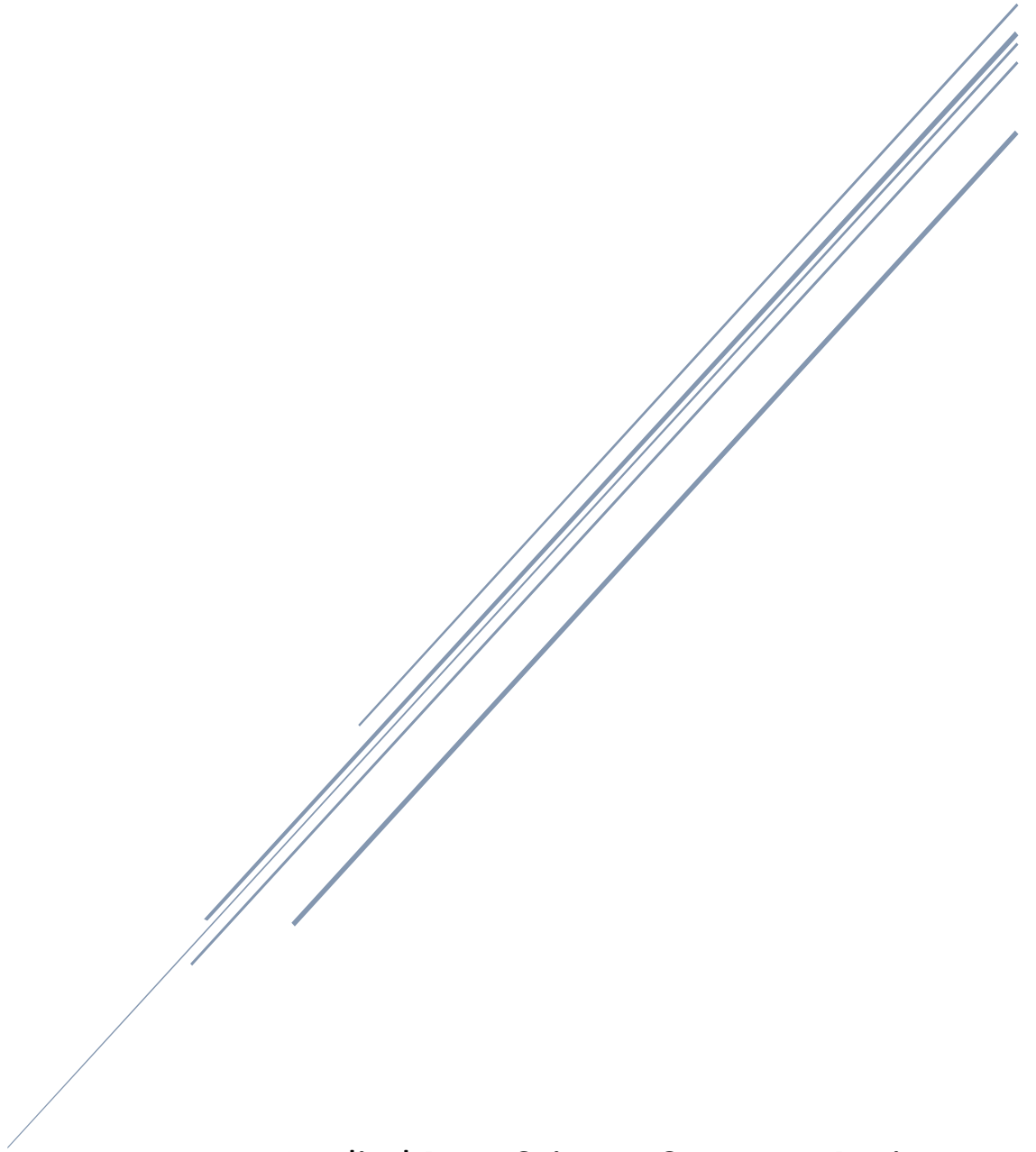


THE BATTLE OF NEIGHBORHOODS

First Step For Starting A New Business in Istanbul



Applied Data Science Capstone Project
Mehmet Mustafa Özalp

1 Introduction & Business Problem

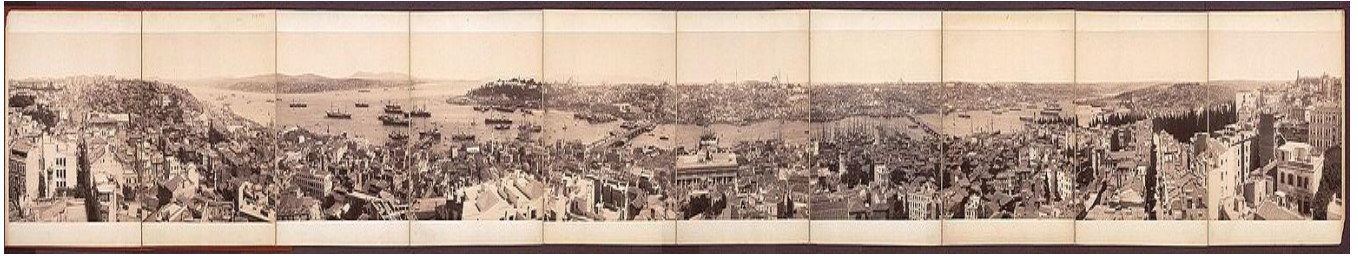


Figure 1: Istanbul: The city paved with gold!

