

Bucharest : The Battle of the Neighborhoods

A. Introduction



Known for its wide, tree-lined boulevards, glorious Belle Époque buildings and a reputation for the high life (which in the 1900s earned its nickname of "Little Paris"), Bucharest, Romania's largest city and capital, is today a bustling metropolis.

Romanian legend has it that the city of Bucharest was founded on the banks of the Dambovită River by a shepherd named Bucur, whose name literally means "joy." His flute playing reportedly dazzled the people and his hearty wine from nearby vineyards endeared him to the local traders, who gave his name to the place [see more here](#).

A.1 Problem description and background discussion

According to [Wikipedia](#), the historical city centre was listed, In 2016, as 'endangered' by the World Monuments Watch. In 2017, Bucharest was the European city with the highest growth of tourists who stay over night, according to the Mastercard Global Index of Urban Destinations. As for the past two consecutive years, 2018 and 2019, Bucharest ranked as the European destination with the highest potential for development according to the same study.

In a 2016 [Telegraph article](#), among the to visit Bucharest, one can read :

- So drinking is incredibly cheap : "The beer prices alone are enough to make you consider leaving London for good. Even in touristy spots, two pints will set you back less than a fiver. And be sure to try out *țuică*, a traditional spirit made from plums"

- It's packed with hipster coffee spots: "Bucharest is home to an abundance of cultured graduates – and where young professionals go, hipster joints follow. Tasty local roasts and trendy waiters can be found at T-Zero, Origo, Diane 4 and The Coffee Factory.



- There are some fabulous summer parks: "Renting a boat at Cișmigiu Gardens is one of the most popular leisure activities in Bucharest, especially on a hot day. The city is not short on green spaces and Herăstrău Park is also a treat.



So in this project we will use Foursquare data in order to identify most common venues in Bucharest today. Then we will analyze these venues in order to propose a business idea in Bucharest.

A.2 Data description

For this problem, we will get the services of Foursquare API to explore the Bucharest data, in terms of neighborhoods. First of all, for this project we will use Bucharest neighborhoods (6 Sectors) as scrapped from [Wikipedia](https://en.wikipedia.org/wiki/Bucharest). In order to geographically situate the neighborhoods, latitudes and longitudes were manually added to the csv file used for the dataframe from [Latitude.to](https://latitude.to/).

Then, data will be extracted from Foursquare in order to include the information about the places around each neighborhood like restaurants, hotels, coffee shops, parks, theaters, art galleries, museums and many more. We will use machine learning technique, "Clustering" to segment the

neighborhoods with similar objects on the basis of each neighborhood data. These objects will be given priority on the basis of foot traffic (activity) in their respective neighborhoods. This will help to locate most common venues in Bucharest but also, will allow to evaluate possible business ideas in tourism area.

B. Methodology and Data Preparation

- Installing and Importing of Libraries
- Scrapping and Cleaning Bucarest Neighborhood Data by use of pandas
- Retrieving and Cleaning Bucharest Venues Data with Foursquare and geopy help
- Exploration of Location Data using Folium
- Encoding and clusterization with KMeans help

C. Results and Discussion

C.1 Scrapping Bucharest Neighborhoods

First of all, for this project we scrapped Bucharest neighborhoods (6 Sectors) as scrapped Wikipedia. In order to geographically situate the neighborhoods, latitudes and longitudes were manually added to the csv file used for the dataframe from Latitude.to. Then, we converted obtained data into a pandas dataframe shown below :

Table 1: Details of Bucharest Neighbhoods (Sectors) as scrapped from Wikipedia

	Neighborhood	Habitants	Surface(km2)	Density(/km2)	Latitude	Longitude
0	Sector-1	238217	68	3 339	44.49239	26.04877
1	Sector-2	362609	32	10 793	44.45280	26.13389
2	Sector-3	395565	32	11 336	44.42340	26.16860
3	Sector-4	170946	34	8 466	44.37536	26.12185
4	Sector-5	382183	29	9 365	44.38766	26.07161
5	Sector-6	376480	38	9 678	44.43543	26.01711

C.2. Retrieving Bucharest Venues from Foursquare

Next we extracted obtain Bucharest Venues data from Foursquare website and inserted into our dataframe.

