

Data Analysis - Trends of Sports Market in Istanbul and Predicting the Next Trend

Date: 31.07.2020

Prepared by: Resat Caner Bas

1 TABLE OF CONTENTS

| | | |
|-------|--|----|
| 2 | Introduction | 2 |
| 2.1 | The Problem | 2 |
| 2.2 | The Interest | 2 |
| 2.3 | Data Collection | 2 |
| 2.3.1 | Database Details | 4 |
| 3 | Exploring the Data..... | 5 |
| 3.1 | The Distribution Graphic of Venues over Time | 6 |
| 3.2 | Distribution Graphics of each Venue Type over Time | 9 |
| 3.3 | Venue Rating Distribution Graphics..... | 10 |
| 3.4 | Half Time Information | 11 |
| 4 | Conclusion | 12 |
| 4.1 | Clustering | 12 |
| 4.2 | Decision Tree..... | 13 |

2 INTRODUCTION

Individual sports become more and more popular all over the globe due to raise of the self-awareness and well-being philosophy is spreading among people.

2.1 THE PROBLEM

As a result, many disciplines of sports become more and more popular. However, not all of them become more popular and almost none of them can keep itself as the most preferred one.

In this report, you will find an analysis of sports market in Istanbul, more specifically, Kadikoy and Besiktas Towns and an prediction for an upcoming trend if possible.

2.2 THE INTEREST

The results of this report may give a deeper insight to the investors planning to find a business targeting the sports market around the Istanbul and in other regions.

2.3 DATA COLLECTION

The Data is collected in the data gathering notebook.

Kadikoy and Besiktas are one of the towns in Istanbul and Turkey that is one of the highest cultural and financial levels in Turkey. This makes them a great candidate to inspect the trends as these regions are generally flag carriers for such trends.

- The neighborhoods of these towns are used. The data of this collected from Turkish Postal Office Database

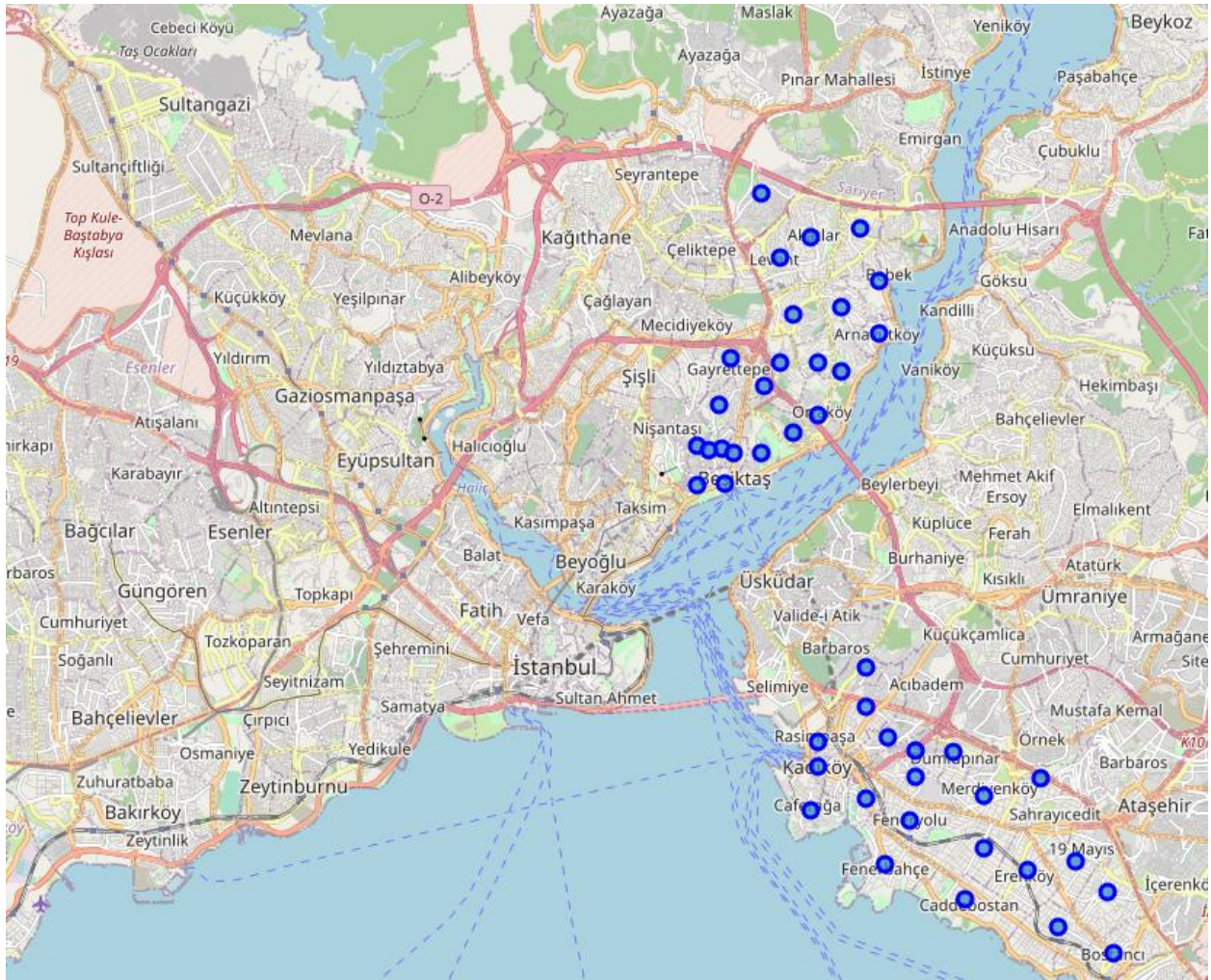


Figure 1-Neighborhood Centers

- Longitude and Latitude data related to the neighborhoods are collected via Google Geocoding API.
- Venue details are collected via **Foursquare Places API**
- Venue details are broken into two databases of Kadikoy neighborhoods and Besiktas Neighborhoods and saved as **kadikoy_venues.csv** and **besiktas_venues.csv**

