

Coursera Applied Data Science Capstone
The battle of neighborhoods (week 2)

Subject: Find a suitable location to establish a burger restaurant in Athens metropolitan area, Greece

Stefanos Bontas, May 2020

1. Introduction

The concept:

Athens metropolitan area is famous for the variety of food and it is an expandable industry.

The objective:

In this project, our objective is to find the most suitable locations for investing in order to establish a new burger restaurant.

The selecting criteria:

We have to define the specifications for a location to be suitable:

- Be close to a metro railway station, in order to be easily accessible by the mass since Athens has a huge problem concerning the car parking
- Total number of similar restaurants with a radius of 1,000 meters around each metro station
- The nearest burger restaurant to the metro line station should also have some impact on the results

Tools that should be used:

We'll use the following tools:

- Watson studio Jupyter notebook with Python 3
- Foursquare developer API
- Install python libraries (conda and geopy)
- Import python libraries:
 - Pandas, for dataset modification
 - Numpy for handling numerical operations
 - Matplotlib, for generating meaningful charts
 - Folium, for map visualizations
- K-means machine learning unsupervised learning algorithm

2. Data

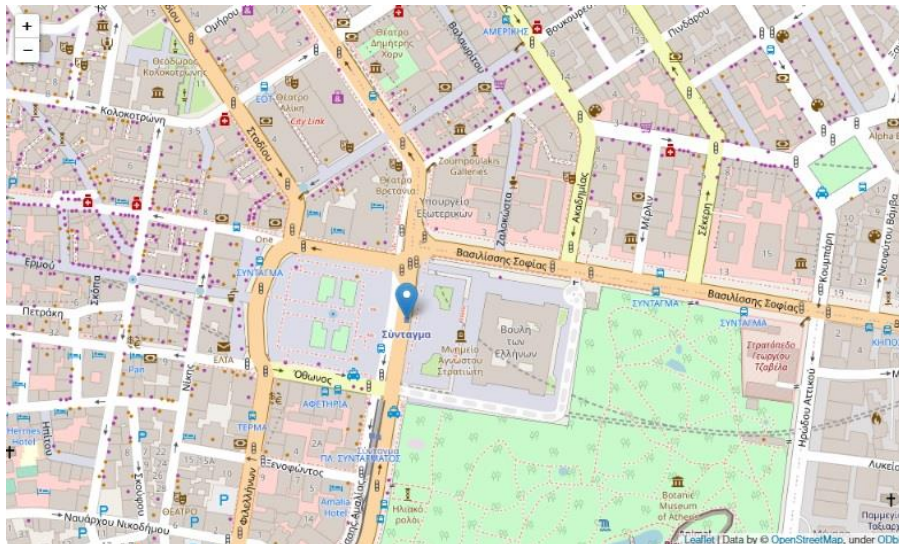
In order to meet the above requirements, we do need to download the following datasets from foursquare developer API:

- Metro stations in Athens and its suburbs. This is the venue category [4bf58dd8d48988d1fd931735](#) in foursquare.
- The burger restaurants around them (1,000 meters radius distance should be enough). This is the venue category [4bf58dd8d48988d1fd931735](#) in foursquare.

For obtaining the correct data set, we should consider that a distance of 20 kilometers radius around the center of Athens (Syntagma square, a point which is being used as a benchmark for all the distances count from/to the city of Athens).

3. Methodology

For the purpose of our study, we have to define a point zero. This will be Syntagma square, the most popular spot in Athens and a transportation hub as well.