Table 2 Bucharest Venues extracted from Foursquare

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Sector-1	44.49239	26.04877	Parcul Bazilescu	44.481900	26.041324	Park
1	Sector-1	44.49239	26.04877	Farmacia M	44.488351	26.050135	Pharmacy
2	Sector-1	44.49239	26.04877	Ground Zero Beer	44.499655	26.068552	Brewery
3	Sector-1	44.49239	26.04877	Stejarii Country Club	44.515902	26.069741	Gym / Fitness Center
4	Sector-1	44.49239	26.04877	Mega Image Concept Store	44.505960	26.080946	Supermarket
5	Sector-1	44.49239	26.04877	Lacul Băneasa	44.496865	26.070336	Lake
6	Sector-1	44.49239	26.04877	18 by the Embassy	44.477370	26.071553	Restaurant
7	Sector-1	44.49239	26.04877	Diplomat Restaurant	44.484538	26.079638	Restaurant
8	Sector-1	44.49239	26.04877	Hard Rock Cafe București	44.475554	26.076278	American Restaurant
9	Sector-1	44.49239	26.04877	Decathlon	44.505962	26.080966	Sporting Goods Shop

It worths to note that there are 136 different unique venue categories for Bucharest. If we extract a Top-10 Venue Category based on this data, one can see that Gym / Fitness Center comes in first place, quite surprisingly.

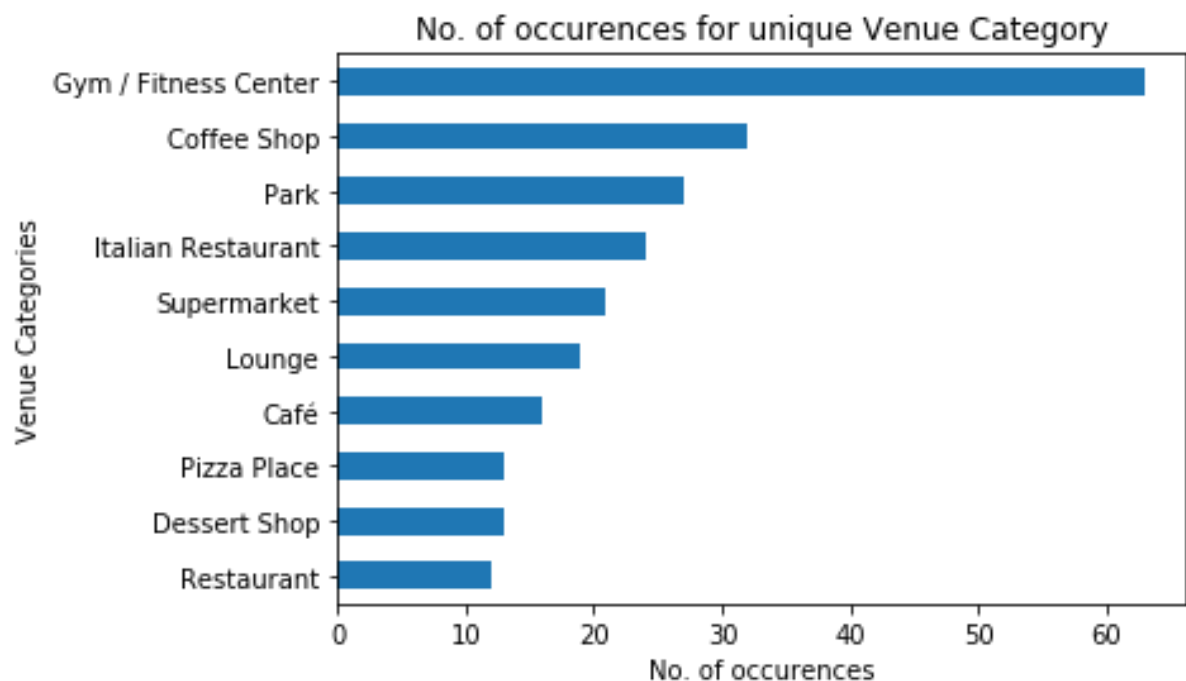


Figure 1 Top-10 of Venue Categories in Bucharest

Combining this observation with the fact that Parks arrive in third position, it seems that having good shape is very important in Bucharest! Coffee shops occupy second position in Top, in agreement with Telegraph article, mentioning that Bucharest *"It's packed with hipster coffee spots"*.

