Athens Fine Dining

Predicting Suitable Areas for investment in the F&D sector

Alexandros Varesis

January 2021

Table of Contents

1. Introduction

1.1 Background

One of the most prolific "*tourism capitals*" of Europe, Athens Greece, displays numerous businesses in the F&D sector, a considerable portion of which is characterized as "*fine dining*" where among others venues such as restaurants, cocktail bars and lounge bars operate successfully. Even after the fiscal distress originated form the economic crisis of 2008 – 2010, Athens displays numerous opportunities for investment in this particular sector mainly due to the increased levels of tourism traffic. The aforementioned correlation is justified, as it is important to consider the fact that, the food and drinks (F&D) sector, has been heavily correlated with the overall industry of tourism. More specifically and according to Sims (2009, p. 321), "*…it is recognized that the kind of foods and drinks on offer for tourists can have major implications for the economic, cultural and environmental sustainability of tourism destinations…*". Therefore, it is rather crucial for one who wishes to invest in the domain of F&D to consider among any other aspects and characteristic of the location, the level of tourism, in order to define the local map of "fine dining".

1.2 Problem

The identification of the best place for an investment in the F&D sector needs significant consideration as the investors need to filter the location in terms of: a) local trends in every neighborhood (e.g., a place dominated by bars or restaurants), b) neighborhoods that favor "fine dining" due to increased local wealth or increased popularity and c) areas of high tourism traffic.

1.3 Interest

Data-driven identification of the most optimal area, through integrated information of tourism and similarity of venues, can help tremendously future investors in order to narrow down the possible locations for their investment. In fact, the overall process provides a unique tool that gives an immediate review of the city and the most popular areas for that type of investment.

2 Data acquisition and cleaning

2.1 Data Sources

For the present project to be successful several data-sets where utilized and processed following a logical sequence. Initially, the first step was for us to find the associated data of the neighborhoods in the capital of Greece. This was achieved by the following *Wiki* page where the municipalities and communities in Athens Metropolitan area where available. In addition to their name the data set, provided information (Table 1) in terms of population and area (km^2). In order to link the several municipalities of Athens with their "*fine-dining*" options the above data (from Table 1) was enhanced in order to attribute location coordinates and through the utilization of the FOURSQUARE API, request the respective revenues in each location.

Table 1 First five row of the acquired tabular data-set for the Athens's Municipalities.

	Municipality	Seat	Population	Area (km2)	Density (/km2)	YPES Code
0	Agia Varvara (Αγία Βαρβάρα)	Agia Varvara (Αγία Βαρβάρα)	30562	2.425	12603	101.0
1	Agia Paraskevi (Αγία Παρασκευή)	Agia Paraskevi (Αγία Παρασκευή)	59500	7.967	7498	102.0
2	Agioi Anargyroi (Άγιοι Ανάργυροι)	Agioi Anargyroi (Άγιοι Ανάργυροι)	32957	3.200	10299	104.0
3	Agios Dimitrios (Άγιος Δημήτριος)	Brahami (Μπραχάμι)	65173	4.949	13169	103.0
4	Aigaleo (Αιγάλεω)	Aigaleo (Αιγάλεω)	74046	6.450	11480	106.0

The final goal in terms of data acquisition was to add the element of tourism trends in the city of Athens. A robust way to do so was to integrate in our project free Airbnb data form *Inside Airbnb*. The data was sourced from publicly available information from the Airbnb site. This gave us the chance to have a clear look of the most popular areas as also the associated prices of the touristic residents. The data was download in a CSV type file by selecting the most recent prices and locations for 2020.

2.2 Data Cleaning and Feature Selection

Each data-set integrated into our project needed certain modifications in order to be utilized successfully and the data cleansing was performed gradually for each one of

them. In this part of the report, we present the overall specific modification applied in each data set.

2.2.1 Athens Municipalities

For the municipalities in Athens, it was mandatory to drop the "*YPES Code*" and the "*Seat*" columns which were irrelevant to our scope while in addition, delete the Greek characters in the parentheses which would have interfered with the overall processes. The following table (Table 2) illustrates the final form of the table after its integration with the "*Nominatim*" function which attributed the respective coordinates in each municipality.

Table 2 First five row of the acquired tabular data-set for the Athens's Municipalities with the attributed relevant coordinates.

