

Coursera Capstone

IBM Applied Data Science Capstone

Identifying real estate opportunities in Madrid, Spain

By Juan Carlos Martinez
May 2020



Introduction

For the average family buying a house is a life-change decision and in most of the cases a one time event. With the increase of population among the cities, and therefore the demand and prices of houses, this decision is getting even tougher.

All this factors are not new, but the recent world pandemic due to the COVID virus is going to affect our lives in a way that we are not even certain yet, including the way of investing in a new house, changing the rules of the game.

Prior to this emergency, we could assume that the closer a house is to the city center the higher the price, as most of the business are in this area and the main venues of the city, but now companies understand that it is indeed possible for a high jobs percentage to work from home, changing the way the current real estate market works.

Business Problem

The purpose of this capstone project is to analyze the different Madrid neighborhoods in order to find similitudes, and to compare the price of the houses, so we can understand which areas are underpriced.

Target Audience of this project

This project is useful for real estate investors as well to individuals who are looking to buy a new home. It has been developed now under the assumption that work conditions are going to drastically change, making possible to live further away from our place of work. This combined to the high rents of the capital of Spain, should be a big enough stimulus to consider moving to other neighborhoods in the city.

Data

To solve this problem we will need the following data:

- List of neighborhoods in Madrid. This defines the scope of this project which is confined to the city of Madrid, the capital city of the country of Spain.
- Latitude and longitude coordinates of those neighborhoods. This is required in order to plot the map and also to get the venue data.
- Venue data for the different neighborhoods, in order to see the similitude of the different areas.
- Housing prices for the different neighborhoods. We will use the price per square meter to compare economic situation of the real estate sector in the different areas.
- Neighborhoods geometry and position, to be able to divide the city in its different parts.

Sources of data and methods to extract them

1.-We will use Wikipedia

(https://es.wikipedia.org/wiki/Anexo:Barrios_administrativos_de_Madrid) to get the list of Madrid city neighborhoods. We will use web scraping techniques to extract the data from the Wikipedia page, with the help of Python requests and wikipedia package.

2.-We will get the house prices per neighborhood from idealista

(<https://www.idealista.com/sala-de-prensa/informes-precio-vivienda/venta/madrid-comunidad/madrid-provincia/madrid/>). Idealista is the principal real estate portal in Spain.

3.- We will get the geographical coordinates of the neighbourhoods using Python Geocoder package which will give us the latitude and longitude coordinates of the neighbourhoods.

4.-We will use the Foursquare API to get the venue data for those neighbourhoods. Foursquare has one of the largest database of 105+ million places and is used by over 125,000 developers.

5.- We will use the geojson given by the Madrid statistics web

(<https://datos.madrid.es/portal/site/egob/menuitem.9e1e2f6404558187cf35cf3584f1a5a0/?vgnextoid=374512b9ace9f310VgnVCM100000171f5a0aRCRD&vgnnextchannel=374512b9ace9f310VgnVCM100000171f5a0aRCRD&vgnnextfmt=default>)

This is a project that will make use of many data science skills, from web scraping (Wikipedia), working with API (Foursquare), data cleaning, data wrangling, to machine learning (K-means clustering) and map visualization (Folium). In the next section, we will present the Methodology section where we will discuss the steps taken in this project, the data analysis that we did and the machine learning technique that was used.

Methodology

Firstly, we need to get the list of neighborhoods in the city of Madrid Fortunately, the list is available in the Wikipedia page (https://es.wikipedia.org/wiki/Anexo:Barrios_administrativos_de_Madrid). We will do web scraping using Python requests and wikipedia package to extract the list of neighbourhoods data. We will also need the housing pricing for the different neighborhoods so we will get the information from the idealista webpage. As the data is spread between different pages, we will manually download the data and import it in our model. The next step is to get the geographical coordinates in the form of latitude and longitude in order to be able to use Foursquare API. To do so, we will use the Geocoder package that will allow us to convert address into geographical coordinates in the form of latitude and longitude. After gathering the data, we will populate the data into a pandas DataFrame to combine all our information in a single table.

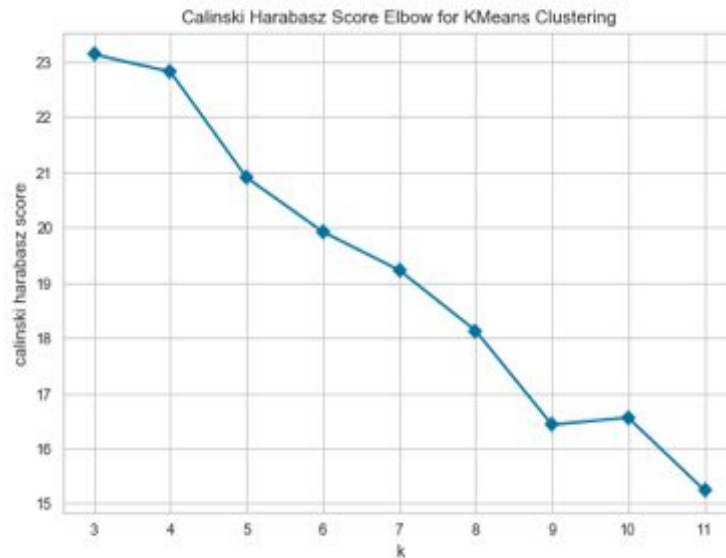
Next, we will use Foursquare API to get the top 100 venues that are within a radius of 2000 meters. We need to register a Foursquare Developer Account in order to obtain the Foursquare ID and Foursquare secret key. We then make API calls to Foursquare passing in the geographical coordinates of the neighborhoods in a Python loop. Foursquare will return the venue data in JSON format and we will extract the venue name, venue category, venue latitude and longitude. With the data, we can check how many venues were returned for each neighborhood and examine how many unique categories can be curated from all the returned venues. Then, we will analyze each neighborhood by grouping the rows by neighborhood and taking the mean of the frequency of occurrence of each venue category.