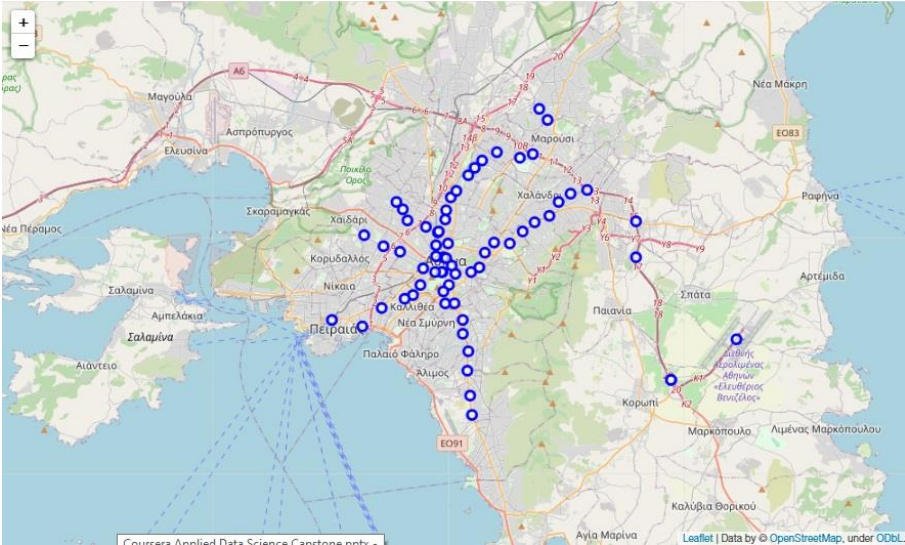
Then we'll use its coordinates to find all of the metro line stations in a radius of 20,000 meters. Below, you may see a snapshot (first 5 rows) from our dataset:

	name	lat	lng	distance	postalCode	venue_type
0	Syntagma Metro Station (Σταθμός Μετρό Σύνταγμα)	37.975235	23.735298	49	105 63	Metro Station
1	Akropoli Metro Station (Σταθμός Μετρό Ακρόπολη)	37.968516	23.730195	917	117 42	Metro Station
2	Evangelismos Metro Station (Σταθμός Μετρό Ευαγγ...	37.976173	23.747163	1004	106 76	Metro Station
3	Monastiraki Metro Station (Σταθμός Μετρό Μονασ...	37.976067	23.725752	878	105 54	Metro Station
4	Panormou Metro Station (Σταθμός Μετρό Πανόρμου)	37.993298	23.764135	3182	115 23	Metro Station

At this point, we have to remove 3 rows which represent metro line parking facilities and no actual stations, since the crowd doesn't have access there.

	name	lat	lng	distance	postalCode	venue_type
13	Θησείο	37.977394	23.720287	1372	NaN	Metro Station
46	Αμαξοστάσιο Μετρό Ελαιώνα	37.986554	23.686795	4467	NaN	Metro Station
63	Αμαξοστάσιο Μετρό Δουκίσσης Πλακεντίας	38.022099	23.835587	10179	NaN	Metro Station

Below you can see a visualization of the metro line stations from the updated dataset:



Then via foursquare API we'll get the fast food restaurants that are 1,000 meters radius from the above metro line stations. The generated list is the below one grouped by the Venue Category:

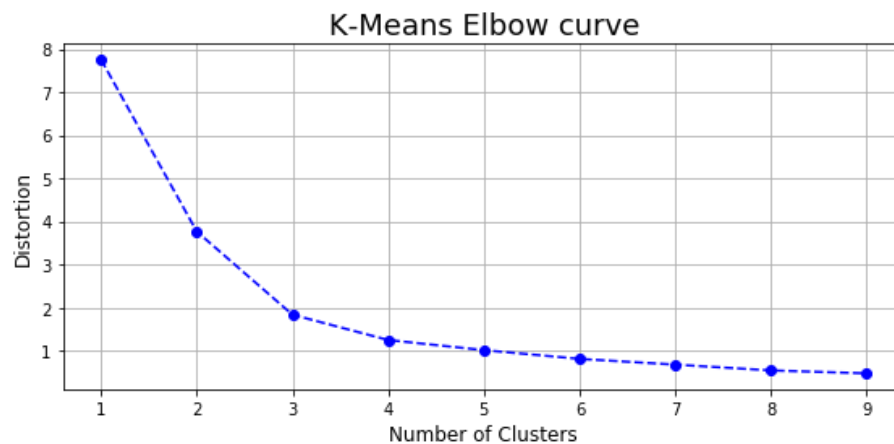
	station	lat	lng	Venue	Venue Latitude	Venue Longitude	Distance from Station
Venue Category							
American Restaurant	5	5	5	5	5	5	5
Bakery	8	8	8	8	8	8	8
Burger Joint	22	22	22	22	22	22	22
Café	14	14	14	14	14	14	14
Chinese Restaurant	4	4	4	4	4	4	4
Cocktail Bar	7	7	7	7	7	7	7
Comfort Food Restaurant	8	8	8	8	8	8	8
Creperie	3	3	3	3	3	3	3
Fast Food Restaurant	286	286	286	286	286	286	286
Frozen Yogurt Shop	8	8	8	8	8	8	8
Greek Restaurant	16	16	16	16	16	16	16
Hot Dog Joint	2	2	2	2	2	2	2
Mexican Restaurant	9	9	9	9	9	9	9
Pizza Place	4	4	4	4	4	4	4
Sandwich Place	6	6	6	6	6	6	6
Snack Place	10	10	10	10	10	10	10
Wine Bar	5	5	5	5	5	5	5

Then we have to **narrow down** our selection so as the type of restaurant to correspond to a Burger house. The selected venue categories should be **American Restaurant, Burger Joint or Fast Food Restaurant**.