	Neighborhood	Population	Area (km2)	Density (/km2)	Latitude	Longitude
0	Agia Varvara	30562	2.425	12603	37.990508	23.659813
1	Agia Paraskevi	59500	7.967	7498	38.006676	23.822694
2	Agioi Anargyroi	32957	3.200	10299	37.965406	23.612490
3	Agios Dimitrios	65173	4.949	13169	37.805011	23.864712
4	Aigaleo	74046	6.450	11480	37.991566	23.681875

2.2.2 Venues Categories

As mentioned in the upper part of the chapter, we utilized the FOURSQUARE API in order to find the venue categories in each municipality. Due to the fact that we registered with the "*free plan*" in the developer mode, the returns of our search were limited to 100 which is the default option. In order to get the venues which are considered to be in the category of "fine dining" we enhanced the search criteria with the aspect of "Price". In FOURSQUARE the price description translates into a scale of 1,2,3,4, 1 being the least expensive, 4 being the most expensive. Even though the correlation is not linear we argue that the "fine dine" choices should be pricier and therefore we searched for the venues attributed with the price tags of "3" and "4". Another consideration in this data acquisition was the radius of search. For the majority of the various municipalities in Athens borders, do not exceed the 1km radius and for some regions even the 500m. Hence our request included a radius of 500m for the returned venues. Finally, we performed the necessary data manipulation in the returned

venues as there was no consistency in some categories. More in particular we attributed the category of “Restaurant” in the venues with the tag “*Tavern, Steakhouse and BBQ Joint*” as also change the tag “*Café*” with “*Coffee Shop*”.

2.2.3 Airbnb Data

The Airbnb data set included several information the majority of which were irrelevant to our objective. The request returned 16 different columns from which we only utilized the ones of "Neighborhood”, “*Latitude*”, “*Longitude*” and “*Price*”. As in the case of the venues we wanted to find the areas which attract tourist with the ability to pay more for their Airbnb residency. As it is clearly illustrated in the following graph (Figure 1) the majority of the selections range among 30€ to 60€. In order to evaluate the areas where the tourists are more likely to spend money in "fine-dining" venues we established the threshold of 70€ per night and more.

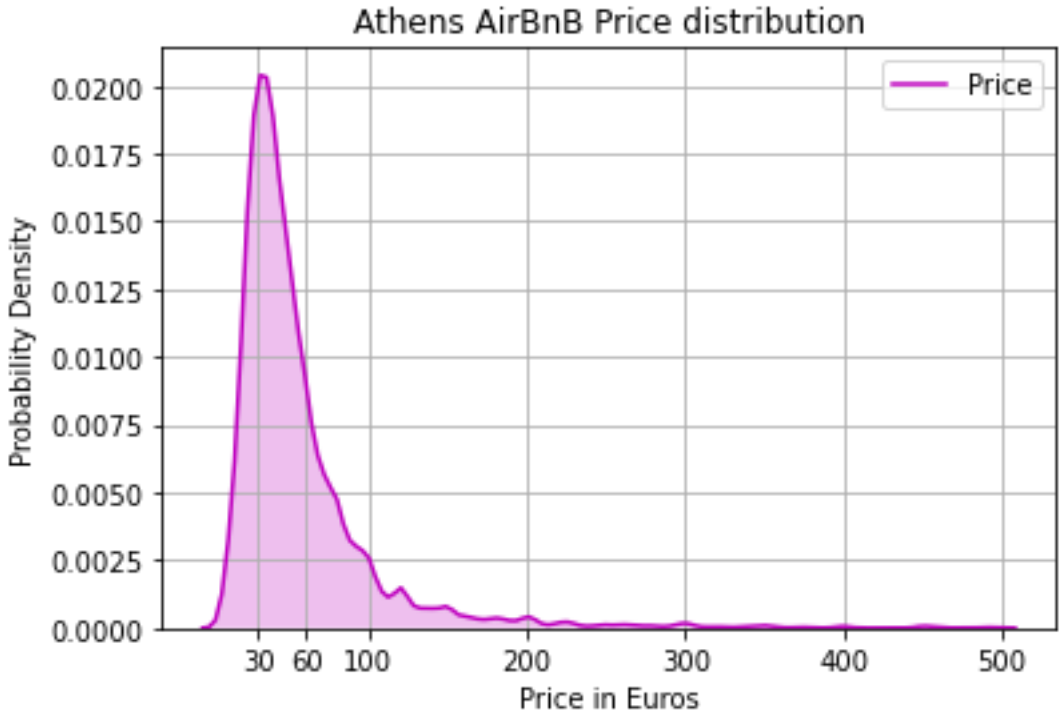

Figure 1 Graph illustrating the price distribution among the most recent Airbnb residencies in Athens.

