Is Buenos Aires a European-like city?

Author: Fabricio Bottalo

July 2021

1.Introduction

1.1 Background

Argentina is a country that has been populated by European immigration mainly, in different waves during the XIX and XX centuries. According to some studies, up to 90% of its population has been linked with European descendants. Nowadays, the immigrants in Argentina only represents approximately 5% of the population, and its origin has migrated from Europe to the neighbor countries. Paraguay, Bolivia, Chile and Perú are the principal origin of the current migrants.

Some studies declare that 63% of the population of Argentina is descendant for at least one Spanish immigrant, and other 60% is it for at least one Italian. So, there is a visible impact of those two countries immigrants in the Argentina's demography. Also, it had a big impact in its culture. Mainly in the capital, Buenos Aires, people have a lot of Spanish or Italian manners, architecture, foods, or slangs. There is a typical saying of the city, that it is said that "it is a part of Europe in South America". So, as the manners could be like ones of the occidental European cultures, the cities should be alike, in terms of way of living, beyond the differences of the economic impact. It would be interesting to see if that is correct, by analyzing the type venues that a city offers. It is also possible that different neighborhoods of Buenos Aires could be like different countries of Europe, as the immigrants, and their cultures, were not perfectly distributed between the cities.

1.2 Objective

This project aims to analyze the venues that each of the neighborhood in Buenos Aires offer and conclude if there are like some European neighborhood or they are more South American. Analyzing every neighborhood of the city could tell us the impact of the immigration of previous generations.

This will be done by doing a classification analysis by separating each location by the clustering method of K-means algorithm.

2.Data acquisition and cleaning

2.1 Data sources

For this project, I will use the data of Foursquare to obtain the different venues of each neighborhood, and the database of the neighborhoods present in Buenos Aires from the cities web-site: https://data.buenosaires.gob.ar/dataset/barrios. The late one contains information about the location, zip codes and names of the neighborhoods.

For the geolocalization of the cities in the European and American continent, I used Google Maps information.

2.2 Data selection

As the Foursquare API gives information about the different venues near the coordinates that you input, I decided to clean the types of venues to considerate. As a previous test let me realize, Buenos Aires' boroughs most common venues were "Argentinian Restaurants" by far.

This creates a great problem with classification, because of how I was dealing with the data. As we only get 100 venues for each borough or city, I use a percentage as a sample of the frequency of each venue for the location. With Argentinian Restaurants, the mean of this kind of venue for each of the Comunas (Buenos Aires boroughs), were significantly bigger that any other venue for other cities. This made difficult to make any other type of analysis, as the classification method of clustering will always group those locations in a group aside the others.

So, deleting this variable, make a richer analyze as enables to see how similar the rest of venues between the cities and Buenos Aires boroughs were.

3. Data analysis

3.1 Locations considerer

For this project, I took all the 15 boroughs of Buenos Aires, which are called "Comunas", and are numbered. As they have different cultural and economic aspect, the hypothesis would be that some may be more like a European city, as most of them could be more like the Latin-American cities.

