REPORT

1.- Introduction/Business Problem

In different countries of the world and in particular in the United States and Canada, there are more and more young people interested in learning other languages, such as Spanish.

Many of them, either individually or with their partners, would like to make the decision to travel, live and work in a country where Spanish is spoken to learn to speak the language well, understand its idioms and also learn about the local culture, history and customs.

Basically this experience could be had in Spain, or in a Latin American country (except Brazil where Portuguese is spoken). Usually Mexico City attracts the attention of young people, but also Madrid, which is the capital of Spain. Unfortunately, information of interest to young people is not always available regarding what these cities offer, which are the best neighborhoods and what services or places of interest they have nearby.

Following these lines, it is proposed to make a comparison between **Madrid** (capital of Spain) and **Mexico City** (capital of Mexico) segmenting and clustering the neighborhoods of both cities, in order to be able to provide information regarding their characteristics and the services they offer. In particular, neighborhoods where there is a greater presence of businesses will be differentiated from those closest to green areas and from others that might offer a more important presence of restaurants, cafes or similar venues.

The idea is to be able to find those particular neighborhoods in the city that offer the closest proximity to these 3 types of services at the same time, thinking that potential travellers might be interested in being close to places where they can find a job, but also where they might be able to have fun and be close to parks.

The abovementioned information and its analysis will be delivered through a web page written in English and linked to Google so that people who speak English and who are interested in studying Spanish and living abroad for some time are able to reach it. It is also possible to contact youth organizations or local governments in the United States and Canada, and share with them the resulting information, so that they can send it respectively to their contacts or disseminate among interested citizens.

2.- Data

A comparison of neighborhoods and segmentation of Mexico City and Madrid will be made, using the Foursquare application, from which those neighborhoods that may meet the conditions sought will be analyzed; particularly neighborhoods with an important presence of businesses, others with cafes and restaurants, as well as neighborhoods close to parks and green areas. Also those that meet all these criteria at the same time or are located at a short distance from these services will be highlighted.

Concerning other sources of information, and knowing that young people may also be interested in learning more about the culture and history of the country, the location of the most important museums and other important landmarks of the city will be identified, which may influence the choice of the more convenient neighborhoods.

Other data of interest will be added, such as country information: GDP income, unemployment level, and other demographic indicators.

3.- Methodology

The same methodology was applied to both Mexico City and Madrid. The analysis carried out on Madrid is explained below:

The analysis began by looking for information regarding the postal codes of Madrid, with the aim of defining its Boroughs and Neighborhoods. This information is available on Wikipedia, from where it was possible to extract the data manually, since there were not so many records. This also allowed the information to be correctly formatted from the start, in a text file. It was necessary to add, for each neighborhood, its Latitude and Longitude, which was obtained from Googlemaps. In some cases, the Googlemaps map a polygon that represents the neighborhood and then the central point of the polygon was taken as latitude and longitude. In the case of not having a polygon, an important point of interest was taken in each Neighborhood, which was indicated by Googlemaps. Finally, everything was recorded in a text file with UTF-8 encoding to avoid problems on GitHub or with reading the text file from the Jupyter Notebook. The information fields of the text file are:

Postal Code – Borough – Neighborhood – Latitude -Longitude

This is a screen shot of the data file in csv format:

