

The battle of neighborhoods, a study to analyze accommodation possibilities in Madrid

Capstone project

by Chary Mariñas.

Business case

Due to the current situation with the coronavirus, the behavior has radically changed so the statistics could not be representative right now.

According to the tourism page of the Community of Madrid, the city of Madrid registered in 2019 the arrival of 10.4 million visitors (+ 2.1% vs 2018), of which 55% were international and 45 % national. 22.6 million overnight stays were generated (+ 5.7% vs 2018). The main source markets for international tourism were the United States, Italy, France, the United Kingdom and Germany, and those that grew the most were the United States, China, Italy, Brazil and Mexico. A total expenditure of 10,451 million (+ 12.6% vs 2018) and an average daily expenditure of 270 euros (+ 11% vs 2018) was reached in the Community of Madrid.([+info](#).)

A tourist who comes to visit Madrid may have a great dilemma with where to stay, so we will analyze the best places based on the offer and services in the Centro district.

Data

Dataset sources for this project will be:

- 1.Madrid apartment listing from the Inside Airbnb site - [+info](#)
- 2.Madrid Neighborhoods Tabulation Areas from Inside Airbnb site - [+info](#)

3.Foursquare API to extract data about venues - food places, museums, galleries, shopping centers, sightseeing attractions, concert halls and so on

Initially, Madrid apartment list from Airbnb has 21255 records so we filtered according to following filters:

- District = 'Centro';
- Availability \geq 100 days/year;
- Number of reviews \geq 10;
- Last Scraped/Reviewed later than 2020-04-01;
- Minimum nights \geq 1;
- Excluded Hostels and Camper/RV;
- Excluded Shared rooms.

This leaves a model with 396 records.

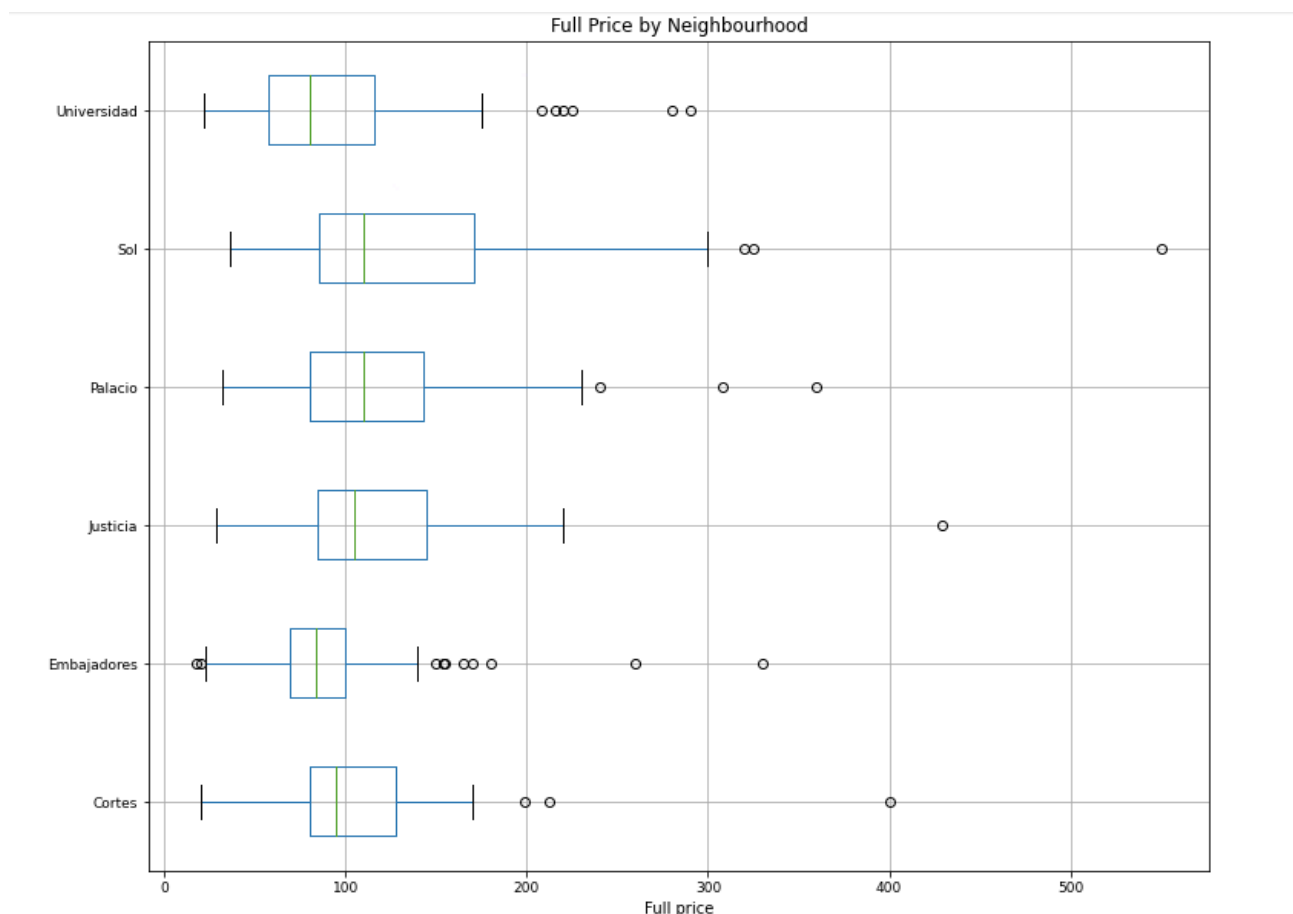
	neighbourhood_group	neighbourhood
0	Arganzuela	Acacias
1	Arganzuela	Atocha
2	Arganzuela	Chopera
3	Arganzuela	Delicias
4	Arganzuela	Imperial
5	Arganzuela	Legazpi
6	Arganzuela	Palos de Moguer
7	Barajas	Aeropuerto
8	Barajas	Alameda de Osuna
9	Barajas	Casco Histórico de Barajas
10	Barajas	Corralejos
11	Barajas	Timón
12	Carabanchel	Abrantes
13	Carabanchel	Buenavista
14	Carabanchel	Comillas
15	Carabanchel	Opañel
16	Carabanchel	Puerta Bonita
17	Carabanchel	San Isidro
18	Carabanchel	Vista Alegre
19	Centro	Cortes