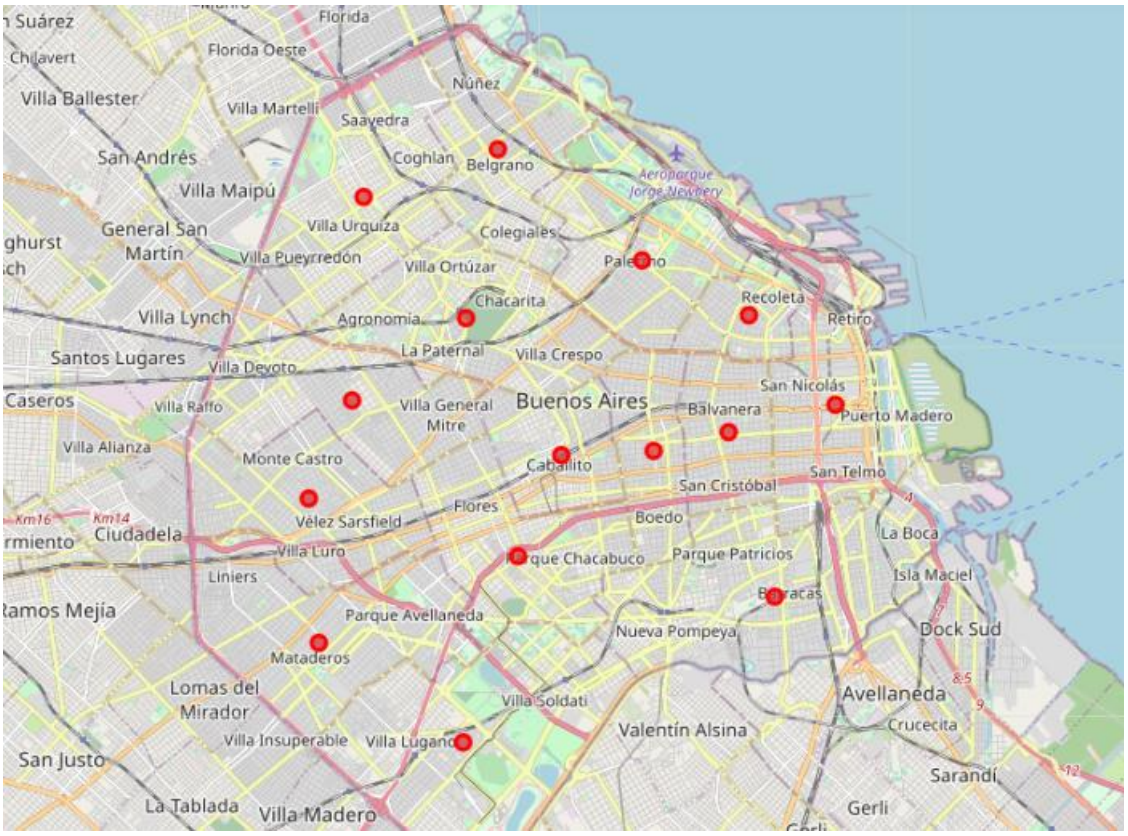

Figure 1. Map of the city of Buenos Aires and its 15 Comunas with red spots.

For comparison, I took 15 cities for Europe, according to the immigration to Argentina in the historical waves. As previously said, most of the population in Argentina descendent from Spanish or Italian. Because of that, I took the most common cities were the immigrants came from, Madrid, Barcelona, Sevilla and Valencia from Spain, and Milan, Rome, and Naples from Italy, being the last one the more common between immigrants.

Some other countries also have an important immigrational link with Argentina, so I took their capital or populated cities to consider. The list completes with Lisbon (Portugal), Paris (France), London (England), Amsterdam (Netherlands), Berlin, Munich (Germany), Warsaw (Poland), Kiev (Ukraine).

For the American continent, I took the most important cities of Latin America, including the two more populated cities in Argentina (after Buenos Aires), being Rosario and Córdoba. The list is Santiago de Chile (Chile), Montevideo (Uruguay), Sao Paulo, Rio de Janeiro, Brasilia (Brazil), Asuncion (Paraguay), Lima (Peru), Quito (Ecuador), Bogota, Medellin (Colombia), City of Mexico, Guadalajara and Monterrey (Mexico).

Figure 2. Map of the cities in Europe and Latin America. In red is Buenos Aires and in blue the rest of them.

3.2 Descriptive analysis

The Foursquare API gives us 100 venues of each city and borough that we input. For these, I have taken the average venue type that appears.

In the following tables, we can see the most common type for each of the Buenos Aires boroughs and for the cities:

Borough / City	Category	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Comuna 1	Buenos Aires borough	Coffee Shop	Café	Theater	Hotel	Hostel
Comuna 2	Buenos Aires borough	Hotel	Coffee Shop	Ice Cream Shop	Italian Restaurant	Café
Comuna 3	Buenos Aires borough	Café	Japanese Restaurar	Bakery	Ice Cream Shop	Spanish Restaurant
Comuna 4	Buenos Aires borough	Japanese Restaurar	Pizza Place	Italian Restaurant	BBQ Joint	Restaurant
Comuna 5	Buenos Aires borough	Ice Cream Shop	Pizza Place	Bakery	Café	Burger Joint
Comuna 6	Buenos Aires borough	Café	Ice Cream Shop	Bakery	Coffee Shop	Pizza Place
Comuna 7	Buenos Aires borough	Ice Cream Shop	Café	Pizza Place	Coffee Shop	Pharmacy
Comuna 8	Buenos Aires borough	Pizza Place	Fast Food Restaurar	Soccer Stadium	Shopping Mall	Soccer Field
Comuna 9	Buenos Aires borough	Pizza Place	Café	BBQ Joint	Dessert Shop	Ice Cream Shop
Comuna 10	Buenos Aires borough	Pizza Place	Café	Deli / Bodega	Plaza	Ice Cream Shop
Comuna 11	Buenos Aires borough	Café	Ice Cream Shop	Pharmacy	Burger Joint	Restaurant
Comuna 12	Buenos Aires borough	Ice Cream Shop	Pizza Place	BBQ Joint	Coffee Shop	Bakery
Comuna 13	Buenos Aires borough	Pizza Place	Café	Coffee Shop	Deli / Bodega	Ice Cream Shop
Comuna 14	Buenos Aires borough	Hotel	Coffee Shop	Bakery	Italian Restaurant	Gym
Comuna 15	Buenos Aires borough	Pizza Place	Coffee Shop	Café	Bakery	Restaurant

Table 1. Most common venue type for each of the Buenos Aires borough.

For the boroughs of Buenos Aires, we can see that most of them have venues like Pizza Place, Coffee Shop, Café (which are very similar o Coffee Shops), Ice Cream Shop and Italian Restaurant as the most popular.

The similarity of them, is that they all are big part of the Italian culture. So, this let us expect that some boroughs may be like the cities in Italy.

Figure 3. Frequencies of the most common venues in the Comunas.

Also, we can see that Bakery is other venue with high frequency in the Comunas, but this could be more associated with French culture, as Argentinian bakery tends to copy French and Nordic bakery.

For the European cities, the Foursquare API returns the venues that seems to be more tourism-related that we can see for the Buenos Aires boroughs.

Borough / City	Category	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Amsterdam	Europe city	Hotel	Restaurant	Coffee Shop	Cocktail Bar	Bar
Barcelona	Europe city	Hotel	Coffee Shop	Tapas Restaurant	Spanish Restaurant	Plaza
Berlin	Europe city	Coffee Shop	Ice Cream Shop	Bookstore	Hotel	Indie Movie Theater
Kiev	Europe city	Hotel	Coffee Shop	Cocktail Bar	Café	Caucasian Restaurant
Lisboa	Europe city	Hotel	Plaza	Scenic Lookout	Restaurant	Garden
London	Europe city	Hotel	Park	Lounge	Art Museum	Coffee Shop
Madrid	Europe city	Hotel	Restaurant	Plaza	Spanish Restaurant	Art Gallery
Milano	Europe city	Boutique	Hotel	Plaza	Ice Cream Shop	Italian Restaurant
Munich	Europe city	Café	Plaza	Hotel	Coffee Shop	German Restaurant
Napoli	Europe city	Pizza Place	Plaza	Historic Site	Italian Restaurant	Hotel
Paris	Europe city	Plaza	Hotel	Cocktail Bar	Art Museum	Bookstore
Roma	Europe city	Plaza	Historic Site	Ice Cream Shop	Monument / Landmar	Hotel
Sevilla	Europe city	Tapas Restaurant	Spanish Restaurant	Plaza	Hotel	Gastropub
Valencia	Europe city	Hotel	Ice Cream Shop	Plaza	Paella Restaurant	Italian Restaurant
Warszawa	Europe city	Café	Cocktail Bar	Park	Plaza	Coffee Shop

Table 2. Most common venue type for each of the European cities.

We can see that for most of them, the Hotel is the most common venue. In all of them, but Warsaw, the hotel venue type is present in the top 5 of most commons. For Buenos Aires borough, the hotels only appear in their top 5 for Comunas 1 and 2, and only four of them have it in their top 10.

Other common venues for European cities are their traditional Plazas and Historical Sites. Also, most of them have a high frequency in Coffee Shops, Cafés, and Art Museums.

Regarding Latin American cities, we can see that the most common venue is more variable between countries, which can be logical.

