CAPSTONE PROJECT REPORT

1. INTRODUCTION

Istanbul, Turkey where I am living in is one of the most crowded cities in the world. Thus, traffic is the biggest issue for the people in this city. On the other side Istanbul is famous with social opportunities. There are hundreds of places in every corner of the city. However, people who want to go outside think twice before deciding where they go because they know they will face with car parking. Even there are so many car parking areas belong Istanbul Metropolitan Municipality, they are usually full. In my project, I want to show there is any relation with these car parking areas and the popularity of venues in Istanbul.

2. DATA

In my project I used three different data sources to get a result.

First, I need the location data of the car parks in Istanbul. For this reason, I used api that was developed by Istanbul Metropolitan Municipality. This data contains id, the name of parks (ParkAdi), borough (Ilce), latitude, longitude, capacity (Kapasitesi), free capacity (BosKapasite), type of park (ParkTipi), distance and free parking time in minutes (UcretsizParklanmaDK).

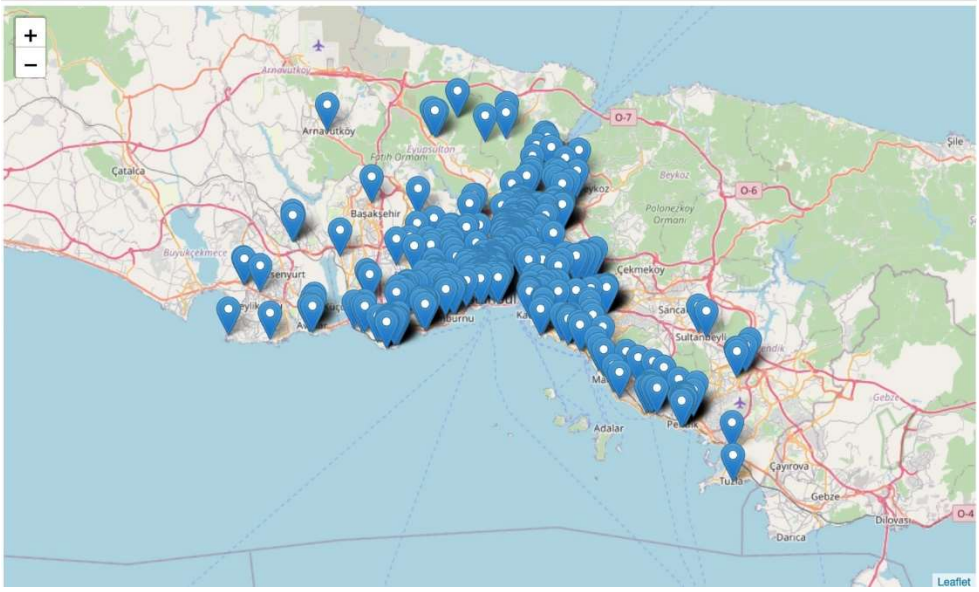
Second one that I need is boroughs and their neighbourhoods in Istanbul. I found this data on one of repository and it is free to use. The data is in an excel file with many sheets, so I create a new file to use in my work. In the end, I created a file with borough (ILCE), neighbourhood (SEMT_ADI_BUYUK) and postal code (POSTA_KODU). However, some rows have NaN cell, so I cleaned the data. Also, I changed the type of POSTA_KODU column from float to int.

I got my third data set via Foursquare API for the details of venues. For this purpose, I created developer account on Foursquare.