Then we read the data to assign a latitude and longitude to each neighborhood:

	Neighborhood	Latitude	Longitude
0	Palacio	40.415129	-3.715618
1	Embajadores	40.409681	-3.701644
2	Cortes	40.414779	-3.697584
3	Justicia	40.423957	-3.695747
4	Universidad	40.425409	-3.705989
5	Sol	40.417827	-3.702222

For the analysis, it is interesting that we calculate the price per person for the Airbnb places and the final one, adding the cleaning fee.

To print out a box plot is the easiest way to visualize the outliers:



Once the outliers are deleted, we still have 360 records to analyze.

Methodology

First I will collect all the information regarding the homes available on AirBnB for Madrid. Then, I will filter them to keep only the neighborhoods belonging to the "Centro" district. This information will be crossed with that available on the geolocation of the neighborhoods and some important calculations will be made such as the price per person (including the cleaning fee). Finally, the information available on Foursquare will be searched for these neighborhoods and they will be divided into clusters.

neighbourhood_cleansed	mean_price_per_person	accommodates
Cortes	25.212212	62
Universidad	26.012543	56
Sol	27.275000	54
Embajadores	27.424786	100
Palacio	27.472275	53
Justicia	32.534444	45

This way, a decision can be made on which one is the best to stay.

Results

First thing we can see is the average price difference grouped by neighborhoods based on the number of available accommodation places:

Based on the Foursquare recommendations, we calculate the top 3 venues for the 5 best places:

	name	Entertainment	Fine Art	Food Place	Shopping	Sightseeing
0	****Studio Centre-Live-In Atocha-WiFi,A/C ****	0.15	0.22	0.48	0.04	0.11
1	***Loft Centre- Reina Sofia-Atocha-WiFi, A/C***	0.15	0.22	0.48	0.04	0.11
2	***Loft Centre-Museum Mile, WiFi, A/C***	0.14	0.21	0.43	0.04	0.18
3	***Studio Centre-Live-In Atocha-WiFi,A/C ***	0.17	0.21	0.54	0.04	0.04
4	**Center Luxurious Apartment. Museum Mile Wifi,A/C	0.11	0.22	0.48	0.04	0.15

