Capstone Project - The Battle of the Neighborhoods (Week 2)

Introduction/Business Problem

In this project I will compare Ankara and Istanbul which are two of the three largest cities in Turkey. Ankara is the capital city of Turkeya and Istanbul is one of the largest metropolises in the world.

While more than 15 million people live in Istanbul, Ankara has a population of about 5 million. Ankara is the political center and Istanbul is the trade center of the country. Also, historical texture and cultural activities make Istanbul a center of attraction.

I will examine and compare the neighborhoods of two cities for representing the skills that I have gained during the IBM Data Science Professional Certificate program.

Data

Second-level Administrative Divisions of the Turkey from Spatial Data Repository of NYU (https://geo.nyu.edu/catalog/stanford-nj696zj1674 (https://geo.nyu.edu/catalog/stanford-nj696zj1674)) will be used for obtaining the boroughs of two cities. This source provide us data related to districts of all cities in Turkey. I will filter the data for creating datasets of boroughs of Ankara and Istanbul.
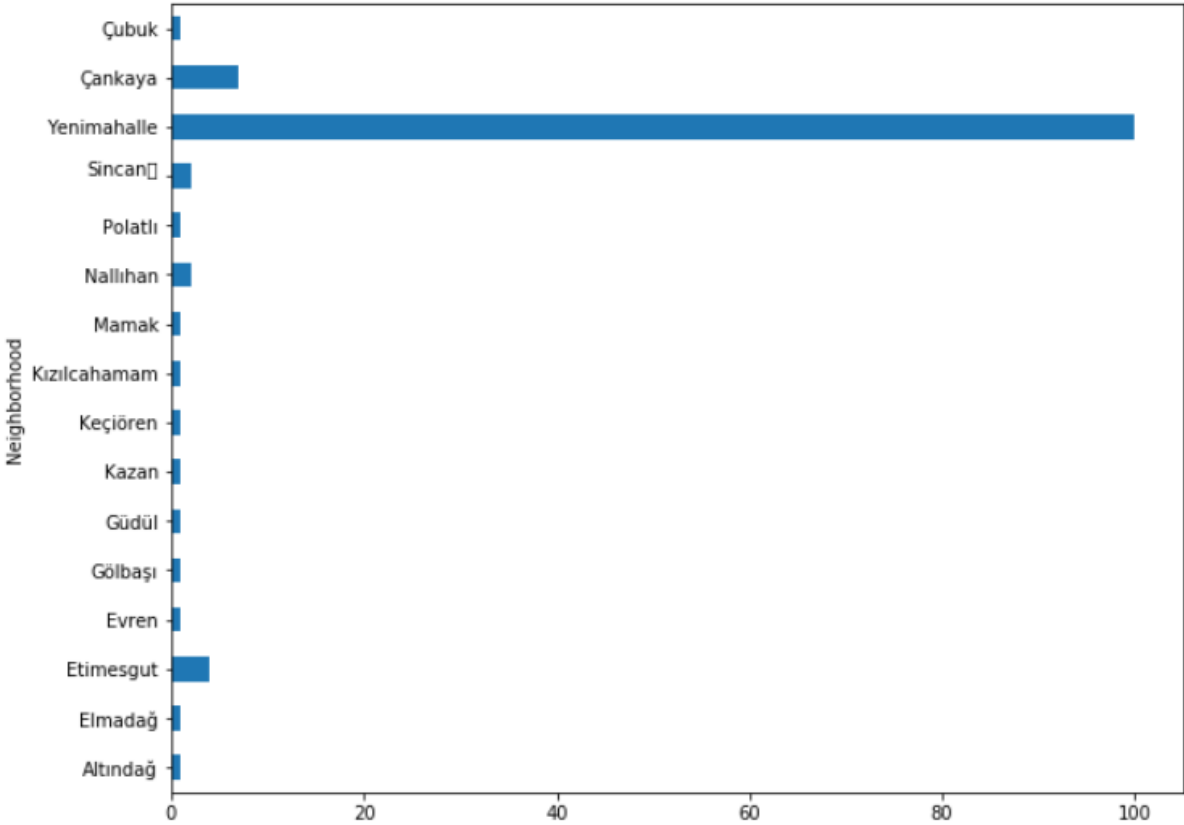
Forsquare API will be used to get the most common venues of given Borough of Ankara and Istanbul.
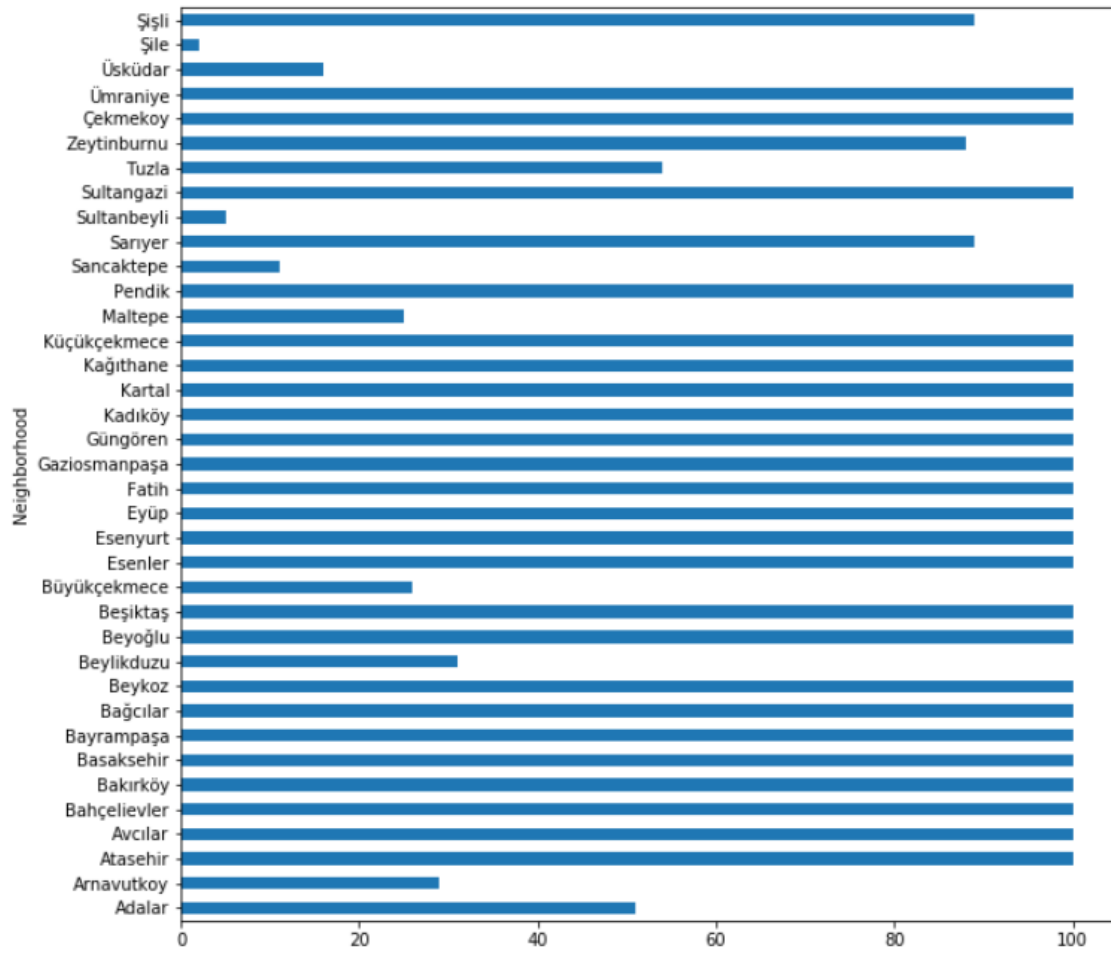
Exploratory Data Analysis

Datasets of boroughs of Ankara and Istanbul have been created by using the data obtained from Second-level Administrative Divisions of the Turkey from Spatial Data Repository of NYU. There are 38 boroughs for Istanbul and 24 boroughs for Ankara.

The most common avenues of given boroughs of Ankara and Istanbul have been obtained by using Foursquare API. Since the coordinates of neighborhoods in boroughs could not be reached, coordinates of boroughs have been used for getting the venues. Radius value has been specified as 1000m. Bar plots of venues in the boooroughs have been shown below graphs.

