

The Battle of Neighborhood

SANTIAGO AND BUENOS AIRES



Introduction

- Santiago and Buenos Aires are two cities in South America, located in opposite areas of the region, the first in the western part and the second in the eastern part of the continent.
- Both cities are very diverse and are the capitals of their respective countries (Chile and Argentina), they have some things in common, multiple panoramas, versatile and cosmopolitan, full of contrasts, combining the traditional history and modernity, and important tourist destination in Latin America.
- Given all of the above, interest arises in discovering how similar they are.

Introduction: Business Problem

- This Project seeks to segment each city and to explore what services are available in each of these divisions, to know which are the most common category in each of them, using the Foursquare API, and to then classify them and perform a clustering analysis to determine how similar or different they are. It will use the k-means clustering algorithm to complete this task.
- This can be useful for stakeholders in moving from one city to another, and who want to maintain a similar lifestyle when changing places of residence.

Some questions to answer are:

- In each city, the amount of venues available and categories are similar?
- What are the 10 most common categories in each city and in each division?
- What characteristics does each cluster have?
- Are there neighborhoods with similar characteristics between both cities?

Data Acquisition

- **Administrative division of each city and its geographical coordinates.** For Santiago de Chile, the data was extracted from: https://es.wikipedia.org/wiki/Anexo:Comunas_de_Chile. While the information from Buenos Aires was extracted from:

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

- **The most common venues in each area.** These will be obtained using the Forsquare API

- **Value of the square meter of the housing.** These were obtained from:

<https://www.latercera.com/pulso/noticia/precio-del-m2-departamentos-supera-los-2-millones-siete-comunas-santiago/654495/>, for Santiago and from: <https://blog.properati.com.ar/reporte-del-mercado-inmobiliario-properati-oikos-bs-as-mayo-2019/>, for Buenos Aires.

- **Files json with the divisions of each city.** For Santiago file was downloaded from:

<https://github.com/jlhonora/geo> and for Buenos Aires from:

<https://data.buenosaires.gob.ar/dataset/barrios/archivo/1c3d185b-fdc9-474b-b41b-9bd960a3806e>.

Data Preparation

-The first thing is to get the data, scrapping every website, already mentioned and transform the data into a pandas data framework.

-Data wrangling is carried out, to obtain clean data frames and only with the necessary information.

	Communes	Province	Superficie(km2)	Población2017	Densidad(hab./km2)	IDH 2005	latitud dms	latitute	longitude	UF/m2	USD/m2
0	Santiago	Santiago	23.2	404495	17435.1	0.807	-33°26'14"	-33.437222	-70.657222	64.4	2261.728
1	Cerrillos	Santiago	21	80832	3849.1	0.743	-33°30'0"	-33.500000	-70.716667	39.7	1394.264
2	Cerro Navia	Santiago	11	132622	12056.5	0.683	-33°25'19.2"	-33.422000	-70.735000	NaN	NaN
3	Conchalí	Santiago	10.7	126955	11865.0	0.707	-33°22'48"	-33.380000	-70.675000	51.5	1808.680
4	El Bosque	Santiago	14.2	162505	11444.0	0.711	-33°34'1.2"	-33.567000	-70.675000	NaN	NaN
5	Estación Central	Santiago	15	147041	9802.7	0.735	-33°27'32.4"	-33.459000	-70.699000	52.8	1854.336
6	Huechuraba	Santiago	44.8	98671	2202.5	0.737	-33°22'4.8"	-33.368000	-70.634000	54.6	1917.552
7	Independencia	Santiago	7	100281	14325.9	0.709	-33°24'46.8"	-33.413000	-70.666000	57.9	2033.448
8	La Cisterna	Santiago	10	90119	9011.9	0.775	-33°31'44.4"	-33.529522	-70.664253	49.8	1748.976
9	La Florida	Santiago	70.2	366916	5226.7	0.804	-33°31'30"	-33.533300	-70.583300	52.9	1857.848
10	La Granja	Santiago	10	116571	11657.1	0.689	-33°31'60"	-33.533333	-70.625000	36.7	1288.904
11	La Pintana	Santiago	30.6	177335	5795.3	0.679	-33°34'58.8"	-33.591389	-70.636423	NaN	NaN