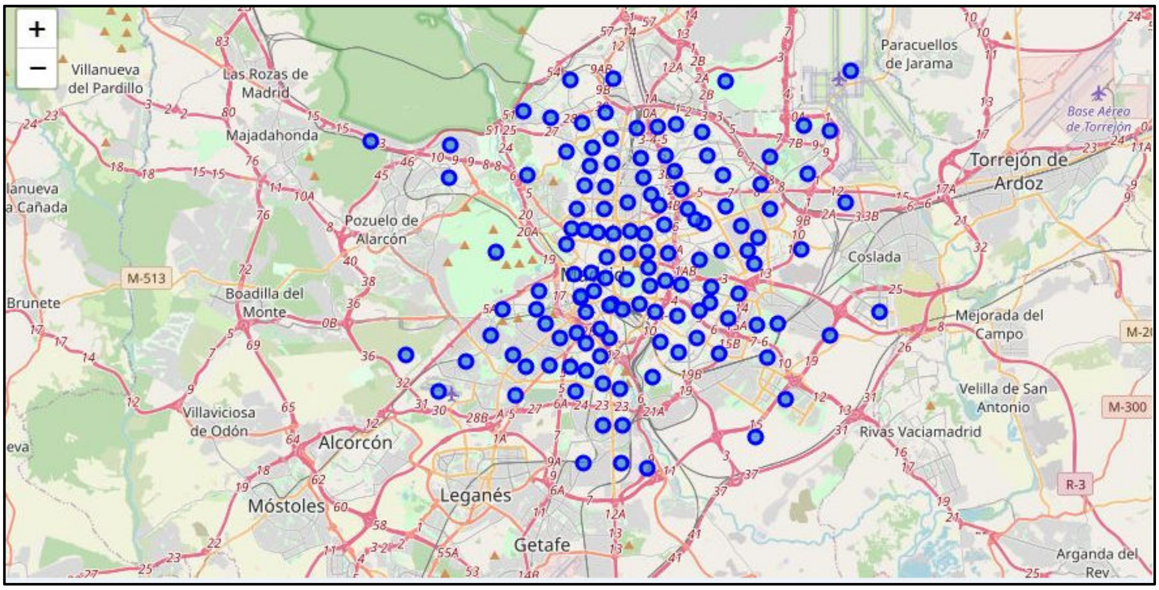
```
Postal Code,Borough,Neighborhood,Latitude,Longitude
11,Centro,Palacio,40.41651402387374,-3.7137077865943455
12,Centro,Embajadores,40.409868731536754,-3.7029723072372907
13,Centro,Cortes,40.41522608259914,-3.697128362755262
14,Centro,Justicia,40.4236599288787,-3.6964691043112525
15,Centro,Universidad,40.40707815787877,-3.709642073861415
16,Centro,Sol,40.41722752794292,-3.704298918157542
21,Arganzuela,Imperial,40.407818340442454,-3.7096733891967735
22,Arganzuela,Acacias,40.40153919179433,-3.707399215686851
23,Arganzuela,Chopera,40.39450792731229,-3.699684673645666
24,Arganzuela,Legazpi,40.3912349877538,-3.6951792614786796
25,Arganzuela,Delicias,40.4046982190728,-3.693278616184685
26,Arganzuela,Palos de Moguer,40.4041349957937,-3.6949432253024352
27,Arganzuela,Atocha,40.40232880453693,-3.6881925141221403
31,Retiro,Pacifico,40.40485102550965,-3.678962689512847
32,Retiro,Adelfas,40.40139501275158,-3.6706887969224606
33,Retiro,Estrella,40.41404196931364,-3.6655241572349486
34,Retiro,Ibiza,40.419355120444095,-3.674507699109315
```

A Jupyter Notebook was prepared to bring this text file to a Pandas Dataframe of the following type:

	Postal Code	Borough	Neighborhood	Latitude	Longitude
0	11	Centro	Palacio	40.416514	-3.713708
1	12	Centro	Embajadores	40.409869	-3.702972
2	13	Centro	Cortes	40.415226	-3.697128
3	14	Centro	Justicia	40.423660	-3.696469
4	15	Centro	Universidad	40.407078	-3.709642

This Pandas has a total of 21 Boroughs and 131 Neighborhoods.

The map of Madrid was then prepared with the Neighborhoods superimposed. The following result was obtained:



Next, an API was prepared to connect with the FourSquare Service, which made it possible to send a request with information about the venues for one of the neighborhoods, to see how everything worked before performing an analysis for all the neighborhoods. The request returned a .json file with the response. As a next step, a function was prepared to extract the category of the venues from that .json file and the .json file was also cleaned and structured to leave it as a Pandas Dataframe, which is shown below. It contains the information of the venues closest to one of the selected neighborhoods:

	name	categories	lat	lng
0	Club del Gourmet Corte Ingles	Gourmet Shop	40.417497	-3.704686
1	LUSH	Cosmetics Shop	40.419012	-3.704898
2	Hotel Liabeny	Hotel	40.418742	-3.703996
3	Kiko store	Cosmetics Shop	40.420128	-3.702317
4	TAKOS	Mexican Restaurant	40.418938	-3.703748

In this case, a total of 80 venues were obtained for one of the neighborhoods.

With this base, the exploration of all the Neighborhoods of Madrid began, for which a function was created to repeat the same previous process, but in all the Neighborhoods. The result was a Pandas Dataframe of 3439 rows with the following structure:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Palacio	40.416514	-3.713708	Palacio Real de Madrid	40.417940	-3.714259	Palace
1	Palacio	40.416514	-3.713708	Santa Iglesia Catedral de Santa María la Real ...	40.415767	-3.714516	Church
2	Palacio	40.416514	-3.713708	Plaza de La Almudena	40.416320	-3.713777	Plaza
3	Palacio	40.416514	-3.713708	Plaza de Oriente	40.418326	-3.712196	Plaza
4	Palacio	40.416514	-3.713708	Zuccaru	40.417179	-3.711674	Ice Cream Shop

Next, it was identified how many unique categories of venues could be found. A total of 277 unique categories were ridentified. With this, a new Pandas Dataframe of 3439 rows (one for each venue) and with the 277 unique categories was prepared to identify the type of venue.

