

# Applied Data Science Capstone Project Data Science Certification (IBM)

Where to set up a café in Sitges, Spain?

Mariana Ferreira



## Introduction: Business Problem

- I will attempt to determine the most optimal location for setting up a new café in the small costal town of Sitges, Spain.
- Location is critical to the overall success of any business.
- The search will focus on finding locations not too crowded by local cafés and locations as close to the city center as possible.
- I will also focus on identifying areas of interest, since they usually signify more foot traffic.
- Competition isn't a bad thing; a few cafés can attract more people to the area, the key is to differentiate our business from the competition.

# DATA

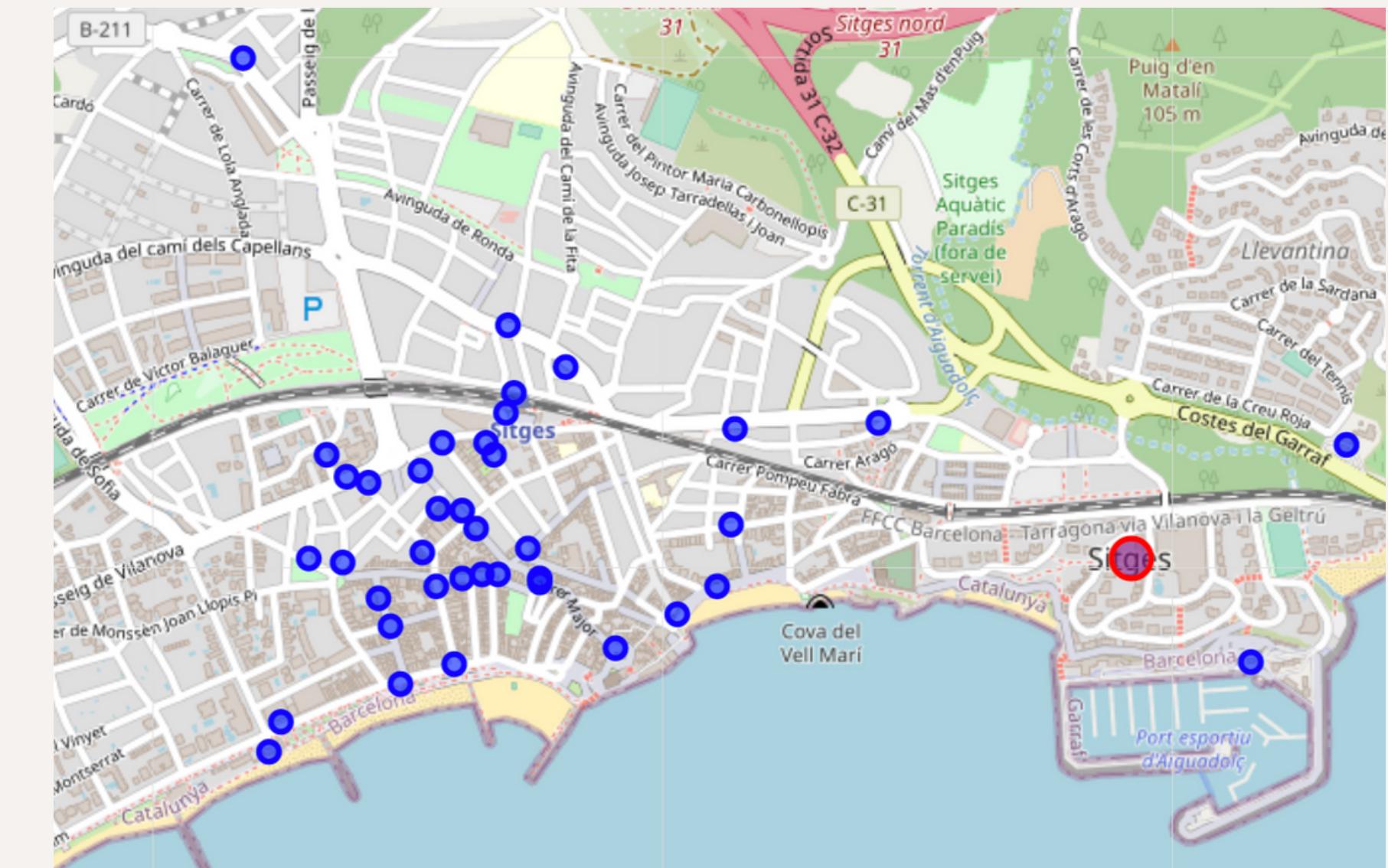
- Due to how small Sitges is, there is no unified criteria for neighborhood division.
- I manually created a csv file including neighborhoods' names, postcodes and coordinates, which I uploaded directly to my project.