[7]:

| | Venue Id | Neighborhood | Neighborhood Latitude | Neighborhood Longitude | Venue | Venue Latitude | Venue Longitude | Venue Distance | Venue Category | Venue Url | Venue Twitter | Venue Instagram | Venue Formatted Phone | Venue Rating | Venue Description | Venue Created | Venue Tips Count | Venue Likes Count |
|----|---------------------------|----------------|-----------------------|------------------------|--------------------------------|----------------|-----------------|----------------|----------------------|-----------------------------------|-----------------|-----------------|-----------------------|--------------|----------------------------|---------------|------------------|-------------------|
| 0 | 4b66f145f964a52075312be3 | Cihannuma mah | 41.047211 | 29.007439 | Turkcell Plaza Fitness Center | 41.048224 | 29.008715 | 155 | Gym | NaN | NaN | NaN | NaN | 0.0 | NaN | 1265037637 | 0 | 0 |
| 1 | 4b703459f964a520830bd2de3 | Bebek mah | 41.077744 | 29.041629 | Mars Athletic Club | 41.075919 | 29.038232 | 350 | Gym / Fitness Center | http://www.marsathletic.com | mars_athletic | NaN | +90 212 363 86 96 | 8.8 | NaN | 1265644633 | 52 | 317 |
| 2 | 4b7aed0ef964a52089462fe3 | Etler mah | 41.087042 | 29.037264 | Hillside City Club | 41.087993 | 29.031091 | 528 | Gym / Fitness Center | https://www.hillside.com.tr | hillsideleisure | NaN | +90 212 352 23 33 | 8.9 | NaN | 1266347278 | 129 | 1088 |
| 3 | 4ba67962f964a520265639e3 | Bebek mah | 41.077744 | 29.041629 | Celebrity Fitness | 41.081342 | 29.035980 | 620 | Gym | NaN | NaN | NaN | NaN | 0.0 | NaN | 1269201250 | 0 | 0 |
| 4 | 4bc89f2192b376b0bdcb513a | Kultur mah | 41.073202 | 29.032900 | Maya sports center | 41.076425 | 29.029158 | 476 | Gym / Fitness Center | NaN | NaN | NaN | NaN | 0.0 | NaN | 1271439137 | 1 | 0 |
| 5 | 4be97d30c5220478ee2aaca | Akat mah | 41.085614 | 29.025626 | Club Sporium | 41.086891 | 29.024003 | 196 | Gym / Fitness Center | http://akattar.clubsporium.com.tr | clubsporium | NaN | +90 212 385 49 10 | 8.4 | Sporla dolu yaglam merkezi | 1273593120 | 236 | 1135 |
| 6 | 4c14c74682a3c9e6b0bcdf0f8 | Kultur mah | 41.073202 | 29.032900 | Mayadrom Sports Center(MSC) | 41.077281 | 29.032698 | 454 | Gym | NaN | NaN | NaN | NaN | 6.3 | NaN | 1276430150 | 24 | 12 |
| 7 | 4c37666c93db0f4737b51f92 | Visnezade mah | 41.041503 | 28.998708 | Mars Athletic Club | 41.043765 | 28.992381 | 587 | Gym / Fitness Center | http://www.marsathletic.com | mars_athletic | NaN | +90 212 332 44 40 | 8.6 | NaN | 1278699116 | 151 | 786 |
| 8 | 4c4ae6c79c8d2d7fe10e316a | Muradiye mah | 41.048450 | 28.998708 | İTÜ Havuz | 41.046354 | 28.995922 | 330 | Gym Pool | http://www.itu.edu.tr | itu1773 | NaN | NaN | 7.9 | NaN | 1279977159 | 6 | 37 |
| 9 | 4c51ab6b1c7ef3be50f1bb9 | Cihannuma mah | 41.047211 | 29.007439 | The Aladdin Club | 41.047260 | 29.007811 | 31 | Gym / Fitness Center | NaN | NaN | NaN | +90 212 327 63 93 | 7.3 | NaN | 1280420715 | 6 | 44 |
| 10 | 4c585cd0d12a201a1a2d768bd | Akat mah | 41.085614 | 29.025626 | Şişli Terakki Spor Salonu | 41.081946 | 29.024032 | 429 | Gym | NaN | NaN | NaN | NaN | 7.1 | NaN | 1280859344 | 6 | 57 |
| 11 | 4c6667d6e75ac9280676f7da | Gayrettepe mah | 41.064196 | 29.006711 | Essporto Health & Fitness Club | 41.065585 | 29.010558 | 357 | Gym / Fitness Center | http://www.essporto.com | essporto | NaN | +90 212 344 06 66 | 6.8 | NaN | 1281779670 | 344 | 941 |

