Clustering urban land candidates for a new restaurant

City of Buenos Aires

The problem: installing a new restaurant

Owners:

- *Is there a good offer of urban land?*
- *Where do I have to install my new restaurant?*
- *Is the price fair?*
- *Are there structural problems, like waterlogging, which can compromise my business?*
- *Will I have enough customers to make my business profitable?*

City Administrators:

- *How can I offer better conditions for investments in the gastromic fields?*

Colecting the data: sources

From the government of the city of Buenos Aires (https://data.buenosaires.gob.ar/):

- List of neighborhoods of Buenos Aires (GEOJSON)
- List of urban land offered in the city (CSV)
- List of waterlogging zones in the city of Buenos Aires (CSV)

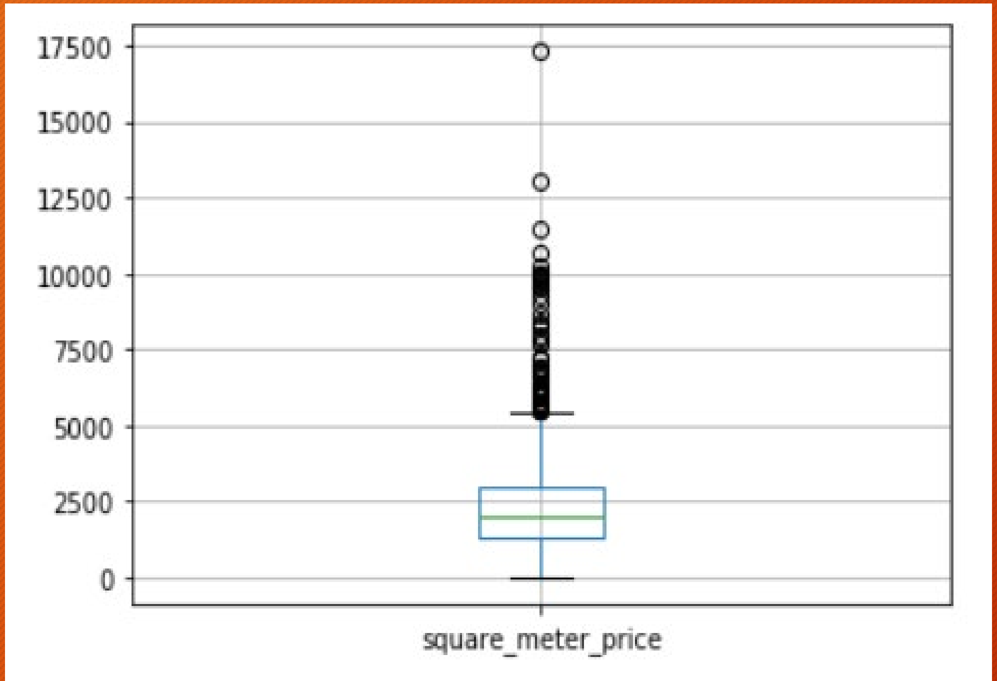
From Foursquare for developers (https://developer.foursquare.com/):