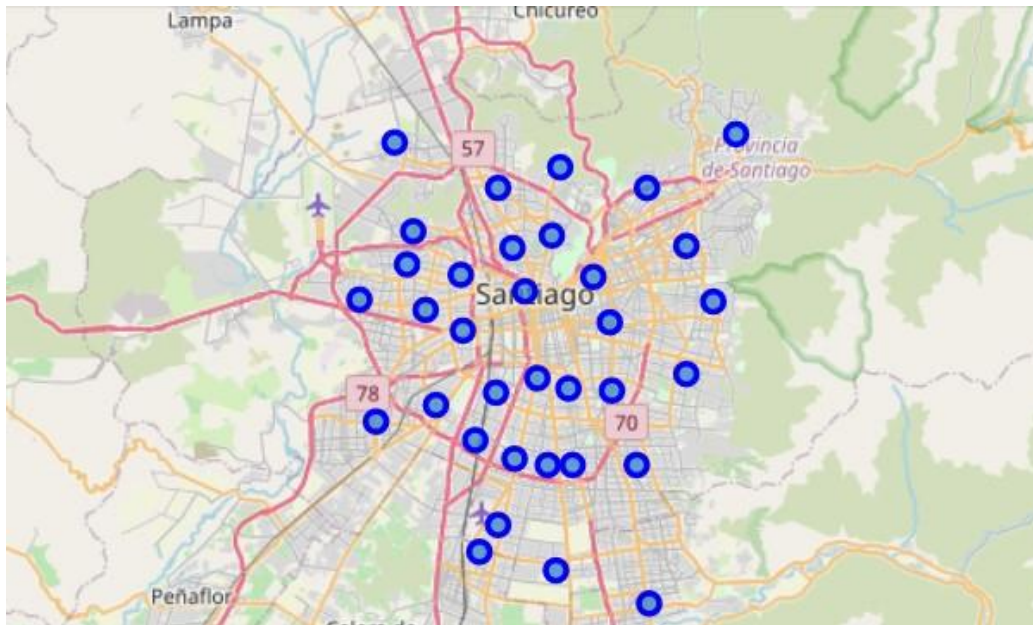
stgo data frame

	Nombre del barrio	Superficie[1]	Habitantes (año 2007)	Densidad de habitantes (año 2007)	Comuna de pertenencia de la Ciudad de Buenos Aires	latitud	longitud
0	Agronomía	2,1 km ²	34.580	8.645	Comuna 15	-34.591516	-58.485385
1	Almagro	4,1 km ²	139.262	33.960	Comuna 5	-34.609988	-58.422233
2	Balvanera	4,4 km ²	152.198	34.950	Comuna 3	-34.609215	-58.403140
3	Barracas	7,6 km ²	77.474	10.194	Comuna 4	-34.645285	-58.387562
4	Belgrano	6,8 km ²	138.942	20.433	Comuna 13	-34.561308	-58.456545

data_ar data frame

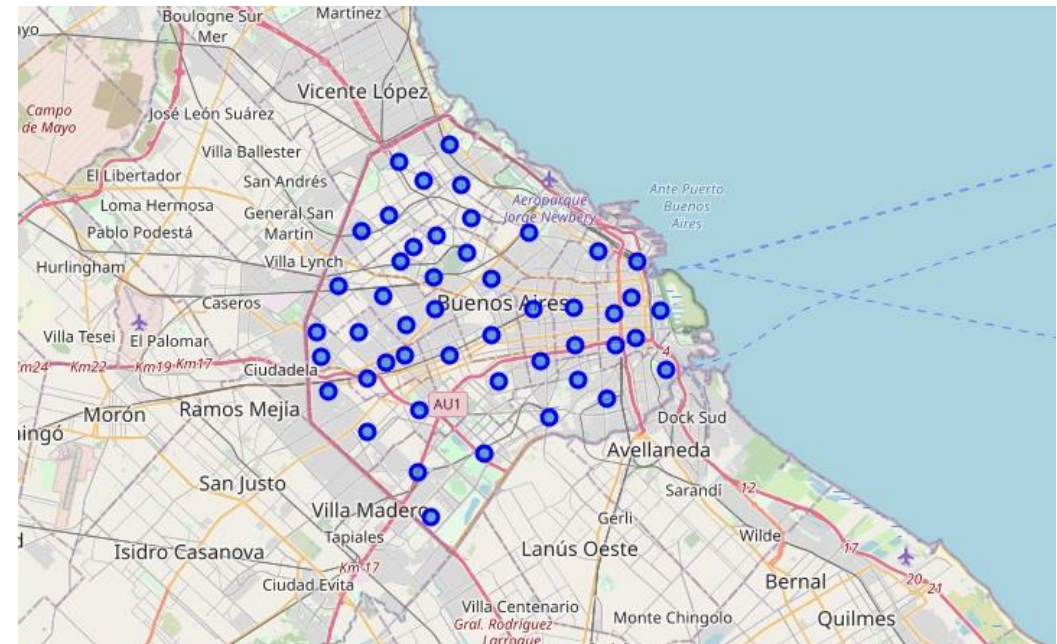
Segmenting each city

SANTIAGO DE CHILE



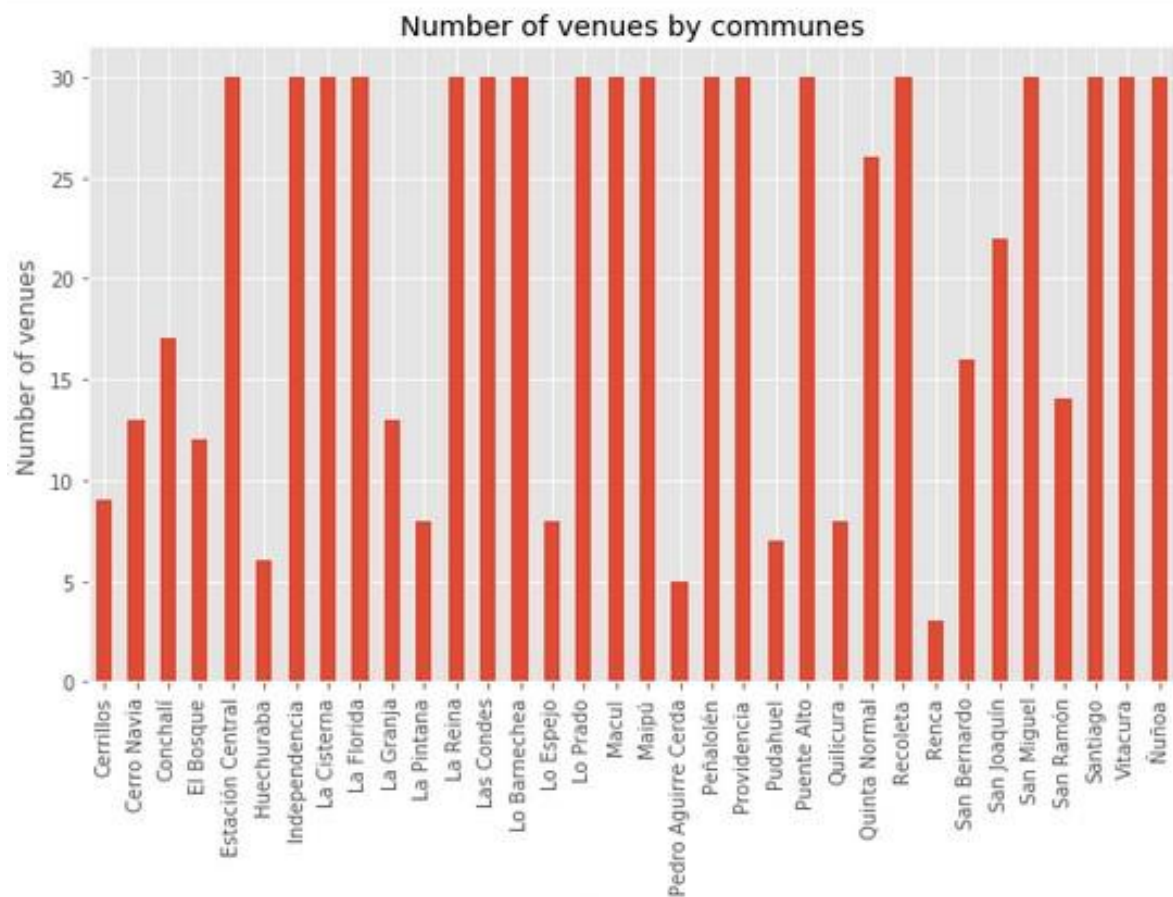
34 Neighborhoods (communes)