Figure 2- Head of besiktas_venues.csv

2.3.1 Database Details

| Column Name | Column Detail |
|------------------------|---|
| Venue Id | ID of venue on Foursquare |
| Neighborhood | Name of the neighborhood |
| Neighborhood Latitude | Latitude of the neighborhood |
| Neighborhood Longitude | Longitude of the neighborhood |
| Venue | Name of the venue |
| Venue Latitude | Latitude of the venue |
| Venue Longitude | Latitude of the venue |
| Venue Distance | The distance in between the venue and the neighborhood center |
| Venue Category | The category of the venue on Foursquare |
| Venue Url | The url of the venue homepage on Foursquare |
| Venue Twitter | The recorded twitter account of the venue on Foursquare |
| Venue Instagram | The recorded instagram account of the venue on Foursquare |
| Venue Formatted Phone | The recorded phone number of the venue on Foursquare |
| Venue Rating | The rating of the venue on Foursquare |
| Venue Description | The description of the venue on Foursquare |
| Venue Created | Venue creation date on Foursquare in Epoch |
| Venue Tips Count | The number of tips of the venue on Foursquare |
| Venue Likes Count | The number of likes of the venue on Foursquare |

3 EXPLORING THE DATA

```
[5]: all_venues.describe(include=['object'])
```

| | Venue Id | Neighborhood | Venue | Venue Category | Venue Uri | Venue Twitter | Venue Formatted Phone | Venue Description | Town |
|--------|--------------------------|-----------------|-------|----------------|-----------------------------|---------------|-----------------------|---|---------|
| count | 1045 | 1045 | 1045 | 1045 | 98 | 65 | 189 | 13 | 1045 |
| unique | 1045 | 44 | 1009 | 13 | 87 | 57 | 168 | 13 | 2 |
| top | 5b156df4b9a5a8002c157540 | Caddebostan mah | Gym | Gym | http://www.fitintime.com.tr | fitintime | +90 216 355 78 43 | Sağlıklı yaşam, zinde bir vücut için "GYM wo... | Kadikoy |
| freq | 1 | 47 | 6 | 431 | 5 | 5 | 4 | 1 | 553 |

Figure 3- Description of Dataframe

According to the description

- 1) Kadikoy is the town with most venues
- 2) Caddebostan is the neighborhood with most venues
- 3) Gym is the most dominant category (41%)

Here is the distribution of categories:

```
[6]:
```

| Venue Type | Count |
|----------------------|-------|
| Gym | 431 |
| Gym / Fitness Center | 315 |
| Yoga Studio | 114 |
| Pilates Studio | 77 |
| Martial Arts Dojo | 51 |
| Gym Pool | 18 |
| Gymnastics Gym | 13 |
| Track | 8 |
| Boxing Gym | 6 |
| Cycle Studio | 5 |
| Outdoor Gym | 3 |
| Weight Loss Center | 3 |
| Climbing Gym | 1 |

