1. Introduction/Business Problem

İzmir is a city on Turkey's Aegean coast. It is the third most populous city in Turkey, where the city had a population of 2,972,900, while İzmir Province had a total population of 4,367,251. Modern İzmir resides in the western extremity of Anatolia and hosts the nearby ancient cities of Ephesus, Pergamon, Sardis and Klazomenai, and centers of international tourism such as Kuşadası, Çeşme, Mordoğan and Foça. So, İzmir is an A tourist attraction is a place of interest where tourists visit, typically for its inherent or an exhibited natural or cultural value, historical significance, natural or built beauty, offering leisure and amusement. For tourists, who prefer vegetarian food, finding a right place to eat can be a daunting challenge, especially in a country like Turkey, where meat heavy dishes like Kebab, Doner are the most popular ones. Turkey is famous for its meat culture and is not quite vegetarian friendly. Luckily, taking advantage of being an Aegean costal city, İzmir has lots of amazing seafood on offer for tourists, and we will discover by studying districts and neighbors applying machine learning tools and try to find best fits for their needs.

The beautiful İzmir

2. Description of the problem

The aim of this project is to explore, cluster, and segment the neighborhoods in the city of İzmir and give recommendations to tourists, immigrants, friends , who are vegetarian and coming from the world to visit İzmir, based on the common restaurant categories in different neighborhoods. This could also partially answer the question 'Where to open restaurants depending on their types in İzmir?' Couple of questions can be answered at the ends of the project. What are the most frequently occurring venues in İzmir Which neighborhood has the higher number of of restaurants depending on its type? Where to find vegetarian-friendly restaurant, sea foods restaurants?

Data

1. We will be using the below datasets for analyzing İzmir. İzmir data, which contains the list of districts, neighborhoods and their latitudes and longitudes.

   I found the data from data source from

   'https://www.atlasbig.com/tr/izmirin-mahalleleri'

   I needed to clean the data and reduced it to city of İzmir and its neighborhoods. I also changed some Turkish words to its English.

   There are 29 districts and 178 neighborhoods in İzmir.

2. We have the coordinates. Foursquare API is used to get the most common venues of given neighborhoods of İzmir. I will use the Foursquare location data to fetch all the venues in each neighborhood by their coordinates. After that, I will refine and group these venues to get only restaurants.

3. Methodology

Collecting data

I found the data from data source from https://www.atlasbig.com/tr/izmirin-mahalleleri

By using beautiful soup, I am able to extract İzmir data, which contains the list of districts, neighborhoods and their latitudes and longitudes.

	Neigbourhood	District	Nüfus	Yüzölçümü (km2)
0	Yalı Mahallesi	Karşıyaka	37.053	1834
1	Erzene Mahallesi	Bornova	36.012	16885
2	Kazımdirik Mahallesi	Bornova	35.318	5004
3	Manavkuyu Mahallesi	Bayraklı	32.068	1287
4	Bostanlı Mahallesi	Karşıyaka	31.738	1489

Data Preparation and Cleansing

Before proceeding with analysis, it is a mandatory step to clean and merge your data and put it into dataframes. Otherwise, the analysis won't be accurate.