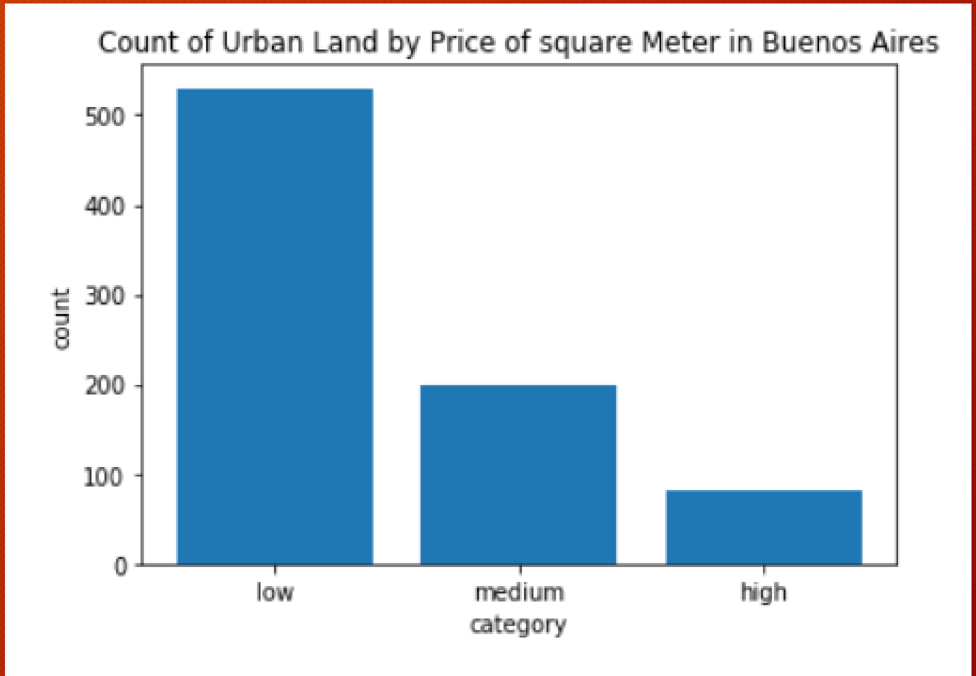
- API to get information about venues in the city

Exploring the data: urban land offer

To obtain insights into the value of urban land in the city and pricing distribution:

- List of urban land offered in the city (CSV)