Borough / City	Category	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Asuncion	Latin American city	Pizza Place	Gym / Fitness Cente	Café	Soccer Stadium	South American Restaura
Bogota	Latin American city	Coffee Shop	Hotel	Bakery	Asian Restaurant	Italian Restaurant
Brasilia	Latin American city	Gym / Fitness Cente	Ice Cream Shop	Burger Joint	Steakhouse	Food Stand
Ciudad de Mexico	Latin American city	Ice Cream Shop	Coffee Shop	Bakery	Mexican Restaurant	Art Museum
Cordoba	Latin American city	Coffee Shop	Sandwich Place	Restaurant	Pizza Place	Hotel
Guadalajara	Latin American city	Mexican Restaurant	Taco Place	Seafood Restaurar	Coffee Shop	Restaurant
Lima	Latin American city	Peruvian Restauran	Bakery	Park	Restaurant	Athletics & Sports
Medellin	Latin American city	Café	Park	Pizza Place	Hotel	Peruvian Restaurant
Monterrey	Latin American city	Mexican Restaurant	Taco Place	Hotel	Ice Cream Shop	Restaurant
Montevideo	Latin American city	Plaza	Coffee Shop	Hotel	Park	Scenic Lookout
Quito	Latin American city	Bakery	Coffee Shop	Dessert Shop	Ice Cream Shop	Italian Restaurant
Rio de Janeiro	Latin American city	Steakhouse	Ice Cream Shop	Park	Mountain	Stadium
Rosario	Latin American city	Park	Fast Food Restaura	Pizza Place	Ice Cream Shop	Hotel
Santiago de Chile	Latin American city	Bakery	Pizza Place	Park	Sandwich Place	Coffee Shop
Sao Paulo	Latin American city	Ice Cream Shop	Pizza Place	Brazilian Restaura	Park	Bookstore

Table 3. Most common venue type for each of the Latin American cities.

Their most common venue type could be Coffee Shops, Ice Cream Shop and Parks, but half of them don't have this type in their top 5, and one third of them neither in their top 10.

As we can see, different countries could have different cultures, between Latin American, so it highly logical that the variation between these cities is higher than the Europeans ones (which has less countries involved and some more cultural like the others) and Buenos Aires boroughs.

4. Modelling

4.1 Model definition

For comparing how much the cities are alike to each other, I have used the K-means clustering algorithm, grouping them by the frequency of their venues type with the Foursquare data. With this method of vector quantization, I aim to partition the observations (15 Comunas and 30 cities) into k clusters in which each observation belongs to the cluster with the nearest mean (cluster centers or cluster centroid), serving as a prototype of the cluster.

For this instance, I've made two clustering model, with two and with three clusters. The objective is to split first the locations into "European" vs "Latin American", so each borough should belong to one of the two clusters, and it can be used to control that the result are according to the logic. The other model, as it would have three clusters, could give some other richer conclusions, in terms of considering boroughs and cities could be like between them in a less extreme way.

The method used is the algorithm of the Python package Sklearn.cluster from KMeans.

4.2 Clustering into two groups

As previously mentioned, I first model the K-mean method with two clusters. The ideal scenario would be that every city would be split according to their respective continent, only considering the frequency of the venue type that they have.

In this first model, the results were according as the expected. With an emphasized difference between the cities of the two continents