Figure 4- The Distribution of Venue Categories

3.1 THE DISTRIBUTION GRAPHIC OF VENUES OVER TIME

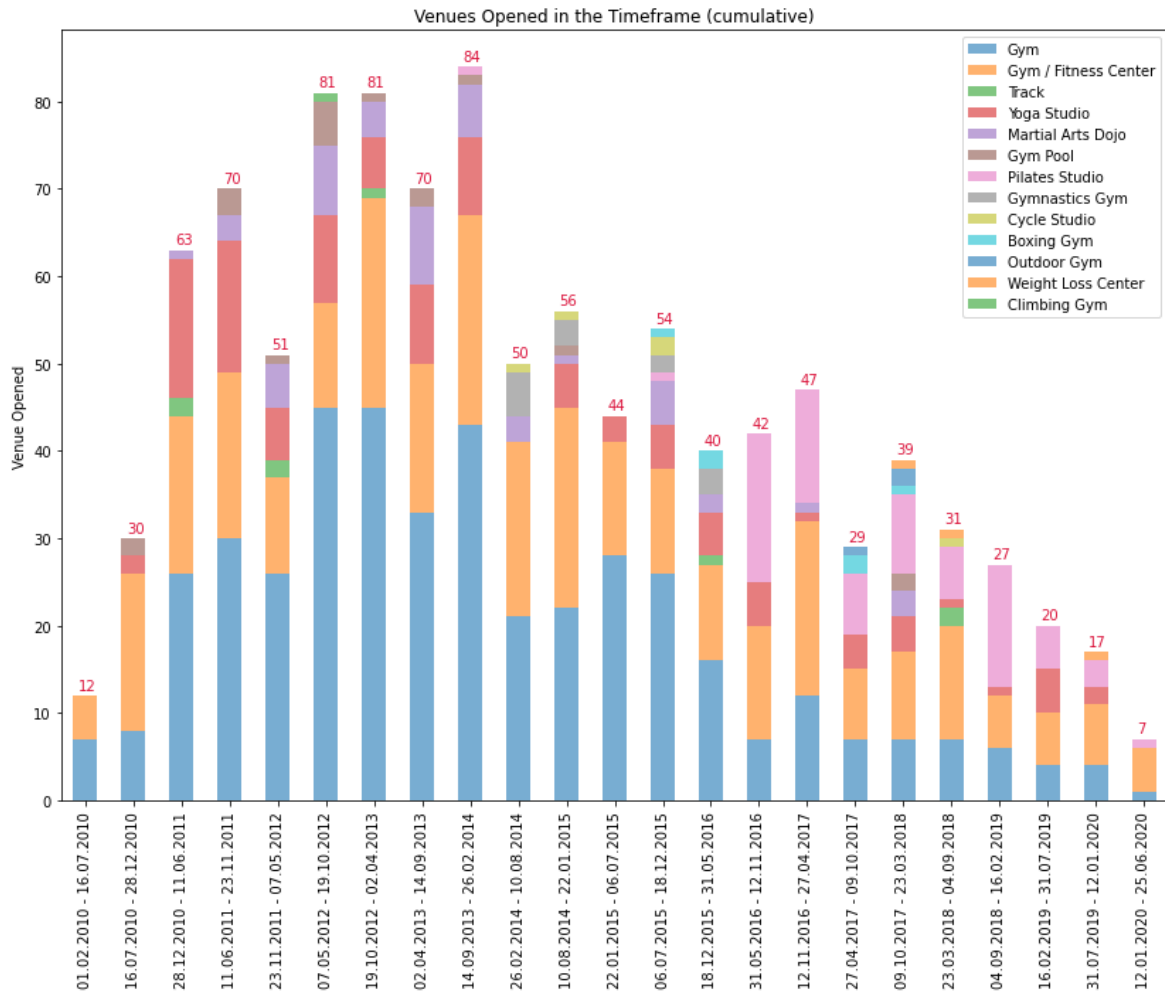


Figure 5- Distribution of Venues over Time in all Neighborhoods

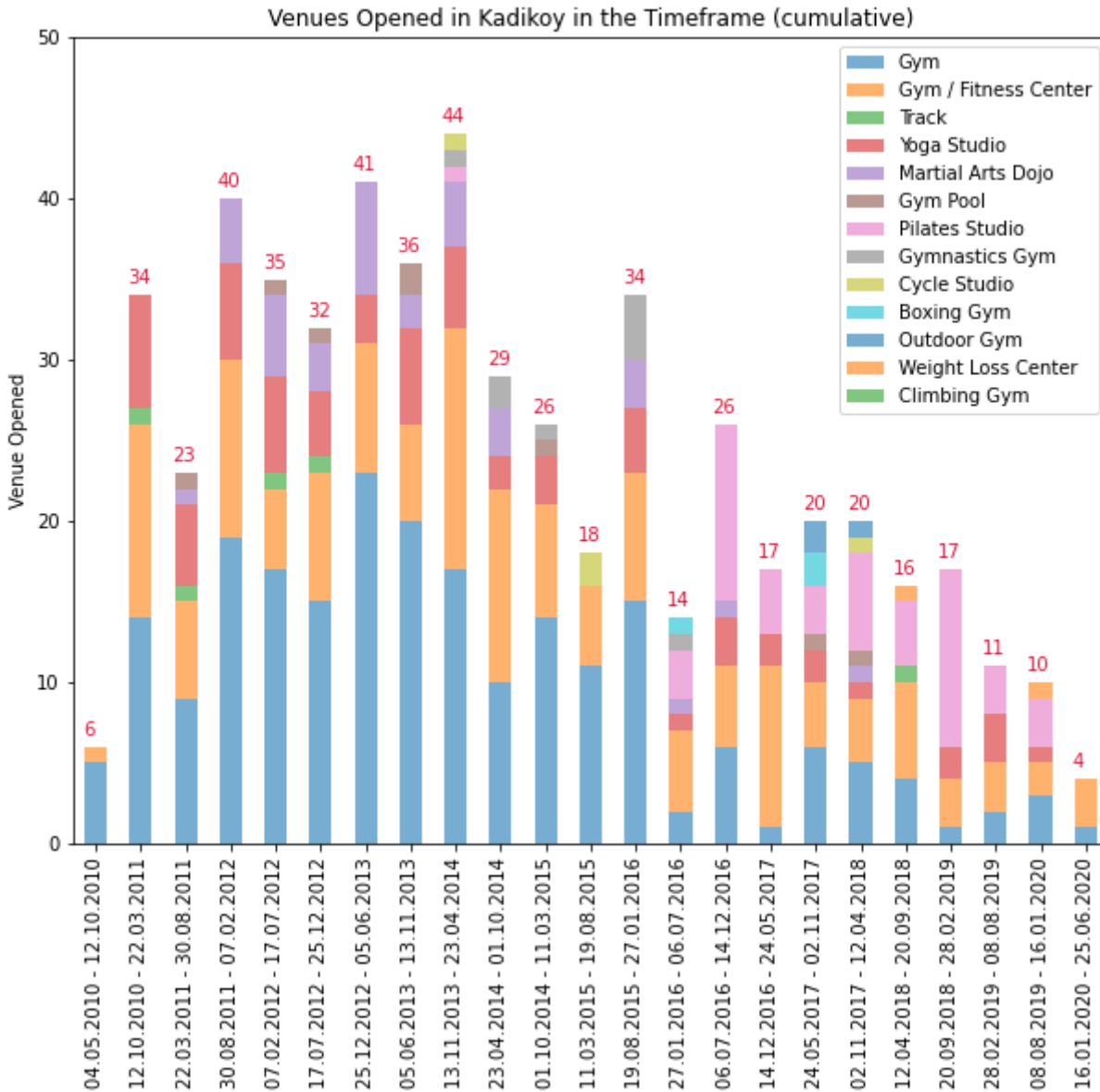


Figure 6- Distribution of Venues over Time in Kadikoy Neighborhoods

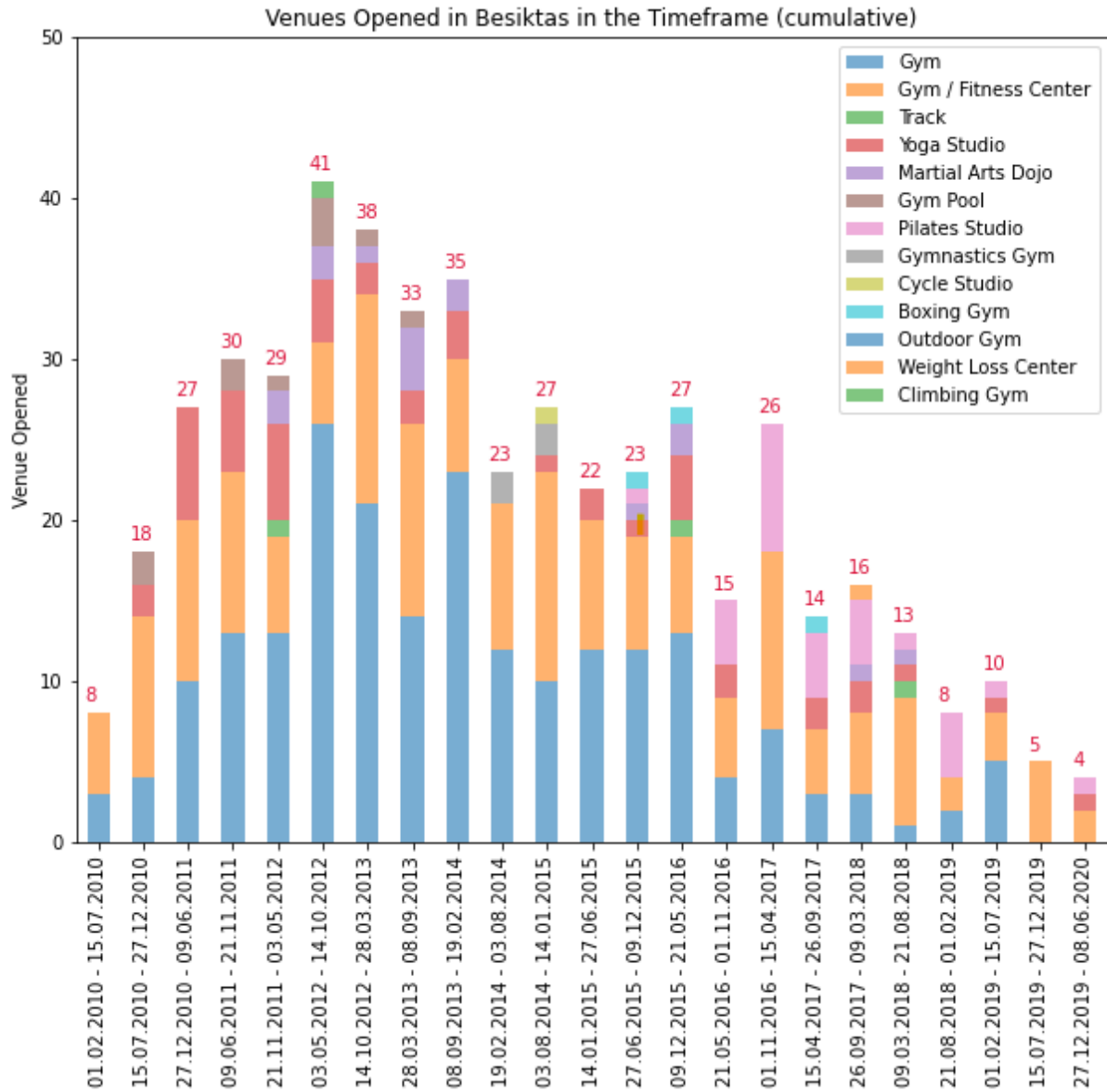


Figure 7- Distribution of Venues over Time in Besiktas Neighborhoods

3.2 DISTRIBUTION GRAPHICS OF EACH VENUE TYPE OVER TIME

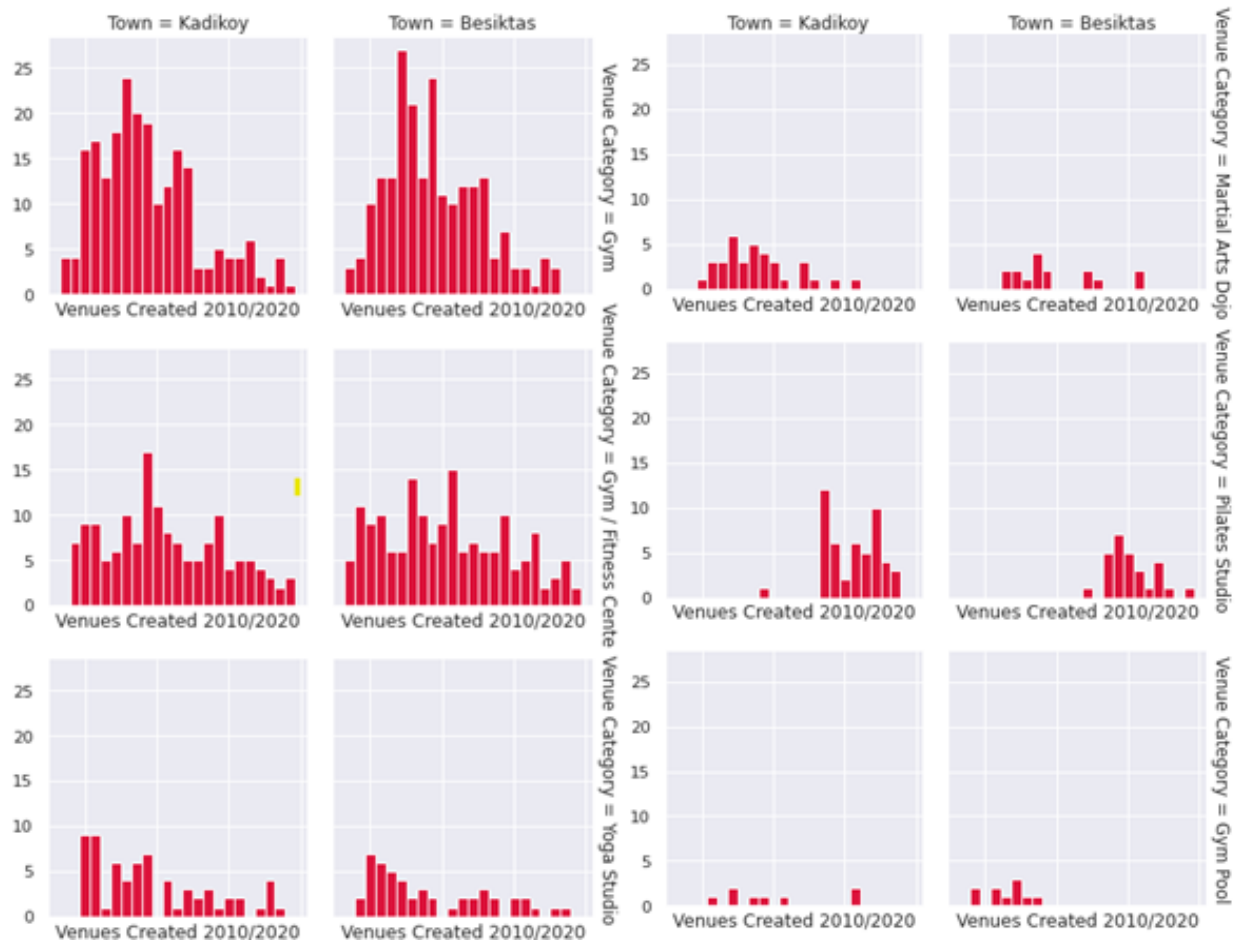


Figure 8- Distribution Graphics of each Venue Type over Time

3.3 VENUE RATING DISTRIBUTION GRAPHICS

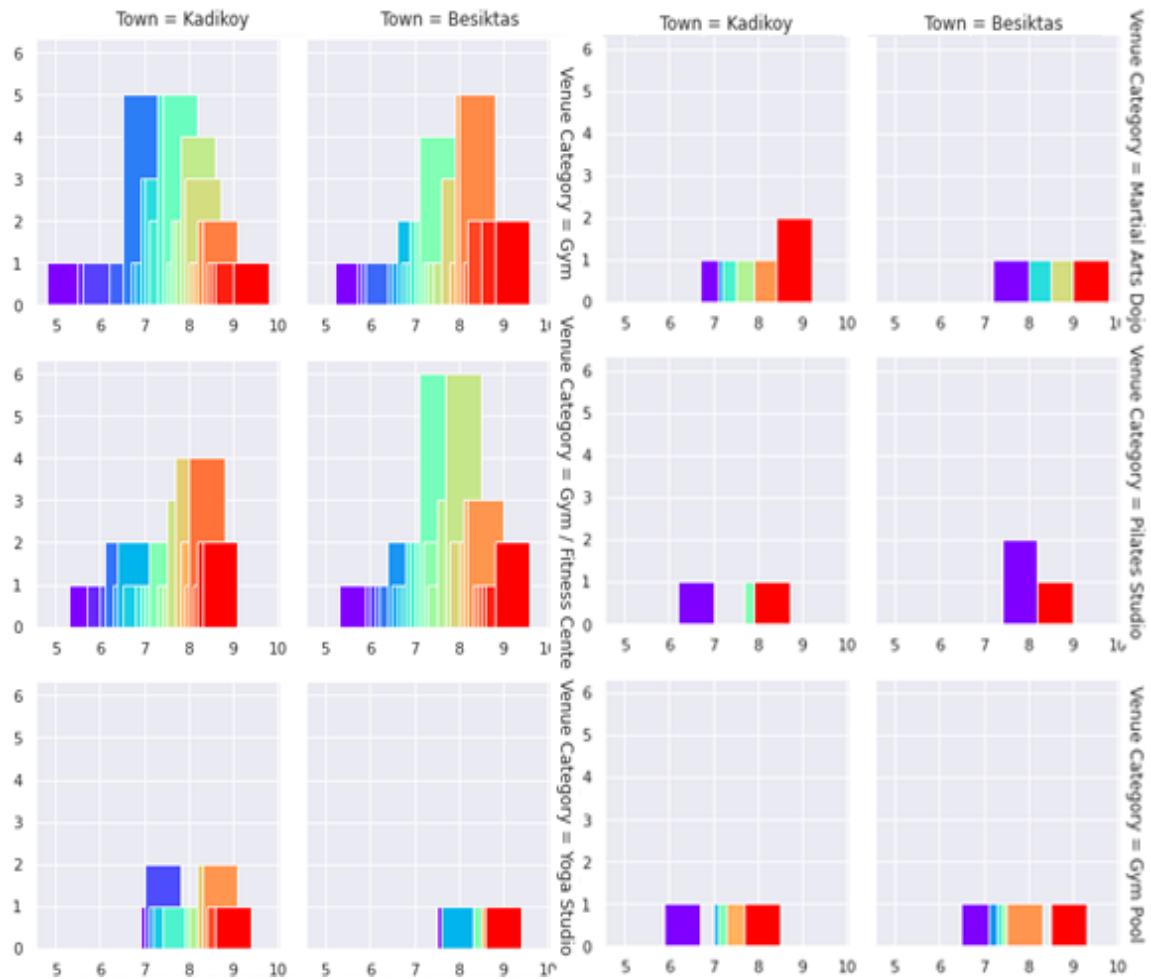


Figure 9- Venue Rating Distribution Graphics

3.4 HALF TIME INFORMATION

```
[16]: l_venues['Venue Category'].unique():
s the total number of {} opened in Kadikoy in between 01.02.2010 and 01.01.2015'.format(len(kadikoy_venue:
s the total number of {} opened in Besiktas in between 01.02.2010 and 01.01.2015'.format(len(besiktas_ven

156 is the total number of Gym opened in Kadikoy in between 01.02.2010 and 01.01.2015