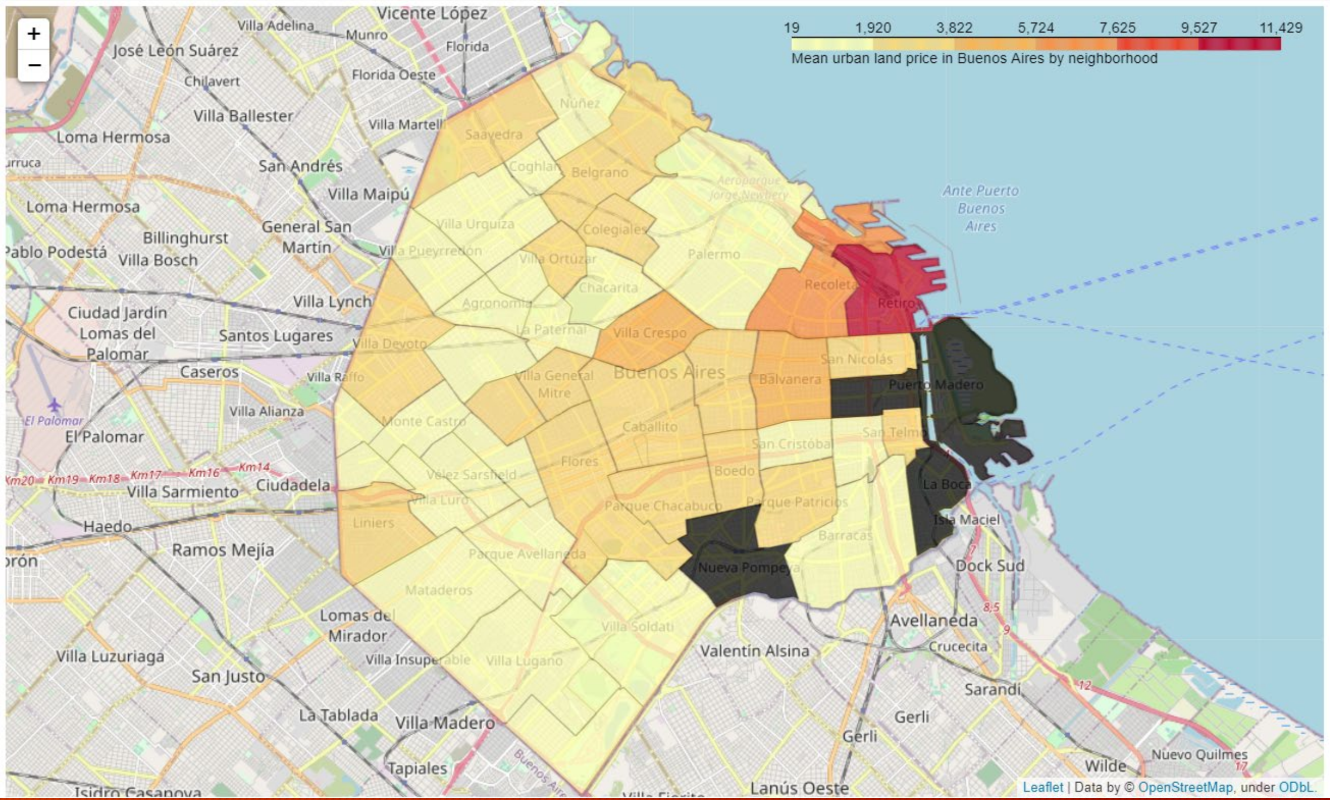




Preparing the data: urban land offer

To identify the offer of urban land in the city, location, prices, and visualization:

- Neighborhoods of Buenos Aires (GEOJSON)
- List of urban land offered in the city (CSV)

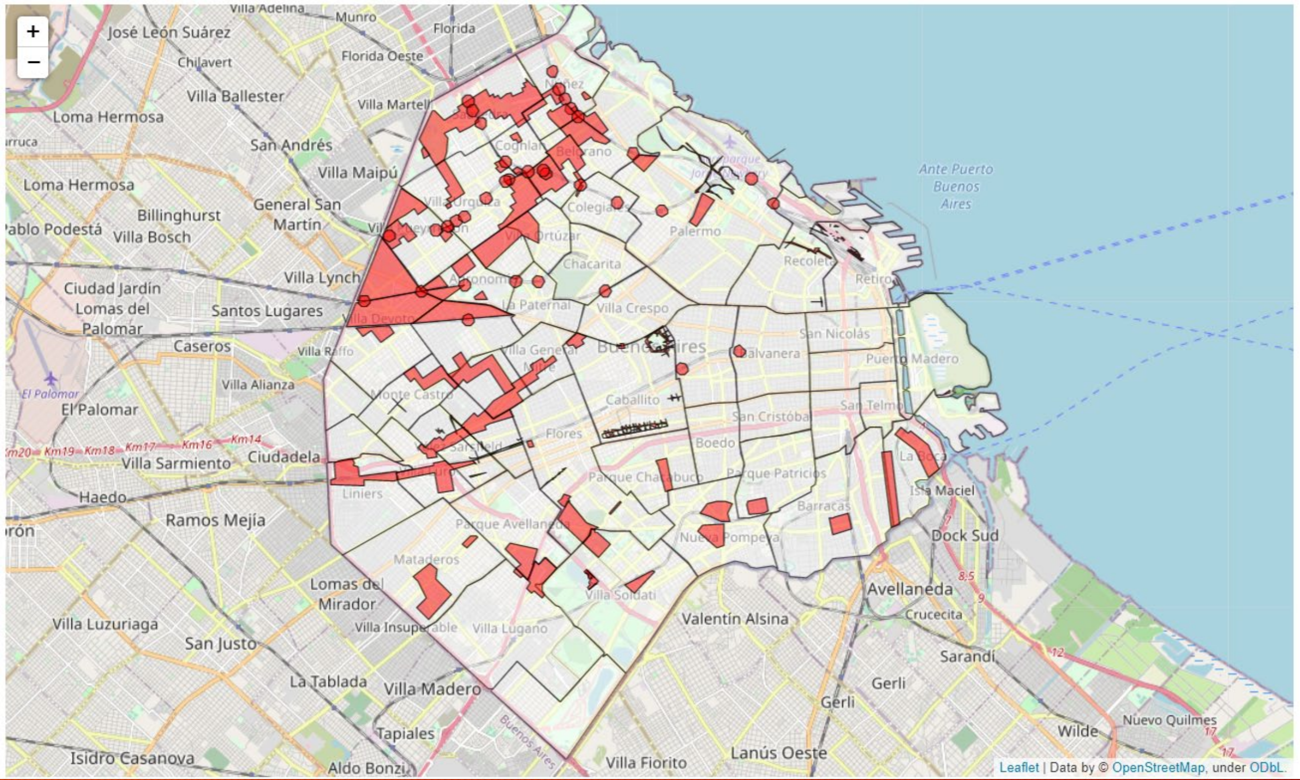


Neighborhoods of Buenos Aires by square meter price

Preparing the data: waterlogging zones

To identify waterlogging zones in the city, location, and criticality:

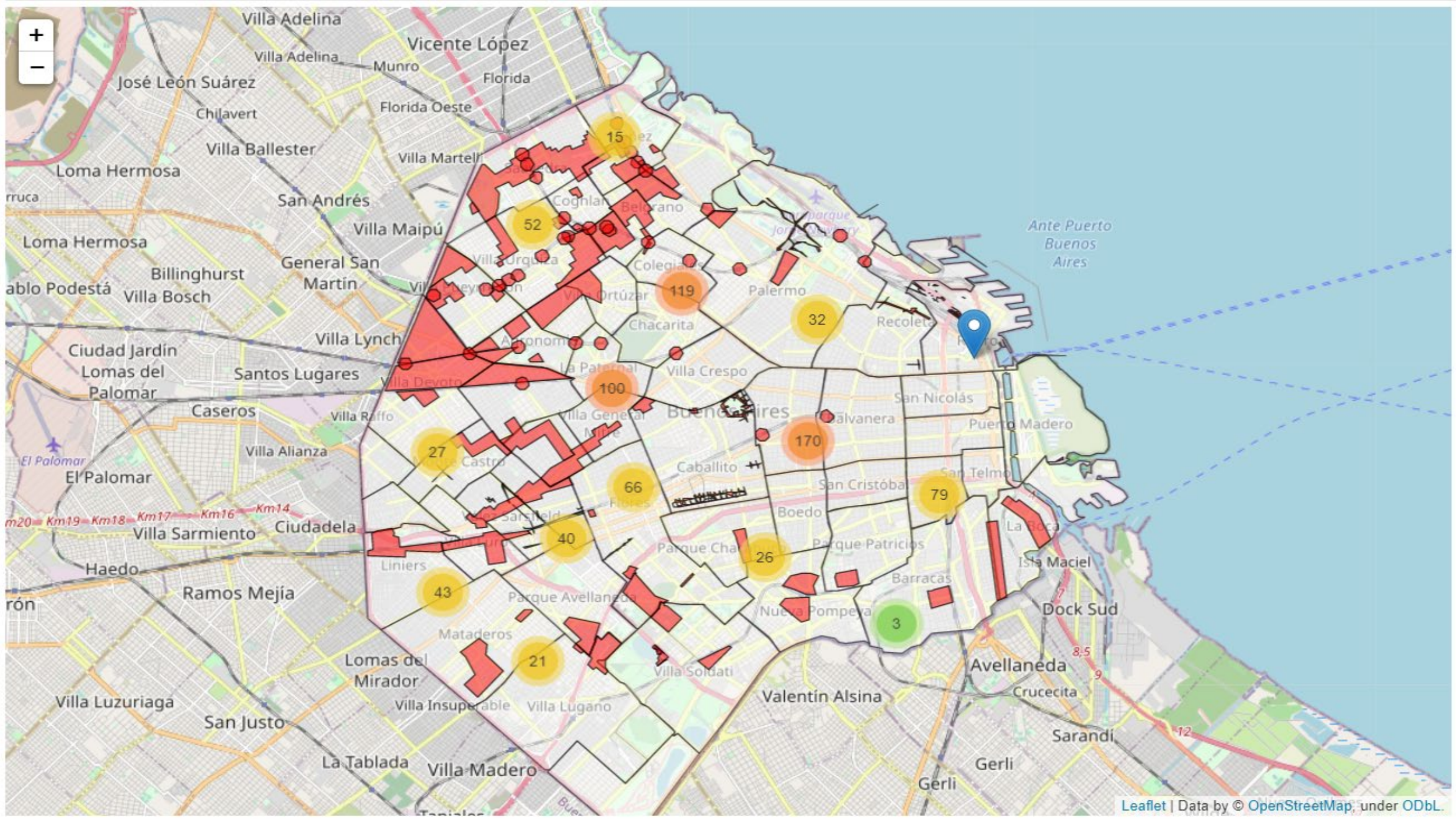
- Neighborhoods of Buenos Aires (GEOJSON)
- List of waterlogging zones in Buenos Aires (CSV)