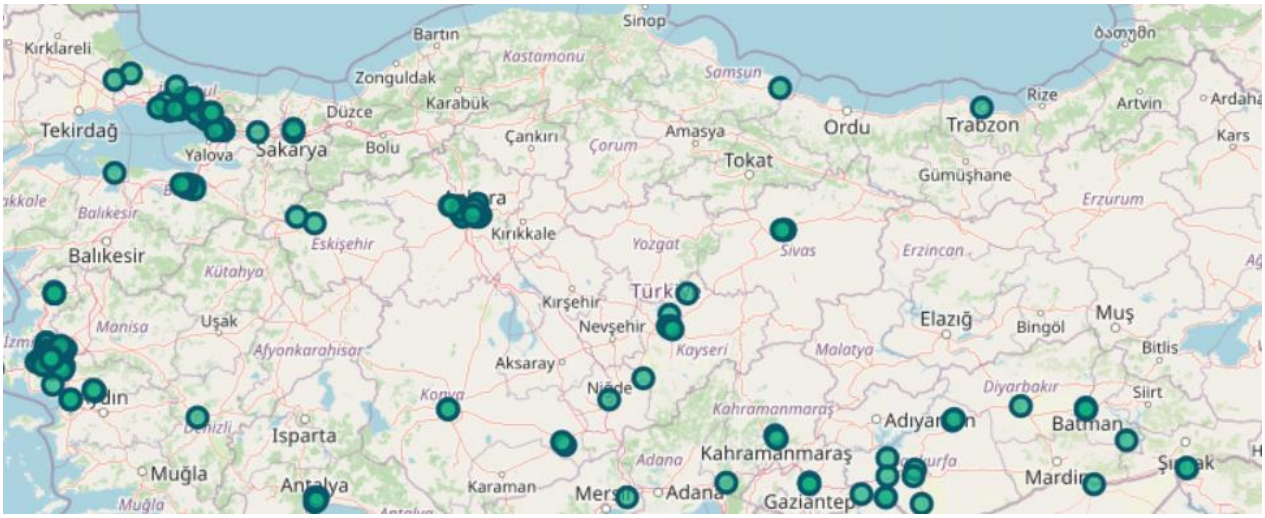
	Neigbourhood	District	Latitude	Longitude
0	Yalı Mahallesi	Karşıyaka	40.895358	29.220435
1	Erzene Mahallesi	Bornova	38.469727	27.229410
2	Kazımdirik Mahallesi	Bornova	38.455146	27.212553
3	Manavkuyu Mahallesi	Bayraklı	38.461078	27.188203
4	Bostanlı Mahallesi	Karşıyaka	38.458185	27.097839

Getting Coordinates of neigbourhoods from **Geopy Client**

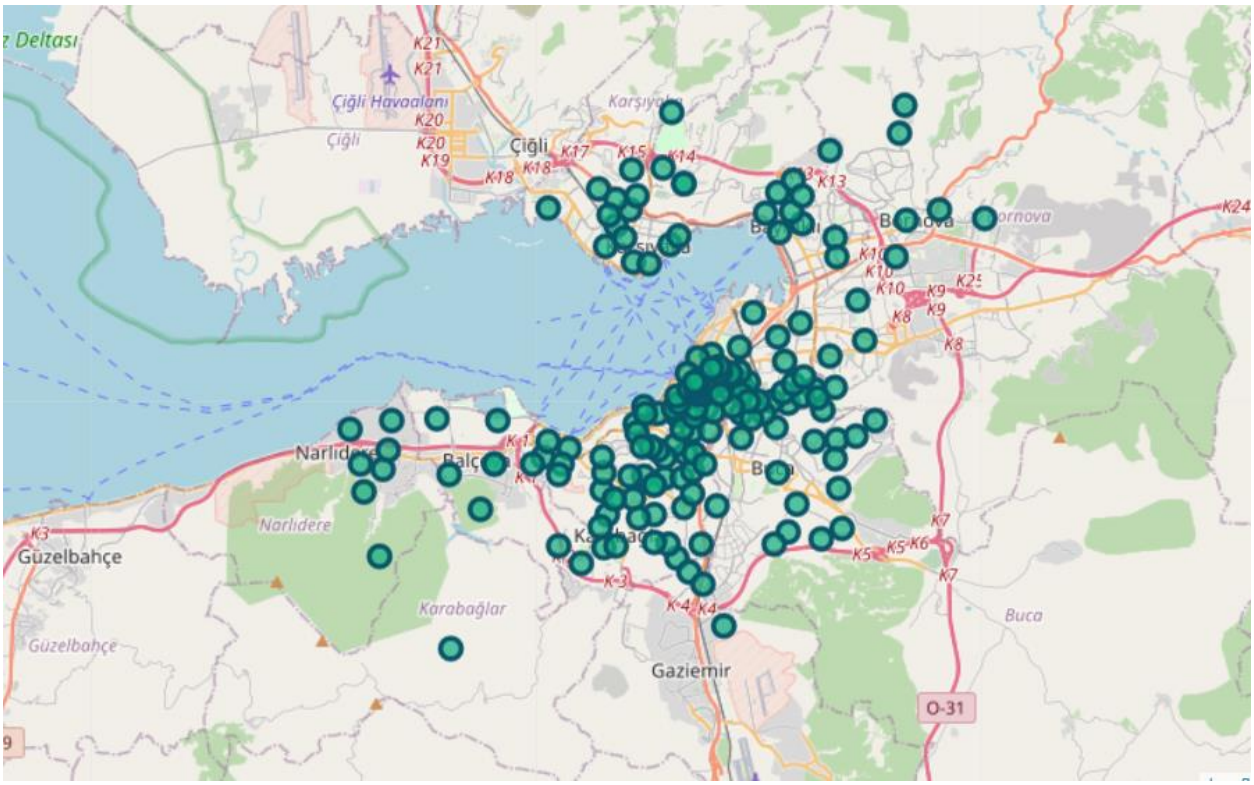
The latitude and longitude values of Izmir are fetched by using geopy library.