Further analysis of this Top-10 indicates "Italian Restaurant" in 4th position and "Pizza Place" in 7th position. Italian food seems to be appreciated in Bucharest! But "Restaurant" item in 9th position indicates that we need further clearing of data: we need to find and group all restaurants for the data.

Let's find all "restaurant occurrences in Venue Category and regroup them:

C.3. Data cleaning

Let's put Top-10 Restaurants on a graph and analyze this data. Once again, one can see that Italian food is highly appreciated in Bucharest, even more than Romanian Cuisine!

Proximity of Turkey have an impact on number of restaurants. We find also asian cuisine as chinese, sushi, or indian. It is remarkable that we did not found French Cuisine in Top-10, when we know that Bucharest is called "Little Paris"! Now let's reorganize this data and aggregate all restaurants under a unique category in order to evaluate impact on Top-10.

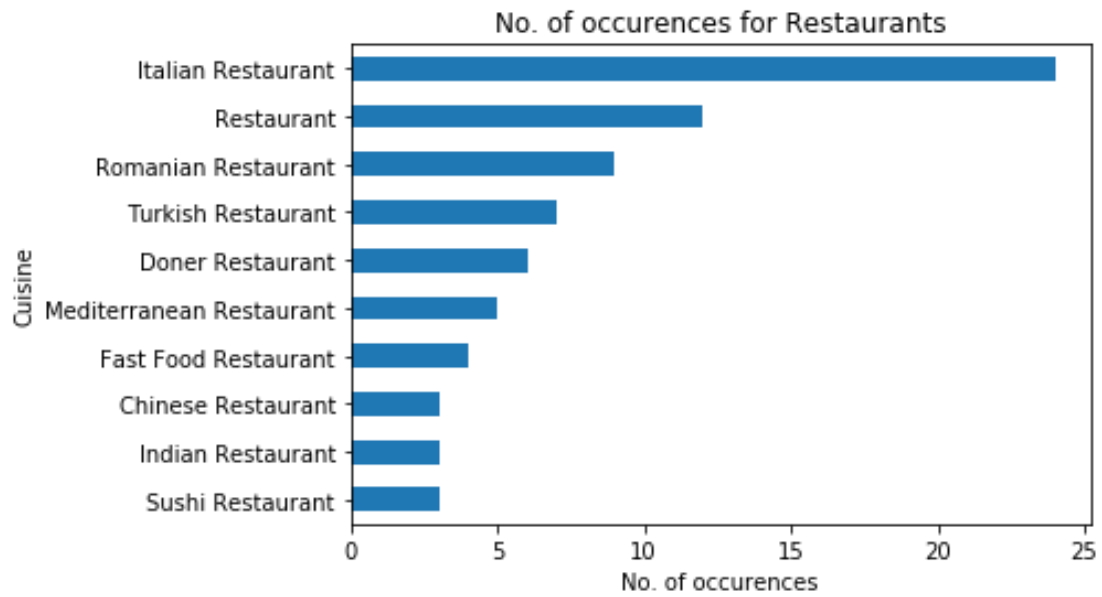


Figure 2 - Top-10 of restaurants in Bucharest

Now let's take a look again on the Top-10: the new graph shows clearly why Fitness Centers and Parks are in Top-10 Venues, the reason is the Restaurant venue in pole-position!