Bar Plot Of Ankara Venues

Bar Plot Of İstanbul Venues

I have obtained 126 venues for Ankara and 2916 venues for Istanbul. Information of venues could not be reached for 8 of 24 boroughs of Ankara and 1 of 38 boroughs of Istanbul. Then, the most common venues have been examined for two cities and tables have been created for the 5 most common venues in the boroughs.

Table of 5 Most Common Venues in Boroughs of Ankara

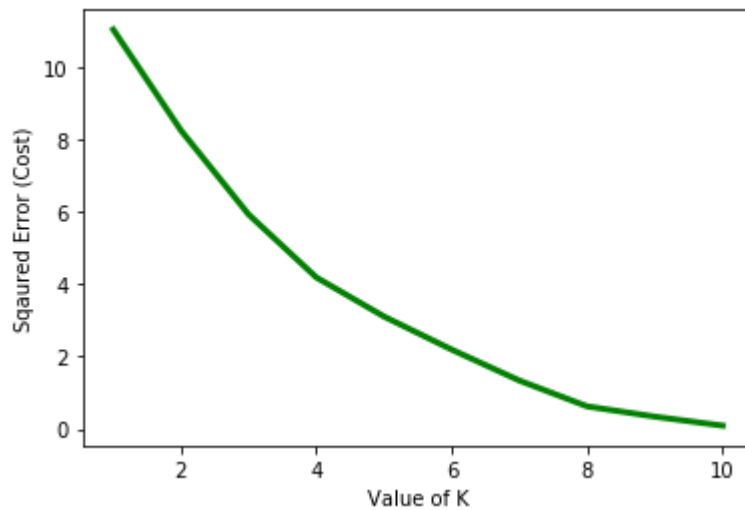|	Neighborhoods	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Altındağ	Garden	Turkish Restaurant	Donut Shop	Hookah Bar	Historic Site
1	Elmadağ	Mountain	Turkish Restaurant	Donut Shop	Historic Site	Herbs & Spices Store
2	Etimesgut	Farm	Mountain	Costume Shop	Turkish Restaurant	Donut Shop
3	Evren	Exhibit	Turkish Restaurant	Donut Shop	Hookah Bar	Historic Site
4	Gölbaşı	Farm	Turkish Restaurant	Donut Shop	Hookah Bar	Historic Site
5	Güdül	Campground	Turkish Restaurant	Donut Shop	Hookah Bar	Historic Site
6	Kazan	Racetrack	Doner Restaurant	Historic Site	Herbs & Spices Store	Gym / Fitness Center
7	Keçiören	Farm	Turkish Restaurant	Donut Shop	Hookah Bar	Historic Site
8	Kızılcahamam	Garden	Turkish Restaurant	Donut Shop	Hookah Bar	Historic Site
9	Mamak	Mountain	Turkish Restaurant	Donut Shop	Historic Site	Herbs & Spices Store
10	Nallıhan	Lake	Campground	Turkish Restaurant	Donut Shop	Historic Site
11	Polatlı	Historic Site	Turkish Restaurant	Donut Shop	Hookah Bar	Herbs & Spices Store
12	Sincan\r\n	Garden	Turkish Restaurant	Donut Shop	Hookah Bar	Historic Site
13	Yenimahalle	Café	Turkish Restaurant	Park	Bakery	Dessert Shop
14	Çankaya	Restaurant	Turkish Restaurant	Paintball Field	Farm	Moving Target
15	Çubuk	Farm	Turkish Restaurant	Donut Shop	Hookah Bar	Historic Site