From there, the venues of the previous Dataframe were grouped for each neighborhood, obtaining a Dataframe with a row for each neighborhood and 277 columns of unique categories. The value in each column indicated the average frequency of occurrence of each venue for each neighborhood. Next, a snapshot of the resulting dataframe is shared:

	Neighborhood	Yoga Studio	Accessories Store	Adult Boutique	Airport Gate	Airport Lounge	Airport Service	American Restaurant
0	Abrantes	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.0
1	Acacias	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.0
2	Adelfas	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.0
3	Aeropuerto	0.0	0.0	0.0	0.058824	0.294118	0.058824	0.0
4	Alameda de Osuna	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.0
...	...	...	...	...	...	...	...	...
122	Ventas	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.0
123	Villaverde Alto	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.0
124	Vinateros	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.0
125	Vista Alegre	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.0
126	Zofio	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.0

With this result, for each neighborhood it is possible to obtain the venues with the highest frequency, according to the following structure:

```
----Abrantes----
                    venue  freq
0                    Park  0.29
1           Metro Station  0.14
2             Pizza Place  0.14
3  Gym / Fitness Center  0.14
4     Athletics & Sports  0.14


----Acacias----
                  venue  freq
0  Spanish Restaurant  0.11
1                   Bar  0.09
2           Supermarket  0.06
3     Tapas Restaurant  0.06
4           Pizza Place  0.06


----Adelfas----
                    venue  freq
0     Spanish Restaurant  0.08
1            Supermarket  0.08
2          Grocery Store  0.08
3                    Bar  0.05
4  Fast Food Restaurant  0.05
```

This information was also put into a Pandas Dataframe to sort the top 10 venues in descending order for each Neighborhood:

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
0	Abrantes	Park	Pizza Place	Athletics & Sports	Metro Station	Fast Food Restaurant	Gym / Fitness Center
1	Acacias	Spanish Restaurant	Bar	Tapas Restaurant	Supermarket	Pizza Place	Asian Restaurant
2	Adelfas	Spanish Restaurant	Grocery Store	Supermarket	Bar	Fast Food Restaurant	Coffee Shop
3	Aeropuerto	Airport Lounge	Spanish Restaurant	Coffee Shop	Sporting Goods Shop	Breakfast Spot	Border Crossing
4	Alameda de Osuna	Smoke Shop	Pizza Place	Bakery	Cocktail Bar	Plaza	Chinese Restaurant