Waterlogging zones in Buenos Aires

Working with the data: Lands in Waterlogging zones

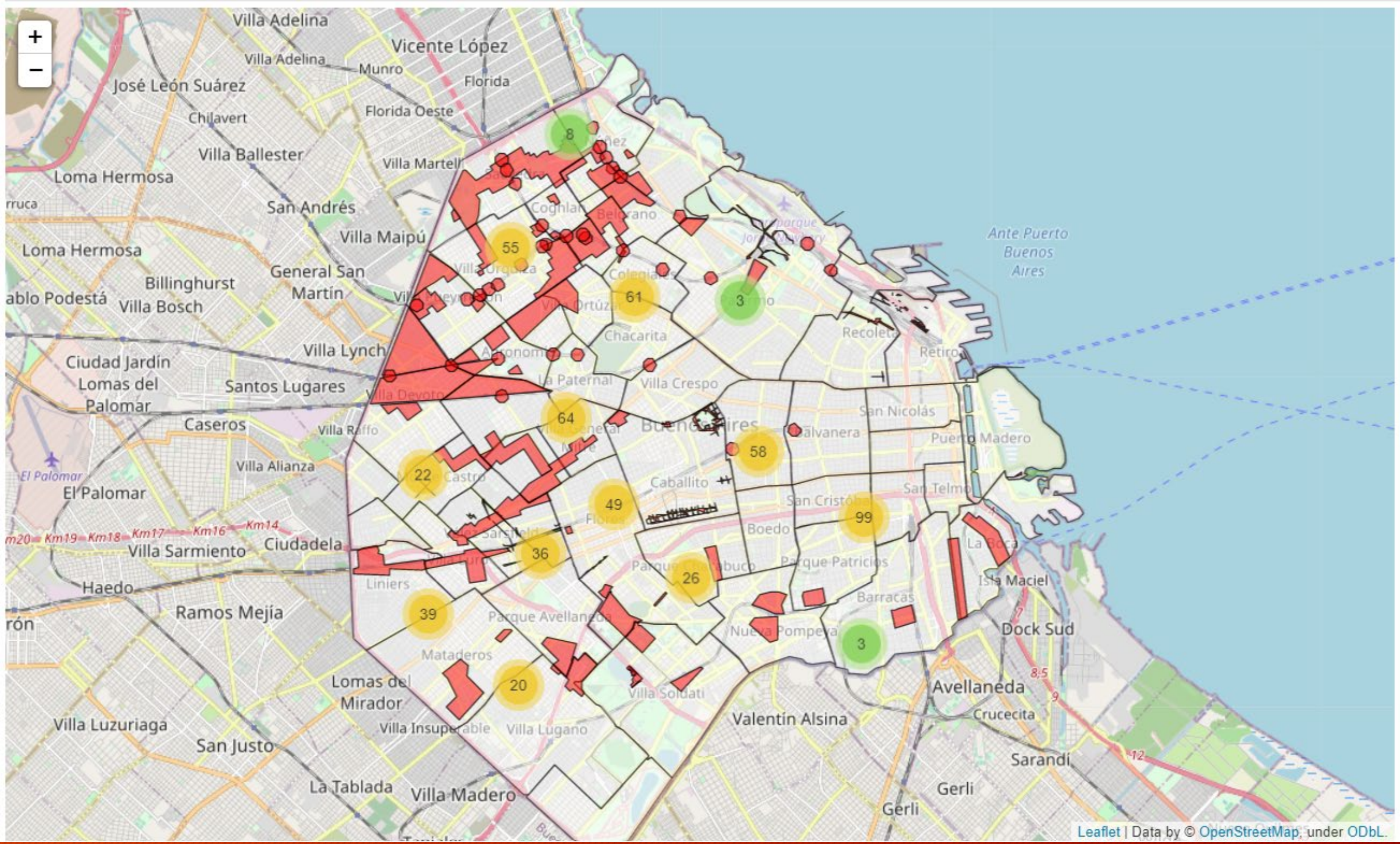
Urban land in waterlogging zones was excluded