BUENOS AIRES, ARGENTINA



48 Neighborhoods

Exploring Every Neighborhood



SANTIAGO DE CHILE

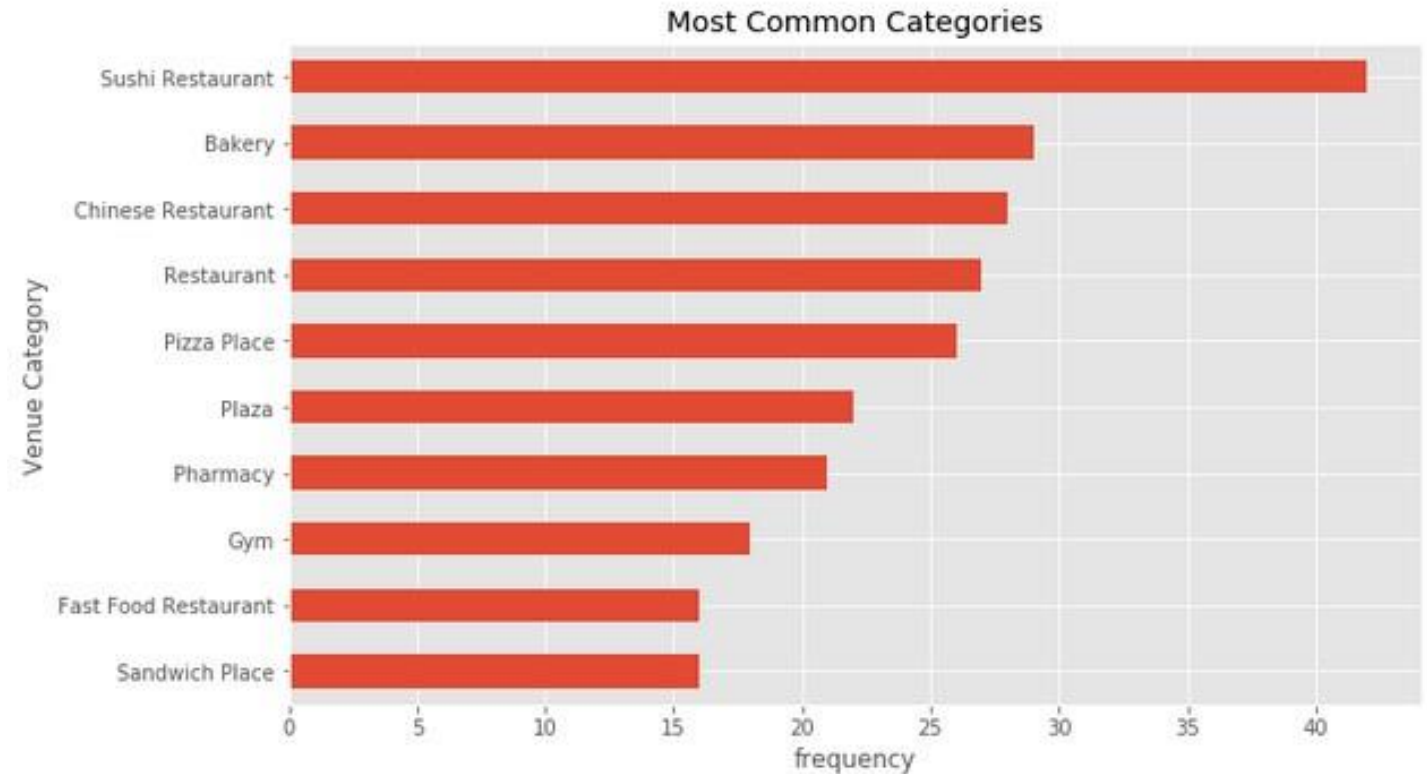
It was set to explore a limit of 30 venues per neighborhood. However, only 18 of the 34 neighborhoods reach this limit, there are even 8 neighborhoods that have less than 10 venues available in their surroundings.