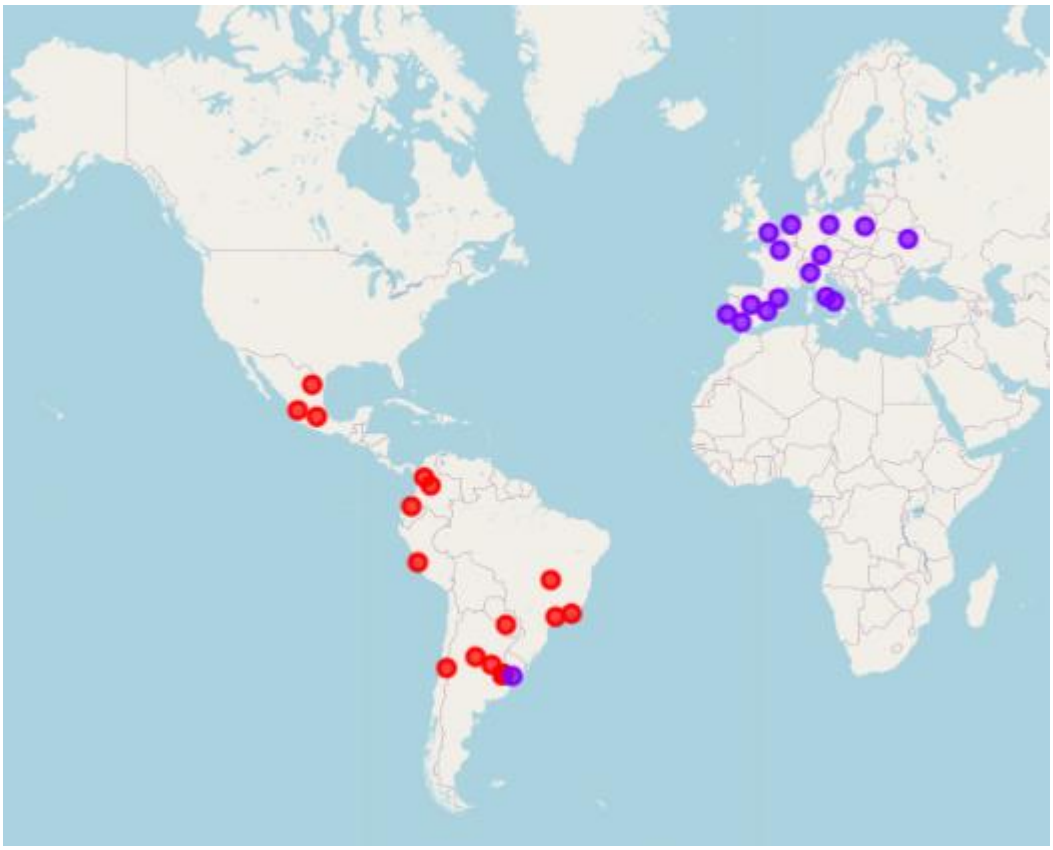

Figure 4. Map of the cities split into two clusters.

We can see the red cluster grouping all the Latin American cities, except for Montevideo, Uruguay, which has been clustered with the Europeans in the purple cluster.

For the Buenos Aires borough, the results are also logical, as most of them tend to be more like the Latin American cities, but for one.

Figure 5. Map of Buenos Aires' Comunas split into two clusters.

We can see that only the Comuna 2, which only groups Recoleta's neighborhood, is more alike to the European cities. This can make sense, as Recoleta is one of the wealthiest neighborhoods in the city of Buenos Aires and it is also the most touristic location.

As previously seen, the European cities tend to have a higher frequency of hotels mainly, and historical places and plazas, with a decent addition of Coffee Stores and Cafés. So, looking of Comuna 2's most common venue type, it can be logical that it has been grouped under a European city. The borough has 11% of hotels as their venue type, with other 8% of Coffee Shops, 4% of Cafés and 4% of Plazas.

Also, for the city of Montevideo, we can see that these three venues type also have the highest frequency of its venues.

4.3 Clustering into three groups

Noticing that the first model gave us some logical results, the next step was model with three cluster. The ideal scenario in this case, would be most Europeans and Latin American cities being grouped into two separated groups, with a third one containing the most Latin-American like cities of Europe and the most European like cities of Latin America.

In this second model, the results were interesting to see as they can give us some point to think about.