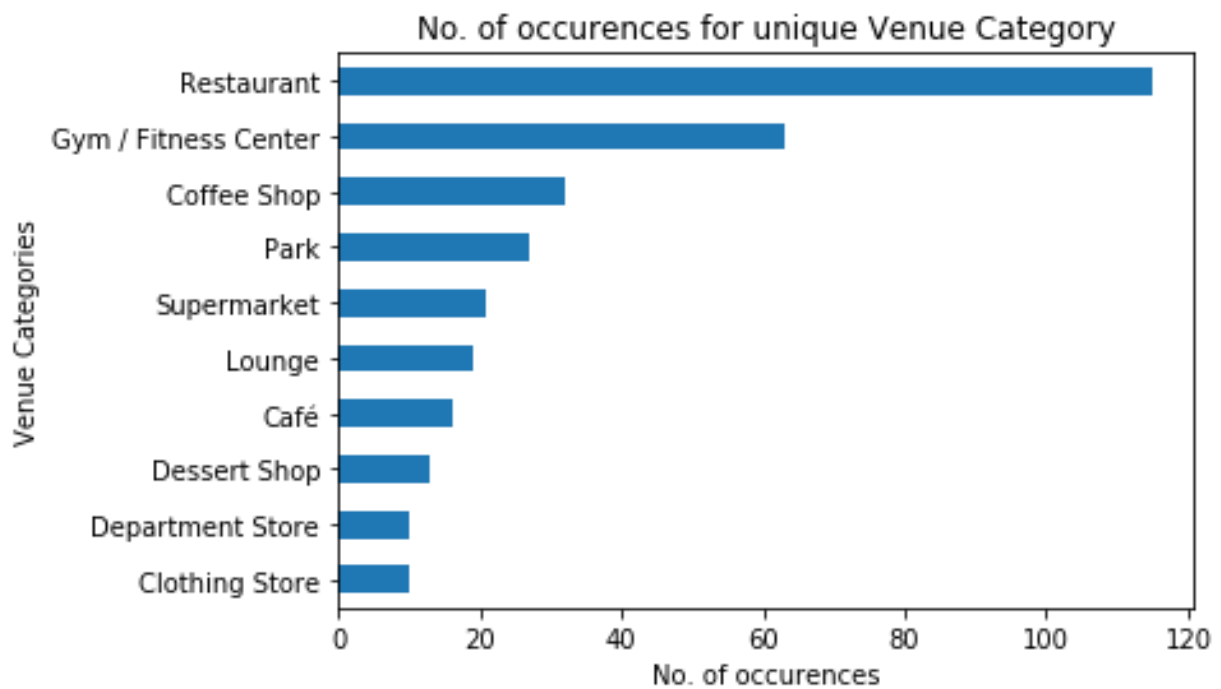


Figure 3 Top-10 of Venues in Bucharest after data cleaning

C.3. Venues localizations

Let's illustrate geographical localizations of these Venues by putting them on Sectors map.