Exploring Every Neighborhood

SANTIAGO DE CHILE

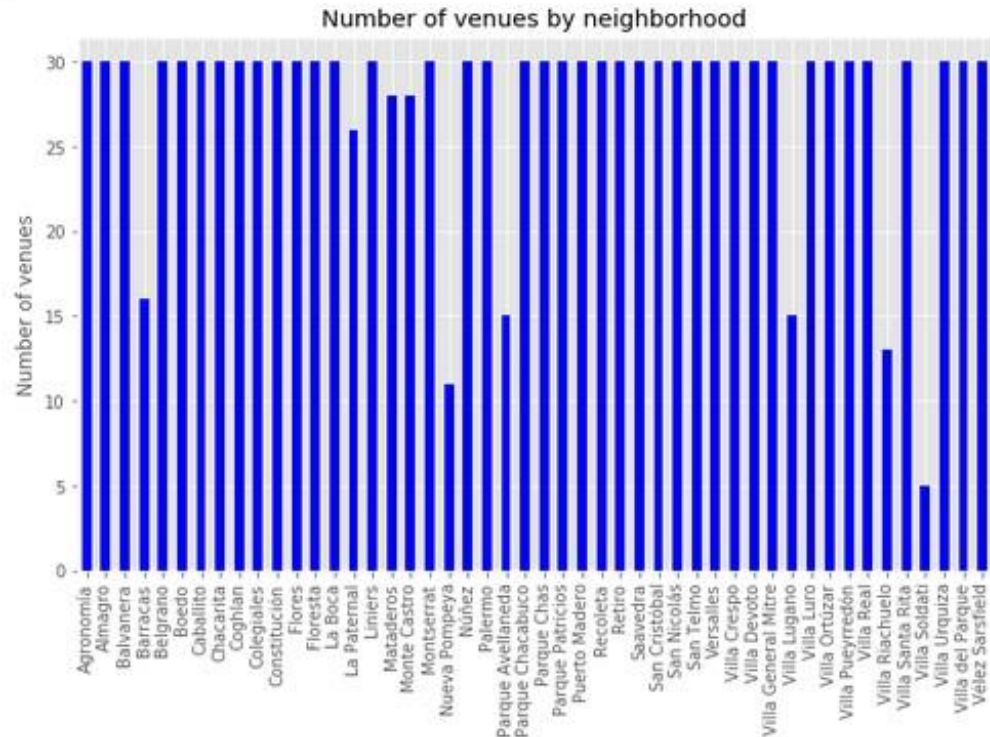
163 unique venues categories are obtained. The most common is sushi restaurant.

	Venue
Venue Category	
Sandwich Place	16
Fast Food Restaurant	16
Gym	18
Pharmacy	21
Plaza	22
Pizza Place	26
Restaurant	27
Chinese Restaurant	28
Bakery	29
Sushi Restaurant	42



Exploring Every Neighborhood

BUENOS AIRES



It was set to explore a limit of 30 venues per neighborhood. Most neighborhoods reach this limit, only 2 neighborhoods have less than 10 venues available in their surroundings.

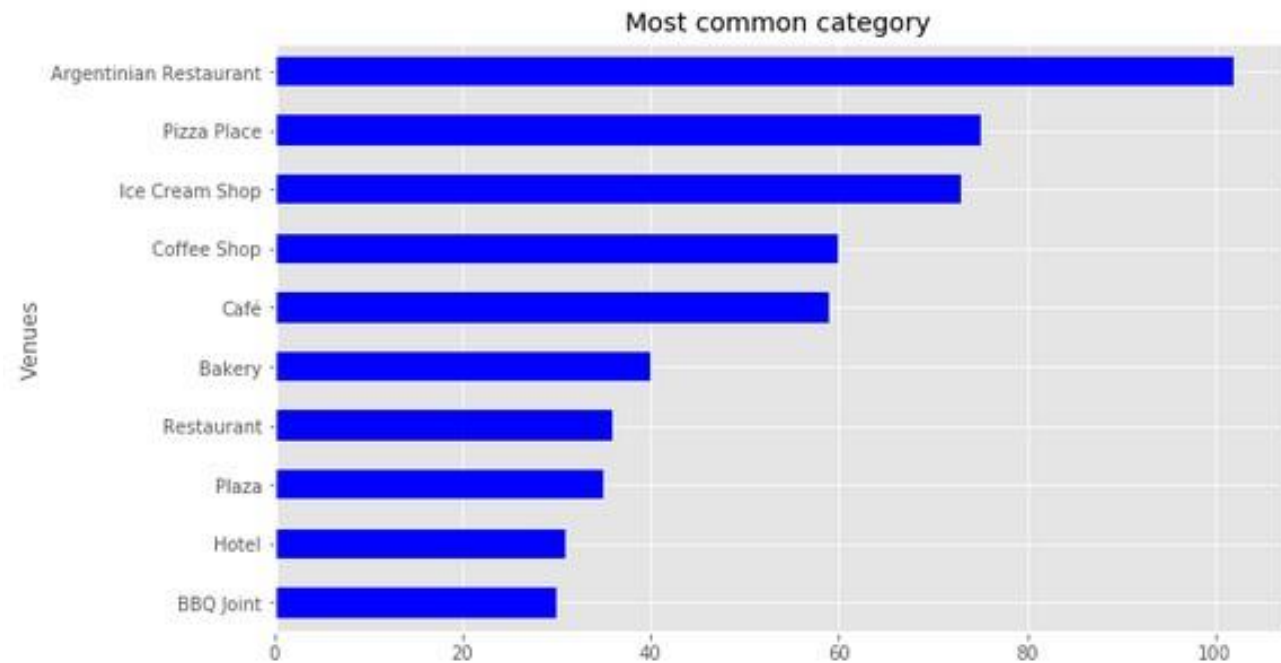
Exploring Every Neighborhood

BUENOS AIRES

In the city there are 200 unique venue categories.

The most common is argentinian restaurant.