Source:

(https://upload.wikimedia.org/wikipedia/commons/thumb/e/ea/Galata_Kulesi_%27nden_panorama.jpg/1550px-Galata_Kulesi_%27nden_panorama.jpg)

1.1 Introduction

Istanbul has always been the city of opportunity; anyone has the right and “theoretical” opportunity to start a business and be successful. The risks of starting your own business can be high, but the rewards can be well worth it.

Sitting in the wheelhouse of Turkey’s economy, the 8,500 year old city, Istanbul, overtook 129 countries with its population exceeding 15 million. Considering that there are 201 countries around the world, Istanbul is home to a population that leaves behind 129 countries, including, Belgium, Greece, Portugal, Austria, Switzerland, Norway, Denmark and Cuba.

Being one of the world’s largest metropolises, Istanbul’s population is also equivalent to the population of 41 countries put together. 41 countries, including Luxembourg, Montenegro, Malta, Iceland, Dominic Republic, Bahamas and the Virgin Islands put together only make one Istanbul.

Alongside its leadership in population density, Istanbul is also the capital of economics and trade. It holds 20 percent of the labor force in Turkey. The business world in Istanbul accounts for 27 percent of national income. Last year Turkey broke its all-time record with a level of 168,1 billion dollars worth of export. 42 percent of this number was from Istanbul. The megacity’s share in exports is 72,1 billion dollars. 37 out of 100 corporate tax payers are located in Istanbul. Therefore, the fact that things are going well in the economy’s locomotive city, which makes us say ‘its paved with gold from its stones to soil,’ is an important illustration of the fact that Turkey’s economy is developing and growing.



Figure 2: Istanbul Socioeconomic Statistics

Source: (https://www.itohaber.com/Archive/2019/7/23/210483_resource/syf_8yying.jpg)

1.2 Business Problem

Starting a new business in Istanbul offers high and fast returns with great risks. Those who think that they want to start a business should decide what they want to do first. To do this, they should conduct a good sector research and analyze the risks and advantages well. Otherwise they can fail and lost the money -maybe they saved for their whole life-.



Figure 3: Starting a New Business

Having knowledge about the market can help you to be successful on starting a new business. You may start to analyze the location, environment, competitors and trending products and etc as a first step.

We stated our business problem as the first step of starting a new business: Analyze the boroughs and neighborhoods and the businesses in those locations. It will lead us what is trend, what people mostly consume and require in that neighborhoods.

2 Data Acquisition and Preprocessing

In this project, three different datasets will be used to solve the problem: Istanbul Neighborhoods, Neighborhoods Socioeconomic Statistics and Foursquare API. After acquiring them from original and reliable sources, they will be wrangled and cleansed into more useful type for our further analysis.

2.1 Istanbul Neighborhood

Table 1: Istanbul Neighborhood

	Neighborhood	County	Area	Search	SearchDesc
0	Anadolu	Arnavutköy	1.5	arnavutkoy-anadolu	Anadolu,Arnavutköy, İstanbul
1	Arnavutköy Merkez	Arnavutköy	6.7	arnavutkoy-arnavutkoy-merkez	Arnavutköy Merkez,Arnavutköy, İstanbul
2	Arnavutköy Yavuzselim	Arnavutköy	2.1	arnavutkoy-arnavutkoy-yavuzselim	Arnavutköy Yavuzselim,Arnavutköy, İstanbul
3	Arnavutköy İmrahor	Arnavutköy	33.2	arnavutkoy-arnavutkoy-imrahor	Arnavutköy İmrahor,Arnavutköy, İstanbul
4	Arnavutköy İslambey	Arnavutköy	1.6	arnavutkoy-arnavutkoy-islambey	Arnavutköy İslambey,Arnavutköy, İstanbul

As shown above, the data consists of 954 unique neighborhoods and 38 districts to which it is affiliated in Istanbul. Search variables are created due to use them in other web scraping activities.

