Optimal Place for Opening New Restaurant in

Istanbul, Turkey

Introduction

Hey everyone! In this project, I am trying find the optimal place for opening a restaurant in Istanbul. Istanbul is the most populous city in Turkey and the country's economic, cultural and historic center [1].

Business Problem

I assume that someone is looking to open a restaurant and trying to find the best place in Istanbul. There are many restaurants and other venues in 39 districts of the city. Finding districts that include fewer restaurants may help to find true place.

Data Description

I have used three data sources for my application. Firstly, scraped a table from "List of districts of Istanbul" page on Wikipedia [2]. This dataset was not include latitude and longitude values of districts. Therefore, I googled coordinates of districts and found a website, in Turkish [3]. After merging this datasets, my final dataset created.

And of course, I have used Foursquare API to find venues in Istanbul.

Methodology

There are many parameters can influence the decision. I think, the most important feature to decide we should open or not new restaurant in spesific distinct, finding how many restaurants there are. Because if there are many restaurants in a distinct, opening a new restaurant requires competitive processes. And competition requires more many, effort and time.

Therefore I have used venues dataset and clustered distincts using KMeans Algorithm. To find an optimal "k" value, iteratively processed the algorithm and used Elbow method. Even it did not give obvious result, I think used a good hyperparameter.

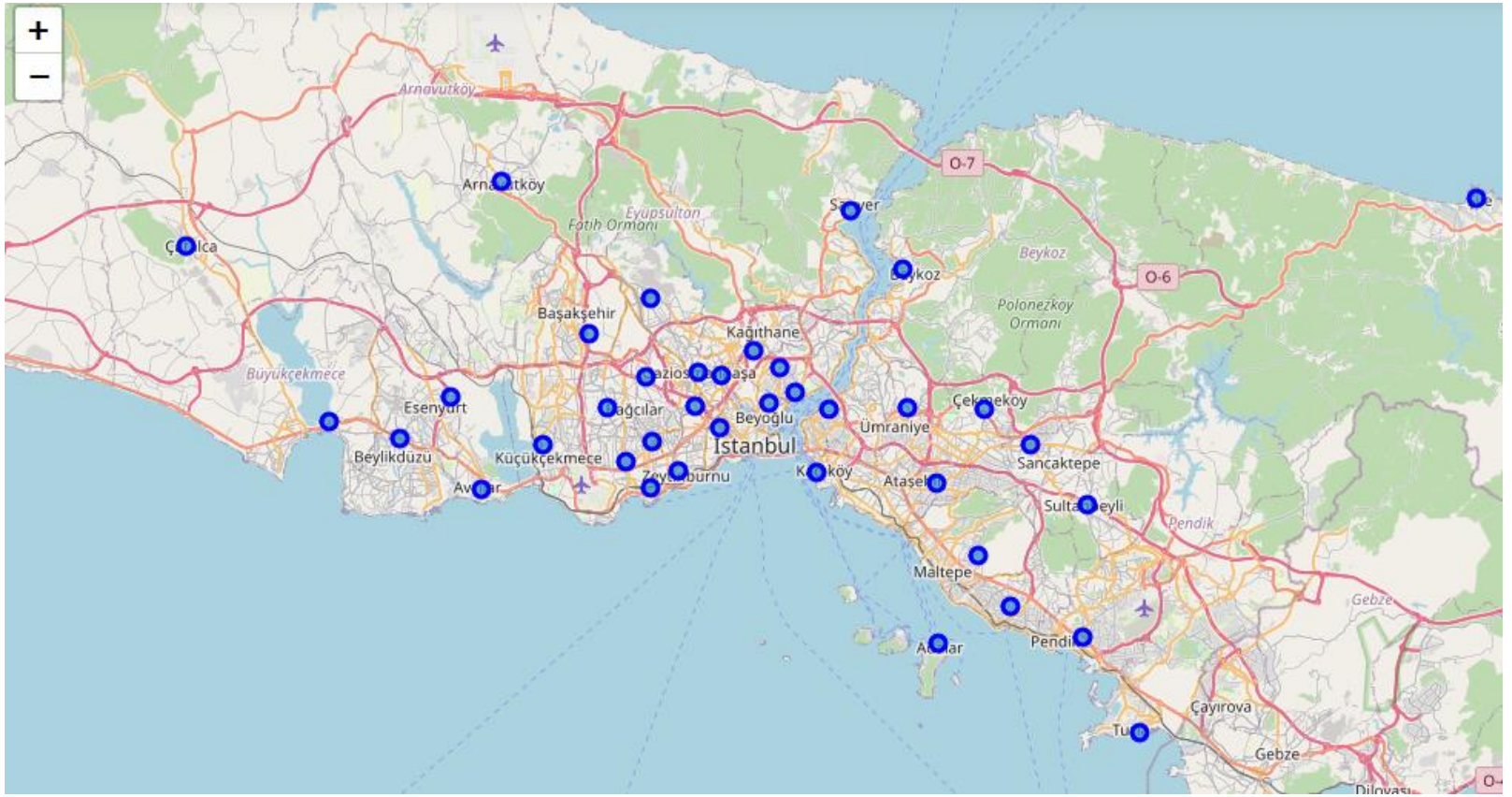
After clustering the distincts, I have visualised clustered distincts in map using Folium. I also added "Annual household income TL(USD)" information to interpret results and make robust decision.