We have also discarded the below metro line stations since there is no burger restaurant close to them but we highly recommend to be investigated further since the foursquare data may be obsolete.

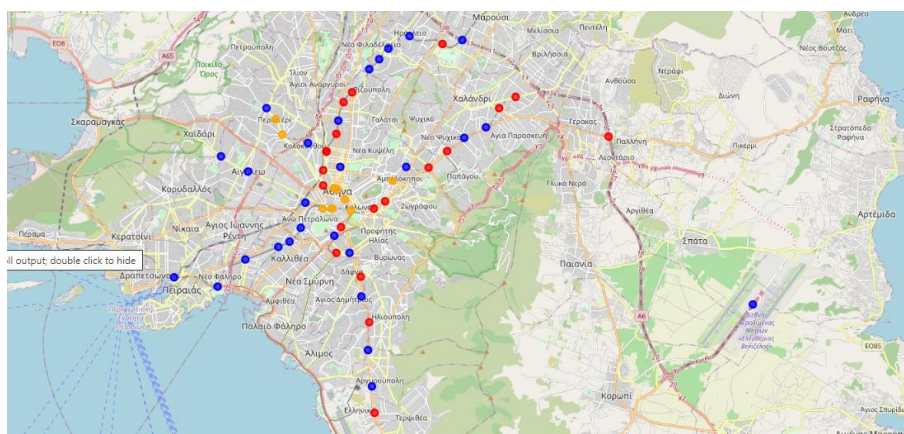
	station	lat_x	lng_x	lat_y	lng_y	Min Distance from Station	Restaurants Count
29	Eleonas Metro Station	37.987818	23.694272	NaN	NaN	NaN	NaN
56	Doukissis Plakentias Metro Station	38.024105	23.833275	NaN	NaN	NaN	NaN
59	Paiania-Kantza Metro Station	37.984707	23.870084	NaN	NaN	NaN	NaN
60	Kifisia ISAP Station	38.071627	23.797488	NaN	NaN	NaN	NaN
61	Koropi Metro Station	37.912707	23.895973	NaN	NaN	NaN	NaN
62	Koropi Suburban Rail Station	37.912917	23.895830	NaN	NaN	NaN	NaN

After data normalization, we ended up with selecting 3 clusters for the K-means clustering algorithm.