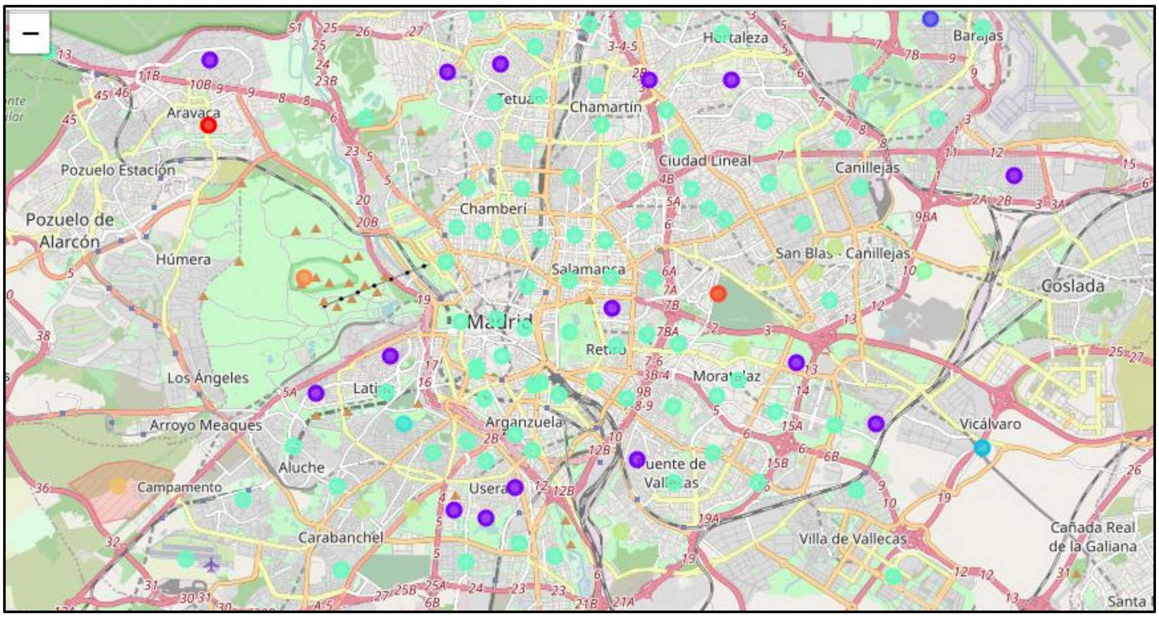
After finishing the abovementioned analysis, all was set to carry out the Clustering of Neighborhoods. A priori it was decided to start a k-means clustering with 12 clusters. The following array was obtained with the assignment of each Neighborhood to a specific cluster.

```
array([ 7,  5,  5,  5,  5,  5,  5,  0,  5,  7,  7,  5, 11,  5,  5,  0,  5,
        5,  5,  5,  7,  8,  0,  5,  9,  5,  3,  5,  5,  5,  5,  5,  5,  5,
        5,  5,  5,  5,  5,  5,  5,  5,  2,  5,  5,  5,  5,  7,  5,  5,  5,
       10,  5,  5,  5,  7,  5,  0,  0,  5,  5,  5,  0,  5,  5,  5,  7,  5,
        5,  0,  7,  5,  5,  5,  5,  5,  5,  5,  5,  5,  5,  5,  5,  5,  5,
        5,  5,  5,  5,  5,  5,  5,  0,  5,  7,  7,  0,  5,  5,  0,  5,  6,
        5,  5,  0,  5,  4,  5,  5,  5,  5,  5,  1,  5,  5,  0,  5,  0,  0,
        0,  5,  5, 10,  5,  5,  5,  0], dtype=int32)
```

This allows the creation of a new dataframe that includes the cluster for each Neighborhood, in addition to the top 10 venues, according to the following structure:

	Postal Code	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue
0	11	Centro	Palacio	40.416514	-3.713708	5.0	Plaza	Tapas Restaurant
1	12	Centro	Embajadores	40.409869	-3.702972	5.0	Hostel	Plaza
2	13	Centro	Cortes	40.415226	-3.697128	5.0	Restaurant	Plaza
3	14	Centro	Justicia	40.423660	-3.696469	5.0	Restaurant	Hotel
4	15	Centro	Universidad	40.407078	-3.709642	5.0	Bar	Tapas Restaurant

Finally, it is possible to obtain a map with the neighborhoods assigned to different clusters:



And also obtain a Dataframe with the information of the Neighborhoods in each of the 12 Clusters. As an example, the result is shown for Cluster 0, which in this case could be called “**Spanish Restaurants**”.

	Borough	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
16	Retiro	0.0	Spanish Restaurant	Restaurant	Tapas Restaurant	Italian Restaurant	Seafood Restaurant
35	Tetuan	0.0	Spanish Restaurant	Gym / Fitness Center	Art Studio	Bookstore	Restaurant
47	Fuencarral-El Pardo	0.0	Spanish Restaurant	Italian Restaurant	Pizza Place	Food & Drink Shop	Tapas Restaurant
53	Moncloa-Aravaca	0.0	Spanish Restaurant	Soccer Field	Asian Restaurant	Supermarket	Women's Store
54	Moncloa-Aravaca	0.0	Spanish Restaurant	Bar	Arcade	Women's Store	Fish & Chips Shop
58	Latina	0.0	Spanish Restaurant	Grocery Store	Park	Supermarket	Paella Restaurant
59	Latina	0.0	Pizza Place	Spanish Restaurant	Tapas Restaurant	Grocery Store	Asian Restaurant
74	Usera	0.0	Spanish Restaurant	Seafood Restaurant	Gastropub	Bar	Bakery
76	Usera	0.0	Spanish Restaurant	Beer Garden	Athletics & Sports	Park	Fish & Chips Shop

Also the result for Cluster 7 that could be called "Park"

