Capstone Project Report

The Battle of Neighbourhoods

Choosing the best location for a new restaurant in Istanbul

Hüseyin Güçlü Çaştaban, September 2019

Introduction: Business Understanding

In this project our aim is to find the best location to open a new restaurant in Istanbul. Our main stakeholder is the management of an international restaurant chain which is planning to open its first venue in the Turkish metropolis. The restaurant will target the high-income segment. Its menu will showcase some of the best examples of world's gastronomy, including popular meals from French, Italian, Japanese, Chinese and Mexican cuisines.

Istanbul is Turkey's largest city with a population of 15 million. It is the centre of economic activity with 31% of the country's GDP coming from the city. Istanbul's per capita income is US$ 18,000. It has a highly skewed distribution with a Gini index of 0.443. Istanbul has active nightlife and historic taverns, a signature characteristic of the city for centuries. Many of the city's most popular and upscale restaurants line the shores of the Bosphorus. Administratively, Istanbul consists of 39 boroughs (*ilçe*) and 782 neighbourhoods (*mahalle*).

The aim of the project is to propose the best location (borough and neighbourhood) to be chosen by the international restaurant chain.

Data

To address the business problem we will need the list of Istanbul's boroughs and neighbourhoods along with their geographical coordinates. We will also need household income data of each borough, and location data concerning the restaurants in Istanbul. Following data sources are used to extract the required information:

- List of Istanbul's boroughs and their geographical coordinates is obtained from Ismail Baskın's Github repository
- The average household income data is acquired by using a daily newspaper website
- The number of restaurants and their type and location is received using Foursquare API
- The population data and average rent data of the top candidate borough is obtained from public websites.
- The latitude and longitude values of the neighbourhoods are acquired by using the "Nominatim" geolocation library.
- Istanbul and its boroughs' GeoJson files were acquired from Yusuf Güven's and Leonardo Iheme's Github repositories.

The acquired data is organised in two CSV tables which are uploaded to the Capstone project's Github repository.

Methodology

The project methodology consists of two stages, namely 1) identifying the best borough amongst 29 other boroughs of Istanbul where the high-income segment is concentrated, 2) identifying the best possible neighbourhood in that borough which will maximise the chances of commercial success.

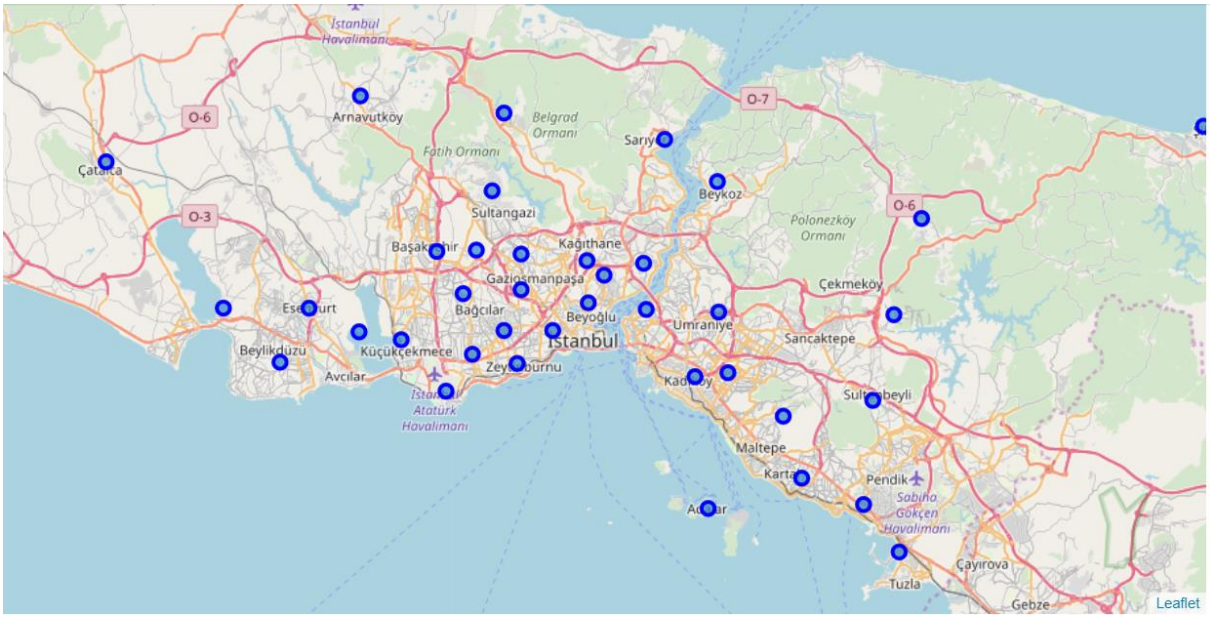
As far as the second step is concerned, three characteristics will be examined in order to choose the best possible neighbourhood(s):