3 Exploratory Data Analysis

3.1 Municipalities | Population and Area

The first task was to understand the population and area (km^2) distribution among the various municipalities in the broad Athens Metropolitan area. Initially, we observe the following:

- The overall population of Athens Metropolitan area is 2.667.440
- The overall area in km^2 of Athens Metropolitan area is 361,3
- The average population of each municipality is 55.572
- The average area in km^2 of each municipality is 7,53
- The Municipality with the lowest population is Penteli with 4.829 residents
- The Municipality with the highest population is Athens with 745.514 residents
- The Municipality with the least area in km^2 in is Nea Chalkidona with 0,8
- The Municipality with the largest area in km^2 is Athens with 38,964

As it is illustrated in the following graphs (Figure 2 & Figure 3) we see that the center of Athens dominating in both population and area. Due to the aforementioned we chose to narrow the data in the top-five and least-five populated municipalities and the respective municipalities in terms of area.

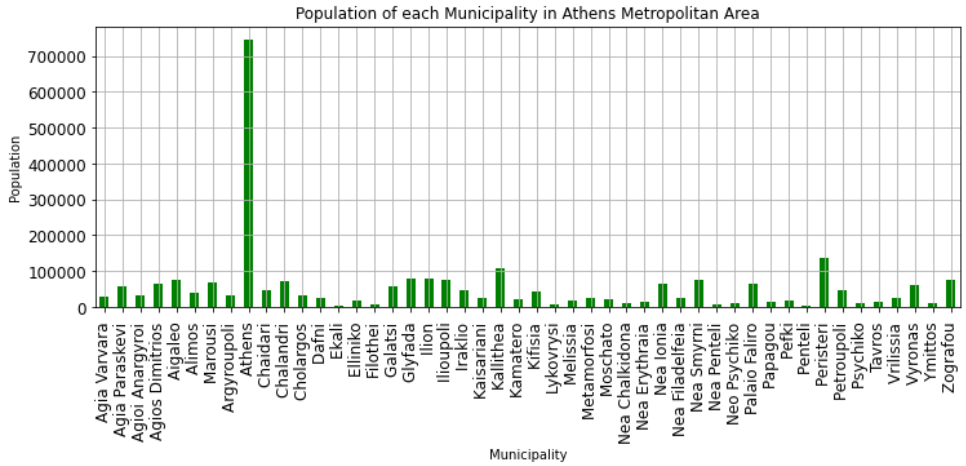

Figure 2 Graph illustrating the population distribution of each municipality in Athens.

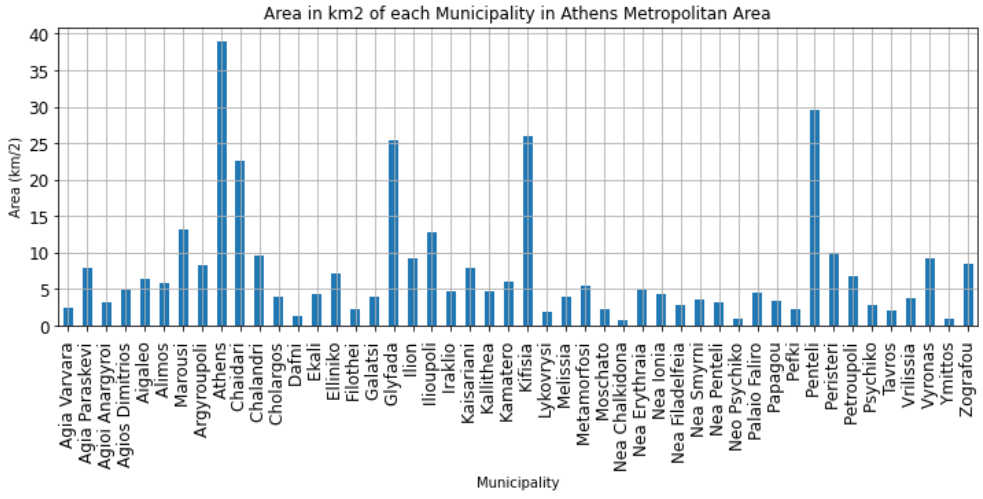

Figure 3 Graph illustrating the area (km^2) distribution of each municipality in Athens.

