# Venues opportunities in Las Condes (Santiago, Chile)

Germán Donoso Henríquez

February 25, 2021

## 1. Introduction: Business Problem

The purpose of this Project is to evaluate new opportunities for future venues in the city of Santiago, Chile. The client is interested in opening a new venue in Santiago, specifically in the commune of Las Condes. They want to know what type of venue are the best option.

The Santiago city divides its territory into several communes. We are going to explore the main communes of Santiago and then compare with Las Condes to analyze which communes are the most similar to the commune in question. With this information we can determine the most frequent venues that are present in the similar communes but not in Las Condes, so these venues are possible niches or future business opportunities.


Figure 1. Communes of Santiago

## 2. Data

The data that we will use to solve the problem are:

1. **Venues of the different communes** (Foursquare API)
2. **Communes parameters:** Latitude, longitude, area and population of the different communes. This information will be obtained from https://es.wikipedia.org/wiki/Anexo:Comunas_de_Chile
3. **UF value per square meter of a property or real estate** (in this case departments) of the communes. The information source is https://www.buenainversion.cl/blog/valor-metro-cuadrado/

**Use of Data:**

1. First, we are going to load the commune's parameters (latitude, longitude, area, population and UF/m²). Then we wrangling this data.
2. We assume that the commune area is circular, so we can obtain a characteristic radius per commune. We will use 3/4 of that radius, limiting the

exploration radius to 5 km, to explore with Foursquare API the several communes and get the venues around them.

3. Then we explore the communes to obtain the venues with the Foursquare API and normalize the venues per commune by the exploration area (circular area, function of exploration radius) to get the different types of venues per km² for each commune.
4. We use the one hot representation of the normalized venues to group the communes into six groups and visualize the results in a folium map.
5. We want to know which are the three most similar communes to Las Condes based on the one hot representation vector and the similarity in cost of living or economic situation in the commune, for this we use the UF/m². We use both indicators to get the most similar communes.
6. Finally, we get the ten most frequent venues for Las Condes and the three similar communes and analyze which venues are present in those communes but not in Las Condes, so those venues are possible niches or business opportunities.

# 3. Methodology and Data Analysis

## 3.1. Load commune parameters

We load the communes of Chile. The data source that we use give us different parameters of the communes. We select the parameters that we want: Name, Province name, Area ($km^2$), Population, Latitude and Longitude.

There are 346 communes in Chile, so we extract only the communes of Santiago. Finally, we convert the format of the latitude and longitude and then transform from degrees minutes seconds to decimal degrees, because that format uses the Foursquare API.

## 3.2. Load UF/$m^2$ of a real state in the communes

We load the UF/$m^2$ of a department for each commune and select only the commune with the data of 2019 (most recent). We convert the format of the commune names to have the same name as the table of the first commune parameters that we loaded.

Then we join both data to obtain one table with the commune's information.

The next table shows the data mentioned above.

Table 1. Data of the communes of Santiago

| Commune | Area (km$^2$) | Population | Latitude | Longitude | UF/m$^2$ |
|---|---|---|---|---|---|
| Santiago | 23,2 | 503.147 | -33,44 | -70,66 | 63,9 |
| Cerrillos | 21 | 88.956 | -33,50 | -70,72 | 41,8 |
| Conchalí | 10,7 | 139.195 | -33,38 | -70,68 | 54,9 |
| Estación Central | 15 | 206.792 | -33,46 | -70,70 | 52,3 |
| Huechuraba | 44,8 | 112.527 | -33,37 | -70,63 | 50 |
| Independencia | 7 | 142.065 | -33,41 | -70,67 | 53 |
| La Cisterna | 10 | 100.434 | -33,53 | -70,66 | 51,8 |
| La Florida | 70,2 | 402.433 | -33,53 | -70,54 | 54,9 |
| La Reina | 23 | 100.252 | -33,44 | -70,53 | 78 |
| Las Condes | 99 | 330.759 | -33,42 | -70,58 | 94,9 |
| Lo Barnechea | 1024 | 124.076 | -33,35 | -70,52 | 92,1 |
| Macul | 12,9 | 134.635 | -33,49 | -70,60 | 58 |
| Maipú | 135,5 | 578.605 | -33,52 | -70,77 | 49,9 |
| Ñuñoa | 16,9 | 250.192 | -33,45 | -70,60 | 79,1 |
| Peñalolén | 54 | 266.798 | -33,49 | -70,53 | 62 |
| Providencia | 14,3 | 157.749 | -33,44 | -70,62 | 91,6 |
| Pudahuel | 197 | 253.139 | -33,43 | -70,72 | 42,9 |
| Quilicura | 58 | 254.694 | -33,36 | -70,73 | 40 |
| Quinta Normal | 13 | 136.368 | -33,43 | -70,70 | 44,5 |
| Recoleta | 16 | 190.070 | -33,41 | -70,64 | 55,1 |
| San Joaquín | 9,7 | 103.485 | -33,49 | -70,63 | 55,2 |
| San Miguel | 10,0 | 133.059 | -33,49 | -70,65 | 55,7 |
| Vitacura | 28,3 | 96.774 | -33,40 | -70,60 | 103 |

In the next map we can see the communes of Santiago. The red point is the commune of Las Condes.