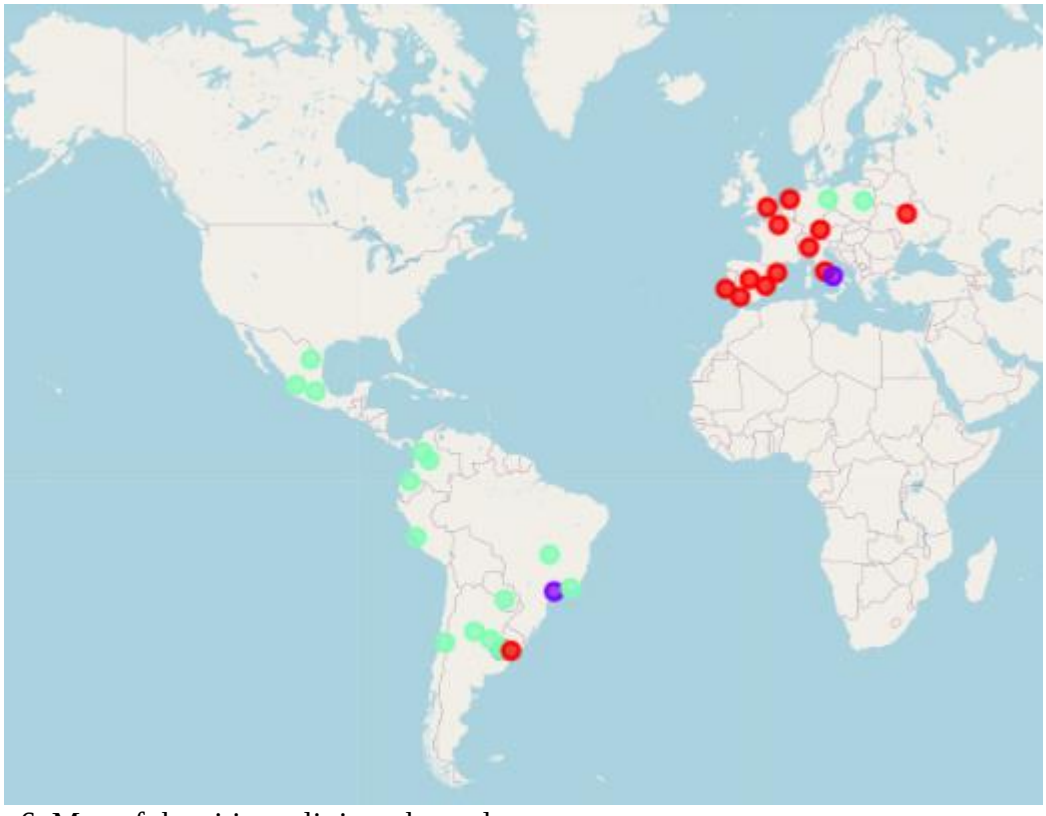

Figure 6. Map of the cities split into three clusters.

This time, we can see a red cluster, grouping most of the European cities, except for Berlin, Warsaw, and Naples, with the addition of Montevideo again. This can be explained as those three cities, along with Rome, where the ones where the tourism may not be as high as the other. Those four cities were the only ones that the hotel type venue wasn't in their top 3 of venues.

But, regarding to Rome, the difference between that city and the other three, is that probably its low frequency of hotels was overcome with its high frequency of plazas, historical places, and art museums, which made it more alike the other Europeans cities.

In cyan, there is the second cluster, of more Latin-American cities, which excludes Sao Paulo, Brazil, and with the addition of Berlin and Warsaw. Again, the addition of these two clusters can be explained in their low frequency of tourism venues and high appearance of Coffee Shops and Cafés.

In the third cluster, in purple, there are the cities that can be more in the middle, including Naples and Sao Paulo. These two cities have high frequency of pizza places, art museums and Italian food.

For the Buenos Aires' boroughs, the results more variable, as most of them are split into the second and third group.