Urban land candidates after waterlogging zone filter

Working with the data: Lands and nearby venues

Urban land with more than four “Food” venues were excluded



Urban land candidates after >4 food venues filter

Working with the data: Clustering

5 clusters were created (K-means), according to square meter price and counters of venues

	long	lat	area	price	square_meter_price	restaurant_counts	hotel_counts	other_counts	Clus_Db
count	29.000000	29.000000	29.000000	2.900000e+01	29.000000	29.000000	29.000000	29.000000	29.0
mean	-58.434736	-34.605149	621.379310	1.376828e+06	2163.341923	2.103448	2.517241	0.275862	0.0
std	0.041687	0.024772	586.457794	1.931002e+06	1327.923271	1.205488	0.870988	0.527565	0.0
min	-58.4								
25%	-58.4								
50%	-58.4								
75%	-58.4								
max	-58.3								

	long	lat	area	price	square_meter_price	restaurant_counts	hotel_counts	other_counts	Clus_Db
count	208.000000	208.000000	208.000000	2.080000e+02	208.000000	208.000000	208.000000	208.000000	208.0
mean	-58.463923	-34.620108	440.392286	5.474886e+05	1527.735525	0.538462	0.182692	0.105769	1.0
std	0.0								
min	-58.5								
25%	-58.4								
50%	-58.4								
75%	-58.4								
max	-58.3								

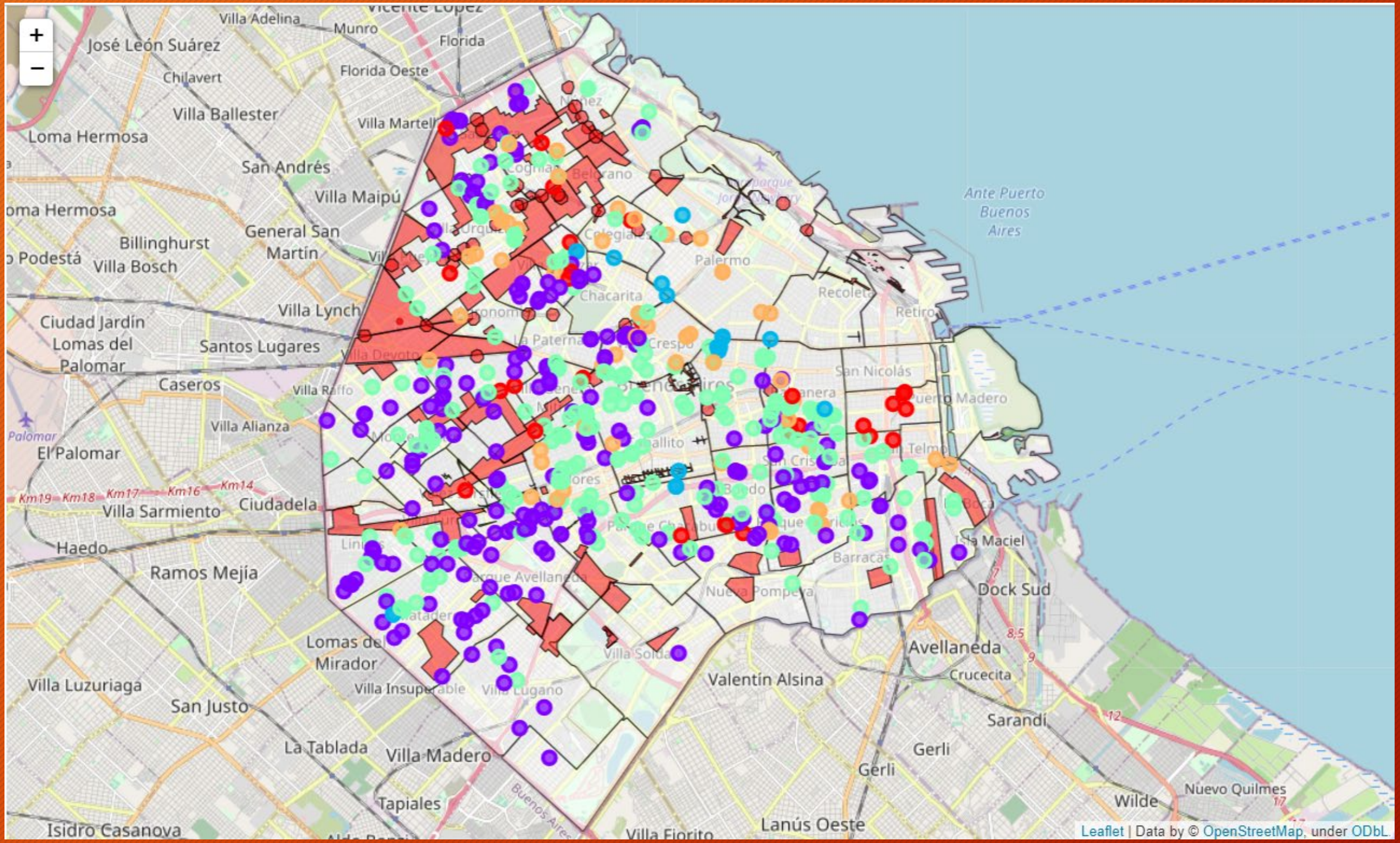
	long	lat	area	price	square_meter_price	restaurant_counts	hotel_counts	other_counts	Clus_Db
count	18.000000	18.000000	18.000000	1.800000e+01	18.000000	18.000000	18.000000	18.000000	18.0
mean	-58.442552	-34.604749	369.607032	8.411111e+05	2876.908611	3.055556	0.333333	2.833333	2.0
std	0.030								
min	-58.512								
25%	-58.450								
50%	-58.436								
75%	-58.425								
max	-58.397								