149 is the total number of Gym opened in Besiktas in between 01.02.2010 and 01.01.2015
87 is the total number of Gym / Fitness Center opened in Kadikoy in between 01.02.2010 and 01.01.2015
100 is the total number of Gym / Fitness Center opened in Besiktas in between 01.02.2010 and 01.01.2015
4 is the total number of Track opened in Kadikoy in between 01.02.2010 and 01.01.2015
1 is the total number of Track opened in Besiktas in between 01.02.2010 and 01.01.2015
45 is the total number of Yoga Studio opened in Kadikoy in between 01.02.2010 and 01.01.2015
32 is the total number of Yoga Studio opened in Besiktas in between 01.02.2010 and 01.01.2015
29 is the total number of Martial Arts Dojo opened in Kadikoy in between 01.02.2010 and 01.01.2015
11 is the total number of Martial Arts Dojo opened in Besiktas in between 01.02.2010 and 01.01.2015
6 is the total number of Gym Pool opened in Kadikoy in between 01.02.2010 and 01.01.2015
10 is the total number of Gym Pool opened in Besiktas in between 01.02.2010 and 01.01.2015
1 is the total number of Pilates Studio opened in Kadikoy in between 01.02.2010 and 01.01.2015
0 is the total number of Pilates Studio opened in Besiktas in between 01.02.2010 and 01.01.2015
3 is the total number of Gymnastics Gym opened in Kadikoy in between 01.02.2010 and 01.01.2015
4 is the total number of Gymnastics Gym opened in Besiktas in between 01.02.2010 and 01.01.2015
1 is the total number of Cycle Studio opened in Kadikoy in between 01.02.2010 and 01.01.2015
0 is the total number of Cycle Studio opened in Besiktas in between 01.02.2010 and 01.01.2015
0 is the total number of Boxing Gym opened in Kadikoy in between 01.02.2010 and 01.01.2015
0 is the total number of Boxing Gym opened in Besiktas in between 01.02.2010 and 01.01.2015
0 is the total number of Outdoor Gym opened in Kadikoy in between 01.02.2010 and 01.01.2015
0 is the total number of Outdoor Gym opened in Besiktas in between 01.02.2010 and 01.01.2015
0 is the total number of Weight Loss Center opened in Kadikoy in between 01.02.2010 and 01.01.2015
0 is the total number of Weight Loss Center opened in Besiktas in between 01.02.2010 and 01.01.2015
0 is the total number of Climbing Gym opened in Kadikoy in between 01.02.2010 and 01.01.2015
1 is the total number of Climbing Gym opened in Besiktas in between 01.02.2010 and 01.01.2015
```

Figure 10- Half Time Cumulative Venue Distribution

On the halfway, the numbers look identical, however Kadikoy is leading for **Gym, Track, Track Yoga Studio, Martial Arts Dojo and Pilates Studio** categories whereas Besiktas is leading the **Gym / Fitness Center, Gymnastic Gym and Gym Pool** categories.

4 CONCLUSION

At first glance, Gyms and Gym fitness center seems to dominate all the regions, however it can't be neglected the Yoga Studio, Pilates Studio and Martial Art Dojo venues have strong trends. According to the trends, Kadıköy seems one step ahead for those popular categories. It could be a good advice to follow trends in Kadıköy and apply them in other regions to catch the trend. 2 different ML methods are applied to reach the best results to predict the most profitable investment out of the dataset.