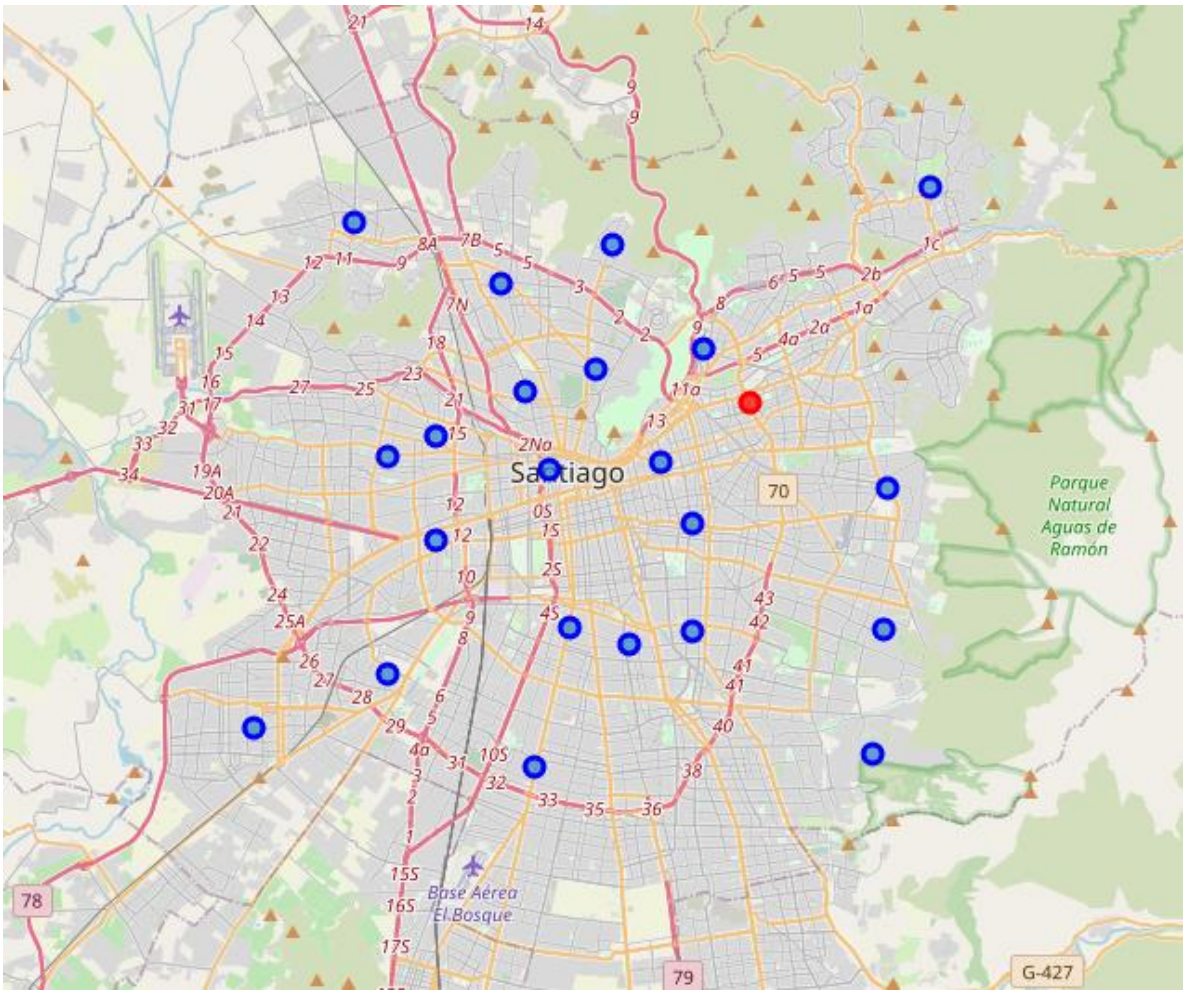


Figure 2. Communes of Santiago

### 3.3. Characteristic radius per commune

We assume that the area of the communes is circular, with this assumption we calculate the characteristic radius per commune, because we need an exploration radius to explore the venues of the communes with the Foursquare API.

We will use ¾ of the obtained radius to explore the communes and limit the radius to 5 km. The exploration radius that we calculate give us also the exploration area, which we will use later on.

