Applied Data Science Capstone Final Project

Sports Facilities in Thessaloniki, Greece

Charalampos Kokkalis

June 11, 2020

1 Introduction/Business Plan

For the final project of the Applied Data Science Capstone, and since we were given the choice of our topic, I decided to do something related to my hometown, Thessaloniki, Greece, and to my main hobby, basketball (and sports in general). In the recent years, I have faced the challenge of finding proper sports facilities, in order to exercise, many times. Although there has been some movement in the market, there is still a great shortage of centers for sports, given the high demand due to the city being famous for its sports teams and thus motivating all its citizens to spend time trying different sports.

2 Aim of the Analysis

The aim of the following analysis is to find suggested sub-areas in the city of Thessaloniki where it would be worth building sports facilities. The outcome of this analysis would be useful both to the Municipality, which could make public sports facilities, and to individuals looking to start their own businesses in the market.

3 Data Used

- To get access to a map of the city, I used the folium Python library.

- To search for sports facilities in the area, I used the Foursquare API.

4 Methodology

All the following steps can be found in the Jupyter Notebook (file name: Final Project.ipynb) in the repository of the project.

1. To begin the analysis in the notebook we need to import the following libraries:

```
1  import json
2  import requests
3  import pandas as pd
4  import folium
```

   json and requests are used to access the Foursquare API, pandas to store and access the data in Dataframes, and folium to display the resulting map.

2. We first make a request to the Foursquare API to get the data for Thessaloniki. The parameters show, including others, where the center of the area we are searching is located, what radius we want to consider, and what we are looking for in the area. Here, I had to use the Greek word for Sports Courts, since it gave more results due to most data being in Greek, making the analysis more meaningful and successful.
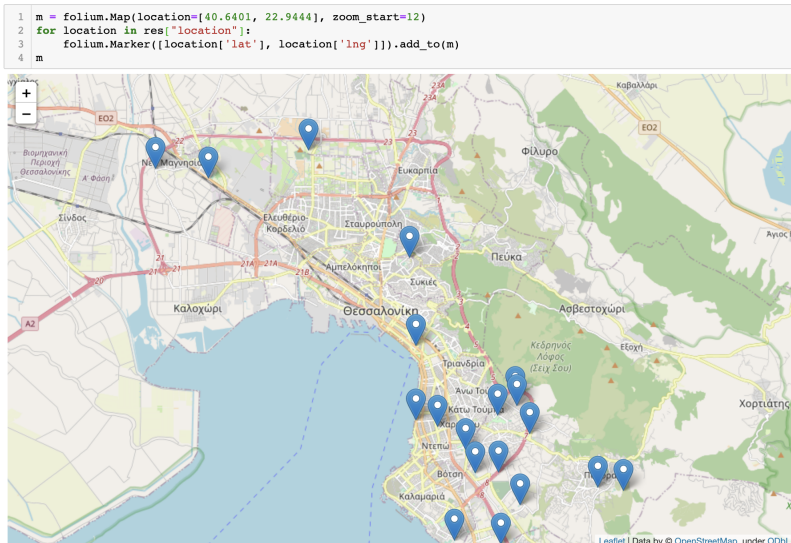
```
1  secret = input("Client secret: ")
2  url = 'https://api.foursquare.com/v2/venues/search'
3
4  params = dict(
5  client_id='DBZCP5M10O4WR1XMUZNMT3PQMLYLIFAFUXYH0RWSZXW32DD0',
6  client_secret=secret,
7  radius=10000,
8  v='20180323',
9  ll='40.6401, 22.9444', # Thessaloniki
10 query='γήπεδα', #Greek Word for Sports Courts - Gave more results
11 limit=50
12 )
13 resp = requests.get(url=url, params=params)
14 data = json.loads(resp.text)
```