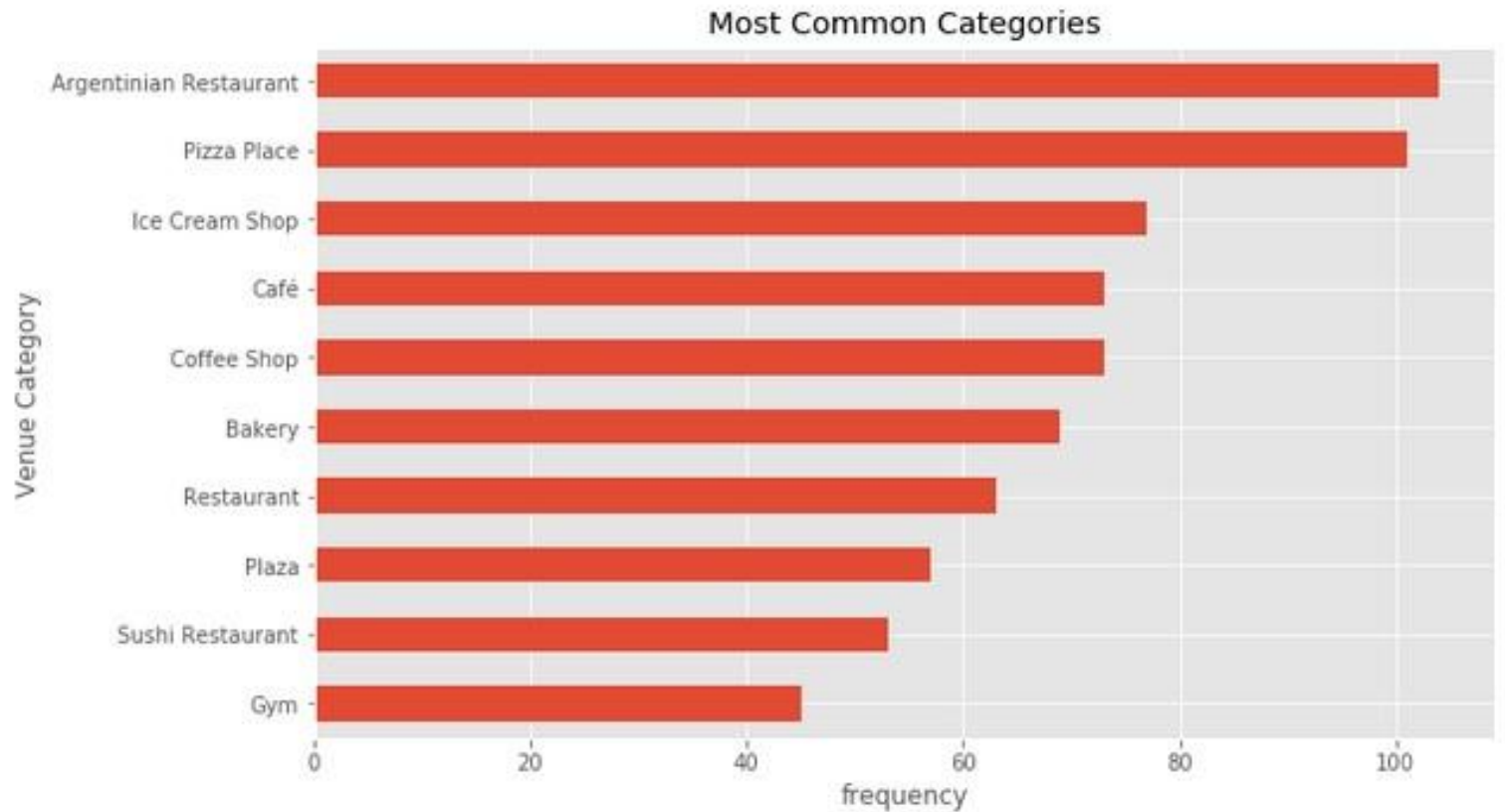
	Venue
Venue Category	
BBQ Joint	30
Hotel	31
Plaza	35
Restaurant	36
Bakery	40
Café	59
Coffee Shop	60
Ice Cream Shop	73
Pizza Place	75
Argentinian Restaurant	102



Exploring Every Neighborhood

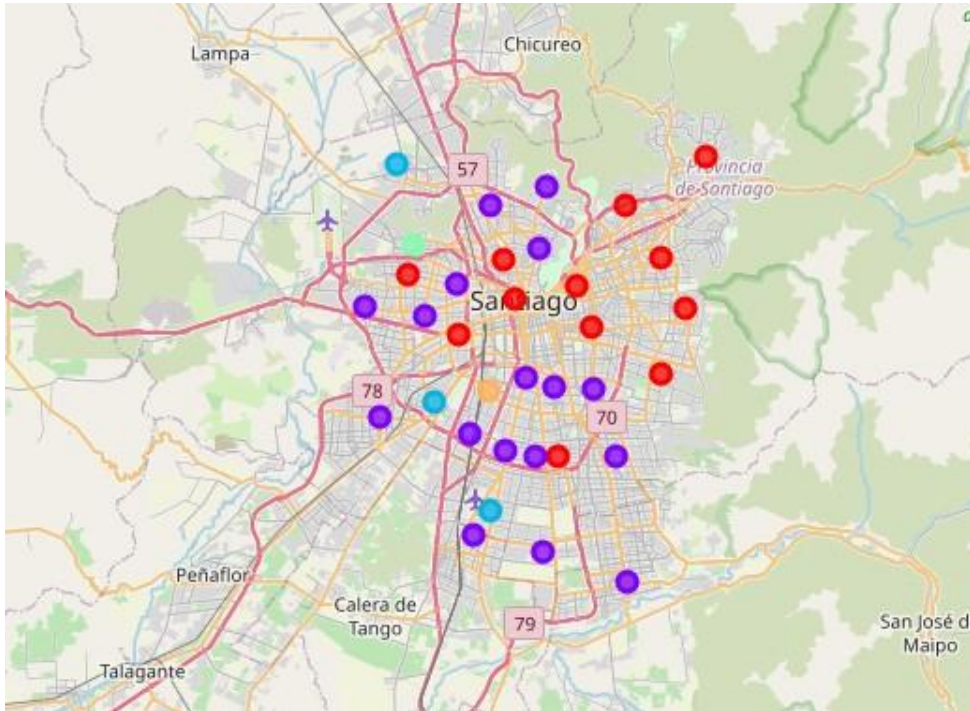
BOTH CITIES

There are 247 unique categories and the most common are argentinian restaurant and pizza places.