### 3.4. Explore the venues with the Foursquare API

Using the Foursquare credentials, the latitude, longitude and radius of the communes, we obtain the venues that are present in each commune. We obtained

a total of 2170 venues for the communes of Santiago. The next table shows the number of venues per commune.

Table 2. Number of venues per commune

| Commune | #Venues |
|---|---|
| Cerrillos | 45 |
| Conchalí | 21 |
| Estación Central | 69 |
| Huechuraba | 43 |
| Independencia | 34 |
| La Cisterna | 53 |
| La Florida | 45 |
| La Reina | 65 |
| Las Condes | 217 |
| Lo Barnechea | 92 |
| Macul | 73 |
| Maipú | 145 |
| Peñalolén | 58 |
| Providencia | 200 |
| Pudahuel | 179 |
| Quilicura | 59 |
| Quinta Normal | 32 |
| Recoleta | 33 |
| San Joaquín | 27 |
| San Miguel | 71 |
| Santiago | 224 |
| Vitacura | 225 |
| Ñuñoa | 160 |

Then we convert the venues per commune to a one hot encoding and group the venues by commune, so we obtain how many venues have each commune for each type of venue. In total there are 263 different venues. Finally, we normalize the venues per commune by the exploration area to get the venues per $km^2$.

### 3.5. K-Means Algorithm

With this information, we use the K-Means Algorithm to group the communes into six groups. The groups that we obtained are showed in the next table.

Table 3. Cluster for each commune

| Commune | Cluster |
|---|---|
| Cerrillos | 1 |
| Conchalí | 1 |
| Estación Central | 1 |
| Huechuraba | 1 |
| Independencia | 4 |
| La Cisterna | 3 |
| La Florida | 1 |
| La Reina | 1 |
| Las Condes | 1 |
| Lo Barnechea | 1 |
| Macul | 3 |
| Maipú | 1 |
| Peñalolén | 1 |
| Providencia | 0 |
| Pudahuel | 1 |
| Quilicura | 1 |
| Quinta Normal | 1 |
| Recoleta | 1 |
| San Joaquín | 1 |
| San Miguel | 3 |
| Santiago | 2 |
| Vitacura | 2 |
| Ñuñoa | 5 |

We can see the cluster of the communes in the following map.