	long	lat	area	price	square_meter_price	restaurant_counts	hotel_counts	other_counts	Clus_Db
count	230.000000	230.000000	230.000000	2.300000e+02	230.000000	230.000000	230.000000	230.000000	230.0
mean	-58.450322	-34.614832	469.895495	6.843039e+05	1723.619470	2.778261	0.195652	0.108696	3.0
std	0.03								
min	-58.52								
25%	-58.47								
50%	-58.45								
75%	-58.41								
max	-58.36								

	long	lat	area	price	square_meter_price	restaurant_counts	hotel_counts	other_counts	Clus_Db
count	58.000000	58.000000	58.000000	5.800000e+01	58.000000	58.000000	58.000000	58.000000	58.0
mean	-58.445423	-34.599804	369.189655	1.713481e+06	4613.270205	2.500000	0.344828	0.086207	4.0
std	0.035074	0.026740	330.749519	1.707705e+06	1101.051535	1.013072	0.479463	0.283121	0.0
min	-58.509985	-34.639637	75.000000	3.100000e+05	3130.081301	1.000000	0.000000	0.000000	4.0
25%	-58.474872	-34.624172	179.750000	7.775000e+05	3718.054728	2.000000	0.000000	0.000000	4.0
50%	-58.447759	-34.597077	253.000000	1.125000e+06	4457.516340	2.500000	0.000000	0.000000	4.0
75%	-58.412631	-34.576762	420.000000	1.712500e+06	4940.497271	3.000000	1.000000	0.000000	4.0
max	-58.364223	-34.544165	1540.000000	7.500000e+06	7684.210526	4.000000	1.000000	1.000000	4.0

The description of each cluster

Working with the data: Result



Clustered "good" candidates

Conclusion

- The study creates clusters of similar “good” candidates for lands to implementing a restaurant
- Only icludes a reduced list of factors: price, existence of venues of a predefined type.
- Improvements:
  - Include factors like attractiveness of each location, security, social and economic dynamics, the volume of tourists, etc.
  - Personalization according to stakeholdes requirements