

# Coursera Capstone Project

Final Project

Felipe Bahamonde

## Introduction/business Problem

Our client, an important wood crafts shop located in Santiago, specifically in Peñalolen, wants to open a new store given the success. For this, he is evaluating different possibilities within Santiago but still doesn't know where to position himself. Currently Peumayen's store is in Peñalolen because it is a sector where people have a high purchasing power, which favors the sales of Peumayen since the products are extremely expensive and not of first necessity.

It is in this context that Peumayen asks for help to locate his store, requesting that sectors or districts could be good for his new store.

## Data

To solve this problem, the location of different points of the Metropolitan region are used, where different districts, small towns, and important sectors within Santiago and the region are located. These data are currently on page [https://carta-natal.es/ciudades/Chile/Region Metropolitana](https://carta-natal.es/ciudades/Chile/Region_Metropolitana), so it will be necessary to perform web-scraping to obtain such data. After obtaining this location, the location scale (DMS-> Degrees) will be changed to enter the locations to the Foursquare API, in order to obtain local profiles and thus locate the best place.

## Methodology

To solve this problem, the procedure will be: Process the data to obtain the locations in a format that is useful, use the Foursquare API to obtain profiles of the different locations, with this information will be clustered to obtain profiles similar to the that already owns the Peumayen store in Peñalolen.

In figure 1 it is possible to see an example of the web page where the data will be obtained.

Países » Chile » Región Metropolitana

### CIUDADES DE REGIÓN METROPOLITANA, CHILE

Chile

▼

Buscar ciudad

Buscar ciudad

Coordenadas geográficas y zona horaria actual de 273 ciudades de Región Metropolitana, Chile.

Se muestran 1 a 50 de un total de 273 ciudades:

1

2

3

4

5

6

»

Ciudad	Latitud	Longitud	Zona horaria
Alfalfalito	34°05'00"S	70°07'00"W	UT-3:00
Alicia de Penafior	33°37'00"S	70°54'00"W	UT-3:00
Alto de Jahuel	33°44'00"S	70°42'00"W	UT-3:00
Angostura	33°54'00"S	70°44'00"W	UT-3:00
Apoquindo	33°24'00"S	70°32'00"W	UT-3:00
Asiento Viejo	33°04'00"S	70°59'00"W	UT-3:00
Bajos de Mena	33°38'00"S	70°37'00"W	UT-3:00
Banos de Colina	33°12'00"S	70°38'00"W	UT-3:00
Banos del Tupungato	33°24'00"S	69°59'00"W	UT-3:00
Barranca de Piche	33°59'00"S	71°04'00"W	UT-3:00

Figure 1

Additionally, to generate the different clusters, the API radius and the distance of the different points obtained from the web page will be adjusted to the center of Santiago, since the page brings data from nearby cities that could interfere in the creation of clusters, as well as that Peumayen is not interested in positioning himself in other cities.

## Results

In the figure it is possible to observe the dataframe created from the web page, where we can see 273 and peak locations with their coordinates in a scale that does not allow us to use Foursquare, so it is decided to change by a small transformation. As shown in figure 2 and 3.

	Ciudad	Latitud	Longitud	Zona horaria
0	Alfalfalito	34°05'00"S	70°07'00"W	UT-3:00
1	Alicia de Penafior	33°37'00"S	70°54'00"W	UT-3:00
2	Alto de Jahuel	33°44'00"S	70°42'00"W	UT-3:00
3	Angostura	33°54'00"S	70°44'00"W	UT-3:00
4	Apoquindo	33°24'00"S	70°32'00"W	UT-3:00
5	Asiento Viejo	33°04'00"S	70°59'00"W	UT-3:00
6	Bajos de Mena	33°38'00"S	70°37'00"W	UT-3:00
7	Banos de Colina	33°12'00"S	70°38'00"W	UT-3:00
8	Banos del Tupungato	33°24'00"S	69°59'00"W	UT-3:00
9	Barranca de Piche	33°59'00"S	71°04'00"W	UT-3:00
10	Barrancas	33°27'00"S	70°46'00"W	UT-3:00
11	Barrio Bellavista	33°25'58"S	70°38'09"W	UT-3:00
12	Barrio Brazil	33°26'18"S	70°39'59"W	UT-3:00
13	Batuco	33°13'00"S	70°47'00"W	UT-3:00
14	Bollenar	33°46'00"S	70°16'00"W	UT-3:00
15	Buin	33°44'00"S	70°45'00"W	UT-3:00
16	Cajon de los Valle	32°57'00"S	70°46'00"W	UT-3:00
17	Calera de Tango	33°39'00"S	70°49'00"W	UT-3:00
18	Caleu	33°00'00"S	71°00'00"W	UT-3:00
19	Caracha	33°04'00"S	71°01'00"W	UT-3:00
20	Carrascal	33°25'00"S	70°43'00"W	UT-3:00
21	Casas de Chacabuco	33°02'00"S	70°41'00"W	UT-3:00
22	Casas de El Retiro	33°37'00"S	70°50'00"W	UT-3:00
23	Casas de La Esperanza	33°33'00"S	70°53'00"W	UT-3:00