Finally we divide the results into 3 clusters:

	review_scores_rating	accommodates	full_price	price_per_person	bathrooms	bedrooms	Cluster Labels
count	39.000000	39.000000	39.000000	39.000000	39.000000	39.000000	39.0
mean	97.743590	3.641026	99.564103	30.044017	1.179487	1.358974	2.0
std	1.312249	1.547276	40.799764	13.476443	0.451419	0.810691	0.0
min	96.000000	1.000000	35.000000	12.666667	1.000000	0.000000	2.0
25%	97.000000	2.000000	69.500000	20.625000	1.000000	1.000000	2.0
50%	98.000000	4.000000	85.000000	27.500000	1.000000	1.000000	2.0
75%	99.000000	4.000000	126.000000	35.000000	1.000000	2.000000	2.0
max	100.000000	8.000000	183.000000	72.500000	3.000000	4.000000	2.0

	review_scores_rating	accommodates	full_price	price_per_person	bathrooms	bedrooms	Cluster Labels
count	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000	31.0
mean	97.451613	3.387097	103.838710	32.977688	1.161290	1.258065	0.0
std	0.960511	1.563701	45.934807	14.128152	0.373878	0.893224	0.0
min	96.000000	1.000000	20.000000	13.875000	1.000000	0.000000	0.0
25%	97.000000	2.000000	77.000000	21.250000	1.000000	1.000000	0.0
50%	97.000000	3.000000	95.000000	28.750000	1.000000	1.000000	0.0
75%	98.000000	4.000000	138.000000	42.050000	1.000000	2.000000	0.0
max	100.000000	8.000000	208.000000	69.500000	2.000000	4.000000	0.0

	review_scores_rating	accommodates	full_price	price_per_person	bathrooms	bedrooms	Cluster Labels
count	30.000000	30.000000	30.000000	30.000000	30.000000	30.000000	30.0
mean	97.466667	3.466667	85.166667	28.205833	1.133333	1.066667	1.0
std	1.166585	1.455864	30.735898	14.098117	0.434172	0.583292	0.0
min	96.000000	1.000000	31.000000	11.500000	1.000000	0.000000	1.0
25%	96.250000	2.250000	69.000000	19.625000	1.000000	1.000000	1.0
50%	97.500000	4.000000	84.000000	23.000000	1.000000	1.000000	1.0
75%	98.000000	4.000000	93.000000	31.250000	1.000000	1.000000	1.0
max	100.000000	8.000000	199.000000	63.000000	3.000000	3.000000	1.0

From the first cluster, better places to stay are:

- Cozy Apartament Close To Plaza Callao In Sol. Full price 139€.
- Luxury Pza Mayor In La Latina . Full price 127€.
- Sophisticated & Refurbished-1Bd 1Bth-Colon-Centre In Justicia. Full price 162€.

From the second cluster:

- Apartamento VIP Madrid Centro III in Embajadores. Full price 115€.
- Vacaciones para 2 "Lavapies" in Embajadores. Full price 54€.
- Lovely Immaculate 2 Br Flat In Old in Embajadores. Full price 84€.

From the third cluster:

- Madrid Centro - Latina Plaza Mayor Apartamento in Embajadores. Full price 80€.
- Apartment In Madrid's Trendy Heart in Justicia. Full price 109€.
- Diseño vanguardista con espectaculares vistas a la Plaza de Santa Ana in Cortes. Full price 120€.

Observations

In the Centro district food places predominate, since this has always been the first venue of all the Airbnb accommodation available. In second place there are generally either entertainment or fine art, so the reason for choosing one site or another will be more based on price.

Based on the top3 places for each cluster, Embajadores is the cheapest neighborhood where stay (54€) and the most expensive is Justicia.

Conclusions

The Downtown district has many places to eat, with very high scores, so it should not be decisive when choosing a place to stay as it is the 1st common venue.

If the client wants to stay in an area where he can enjoy fine art, he should choose a hotel in the Embajadores area, since that is where the museums are.

In case the client wants entertainment, the best option is to choose an accommodation in Sol.