	Borough	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
68	Carabanchel	7.0	Plaza	Pizza Place	Bar	Park	Falafel Restaurant
70	Carabanchel	7.0	Park	Pizza Place	Athletics & Sports	Metro Station	Fast Food Restaurant
78	Puente de Vallecas	7.0	Grocery Store	Business Service	Music Venue	Park	Gym / Fitness Center
86	Moratalaz	7.0	Bar	Brewery	Park	Nightlife Spot	Dessert Shop
91	Ciudad Lineal	7.0	Park	Metro Station	Supermarket	Bar	Fast Food Restaurant
103	Hortaleza	7.0	Grocery Store	Park	Soccer Field	Metro Station	Women's Store
106	Villaverde	7.0	Plaza	Sports Bar	Grocery Store	Park	Falafel Restaurant
108	Villaverde	7.0	Spanish Restaurant	Bar	Grocery Store	Park	Women's Store
115	San Blas-Canillejas	7.0	Snack Place	Park	Music Venue	Gym	Women's Store
116	San Blas-Canillejas	7.0	Park	Food	Snack Place	Spanish Restaurant	Metro Station

Finally, the above information was complemented with other data of interest, such as the location of important landmarks (information retrieved from Google Maps), as well as general information about the country such as per capita income or unemployment level, among others. which were obtained from Wikipedia.

4.- Results

In the case of Madrid, the result of the Clustering yields the following thematic organization of the Boroughs and some of their Neighborhoods:

Cluster Number	Cluster Type/Name	Boroughs	Neighborhood
11	Athletics / Gym	**Moncloa-Aravaca**	56
10	Soccer Field	Fuencarral-El Pardo Ciudad Lineal	44 90
9	Trail	**Moncloa-Aravaca**	50
8	Golf Course	Latina	61
7	Park	Carabanchel Puente de Vallecas Moratalaz Ciudad Lineal Hortaleza Villa Verde **San Blas-Canillejas**	68, 70 78 86 91 103 106,108 115-116
6	Electronics Stores	**San Blas Canillejas**	118
5	Hotels and Restaurants	Centro **San Blas-Canillejas** Barajas	0,1,2,3,4 121 122-123-124-126
4	Bed & Breakfast	Carabanchel	66
3	Mediterranean Restaurant	Vicalvaro	111
2	Canal Lock	Fuencarral-El Pardo	43
1	Bakery	Barajas	125
0	Spanish Restaurant	Retiro Tetuán	16 35

		Fuencarral-El Pardo	47
		Moncloa-Aravaca	53-54
		Latina	58-59
		Usera	74,76,77
		Puente de Vallecas	79
		Moratalaz	85
		Ciudad Lineal	97
		Hortaleza	101
		Vicalvaro	113
		San Blas-Canillejas	119

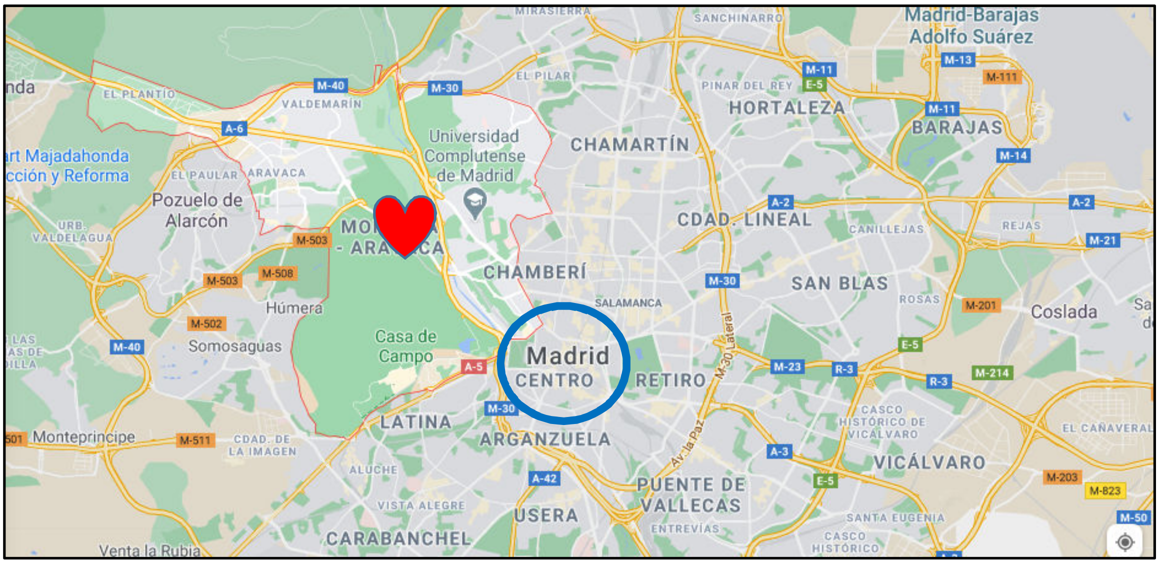
From the previous data, it is possible to see that some Boroughs are repeated in several categories, which is precisely what we want to look for, Boroughs that can offer different services and entertainment at the same time, for someone who wants to live in the city for a while, with the goal of learning Spanish. The Boroughs that stand out and their categories are as follows:

Moncloa-Aravaca: Athletics / Gym; Trail; Spanish Restaurant

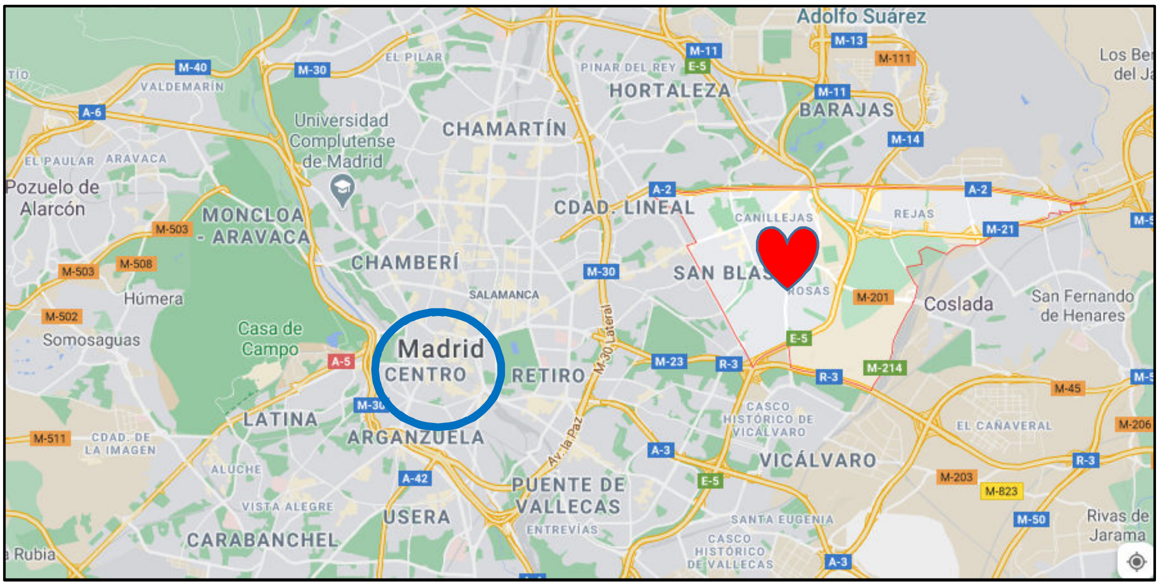
San Blas- Canillejas: Parks, Electronics Stores, Hotels and Restaurants, Spanish Restaurants

Both Boroughs look interesting, however the analysis was complemented to decide which one of them is the more convenient.

According to Googlemaps, this is where **Moncloa-Aravaca** is located (signaled with a red heart on the map of Madrid). As a first comment, it is a sector of the city that is very close to the city center (blue circle) and it is also worth mentioning that there are universities in the area.



Next, this is where **San Blas-Canillejas** is located (signaled with a red heart on the map of Madrid). As a first comment, it is an area that, although close to the airport, it is unfortunately farther from the city center (blue circle).



By enlarging the map around both places, the proximity to other landmarks was analyzed. The result was the following:

Moncloa-Aravaca

- Hospital
- Complutense University of Madrid
- Polytechnic University of Madrid

- The Museum of America
- Madrid Amusement Park
- Very close to the downtown area of Madrid.

San Blas-Canillejas

- Parks
- Airport

Additionally, Both Boroughs are equidistant from the new Chamartín Business District.

As a result of the previous analysis, Moncloa-Aravaca is proposed as the best candidate to live in Madrid for people who want to study Spanish. Besides, the Google Maps description for the Moncloa-Aravaca Borough is the following:

"The tree-lined Moncloa-Aravaca neighborhood is home to the large Casa de Campo park, where the Cable Car takes visitors a few steps from the Madrid Zoo Aquarium and the Amusement Park. Quiet terraced cafes line the Manzanares River and the Parque del Oeste promenade. In the afternoon, crowds enjoy sunset views from the Egyptian Temple of Debod. The students hang out in the bars near the Moncloa universities. "

In addition, the *Life Quality Score for Madrid* should also be considered (Source: https://teleport.org/cities/madrid/). In this analysis such criteria as Safety, Healthcare, Leisure & Culture and Tolerance stand out, which may be important aspects to consider.