Application

Here is the scraped dataset:

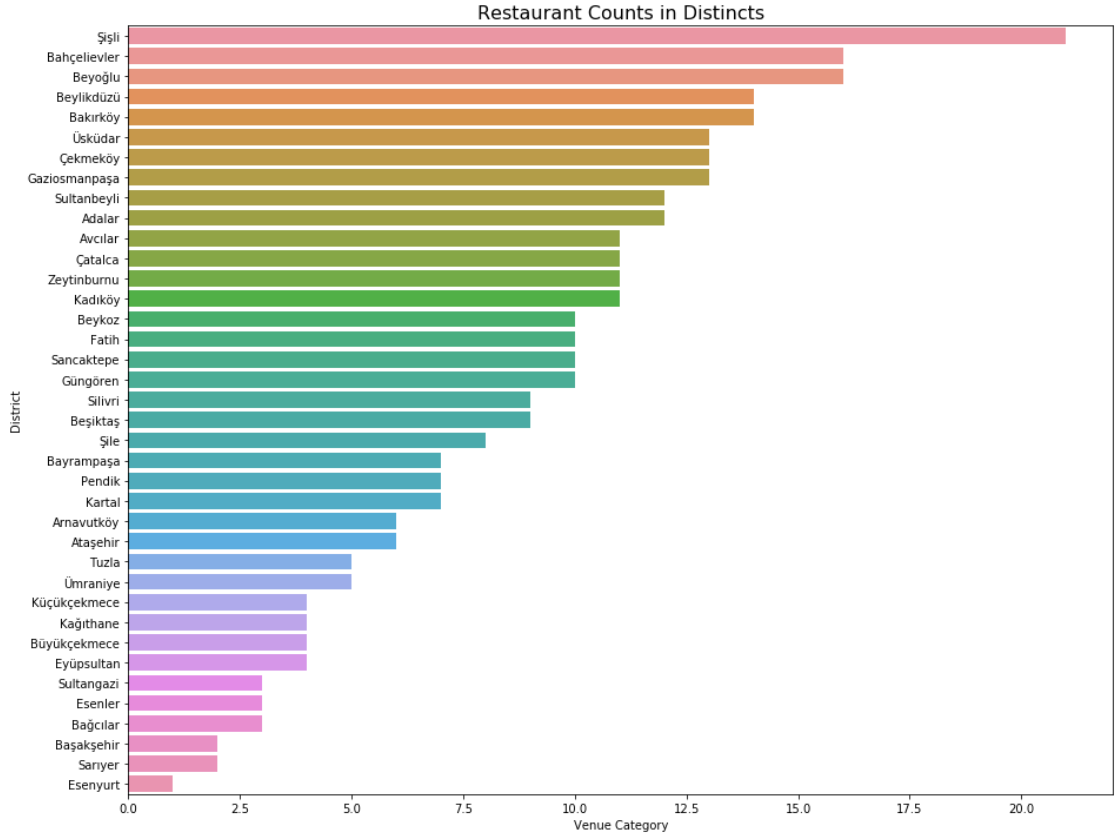
	District	Population (2019)	Area (km²)	Density (per km²)	Mensual household income TL(USD)	Annual household income TL(USD)	Latitude	Longitude
0	Adalar	15238	11.05	1379	6.652₺ (918$)	79.821₺ (10,978$)	40.8747	29.1294
1	Arnavutköy	282488	450.35	627	2.030₺ (279$)	24.360₺ (3,350$)	41.1864	28.7389
2	Ataşehir	425094	25.23	16849	6.577₺ (904$)	78.924₺ (10,854$)	40.9833	29.1278
3	Avcılar	448882	42.01	10685	3.662₺ (503$)	43.938₺ (6,064$)	40.9792	28.7214
4	Bağcılar	745125	22.36	33324	3.197₺ (441$)	38.367₺ (5,295$)	41.0341	28.8330

Application (continued)



Application (continued)

Restaurant counts of distincts gives a general intuition.



