Final Report Coursera Capstone Project – The Battle of the Neighborhoods

Athens, Greece

Athens, the capital and largest city of Greece, is home to about 5 million people, almost half of the total population, living either in the Centre or in the suburbs.

1. Introduction

1.1. Background

Athens is a highly industrialized city which also provides lots of fun activities and options for all kinds of taste. As such there are numerous places for people to eat, drink, relax, shop etc. The largest number of tourists visiting Greece will certainly pay a visit to the capital of western civilization. One way or another they will land on Eleftherios Vezinelos International Airport and most likely will have to go to the port of Piraeus would they want to go to the islands.

Being not only a busy but also 'crossroad' city, Athens hosts a great number of restaurants in many different areas, where local people and tourists feel welcomed. It would be wise to list and visualize areas for possible investors and/or businessmen in case they would like to start an enterprise such as a restaurant or a café.

As the Greek economy is strongly dependent on tourism, there are a lot of places for someone to enjoy themselves. Fundamentally Athens could not be otherwise. There is a plethora of night clubs, bars, cafes, beach bars, restaurants (Greek and foreign-themed), and many other places. Consequently, would an investor be interested into trying something new in this city, how could we advise him to prepare and focus his business plan?

1.2. Problem

Data that might contribute to determining district availability for starting a new brand might include area population, similar venues, dissimilar venues and in general other coexisting venues. This project aims to suggest whether and or not an area is suitable for that.

1.3. Interest/Target Audience

Consequently, investors, businessmen but also many tourists might be interested into finding out which districts of Athens offer what kind of venue.

2. Data acquisition and cleaning

Data sources
For this project we will need the following data

2.1. Athens data that contain the municipalities in Athens Prefecture taken from the Wikipedia. Data source: https://en.wikipedia.org/wiki/Athens_Prefecture

2.2. Data specifically for the municipality of Athens, since it is the most important region of Athens Prefecture, that we will use to analyze the neighborhoods of Athens Centre. Data source: https://en.wikipedia.org/wiki/Category:Neighbourhoods_in_Athens

2.3. All kind of venues in Athens: Data source: Foursquare APIs

3. Methodology

Data Cleaning - Exploratory Data

The first thing needs to be done is to scrape the municipalities of Athens Prefecture as illustrated into the link provided above and to scrape the neighborhoods of Central Athens, since it is the most important area of the whole city. Scraping the municipalities gave me a nice data frame with the names, seats, population, area, density and YPES code of each municipality. The next step following scraping is adding the coordinates using the geopy library with geolocator.

After some cleaning and manipulating I had a beautiful data frame with the coordinates as well.

Scraping for the neighborhoods I only got the names – the other information is not provided in the web. Again, after some cleaning and manipulating I got a nicely executed data frame.

Below we can get a glimpse the final form of both data frames:

	Municipality	Seat	Population	Area (km2)	Density (/km2)	Latitude	Longitude
0	Agia Varvara (Αγία Βαρβάρα)	Agia Varvara (Αγία Βαρβάρα)	30562	2.425	12603	37.995370	23.661860
1	Agia Paraskevi (Αγία Παρασκευή)	Agia Paraskevi (Αγία Παρασκευή)	59500	7.967	7498	38.013560	23.821750
2	Agioi Anargyroi (Άγιοι Ανάργυροι)	Agioi Anargyroi (Άγιοι Ανάργυροι)	32957	3.200	10299	38.026480	23.717880
3	Agios Dimitrios (Άγιος Δημήτριος)	Brahami (Μπραχάμι)	65173	4.949	13169	37.939780	23.731470
4	Aigaleo (Αιγάλεω)	Aigaleo (Αιγάλεω)	74046	6.450	11480	37.992740	23.680430
5	Alimos (Άλιμος)	Kalamaki (Καλαμάκι)	38047	5.909	6439	37.914670	23.705440
6	Marousi (Μαρούσι)	Marousi (Μαρούσι)	69470	13.093	5369	38.048260	23.797180
7	Argyroupoli (Αργυρούπολη)	Argyroupoli (Αργυρούπολη)	33158	8.228	4030	37.915570	23.750080
8	Athens (Αθήνα)	Athens (Αθήνα)	745514	38.964	19133	37.976830	23.709890
9	Chaidari (Χαϊδάρι)	Chaidari (Χαϊδάρι)	46276	22.655	2043	38.003030	23.663760
10	Chalandri (Χαλάνδρι)	Chalandri (Χαλάνδρι)	71684	9.629	6634	38.022420	23.794140
11	Cholargos (Χολαργός)	Cholargos (Χολαργός)	32166	3.950	8143	38.003610	23.798460
12	Dafni (Δάφνη)	Dafni (Δάφνη)	23674	1.375	17155	37.947516	23.772295
13	Ekali (Εκάλη)	Ekali (Εκάλη)	5190	4.332	1198	38.088017	23.814780
14	Elliniko (Ελληνικό)	Elliniko (Ελληνικό)	16740	7.127	2349	37.881530	23.728390
15	Filothei (Φιλοθέη)	Filothei (Φιλοθέη)	7310	2.301	3177	38.042556	23.771567
16	Galatsi (Γαλάτσι)	Galatsi (Γαλάτσι)	58042	4.026	14417	38.013060	23.752650
17	Glyfada (Γλυφάδα)	Glyfada (Γλυφάδα)	80409	25.366	3170	37.869650	23.738910
18	Ilion (Ίλιον)	Nea Liosia (Νέα Λιόσια)	80859	9.250	8554	38.030510	23.712640
19	Ilioupoli (Ηλιούπολη)	Ilioupoli (Ηλιούπολη)	75904	12.724	5965	37.936720	23.742980
20	Irakleio (Ηράκλειο)	Neo Irakleio (Νέο Ηράκλειο)	45926	4.652	9902	35.334280	25.123050
21	Kaisariani (Καισαριανή)	Kaisariani (Καισαριανή)	26419	7.841	3369	37.969090	23.754840
22	Kallithea (Καλλιθέα)	Kallithea (Καλλιθέα)	109609	4.749	23080	37.955010	23.711620
23	Kamatero (Καματερό)	Kamatero (Καματερό)	22234	5.950	3737	38.060375	23.707807
24	Kifisia (Κηφισιά)	Kifisia (Κηφισιά)	43929	25.937	1694	38.071719	23.811310
25	Lykovrysi (Λυκόβρυση)	Lykovrysi (Λυκόβρυση)	8116	1.950	4162	38.069440	23.781940
26	Melissia (Μελίσσια)	Melissia (Μελίσσια)	19526	3.906	4936	38.051140	23.836120
27	Metamorfosi (Μεταμόρφωση)	Koukouvaounes (Κουκουβάουνες)	26448	5.502	4807	38.059890	23.757197
28	Moschato (Μοσχάτο)	Moschato (Μοσχάτο)	23153	2.325	9958	37.956283	23.680874
29	Nea Chalkidona (Νέα Χαλκηδόνα)	Nea Chalkidona (Νέα Χαλκηδόνα)	10112	0.800	12640	38.025840	23.728470

|	Neighborhood	Latitude	Longitude
0	Aerides	37.976140	23.736400
1	Agios Eleftherios	38.019970	23.726270
2	Agios Panteleimonas	37.962660	23.625720
3	Akadimia Platonos	37.990950	23.707180
4	Akadimia	37.986750	23.711030
5	Ampelokipoi	37.991060	23.764290
6	Anafiotika	37.972240	23.728520
7	Ano Petralona	37.969480	23.708970
8	Asteroskopeio	37.974380	23.719200
9	Asyrmatos	38.006780	23.881790
10	Attiki	37.976140	23.736400
11	Elaionas	37.966670	23.669400
12	Ellinoroson	37.997930	23.775160
13	Erythros Stavros	37.992680	23.767630
14	Evangelismos	37.977300	23.747260
15	Exarcheia	37.988170	23.733610
16	Gazi	37.974620	23.764545
17	Girokomeio	40.290470	21.775090
18	Goudi	37.982920	23.770980
19	Gouva	37.959980	23.740830
20	Gyzi	38.020475	23.845253
21	Ilisia	37.976410	23.755570
22	Kallimarmaro	37.966910	23.740220
23	Kallimarmaro	37.966910	23.740220
24	Kerameikos	37.979190	23.719200
25	Kolokynthou	37.996760	23.706090
26	Kolonaki	37.977620	23.743590
27	Kolonos	37.993340	23.716370
28	Koukaki	37.963720	23.723530
29	Kountouriotika	37.988560	23.735324
30	Kynosargous	38.023995	23.673820