Folium library will be used to visualize geographic details of Izmir and its boroughs

I used FourSquare data to find venues and all restaurants. I needed to filter them by their coordinates.



There are places with the same neighborhood and districts name all around Turkey. I have to eliminate those places that are not in Izmir. I am doing it by finding the maximum and minimum Longitude and Latitude values of Izmir. This method will ensure me that I am getting exact locations of Izmir's neighborhoods



Nice! Now we can see all the neighbourhoods of Izmir visualized on the map.

Applying k-means algorithm

To be able to analyze the data needs to be transformed from categorical data into numerical data for Machine Learning algorithms.This is done by appliyng One hot encoding. a data-frame for the venue categories was created.

Then the rows are grouped by by uing groupby function on neighborhood column and By calculating the average of the frequency of occurrence of each venue category.

```
# one hot encoding
izmir_onehot = pd.get_dummies(izmir_venues_only_restaurant[['Venue Category']], prefix="", prefix_sep="")

# add neighborhood column back to dataframe
izmir_onehot['Borough'] = izmir_venues_only_restaurant['Borough']

# move neighborhood column to the first column
col_name="Borough"
first_col = izmir_onehot.pop(col_name)
izmir_onehot.insert(0, col_name, first_col)
izmir_onehot
```

This technique made the data much simplier to analyze.

	Borough	African Restaurant	American Restaurant	Asian Restaurant	Bosnian Restaurant	Cambodian Restaurant	Chinese Restaurant	Comfort Food Restaurant	Doner Restaurant	Eastern European Restaurant	...	Middle Eastern Restaurant	Modern European Restaurant	I Ameri Restau
1	Erzene Mahallesi	0	0	0	0	0	0	0	0	0	...	0	0	
2	Erzene Mahallesi	0	0	0	0	0	0	0	0	0	...	0	0	
3	Erzene Mahallesi	0	0	0	0	0	0	0	0	0	...	0	0	
4	Kazımdirik Mahallesi	0	0	0	0	0	0	0	0	0	...	0	0	
5	Kazımdirik Mahallesi	0	0	0	0	0	0	0	0	0	...	0	0	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	
795	Uğur Mahallesi	0	0	0	0	0	0	0	0	0	...	0	0	
796	Uğur Mahallesi	0	0	0	0	0	0	0	0	0	...	0	0	
797	Uğur Mahallesi	0	0	0	0	0	0	0	0	0	...	0	0	
798	Uğur Mahallesi	0	0	0	0	0	0	0	0	0	...	0	0	
799	Uğur Mahallesi	0	0	0	0	0	0	0	0	0	...	0	0	

'99 rows × 31 columns

We are using prescriptive analytics approach to find a location and VenueCategory to open a restaurant. To cluster the neighbourhoods of Izmir based on the venue categories I used K-Means clustering algorithm.

