

Puntajes PSU (<http://www.t13.cl/noticia/nacional/te-puede-servir/ranking-colegios-mejor-promedio-psu-2016>)

```
In [3]: from robobrowser import RoboBrowser
import geocoder
import folium
import pandas
import os
cwd = os.getcwd()

browser = RoboBrowser(history=True, parser='html5lib')
url = 'http://www.tl3.cl/noticia/nacional/te-puede-servir/ranking-colegios-mejor-promedio-psu-2016'
browser.open(url)
```

```
In [4]: table=browser.select('.article-content table')[0]
        t=table.__str__()
        t[:1000]
```

[illegible]

```
In [5]: t=t.replace(',','.')
df=pandas.read_html(t, header=0)[0]
df.head(10)
```

Out[5]:

	Posición en el ranking	Tipo de establecimiento	Comuna	Colegio	Alumnos que rindieron la prueba	Promedio PSU
0	1	Particular	Providencia	Cambridge College	29	704.4
1	2	Particular	Las Condes	Cordillera	63	699.0
2	3	Particular	Concón	Montemar	22	694.0
3	4	Particular	Maipú	Internacional Alba	26	693.4
4	5	Particular	Vitacura	Los Andes	69	692.9
5	6	Particular	Puerto Montt	Instituto Alemán	46	690.9
6	7	Particular	Lo Barnechea	Maimonides School	10	689.7
7	8	Particular	Chiguayante	Pinares	16	685.5
8	9	Particular	Angol	San José	19	683.7
9	10	Particular	La Reina	The Grange School	133	682.6

```
In [6]: g=geocoder.google('colegio the grange school, la reina, chile')
print(g)
print(g.latlng)
```

```
<[ERROR - No results found] Google - Geocode [empty]>
```

```
-----
ValueError                                Traceback (most recent call last)
```

```
<ipython-input-6-361371feab2d> in <module>()
```

```
1 g=geocoder.google('colegio the grange school, la reina, chile')
```

```
2 print(g)
```

```
----> 3 print(g.latlng)
```

```
~/anaconda/envs/dataviz2/lib/python3.6/site-packages/geocoder/base.py in __getattr__(self, name)
```

```
917     """
```

```
918         if not self.ok:
```

```
--> 919             raise ValueError(self.status)
```

```
920
```

```
921         if self.current_result is None:
```

```
ValueError: ERROR - No results found
```

```
In [ ]: g.latlng
```

```
-36.8883627, -73.0351594
```

```
(https://www.google.cl/maps/@-33.4382165,-70.5686165,3a,75y,3.98h,101.74t/data=!3m7!1e1!3m5!1sxP13kKF1tvm6pau-
```

```
DNnsPA!2e0!6s%2F%2Fgeo1.ggpht.com%2Fcbk%3Fpanoid%3DxP13kKF1tvm6pau-
```

```
DNnsPA%26output%3Dthumbnail%26cb_client%3Dmaps_sv.tactile.gps%26thumb%3D2%26w%3D203%26h%3D100%26yaw%3D33.01:
```

```
In [25]: from joblib import Memory
import warnings
warnings.filterwarnings('ignore')
memory = Memory(cachedir='colegios', verbose=0)
import time
from memorize.memorize import Memorize

import json

def persist_to_file(file_name):
    def decorator(original_func):

        try:
            cache = json.load(open(file_name, 'r'))
        except (IOError, ValueError):
            cache = {}

        def new_func(param):
            if param not in cache:
                cache[param] = original_func(param)
                json.dump(cache, open(file_name, 'w'))
            return cache[param]

        return new_func

    return decorator

@persist_to_file('cache.dat')
def get_latlng(query):
    providers = geocoder.osm, geocoder.google
    for provider in providers:
        #print(provider)
        try:
            g=provider(query)
        except ValueError:
            continue
        if g:
            return g.latlng
    return None, None

def get_latlng_nocache(query):
    providers = geocoder.osm, geocoder.google
```

```

for provider in providers:
    #print(provider)
    try:
        g=provider(query)
    except ValueError:
        continue
    if g:
        return g.latlng
return None, None

for index,row in df.iterrows():
    stopwords = 'colegio', 'liceo', 'school', 'instituto', 'college'
    query = row.Colegio + ", "+ row.Comuna + ", chile"
    query = query.lower()
    for i in stopwords:
        if i in query:
            break
    else:
        query = 'colegio ' + query
    print(query),
    g=get_latlng(query)
    time.sleep(0.1)
    if g ==[None, None]:
        g = get_latlng_nocache(query)
        if not g:
            print('ERROR', query)
            continue
    if g:
        lat, lng = g
        df.set_value(index,'lat', lat)
        df.set_value(index,'lng', lng)
        print(g)
    else:
        print(g, 'ERROR', query)