For the analysis I used FourSquare to get data of venues in it, as well to find count of restaurants specifically. Python Bar Plot was used to visualize the count of restaurants in each neighborhood, because it's easy to compare watching it.
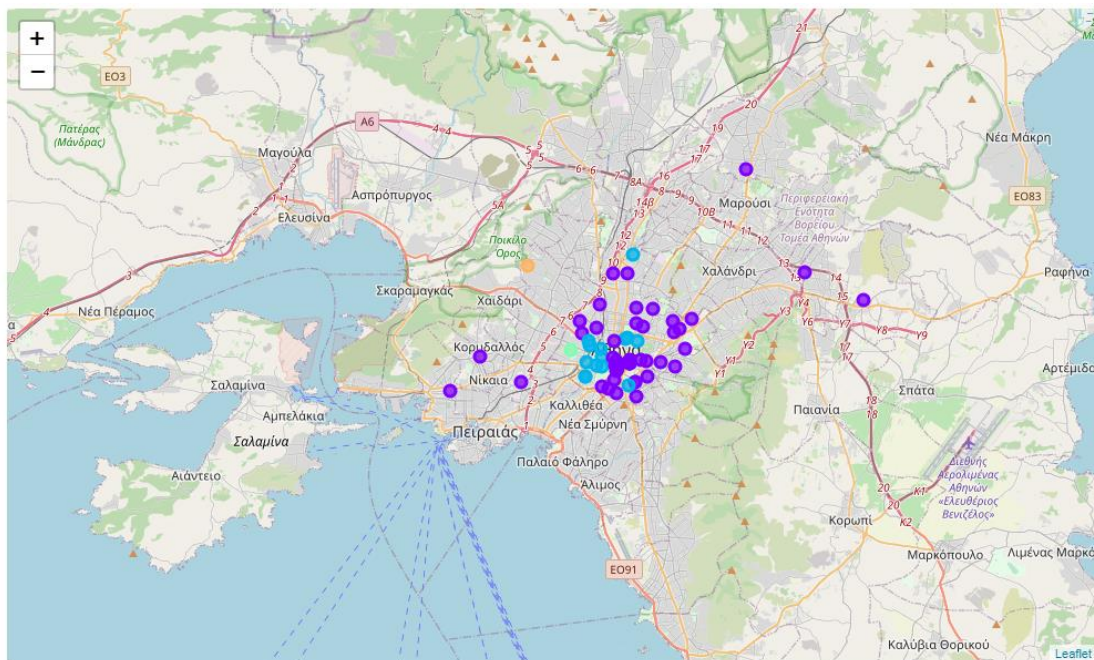
4. Results