2.2 Demographics & Real Estate Information

Table 2: Istanbul Neighborhoods' Socioeconomic Statistics

	Neighborhood	County	ReportName	AvgSalePrice	MinSalePrice	MaxSalePrice	AvgRentPrice	MinRentPrice	MaxRentPrice	ReturnPeriod	SalePrice
0	Anadolu	Arnavutköy	Anadolu, Arnavutköy, İstanbul Bölge Raporu	247932.0	185949.0	309915.0	925.0	694.0	1156.0	22.0	
1	Arnavutköy Merkez	Arnavutköy	Arnavutköy Merkez, Arnavutköy, İstanbul Bölge Raporu	265134.0	198851.0	331418.0	1083.0	812.0	1354.0	20.0	
15	Hadımköy	Arnavutköy	Hadımköy, Arnavutköy, İstanbul Bölge Raporu	283402.0	212552.0	354253.0	1011.0	758.0	1264.0	23.0	
38	Atatürk	Ataşehir	Atatürk, Ataşehir, İstanbul Bölge Raporu	710812.0	533109.0	888515.0	2835.0	2126.0	3544.0	21.0	
39	Aşıkveysel	Ataşehir	Aşıkveysel, Ataşehir, İstanbul Bölge Raporu	379669.0	284752.0	474586.0	1577.0	1183.0	1971.0	20.0	
40	Barbaros	Ataşehir	Barbaros, Ataşehir, İstanbul Bölge Raporu	899479.0	674609.0	1124349.0	3492.0	2619.0	4365.0	21.0	
41	Esatpaşa	Ataşehir	Esatpaşa, Ataşehir, İstanbul Bölge Raporu	426822.0	320117.0	533528.0	1828.0	1371.0	2285.0	19.0	
43	Fatih	Ataşehir	Fatih, Ataşehir, İstanbul Bölge Raporu	455820.0	341865.0	569775.0	1761.0	1321.0	2201.0	22.0	
44	Kayışdağı	Ataşehir	Kayışdağı, Ataşehir, İstanbul Bölge Raporu	409786.0	307340.0	512233.0	1791.0	1343.0	2239.0	19.0	
45	Küçükbakkalköy	Ataşehir	Küçükbakkalköy, Ataşehir, İstanbul Bölge Raporu	556521.0	417391.0	695651.0	2241.0	1681.0	2801.0	21.0	

These data -as shown above- consists of neighborhoods' socioeconomic statistics which is scrapped from a popular online real estate firm's website. Due to the legal concerns, I will not share the name of the website. The data includes as below;

- ❖ Average Sales Price
- ❖ Min Sales Price
- ❖ Max Sales Price
- ❖ Average Rent Price
- ❖ Min Rent Price
- ❖ Max Rent Price
- ❖ Average Return Period of a House
- ❖ Sales Price Change In Last Month
- ❖ Sales Price Change In 3 Years
- ❖ Sales Price Change In 5 Years
- ❖ Rent Price Change In Last Month
- ❖ Rent Price Change In 3 Years
- ❖ Rent Price Change In 5 Years
- ❖ Population
- ❖ Married Percentage of Population
- ❖ Single Percentage of Population
- ❖ Socioeconomic Class

2.3 Foursquare API

Finally, due to the limitations of foursquare api, 3 most popular county is chosen for further analysis and foursquare api is called for these counties' neighborhoods. And we acquired data as show below;

Table 3: Foursquare Venues

Neighborhood	County	NeighborhoodLatitude	NeighborhoodLongitude	VenueId	VenueName	VenueLatitude	VenueLongitude	VenueCategory
Ataköy 1. kısım	Bakırköy	40.974646	28.870076	55681c5f498ed3e60be0904d	Happy Moon's	40.974813	28.870578	Restaurant
Ataköy 1. kısım	Bakırköy	40.974646	28.870076	555b40b1498eec90b9a0999f	Joker Baby	40.974578	28.867757	Baby Store
Ataköy 1. kısım	Bakırköy	40.974646	28.870076	4ddf8cf018388714bb6a0f79	Cookshop	40.974948	28.869112	Restaurant
Ataköy 1. kısım	Bakırköy	40.974646	28.870076	4e1e990bd4c0fc6e343aae5f	Starbucks	40.974969	28.868663	Coffee Shop
Ataköy 1. kısım	Bakırköy	40.974646	28.870076	4cbc55bc7a5d9eb0a95c31e9	Sheraton Istanbul Ataköy Hotel	40.972817	28.871142	Hotel

3 Methodology

3.1 Exploratory Analysis

After merging and cleansing datasets, some visualizations are created to interpret the data better.