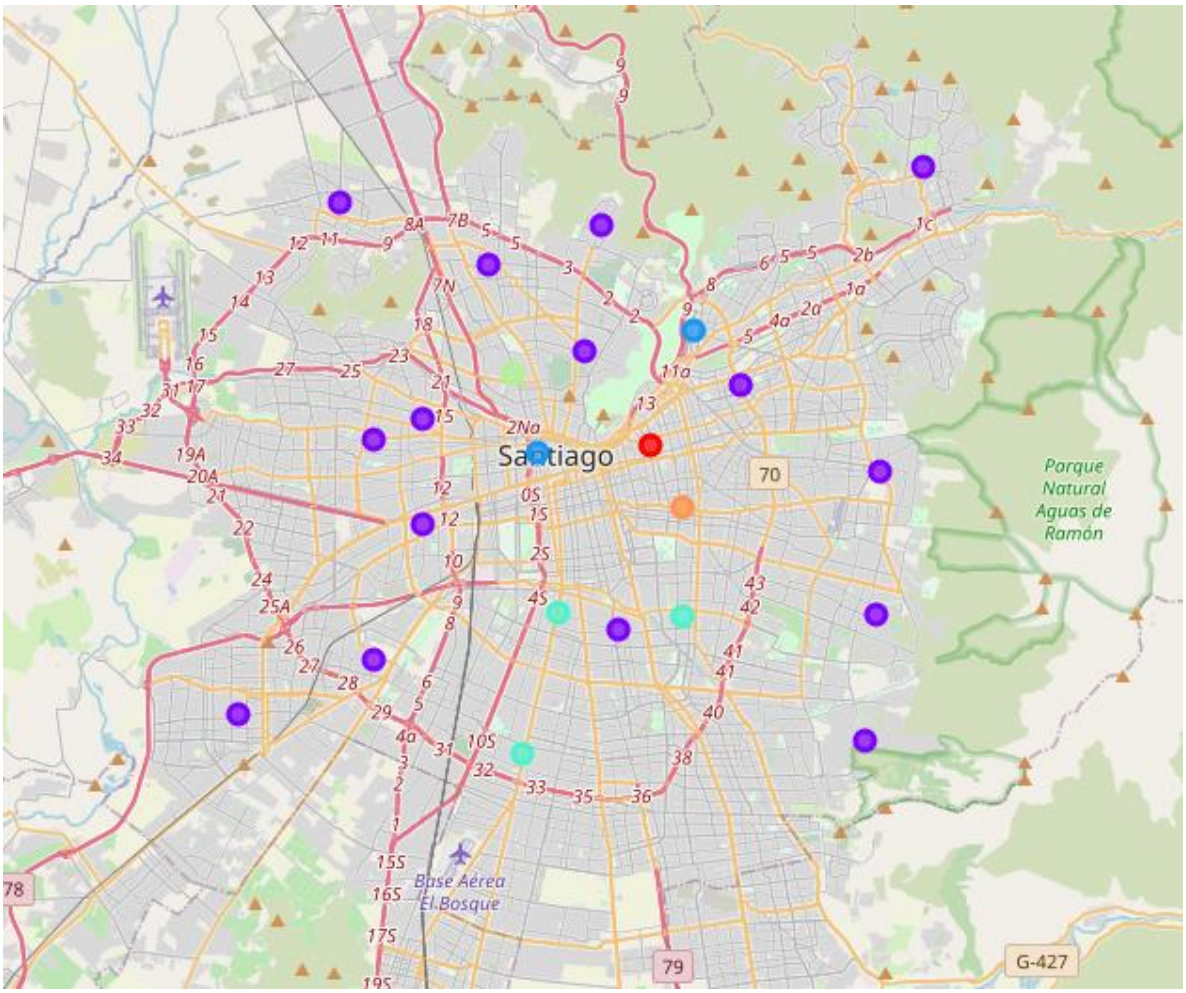


Figure 3. Cluster of the communes

We observe that Las Condes belong to the most popular cluster (number 1), with other 14 communes in the same cluster.

### 3.6. Euclidean Distance

We calculate the Euclidean distance between Las Condes and the other communes using the normalized vectors we used to get the cluster. We obtained that the communes of Las Condes cluster are the same as the 14 most similar communes to Las Condes according the Euclidean distance. We select the 8 most similar communes. The next figure shows the results.

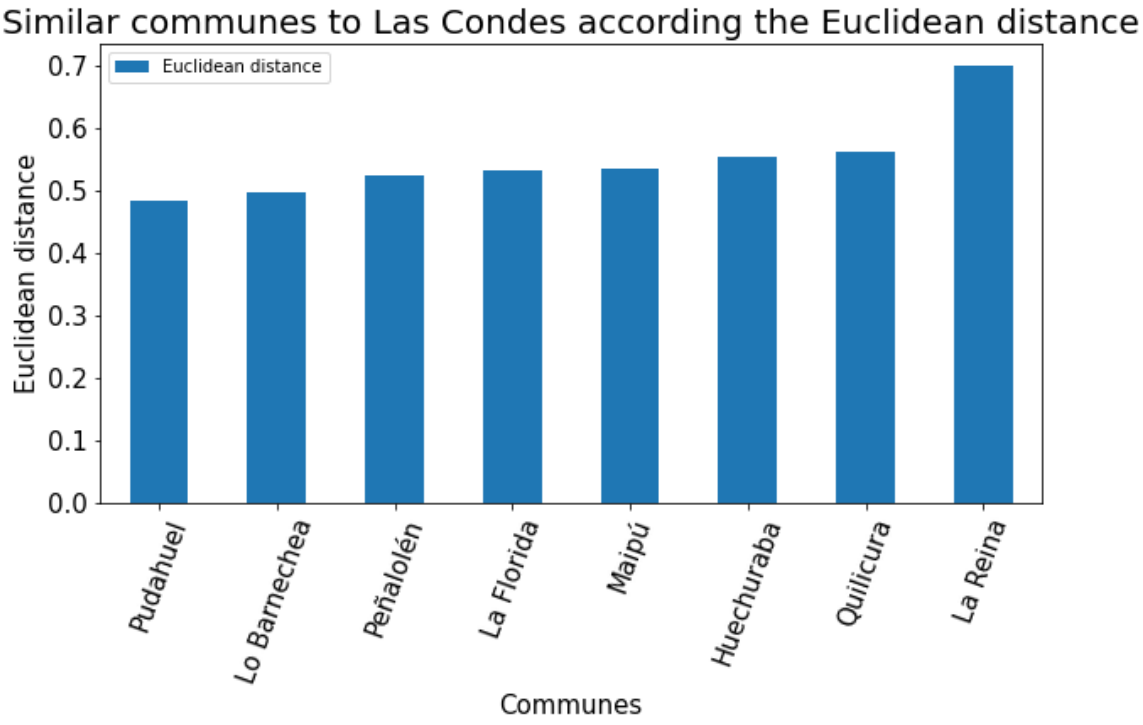


Figure 4. Most similar communes to Las Condes using Euclidean distance

We see that the Euclidean distance is similar in the communes, with exception of the last commune (La Reina), which has a greater distance.

### 3.7. UF/m$^2$ difference

We also use the UF/m² difference because this gives us a reference of the similarity in the economic situation of the people who live in the communes. We assume that the target audience of the future venue has similar preferences depending on the economic situation. The next chart shows the results.