Then, we will perform clustering on the data by using k-means clustering. K-means clustering algorithm identifies k number of centroids, and then allocates every data point to the nearest cluster, while keeping the centroids as small as possible. It is one of the simplest and popular unsupervised machine learning algorithms and is particularly suited to solve the problem for this project. We will use the elbow method to detect the optimal cluster number to divide the neighborhoods. The results will allow us to identify similar neighborhoods in terms of venues.

Finally we will display the results in a map using Folium, dividing the map by neighborhoods and coloring the map using the choropleth tool, so it is easier to compare price/type of neighborhood for the different clusters.

Results

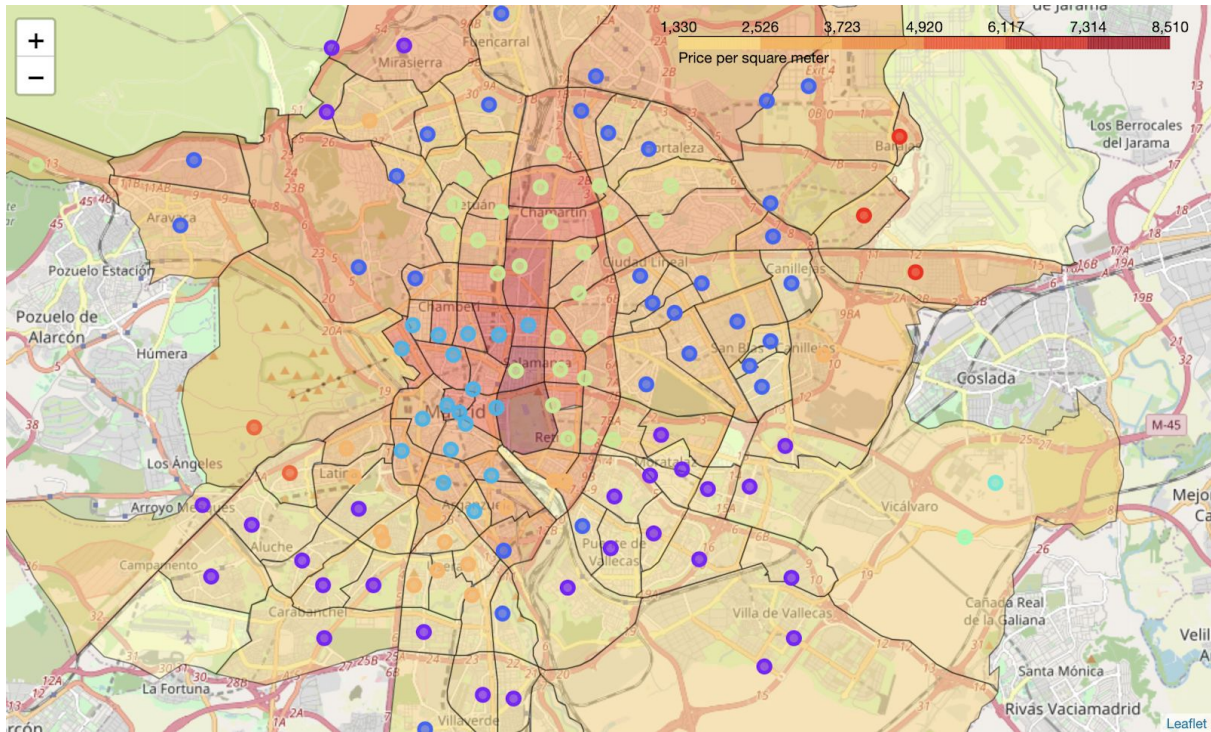
First, we used the elbow method to determine the optimal number of clusters obtaining 9:



- Cluster 0: 'Airport' neighbourhoods (Red in the map)
- Cluster 1: Outdoor South neighbourhoods. Residential areas with a lot of grocery stores, supermarkets and bars. (Purple in the map)
- Cluster 2: Outdoor North neighborhoods. Residential areas with supermarkets, Spanish restaurants and green areas (Dark blue in the map)
- Cluster 3: Center neighborhoods. Area with a lot of hotels, restaurants and squares. (Light blue in the map)
- Cluster 4: Only one neighborhood: El Cañaveral - Los Berrocales. The algorithm made the cluster considering a lot of venues related to tolls and road businesses. (Tourquoise in the map)
- Cluster 5: Only one neighborhood: Casco Histórico de Vicálvaro. The algorithm made the cluster considering a lot of venues related to tolls and the presence of the zoo. (Tourquoise-green in the map)
- Cluster 6: North neighborhoods. Area with a lot of different cuisine restaurants, cafes, and boutiques. (Light green in the map)
- Cluster 7: South neighborhoods: Area with a lot of parks, grocery stores and Spanish restaurants. (Orange in the map)

- Cluster 8: 'Theme park neighborhoods'. Main venues related to the theme park. (Dark orange in the map).

We also added the variable housing price into the map, we can see a heat map with the clusters information:



Considering only the neighborhood type we can see that a couple of neighborhoods in the north area have a lower price than their surroundings. If we compare them to the mean price of the cluster:

	Neighborhood	Cluster Labels	Price	Max	Max_period	Year change	Cluster mean price	Price/Max
0	Fuentelarreina	1	3311	4296	abr 2011	-0.044	2280.758621	0.770717
1	La Paz	2	3476	3825	dic 2010	-0.055	3103.000000	0.908758
2	Tres Olivos - Valverde	2	2698	2895	dic 2010	0.035	3103.000000	0.931952
3	Apóstol Santiago	2	2890	2963	mar 2020	0.008	3103.000000	0.975363
4	Pinar del Rey	2	2863	3238	nov 2010	-0.047	3103.000000	0.884188
5	Valdezarza	2	3090	3448	mar 2019	-0.075	3103.000000	0.896172
6	Las Tablas	6	4192	4324	jul 2019	0.006	4478.103448	0.969473
7	Canillas	6	3242	3692	nov 2009	0.005	4478.103448	0.878115
8	Bellas Vistas	6	3513	3688	may 2019	-0.024	4478.103448	0.952549
9	Berruguete	6	3253	3330	nov 2009	0.022	4478.103448	0.976877
10	Valdeacederas	6	3252	3541	oct 2008	0.027	4478.103448	0.918385
11	Ventilla-Almenara	6	3383	3494	feb 2020	-0.014	4478.103448	0.968231
12	Peñagrande	7	3228	3428	abr 2010	-0.007	2754.250000	0.941657