Figure 2

	Ciudad	Latitud	Longitud	Zona horaria	A	B	C	D	E	F	lat	long
0	Alfalfalito	34°05'00"S	70°07'00"W	UT-3:00	70	07'00"W	07	00"W	00	W	-34.083333	-70.116667
1	Alicia de Penafior	33°37'00"S	70°54'00"W	UT-3:00	70	54'00"W	54	00"W	00	W	-33.616667	-70.900000
2	Alto de Jahuel	33°44'00"S	70°42'00"W	UT-3:00	70	42'00"W	42	00"W	00	W	-33.733333	-70.700000
3	Angostura	33°54'00"S	70°44'00"W	UT-3:00	70	44'00"W	44	00"W	00	W	-33.900000	-70.733333
4	Apoquindo	33°24'00"S	70°32'00"W	UT-3:00	70	32'00"W	32	00"W	00	W	-33.400000	-70.533333
5	Asiento Viejo	33°04'00"S	70°59'00"W	UT-3:00	70	59'00"W	59	00"W	00	W	-33.066667	-70.983333
6	Bajos de Mena	33°38'00"S	70°37'00"W	UT-3:00	70	37'00"W	37	00"W	00	W	-33.633333	-70.616667
7	Banos de Colina	33°12'00"S	70°38'00"W	UT-3:00	70	38'00"W	38	00"W	00	W	-33.200000	-70.633333
8	Banos del Tupungato	33°24'00"S	69°59'00"W	UT-3:00	69	59'00"W	59	00"W	00	W	-33.400000	-69.983333
9	Barranca de Piche	33°59'00"S	71°04'00"W	UT-3:00	71	04'00"W	04	00"W	00	W	-33.983333	-71.066667
10	Barrancas	33°27'00"S	70°46'00"W	UT-3:00	70	46'00"W	46	00"W	00	W	-33.450000	-70.766667
11	Barrio Bellavista	33°25'58"S	70°38'09"W	UT-3:00	70	38'09"W	38	09"W	09	W	-33.432778	-70.635833
12	Barrio Brazil	33°26'18"S	70°39'59"W	UT-3:00	70	39'59"W	39	59"W	59	W	-33.438333	-70.666389
13	Batuco	33°13'00"S	70°47'00"W	UT-3:00	70	47'00"W	47	00"W	00	W	-33.216667	-70.783333
14	Bollenar	33°46'00"S	70°16'00"W	UT-3:00	70	16'00"W	16	00"W	00	W	-33.766667	-70.266667
15	Buin	33°44'00"S	70°45'00"W	UT-3:00	70	45'00"W	45	00"W	00	W	-33.733333	-70.750000
16	Cajon de los Valle	32°57'00"S	70°46'00"W	UT-3:00	70	46'00"W	46	00"W	00	W	-32.950000	-70.766667
17	Calera de Tango	33°39'00"S	70°49'00"W	UT-3:00	70	49'00"W	49	00"W	00	W	-33.650000	-70.816667
18	Caleu	33°00'00"S	71°00'00"W	UT-3:00	71	00'00"W	00	00"W	00	W	-33.000000	-71.000000

Figure 3

Once the change is made, a map with all the extracted localities is generated, as shown in Fig. 4.