4.1 CLUSTERING

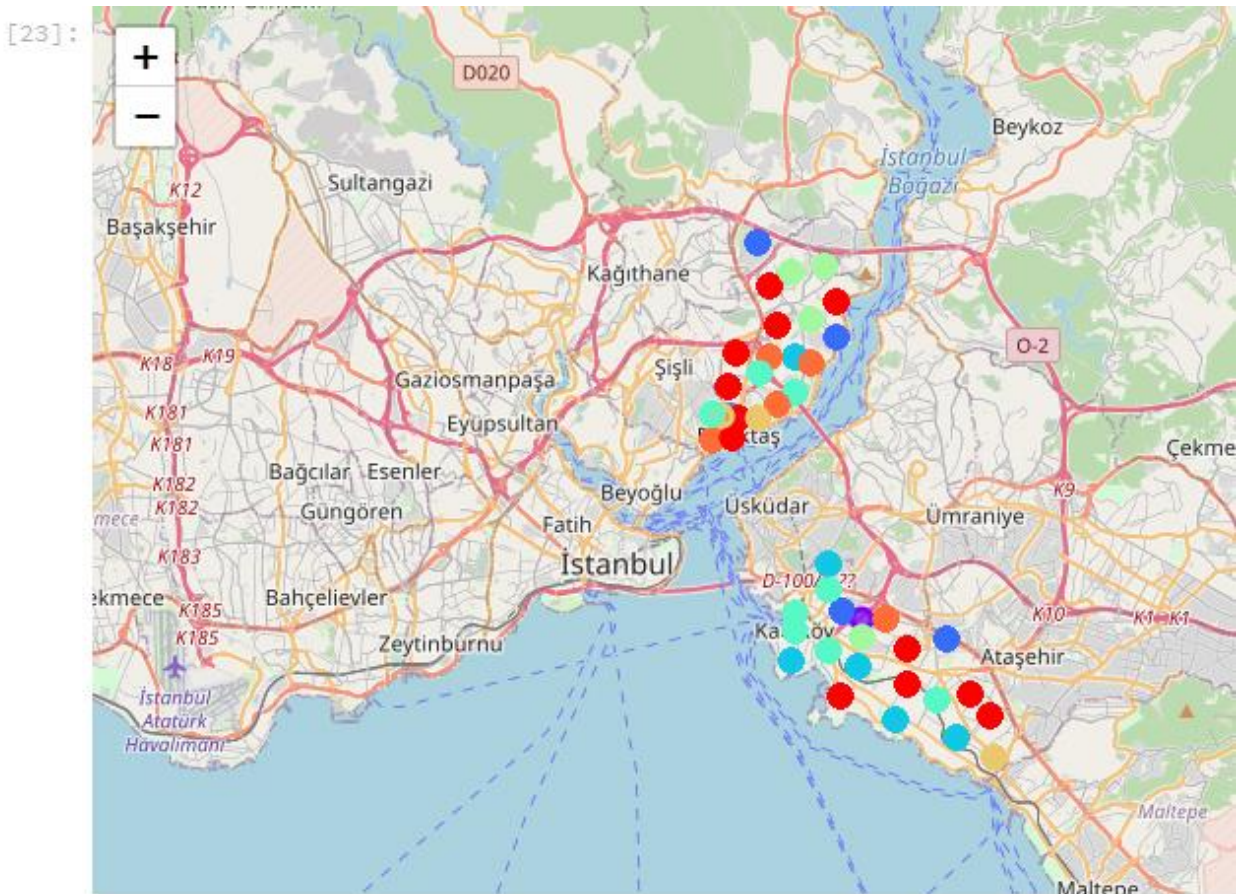


Figure 11- Cluster Map

Clustering gives a great overview to see where the deviations for categories are. For example, in cluster 6, Yoga Studio is the 2nd most common venue, and it might be good to consider an investment on Turkali Mah. neighborhood. The *frequency of the Venue Category* is the base of the model.

| | | | | | | | |
|--------------|---|----------|-----|----------------------|----------------------|--------------------|--------------------|
| Bostanci mah | 6 | Kadikoy | Gym | Yoga Studio | Gym / Fitness Center | Pilates Studio | Gymnastics Gym |
| Yildiz mah | 6 | Besiktas | Gym | Yoga Studio | Gym / Fitness Center | Weight Loss Center | Track |
| Turkali mah | 6 | Besiktas | Gym | Gym / Fitness Center | Yoga Studio | Pilates Studio | Weight Loss Center |

Figure 12 - Cluster 6 of Neighborhoods

Similar deductions can be made depending on the clustering.

4.2 DECISION TREE

Decision Tree is created based on following features:

Venue Created : Date

Neighborhood : Location

Due to the domination of **Gym** and **Gym Fitness Center** categories, the accuracy of the Decision Tree is affected. It tends to locate the result in between those two categories.

However, it is acceptable with the dataset except those categories. According to the result of the tree, it could be the best to invest on a **Pilates Studio in Kadikoy-Caddebostan, in September**. Similar predictions can be made using the model.



```

dates=np.arange(30)
df_caddebostan=pd.DataFrame(columns=['Neighborhood','Date','Result'])
neighborhood_code=8 #8 is caddebostan's code

for i in range(1,31):
    #timestamp for Sep2020 days
    dates[i-1]=int(datetime.datetime(2020,9,i,12,0).timestamp())
    df_caddebostan['Date']=dates
    df_caddebostan['Neighborhood']=neighborhood_code
    df_caddebostan['Result']=CatTree2.predict(df_caddebostan[['Date','Neighborhood']].values)

print('{} would be the best choice to open in Caddebostan'.format(df_caddebostan.groupby('Result').sum().index[0]))
Pilates Studio would be the best choice to open in Caddebostan

```