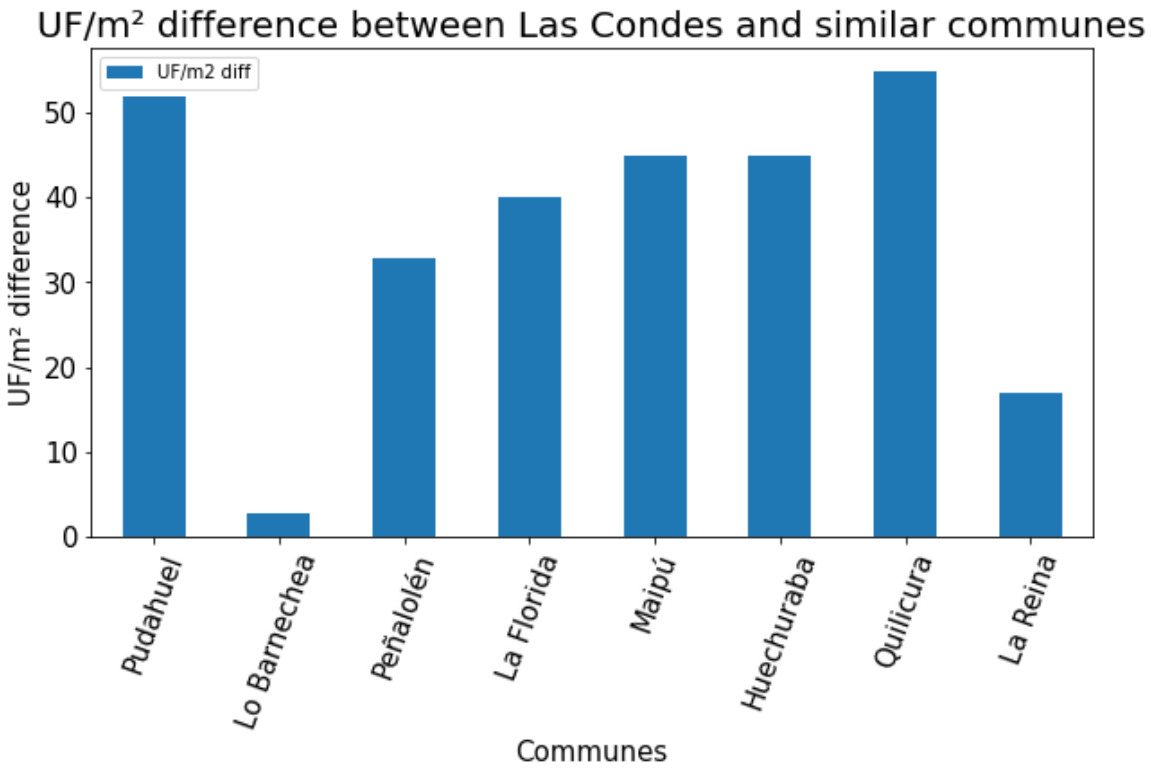


Figure 5. UF/m$^2$ difference

If we analyze the percentage difference, we can see that six communes exceed 35% and some reach values close to 55%. Lo Barnechea and La Reina have values lower than 20%.

### 3.8. Final indicator

To obtain the most similar communes to Las Condes, we multiply both indicators (Euclidean distance and UF/m$^2$ difference) to obtain only one indicator.