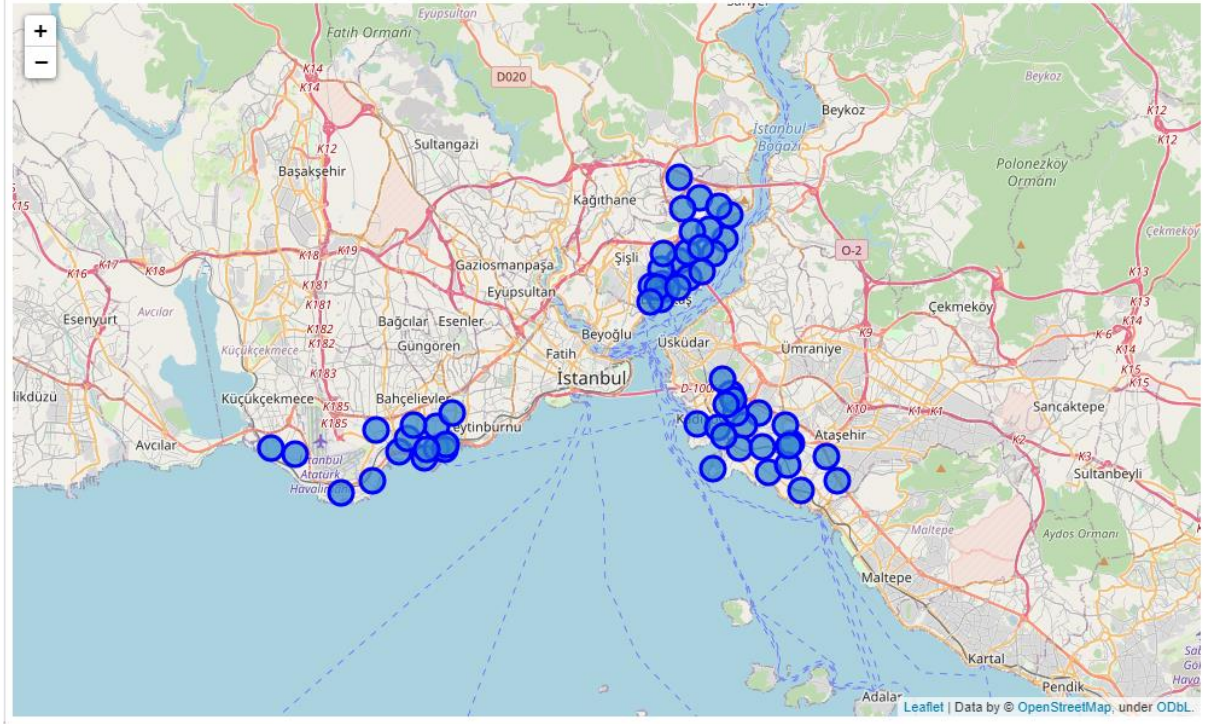


Figure 4: Selected Neighborhoods (Beşiktaş, Bakırköy and Kadıköy counties)

As shown above, the selected neighborhoods are located in the most popular counties of Istanbul. The main reason for choosing this inner district is that they are similar in many ways, and if desired, the analysis can be arranged for any district to be selected.