```

```

cambridge college, providencia, chile
[-33.4297301, -70.6171096621219]
colegio cordillera, las condes, chile
[-33.39712965, -70.5143209449632]
colegio montemar, concón, chile
[-32.95104785, -71.5354626703725]
colegio internacional alba, maipú, chile
[-33.4805743, -70.74499]

```

colegio los andes, vitacura, chile
[-33.37579495, -70.5270086976089]
instituto alemán, puerto montt, chile
[-41.4703149, -72.9327801328877]
maimonides school, lo barnechea, chile
[-33.362362, -70.485016]
colegio pinares, chiguayante, chile
[-36.9329182, -73.0240795401104]
colegio san josé, angol, chile
(None, None)
the grange school, la reina, chile
[-33.43693825, -70.5678364497786]
instituto hebreo, lo barnechea, chile
[-33.36782585, -70.50897645]
colegio trebulco, talagante, chile
[-33.68492175, -70.9181438684788]
colegio tabancura, vitacura, chile
[-33.3803347, -70.5348377]
the kent school, providencia, chile
[-33.442454, -70.606202]
colegio la girouette, las condes, chile
[-33.42014435, -70.5684437802117]
colegio arauco, arauco, chile
[-37.24882155, -73.3061872870916]
colegio los alerces, lo barnechea, chile
[-33.36445855, -70.5183101905381]
the newland school, lo barnechea, chile
[-33.3599392, -70.5327805]
colegio francisco encina, ñuñoa, chile
[-33.4637602, -70.60044585]
colegio josefino santísima trinidad, providencia, chile
[-33.43587615, -70.6188793531496]
andree english school, la reina, chile
[-33.439611, -70.551921]
colegio puerto varas, puerto varas, chile
[-41.3084803, -72.9931364]
colegio particular montessori, talca, chile
(None, None)
colegio albamar, viña del mar, chile
[-32.96045675, -71.5406954610588]
colegio nido de águilas, lo barnechea, chile
[-33.3511194, -70.5031656]
colegio itahue, chiguayante, chile

[-36.8883627, -73.0351594]
colegio cumbres, las condes, chile
[-33.3933483, -70.5021064]
the southern cross school, las condes, chile
(None, None)
trewhela's english school, providencia, chile
[-33.4226143, -70.5973636]
instituto alemán c. anwandter, valdivia, chile
[-39.8121837, -73.2529625]
saint george's college, vitacura, chile
[-33.38274525, -70.59655052429]
colegio ss.cc de apoquindo, las condes, chile
[-33.410922, -70.549691]
instituto alonso de ercilla, santiago, chile
[-33.43709455, -70.6666589741227]
liceo augusto d'halmar, ñuñoa, chile
[-33.46016265, -70.5812653049313]
colegio huelén, vitacura, chile
[-33.37760185, -70.5751837243297]
colegio del verbo divino, las condes, chile
[-33.420037, -70.58666889999999]
saint gabriel's school, providencia, chile
[-33.4271401, -70.59261055]
colegio san jorge, laja, chile
[-37.2825293, -72.7091183]
colegio villa maría academy, las condes, chile
[-33.4210241, -70.5889657966817]
colegio dunalastair, las condes, chile
[-33.37493645, -70.5206063232946]
santiago college, lo barnechea, chile
[-33.3395149, -70.5450678256533]
colegio san benito, vitacura, chile
[-33.3713969, -70.5808853281962]
colegio internacional sek chile, las condes, chile
[-33.40304, -70.564276]
colegio san ignacio el bosque, providencia, chile
[-33.43500655, -70.5942887]
colegio alianza francesa j. mermoz, curicó, chile
(None, None)
colegio amalia errázuriz, ovalle, chile
[-30.61201255, -71.2171419776463]
colegio charles de gaulle, concepción, chile
[-36.829532, -73.0464155]