3. From the data returned in the previous step, we access the list of venues. We display the head() of these to get an idea about what our data look like:

```python
res=pd.DataFrame(data['response']['venues'])
res.head()
```
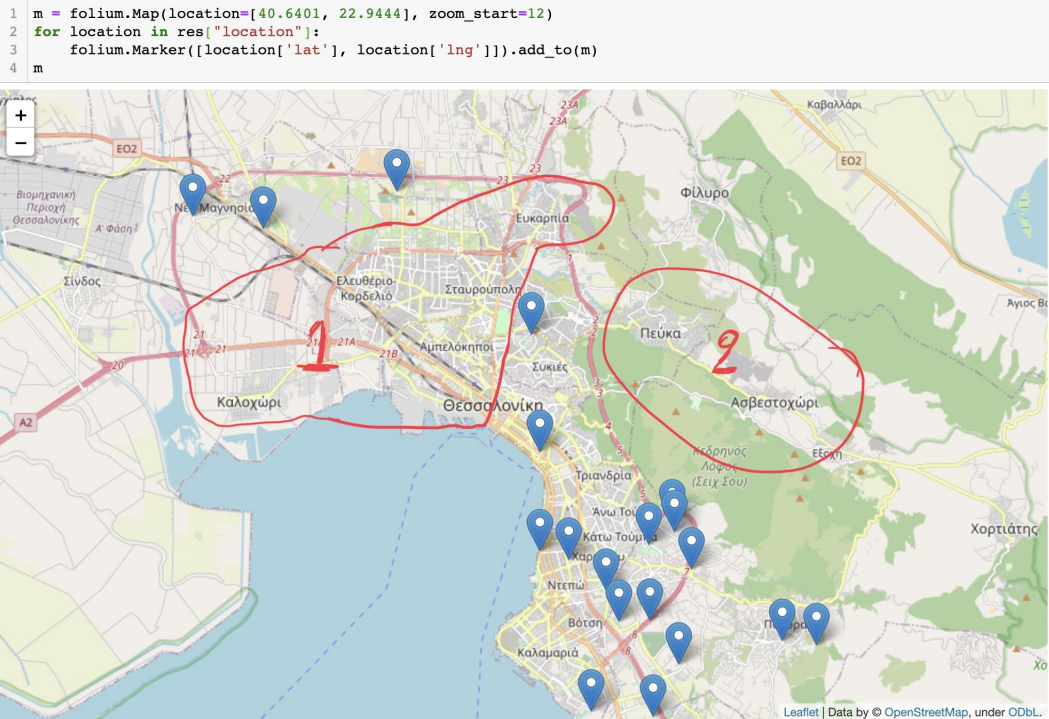
	id	name	contact	location	categories	verified	stats	beenHere	hereNow
0	4f48e083e5e81cf038a2a9ba	Γήπεδα Τέννις ΧΑΝΘ	{}	{'address': 'Πλατεία ΧΑΝΘ', 'lat': 40.62846611...	[{'id': '4e39a956bd410d7aed40cbc3', 'name': 'T...	False	{'tipCount': 0, 'usersCount': 0, 'checkinsCoun...	{'count': 0, 'lastCheckinExpiredAt': 0, 'marke...	{'count': 0, 'summary': 'Nobody here', 'groups...
1	4f9eb78be4b0ef2f7274d85f	Δημοτικά Γήπεδα Τένις	{}	{'address': 'Νέα Παραλία Θεσσαλονίκης', 'lat':...	[{'id': '4e39a956bd410d7aed40cbc3', 'name': 'T...	False	{'tipCount': 0, 'usersCount': 0, 'checkinsCoun...	{'count': 0, 'lastCheckinExpiredAt': 0, 'marke...	{'count': 0, 'summary': 'Nobody here', 'groups...
2	4e03562662e189f23d68a6fb	Γαλαξίας Δημοτικά Γήπεδα Τένις	{}	{'lat': 40.602246144095425, 'lng': 22.95951286...	[{'id': '4e39a956bd410d7aed40cbc3', 'name': 'T...	False	{'tipCount': 0, 'usersCount': 0, 'checkinsCoun...	{'count': 0, 'lastCheckinExpiredAt': 0, 'marke...	{'count': 0, 'summary': 'Nobody here', 'groups...
3	4d7e80913633a093bdbb7f9d	Γηπεδα Σαββίδης	{}	{'lat': 40.69206707742363, 'lng': 22.903922437...	[{'id': '4cce455aebf7b749d5e191f5', 'name': 'S...	False	{'tipCount': 0, 'usersCount': 0, 'checkinsCoun...	{'count': 0, 'lastCheckinExpiredAt': 0, 'marke...	{'count': 0, 'summary': 'Nobody here', 'groups...
4	5274a5a5498e3d83807b5651	Γήπεδα Σάββας Κωφίδης	{}	{'lat': 40.56386356932276, 'lng': 22.986541699...	[{'id': '4cce455aebf7b749d5e191f5', 'name': 'S...	False	{'tipCount': 0, 'usersCount': 0, 'checkinsCoun...	{'count': 0, 'lastCheckinExpiredAt': 0, 'marke...	{'count': 0, 'summary': 'Nobody here', 'groups...

4. Finally, we access (using the folium Python library) the map of Thessaloniki, and add markers at all locations where we found sports centers. The result is the following:

```python
m = folium.Map(location=[40.6401, 22.9444], zoom_start=12)
for location in res["location"]:
    folium.Marker([location['lat'], location['lng']]).add_to(m)
m
```

5   Results

The resulting map that we ended up with after the analysis of the data, makes it obvious that some parts of the city have a greater need for sports facilities than others.

```python
m = folium.Map(location=[40.6401, 22.9444], zoom_start=12)
for location in res["location"]:
    folium.Marker([location['lat'], location['lng']]).add_to(m)
m
```


As we can see, although areas 1 and 2 in the map are quite central, they have zero sports centers. Therefore, a new related business in the area would be likely to be profitable, as there will be great demand for its services. Also, the Municipality of Thessaloniki could consider building open courts, accessible to everyone.

6   Discussion

Even though the analysis came to a clear result, it could have been even more meaningful had there been enough data for the area. Unfortunately, lots of sports facilities were not identified by the Foursquare API. Also, not a single one of them had statistics about people that have visited the place, reviews, and rankings, which would have made it possible to classify "more complete" sports centers separately from simple courts.

7   Conclusion

The outcome of this analysis suggests two big areas of the city that would support investment in the field of Sports Centers. Although there were some setbacks, mostly concerning the lack of detailed and complete data, the Foursquare API was still very helpful and allowed for us to get a resulting map making it obvious to the analyst where the courts should be built.