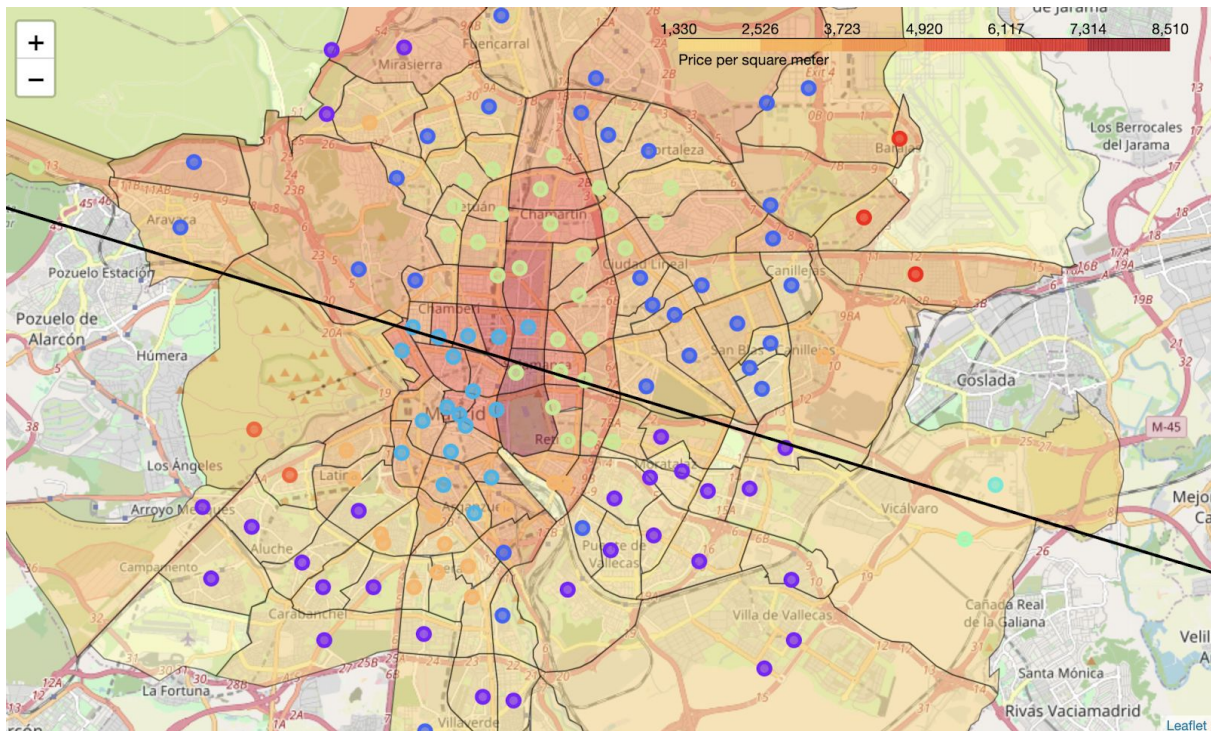
We can observe that from the table only the following neighborhood prices are under the cluster mean:

	Neighborhood	Cluster Labels	Price	Max	Max_period	Year change	Cluster mean price	Price/Max
2	Tres Olivos - Valverde	2	2698	2895	dic 2010	0.035	3103.000000	0.931952
3	Apóstol Santiago	2	2890	2963	mar 2020	0.008	3103.000000	0.975363
4	Pinar del Rey	2	2863	3238	nov 2010	-0.047	3103.000000	0.884188
5	Valdezarza	2	3090	3448	mar 2019	-0.075	3103.000000	0.896172
6	Las Tablas	6	4192	4324	jul 2019	0.006	4478.103448	0.969473
7	Canillas	6	3242	3692	nov 2009	0.005	4478.103448	0.878115
8	Bellas Vistas	6	3513	3688	may 2019	-0.024	4478.103448	0.952549
9	Berruguete	6	3253	3330	nov 2009	0.022	4478.103448	0.976877
10	Valdeacederas	6	3252	3541	oct 2008	0.027	4478.103448	0.918385
11	Ventilla-Almenara	6	3383	3494	feb 2020	-0.014	4478.103448	0.968231

Discussion

At first sight, it might seem odd that the clusters are that well distributed, but considering how the city has been expanding through the years it makes sense that the clusters adjust to this growth of the city.

We can clearly appreciate 2 distributions in regards to the clusters positioning, first from the center and outwards, and second north and south. As we stated before, the first factor of the distribution, adjust to the normal growth of a city.



The central area is where the city historic quarter is located, and from that point the area has been growing absorbing the surroundings and adapting them to the needs of the time. The blue and purple clusters are residential areas that have been developed recently in order to absorb the housing demand of the city.

The second factor cannot be analyzed using only the data employed in this analysis and is out of the scope of the project. With that said, we considered the clusters separately and compared them to check which one could be underpriced.

As shown in the results we found 10 neighborhoods which are subject to investment according to the study performed. We can also divide this result depending on 2 factors:

- **Max_period:** This variable states the period where the maximum price was reached. In our analysis we can distinguish 2 main groups, neighborhoods where this price peaked 10 years ago and from 2019 to 2020.
- **Year change:** This variable shows the change in the price from the previous year to this one. Is a good indicator on how the people are investing in the neighborhood. The neighborhood price can be positive or negative depending on this.