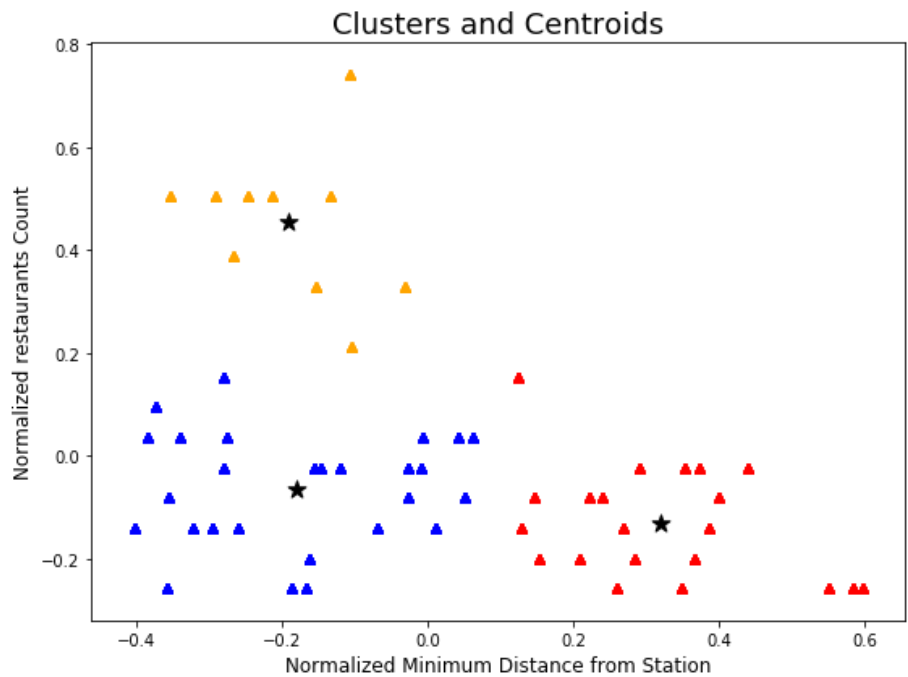


4. Results

After executing the K-means, we ended up with clustering the metro line stations with 3 colors, 1 for each cluster. You may find below the updated map:



Below you may also found a scatter plot showing the distribution of the clusters created:



Cluster label 2 – High potential (marked with blue color in the map):

In this cluster, the number of restaurants is less than the average and the minimum distance from station higher than the respective average as well.

	station	Cluster Label	lat	lng	Min Distance from Station	Restaurants Count	Norm Min Distance from Station	Norm Restaurants Count
37	Panormiou Metro Station	2	37.993298	23.764135	121	8	-0.279076	0.153144
59	Sygrou-Fix Metro Station	2	37.964919	23.726618	44	1	-0.358376	-0.258621
60	Aghios Dimitrios Metro Station	2	37.940089	23.740915	106	3	-0.294524	-0.140974
109	Kerameikos Metro Station	2	37.978550	23.711564	30	7	-0.372794	0.094320
144	Egaleo Metro Station	2	37.991428	23.681813	326	3	-0.067953	-0.140974
147	Aghios Ioannis Metro Station	2	37.958003	23.734743	230	1	-0.166821	-0.258621
148	Victoria ISAP Station	2	37.993158	23.729811	18	6	-0.385152	0.035497
169	Petralona ISAP Station	2	37.968341	23.709012	250	5	-0.146223	-0.023327
182	Tavros ISAP Station	2	37.962494	23.703367	385	6	-0.007191	0.035497
192	Sepolia Metro Station	2	38.002892	23.713120	433	6	0.042242	0.035497
198	Kallithea ISAP Station	2	37.960339	23.697373	366	4	-0.026759	-0.082150
207	Kato Patisia ISAP Station	2	38.012168	23.728596	1	3	-0.402660	-0.140974
226	Moschato ISAP Station	2	37.955221	23.680209	453	6	0.062840	0.035497
245	Anthoupoli Metro Station	2	38.017143	23.691265	63	6	-0.338808	0.035497
251	Holargos Metro Station	2	38.005225	23.794296	403	3	0.011346	-0.140974
254	Alimos Metro Station	2	37.918175	23.744169	121	5	-0.279076	-0.023327
259	Petrossos ISAP Station	2	38.033124	23.744857	124	6	-0.275986	0.035497
265	Agia Marina Metro Station	2	37.997628	23.667484	210	1	-0.187418	-0.258621
266	Pefkakia ISAP Station	2	38.037224	23.749974	367	5	-0.025729	-0.023327
271	Neo Faltro ISAP Station	2	37.944256	23.666055	441	4	0.050481	-0.082150
275	Nomismatokopio Metro Station	2	38.009402	23.805635	140	3	-0.259509	-0.140974
278	Nea Ionia ISAP Station	2	38.041531	23.755007	47	4	-0.355286	-0.082150
282	Piraeus ISAP Station	2	37.948087	23.643211	275	5	-0.120477	-0.023327
287	Argyroupoli Metro Station	2	37.903425	23.746266	383	5	-0.009251	-0.023327
296	Irakleio ISAP Station	2	38.046674	23.765953	79	3	-0.322330	-0.140974
303	Neratziotissa ISAP Station	2	38.045224	23.793136	242	5	-0.154462	-0.023327
311	Airport Metro Station	2	37.936777	23.944811	235	2	-0.161671	-0.199797

Cluster label 0 – average potential (marked with orange color in the map):

There are already a number of similar restaurants in the neighborhood but the respective investment is not prohibited to open a new one since the average distance from metro line station is high.