Clustering the dataset I obtained and manipulating them into my area of interest, and using the Folium map, I obtained the following maps about: venues in Athens prefecture, restaurants in Athens prefecture, venues in neighborhoods of Athens Centre, restaurants in neighborhoods of Athens Centre.
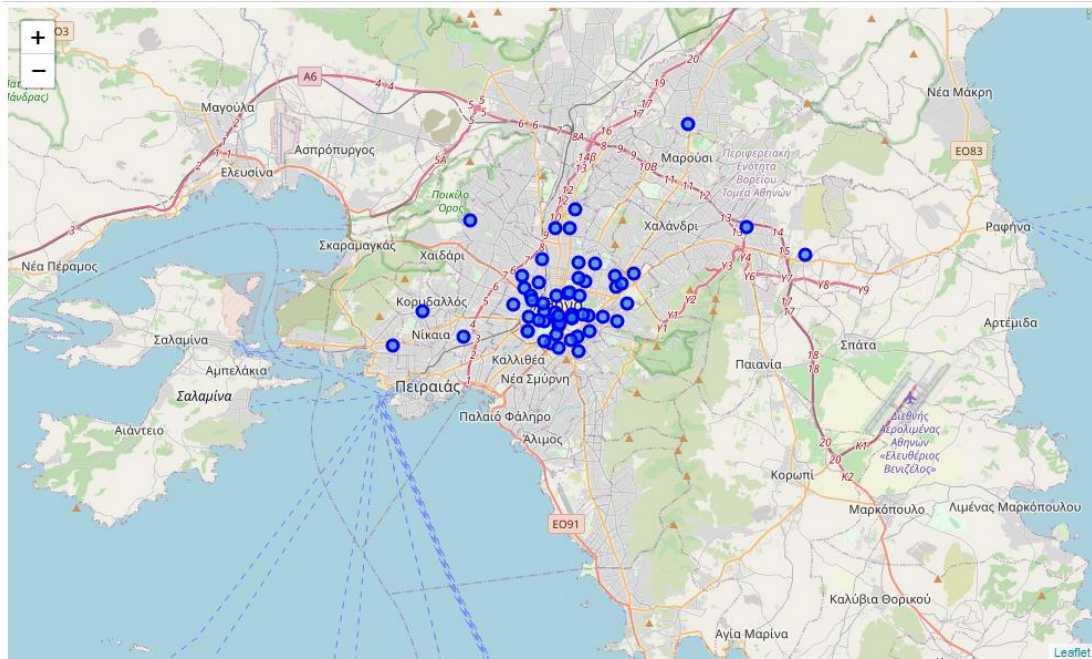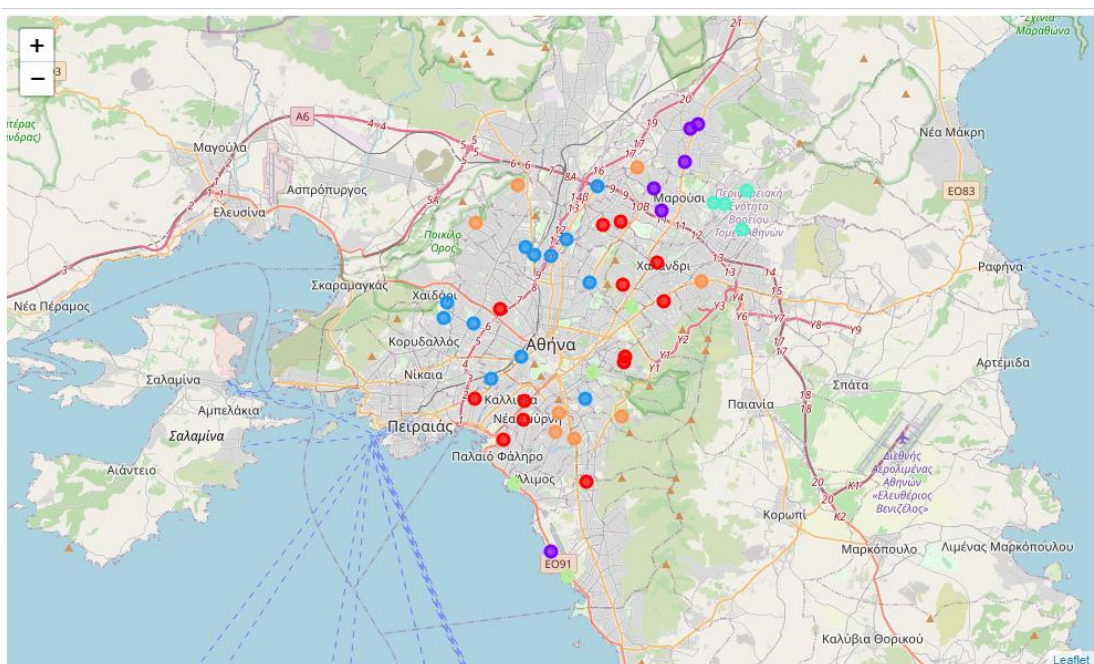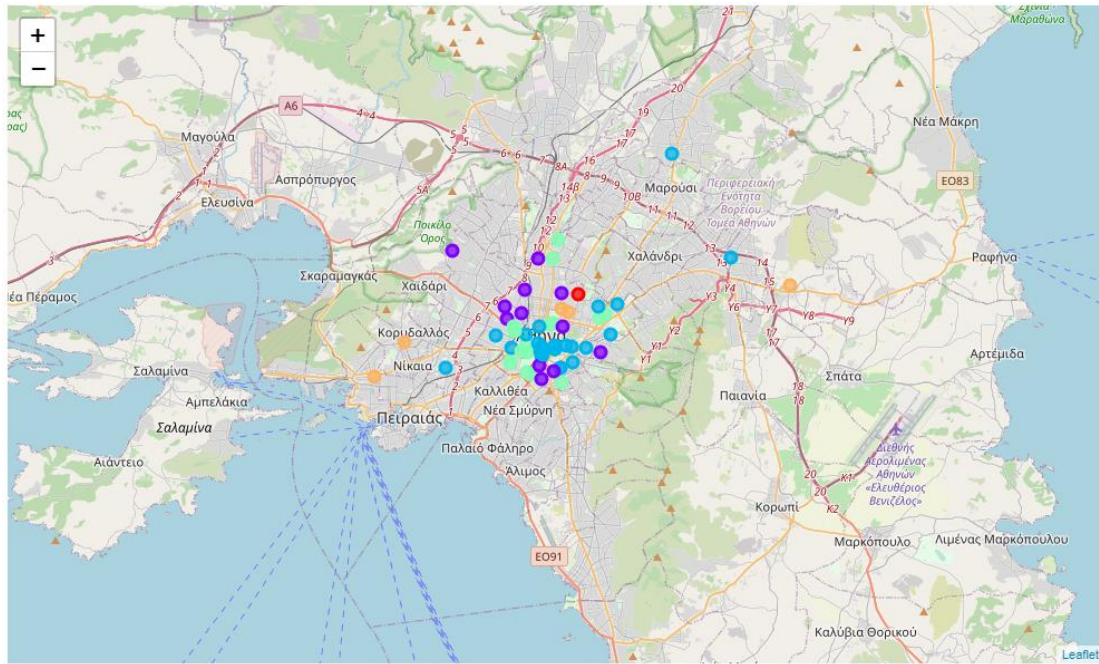