## Factors that will influence the decision are:

- number of and distance to existing cafés in the neighborhood
  - number of points of interest (universities, schools, public offices, etc)
  - distance from neighborhood to city center
- 
- To find the number of existing cafés and landmarks and their location per neighborhood, I will be using the Foursquare API.

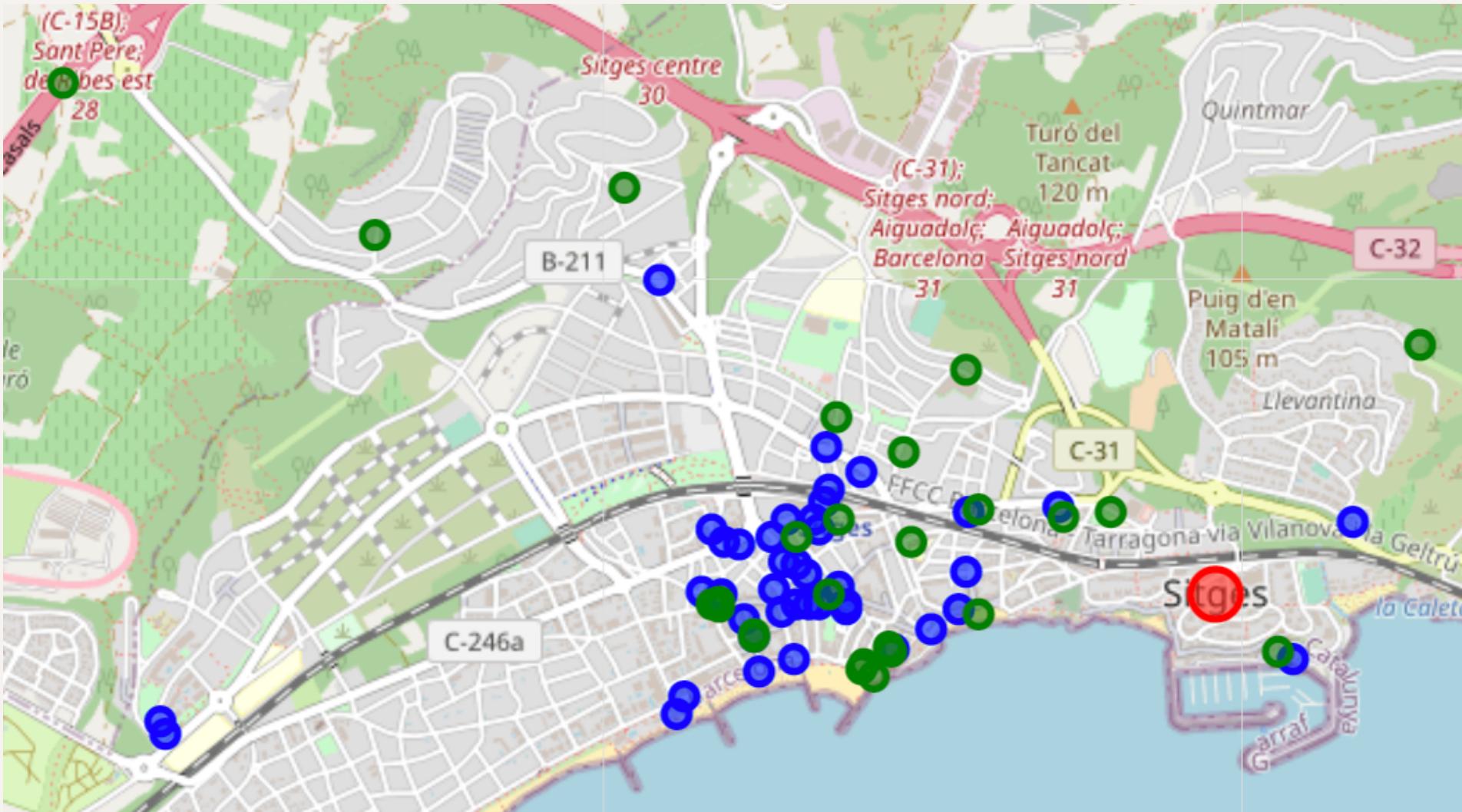
## Manually obtained neighborhood database:

	Neighborhood	Postal Code	Latitude	Longitude
0	Downtown Sitges/Sant Sebastià	08870	41.2372	1.8059
1	Aiguadolç	08870	41.2348	1.8254
2	Urbanització Levantina	08870	41.2403	1.8249
3	Urbanització Montgavina	08870	41.2383	1.8313
4	Quint Mar	08870	41.2495	1.8661
5	Urbanització Vallpineda	08870	41.2459	1.7962
6	Poble Sec	08870	41.2439	1.8088
7	Els Molins	08870	41.2392	1.8170
8	La Plana	08870	41.2387	1.8097
9	Vinyet	08870	41.2344	1.7911
10	Urbanització Rocamar	08870	41.2352	1.7817
11	Terramar	08870	41.2279	1.7822



Then I used the Foursquare API to get the top 100 cafés in Sitges within a radius of 3000 meters from Downtown. With that new dataframe, I created a map to visualize all cafés.

Then, with Foursquare API, I located points of interest within the same radius and mapped them along with the cafés, to get a sense of where most venues were located.



- Cafés
- Points of interest

This map shows what my previous research suggested; areas with hospitals, schools, post offices and/or parks, are preferred when it comes to setting up a café.

# METHODOLOGY

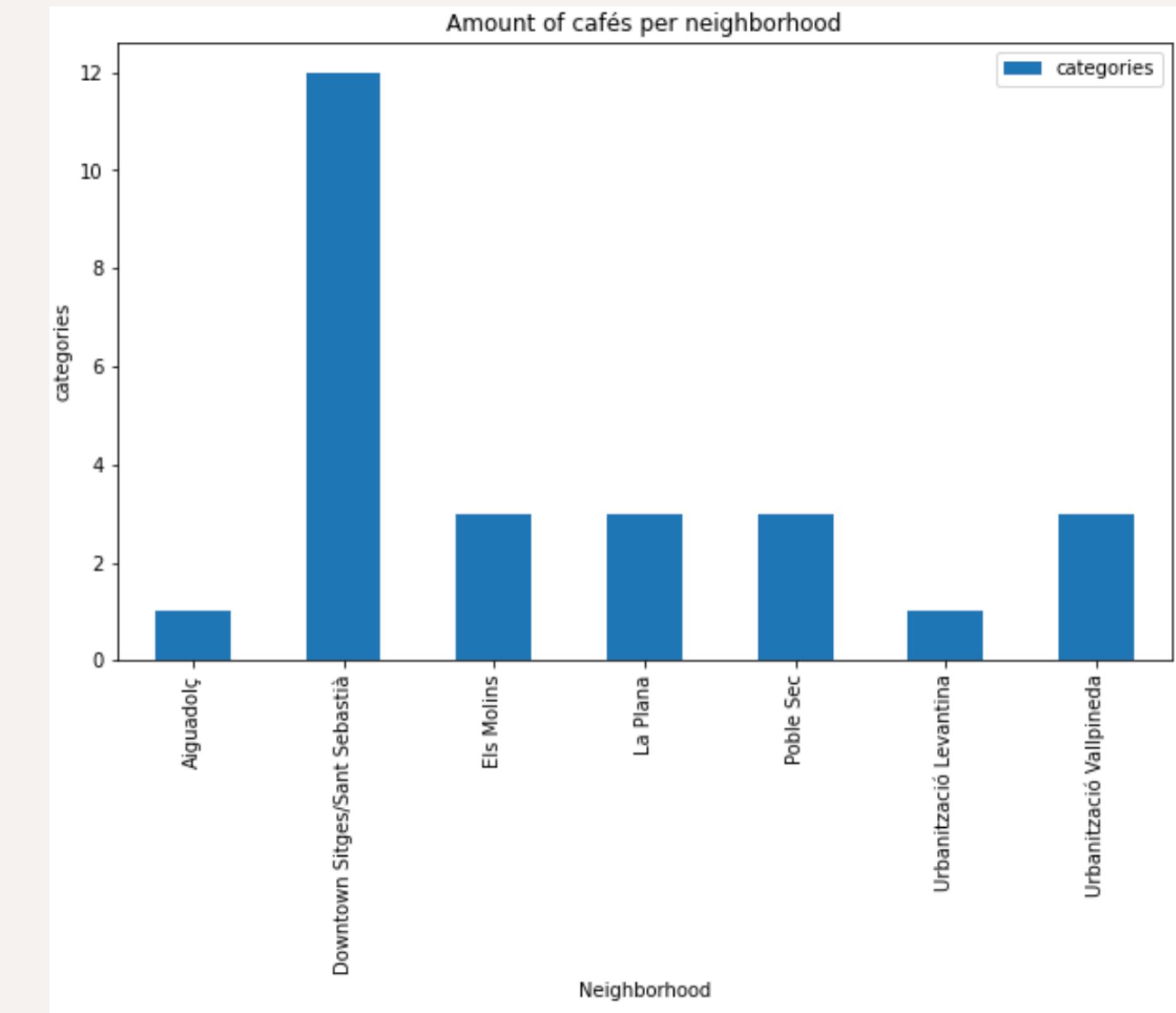
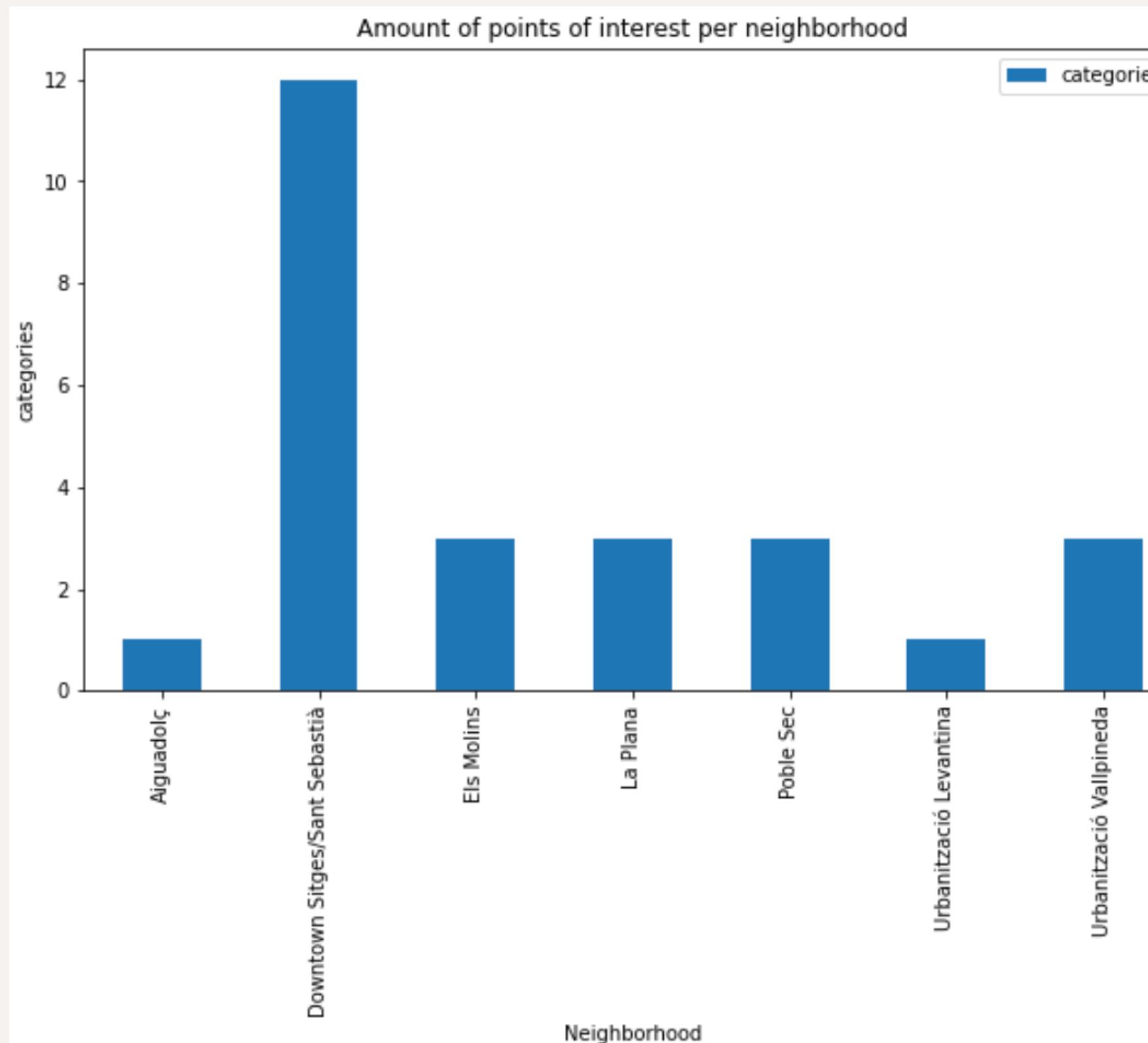
- Collect the required data.
- Detect areas where there are points of interest, specially areas with low number of cafés.
- Determine 'venue frequency' across Sitges.
- Focus on areas where there are at least two points of interest and less than four cafés (competition).
- Focus on most promising areas and mark on the map, specific locations where we could set up our business.

