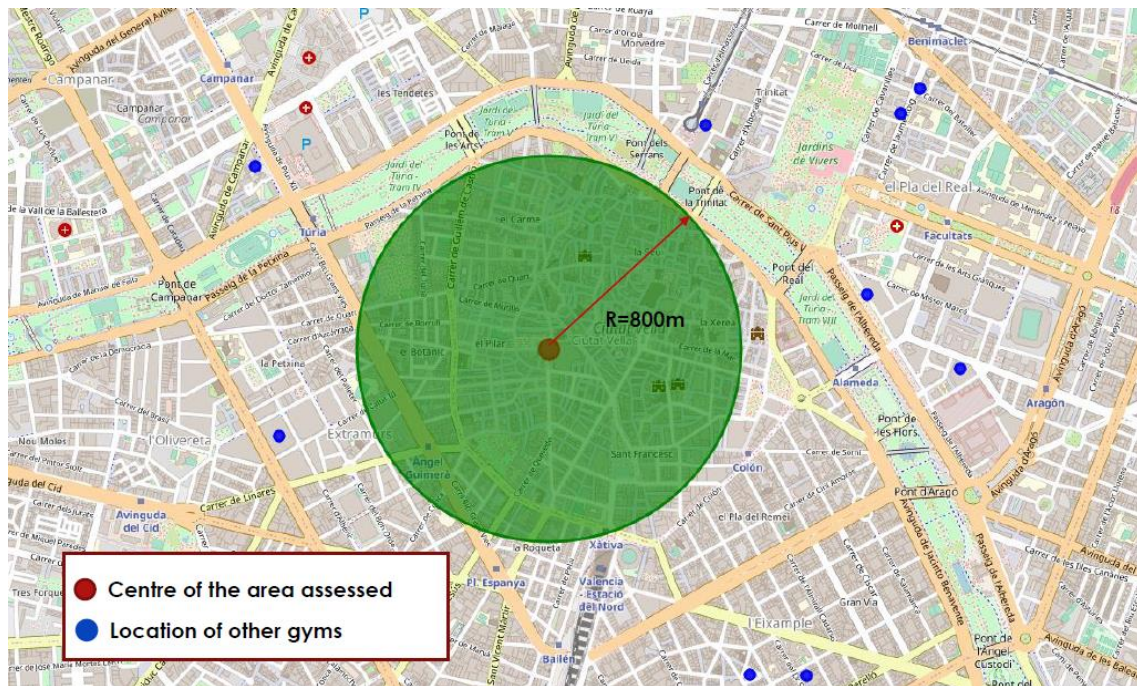


Figure 4. Map of the locations of the Gyms nearby the City Centre of Valencia

5. Conclusions

The analysis carried out shows that the city centre of Valencia is, potentially, a good place to establish a gym business. There is a **lack of gyms nearby** the area assessed, so that clearly exists a business opportunity. Besides, it is a neighbourhood with a lot of activity, which means potential clients for our business.

However, there are other parameters to be assessed to establish a new gym business that are out of the scope of this work. In this sense, **the problem should continue by exploring**, at least, following items:

Venues for establishing the gym that match with the following characteristics:

- Enough venue size (square meters)
- Good price for the venue (euros per square meter)
- Good accessibility for potential clients
- Enough parking places in the surroundings
- Etc.

Market analysis:

- Closest gyms nearby the potential locations finally decided
- Price benchmarking
- Potential clients profile
- Etc.