colegio champagnat, villa alemana, chile
[-33.0466949, -71.3647156776244]
colegio alemán de concepción, san pedro de la paz, chile
[-36.8506206, -73.0927243]
colegio lincoln international academy, lo barnechea, chile
[-33.3695059, -70.5111826825731]
colegio tabor y nazareth, lo barnechea, chile
[-33.32443, -70.5414538]
colegio alemán de santiago, las condes, chile
[-33.39849825, -70.5698888938327]
colegio alemán sankt thomas morus, providencia, chile
[-33.42625505, -70.6107939984346]
colegio de los sagrados corazones, vitacura, chile
[-33.3784759, -70.5501737]
colegio pedro de valdivia, providencia, chile
[-33.440636, -70.6077678]
colegio akros, ñuñoa, chile
[-33.4551968, -70.5891055]
the thomas jefferson school, talcahuano, chile
[-36.784576, -73.060369]
colegio lycée antoine de saint-exupéry, vitacura, chile
[-33.3947866, -70.5820801]
colegio san francisco javier de huechuraba, huechuraba, chile
[-33.35034715, -70.6839407930914]
colegio sagrada familia, viña del mar, chile
[-32.953214, -71.5375367524631]
colegio everest, lo barnechea, chile
[-33.3335136, -70.5552317519729]
colegio serena, la serena, chile
[-29.9176152, -71.2551169932687]
colegio santa úrsula, vitacura, chile
[-33.3933445, -70.5990144520582]
colegio san isidro, buin, chile
[-33.7627216, -70.7464091]
colegio el salvador, san vicente, chile
[-34.43929895, -71.0847090631952]
instituto o'higgins, rancagua, chile
[-34.1736889, -70.7421855124169]
colegio calasanz, ñuñoa, chile
[-33.4489771, -70.5848915]
colegio pedro de valdivia agustinas, santiago, chile
[-33.4419773, -70.6645155]
colegio particular san lorenzo, copiapó, chile

[-27.38770715, -70.296433063246]
colegio sagrados corazones, hualpén, chile
[-36.8003046, -73.0859965]
colegio cruz del sur, punta arenas, chile
[-53.1554773, -70.9050057]
colegio santa cruz de chicureo, colina, chile
[-33.28656355, -70.6933838351205]
colegio alemán, temuco, chile
[-38.7370222, -72.6137053]
colegio padre hurtado y juanita de los andes, las condes, chile
[-33.4136656, -70.5124946]
colegio suizo de santiago, ñuñoa, chile
[-33.4566587, -70.6084014]
colegio pumahue, peñalolén, chile
[-33.4970005, -70.5467595888828]
colegio scuola italiana alcide de gasperi, la serena, chile
[-29.9344285, -71.2657064]
colegio pedro de valdivia - peñalolén, peñalolén, chile
[-33.49310465, -70.5765969648749]
instituto inglés rancagua, rancuagua, chile
[-34.1774722, -70.744939]
colegio wenlock school, las condes, chile
[-33.410778, -70.5220646]
colegio alemán de valparaíso, viña del mar, chile
[-33.0390072, -71.5260857]
colegio british royal school, la reina, chile
(None, None)
liceo maría luisa bombal, vitacura, chile
[-33.388298, -70.5796385]
colegio craighouse, lo barnechea, chile
[-33.3630003, -70.5101596]
colegio inglés, talca, chile
[-35.4509891, -71.5969145]
colegio constitución, constitución, chile
[-35.32940015, -72.4131335016564]
colegio gerónimo rendic, la serena, chile
[-29.90419045, -71.2437038239713]
colegio saint gaspar college, ñuñoa, chile
[-33.4599669, -70.6050505]
colegio internacional antofagasta limitada, antofagasta, chile
[-23.7282478, -70.43684449999999]
colegio francisco de asis, las condes, chile
[-33.3919505, -70.5094592]

```

green house school, temuco, chile
[-38.7411643, -72.6377808496584]
colegio pedro de valdivia las condes, las condes, chile
[-33.37071395, -70.5080432049527]
colegio inglés saint john, rancagua, chile
[-34.1754052, -70.7231044]
colegio san ignacio, calama, chile
[-22.4403566, -68.92432509999999]
colegio británico st. margaret´s s.a., concón, chile
[-32.9396604, -71.5305072]
colegio campanario, buin, chile
[-33.73853195, -70.6924615460235]
colegio san francisco de la selva, copiapó, chile
[-27.387174, -70.305078]
colegio dunalastair valle norte, colina, chile
[-33.30656115, -70.6664782184677]
colegio escuela el maitén, la unión, chile
[-40.3071504, -73.08733509999999]
sun valley college, calle larga, chile
[-32.8600102, -70.589402]