In Figure 4, it is easy to identify the top-five and least-five areas the main goal of this distinction is to add a qualitative approach in the process of underrating the urban development in the broader area of Athens. This approach tries to give a brief image of the current urbanization of each municipality as "*… land is the spatial carrier of urban populations and socioeconomic and ecological environments…*" (Zhang and Xie, 2019). The idea is that in urban expansion, the social economy plays a leading role and as a main carrier of urban development, construction land is the main driving force for urban economic growth in a certain stage (Zhang, Yang and Huo, 2016). The present graphs will give a quick image towards the investors in order to further evaluate their selection for a new venue in the end.

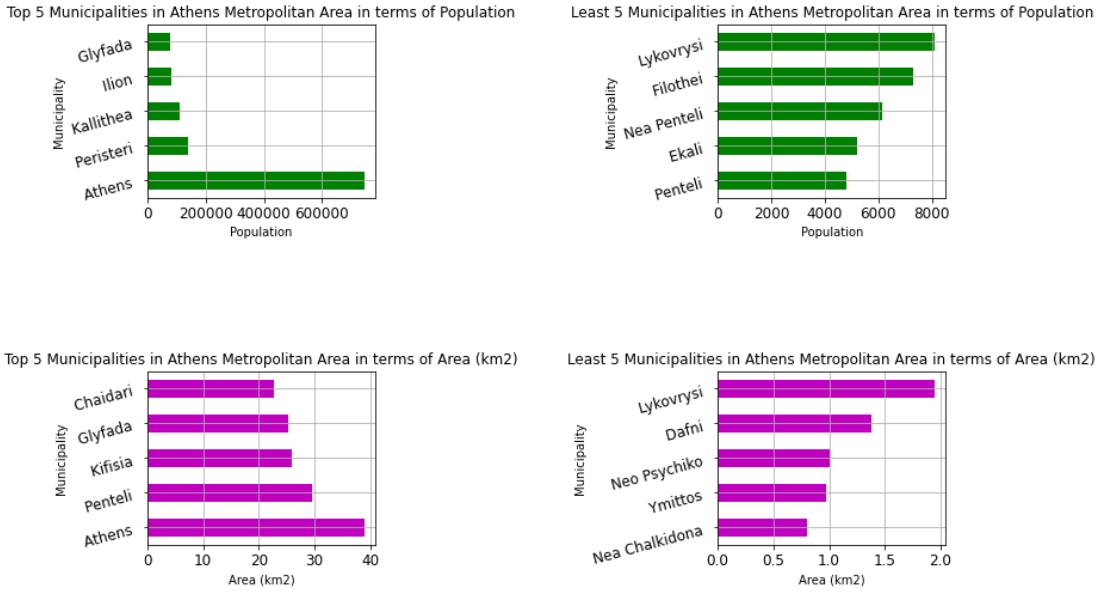

Figure 4 Set of graphs illustrating the top-five and least-five municipalities in terms of population and area (km^2).

3.2 Municipalities | Locations and Venues

After the successful integration of the municipalities with their corresponding coordinates (Figure 5) we utilized the FOURSQUARE API in order to return the respective venues. Overall, 56 venues were returned and as presented in the following graphs (Figure 6 & Figure 7), it is easy to identify the following facts:

- Overall, seven categories of venues where identified.
- Restaurants and Cocktail Bars are by far the most popular venue categories of "*fine dining*" in Athens.
- The municipalities of Chalandri, Glyfada and Petroupoli are the ones which illustrate the greatest number of venues.