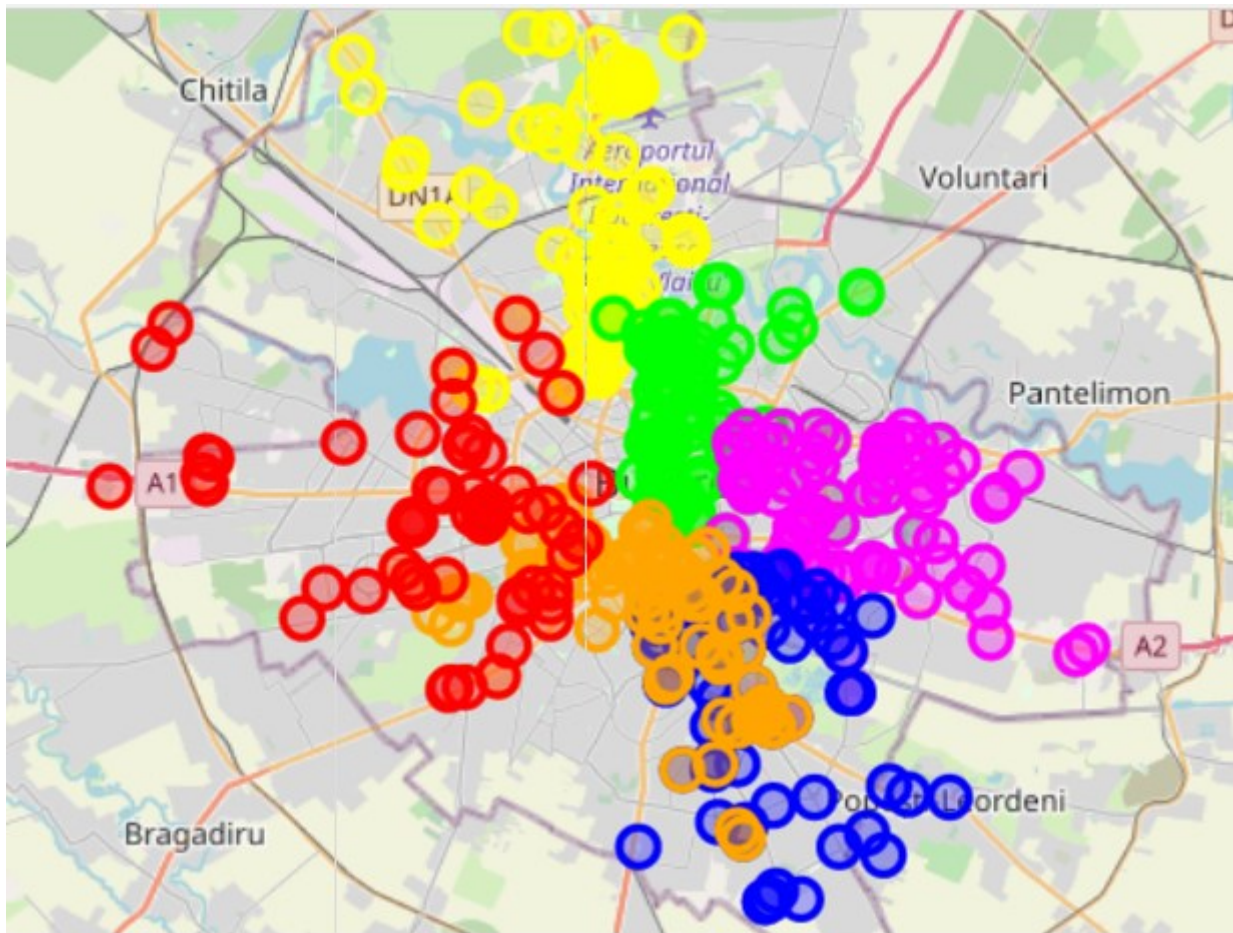


Figure 4 Venues Localizations on Bucharest Neighborhoods map

C.4. Venues localizations

After onehot encoding and frequency analysis, we create a dataframe of Top-5 Venues by Neighborhood :

Table 3 Top-5 Venue Categories by Neighborhood

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Sector-1	Restaurant	Dessert Shop	Gym / Fitness Center	Coffee Shop	Park
1	Sector-2	Restaurant	Coffee Shop	Gym / Fitness Center	Park	Dessert Shop
2	Sector-3	Restaurant	Gym / Fitness Center	Café	Lounge	Park
3	Sector-4	Restaurant	Gym / Fitness Center	Supermarket	Department Store	Lounge
4	Sector-5	Restaurant	Gym / Fitness Center	Park	Coffee Shop	Supermarket
5	Sector-6	Restaurant	Gym / Fitness Center	Coffee Shop	Park	Burger Joint

We can see that Restaurants are most common venue for each neighborhood. In second place we find Gym center, except for Sector-1 and Sector-2 where a Dessert Shop and Coffee Shop, respectively, appears.

C.5 Cluster Neighborhoods

We used then k-means in order to cluster the neighborhood into 4 clusters.