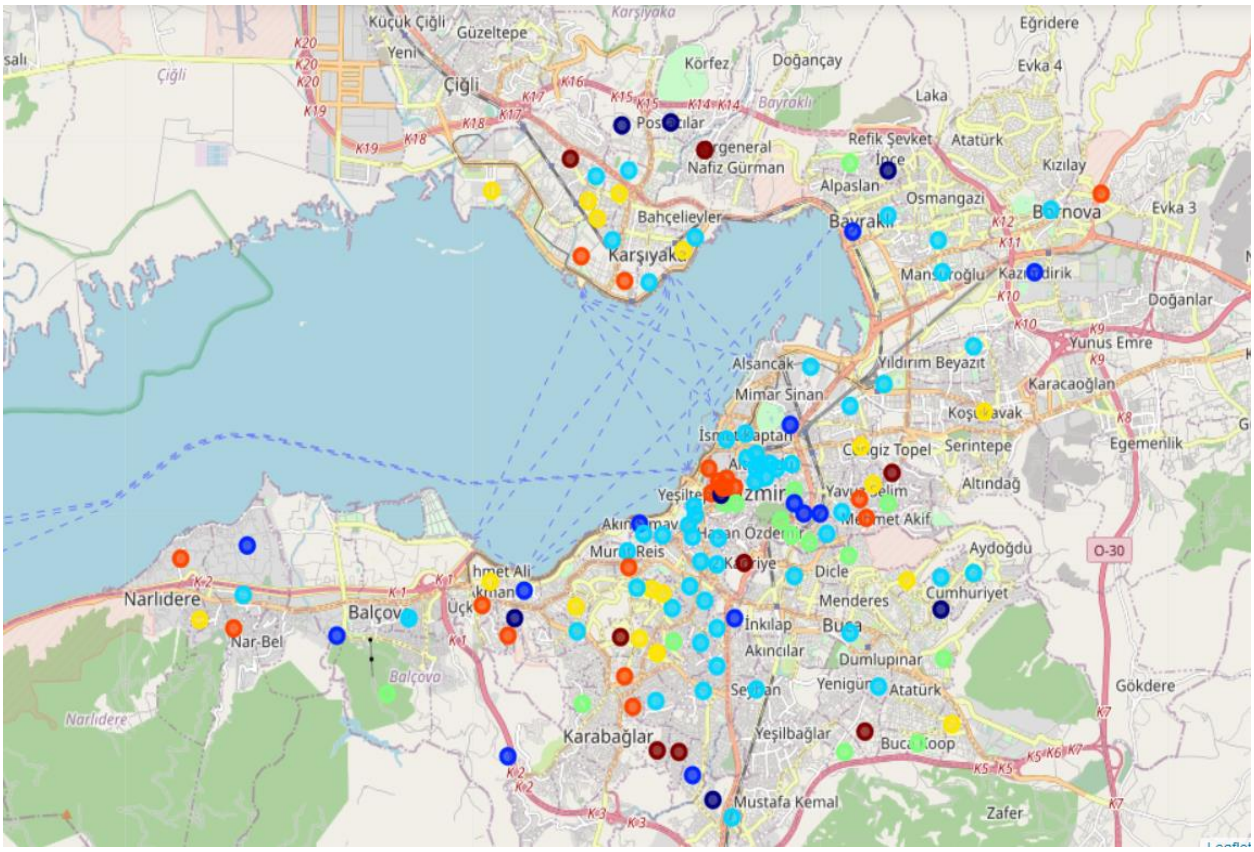
k-means clustering is a method of unsupervised learning, that aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean.

The elbow method helps us to find the optimal k of k-means. In our case, I will run K-Means to cluster the neighbourboroughs into 7 clusters.