```

```
In [26]: df[df.isnull().any(axis=1)]
```

```
Out[26]:
```

	Posición en el ranking	Tipo de establecimiento	Comuna	Colegio	Alumnos que rindieron la prueba	Promedio PSU	lat	Ing
8	9	Particular	Angol	San José	19	683.7	NaN	NaN
22	23	Particular	Talca	Particular Montessori	53	669.8	NaN	NaN
27	28	Particular	Las Condes	The Southern Cross School	31	668.6	NaN	NaN
44	45	Particular	Curicó	Alianza Francesa J. Mermoz	28	660.1	NaN	NaN
81	82	Particular	La Reina	Colegio British Royal School	63	648.0	NaN	NaN

```
In [27]: df=df.dropna()
df.head(25)
```

```
Out[27]:
```

	Posición en el ranking	Tipo de establecimiento	Comuna	Colegio	Alumnos que rindieron la prueba	Promedio PSU	lat	lng
0	1	Particular	Providencia	Cambridge College	29	704.4	-33.429730	-70.617110
1	2	Particular	Las Condes	Cordillera	63	699.0	-33.397130	-70.514321
2	3	Particular	Concón	Montemar	22	694.0	-32.951048	-71.535463
3	4	Particular	Maipú	Internacional Alba	26	693.4	-33.480574	-70.744990
4	5	Particular	Vitacura	Los Andes	69	692.9	-33.375795	-70.527009
5	6	Particular	Puerto Montt	Instituto Alemán	46	690.9	-41.470315	-72.932780
6	7	Particular	Lo Barnechea	Maimonides School	10	689.7	-33.362362	-70.485016
7	8	Particular	Chiguayante	Pinares	16	685.5	-36.932918	-73.024080
9	10	Particular	La Reina	The Grange School	133	682.6	-33.436938	-70.567836
10	11	Particular	Lo Barnechea	Instituto Hebreo	76	681.0	-33.367826	-70.508976
11	12	Particular	Talagante	Trebulco	22	680.8	-33.684922	-70.918144
12	13	Particular	Vitacura	Tabancura	107	679.4	-33.380335	-70.534838
13	14	Particular	Providencia	The Kent School	61	678.5	-33.442454	-70.606202
14	15	Particular	Las Condes	La Girouette	55	678.2	-33.420144	-70.568444
15	16	Particular	Arauco	Arauco	21	676.0	-37.248822	-73.306187
16	17	Particular	Lo Barnechea	Los Alerces	61	675.6	-33.364459	-70.518310