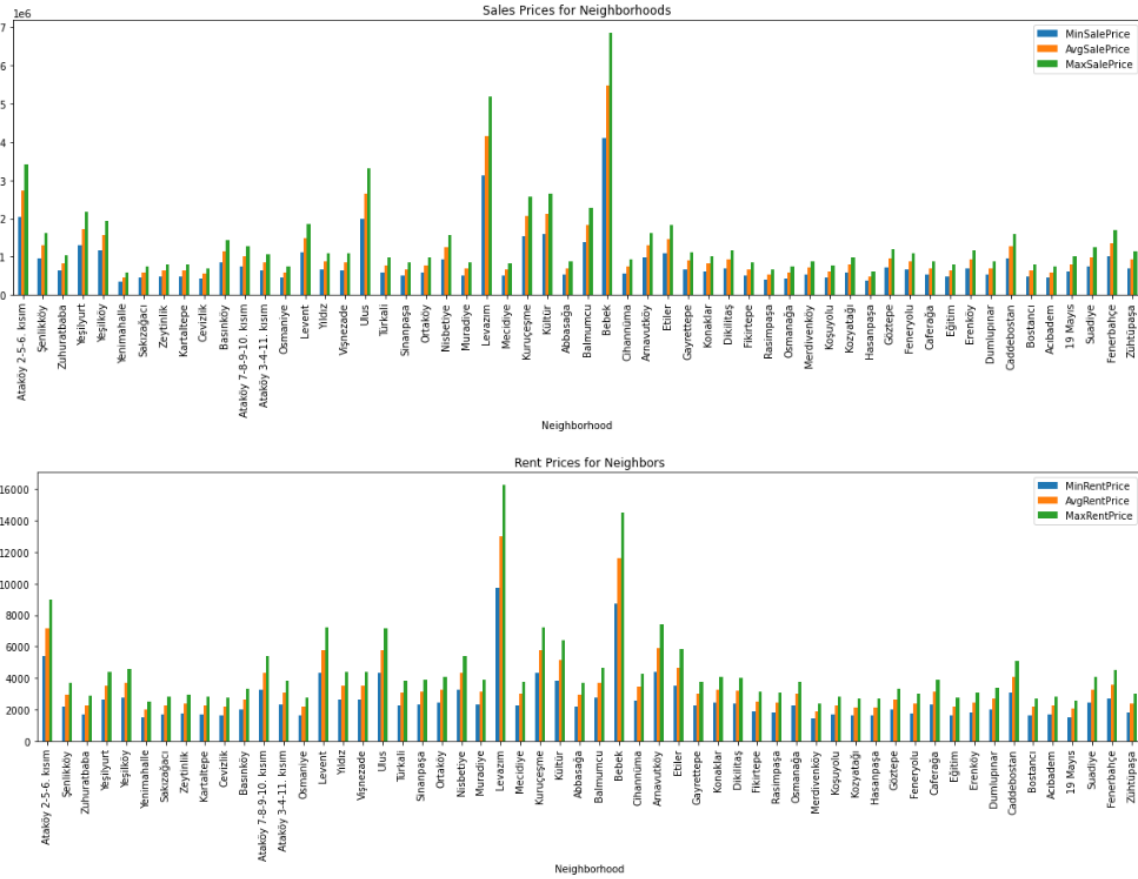


Figure 5: Sales and Rent Prices for Neighborhoods

These bar charts displaying the neighborhoods sales and rent prices (average, min, max). As shown in the Bebek, Levazım, Ataköy 2-5-6. Kısım, Ulus, Levent, Arnavutköy ve Etiler neighborhoods are leaders on both sales and rent prices.

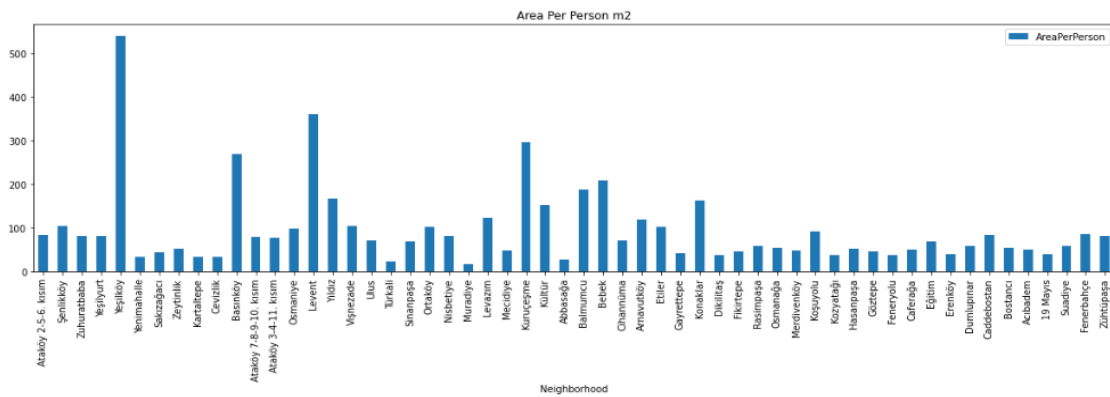


Figure 6: Area per person (meter square)

As shown above, the area per person charts show that Yeşilköy, Başınköy, Levent and Kuruçeşme are leading on area per person. The type and level of construction plays an important role in this.

3.2 Cluster Analysis

Afterwards, K-means clustering was conducted in order to group the neighborhoods according to their socioeconomic statistics and most frequent venues they have. As the first step of the cluster analysis encoding and scaling conducted to prepare data for the model and the data shape was (56, 87).

To find the optimal cluster number; inertia, silhouette_score and calinski_harabasz_score are used.