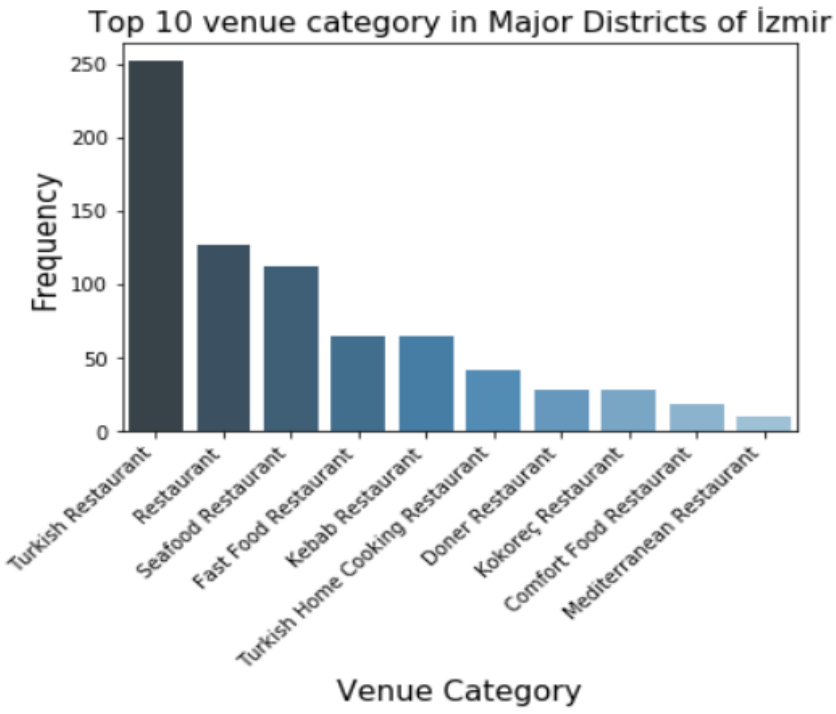
	Cluster Labels	Borough	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	0	2. Kadriye Mahallesi	Kebab Restaurant	Turkish Restaurant	American Restaurant	Asian Restaurant	Bosnian Restaurant	Chinese Restaurant	Comfort Food Restaurant	Doner Restaurant	Eastern European Restaurant	Empanada Restaurant
1	6	26 Ağustos Mahallesi	Kebab Restaurant	Seafood Restaurant	Turkish Restaurant	Italian Restaurant	American Restaurant	Asian Restaurant	Bosnian Restaurant	Chinese Restaurant	Comfort Food Restaurant	Doner Restaurant
2	0	Abdi İpekçi Mahallesi	Kebab Restaurant	Turkish Restaurant	American Restaurant	Asian Restaurant	Bosnian Restaurant	Chinese Restaurant	Comfort Food Restaurant	Doner Restaurant	Eastern European Restaurant	Empanada Restaurant
3	4	Adatepe Mahallesi	Turkish Home Cooking Restaurant	Vegetarian / Vegan Restaurant	Italian Restaurant	American Restaurant	Asian Restaurant	Bosnian Restaurant	Chinese Restaurant	Comfort Food Restaurant	Doner Restaurant	Eastern European Restaurant
4	5	Adnan Süvari Mahallesi	Kebab Restaurant	Restaurant	Fast Food Restaurant	Turkish Restaurant	Italian Restaurant	American Restaurant	Asian Restaurant	Bosnian Restaurant	Chinese Restaurant	Comfort Food Restaurant
...	...	...	...	...	...	...	...	...	...	...	...	...
131	3	İsmet Paşa Mahallesi	Kebab Restaurant	Turkish Restaurant	Comfort Food Restaurant	American Restaurant	Asian Restaurant	Bosnian Restaurant	Chinese Restaurant	Doner Restaurant	Eastern European Restaurant	Empanada Restaurant
132	3	İzkent Mahallesi	Turkish Restaurant	Seafood Restaurant	Kokoreç Restaurant	Vegetarian / Vegan Restaurant	Halal Restaurant	American Restaurant	Asian Restaurant	Bosnian Restaurant	Chinese Restaurant	Comfort Food Restaurant
133	1	Şehit Nedim Tuğaltay Mahallesi	Turkish Restaurant	Vegetarian / Vegan Restaurant	American Restaurant	Asian Restaurant	Bosnian Restaurant	Chinese Restaurant	Comfort Food Restaurant	Doner Restaurant	Eastern European Restaurant	Empanada Restaurant
134	0	Şemikler Mahallesi	Kebab Restaurant	Kokoreç Restaurant	Italian Restaurant	American Restaurant	Asian Restaurant	Bosnian Restaurant	Chinese Restaurant	Comfort Food Restaurant	Doner Restaurant	Eastern European Restaurant
135	5	Şirinkapı Mahallesi	Restaurant	Fast Food Restaurant	Vegetarian / Vegan Restaurant	Italian Restaurant	American Restaurant	Asian Restaurant	Bosnian Restaurant	Chinese Restaurant	Comfort Food Restaurant	Doner Restaurant