	Posición en el ranking	Tipo de establecimiento	Comuna	Colegio	Alumnos que rindieron la prueba	Promedio PSU	lat	lng
17	18	Particular	Lo Barnechea	The Newland School	82	675.6	-33.359939	-70.532781
18	19	Particular	Ñuñoa	Francisco Encina	24	675.6	-33.463760	-70.600446
19	20	Subvencionado	Providencia	Josefino Santísima Trinidad	27	673.5	-33.435876	-70.618879
20	21	Particular	La Reina	André English School	138	672.6	-33.439611	-70.551921
21	22	Particular	Puerto Varas	Puerto Varas	46	671.8	-41.308480	-72.993136
23	24	Particular	Viña del Mar	Albamar	31	669.7	-32.960457	-71.540695
24	25	Particular	Lo Barnechea	Nido de Águilas	27	669.7	-33.351119	-70.503166
25	26	Particular	Chiguayante	Itahue	7	669.6	-36.888363	-73.035159
26	27	Particular	Las Condes	Cumbres	202	668.8	-33.393348	-70.502106

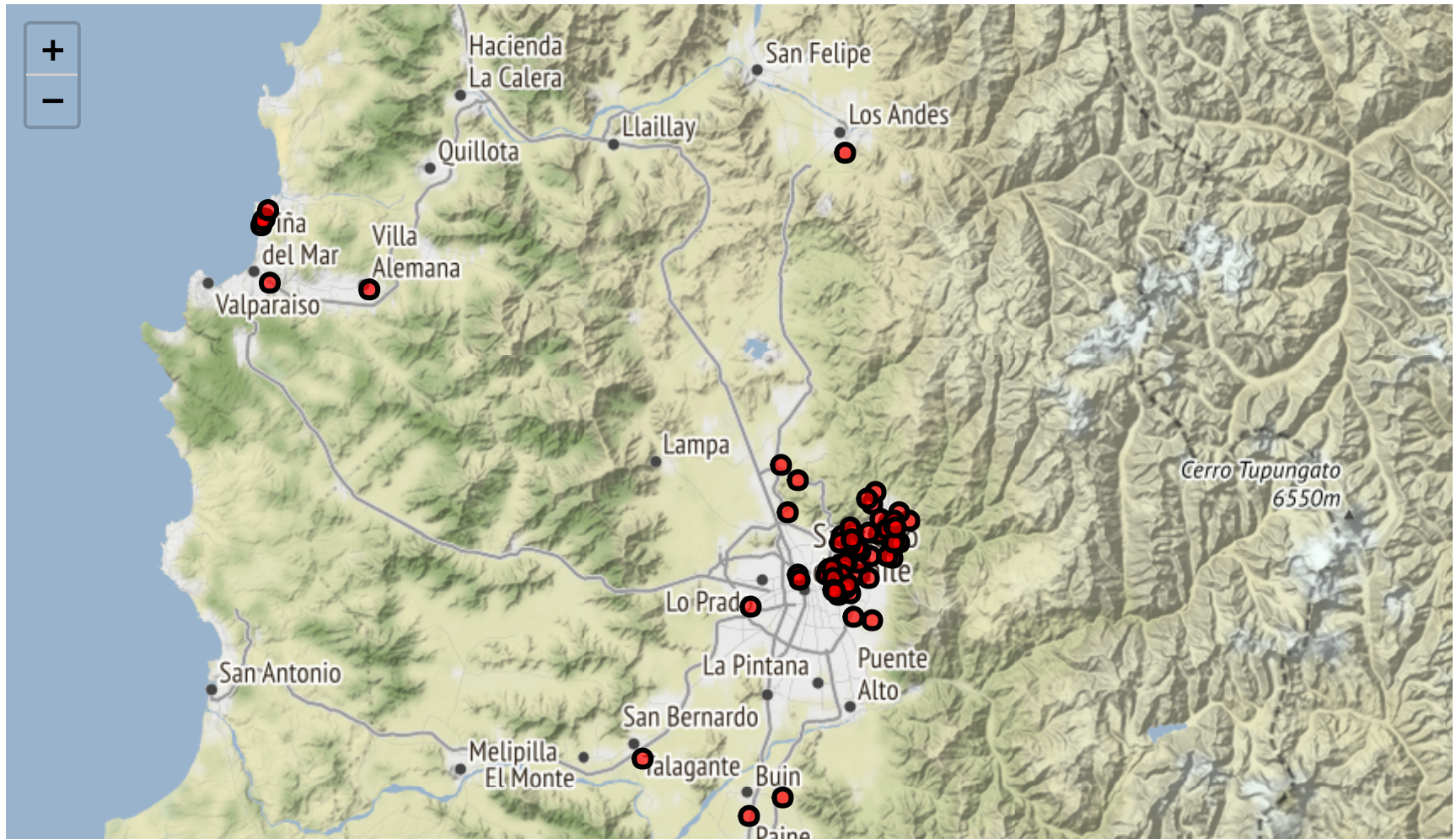
```

In [28]: map = folium.Map(location=[-33.39, -70.57], zoom_start=8, tiles='Stamen Terrain')
i=0
for index, row in df.iterrows():
    map.add_child(
        folium.CircleMarker([row.lat,row.lng],
                             color='#000000',
                             fill_color='#ff0000',
                             fill=True,
                             fill_opacity=0.7,
                             radius=5,
                             popup=u"%s %s %s" % (row.Colegio, row['Promedio PSU'], row['Comuna'])))

map.save('colegios.html')
map