With respect to other complementary information, we have:

- GDP per capita : **USD 33,711 (2019)**
- Unemployment level in Spain : **13.96% (2019)**
- Inhabitants of the city of Madrid : **6.6 million (2019)**
- Number of Inhabitants in Moncloa- Aravaca: **116,903 (2017)**

In the case of Mexico City, the result of the Clustering yields the following thematic organization for the Boroughs and some of their Neighborhoods:

Cluster Number	Cluster Type/Name	Boroughs	Neighborhood
11	Office	**Coyoacán**	21,22
10	Mixed		
9	Pet Store	Azcapotzalco	9
8	History Museum	**Miguel Hidalgo**	53
7	Mobile Phone Shops	**Del Iztapalapa**	45
6	Mixed venues		
5	Flea Market	Tiahuac	67
4	Mixed venues		
3	Gym /Sports	**Iztacalco** **Del. Iztapalapa**	39 49
2	Coffee Shop / Bakery	**Del. Iztapalapa**	47
1	Rental Car	Azcapotzalco	4
0	Mexican Restaurant / Taco Place	Alvaro Obregon Azcapotzalco Benito Juárez Coyoacán Cuauhtemoc Gustavo a. Madero **Iztacalco** **Del. Iztapalapa** Mag. Contreras **Miguel Hidalgo** Venustiano Carr.	1,2 5,6 16 20 29 35 37, 38, 40 42, 43, 46, 48 50, 51 54 70, 73

From the previous data, it is possible to see that some Boroughs are repeated in several categories, which is precisely what we want to look for: Boroughs that can offer different services and entertainment at the same time, which is deemed important for someone wanting to live in the city for a while, with the goal of learning Spanish.

The Boroughs that stand out and their categories are as follows:

Iztapalapa: Mobile Phone Shops; Gym-Sports; Coffee Shops-Bakeries; Mexican Restaurant.

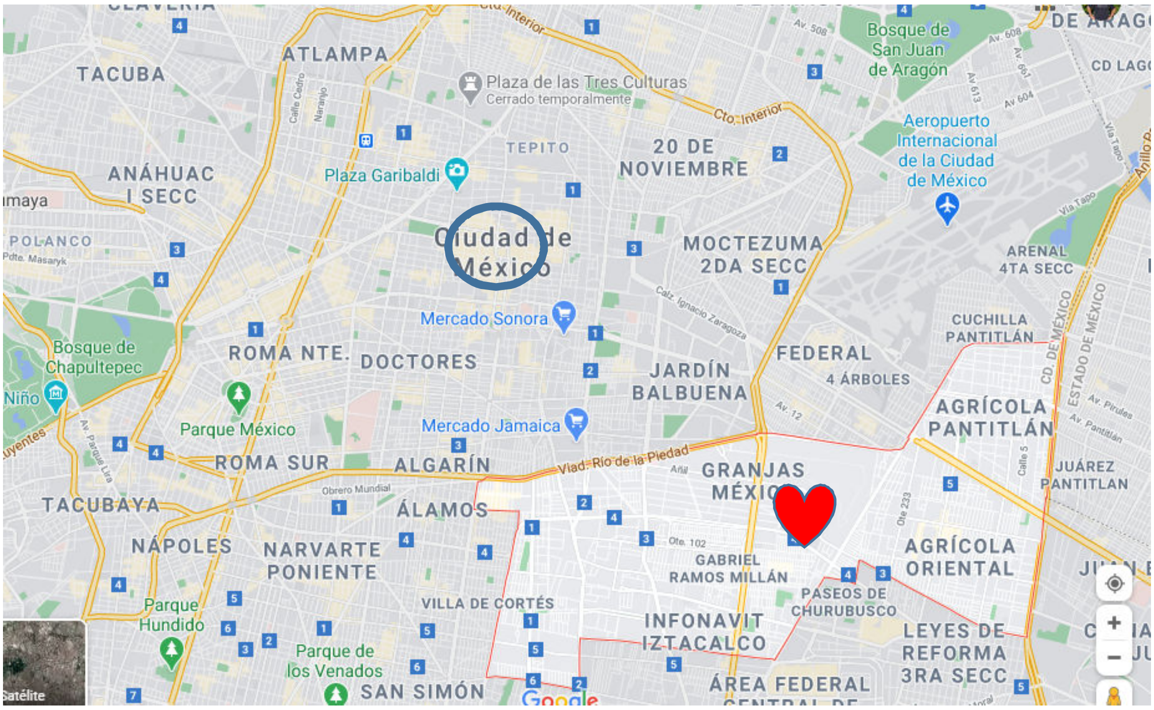
Iztacalco: Gym-Sports; Mexican Restaurant

Miguel Hidalgo: Mexican Restaurant, History Museum

Of the above, the Borough with the most services is **Iztapalapa**, however it is necessary to complement the analysis to decide which one is better.