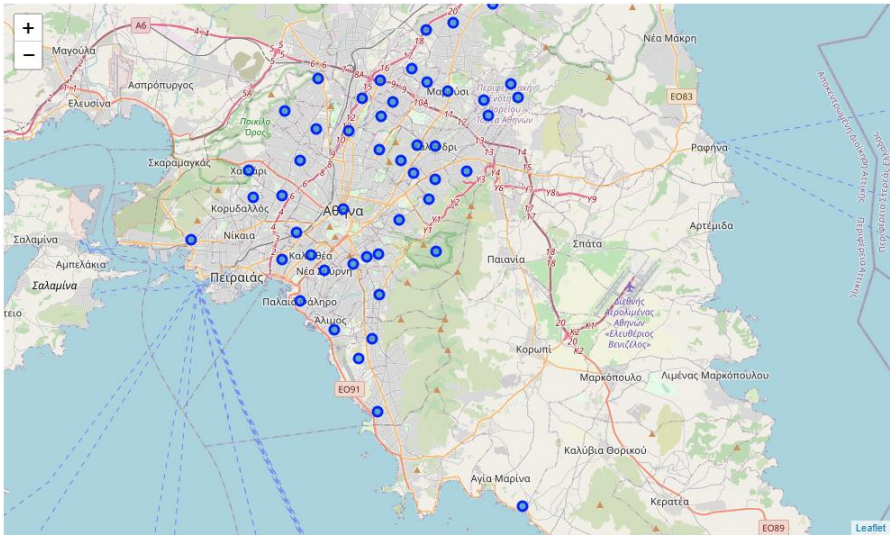

Figure 5 Map of Athens, Greece with the various municipalities plotted, indicated with blue circles.

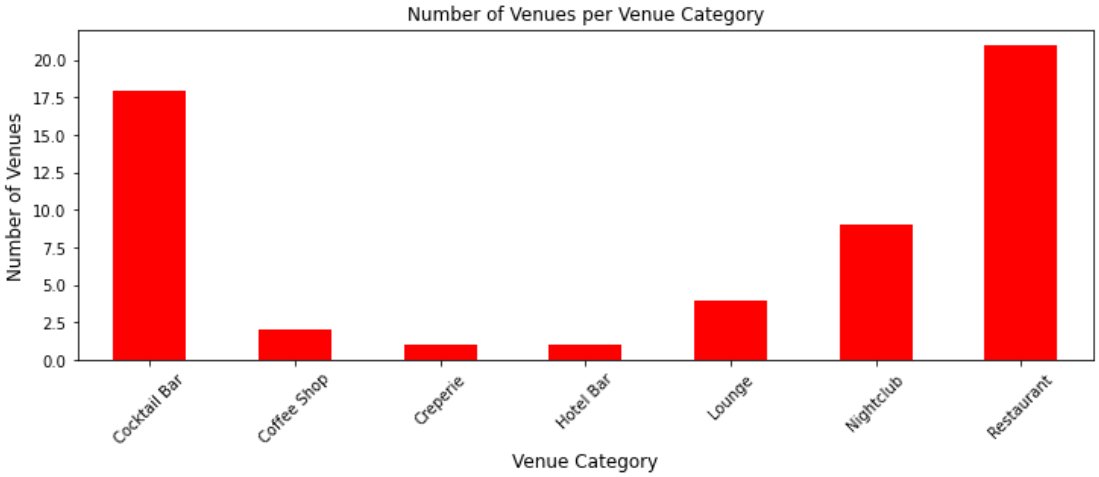

Figure 6 Map of Athens, Greece with the various municipalities plotted, indicated with blue circles.

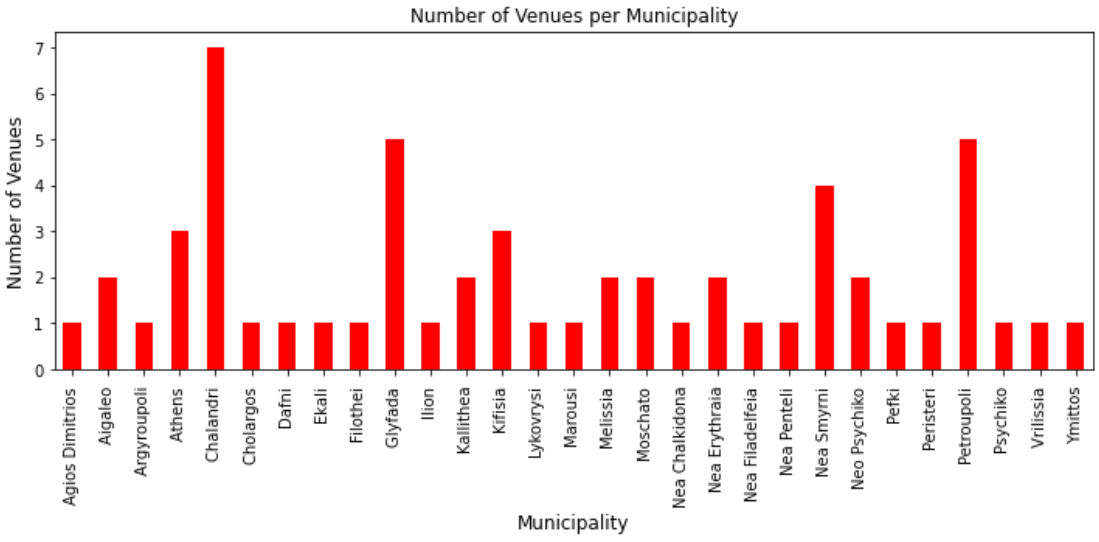

Figure 7 Map of Athens, Greece with the various municipalities plotted, indicated with blue circles.