1. **Population**: the population of the neighbourhood of choice should at least 10000, so that the condition of a critical mass in the vicinity is satisfied.
2. **Average House Rent**: The higher the average house rent, the 'posher' is the neighbourhood, hence the target segment concentration.
3. **Restaurant Density:** The lower the restaurant density, the milder is competition within the neighbourhood, where density is defined as the number of restaurants per 10,000 residents.

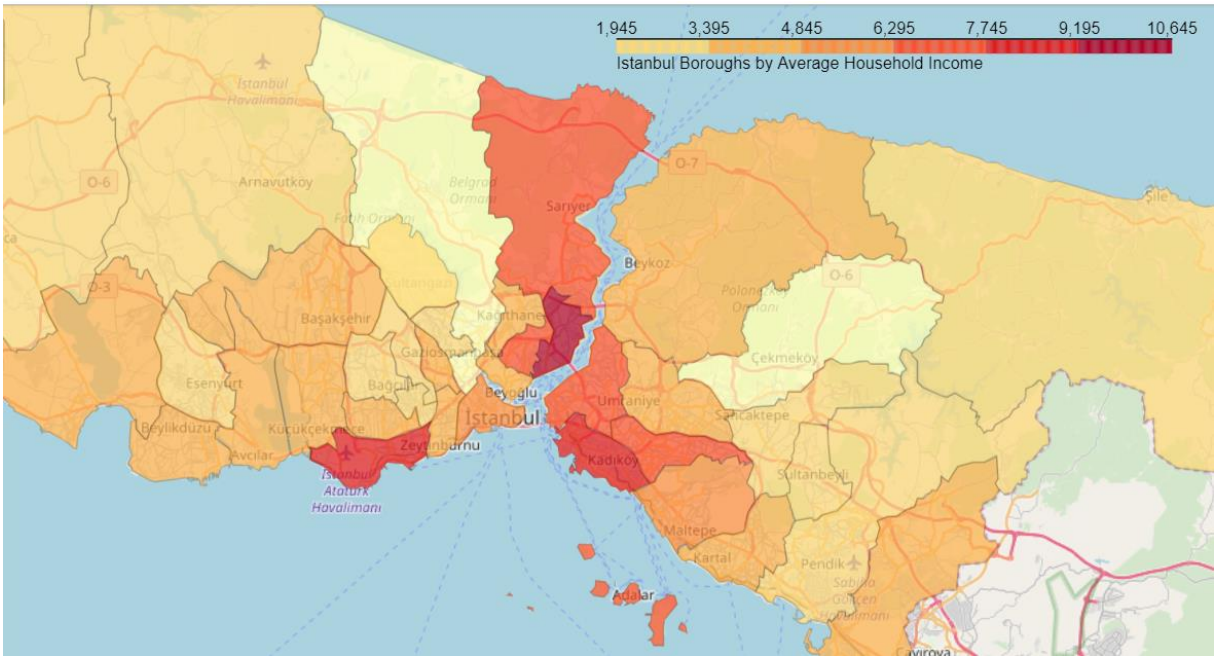
Finally, a k-means clustering analysis will be conducted to explore any hidden similarities between neighbourhoods.

