**IBM Data Science Professional Certificate – Capstone Project**

|  |
| --- |
|  |
| The Battle of Neighborhoods |
|  |
| **Opening new Restaurant in Bucharest, Romania** |
|  |

**1. Introduction/Business Problem**

**1.1 Background**

Bucharest is the capital and largest city of Romania. It is located in the southeast of the country, at coordinates: 44°25′57″N 26°06′14″E, north of the Danube River and the Bulgarian border.

Bucharest occupied the **109th** place in 2019 for **quality of life** by the Mercer international surveys. Bucharest, at number 109, ranks low in the European list but better than other European cities such as Sofia (116th) and Belgrade (139th).

Bucharest is the center of the Romanian economy and industry, accounting for around 24% (2017) of the country's GDP and about one-quarter of its industrial production, while being inhabited by 9% of the country's population.

Bucharest has **2 million** populations and consists of **6 sectors**. Neighborhoods and populations in sectors are as below:

1. **Sector 1** *(population 227,717):* Dorobanți, Băneasa, Aviației, Pipera, Primăverii, Dămăroaia, Grivița
2. **Sector 2** *(population 357,338):* Pantelimon, Colentina, Iancului, Tei, Floreasca, Moșilor, Obor, Fundeni,
3. **Sector 3** *(population 399,231):* Vitan, Dudești, Titan, Centrul Civic, Dristor, Lipscani
4. **Sector 4** *(population 300,331):* Berceni, Olteniței, Progresul, Văcărești, Tineretului
5. **Sector 5** *(population 288,690):* Rahova, Ferentari, Giurgiului, Cotroceni, 13 Septembrie, Dealul Spirii
6. **Sector 6** *(population 371,060):* Giulești, Crângași, Drumul Taberei, Militari, Grozăvești(Regie), Ghencea

**1.2. Problem Statement**

In this project, we will investigate that if there is any good location in Bucharest for opening a restaurant.

**1.3. Interest**

Anyone who wants to open a new venue in any location may be interested in this project by customizing and changing the search criteria according to their needs.

**2. Data**

**2.1 Data sources**

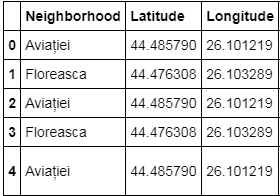
We can list the data as below for the chosen problem statement:

* We get the neighborhood data of Bucharest from Wikipedia
* We use python geocoder library to get geographical coordinates of neighborhoods
* We use **Foursquare API venues method** to get ranks and likes of restaurants by given venue id.
* We use **Foursquare API** **venues explore** method to get the venues of given neighborhoods of Bucharest.

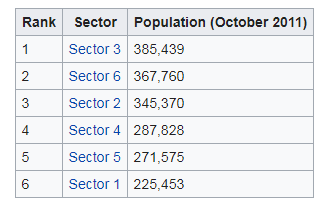
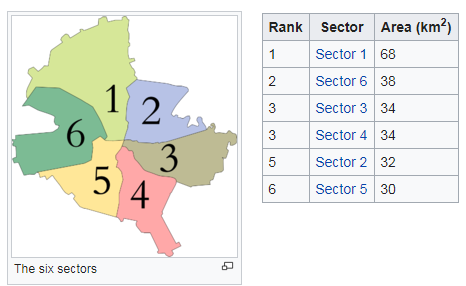
**2.2 Feature Selection and Data Usage**

We will use neighborhood location values to analyze Bucharest geography.

Folium library of python will be used to draw maps by using given latitude and longitudes of neighborhoods. Selected features will be as below



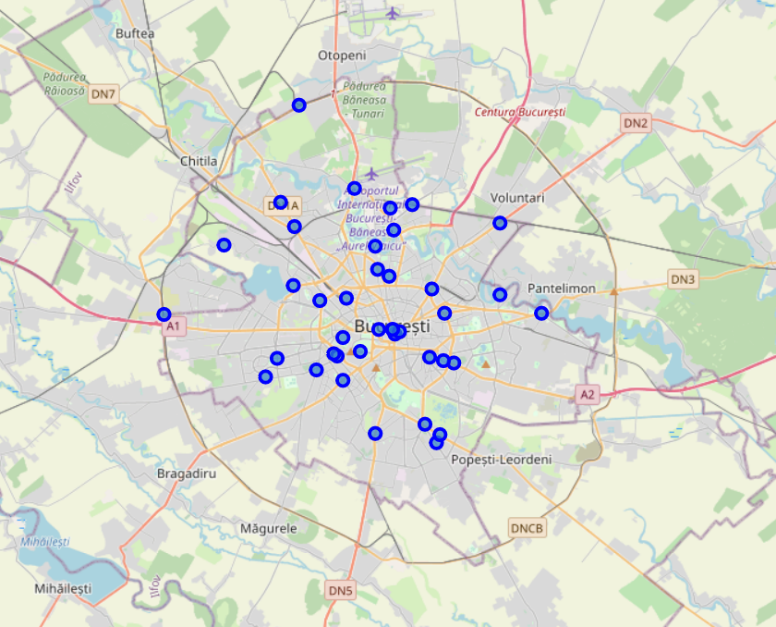
We will use populations of neighborhoods which has a large population and a new restaurant will be good for the neighborhood and the business.