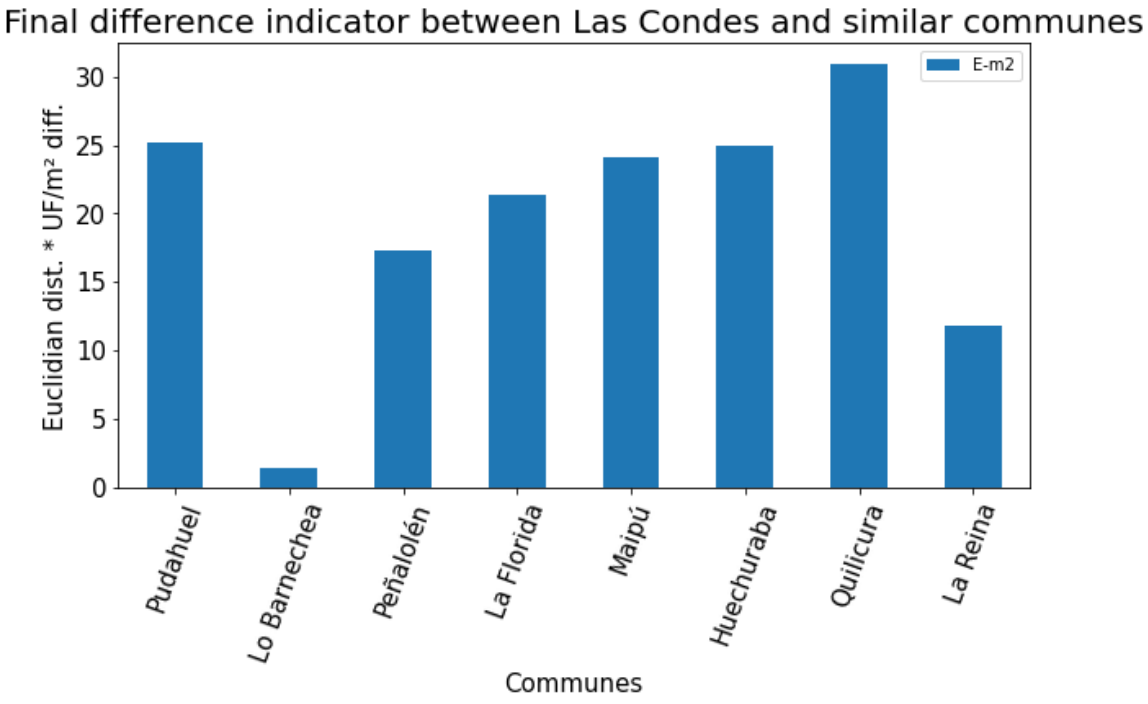


Figure 6. Final difference indicator

The three most similar communes to Las Condes according to the new indicator are:

1. **Lo Barnechea**
2. **La Reina**
3. **Peñalolén**

## 3.9. Ten most frequent venues

First, we select the 15 most frequent venues. To clean the data, we group the venues that are the same but in different language (English and Spanish), for example 'Coffee Shop' and 'Café'. We delete the venues that are useless to the analysis, for example the 'Bus Station' or 'Park'. After we clean the data, we select the ten most frequent venues per commune and join the venues of the four communes into one table.

On the other hand, we divide the values of each commune by the maximum value (frequency) in that commune. In this way we can compare how many venues there are in relation to the most frequent venue per commune. These values allow us understand the importance of some venues in a specific commune and compare them with the other communes.

# 4. Results and Discussion

The following chart shows the ten most frequent venues per commune and its frequency.