Clustering Neighborhoods

SANTIAGO DE CHILE

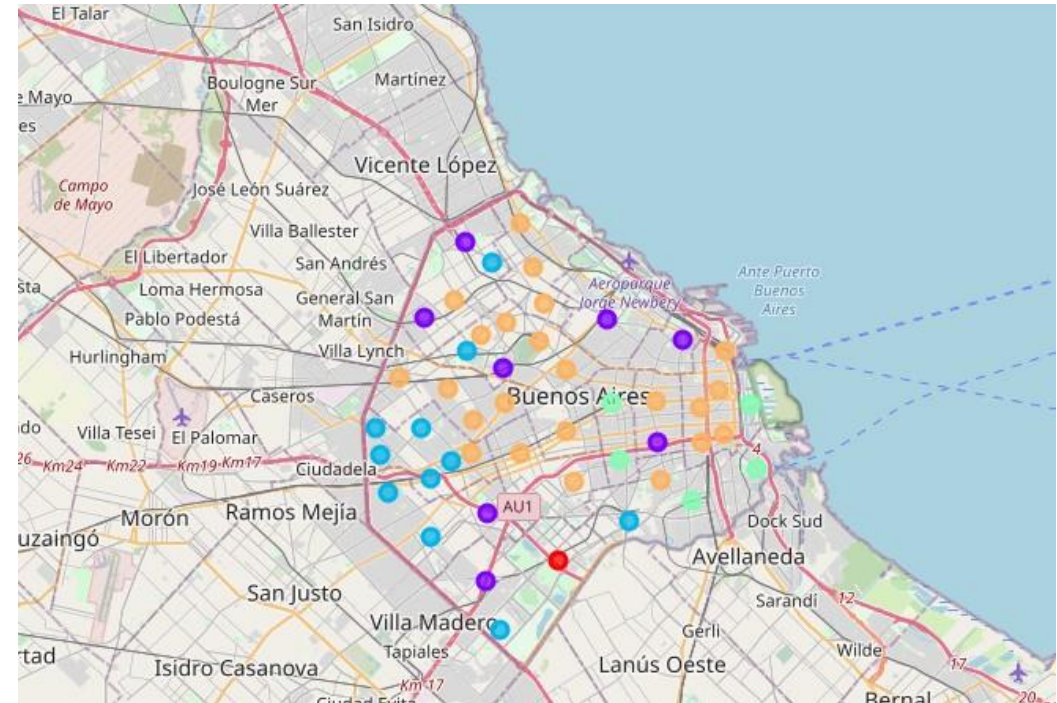


After applying the k-means clustering algorithm, with a number of clusters equal to 5, the most of the neighborhoods belong to cluster 0 (red) or cluster 1 (purple), the other clusters have 3 or fewer neighborhoods, which is why they can be defined as outliers, that is, unusual areas in the city.

Clustering Neighborhoods

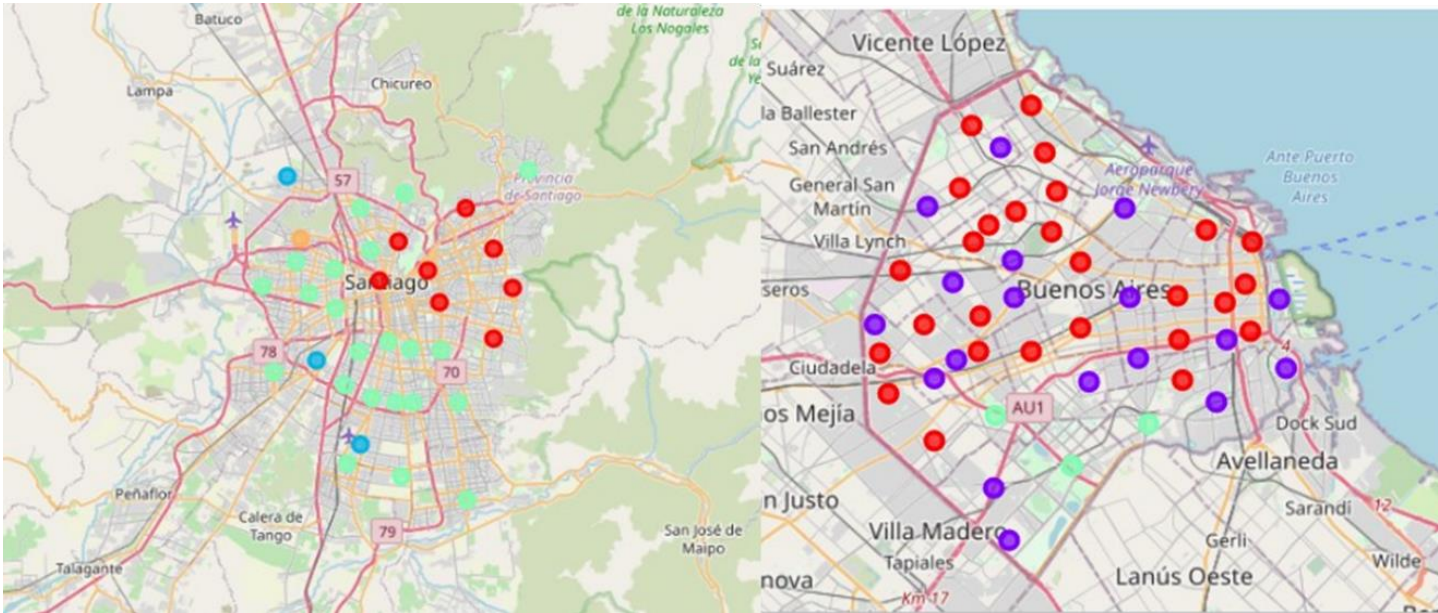
BUENOS AIRES

After applying the k-means clustering algorithm, with a number of clusters equal to 5, the most unusual neighborhood is Villa Soldati in Cluster 0 (red). Cluster 4 (orange) is the one with the highest number of neighborhoods and it is distributed throughout almost the entire city. Cluster 2 (turquesa) is more frequent in the western part of the city, while cluster 3 (green) is more frequent in the southeast and closely resembles cluster 4 and cluster 1 (purple).