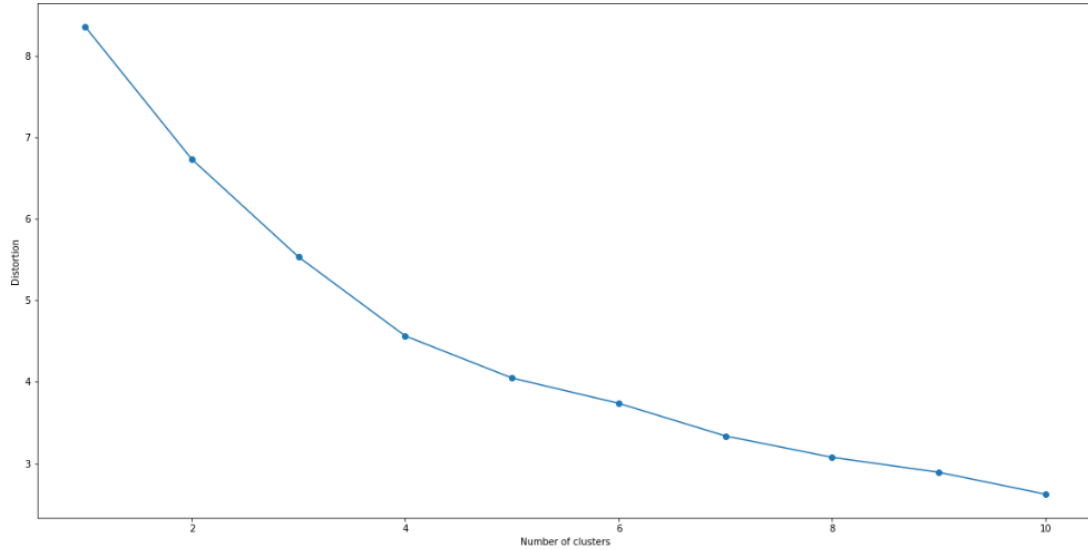


Figure 7: Inertia over cluster numbers

Table 4: Scores over cluster numbers

	cluster	inertia	silhouette_score	ch_score
0	1	8.360492	NaN	NaN
1	2	6.733137	0.269405	13.051441
2	3	5.534035	0.247638	13.534629
3	4	4.563612	0.264687	14.421158
4	5	4.047806	0.278314	13.584330
5	6	3.736948	0.229268	12.372514
6	7	3.336433	0.237668	12.297507
7	8	3.074267	0.229503	11.790905
8	9	2.890200	0.237308	11.119633
9	10	2.621817	0.236832	11.187279

Here we can see the drop in the sum of squared distance starts to slow down after $k=8$. Hence 6 is the optimal number of clusters for our analysis. We can verify this by calculating the silhouette coefficient and CH score for $k=6$. Both the values are higher than they were for our earlier clusters. We can conclude that $k=6$ is our optimal number of clusters.

4 Results

Upon different analysis, we are able to discover the best neighborhoods based on the socioeconomic status and at the same time which kind of venues are suitable for those neighborhoods.

4.1 1st Cluster

First cluster consists of one 3 neighborhoods with 546K sale price, 44.6 square meter area per person. This cluster is the lowest cluster according to socioeconomics.

```
AreaPerPerson      44.60
AvgSalePrice       546,362.33
SalePrice5YearChange 0.51
AvgRentPrice       2,432.67
RentPrice5YearChange 0.53
Name: mean, dtype: float64
```

	Neighborhood	County	Cluster	SocioeconomicClas	AreaPerPerson	AvgSalePrice	SalePrice5YearChange	AvgRentPrice	RentPrice5YearChange	1st Most Common Venue
8	Sakızağacı	Bakırköy	0	B+	44.98	589,433.00	0.71	2,257.00	0.58	Restaurant
9	Yenimahalle	Bakırköy	0	B+	33.50	483,840.00	0.28	2,027.00	0.30	Restaurant
54	Osmanağa	Kadıköy	0	A+	55.32	588,014.00	0.52	3,014.00	0.70	Café

4.2 2nd Cluster

Second cluster consists of one 13 neighborhoods with 860K sale price, 91.8 square meter area per person. This cluster is the fourth cluster according to socioeconomics.

```
AreaPerPerson      91.81
AvgSalePrice       858,628.46
SalePrice5YearChange 0.53
AvgRentPrice       2,762.69
RentPrice5YearChange 0.63
Name: mean, dtype: float64
```

	Neighborhood	County	Cluster	SocioeconomicClas	AreaPerPerson	AvgSalePrice	SalePrice5YearChange	AvgRentPrice	RentPrice5YearChange	1st Most Common Venue
2	Ataköy 3-4-11. kısım	Bakırköy	1	A	77.54	847,148.00	0.45	3,072.00	0.58	Café
3	Ataköy 7-8-9-10. kısım	Bakırköy	1	A+	78.51	1,006,960.00	0.41	4,344.00	0.92	Café
4	Basınköy	Bakırköy	1	A	269.13	1,145,384.00	0.76	2,852.00	0.59	Café
13	Zuhuratbaba	Bakırköy	1	A+	82.16	835,380.00	0.90	2,297.00	0.50	Pizza Place
14	Şenlikköy	Bakırköy	1	A+	104.34	1,291,062.00	0.61	2,960.00	0.32	Café
24	Konaklar	Beşiktaş	1	A+	163.72	816,129.00	0.27	3,294.00	0.46	Park
40	Bostancı	Kadıköy	1	A+	54.23	642,322.00	0.44	2,180.00	0.43	Café
43	Dumlupınar	Kadıköy	1	A	59.52	699,442.00	0.19	2,727.00	0.93	Café
45	Eğitim	Kadıköy	1	A	68.93	639,080.00	0.82	2,193.00	1.06	Café
46	Fenerbahçe	Kadıköy	1	A+	88.43	1,362,835.00	0.88	3,599.00	0.48	Café
48	Fikirtepe	Kadıköy	1	A-	46.39	678,831.00	0.11	2,531.00	1.13	Café
50	Hasanpaşa	Kadıköy	1	A-	52.45	488,320.00	0.50	2,165.00	0.48	Café
53	Merdikenköy	Kadıköy	1	A+	49.21	708,299.00	0.82	1,901.00	0.31	Café