Table of 5 Most Common Venues in Boroughs of İstanbul

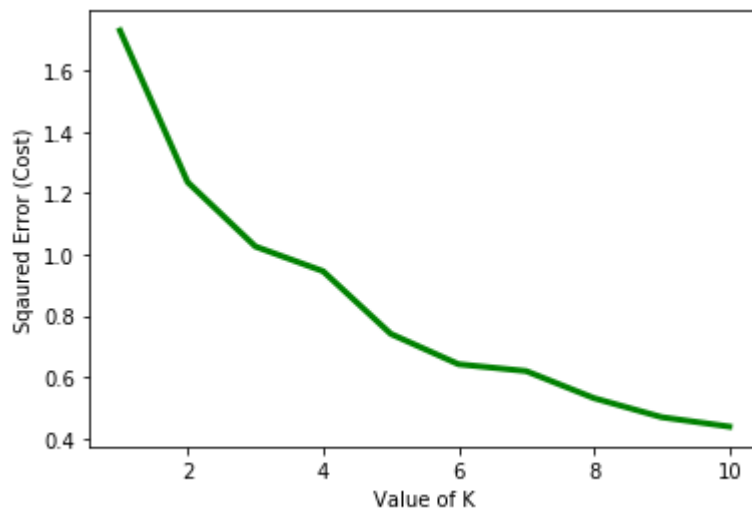|	Neighborhoods	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Adalar	Café	Seafood Restaurant	Dessert Shop	Beach	Boat or Ferry
1	Arnavutkoy	Turkish Restaurant	Farm	Wings Joint	Tea Room	Plaza
2	Atasehir	Café	Park	Coffee Shop	Turkish Restaurant	Restaurant
3	Avcılar	Café	Turkish Restaurant	Dessert Shop	Steakhouse	Restaurant
4	Bahçelievler	Turkish Restaurant	Café	Gym / Fitness Center	Gym	Bakery
5	Bakırköy	Café	Turkish Restaurant	Dessert Shop	Steakhouse	Restaurant
6	Basaksehir	Coffee Shop	Hotel	Clothing Store	Sporting Goods Shop	Steakhouse
7	Bayrampaşa	Café	Coffee Shop	Clothing Store	Cosmetics Shop	Park
8	Bağcılar	Café	Turkish Restaurant	Gym / Fitness Center	Gym	Clothing Store
9	Beykoz	Café	Seafood Restaurant	Breakfast Spot	Restaurant	Historic Site
10	Beylikduzu	Turkish Restaurant	Boat or Ferry	Restaurant	Food	Bakery
11	Beyoğlu	Hotel	Restaurant	Coffee Shop	Soccer Stadium	Café
12	Beşiktaş	Café	Fast Food Restaurant	Gym	Restaurant	Steakhouse
13	Büyükçekmece	Beach	Hotel	Surf Spot	Resort	Motel
14	Esenler	Café	Coffee Shop	Clothing Store	Cosmetics Shop	Park
15	Esenyurt	Café	Turkish Restaurant	Electronics Store	Coffee Shop	Clothing Store
16	Eyüp	Café	Restaurant	Turkish Restaurant	Bakery	Historic Site
17	Fatih	Café	Restaurant	Art Gallery	Coffee Shop	Steakhouse
18	Gaziosmanpaşa	Café	Turkish Restaurant	Restaurant	Gym / Fitness Center	Coffee Shop
19	Güngören	Clothing Store	Turkish Restaurant	Gym / Fitness Center	Café	Hotel
20	Kadıköy	Restaurant	Coffee Shop	Gym / Fitness Center	Kebab Restaurant	Turkish Restaurant
21	Kartal	Café	Restaurant	Bar	Turkish Restaurant	Seafood Restaurant
22	Kağıthane	Turkish Restaurant	Café	Hotel	Hookah Bar	Restaurant
23	Küçükçekmece	Café	Clothing Store	Turkish Restaurant	Fast Food Restaurant	Coffee Shop
24	Maltepe	Pet Store	Restaurant	BBQ Joint	Steakhouse	Buffet
25	Pendik	Café	Restaurant	Bar	Turkish Restaurant	Seafood Restaurant

Clustering

Firstly, I have applied elbow method for specifying the number of clusters for K-means clustering. According to the plots I have specified the number of clusters for Ankara as 8 and for Istanbul as 6.