With all this metrics, it depends on the interest of the stakeholders to invest in one or another neighborhood. If we are talking about families, probably we will need to check the presence of schools and green areas, whereas if we consider big property investors, they will be more interested in an area which is developing.

Conclusion

In this project, we have gone through the process of identifying the business problem, specifying the data required, extracting and preparing the data, performing machine learning by clustering the data \ and lastly providing recommendations to the relevant stakeholders.

We found 10 neighborhoods that are subject of investment due to the price relationship between the other elements of the clusters and given recommendations depending on the metrics employed.

For future research we suggest analyzing in more detail the neighborhoods found, make a market analysis and give a more specific recommendation depending on the stakeholder type.

References

Category:Neighborhoods in Madrid. *Wikipedia*

https://es.wikipedia.org/wiki/Anexo:Barrios_administrativos_de_Madrid

Foursquare Developers Documentation. *Foursquare*

<https://developer.foursquare.com/docs>

Madrid statistics web data <https://datos.madrid.es>

Idealista portal

<https://www.idealista.com/sala-de-prensa/informes-precio-vivienda/venta/madrid-comunidad/madrid-provincia/madrid>

Appendix

Clusters

```
madrid_merged.loc[madrid_merged['Cluster Labels']==0]
```

	Neighborhood	Price	Mensual change	3month change	Year change	Max	Max variation	Max_period	Latitude	Longitude	...	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
6	Alameda de Osuna	3340	0.008	0.020	0.016	3340	0.000	abr 2020	40.45818	-3.58953	...	Hotel	Spanish Restaurant	Restaurant	Tapas Restaurant	Clothing Store
8	Casco Histórico de Barajas	2992	-0.001	-0.035	-0.029	3203	-0.066	dic 2019	40.47482	-3.57951	...	Hotel	Spanish Restaurant	Coffee Shop	Duty-free Shop	Air Line
101	Rejas	3026	0.011	0.001	-0.012	3315	-0.087	ene 2009	40.44629	-3.57489	...	Hotel	Spanish Restaurant	Restaurant	Hotel Bar	Clothing Store

```
madrid_merged.loc[madrid_merged['Cluster Labels']==1]
```

	Neighborhood	Price	Mensual change	3month change	Year change	Max	Max variation	Max_period	Latitude	Longitude	...	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
10	Abrantes	1938	0.015	0.030	0.012	2615	-0.259	may 2010	40.379800	-3.726360	...	Fast Food Restaurant	Grocery Store	Tapas Restaurant	Bar
11	Buena Vista	2042	-0.009	0.030	0.027	2460	-0.170	dic 2010	40.368600	-3.740050	...	Fast Food Restaurant	Spanish Restaurant	Park	Pizza Place
15	Puerta Bonita	1980	0.002	0.020	0.013	2696	-0.265	jun 2009	40.379820	-3.740460	...	Fast Food Restaurant	Bar	Tapas Restaurant	Restaurant
16	San Isidro	2221	-0.004	-0.035	-0.03	2908	-0.236	abr 2009	40.395880	-3.730480	...	Grocery Store	Pizza Place	Park	Fast Food Restaurant
17	Vista Alegre	2019	-0.023	-0.012	-0.015	2528	-0.201	nov 2009	40.384920	-3.746350	...	Tapas Restaurant	Bar	Pizza Place	Spanish Restaurant
45	Fuentealarreina	3311	-0.020	-0.037	-0.044	4296	-0.229	abr 2011	40.480200	-3.739340	...	Athletics & Sports	Pizza Place	Bakery	Bar
48	Mirasierra	3759	-0.002	0.031	0.057	3963	-0.051	mar 2011	40.494270	-3.717690	...	Bar	Tapas Restaurant	Restaurant	Burger Joint
49	Montecarmelo	4576	-0.001	0.000	0.079	4630	-0.012	oct 2019	40.493820	-3.737941	...	Restaurant	Spanish Restaurant	Bar	Athletics & Sports
61	Águilas	2247	0.048	0.057	0.065	2648	-0.152	mar 2010	40.381660	-3.771880	...	Bar	Spanish Restaurant	Italian Restaurant	Fast Food Restaurant
62	Aluche	2246	0.021	0.015	0.031	2658	-0.155	ago 2009	40.392710	-3.760320	...	Bar	Pizza Place	Italian Restaurant	Park
63	Campamento	2296	-0.020	0.027	0.009	2496	-0.080	dic 2010	40.396660	-3.774080	...	Italian Restaurant	Bar	Exhibit	Spanish Restaurant
74	Fontarrón	2196	-0.001	-0.005	0	2500	-0.121	ene 2011	40.403010	-3.649100	...	Bar	Park	Grocery Store	Pizza Place
75	Marroquina	2917	0.011	0.017	0.038	2995	-0.026	jul 2019	40.411760	-3.646180	...	Park	Bar	Pizza Place	Supermarket
77	Pavones	2160	-0.011	-0.027	n.d.	2339	-0.077	jul 2011	40.400040	-3.633000	...	Exhibit	Pizza Place	Park	Fast Food Restaurant
78	Vinateros	2365	0.008	-0.054	-0.015	2600	-0.090	feb 2011	40.404440	-3.640290	...	Fast Food Restaurant	Park	Pizza Place	Spanish Restaurant
79	Entrevías	1541	-0.010	-0.011	0.064	1871	-0.177	jul 2010	40.379250	-3.672120	...	Spanish Restaurant	Pizza Place	Supermarket	Park
80	Numancia	2053	-0.010	-0.014	-0.006	2830	-0.275	oct 2008	40.398510	-3.659010	...	Spanish Restaurant	Bar	Supermarket	Pizza Place
81	Palomeras Bajas	2042	0.007	-0.031	0.068	2569	-0.205	jun 2010	40.387560	-3.660290	...	Spanish Restaurant	Pizza Place	Park	Bar
82	Palomeras sureste	2073	-0.004	0.031	0.058	2784	-0.255	sep 2009	40.385370	-3.635300	...	Spanish Restaurant	Fast Food Restaurant	Bar	Grocery Store
83	Portazgo	1887	0.007	0.033	0.036	2365	-0.202	jul 2010	40.390870	-3.648100	...	Bar	Spanish Restaurant	Fast Food Restaurant	Supermarket
114	Orcasitas	1807	-0.001	-0.002	0.033	2376	-0.240	abr 2011	40.369850	-3.712310	...	Fast Food Restaurant	Bar	Spanish Restaurant	Seafood Restaurant
118	Ambroz	2022	-0.004	-0.034	-0.011	2468	-0.181	ene 2011	40.409420	-3.611350	...	Spanish Restaurant	Pizza Place	Grocery Store	Exhibit