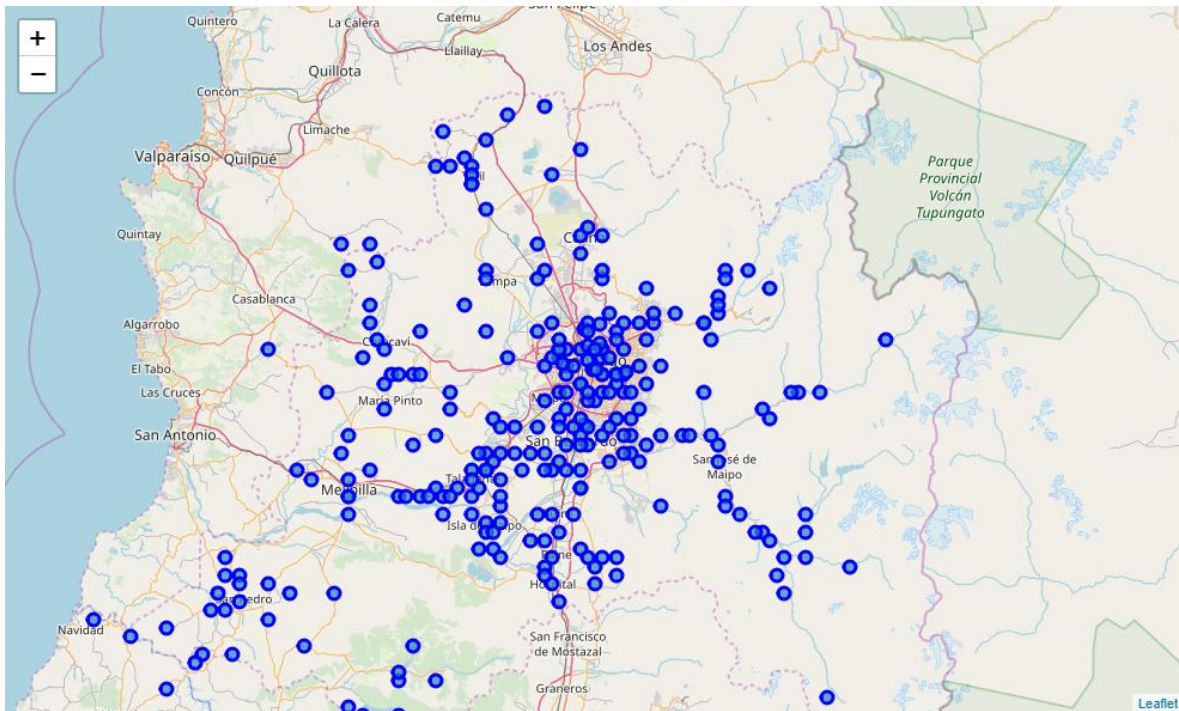


Figure 4

It is possible to see that many points are outside of Santiago since this involves information from the entire metropolitan region. To improve this, we decided to calculate the distance (Euclidean) to the center of Santiago (obtained with Nominatim) leaving those with a distance better than 0.3. as shown in figure 5 (86 places)

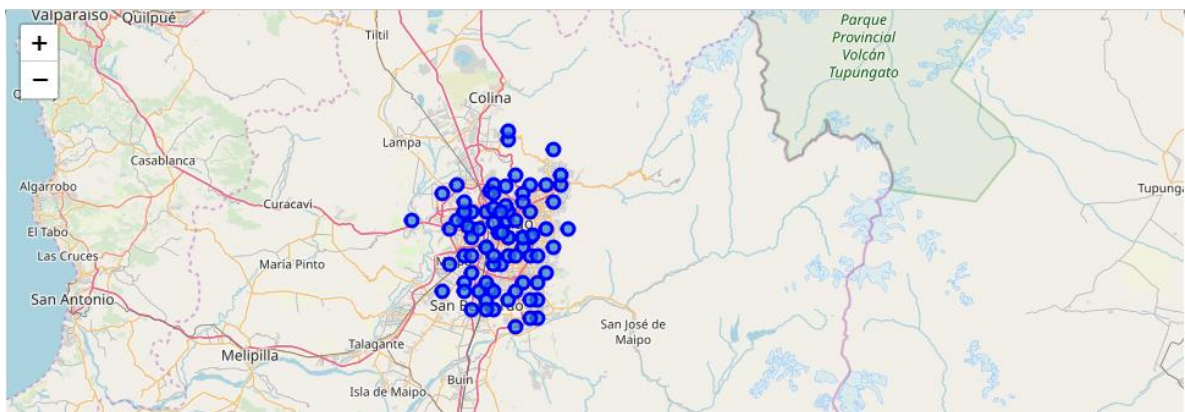


Figure 5

Now, we proceed to clustering to obtain the Peñalolen-like sectors as shown in fig 6