Clustering Neighborhoods

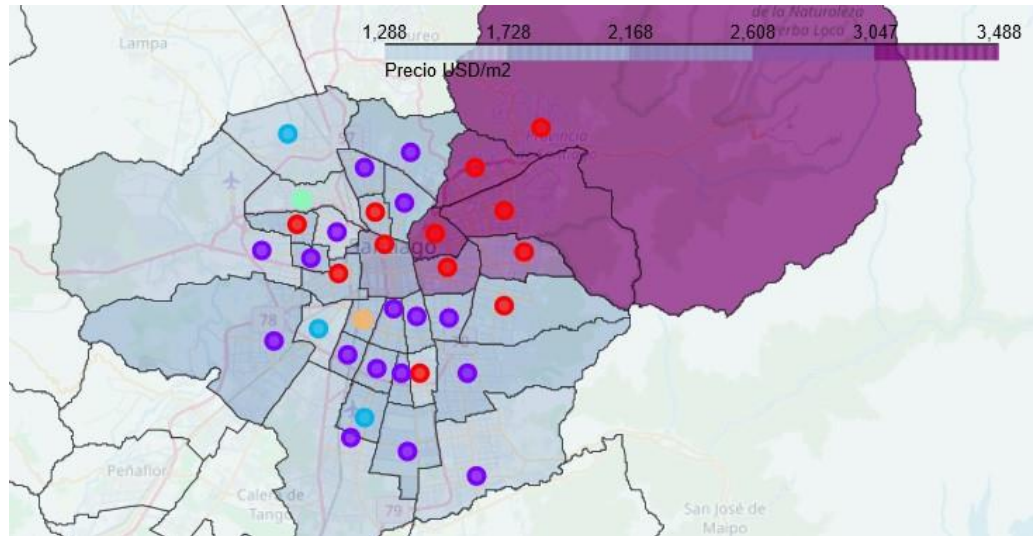
BOTH CITIES



After applying the k-means clustering algorithm, with a number of clusters equal to 5, only clusters 3 (green) and 0 (red) are present in both cities.

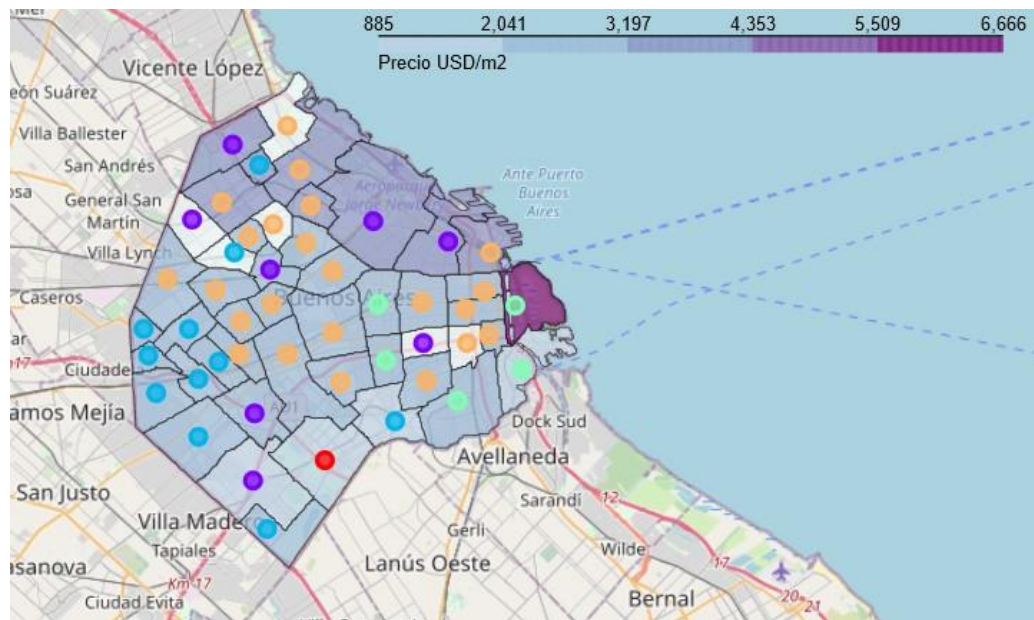
Mapping Cities

SANTIAGO



When making the choropleth map of the average house prices per square meter, the highest prices are located in the eastern sector of the city, exceeding 2100 dollars per m², in addition, these neighborhoods belong to cluster 0 (red) of the city.