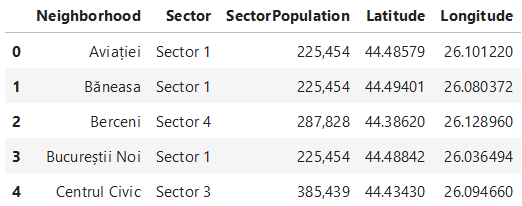
 

We will also classify neighborhoods by using venues distribution and counts. In this way, we will find similarities of neighborhoods which will help us to choose location for opening a new restaurant. Similar structured neighborhoods may handle same type of venue.

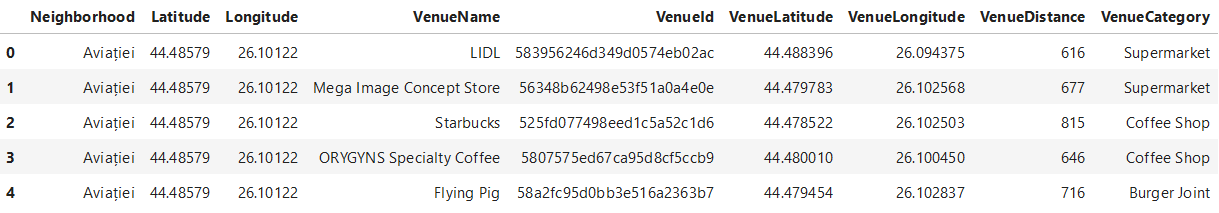


# 3. Methodology

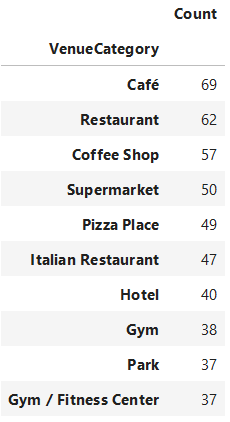
* The Bucharest data is available with the neighborhood name in Wikipedia. Location information (latitude and longitude) of neighborhoods are taken from geocoder library.
* Took neighborhood from Wikipedia and put blue dots on Bucharest map to see centers of neighborhoods. There are **40** neighborhoods in Bucharest
* We also got sector names of neighborhoods and populations of sectors. There are 6 sectors in Bucharest. I also merged Sector and Neighborhood data to see in which sector neighborhood places.