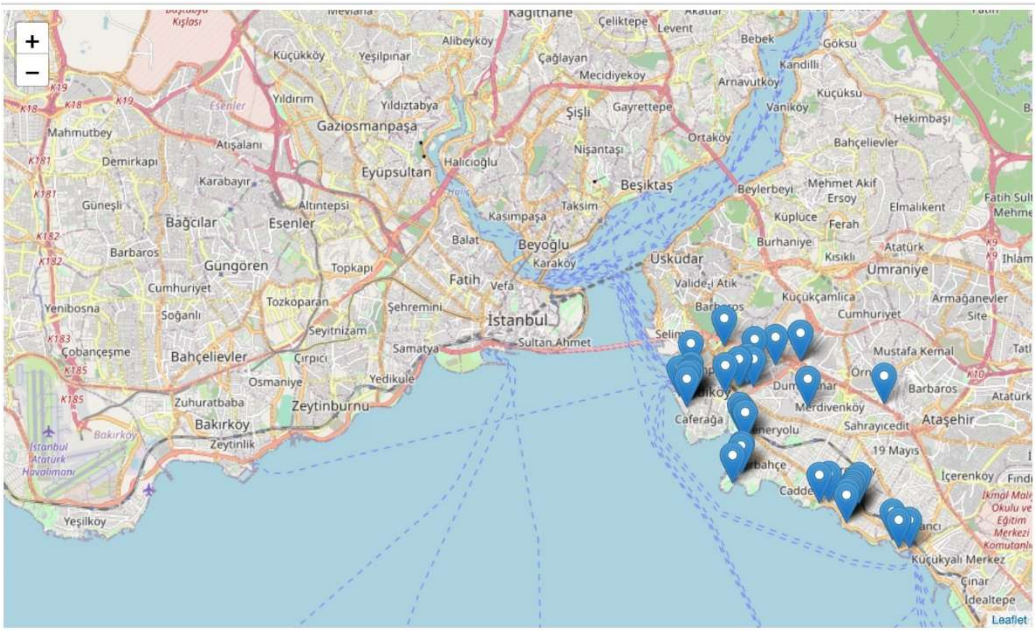
3. METHODOLOGY

In my project I used a simple K-Mean Clustering. Before getting the result, I had to clean, prepare data in the correct form. Thus, I remove the null data and correct the latitude and longitude data for some neighbourhoods.

After that, I first showed the parking in Istanbul because I wanted to see the distribution of the parking areas in the city.

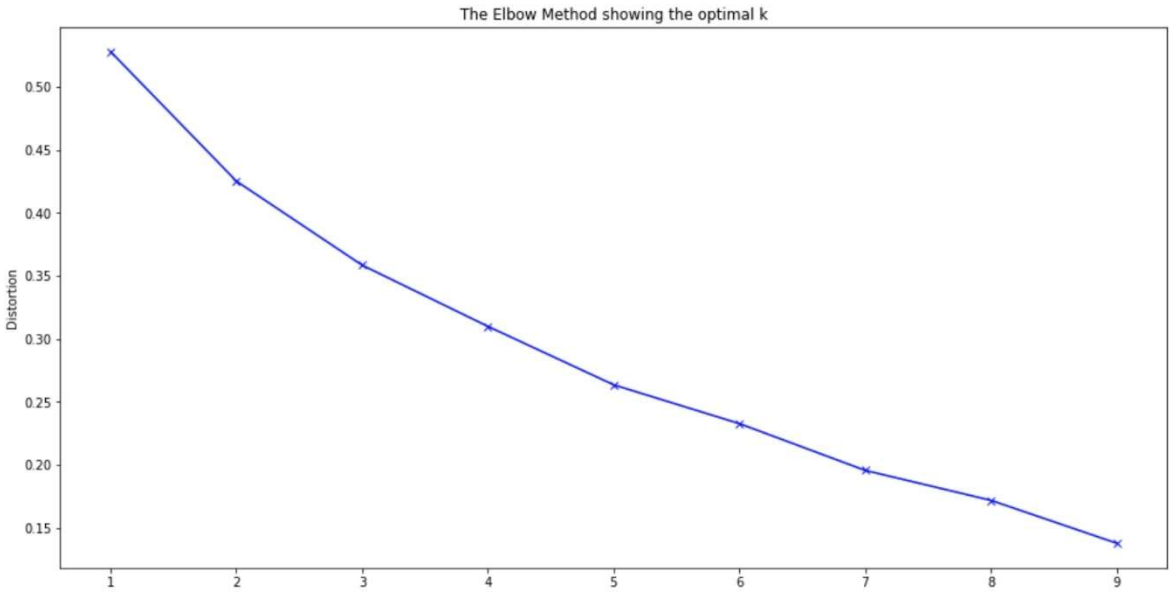


Then I chose the brough Kadikoy for examination because it is one of the most crowded town and there are so many popular places, and I created a similar map for Kadikoy to show the car parking area in this town.