4.3 3rd Cluster

Third cluster consists of one 10 neighborhoods with 2202K sale price, 102.9 square meter area per person. This cluster is the first cluster according to socioeconomics.

```
AreaPerPerson      102.95
AvgSalePrice      2,202,593.90
SalePrice5YearChange    0.95
AvgRentPrice       5,686.90
RentPrice5YearChange    0.59
Name: mean, dtype: float64
```

	Neighborhood	County	Cluster	SocioeconomicClas	AreaPerPerson	AvgSalePrice	SalePrice5YearChange	AvgRentPrice	RentPrice5YearChange	1st Most Common Venue
1	Ataköy 2-5-8. kısım	Bakırköy	2	A+	84.70	2,738,520.00	0.82	7,188.00	0.78	Café
11	Yeşilyurt	Bakırköy	2	A+	80.74	1,735,690.00	1.88	3,499.00	0.59	Café
15	Abbasağa	Beşiktaş	2	A	27.39	698,687.00	0.55	2,956.00	0.39	Coffee Shop
18	Balmumcu	Beşiktaş	2	A+	188.13	1,825,730.00	0.81	3,698.00	0.38	Café
19	Bebek	Beşiktaş	2	A+	209.28	5,476,497.00	1.34	11,606.00	0.96	Café
21	Dikilitaş	Beşiktaş	2	A+	37.89	936,628.00	0.67	3,214.00	0.50	Coffee Shop
27	Levazım	Beşiktaş	2	A+	123.26	4,150,260.00	1.29	12,988.00	1.11	Boutique
35	Ulus	Beşiktaş	2	A+	71.01	2,658,331.00	0.77	5,750.00	0.26	Music Venue
37	Yıldız	Beşiktaş	2	A+	167.35	872,685.00	0.73	3,519.00	0.55	Café
44	Erenköy	Kadıköy	2	A+	39.72	932,911.00	0.66	2,453.00	0.35	Bakery

4.4 4th Cluster

Fourth cluster consists of one 15 neighborhoods with 900K sale price, 110.2 square meter area per person. This cluster is the third cluster according to socioeconomics.

```

AreaPerPerson      110.20
AvgSalePrice       900,995.27
SalePrice5YearChange  0.60
AvgRentPrice       2,854.73
RentPrice5YearChange  0.44
Name: mean, dtype: float64

```

	Neighborhood	County	Cluster	SocioeconomicClas	AreaPerPerson	AvgSalePrice	SalePrice5YearChange	AvgRentPrice	RentPrice5YearChange	1st Mo: Commo Venu
5	Cevizlik	Bakırköy	3	A-	34.20	580,005.00	0.83	2,222.00	0.43	Caf
7	Osmaniye	Bakırköy	3	A+	98.02	589,981.00	0.86	2,204.00	0.82	Caf
10	Yeşilköy	Bakırköy	3	A+	539.14	1,555,450.00	0.72	3,693.00	0.47	Seafoo Restaur
12	Zeytinlik	Bakırköy	3	A	53.46	629,005.00	0.89	2,391.00	0.51	Caf
23	Gayrettepe	Beşiktaş	3	A+	41.62	892,504.00	0.57	2,997.00	0.43	Hot
28	Levent	Beşiktaş	3	A+	361.25	1,479,065.00	0.26	5,781.00	0.45	Art Galler
38	19 Mayıs	Kadıköy	3	A+	40.34	804,898.00	0.78	2,051.00	0.33	Caf
39	Acıbadem	Kadıköy	3	A+	49.62	594,012.00	0.42	2,269.00	0.35	Caf
41	Caddebostan	Kadıköy	3	A+	83.12	1,263,347.00	0.79	4,078.00	0.66	Coffe Sho
47	Feneryolu	Kadıköy	3	A+	38.87	870,300.00	0.59	2,397.00	0.34	Caf
49	Göztepe	Kadıköy	3	A+	45.19	950,388.00	0.74	2,656.00	0.52	Baker
51	Kozyatağı	Kadıköy	3	A+	37.65	785,231.00	0.76	2,149.00	0.34	Caf
52	Koşuyolu	Kadıköy	3	A-	91.35	622,492.00	0.56	2,247.00	0.57	Caf
57	Suadiye	Kadıköy	3	A+	57.94	995,114.00	0.56	3,275.00	0.52	Coffe Sho
58	Zühtüpaşa	Kadıköy	3	A	81.20	923,157.00	0.32	2,411.00	0.10	Socce Stadiu