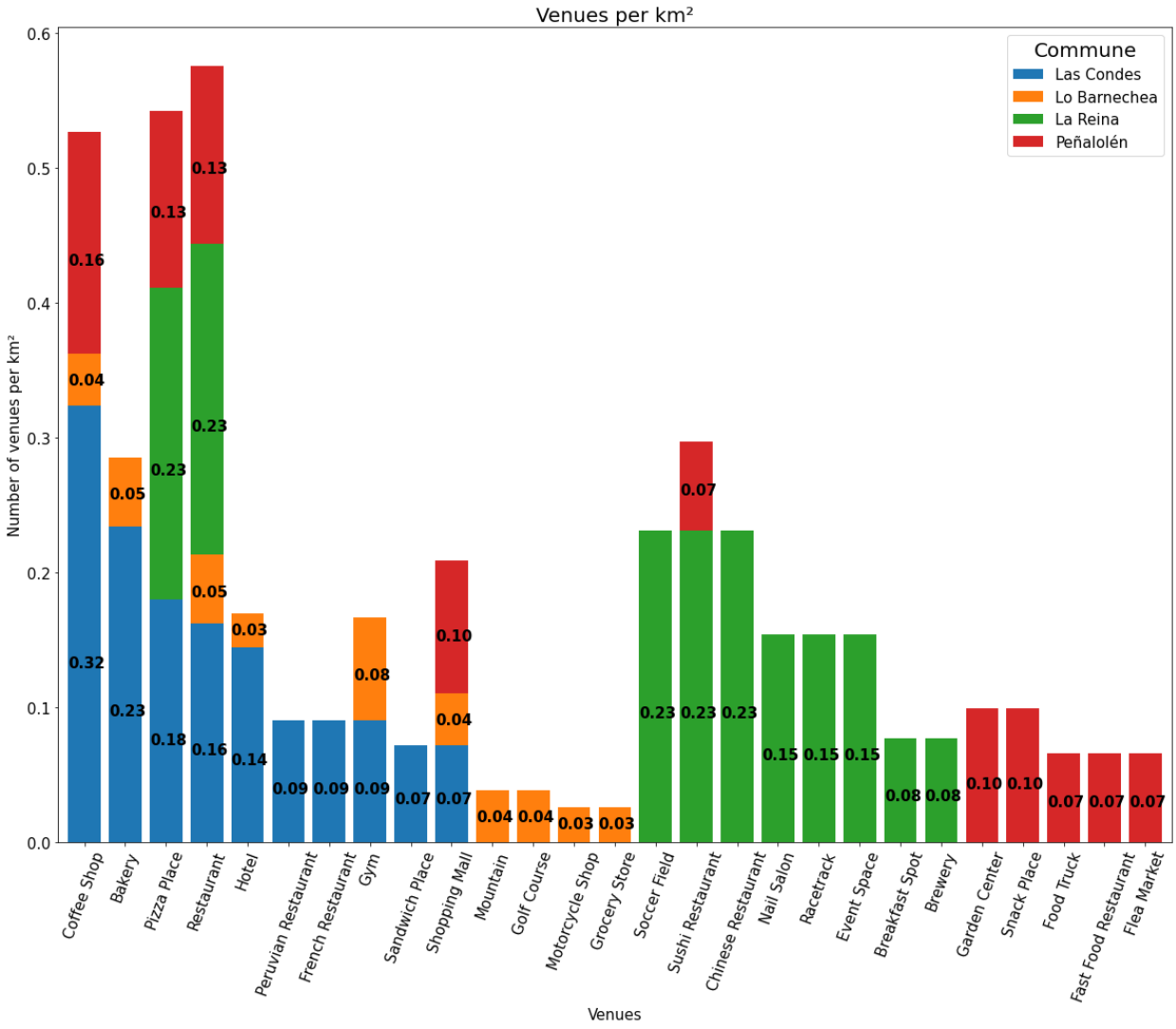


Figure 7. Ten most frequent venues per commune

The previous chart shows us the ten most frequent venues per commune stacked. We can see that several frequent venues of Las Condes are also present in other communes, with some exceptions.

The three most frequent stacked venues are: Restaurant, Pizza Place and Coffee Shop, and they are three of the four most frequent venues of Las Condes, so maybe these venues are not a good option.

We want to find venues that are frequent in the similar communes but not in Las Condes. If we continue seeing the venues of Las Condes, there is one case that is more attractive, this is the Shopping Mall. This venue is frequent in Las Condes and it seems to be relatively frequent for the commune of Lo Barnechea and Peñalolén. The next chart shows the relative frequency according to the max frequency of the venues for each commune.