3.3 Municipalities | Most Common Venues

The final step in the analysis of the various municipalities and the venues, was the evaluation of the most common types of venues per category. This was achieved by grouping each municipality and take the mean of frequency of occurrence of each venue category. As it is presented in the following table, we limited the analysis in the three most common venues per municipality.

Table 3 First five row of the three most common venues per municipality.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
0	Agios Dimitrios	Cocktail Bar	Restaurant	Nightclub
1	Aigaleo	Restaurant	Cocktail Bar	Nightclub
2	Argyroupoli	Nightclub	Restaurant	Lounge
3	Athens	Cocktail Bar	Hotel Bar	Restaurant
4	Chalandri	Cocktail Bar	Restaurant	Nightclub

4 Predictive Modeling

4.1 Clustering for understanding

According to Wu (2012), the application of clustering serve two purposes: a) the understanding of the data – set and b) the utilization of the data -set. In the present report and initiative, the purpose of understanding translates into the automatic identification of conceptually meaningful groups of objects that share common characteristics which in our case are “fine-dining” options in the broad area of Athens.

A simple definition of clustering is the process of organizing patterns (observations, data, or features) into clusters, where the members of a group are similar to each other according to a defined criterion. We chose to apply the algorithm of k-means clustering while we took under consideration the fact that clustering is a process where the result of the is fed back to the system, which by combining this result with the other inputs, proceeds to extract features and calculates the similarity relations, with key objective to provide the final extraction of the groups.

4.2 Algorithm Parameters

More specifically, regarding the attributes of the algorithm we chose to define the number of the times the k-means will be run with different centroid seeds to be “50” as also the maximum number of iterations for a single run to be “500”. Moreover, due to

the way in which the original centroids are defined which is not clear and therefore a fairly popular way of choosing the original centroids, is to be chosen at a random state.

4.3 Optimal Number of Clusters

One of the most common methods used to determine the number of the most optimal number of clusters is the "*Elbow Method*". It is a graphical method that depicts the variation of observations as a function of the number of groups and allows the visual identification of the point where the marginal gain from further grouping decreases. Thus, the optimal number of k is chosen after which no abrupt change of inclination of the curve is detected. Regarding our project and as shown in the following figure, the most optimal choice for our data set is either 5, 6 or 7 clusters.