Istanbul and its boroughs

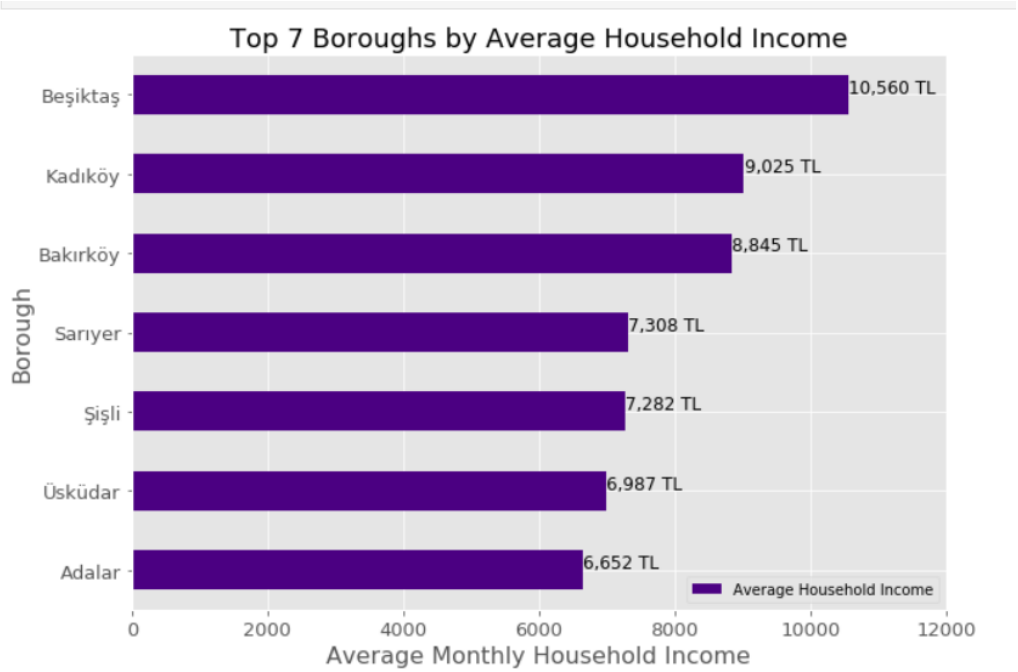
Istanbul consists of 39 boroughs (*ilçe*) scattered on both sides of the Bosphorus strait.



In order to identify the best possible borough we will analyse the results of a household income survey published by a daily newspaper. Below there is a choropleth chart created using **average monthly household income** figures of each borough.



As one can see on the above map, Istanbul's boroughs widely differ in average household income, where one can spot a few with a high income. Let's further visualise the data by means of a horizontal bar chart.



Beşiktaş is by far the wealthiest borough with an average household income of **10,560 TL**. This is where we have the highest concentration of **high income segment**, therefore it is advisable that the restaurant chain opens its first venue in Beşiktaş.

The neighbourhoods of Beşiktaş

Foursquare Data

Now that we decided to focus on Beşiktaş borough, let's use Foursquare API to get information on restaurants in each neighborhood. We searched for venues in 'food' category, but only for those that are restaurants - coffee shops, pizza places, bakeries etc. are not relevant so we don't care about those. So, we included in our list only venues that have 'restaurant' and similar words in category name. The Foursquare API returned 969 restaurants in 40 unique categories.

```
besiktas_restaurants.head()
```

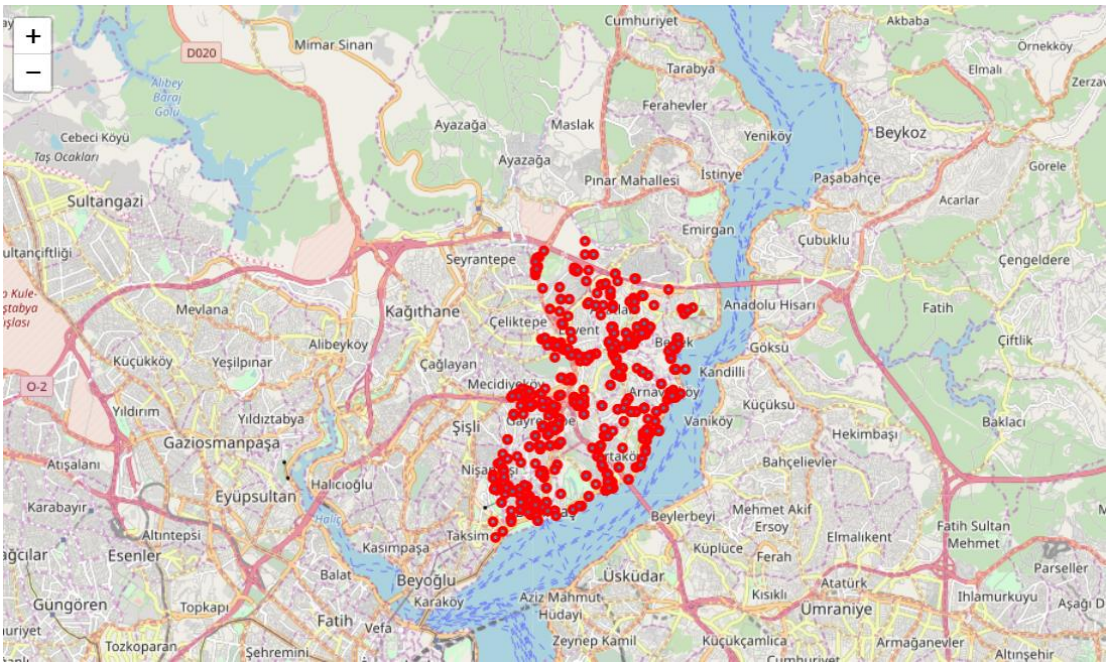
```
Obtaining venues around candidate locations: . . . . . done.
Total number of restaurants in Beşiktaş:  (969, 7)
```