4.Results and Discussion

The aim of this project was to explore, cluster, and segment the neighborhoods in the city of İzmir and give recommendations to tourists, immigrants, friends , who are vegetarian and coming from the world to visit İzmir, based on the common restaurant categories in different neighborhoods. Turkey is famous for its meat culture and is not quite vegetarian friendly but being an Aegean costal city, İzmir has lots of amazing seafood and vegatarian foods on offer for tourists. We discovered districts and neighbors by applying machine learning tools and tried to find best fits for their needs. To achieve the most common venues in each neighborhood, I employed the k-means clustering algorithm and was able to sort the popular neighborhoods by their types restaurants and with most Vegan friendly restaurants.



Visulaization of Izmir's restaurants by clusters



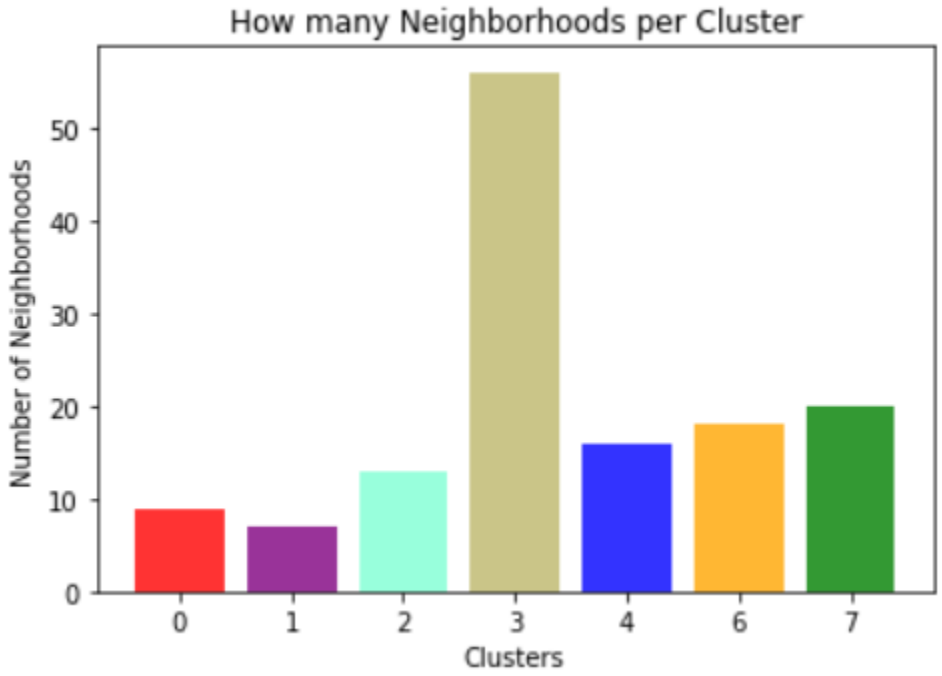
Obiviously, Turkish restaurants are the most popular one alongside with Kebab, Doner, Kokorec. Since Izmir is a coastal Aegean city, Seafood restaurants are popular too. Some of the seafood restaurants might be vegetarian or vegan restaurants.

