DATA Section.

The datasets that are going to be used for this study/analysis come from 2 main sources:

a) Figure 1 has been extracted from the following web:
https://www.tinsa.cl/sala-de-prensa/notas-de-prensa/precio-m2-comunas-santiago/

This image had to be converted to a csv file in order to be readable, so the following service was used:
https://convertio.co/es/ocr/

In [126]:
```
### Retrieveng and Wrangling data
import pandas as pd
file1 = '/Users/diegop/Desktop/Capstone_Project/new.csv'
df_new = pd.read_csv(file1, decimal=',')
df_new.head()
```

Out[126]:

	Departamentos nuevos
Comuna	UF por m2
Vitacura	106,9
Las Condes	100,6
Providencia	97,5
Lo Barnechea	93,2

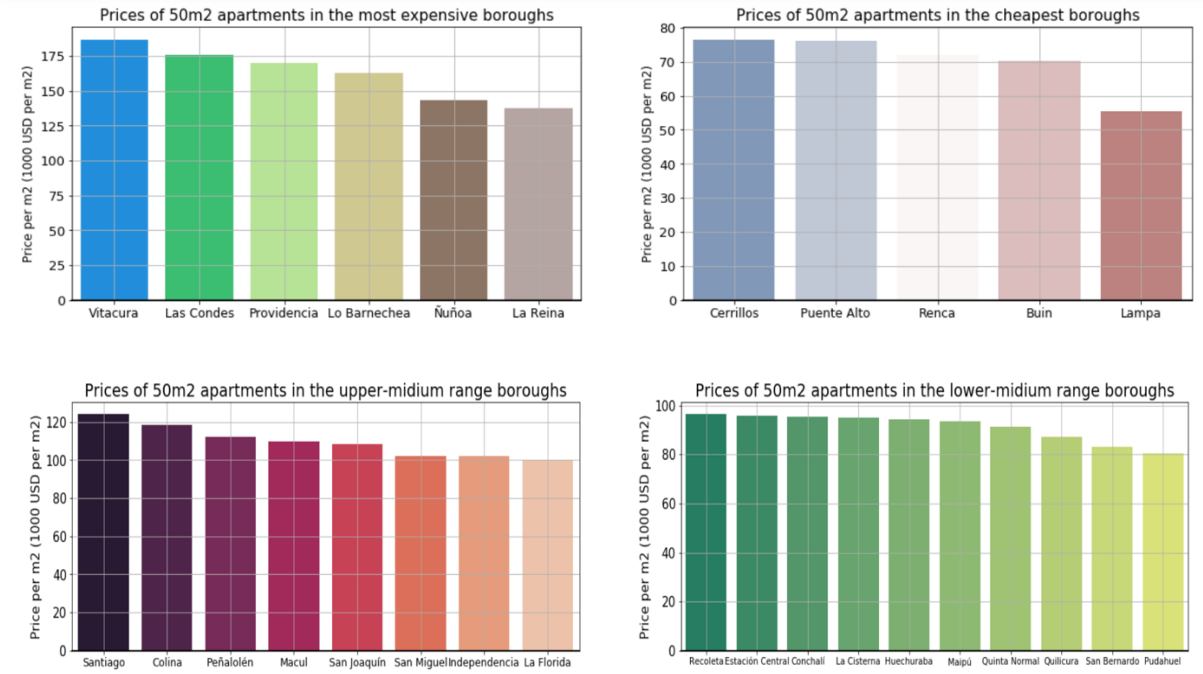
Once the file was successfully converted, the UF was transformed to USD.
A thorough Data wrangling process allowed to obtain a clean and neat dataframe:

In [133]:
```
# Adding prices of new apartments with different square meters.

df_new['30m2'] = df_new['USD/m2']*30
df_new['50m2'] = df_new['USD/m2']*50
df_new['80m2'] = df_new['USD/m2']*80
df_new['100m2'] = df_new['USD/m2']*100

df_new.style.set_caption("Prices in Thousand dollars")
df_new.head()
```

Out[133]:

	USD/m2	30m2	50m2	80m2	100m2
Vitacura	3.73081	111.9243	186.5405	298.4648	373.081
Las Condes	3.51094	105.3282	175.5470	280.8752	351.094
Providencia	3.40275	102.0825	170.1375	272.2200	340.275
Lo Barnechea	3.25268	97.5804	162.6340	260.2144	325.268
Ñuñoa	2.86180	85.8540	143.0900	228.9440	286.180

Some clean plots were obtained after arranging the dataframe:



b) FOURSQUARE API usage: In order to complement my findings, Foursquare data will be used. Specifically to visualize what was stated on the introduction section regarding Services and Connectivity within the communes.

c) Machine Learning algorithms such as k means will also be part of the data set.

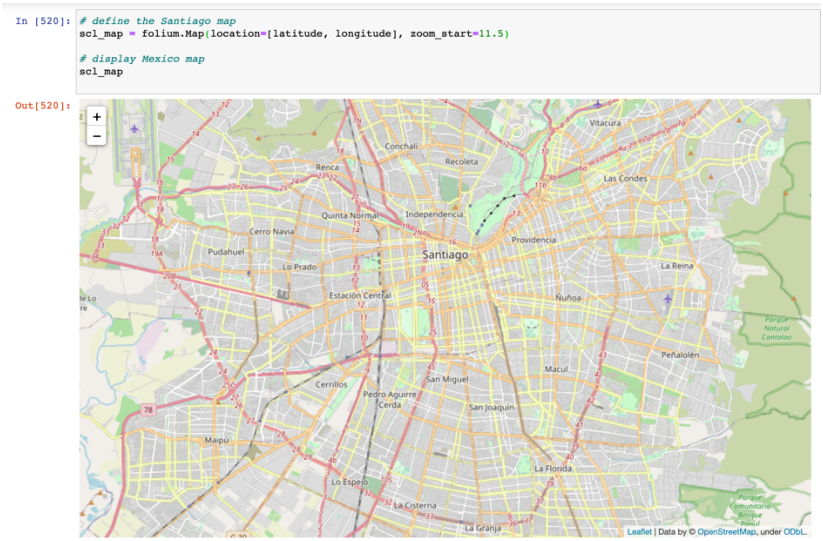


Figure 4. Santiago map obtained using Folium library.

The whole set of dataframes, plots and maps won't be added to this section, otherwise the document would become too extensive and boring to the reader.