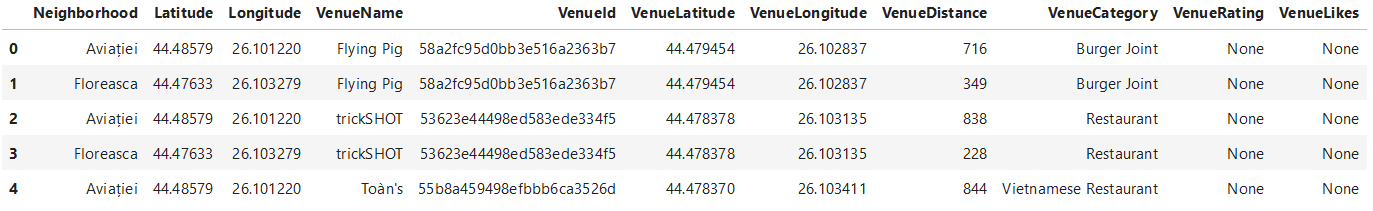


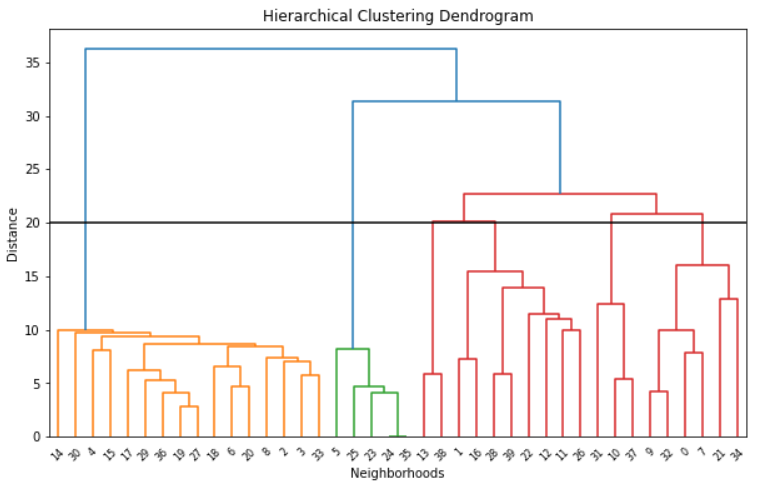
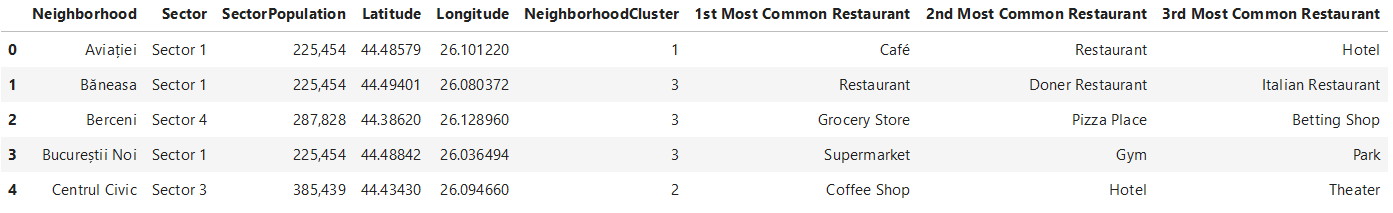
* By using Foursquare API we got venues **1 km** around center of each neighborhood with limit **100 venue**. We merged data with Neighborhood data as below

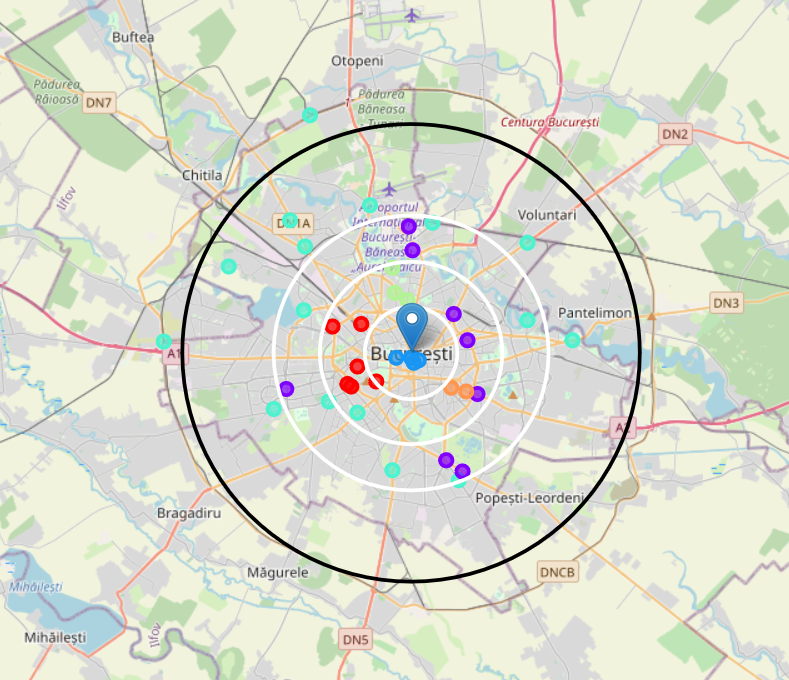
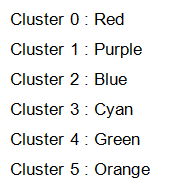
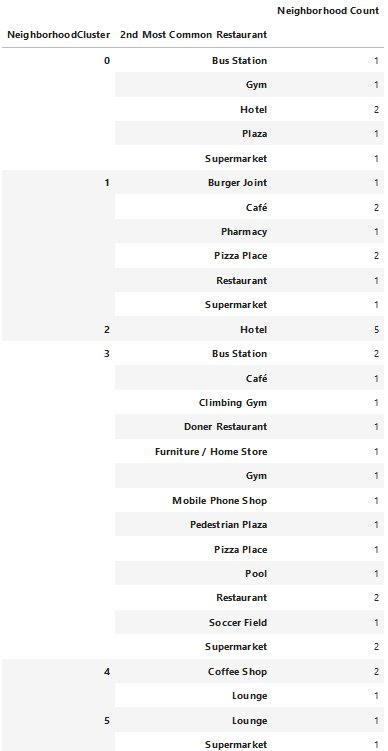