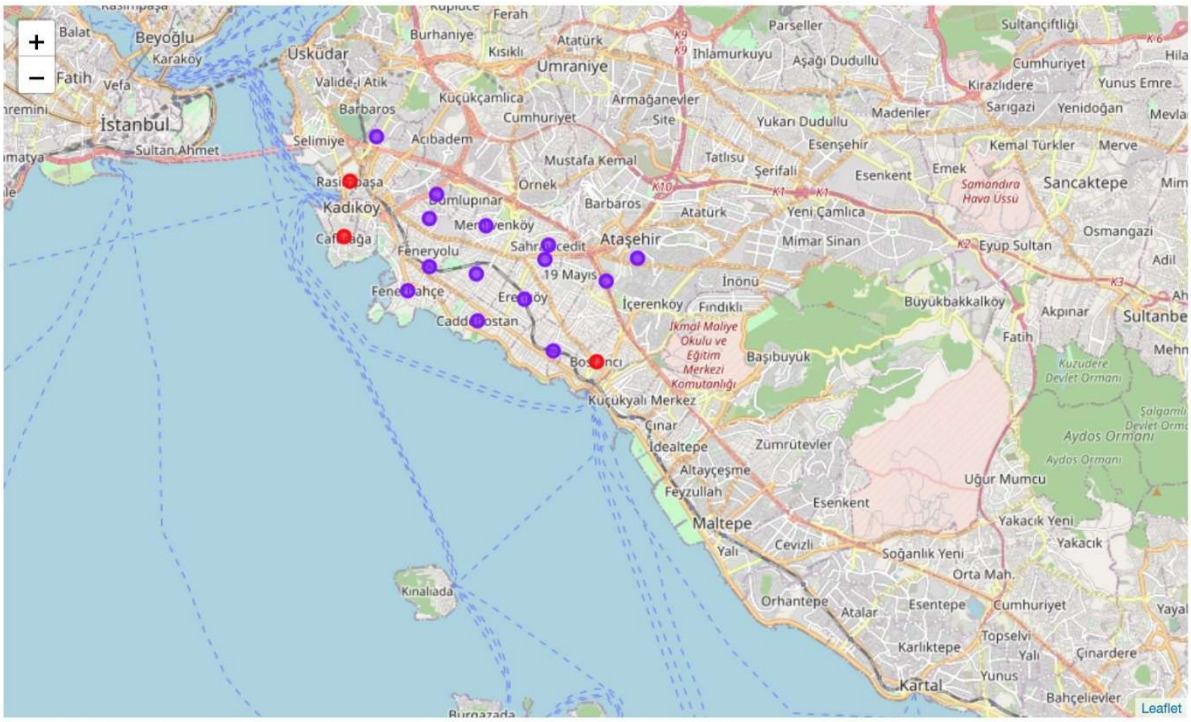


As seen in Kadikoy map, some neighbourhood have more car parking areas than the other ones so there is no homogeneous distribution.

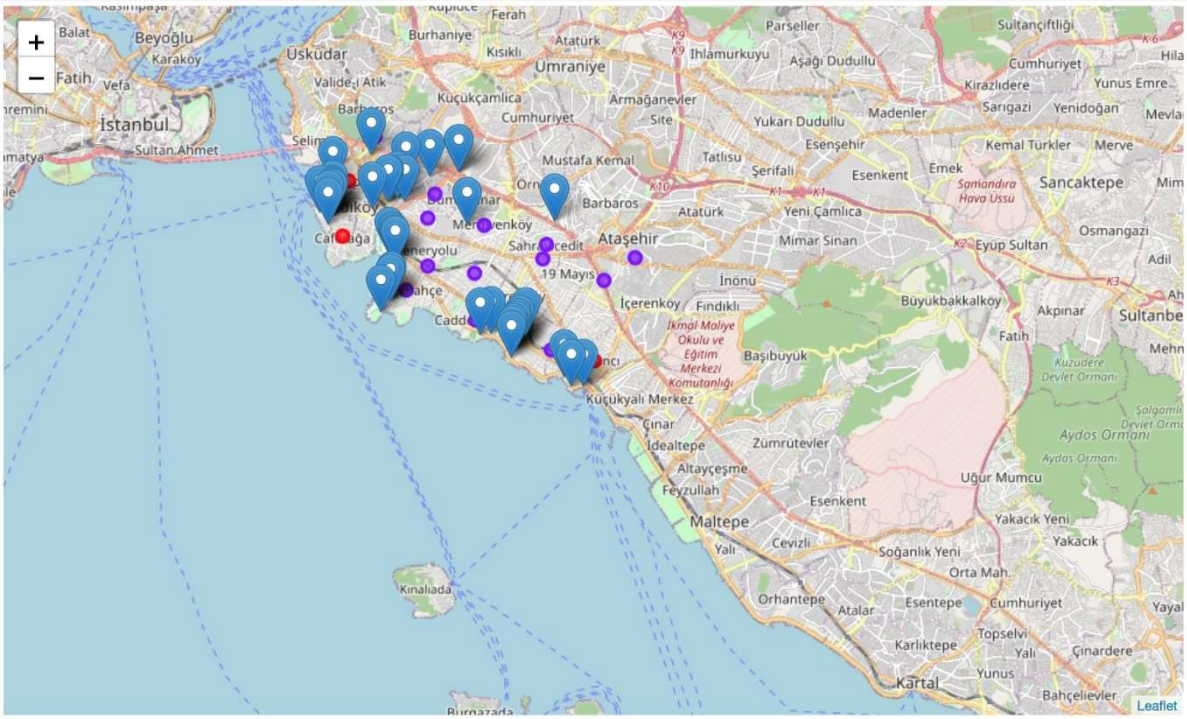
For examining my data, I used K-Means clustering but I used the elbow method to find the optimal k. Actually, I didn't notice a meaningful elbow joint on the graphic but 2 looks like more optimal than the others.



After that I ran K-Mean clustering, and I got a result as below.



To observe the relation between car parking area and the neighbourhood, I added the car parking area on the map.



4. RESULT

Regarding to my observation, the popularity of the venues has no relation with the car parking area. Even though some neighbourhoods have almost no car parking area but they are still in the same group with the other neighbourhoods have many car parking areas.

5. DISCUSSION

Even if people don't want to go outside because of limited parking area why are the popularities of the neighbourhoods are similar?

6. CONCLUSION

Istanbul is one of the cosmopolitan cities around the world. Even traffic is a kind of nightmare, people still don't give up going outside because the life is more stressful than the other cities in Turkey.

However, this problem shows us that we still examine why these places are popular despite to the traffic problem.