

Capstone Project — The Battle of Neighborhoods Santiago

BY ANÍBAL FERNÁNDEZ



Analyze and
determine where
could be a good
place to put a
restaurant in
Santiago

SANTIAGO IS THE CAPITAL CITY
AND THE LARGEST OF CHILE

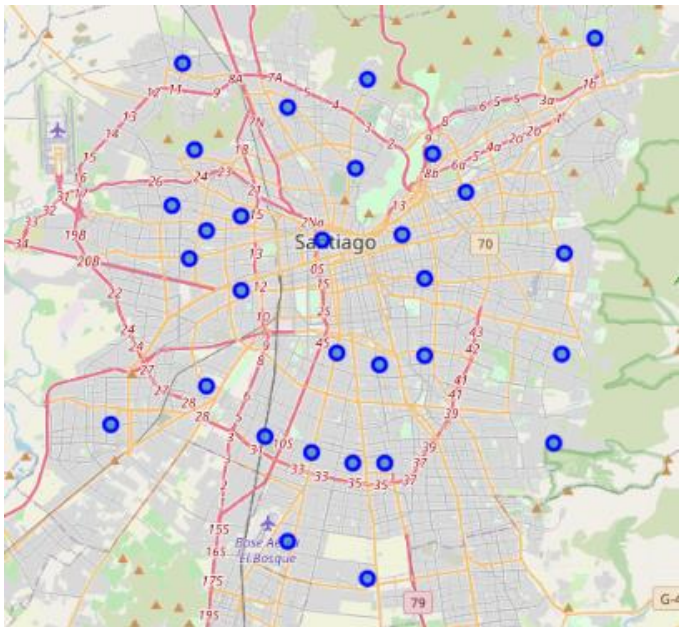
Data acquisition and cleaning

We will Scrap Santiago communes* table from Wikipedia and transform the coordinates of the major communes using in google maps

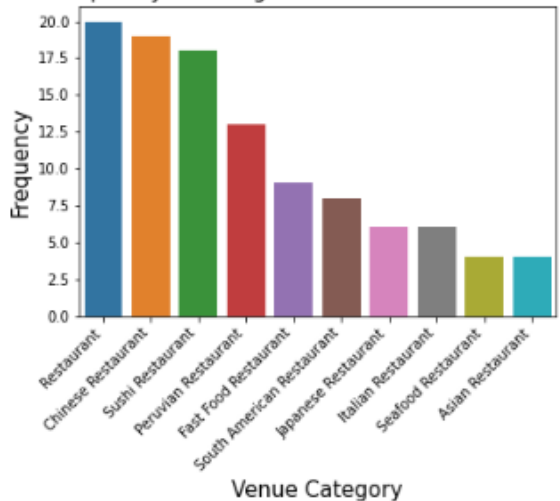
Datasource:
https://es.wikipedia.org/wiki/Anexo:Comunas_de_Chile_por_poblaci%C3%B3n

*The communes are the minor and basic administrative division of Chile.

0	Santiago	23.2	404495	17435.1	-33.437222	-70.857222
1	Cerrillos	21.0	80832	3849.1	-33.500000	-70.716667
2	Cerro Navia	11.0	132622	12056.5	-33.422000	-70.735000
3	Conchalí	10.7	126955	11865.0	-33.380000	-70.675000
4	El Bosque	14.2	162505	11444.0	-33.567000	-70.675000
5	Estación Central	15.0	147041	9802.7	-33.459000	-70.699000
6	Huechuraba	44.8	98671	2202.5	-33.368000	-70.634000
7	Independencia	7.0	100281	14325.9	-33.413000	-33.413000
8	La Cisterna	10.0	90119	9011.9	-33.529000	-70.663000
9	La Florida	70.2	366916	5226.7	-33.525000	-70.538000
10	La Granja	10.0	116571	11657.1	-33.533333	-70.625000
11	La Pintana	30.6	177335	5795.3	-33.583000	-70.634000
12	La Reina	23.0	92787	4034.2	-33.443000	-70.532000
13	Las Condes	99.0	294838	2978.2	-33.416667	-70.583333
14	Lo Barnechea	1024.0	105833	103.4	-33.350000	-70.516667
15	Lo Espejo	7.0	98804	14114.9	-33.522000	-70.687000
16	Lo Prado	7.0	96249	13749.9	-33.445000	-70.726000
17	Macul	12.9	116534	9033.6	-33.487000	-70.604000
18	Maipú	135.5	521657	3849.6	-33.516667	-70.766667
19	Nuñoa	16.9	208237	13321.7	-33.454000	-70.604000
20	Pedro Aguirre Cerda	10.0	101174	10117.4	-70.604000	-70.633667
21	Peñalolén	54.0	241599	4474.1	-33.486189	-70.533333
22	Providencia	14.3	142079	9935.6	-33.435000	-70.616000
23	Pudahuel	197.0	230293	1197.0	-33.433333	-70.716667
24	Quilicura	58.0	210410	3627.8	-33.361000	-70.729000
25	Quinta Normal	13.0	110026	8463.5	-33.427000	-70.699000
26	Recoleta	16.0	157851	9865.7	-33.406000	-70.640000
27	Renca	24.0	147151	6131.3	-33.398000	-70.723000
28	San Joaquín	9.7	94492	9741.4	-33.491000	-70.628000
29	San Miguel	10.0	107954	10795.4	-33.485833	-70.649444
30	San Ramón	7.0	82900	11842.8	-33.533333	-70.641667
31	Vitacura	28.3	85384	3017.1	-33.400000	-70.600000