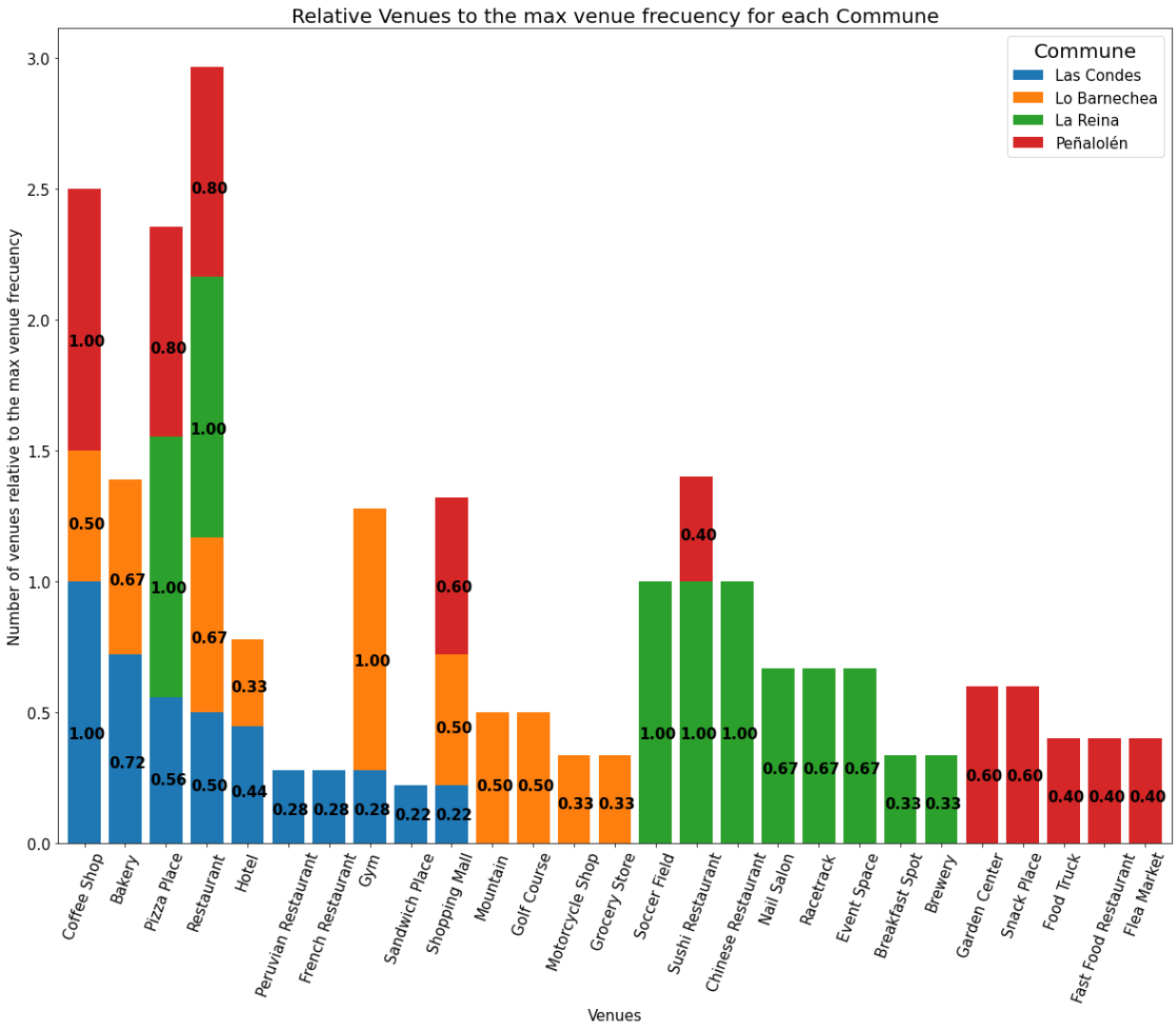


Figure 8. Relative frequency for each venue

In this chart we can see how many venues there are in relation to the most frequent venue per commune. So, if a value is 0.5 for example, this say that there are 0.5 venues for each most frequent venue in that commune. This gives us a better understand of the relative importance of the venues in each commune.

We continue with the Shopping Mall. Now we see that Lo Barnechea and Peñalolén have more than Las Condes, so maybe this tell us that there is space to put another shopping mall, it could be an option to a great stakeholder with a good economic situation.

The same with the Gym, in Lo Barnechea this is the most frequent venue (1 value), in contrast to Las Condes there are 0.28 gyms for each most common venue (Coffee Shop), so the Gym is another option for future venues.

Then we observe the venues that are not present in Las Condes but are frequent in the similar communes. If we see the venues of **Lo Barnechea**, there are four possible venues: Mountain, Golf Course, Motorcycle Shop and Grocery Store. We will focus in venues that are possible shop or do not require very large spaces, so we rule out the Mountain because this depend of the geography of the commune, and the Golf Course because we need a large area for this one. Then we have two possible venues: Grocery Store and Motorcycle Shop.

We continue with **La Reina**. This commune shows three main options: Soccer field, Sushi Restaurant and Chinese Restaurant. The Soccer Field require a considerable area compared to shops, but it can be located within the city, so it does not rule out as an option, this venue is also one of the most frequent venues of La Reina. The Sushi Restaurant is frequent also in Peñalolén, so this is a great possible venue. Finally, the Chinese Restaurant is another of the most frequent venues of the commune, so it could be a good option. Other options but not as frequent as the previous mentioned are: Nail Salon, Racetrack, Event Space, Breakfast Spot and Brewery.

Finally, we have the venues of **Peñalolén**. In this commune we have two main venues that are not present in Las Condes: Garden Center and Snack Place, both of them are feasible option. Then we have other three frequent venues: Food Truck, Fast Food Restaurant and Flea Market.

The client or stakeholder have different option, then they can rule out the venues that are out of their budget, for example the Shopping Mall could exceed the client's budget.

Perhaps the client have some possible location point where they'd want to place the venue. Then we can evaluate these points and explore for different option venues that the client want to analyze (with the Foursquare API) and compare which location point have less of the same category venue that we want to place. So, for example, if the client wants to place a Chinese Restaurant, then we can look for strategic points that have fewer Chinese Restaurants around them.

# 5. Conclusion

The purpose of this Project is to evaluate new opportunities for future venues in the city of Santiago, Chile. The client is interested in opening a new venue in Santiago, specifically in the commune of Las Condes.

So, we analyzed the data and obtained that the three most similar communes to Las Condes are: **Lo Barnechea, La Reina and Peñalolén**. Then we extracted the ten most frequent venues per commune and search the venues that are frequent in the similar communes but not in Las Condes. We analyzed these venues and got the following results:

- The **Shopping Malls and Gyms** are venues that are frequent in Las Condes but it could be space to place one of them if we compare in a relative way with the other communes.
- **Lo Barnechea:** the main possible venues are **Grocery Store and Motorcycle Shop.**
- **La Reina:** this commune suggests three main options; **Soccer field, Sushi Restaurant and Chinese Restaurant.** Other less frequent options are; Nail Salon, Racetrack, Event Space, Breakfast Spot and Brewery.
- **Peñalolén:** the two main venues are **Garden Center and Snack Place**. Other good options are Food Truck, Fast Food Restaurant and Flea Market.

We presented seven main venues according the three communes and other ten more options that are good options but not as frequent as the main options. We have a total of **17 possible venues** that the client can choose.

Then the client can rule out the venues that are out of their budget. Finally, we can continue the analysis looking for strategic points in Las Condes that have fewer venues nearby of the same category that the client wants.