Here are the some interesting insights which might be useful to travelers as well as people with business interests.

· Turkish restaurants top the charts of most common venues all districts following by Seafood restaurants

· Konak , Karşıyaka, and Karabağlar have the maximum number of restaurants.

· Cluster6 is dominated by Seafood restaurants and Vegetarian / Vegan Restaurants. So if you are a Vegetarian, these are the neighbourhoods that you want to visit. I recommend going somewhere close to the sea side like Bostanlı Karşıyaka 😊



cluster 3 which is the light blue circle on the map has the most neighborhoods. This is due to fact that this particular area is the most popular place to visit for locals in Izmir

	Neigbourhood	District	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
0	Erzene Mahallesi	Bornova	38.469727	27.229410	6.0	Seafood Restaurant	Restaurant	Turkish Restaurant	Vegetarian / Vegan Restaurant	Halal Restaurant	American Restaurant	Asian Restaurant	Bosnian Restaurant
3	Bostanlı Mahallesi	Karşıyaka	38.458185	27.097839	6.0	Seafood Restaurant	Restaurant	Vegetarian / Vegan Restaurant	Turkish Restaurant	Kumru Restaurant	Kebab Restaurant	Turkish Home Cooking Restaurant	New American Restaurant
26	Aksoy Mahallesi	Karşıyaka	38.453675	27.108870	6.0	Turkish Restaurant	Turkish Home Cooking Restaurant	Seafood Restaurant	Vegetarian / Vegan Restaurant	Halal Restaurant	American Restaurant	Asian Restaurant	Bosnian Restaurant
32	Ali Fuat Cebesoy Mahallesi	Karabağlar	38.380457	27.108602	6.0	Seafood Restaurant	Kebab Restaurant	Doner Restaurant	Italian Restaurant	American Restaurant	Asian Restaurant	Bosnian Restaurant	Chinese Restaurant
44	Arap Hasan Mahallesi	Karabağlar	38.400710	27.109956	6.0	Seafood Restaurant	Restaurant	Turkish Restaurant	Turkish Home Cooking Restaurant	Middle Eastern Restaurant	Fast Food Restaurant	Vegetarian / Vegan Restaurant	Greek Restaurant
46	Üçkuyular Mahallesi	Karabağlar	38.393574	27.072844	6.0	Seafood Restaurant	Turkish Restaurant	Restaurant	Vegetarian / Vegan Restaurant	Halal Restaurant	American Restaurant	Asian Restaurant	Bosnian Restaurant
50	Ulubatlı Mahallesi	Konak	38.409803	27.169766	6.0	Turkish Home Cooking Restaurant	Seafood Restaurant	Turkish Restaurant	Vegetarian / Vegan Restaurant	Halal Restaurant	American Restaurant	Asian Restaurant	Bosnian Restaurant
71	Çamtepe Mahallesi	Narlıdere	38.389342	27.009759	6.0	Seafood Restaurant	Turkish Restaurant	Vegetarian / Vegan Restaurant	Italian Restaurant	American Restaurant	Asian Restaurant	Bosnian Restaurant	Chinese Restaurant

Cluster6 is dominated by Seafood restaurants and Vegetarian / Vegan Restaurants. So if you are a Vegetarian, these are the neighbourhood you want to visit.

It is highly possible that the accuracy of our analysis has been affected by following limitations

A significant number of restaurants in Foursquare are labeled as “Restaurant”

instead of a specific restaurant category, we got a large number of restaurants labeled as “Restaurant” from Foursquare

The project can be improved by getting more accurate information and by appliyng distinctive ML techniques

5. Conclusion

As a result, we have explored Izmir’s neighborhoods and its restaurants and and later categorized them into different boroughs, clusters, and segment the neighborhoods in the city based on the common restaurant categories in different neighborhoods. We further shortlisted the neighborhoods based on the common venues and decied a neighborhood which best suits to our problem.