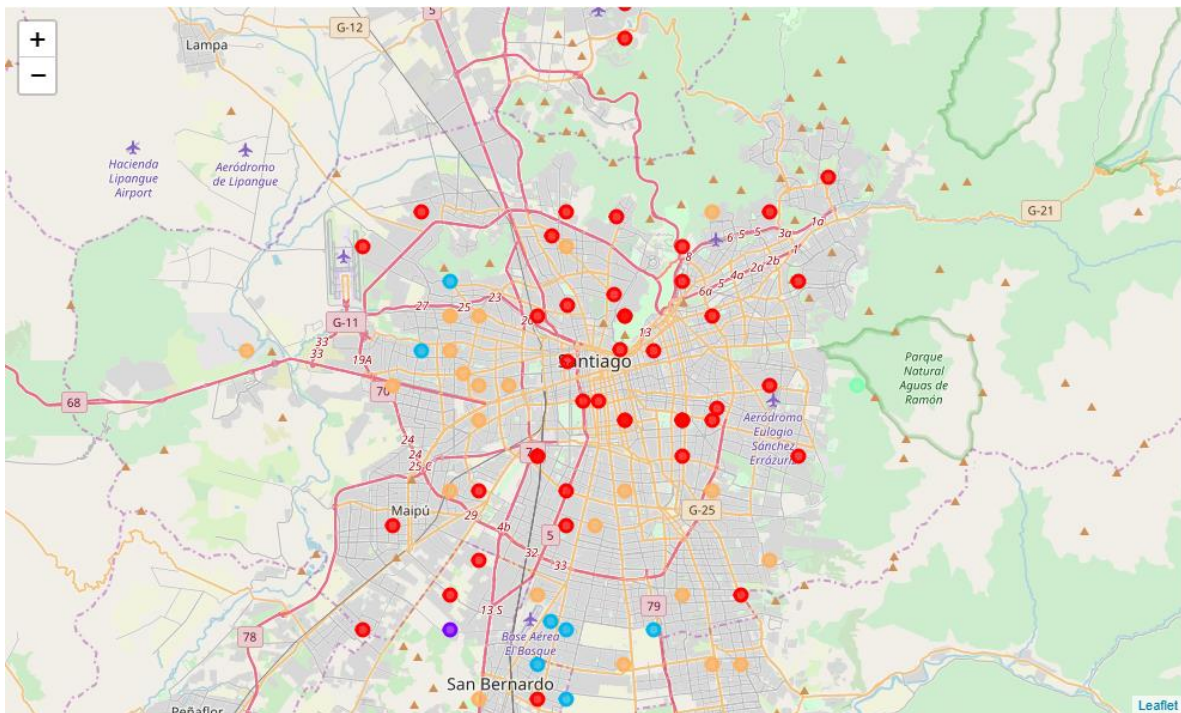


Figure 6

From here it is possible to extract that the northeastern zone has a similar profile to Peñalolen, where more stores and place of consumption are seen, not so the west or south sector. Which have a more urban profile and places like fields or farms.

## Discussion

As mentioned in the previous section, the northeastern areas resemble the peñalolen profile, indicating a greater number of consumer stores than other areas. From this it follows as a suggestion to move your trade in that direction, trying to position yourself in the communes of the counts or providence where the largest number of stores are concentrated.

## Conclusion

From this exercise it was possible to obtain results that agree with the reality of Santiago where the northeastern sector has a greater purchasing power, which is what Peumayen was looking for in this fictitious case. On the other hand, you have that the foursquare's API is not very used here in Chile, so it is suggested for future activities at the same time in this area, replaced by Google API.

For a better detail it could be excluded from the other clusters and repeat the exercise only on the cluster of the north-east zone, to include another level of detail and specify more possible location for Peumayen.