Elbow Method for Ankara
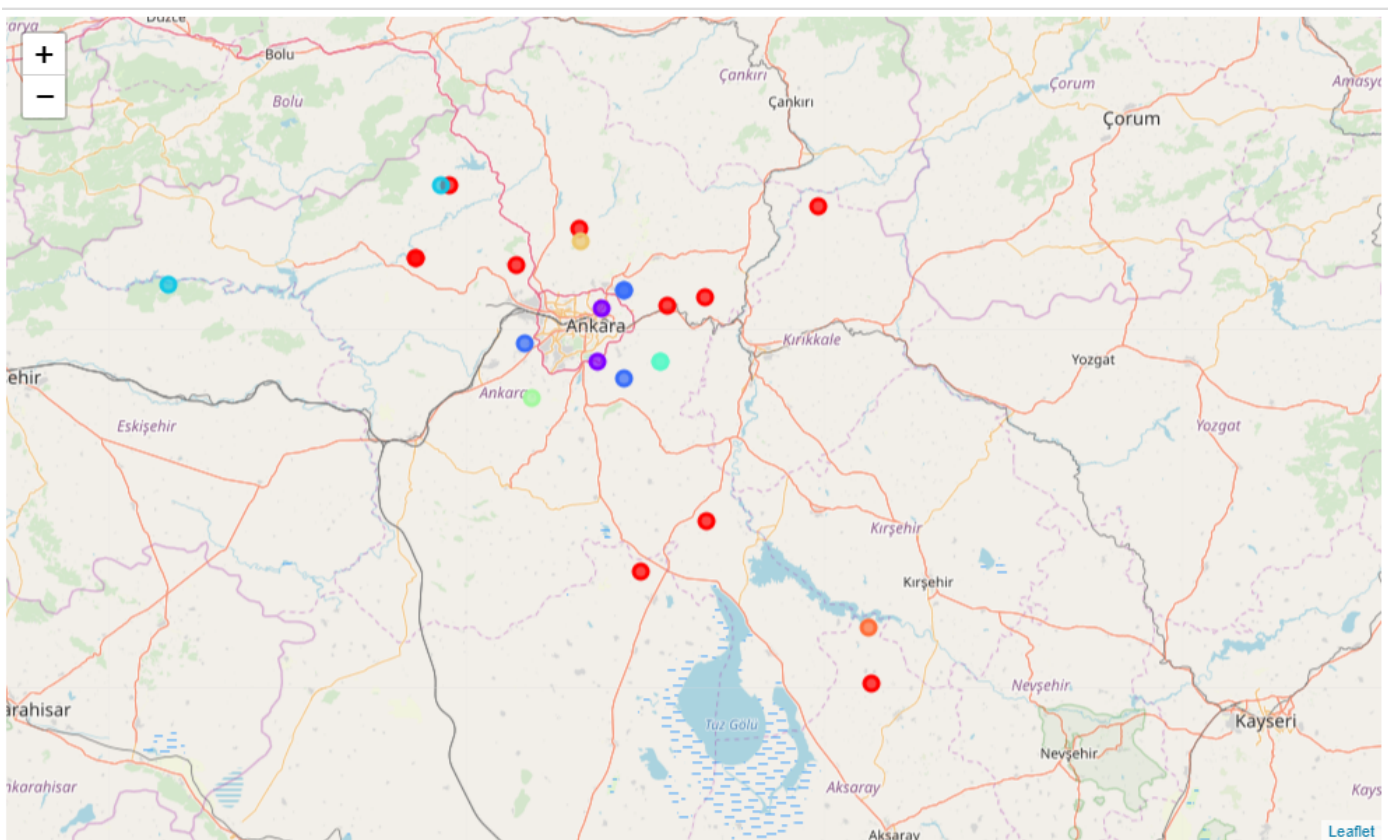

Elbow Method for Istanbul

Then k-means clustering is applied with the given number of clusters. Then the tables were merged with the cluster numbers. Lastly, the clusters have been examined on the maps.
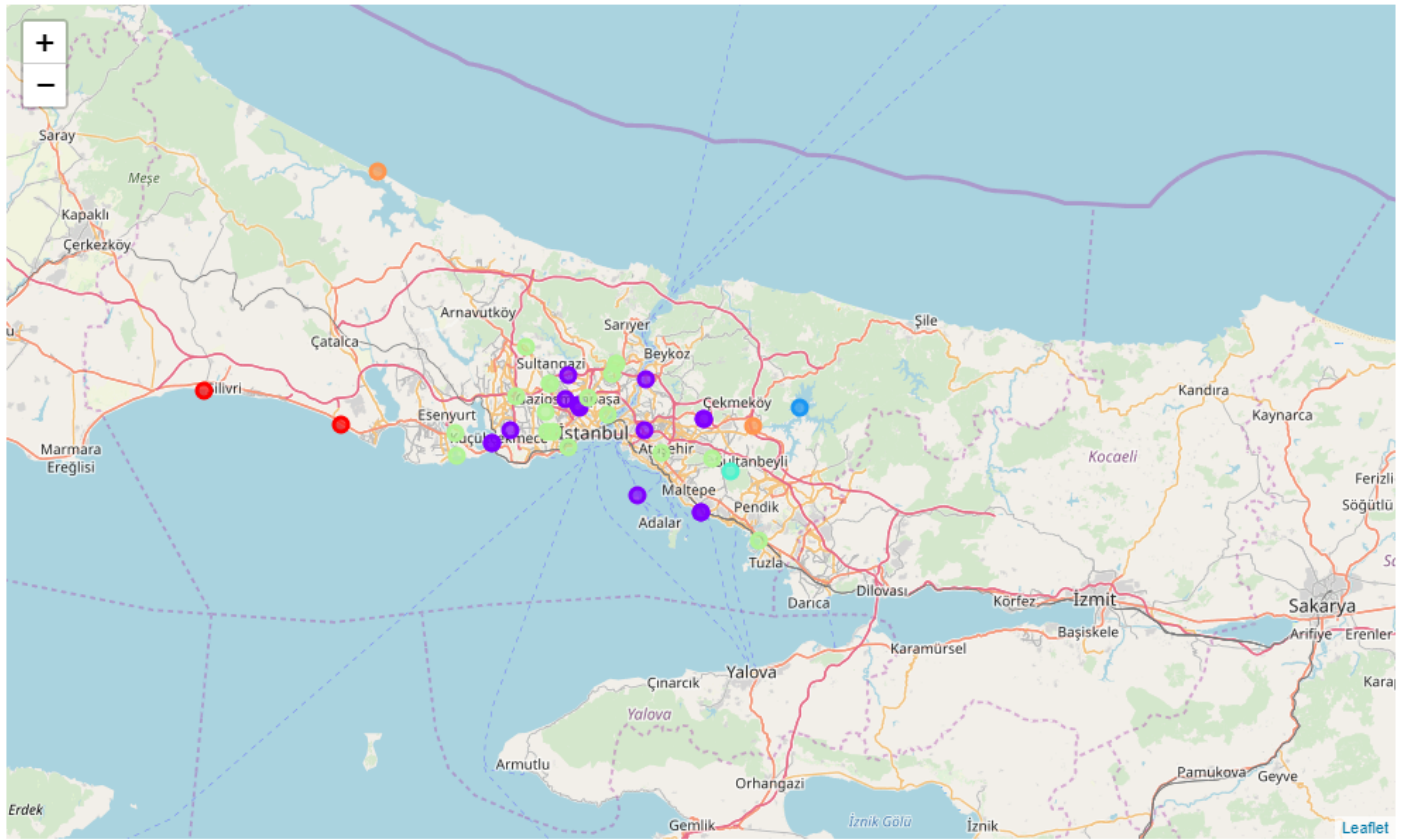
Results

When the number of venues in the boroughs of two cities are compared it is seen that there is a big difference between the cities. While most of the boroughs of Istanbul include more than 10 venues in Ankara only Yenimahalle has venues more than 100. In İstanbul most common venues are cafes, restaurants, coffee shops and stores. Turkish restaurants, farms, gardens, donut shops are the most common venues in Ankara. The groups could be seen in the following maps.

Maps of Cluester of Venues in Boroughs of Ankara

Maps of Cluester of Venues in Boroughs of Istanbul

Discussion

Firstly, since the coordinates of the neighborhoods in the boroughs could not be reached venues could not be reached effectively. Although Çankaya is known to be the center of the city there are not as much places as in Yenimahalle. If the neighborhood coordinates are obtained the venues in Çankaya can be obtained more effectively. This can effect the results of clustering.

As I mentioned before Istanbul is more crowded than Ankara.For that reason venues are seem to be more intensed in IStanbul boroughs. Also, the venues in Ankara is gathered in the center of the city but in Istanbul venues are distributed almost equally among all the boroughs. There are lots of cafes and restaurants in two cities.

Conclusion

There are big differences in terms of the distribution of the venues between two cities as expected. These analysis should be expanded with the neighborhood coordinates. Investors and people who think about working in these cities can benefit from these kind of analysis.