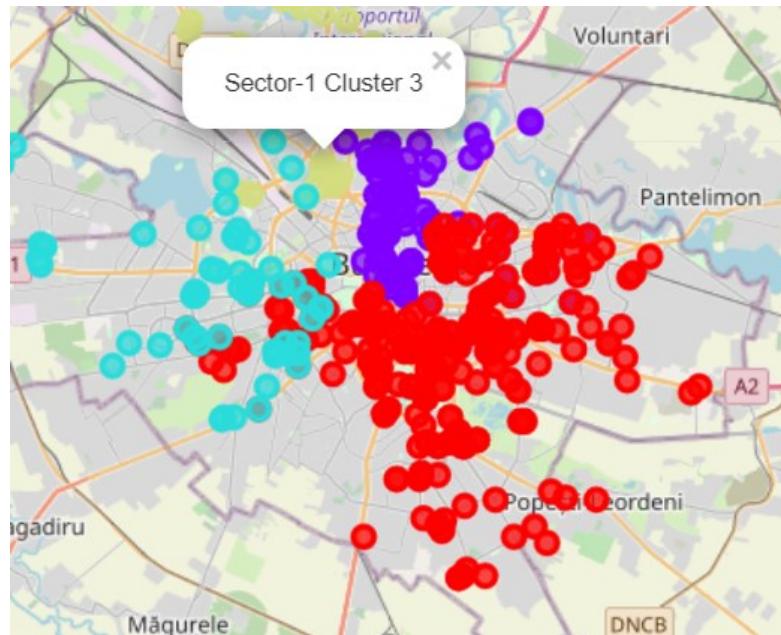


Figure 5 Folium visualization of the 4 clusters of most common venues in Bucharest

Finally we extracted only the most common restaurants localization for the 4 clusters in order to identify possible zone for a new, French cuisine, restaurant

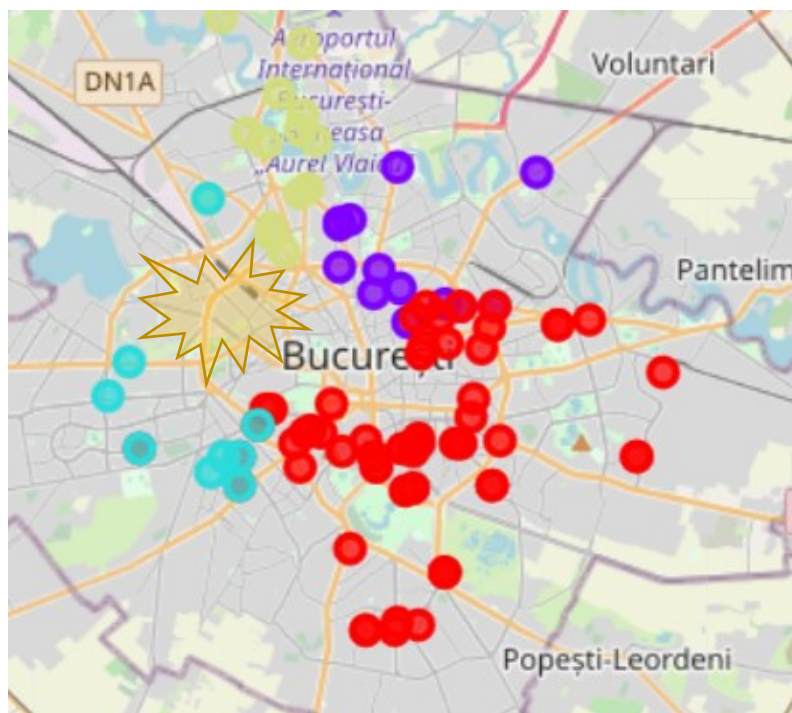


Figure 6 : Folium visualization of most common venue for restaurants in Bucharest

We can see here above that center and center-west of Bucharest not very dense in restaurants. Coroborated with previous observation that there's no French restaurant in Top-10 venues (in fact

there's a only one French restaurant in data we've obtained from Foursquare) we can say that we recommend to open a French cuisine restaurant in that area.

D. Conclusions and Perspectives

This study was inspired by a Telegraph article mentioning that in Bucharest there are a lot of coffee shops and parks, and also, that beer is cheap .

After scrapping Bucharest neighborhoods from Wikipedia, we inserted them in a dataframe together with most common venues from Foursquare. At first sight, the most common venue in Bucharest was Gym/fitness Centers, which was quite amazing! After data cleaning in fact the most common venue in Bucharest are restaurants. Their important number could be an explanation for cheap beer, but also for need of large number of Fitness centers 😊 What was surprising was there's no French cuisine in Top-10 restaurants in Bucharest, especially when we know that Bucharest is called the Little Paris so this is a nice business opportunity in Bucharest.

In agreement with Telegraph article, we found than in Top-5 venues the coffee shops and the parks.

Finally kmeans clusterization and map positioning visualization with Folium allows us to recommend to choose the area of center-west for opening a French cuisine restaurant in Bucharest.