According to Googlemaps, this is where **Del. Iztapalapa** is located (signaled with a red dot on the map of Mexico City). As a first comment, it is an area of the city that is very far from the city center (Blue Circle).

According to Googlemaps, this is where **Iztacalco** is located (signaled with a red heart on the Map of Mexico City). As a first comment, it is a Borough that is quite close to the city center.



Finally, according to Googlemaps, this is where **Miguel Hidalgo** is located (signaled with a red heart on the map of Mexico City). As a first comment, it is a Borough very close to the city center, which also has a museum and a large park.

By enlarging the map, the proximity to other landmarks was analyzed. The result was as following:

Iztacalco

- Food Market

Miguel Hidalgo

- Anthropology National Museum
- Park
- Spanish Hospital
- Military Center Hospital
- Soumaya Museum
- Hippodrome

As a result of the previous analysis, **Miguel Hidalgo** is proposed as the best candidate to live in Mexico City for people who want to study Spanish and learn about history, culture and also enjoy free time.

The description given by Google Maps of Miguel Hidalgo is as follows:

"Miguel Hidalgo mixes exclusive areas like Polanco, characterized by its sophisticated gastronomy and luxury brand stores, with simple residential districts like Escandon. The elegant Soumaya Museum houses works by great masters and sculptures by Rodin, while the National Museum of Anthropology displays artifacts from the Mayan period onwards. Within the extensive Bosque de Chapultepec park, the top of Castillo de Chapultepec (Chapultepec Castle) offers views of the entire city. "

In addition, the *Life Quality Score for Mexico City* should also be considered (Source: https://teleport.org/cities/mexico-city/). In this analysis such criteria as Living Costs as well as Leisure & Culture may be important aspects to consider.

With respect to other complementary information, we have:

- GDP per capita : **USD 17,881 (2019)**
- Unemployment level in Spain : **5% (2019)**
- Inhabitants of the city of Madrid : **8.8 million (2015)**
- Number of Inhabitants in Moncloa- Aravaca: **364,000 (2015)**

5.- Discussion

Based on the data provided in the previous section, a Borough of Madrid (Moncloa-Aravaca) and another of Mexico City (Miguel Hidalgo) were suggested for those who want to live and work for a time in a capital city where Spanish is spoken.

It is not easy to compare countries or cities. Rather, the analysis carried out shows the services and activities that a person could access in any of the two places indicated, but it really is difficult to say which is better.

Everything will depend on the personal taste. On the one hand, if someone prefers travelling to a developed country located in the old continent with a higher per capita income and with better Safety, Healthcare, Leisure & Culture and Tolerance scores, probably Madrid should be chosen (and the Moncloa-Aravaca Borough). On the other hand, if someone would rather visit a country closer to the United States, a developing country, with a lower per capita income, then Mexico City and the Miguel Hidalgo Borough should be chosen. Notwithstanding the above, it is clear that Madrid has more category wins according to the Life Quality Score.

If someone wants to consider other variables to compare both cities in other aspects, the website https://teleport.org/compare/madrid-and-mexico-city/ provides additional data to make a comparison. Here is the information for Madrid and Mexico City:

	Madrid	**Mexico City**
Small apartment median rent/month	$870	**$520**
Medium apartment median rent/month	$1.200	**$710**
Cost of broadband internet connection	$35	**$21**
Cost of movie ticket	$11	**$4**
Monthly fitness club membership	**$46**	$64
Monthly Public Transport	$71	**$18**
Traffic! (the higher the index, the better)	**0.68**	0.24
Lunch	$13	**$6**
Air quality	**0.31**	0.15
Cleanliness	**0.51**	0.28

In aspects related to Leisure and Culture, both cities are quite similar, and that is probably the main reason that makes it difficult to choose one city over the other:

Venue	Index score Madrid	Index Score Mexico City
Museums	0.88	1.00
Historical Sites	0.99	0.99
Sport venues	0.82	1.00
City Ranking	10/264	4/264

6.- Conclusion

The objective of defining the best Boroughs to live, study and learn about the culture, as well as the language, was carried out for the capital cities of two Spanish-speaking countries: Madrid (Spain) and Mexico City (Mexico). It is not easy to compare countries, nor cities, however information was provided as to allow making the best choice considering personal preferences. Without a doubt, both countries and cities deliver unique experiences.

Most of the analysis was carried out thanks to Data Science and the use of Python, Jupyter notebooks and data frames, as well as clustering. The use of different or more diverse or advanced tools might provide more insight into the the question that was analysed.