```

Out[28]:



In [27]: `df.head()`

Out[27]:

	Posición en el ranking	Tipo de establecimiento	Comuna	Colegio	Alumnos que rindieron la prueba	Promedio PSU	lat	lng
0	1	Particular	Providencia	Cambridge College	29	704.4	-33.429787	-70.616853
1	2	Particular	Las Condes	Cordillera	63	699.0	-33.396461	-70.513284
2	3	Particular	Concón	Montemar	22	694.0	-32.950793	-71.535073
3	4	Particular	Maipú	Internacional Alba	26	693.4	-33.480762	-70.745021
4	5	Particular	Vitacura	Los Andes	69	692.9	-33.375544	-70.526525

```

In [44]: map = folium.Map(location=[-33.39, -70.57], zoom_start=8, tiles='Stamen Toner')
i=0
for index, row in df2.iterrows():
    map.add_child(
        folium.CircleMarker([row.lat,row.lng],
                             color='#ff0000',
                             fill_color='#ff0000',
                             fill=True,
                             fill_opacity=0.7,
                             radius=5*row['Alumnos que rindieron la prueba']/100.0,
                             popup=u"%s %s %s" % (row.Colegio, row['Promedio PSU'], row['Comuna'])))

map.save('colegios.html')
map

```

Out[44]:



```
In [28]: df.head()
```

...

```
In [13]: df3 = df[df['Tipo de establecimiento'] != 'Particular']
df3
```

```
Out[13]:
```