121	Valdebernardo - Valderribas	2692	-0.010	-0.070	-0.036	2895	-0.070	ene 2020	40.400720	-3.621330	...	Spanish Restaurant	Pizza Place	Exhibit	Grocery Store
122	Casco Histórico de Vallecas	1993	-0.002	-0.013	0.015	2538	-0.215	mar 2010	40.362510	-3.617100	...	Clothing Store	Fast Food Restaurant	Supermarket	Restaurant
123	Ensanche de Vallecas - La Gavia	2655	0.001	-0.015	-0.006	3278	-0.190	oct 2010	40.368575	-3.609006	...	Restaurant	Fast Food Restaurant	Supermarket	Clothing Store
124	Santa Eugenia	2144	-0.012	-0.062	0	2550	-0.159	jun 2010	40.381210	-3.609720	...	Spanish Restaurant	Fast Food Restaurant	Gym	Exhibit
126	Los Ángeles	1737	-0.022	-0.057	-0.039	2487	-0.302	abr 2010	40.356410	-3.695980	...	Pizza Place	Spanish Restaurant	Train Station	Fast Food Restaurant
127	Los Rosales	1823	-0.003	-0.024	0.007	2299	-0.207	ago 2010	40.355670	-3.687220	...	Train Station	Pizza Place	Spanish Restaurant	Hotel
129	San Cristóbal	1400	-0.022	-0.045	0.141	1727	-0.190	nov 2010	40.340490	-3.687740	...	Train Station	Pizza Place	Spanish Restaurant	Restaurant

```
madrid_merged.loc[madrid_merged['Cluster Labels']==2]
```

	Neighborhood	Price	Mensual change	3month change	Year change	Max	Max variation	Max_period	Latitude	Longitude	...	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
4	Legazpi	4436	-0.010	0.010	-0.013	4714	-0.059	jun 2019	40.387020	-3.689900	...	Spanish Restaurant	Restaurant	Coffee Shop	Ci
7	Campo de las Naciones-Corralejos	3590	-0.002	0.033	0.089	3901	-0.080	feb 2011	40.460890	-3.615630	...	Spanish Restaurant	Park	Hotel	Restaur
9	Timón	2998	-0.027	-0.029	-0.043	3251	-0.078	jun 2011	40.485680	-3.604720	...	Airport Lounge	Restaurant	Coffee Shop	Span Restaurant
35	Vallehermoso	4640	-0.005	0.021	0.101	4665	-0.005	mar 2020	40.444800	-3.714830	...	Spanish Restaurant	Restaurant	Tapas Restaurant	E
38	Concepción	3189	-0.020	-0.022	-0.019	3474	-0.082	may 2010	40.439730	-3.648500	...	Spanish Restaurant	Bar	Restaurant	Pr
39	Costillares	4195	0.000	-0.017	-0.002	4306	-0.026	jun 2019	40.480410	-3.668270	...	Spanish Restaurant	Restaurant	Park	Supermar
40	Pueblo Nuevo	2530	-0.010	0.014	0.026	3050	-0.170	ene 2009	40.428950	-3.638260	...	Spanish Restaurant	Bar	Restaurant	Supermar
41	Quintana	2847	0.019	0.007	-0.006	3094	-0.080	ago 2010	40.437500	-3.642222	...	Spanish Restaurant	Bar	Café	Pr
43	San Pascual	3532	0.029	0.038	-0.001	4126	-0.144	ago 2019	40.445500	-3.651810	...	Spanish Restaurant	Bar	Park	Ho
44	Ventas	2718	0.010	0.020	0.002	3015	-0.099	mar 2009	40.422380	-3.650200	...	Bar	Spanish Restaurant	Restaurant	Pr
46	La Paz	3476	-0.004	-0.029	-0.055	3825	-0.091	dic 2010	40.481580	-3.694270	...	Spanish Restaurant	Restaurant	Hotel	Itali Restaurant
51	Pilar	3261	0.007	-0.013	-0.011	3333	-0.022	ene 2011	40.475430	-3.711300	...	Spanish Restaurant	Hotel	Restaurant	Pr
52	Tres Olivos - Valverde	2698	0.017	-0.025	0.035	2895	-0.068	dic 2010	40.501095	-3.690618	...	Restaurant	Spanish Restaurant	Bar	Supermar
53	Apóstol Santiago	2890	-0.025	0.014	0.008	2963	-0.025	mar 2020	40.475570	-3.660710	...	Spanish Restaurant	Bar	Restaurant	Pr
56	Palomas	3885	0.009	0.041	0.053	4522	-0.141	jun 2011	40.453830	-3.615010	...	Spanish Restaurant	Restaurant	Hotel	Pr
57	Pinar del Rey	2863	-0.007	-0.008	-0.047	3238	-0.116	nov 2010	40.472250	-3.649430	...	Spanish Restaurant	Park	Bar	Chine Restaur
59	Valdebebas - Valdefuentes	3835	-0.001	0.043	0.044	3873	-0.010	feb 2020	40.482448	-3.616716	...	Spanish Restaurant	Soccer Field	Restaurant	Asi Restaur
60	Virgen del Cortijo - Manoteras	3612	0.039	0.031	0.11	3914	-0.077	oct 2018	40.487810	-3.664190	...	Spanish Restaurant	Bar	Supermarket	Pr
67	Aravaca	3654	-0.008	-0.024	0.005	3742	-0.024	ene 2020	40.456200	-3.780420	...	Spanish Restaurant	Italian Restaurant	Restaurant	Supermar
70	Ciudad Universitaria	3805	-0.001	0.042	-0.014	4207	-0.096	abr 2010	40.447170	-3.730650	...	Spanish Restaurant	Restaurant	Bar	Pr
72	Valdemarín	4336	-0.018	-0.016	0.011	4864	-0.108	mar 2010	40.469960	-3.776410	...	Restaurant	Spanish Restaurant	Hotel	Supermar