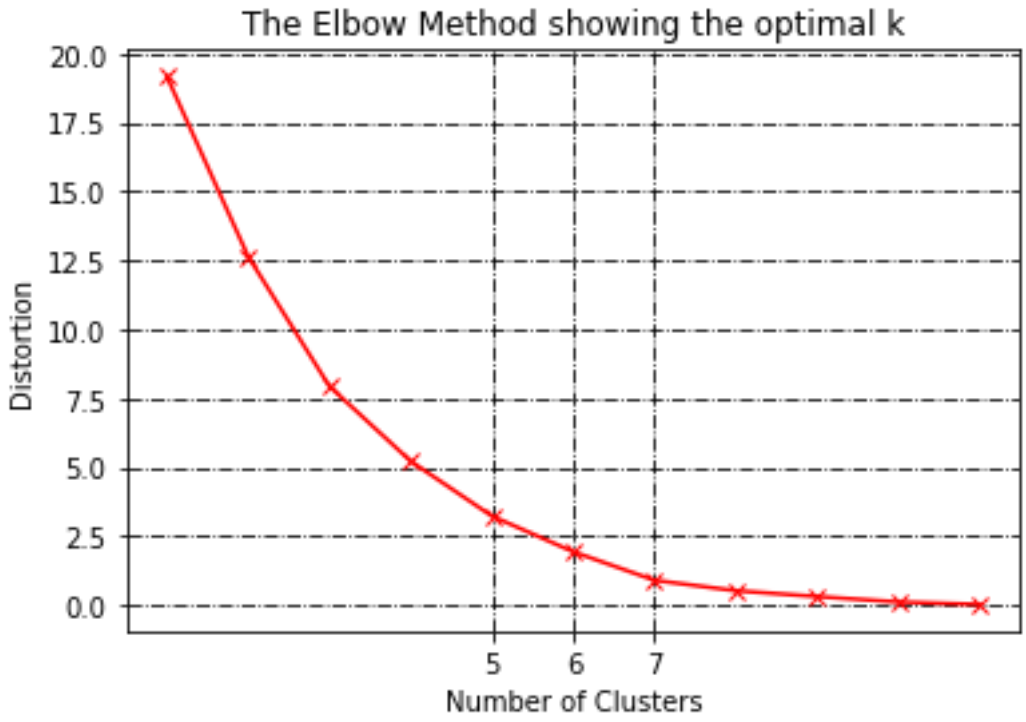

Figure 8 Graph illustrating the elbow method and the most optima k number for the present model.

5 Results

After the selection of the most optimal cluster number (i.e., k=6), the algorithm provided a data frame with each municipality and the correlative Venue Category, Cluster Label as also the three most common venues in the area. The distribution of the aforementioned clusters were plotted in a map resulting in the following figure.

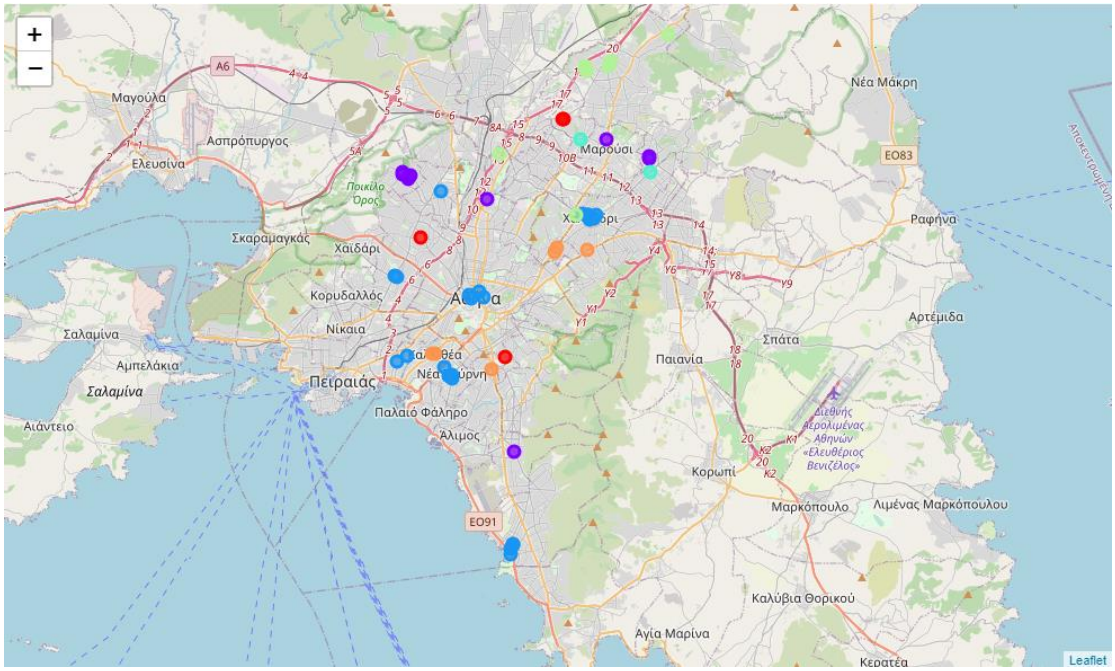

Figure 9 Map of Athens, Greece with the various municipalities plotted, indexed with the correlative cluster label.

Table 4 Table listing the corresponding cluster labels and their colors in the map with the respective venue category.

Cluster Label	Color	Venue Type
Cluster 1	Red	Lounge
Cluster 2	Magenta	Nightclub / Restaurant
Cluster 3	Blue	Cocktail Bar / Restaurant
Cluster 4	Blue-Green	Coffee Shop
Cluster 5	Green	Restaurant
Cluster 6	Orange	Cocktail Bar

Finally, in order to enhance the above map, we chose to superimpose the spatial distribution of the most expensive Airbnb choices in the city of Athens (Figure 10). What is interesting to observe is the fact that the majority of the Airbnb places are located in the center of Athens which is completely justified if one considers the fact that the most popular historical and archeological sites are near this area.