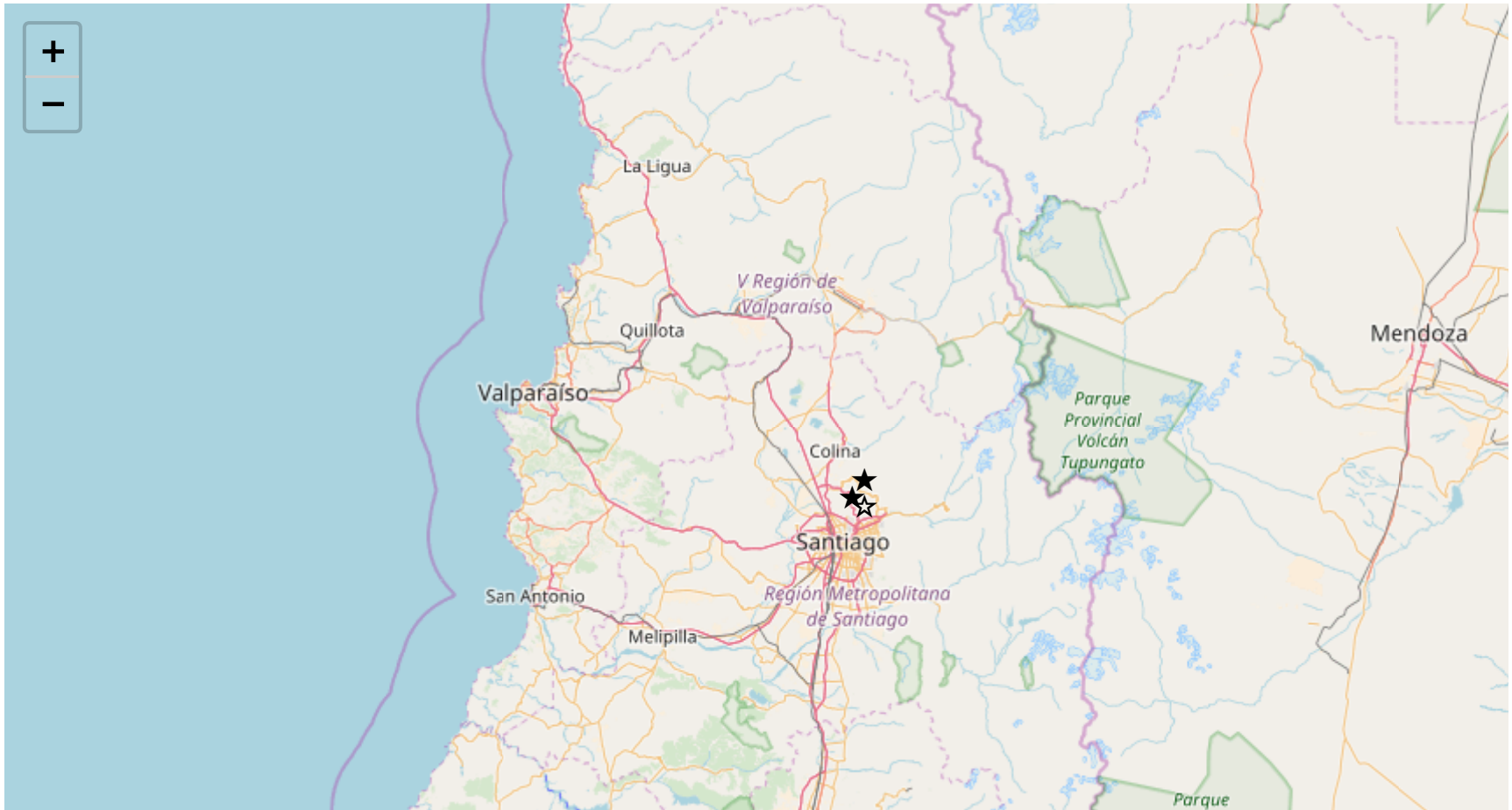
	Posición en el ranking	Tipo de establecimiento	Comuna	Colegio	Alumnos que rindieron la prueba	Promedio PSU	lat	lng
19	20	Subvencionado	Providencia	Josefino Santísima Trinidad	27	673.5	-33.435800	-70.619132
33	34	Municipal	Ñuñoa	Liceo Augusto D 'Halmar	60	665.0	-33.460248	-70.580760
64	65	Subvencionado	San Vicente	Colegio El Salvador	65	652.5	-34.439336	-71.085276
82	83	Subvencionado	Vitacura	Liceo María Luisa Bombal	35	648.0	-33.388045	-70.579640
86	87	Subvencionado	La Serena	Colegio Gerónimo Rendic	112	647.5	-29.904211	-71.243276
96	97	Subvencionado	Copiapó	Colegio San Francisco de la Selva	15	643.2	-27.387174	-70.305078
98	99	Municipal	La Unión	Escuela El Maitén	1	643.0	-40.307150	-73.087335


```

In [72]: map = folium.Map(location=[-33.39, -70.57], zoom_start=8, tiles='OpenStreetMap')
        colors = {
            'Particular': ('green', ''),
            'Subvencionado': ('red', 'star'),
            'Municipal': ('red', 'star-empty'),
        }
        for index, row in df.iterrows():
            color, icon = colors[row['Tipo de establecimiento']]
            folium.Marker(
                location=[row.lat, row.lng],
                popup=u"%s - %s - %s - %s" % (row['Tipo de establecimiento'], row.Colegio, row['Promedio PSU'],
                icon=folium.Icon(color=color, icon=icon)
            ).add_to(map)
        map.save('colegios.html')
        map

```

Out[72]:





[mapa \(colegios.html\)](#)