73	Valdezarza	3090	0.057	0.034	-0.075	3448	-0.104	mar 2019	40.466660	-3.719740	...	Spanish Restaurant	Bar	Supermarket	Pr
84	San Diego	1927	0.013	-0.001	0.01	3143	-0.387	abr 2008	40.392310	-3.668000	...	Spanish Restaurant	Bar	Italian Restaurant	Restaur
97	Amposta	2015	0.001	-0.028	0.036	2078	-0.031	dic 2019	40.426430	-3.621310	...	Supermarket	Spanish Restaurant	Restaurant	E
98	Arcos	2243	-0.004	0.053	0.095	2421	-0.073	ago 2010	40.421780	-3.617840	...	Spanish Restaurant	Supermarket	Restaurant	Cr
99	Canillejas	2339	-0.009	-0.009	0.062	2604	-0.102	nov 2010	40.443730	-3.609770	...	Spanish Restaurant	Park	Restaurant	Cr
100	Hellín	1992	0.013	0.035	0.056	2112	-0.057	may 2011	40.431640	-3.615500	...	Restaurant	Spanish Restaurant	Supermarket	Ho
103	Salvador	3284	0.035	0.068	-0.077	3967	-0.172	ene 2011	40.443840	-3.634770	...	Spanish Restaurant	Park	Restaurant	E
104	Simancas	2632	0.018	0.012	0.064	3268	-0.194	feb 2009	40.435770	-3.624880	...	Spanish Restaurant	Restaurant	Bar	Supermar
116	San Fermín	2025	0.004	-0.009	0.047	2581	-0.216	sep 2011	40.373690	-3.690260	...	Spanish Restaurant	Grocery Store	Park	Chine Restaur
128	San Andrés	1656	0.001	0.026	0.021	2691	-0.385	ago 2009	40.349220	-3.712110	...	Spanish Restaurant	Clothing Store	Metro Station	Restaur

```
madrid_merged.loc[madrid_merged['Cluster Labels']==3]
```

	Neighborhood	Price	Mensual change	3month change	Year change	Max	Max variation	Max_period	Latitude	Longitude	...	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Co
0	Acacias	3920	-0.018	-0.064	-0.056	4273	-0.083	dic 2018	40.401370	-3.706690	...	Tapas Restaurant	Spanish Restaurant	Art Gallery	Park	Rest
1	Chopera	3578	-0.001	-0.013	-0.047	3843	-0.069	jun 2019	40.395360	-3.698330	...	Restaurant	Art Gallery	Spanish Restaurant	Coffee Shop	
3	Imperial	4099	0.030	0.045	-0.005	4171	-0.017	feb 2019	40.408330	-3.718650	...	Tapas Restaurant	Plaza	Spanish Restaurant	Park	
5	Palos de Moguer	3986	0.028	0.011	0.005	4128	-0.034	dic 2018	40.403010	-3.693580	...	Hotel	Plaza	Art Gallery	Restaurant	Mi
18	Chueca-Justicia	5983	0.021	0.015	-0.003	6120	-0.022	jul 2019	40.421350	-3.698730	...	Plaza	Restaurant	Hotel	Bookstore	Ice
19	Huertas-Cortes	5159	-0.021	-0.021	-0.005	5481	-0.059	jun 2018	40.413968	-3.700581	...	Plaza	Spanish Restaurant	Tapas Restaurant	Hotel	Rest
20	Lavapiés-Embajadores	4455	0.002	-0.005	0.015	4489	-0.008	feb 2020	40.408530	-3.704960	...	Tapas Restaurant	Plaza	Spanish Restaurant	Art Gallery	
21	Malasaña-Universidad	5337	0.008	0.010	0.03	5337	0.000	abr 2020	40.428726	-3.704123	...	Plaza	Restaurant	Hotel	Tapas Restaurant	Si Rest
22	Palacio	4800	0.026	-0.011	-0.041	5073	-0.054	mar 2019	40.415170	-3.712730	...	Tapas Restaurant	Plaza	Spanish Restaurant	Bar	
23	Sol	5003	0.008	-0.021	-0.061	5468	-0.085	ago 2019	40.418020	-3.705770	...	Plaza	Tapas Restaurant	Restaurant	Spanish Restaurant	
30	Almagro	6648	-0.011	-0.012	0.058	6728	-0.012	dic 2019	40.432960	-3.691530	...	Spanish Restaurant	Restaurant	Plaza	Hotel	Jap Rest
31	Arapiles	5027	0.000	0.006	-0.064	5435	-0.075	sep 2018	40.432860	-3.708400	...	Restaurant	Spanish Restaurant	Bookstore	Hotel	
32	Gaztambide	4867	-0.002	-0.006	0.003	5000	-0.027	ago 2019	40.434900	-3.715510	...	Spanish Restaurant	Café	Ice Cream Shop	Bookstore	Rest
34	Trafalgar	5949	0.030	0.068	0.079	5949	0.000	abr 2020	40.433030	-3.700030	...	Restaurant	Tapas Restaurant	Plaza	Spanish Restaurant	
58	Sanchinarro	4419	-0.010	0.000	-0.013	4538	-0.026	mar 2019	40.416500	-3.702560	...	Plaza	Restaurant	Tapas Restaurant	Spanish Restaurant	
68	Argüelles	4957	0.000	0.000	0.01	4970	-0.003	dic 2019	40.429920	-3.718420	...	Plaza	Hotel	Restaurant	Ice Cream Shop	Boc
88	Jerónimos	7515	0.034	0.115	0.227	7515	0.000	abr 2020	40.417290	-3.692230	...	Restaurant	Plaza	Hotel	Art Museum	Si Rest
91	Castellana	6922	0.000	-0.028	-0.011	7120	-0.028	dic 2019	40.434940	-3.683190	...	Plaza	Hotel	Ice Cream Shop	Spanish Restaurant	T

madrid_merged.loc[madrid_merged['Cluster Labels']==4]