# ANALYSIS

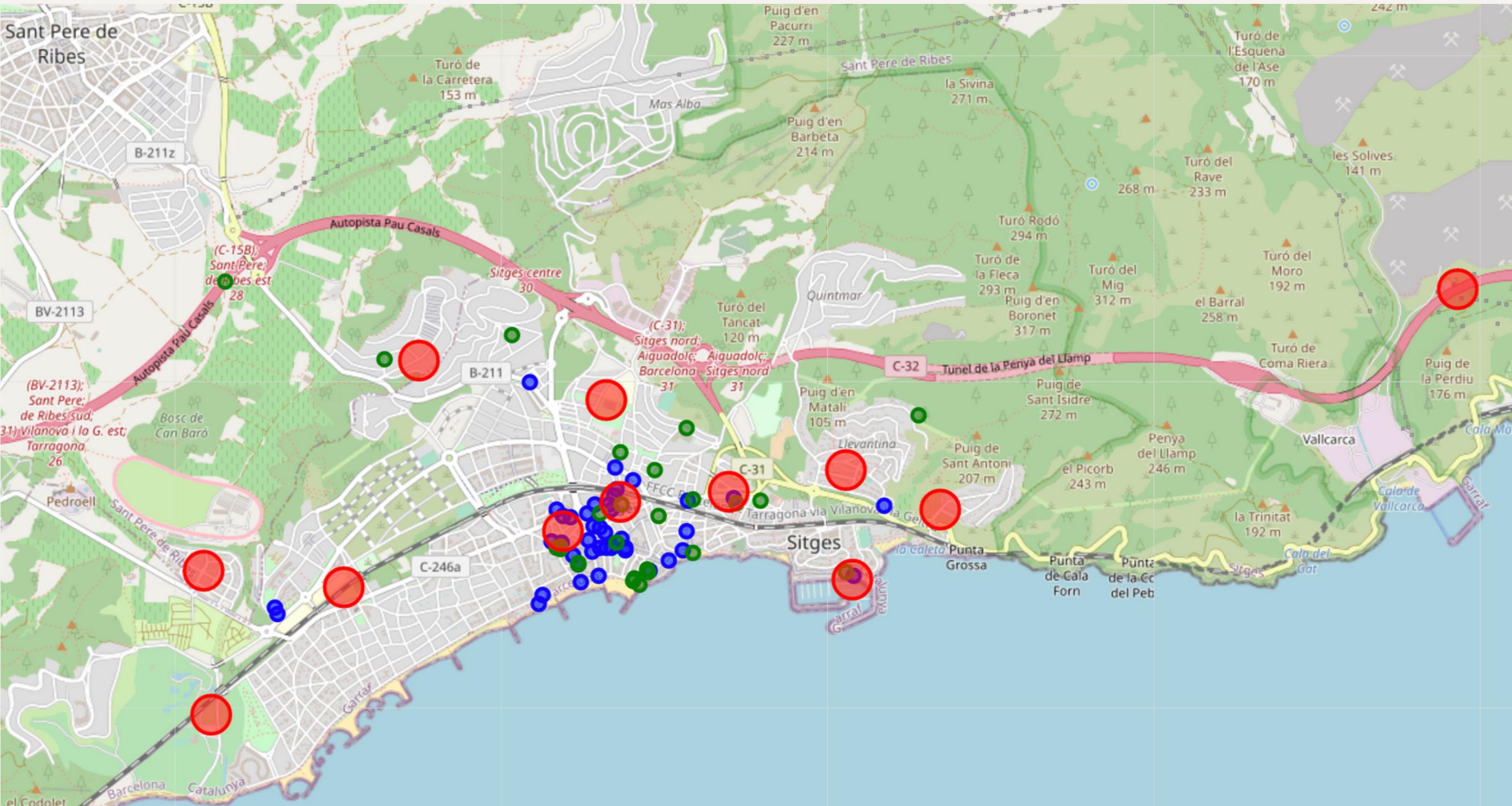
I merged all dataframes to create one dataset with neighborhoods and all venues; cafés and points of interest.