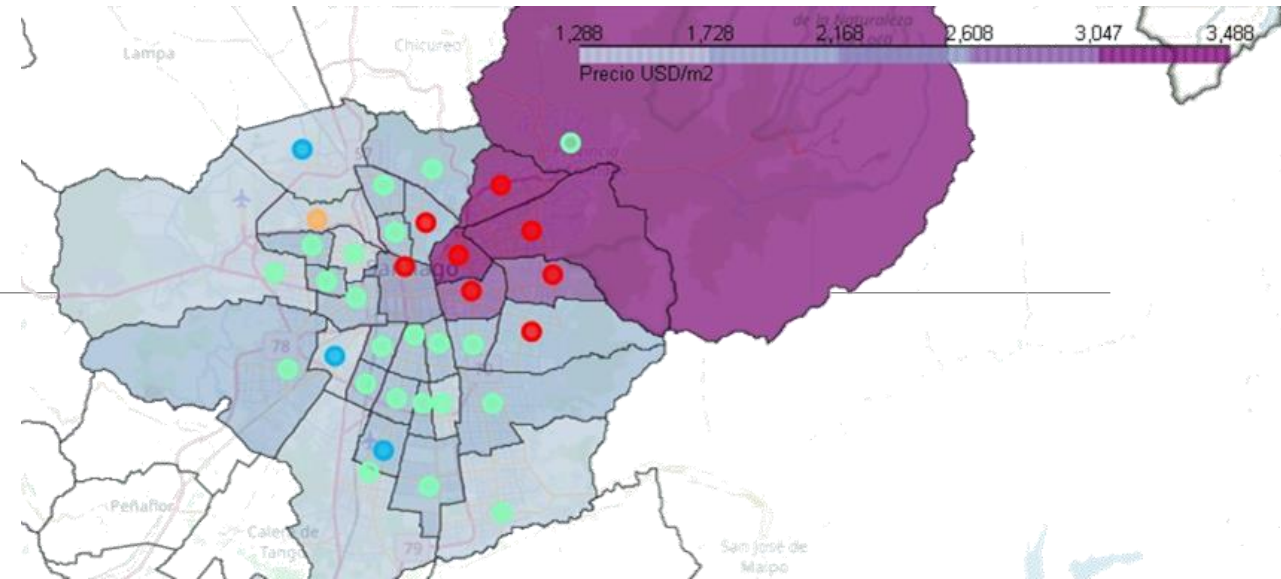
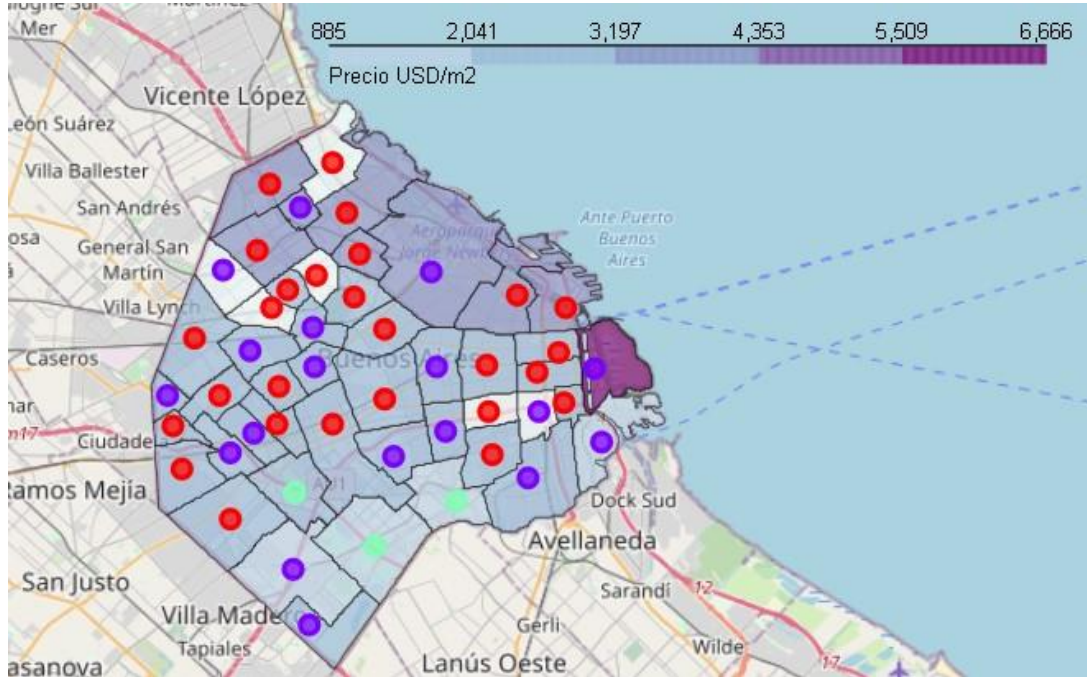
BUENOS AIRES



When making the choropleth map of the average house prices per square meter, the highest prices are located in Puerto Madero with 6666 dollars per m2, however, in general, a homogeneous price range per m2 is observed for the city, with the range between 2141 and 3197 dollars being common, whose neighborhoods belong to cluster 4 (orange) of the city in greater quantity.

Mapping Cities

BOTH CITIES



When making the choropleth map of the average house prices per square meter, the highest prices are located in Puerto Madero with 6666 dollars per m2, however, this neighborhood belongs to cluster 1(purple), only present in Buenos Aires.

For cluster 0(red), the average house price per m2 for neighborhoods in Buenos Aires ranges between 2041 and 4353 dollars, while for neighborhoods in Santiago it ranges between 1728 and 3488 dollars, both ranges are similar, for which a person at Moving from one city to the next can find a similar place to where you leave and maintain a similar lifestyle.

For Cluster 3(green), the price ranges for the neighborhoods of Buenos Aires are between 885 and 3197 dollars, while for the neighborhoods of Santiago these range between 1228 and 3488 dollars, for which it is also possible to find a home with similarly priced in a similar neighborhood.

Discussion and Conclusion

- The most common categories of venues are related to food. For Santiago it corresponds to Sushi restaurants and for Buenos Aires, Argentine restaurants. When analyzing both cities, the most common places are Argentine restaurants, pizza places, ice cream shop and coffee shop. When analyzing both cities if there are similar neighborhoods, by clustering analysis two clusters were obtained, cluster 0 with the most common venues of Restaurant, Coffee Shop (Bakery, Pizza Place, Bar) and made up of the neighborhoods of: Santiago, La Reina , Las Condes, Nuñoa, Peñalolén, Providencia, Recoleta, Vitacura, Agronomia, Balvaner, Belgrano, Caballito, Chacarita, Colegiales, Flores, Floresta, Liniers, Mataderos, Montserrat, Monte Castro, Núñez, Parque Chas, Parque Patricios, Recoleta, Retiro , Saavedra, San Cristóbal, San Nicolás, San Telmo, Versalles, Villa Crespo, Villa Devoto, Villa Ortúzar, Villa Santa Rita, Villa Urquiza and cluster 3, with the most common venues of sushi restaurant, Chinese restaurant and Fast Food Restaurant (Farmers Market), the neighborhoods that comprise it correspond to: Cerro Navia, Conchalí, Central Station, Huechuraba, Independencia, La Cisterna, Florida, La Granja, La Pintana, Lo Barnechea, Lo Espejo, Lo Prado, Macul, Maipú, Pedro Aguirre Cerda , Pudahuel, Quinta Normal, San Joaquín, San Miguel, San Ramón, Puente Alto, San Bernardo, Nueva Pompeya, Avellaneda Park, Villa Soldati. By adding the average price factor of the square meter of housing, the ranges are similar for all the neighborhoods present in each cluster.

Other interesting data to add to future analyzes would be the crime rate for each neighborhood and the distance to your workplace or from the city center.

In conclusion, it is possible to generate some useful information for people who are considering moving among Buenos Aires and Santiago or vice versa.

Thanks for Reading!