Max	Max variation	Max_period	Latitude	Longitude	...	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
2420	-0.001	ene 2020	40.401433	-3.552617	...	Toll Booth	Park	Burger Joint	Bar	Music Venue	Deli / Bodega	Exhibit	Food Court	Food & Drink Shop	Food

madrid_merged.loc[madrid_merged['Cluster Labels']==5]

	Neighborhood	Price	Mensual change	3month change	Year change	Max	Max variation	Max_period	Latitude	Longitude	...	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
119	Casco Histórico de Vicalvaro	2139	0.01	-0.003	-0.017	2441	-0.124	mar 2011	40.38999	-3.56123	...	Mediterranean Restaurant	Toll Booth	Zoo	Fast Food Restaurant	

madrid_merged.loc[madrid_merged['Cluster Labels']==6]

	Neighborhood	Price	Mensual change	3month change	Year change	Max	Max variation	Max_period	Latitude	Longitude	...	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
24	Bernabéu-Hispanoamérica	5238	0.004	0.013	0.013	5238	0.000	abr 2020	40.456880	-3.676980	...	Spanish Restaurant	Restaurant	Tapas Restaurant	Hotel
25	Castilla	4173	-0.007	-0.046	-0.004	4393	-0.050	dic 2018	40.471310	-3.676000	...	Spanish Restaurant	Restaurant	Hotel	Plaza
26	Ciudad Jardín	4678	0.032	0.025	0.053	4820	-0.029	nov 2018	40.450460	-3.667710	...	Spanish Restaurant	Hotel	Restaurant	Japanese Restaurant
27	El Viso	6167	0.006	-0.015	-0.066	6614	-0.068	feb 2019	40.447460	-3.685430	...	Spanish Restaurant	Japanese Restaurant	Restaurant	Burger Joint
28	Nueva España	5448	0.001	0.029	0.007	5663	-0.038	may 2010	40.464120	-3.679800	...	Spanish Restaurant	Restaurant	Hotel	Plaza
29	Prosperidad	4234	-0.007	-0.013	-0.017	4475	-0.054	nov 2018	40.441800	-3.669250	...	Spanish Restaurant	Restaurant	Japanese Restaurant	Coffee Shop
33	Nuevos Ministerios-Ríos Rosas	5049	-0.002	0.010	0.01	5204	-0.030	nov 2018	40.445960	-3.691710	...	Spanish Restaurant	Restaurant	Japanese Restaurant	Gym / Fitness Center
36	Atalaya	4305	0.000	0.000	0.096	4398	-0.021	nov 2018	40.464570	-3.663350	...	Spanish Restaurant	Restaurant	Seafood Restaurant	Indian Restaurant
37	Colina	4244	-0.025	-0.089	-0.102	4752	-0.107	nov 2019	40.458610	-3.660120	...	Spanish Restaurant	Restaurant	Bakery	Plaza
42	San Juan Bautista	4190	-0.012	-0.023	0.01	4407	-0.049	jun 2019	40.451590	-3.655910	...	Spanish Restaurant	Bar	Hotel	Restaurant
47	Las Tablas	4192	0.004	0.008	0.006	4324	-0.030	jul 2019	40.506490	-3.672350	...	Spanish Restaurant	Coffee Shop	Restaurant	Burger Joint
54	Canillas	3242	0.000	0.000	0.005	3692	-0.122	nov 2009	40.464480	-3.643380	...	Spanish Restaurant	Park	Bar	Bakery
55	Conde Orgaz-Piovera	4201	0.006	-0.017	-0.015	5203	-0.193	mar 2009	40.457130	-3.647570	...	Spanish Restaurant	Bakery	Park	Argentinian Restaurant
71	El Plantío	2654	-0.003	0.006	0.001	3410	-0.222	nov 2010	40.468930	-3.820610	...	Spanish Restaurant	Restaurant	Café	Tennis Court
76	Media Legua	2568	0.004	-0.036	-0.113	2894	-0.113	abr 2019	40.410740	-3.659520	...	Spanish Restaurant	Tapas Restaurant	Restaurant	Park
86	Estrella	4114	0.000	-0.010	-0.011	4300	-0.043	ago 2019	40.411170	-3.665930	...	Spanish Restaurant	Restaurant	Tapas Restaurant	Museum
87	Ibiza	5413	0.017	-0.004	0.024	5578	-0.030	jul 2019	40.417940	-3.676260	...	Spanish Restaurant	Restaurant	Hotel	Tapas Restaurant
89	Niño Jesús	4764	-0.004	-0.006	-0.047	5036	-0.054	oct 2019	40.410950	-3.672300	...	Spanish Restaurant	Hotel	Restaurant	Tapas Restaurant
92	Fuente del Berro	4564	0.012	0.020	0.048	4578	-0.003	sep 2018	40.423760	-3.667540	...	Spanish Restaurant	Restaurant	Tapas Restaurant	Boutique
93	Goya	5772	0.002	-0.017	0.002	5970	-0.033	ago 2019	40.425470	-3.674180	...	Spanish Restaurant	Restaurant	Japanese Restaurant	Tapas Restaurant
94	Guindalera	4275	0.014	-0.001	0.027	4439	-0.037	oct 2019	40.432410	-3.666000	...	Spanish Restaurant	Tapas Restaurant	Restaurant	Bar