id	categories	Neighborhood
<b>4bed37eb6e8c20a1c2017061</b>	Café	Downtown Sitges/Sant Sebastià
<b>4e66200a62e13e3bd03011d7</b>	Café	Downtown Sitges/Sant Sebastià
<b>53836f54498e15919e982ae4</b>	Café	Downtown Sitges/Sant Sebastià
<b>51727d54e4b02cd2ba327959</b>	Café	Downtown Sitges/Sant Sebastià
<b>4bfd49e8e529c928f5abba8c</b>	Café	Downtown Sitges/Sant Sebastià
<b>5b931f6bbcbf7a002ceca396</b>	Café	Downtown Sitges/Sant Sebastià
<b>4d496a925129a35d4b6f39ac</b>	Café	Downtown Sitges/Sant Sebastià
<b>5e0628f0a5f58500073d4d68</b>	Café	Downtown Sitges/Sant Sebastià
<b>4dc6e9dc45dd264552959480</b>	Café	Downtown Sitges/Sant Sebastià
<b>518e6961498eac2b648aa51d</b>	Café	La Plana
<b>4bfa67b5508c0f475c064031</b>	Café	Downtown Sitges/Sant Sebastià

Measured the frequency of cafés and points of interest per neighborhood and created bar charts to visualize which neighborhoods feature the most venues.



Also mapped all neighborhoods, cafes and neighborhoods:

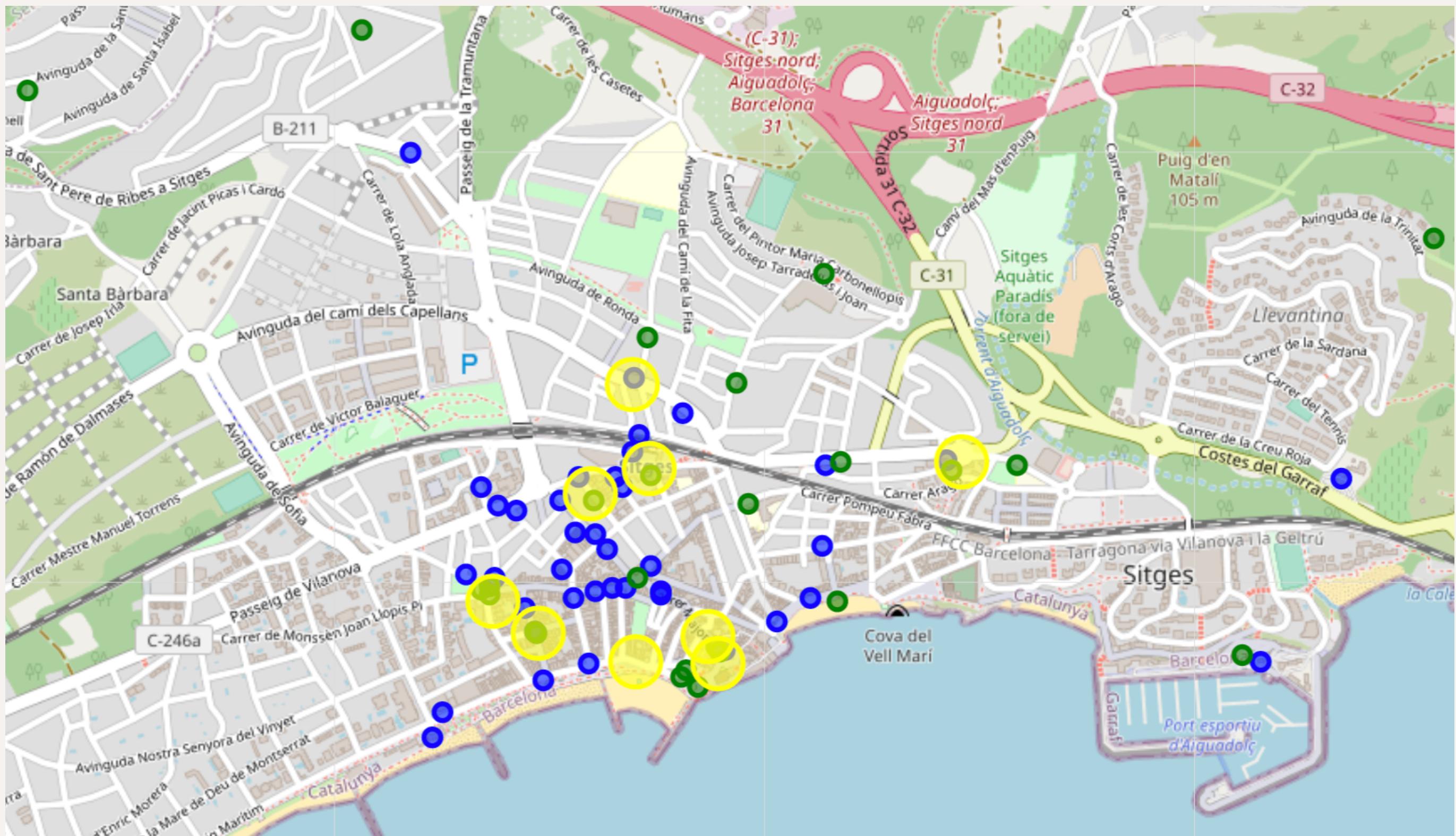


- Cafés
- Points of interest
- Neighborhoods

- Based off of the charts and the map, I determined which neighborhoods feature the highest venue density, Downtown being the main one.
- Created a dataframe with potential locations and their coordinates.

	PL	Latitude	Longitude
0	Location 1	41.235474	1.811303
1	Location 2	41.234948	1.811558
2	Location 3	41.234994	1.809444
3	Location 4	41.235530	1.806950
4	Location 5	41.238215	1.808291
5	Location 6	41.238697	1.809786
6	Location 7	41.240330	1.809360
7	Location 8	41.236170	1.805771
8	Location 9	41.238836	1.817763

- Created a final map featuring those potential locations for setting up the business, which are marked in yellow.



## RESULTS AND DISCUSSION

- Sitges is a small town in Spain so not a lot of information available online.
- Manually created a database by analyzing local real state and tourism information.
- Used the Foursquare API to find all cafés and points of interest in town.
- Created a map and bar charts with all the venues, to determine neighborhoods with the highest venue density.
- Most venues are concentrated around Downtown and La Plana.
- Other neighborhoods are mostly residential areas, less likely to be the good options with high foot traffic.
- Despite the limitations, there are still areas near points of interest (with high foot traffic), where we could set up a new café.
- **The result of this analysis is nine potential candidates.**
- These areas are, mostly around **Downtown and La Plana**, near at least two points of interest and around less than four competitors.

## CONCLUSION

Purpose was to identify Sitges' areas close to the center, with low number of cafés and potentially high foot traffic, to narrowing down the search for optimal location for a new café.

Using Foursquare data, determined venue density distribution across town using Foursquare data.

Identified the neighborhoods with the most potential.

Created a database of nine locations, to be used as starting points for final exploration.

Final decision on optimal café location will be made by future business owners based on specific characteristics of neighborhoods and locations in every recommended zone.

Taking into consideration proximity to major roads, real estate availability, prices, distinctive characteristics of the competitors and the new business, among others.

# REFERENCES

<https://www.idealista.com>

<https://www.sevenmiles.com.au/editorial/cafe-location/>

<https://articles.bplans.com/how-to-choose-a-business-location/>

<https://www.sitges-tourist-guide.com/en/areas/sitges-districts.html>