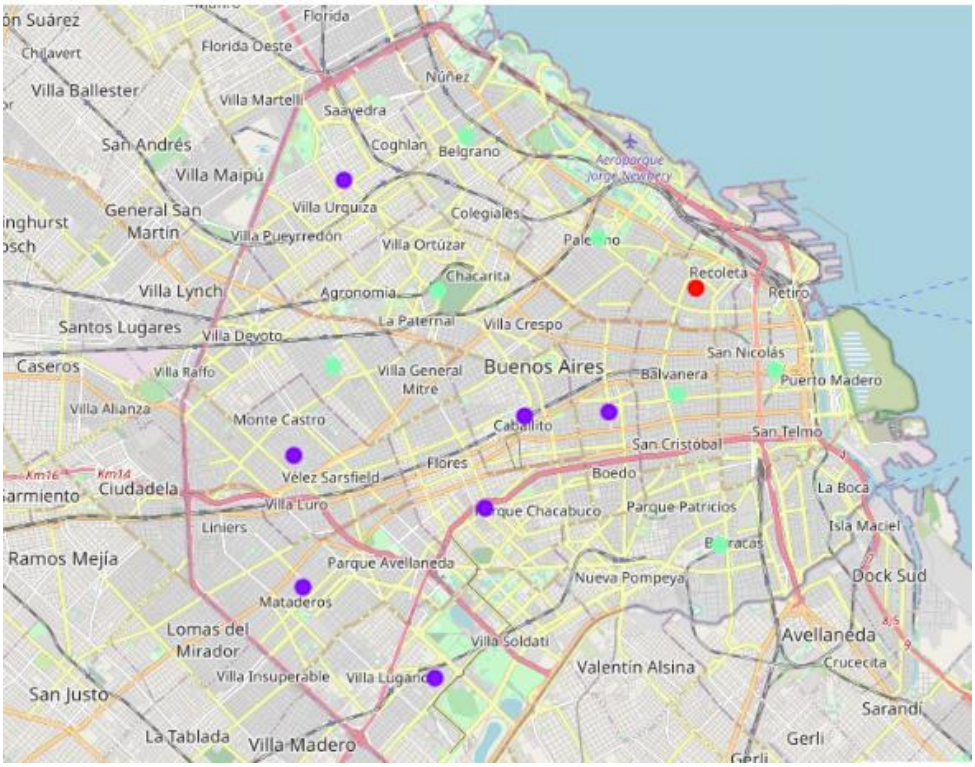

Figure 7. Map of Buenos Aires' Comunas split into three clusters.

We can see that again the Comuna 2, of Recoleta's neighborhood, is more alike to most of western European cities (the red cluster). This also make sense as their venue types are more similar as in the previous model.

The difference this time is that the remaining 14 Comunas were split in halves for the two other clusters. The Comunas 1, 3, 4, 11, 13, 14 and 15 were included in the cyan cluster (with most of Latin American cities) and the Comunas 5, 6, 7, 8, 9, 10 and 12 were included in the purple cluster with Naples and Sao Paulo.

It is also remarkable of how the clustering split them by geography, without considering its actual location. The north-easter boroughs seems to be more related to the cyan cluster whether the south-wester borough tend to be more like the purple one.

It was expectable that in way cities like Naples could be more like some boroughs of Buenos Aires, as its culture is some ways different from the ones in the north of Europe, and southern boroughs of Buenos Aires have been highly influenced by southern Italy immigrants. In fact, we can see that these boroughs have a high frequency of pizza places, ice cream shops and cafés. Despite, those venues were frequent in most of the Comunas, those one has the highest percentages of these venues that have Italian origins.

The purple cluster includes the Comunas with the lowest incomes, so they can be partially influenced for this economic impact, despite that this variable was not included directly, but can impact of the venue types that a location has.

Previously of the results, we could be expecting that the Comuna 4, that includes Barracas and La Boca, were more alike Naples and the purple cluster. Despite that, the Comuna 4 has a big influence of different types of immigrants, which made its venues type more variable than the previous group.

In the other hand, we could expect that Comuna 12, including Villa Urquiza neighborhood, would be included in the mid-income group, with the cyan cluster. This borough it is in practice, like some of the neighborhoods included in the Comunas 11 and 15, which belongs to the cyan group.

For the cyan cluster, we can a see that the boroughs have more variable types of venues, as the different cities in Latin America have. This can be explained as the impact of different immigrants from Spain, France, Asia and not as big from the Italian community, despite also the ice cream shops and coffee stores being popular.

5. Conclusion

In this study, I analyzed the similarities of the fifteen Comunas of Buenos Aires between some cities of Europe and Latin America, according to the types of venues that they have. Under the hypothesis of that the frequency of some venues depend much on the culture of a country or area, I could explain if a certain borough has a similar culture and lifestyle of a European city. I have work with a K-mean clustering algorithm to find that similarities.

I identified that the Comuna 2, which only includes the neighborhood of Recoleta, is the only one that can be compared to a European city in terms of lifestyle, without considering any variable of economic impact.

With the model of three cluster, the idea was to identify if there was any other borough that could have some similarity to other city, but not as high as for Recoleta. In this case, the model split Buenos Aires is almost perfect two parts. With the north-east being more like the other Latin American cities (and maybe the Easter Europeans, despite Kiev was not included), and with the south-west being like the cities like Naples or Sao Paulo, with high impact of South Italian culture.

As for the initial question. No, Buenos Aires does not seem to be any similar to a European city. As much, it can have some parallelism with some Eastern-European cities.