* There are 219 unique venue categories in Bucharest and top 10 venue counts is as below



* Used foursquare to get restaurant ratings by using venuid and merged with venue data

* Used Agglomerative Clustering to cluster neighborhoods according to count of venue categories in each neighborhood. According to the below graph, I decided to separate our neighborhoods into 6 clusters (cut at distance of 20, horizontal black line).



# 4. Results

* Total number of restaurants: 477
* Total number of Turkish restaurants: 22

From heat maps, we can see that there are

* Too many restaurants in 4 km to the center
* Enough restaurants between 4-6 km to the center
* There are **not many**restaurants after **6 km to center**
* Turkish restaurants are located in South, North , East, between 2-4 Km distance around center
* There are **not** Turkish Restaurants in West and North West part
* Some Turkish restaurants gets bad points. That means people do not prefer these restaurant too much and any turkish restaurant may be opened around

# 5. Discussion

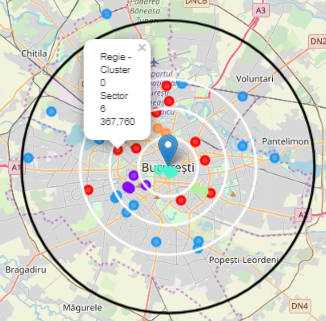
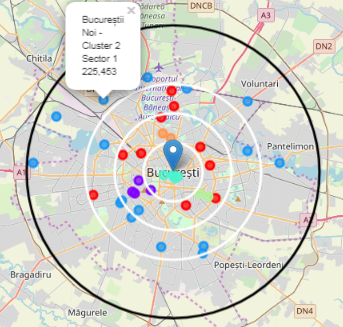
Cluster 2 suffers from restaurant. Especially there is no Turkish restaurant in west part of Cluster 2 . We can think about Cluster 2 , West part

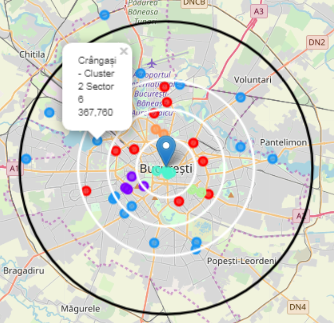
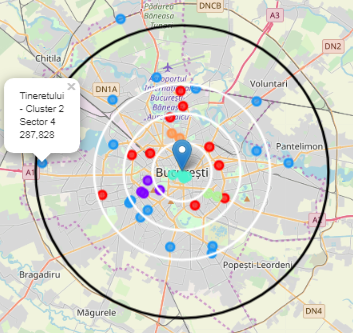
Cluster 0 especially likes Italian tastes , may be these neighborhoods will not like Turkish tastes . We can only think about 1 location may be.

Cluster 1, 2, 3 are centers full of restaurants as we saw in the heat map. . So we cannot think about these areas

# 6. Conclusion

We have chosen 4 locations for opening Turk restaurant. Below are locations of neighborhoods. Hope this work will help somebody who wants to open a new ethnic restaurant in Bucharest.

* Neighborhood : Regie - Cluster 0, Sector 6, 367.760

I choosed Regie , since west part of Bucharest has not Turkish restaurant and Regie is close to center . Regie is red cluster which is same with the other 2 red markers having Turk restaurant. Red clusters populations may like Turkish food.

* Neighborhood : Bucureștii Noi , Cluster 2, Sector 1 , Population : 225453

I choosed Bucharest-noi , since west part of Bucharest has not Turkish restaurant and even not too much restaurant after 6 km distance to center . One restaurant will be good for this area.

* Neighborhood : Crângași , Cluster 2, Sector 6 , Population : 367760

I choosed Crangasi , since Crangasi and west part of Bucharest has not Turkish restaurant and from bar chart we can see that this neighborhood has not too much restaurant. Also this neighborhood is in Sector 6 and this sector has biggest population which will handle one more restaurant

* Neighborhood : Tineretului , Cluster 2, Sector 4 , Population : 287828

I choosed Tineretului , since Tineretului has not Turkish restaurant and from bar chart we can see that this neighborhood has not too much restaurant.