	station	Cluster Label	lat	lng	Min Distance from Station	Restaurants Count	Norm Min Distance from Station	Norm Restaurants Count	
0	Syntagma Metro Station	0	37.975235	23.735298		133	12	-0.266718	0.388438
23	Monastiraki Metro Station	0	37.976067	23.725752		109	14	-0.291434	0.506085
45	Omonia Metro Station	0	37.984602	23.726186		186	14	-0.212135	0.506085
63	Panepistimio Metro Station	0	37.980004	23.732354		288	18	-0.107088	0.741379
81	Monastiraki ISAP Station	0	37.976044	23.725204		152	14	-0.247150	0.506085
95	Omonoia ISAP Station	0	37.984100	23.728071		50	14	-0.352196	0.506085
116	Thisseio ISAP Station	0	37.976166	23.720449		263	14	-0.132835	0.506085
158	Ampelokipi Metro Station	0	37.987439	23.757075		243	11	-0.153432	0.329615
210	Aghios Antonios Metro Station	0	38.006380	23.699421		290	9	-0.105029	0.211968
233	Peristeri Metro Station	0	38.012749	23.696161		361	11	-0.031908	0.329615

Cluster label 1 – low potential (marked with red color in the map):

There are already a number of similar restaurants in the neighborhood as well as close to the metro line station.

	station	Cluster Label	lat	lng	Min Distance from Station	Restaurants Count	Norm Min Distance from Station	Norm Restaurants Count	
12	Akropoli Metro Station	1	37.968516	23.730195		513	8	0.124632	0.153144
20	Evangelismos Metro Station	1	37.976173	23.747163		518	3	0.129781	-0.140974
130	Attiki Metro Station	1	37.999452	23.722549		754	5	0.372829	-0.023327
135	Megaro Moussikis Metro Station	1	37.979014	23.753173		534	4	0.146259	-0.082150
139	Metaxourghio Metro Station	1	37.985549	23.720734		674	5	0.290440	-0.023327
154	Neos Kosmos Metro Station	1	37.957934	23.727630		749	2	0.367680	-0.199797
156	Larisis Metro Station	1	37.991851	23.720942		541	2	0.153468	-0.199797
174	Attiki ISAP Station	1	37.999735	23.722709		736	5	0.354292	-0.023327
179	Dafni Metro Station	1	37.948246	23.740441		768	3	0.387247	-0.140974
188	Agios Nikolaos ISAP Station	1	38.006748	23.727649		607	4	0.221439	-0.082150
202	Katehaki Metro Station	1	37.992993	23.775953		819	5	0.439771	-0.023327
219	Agios Eleftherios ISAP Station	1	38.019715	23.731599		653	3	0.268813	-0.140974
222	Ethniki Amyna Metro Station	1	37.999722	23.785535		594	2	0.208051	-0.199797
224	Ilioupoli Metro Station	1	37.929474	23.744744		669	2	0.285291	-0.199797
232	Ano Patisia ISAP Station	1	38.023857	23.735795		730	1	0.348113	-0.258621
244	Halandri Metro Station	1	38.021967	23.820975		960	1	0.584982	-0.258621
292	Agia Paraskevi Metro Station	1	38.017153	23.812426		780	4	0.399606	-0.082150
299	Eirini ISAP Station	1	38.043475	23.783093		624	4	0.238947	-0.082150
308	Elliniko Metro Station	1	37.892466	23.747613		645	1	0.260574	-0.258621
309	Pallini Metro Station	1	38.006599	23.869600		927	1	0.550996	-0.258621
310	KAT ISAP Station	1	38.065601	23.803997		972	1	0.597340	-0.258621

5. Discussion (observations & recommendations)

In order the analysis to be valid, the selected location should be also checked for potential venues not listed in foursquare.

Since Athens metropolitan railway is expanding, new neighborhoods are being added as business opportunities so this analysis should be re-executed.

Some stations appear to have no similar restaurants nearby, which is quite worrying and further analysis – investigation is being needed before investing.

Other socioeconomic factors should be taken into consideration such as:

- The actual traffic of each location since some stations are being served as transportation hubs (e.g. Syntagma square, Piraeus port, etc.) or are close to touristic attractions (e.g. Monastiraki & Acropolis stations)
- Total number of restaurants established nearby can also be an indication if the location is good or not for dining
- Real estate indexes can also distinguish an emerging location from an emerged one

6. Conclusion

This analysis can be a good starting point for digging further into detail since there are other factors that should be taken into consideration apart from the number of similar burger restaurants located close to the metro line stations of Athens metropolitan area as well as their average distance to each metro lie station.