Map of Athens
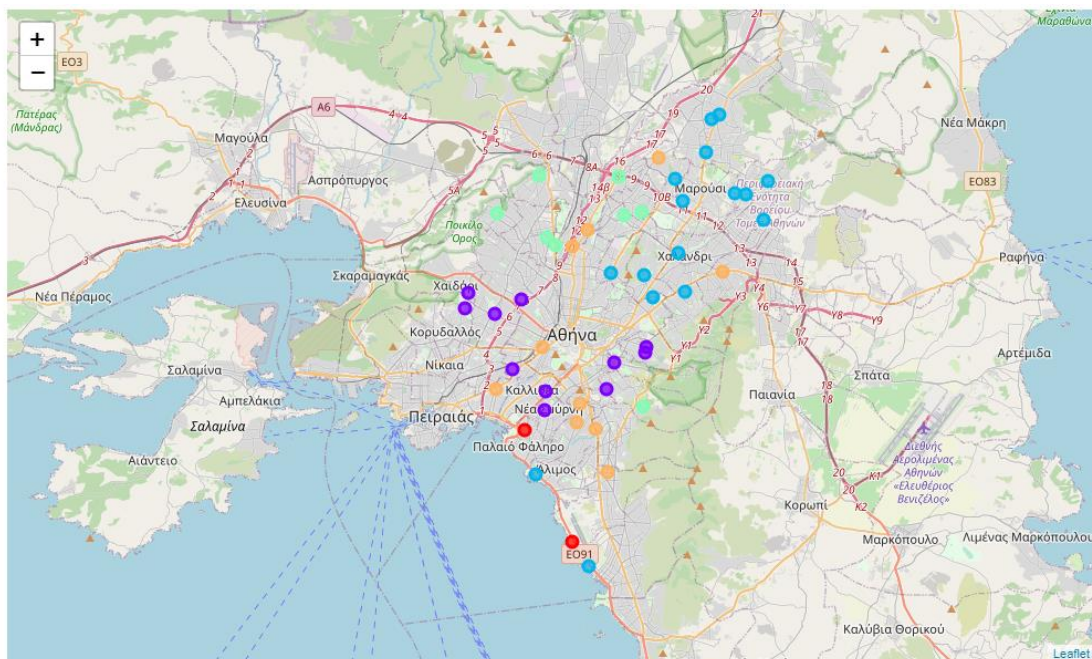
Map of Athens Prefecture

Map of Athens Centre

Restaurants in Athens Prefecture

Restaurants in Athens Centre

5. Discussion

- We got a glimpse of the venues and restaurants in Athens and were able to find out some interesting insights which might be useful to travelers as well as people with business interests. Let's summarize our findings:
- Greeks like to drink a lot of coffee as there are too many cafes and coffee shops.
- Greek restaurants top every other kind of restaurant.
- In the center of the city where more than half a million people live and most tourists stay, there is a differential on the kind of restaurants there is. This shows up as being a multicultural area.

The clustering is based on the most common venues obtained from Foursquare data and the population density of each area.
However, in my analysis, I have ignored other factors like distance of the venues from closest stations, range of prices of restaurants, Michelin Restaurants and so on, since I do not have such data and it would be difficult to farm it for a small exploratory study like this. Hence, this analysis only helps travelers and investors to get an overview of venue distribution.
Furthermore, this results also could potentially vary if we use some other clustering techniques like DBSCAN.

6. Conclusion

In a fast-moving world, there are many real-life problems or scenarios where data can be used to find solutions to those problems. Like seen above, data was used to cluster neighborhoods in Athens based on the venues and restaurants in over 40 municipalities and more than 60 neighborhoods. The results can help an investor or a traveler to decide about the district that fit the most his needs.

I have made use of some frequently used python libraries to scrap web-data, use Foursquare API to explore the major districts of Tokyo and saw the results of segmentation of districts using Folium leaflet map. Similarly, data can also be used to solve other problems, which most people face in metropolitan cities. Potential for this kind of analysis in a real-life problem is discussed in great detail. Also, some of the drawbacks and chance for improvements to represent even more realistic pictures are mentioned.