4.5 5th Cluster

Fifth cluster consists of one 8 neighborhoods with 1191K sale price, 109.4 square meter area per person. This cluster is the second cluster according to socioeconomics.

```
AreaPerPerson      109.44
AvgSalePrice      1,191,230.38
SalePrice5YearChange 0.39
AvgRentPrice       3,867.38
RentPrice5YearChange 0.27
Name: mean, dtype: float64
```

	Neighborhood	County	Cluster	SocioeconomicClas	AreaPerPerson	AvgSalePrice	SalePrice5YearChange	AvgRentPrice	RentPrice5YearChange	1st Most Common Venue
6	Kartalteppe	Bakırköy	4	A+	34.11	648,267.00	0.45	2,263.00	0.49	Baker
22	Etiler	Beşiktaş	4	A+	102.54	1,465,654.00	0.03	4,679.00	0.13	Steakhous
25	Kuruçesme	Beşiktaş	4	A+	296.15	2,061,468.00	0.63	5,760.00	0.23	Restaurar
26	Kültür	Beşiktaş	4	A+	151.42	2,127,174.00	0.41	5,137.00	0.07	Caf
29	Mecidiye	Beşiktaş	4	A+	47.80	664,343.00	0.35	3,010.00	0.34	Caf
31	Nispetiye	Beşiktaş	4	A+	81.55	1,256,279.00	0.25	4,322.00	0.17	Gym Fitnes Cente
32	Ortaköy	Beşiktaş	4	A+	102.75	781,700.00	0.48	3,290.00	0.36	Caf
55	Rasimpaşa	Kadıköy	4	A+	59.22	524,958.00	0.49	2,478.00	0.39	Caf

4.6 6th Cluster

Sixth cluster consists of one 8 neighborhoods with 820K sale price, 65 square meter area per person. This cluster is the fifth cluster according to socioeconomics.

```
AreaPerPerson      65.00
AvgSalePrice      820,849.71
SalePrice5YearChange 0.29
AvgRentPrice       3,616.00
RentPrice5YearChange 0.38
Name: mean, dtype: float64
```

	Neighborhood	County	Cluster	SocioeconomicClas	AreaPerPerson	AvgSalePrice	SalePrice5YearChange	AvgRentPrice	RentPrice5YearChange	1st Most Common Venue
17	Arnavutköy	Beşiktaş	5	A+	118.80	1,298,895.00	-0.10	5,905.00	0.16	Seafood Restaurant
20	Cihannüma	Beşiktaş	5	A+	72.13	745,628.00	0.30	3,440.00	0.55	Coffee Shop
30	Muradiye	Beşiktaş	5	A	16.52	684,933.00	0.37	3,139.00	0.53	Coffee Shop
33	Sinanpaşa	Beşiktaş	5	A+	68.55	672,212.00	0.41	3,136.00	0.43	Coffee Shop
34	Türkali	Beşiktaş	5	A+	23.17	780,491.00	0.58	3,057.00	0.44	Coffee Shop
36	Vişnezade	Beşiktaş	5	A+	104.48	863,598.00	0.08	3,517.00	0.00	Coffee Shop
42	Caferağa	Kadıköy	5	A+	51.32	700,191.00	0.41	3,118.00	0.57	Café

5 Discussion

Results show that socioeconomics are effective on created clusters. Moreover, venue categories and venues in these clusters vary luxury levels. For example cluster 3 and cluster 5 consists of neighborhoods with high socioeconomic values. Mostly luxury cafes and restaurants (different concepts, different kitchens), gym, spa, boutique and art galleries come forwards.

Rest of the clusters consists of neighborhoods of which mostly white collar workers live and work, mostly serve their requirements. Pubs, restaurants gyms, public pools, coffee shops, clothing stores and etc.

Since this is an introductory work for location analysis for a new business, it can be considered to narrow analysis concept.

6 Conclusion

From this analysis, we have found that socioeconomic data is important at clustering neighborhoods and venue categories. As shown in the selected counties have high seriocomics, the level and importance of luxury stands out.

Since this is an introductory work for location analysis for a new business, it can be considered to narrow analysis concept. For example if you want to open a pizza point, you can analyze pizza venues, delivery points of neighborhoods and their socioeconomics for the level of services.