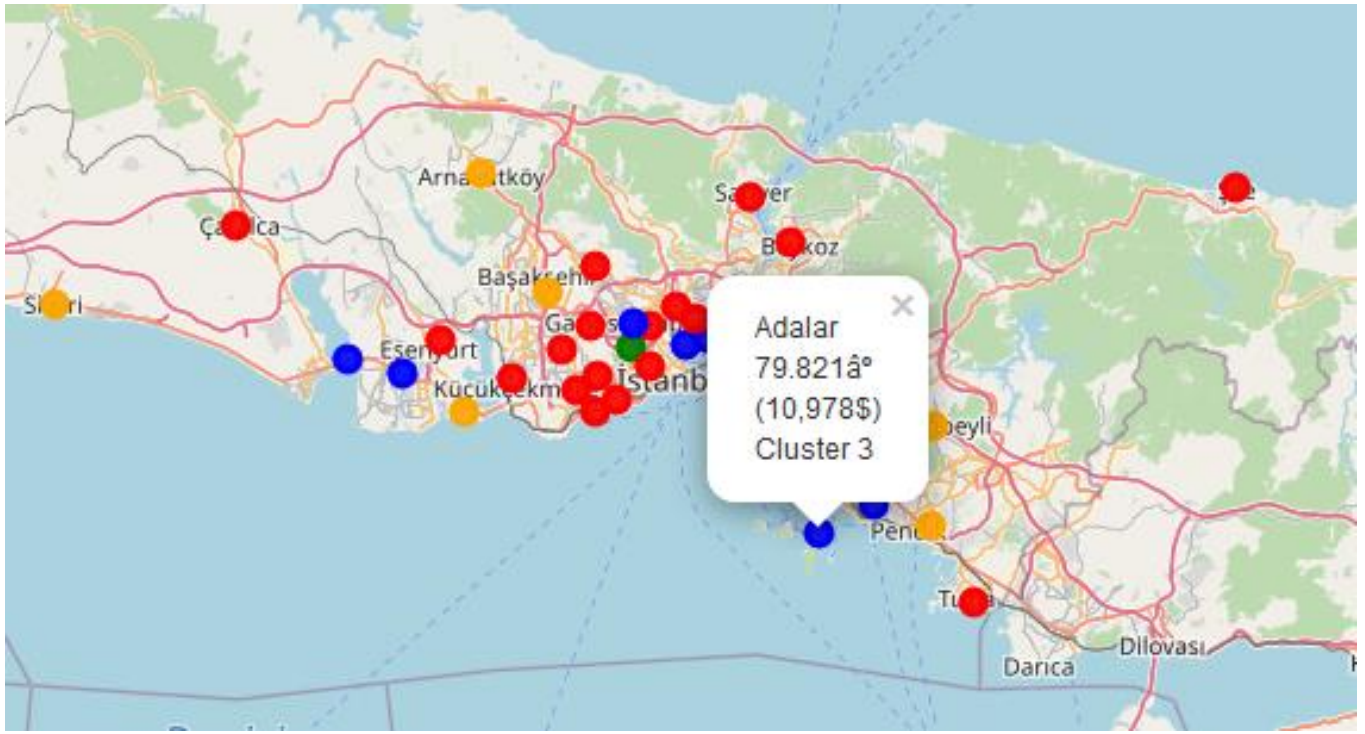
Application (continued)

Distincts in Cluster 3 seems have less restaurants, you can see below:

Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
0	Restaurant	Toll Plaza	Business Service	Gastropub
1	Café	Turkish Restaurant	Dessert Shop	Turkish Restaurant
2	Café	Turkish Restaurant	Dessert Shop	Park
3	Café	Café	Breakfast Spot	Turkish Home Cooking Restaurant

Application (continued)

Cluster 3 (blue) is optimal distinct for opening a new restaurant



Results

I have clustered 39 distincts of Istanbul. When we look the most common venues of clusters:

- Cluster 0: The most common venue is restaurant and this distincts will be a bad choice for opening a new restaurant.
- Cluster 1-2: Even the 1st most common venue is not restaurant, these cluster also includes many restaurants and doesn't seem a good choice.
- Cluster 3: This is the optimal distincts for opening a restaurant as result. In the map above, blue markers show Cluster 3 and their annual household income.

Discussion

As noticed before, finding the optimal place for opening a restaurant have many parameters. In this project's results, I found a general result about finding the optimal distinct by using venues data and unsupervised learning algorithm. It is possible to get better results by using given specific parameters.

Conclusion

This project is just an example of using machine learning algorithm to make decision for real-life problems. Even so, while designing all parts and coding my ideas, I learned many things! It is possible to create nice datasets with scraping, use these datasets to build machine learning models and decision-making.

Thank you so much! :)

Tolgahan Çepel

Sources

[1] https://en.wikipedia.org/wiki/Istanbul
[2] https://en.wikipedia.org/wiki/List_of_districts_of_Istanbul
[3] http://dunya.arztalep.com/tr-TR/istanbul/istanbul-haritasi.aspx