95	Lista	5944	0.022	0.040	0.006	5944	0.000	abr 2020	40.432310	-3.675080	...	Spanish Restaurant	Tapas Restaurant	Restaurant	Japanese Restaurant
96	Recoletos	8440	-0.020	0.000	0.033	8652	-0.025	feb 2020	40.425300	-3.686510	...	Spanish Restaurant	Restaurant	Tapas Restaurant	Coffee Shop
105	Bellas Vistas	3513	0.002	0.044	-0.024	3688	-0.048	may 2019	40.454570	-3.705520	...	Spanish Restaurant	Japanese Restaurant	Tapas Restaurant	Hotel
106	Berruguete	3253	0.011	-0.003	0.022	3330	-0.023	nov 2009	40.460460	-3.703870	...	Spanish Restaurant	Restaurant	Bar	Supermarket
107	Cuatro Caminos	4344	0.002	-0.004	-0.002	4480	-0.030	oct 2018	40.452970	-3.697270	...	Spanish Restaurant	Tapas Restaurant	Restaurant	Japanese Restaurant
108	Cuzco-Castillejos	4251	0.000	0.000	-0.007	4399	-0.034	jun 2019	40.459089	-3.690612	...	Spanish Restaurant	Restaurant	Japanese Restaurant	Pizza Place
109	Valdeacederas	3252	0.010	0.050	0.027	3541	-0.082	oct 2008	40.465930	-3.700960	...	Spanish Restaurant	Restaurant	Hotel	Seafood Restaurant
110	Ventilla-Almenara	3383	-0.015	-0.010	-0.014	3494	-0.032	feb 2020	40.468400	-3.693160	...	Spanish Restaurant	Restaurant	Japanese Restaurant	Hotel

madrid_merged.loc[madrid_merged['Cluster Labels']==7]

	Neighborhood	Price	Mensual change	3month change	Year change	Max	Max variation	Max_period	Latitude	Longitude	...	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
2	Delicias	3809	0.008	-0.024	-0.092	4193	-0.092	abr 2019	40.39613	-3.68946	...	Plaza	Art Gallery	Restaurant	Café	
12	Comillas	2634	-0.002	-0.025	-0.011	2833	-0.070	may 2010	40.39495	-3.70976	...	Tapas Restaurant	Spanish Restaurant	Art Gallery	Park	
13	Opañel	2405	-0.001	0.028	0.049	3045	-0.210	mar 2009	40.38915	-3.72375	...	Park	Spanish Restaurant	Coffee Shop	Fast Food Restaurant	
14	Pau de Carabanchel	2998	-0.006	-0.009	-0.019	3603	-0.168	jul 2010	40.39094	-3.72420	...	Park	Coffee Shop	Clothing Store	Fast Food Restaurant	Re
50	Peñagrande	3228	0.003	-0.017	-0.007	3428	-0.059	abr 2010	40.47839	-3.72735	...	Spanish Restaurant	Park	Bakery	Restaurant	Re
64	Los Cármenes	2662	-0.031	0.110	0.136	2780	-0.042	nov 2010	40.40280	-3.73178	...	Theme Park Ride / Attraction	Park	Spanish Restaurant	Tapas Restaurant	
66	Puerta del Ángel	2745	0.018	0.099	0.085	2951	-0.070	sep 2009	40.40886	-3.73392	...	Park	Theme Park Ride / Attraction	Spanish Restaurant	Tapas Restaurant	
85	Adelfas	4242	0.007	0.007	-0.025	4353	-0.025	may 2019	40.40173	-3.67288	...	Spanish Restaurant	Park	Restaurant	Museum	
90	Pacífico	4121	-0.018	-0.007	0.033	4197	-0.018	oct 2019	40.40191	-3.67603	...	Spanish Restaurant	Tapas Restaurant	Restaurant	Park	
102	Rosas	3114	0.012	-0.002	-0.007	3394	-0.082	jul 2010	40.42869	-3.60086	...	Park	Sandwich Place	Soccer Field	Chinese Restaurant	Re
111	12 de Octubre-Orcasur	1578	-0.053	-0.087	-0.059	2466	-0.360	mar 2012	40.37771	-3.69891	...	Spanish Restaurant	Grocery Store	Bar	Park	F. Re
112	Almendrales	2089	-0.001	-0.042	0.075	2730	-0.235	mar 2010	40.38428	-3.69991	...	Park	Spanish Restaurant	Bar	Restaurant	
113	Moscardó	2322	-0.023	0.051	-0.001	2420	-0.041	oct 2010	40.38888	-3.70666	...	Spanish Restaurant	Park	Restaurant	Art Gallery	
115	Pradolongo	2118	0.000	0.010	0.01	2415	-0.123	nov 2010	40.38297	-3.70865	...	Spanish Restaurant	Park	Bar	Grocery Store	
117	Zofio	1821	-0.017	-0.036	-0.089	2319	-0.215	oct 2010	40.37987	-3.71495	...	Bar	Park	Grocery Store	Spanish Restaurant	
125	Butarque	2182	0.015	0.018	0.022	2472	-0.117	jun 2010	40.33935	-3.76848	...	Tapas Restaurant	Spanish Restaurant	Park	Supermarket	

madrid_merged.loc[madrid_merged['Cluster Labels']==8]

	Neighborhood	Price	Mensual change	3month change	Year change	Max	Max variation	Max_period	Latitude	Longitude	...	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
65	Lucero	2353	0.008	-0.012	0.017	2727	-0.137	may 2010	40.40358	-3.74960	...	Theme Park Ride / Attraction	Grocery Store	Pizza Place	Restaurant	Pa
69	Casa de Campo	3586	-0.001	-0.032	-0.029	3944	-0.091	abr 2018	40.41336	-3.75982	...	Theme Park Ride / Attraction	Exhibit	Restaurant	Tapas Restaurant	Piz: Plac