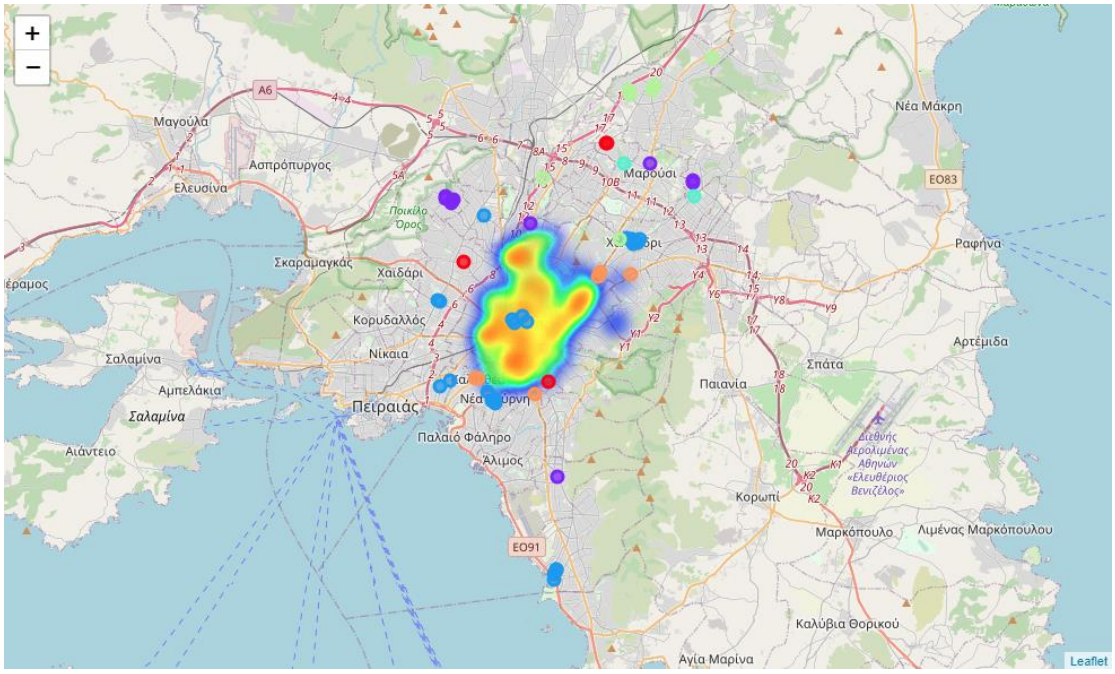

Figure 10 Map of Athens, Greece with the various municipalities plotted, indexed with the correlative cluster label.

6 Conclusions

The city of Athens indicates a rather dense grid of areas with increased potential for future investments in the sector of "fine-dining". According to the municipalities where we locate the most available venues are *Chaidari, Glyfada and Petroupoli* (Figure 7) while the most popular venue types are *Restaurants, Nightclub* and *Cocktail Bars* (Figure 6). The parameter of tourism which may be one of the most critical factors in the decision of an investor towards the F&D sector indicates that the best location is the down town center of Athens. Nevertheless, the generated results of our project provide adequate information for the initial planning on an investment in the sector as also a through understanding of the "*nature"* of each municipality.

7 Future directions

The process of identifying the venue categories and the correlative character of the various municipalities in Athens proves to be a robust way that can serve the initial planning of an investment team in order to get the "big picture" of the F&D setting of the city. Due to limitations of the data set as also the return results limit of the FOURSQUARE API we firmly believe that without these restrictions, the process can be enhanced and elevated in terms of precision and result fidelity.

References

Sims, R. (2009) ‘Food, place and authenticity: local food and the sustainable tourism experience’, *Journal of Sustainable Tourism*, 17(3), pp. 321–336. doi: 10.1080/09669580802359293.

Wu, J. (2012) *Advances in K-means clustering: a data mining thinking*. Springer Science & Business Media.

Zhang, W., Yang, D. and Huo, J. (2016) ‘Studies of the Relationship between City Size and Urban Benefits in China Based on a Panel Data Model’, *Sustainability*, 8(6), p. 554. doi: 10.3390/su8060554.

Zhang, Y. and Xie, H. (2019) ‘Interactive Relationship among Urban Expansion, Economic Development, and Population Growth since the Reform and Opening up in China: An Analysis Based on a Vector Error Correction Model’, *Land*, 8(10), p. 153. doi: 10.3390/land8100153.