10 Most Frequently Occuring Venues in the Communes of Santiago



Using Folium to visualize the geographic locations of the different communes

We look the most popular type of restaurant

```
[20]: Santiago_5_Commune_Venues_Top10 = Santiago_Venues_only_restaurant['Venue Category'].value_counts()[0:10].to_frame(name='frequency')
Santiago_5_Commune_Venues_Top10=Santiago_5_Commune_Venues_Top10.reset_index()
#Santiago_5_Comm_Venues_Top10

Santiago_5_Commune_Venues_Top10.rename(index=str, columns={"index": "Venue_Category", "frequency": "Frequency"}, inplace=True)
Santiago_5_Commune_Venues_Top10
```

[20]:

	Venue_Category	Frequency
0	Restaurant	20
1	Chinese Restaurant	19
2	Sushi Restaurant	18
3	Peruvian Restaurant	13
4	Fast Food Restaurant	9
5	South American Restaurant	8
6	Japanese Restaurant	6
7	Italian Restaurant	6
8	Seafood Restaurant	4
9	Asian Restaurant	4

```
[24]: Santiago_Venues_restaurant
```

```
[24]: Commune
Cerrillos      2
Conchalí      1
Estación Central  3
La Cisterna    6
La Pintana    1
La Reina      3
Las Condes    11
Lo Barnechea  5
Lo Prado      5
Macul         3
Maipú         2
Peñalolén     2
Providencia   6
Pudahuel      2
Quilicura     7
Recoleta      2
San Miguel    10
Santiago      27
Vitacura      15
Ñuñoa         9
Name: Venue Category, dtype: int64
```

Which communes have more restaurants

We try to cluster the communes based on the venue categories and use K-Means clustering

```
[33]: # set number of clusters
kclusters = 5

Santiago_grouped_clustering = Santiago_grouped.drop('Commune', 1)

# run k-means clustering
kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(Santiago_grouped_clustering)

# check cluster labels generated for each row in the dataframe
kmeans.labels_[0:10]

[33]: array([4, 0, 4, 1, 2, 1, 4, 1, 2, 4], dtype=int32)

[34]: # add clustering labels
Commune_venues_sorted.insert(0, 'Cluster Labels', kmeans.labels_)

Santiago_merged = Santiago

# merge toronto_grouped with toronto_data to add Latitude/Longitude for each neighborhood
Santiago_merged = Santiago_merged.join(Commune_venues_sorted.set_index('Commune'), on='Commune')

Santiago_merged# check the last columns!
```

```
[34]:
```

	Commune	Surface (km2)	Population	Density (hab./km2)	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Santiago	23.2	404495	17435.1	-33.437222	-70.657222	1.0	Peruvian Restaurant	Japanese Restaurant	Sushi Restaurant	Asian Restaurant	Chinese Restaurant	South American Restaurant	Restaurant	Italian Restaurant	Falafel Restaurant	Fast Food Restaurant
1	Cerrillos	21.0	80832	3849.1	-33.500000	-70.716667	4.0	Restaurant	Fast Food Restaurant	Thai Restaurant	Italian Restaurant	Argentinian Restaurant	Asian Restaurant	Chinese Restaurant	Falafel Restaurant	French Restaurant	Greek Restaurant
2	Cerro Navia	11.0	132622	12056.5	-33.422000	-70.735000	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	Conchalí	10.7	126955	11865.0	-33.380000	-70.675000	0.0	Fast Food Restaurant	Thai Restaurant	Italian Restaurant	Argentinian Restaurant	Asian Restaurant	Chinese Restaurant	Falafel Restaurant	French Restaurant	Greek Restaurant	Indian Restaurant
4	El Bosque	14.2	162505	11444.0	-33.567000	-70.675000	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
5	Estación Central	15.0	147041	9802.7	-33.459000	-70.699000	4.0	Japanese Restaurant	Argentinian Restaurant	Restaurant	Italian Restaurant	Asian Restaurant	Chinese Restaurant	Falafel Restaurant	Fast Food Restaurant	French Restaurant	Greek Restaurant

Conclusion

- We determine that the most tentative communes to put a restaurant are Santiago, Vitacura and Las condes
- The largest frequency of restaurant are restaurant (not classified as a specific cuisine), Chinese restaurant and sushi restaurant.
- We not have information about Cerro Navia, El bosque, Huechuraba, Independencia, La Florida, La Granja, Lo espejo, Pedro Aguirre Cerda, Quinta Normal, Renca, San Joaquín, San Ramón