[7]:

	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Abbasağa	41.049574	29.005171	Nisa Cafe Ev Yemekleri	41.049978	29.007086	Restaurant
1	Abbasağa	41.049574	29.005171	Kamos	41.050900	29.005947	Greek Restaurant
2	Abbasağa	41.049574	29.005171	Edirne Ciğercisi Naci Usta	41.048101	29.002113	Turkish Restaurant
3	Abbasağa	41.049574	29.005171	Çiğköftem	41.045906	29.002836	Turkish Restaurant
4	Abbasağa	41.049574	29.005171	Chinese Time	41.050894	29.007095	Chinese Restaurant

[23]:
```
print('There are {} uniques categories.'.format(len(besiktas_restaurants['Venue Category'].unique())))
```

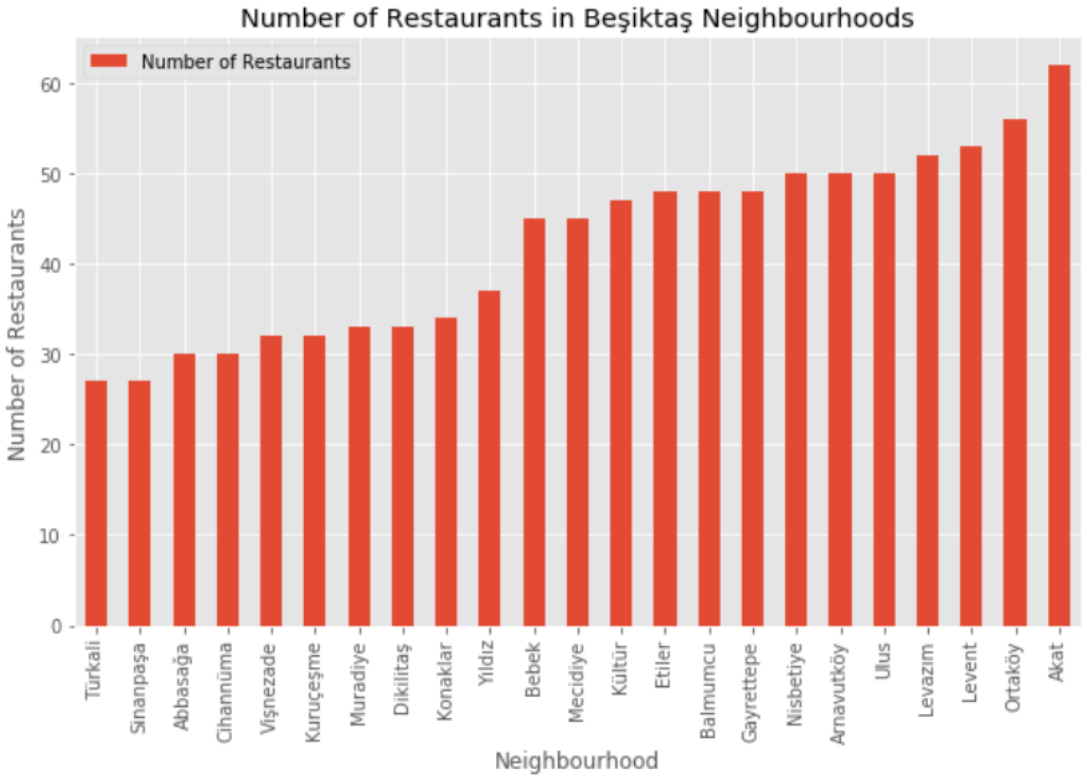
```
There are 40 uniques categories.
```



The above map shows all the restaurant venues in Beşiktaş. One can see that there is concentration along the shore of Bosphorus as well along main avenues and streets.

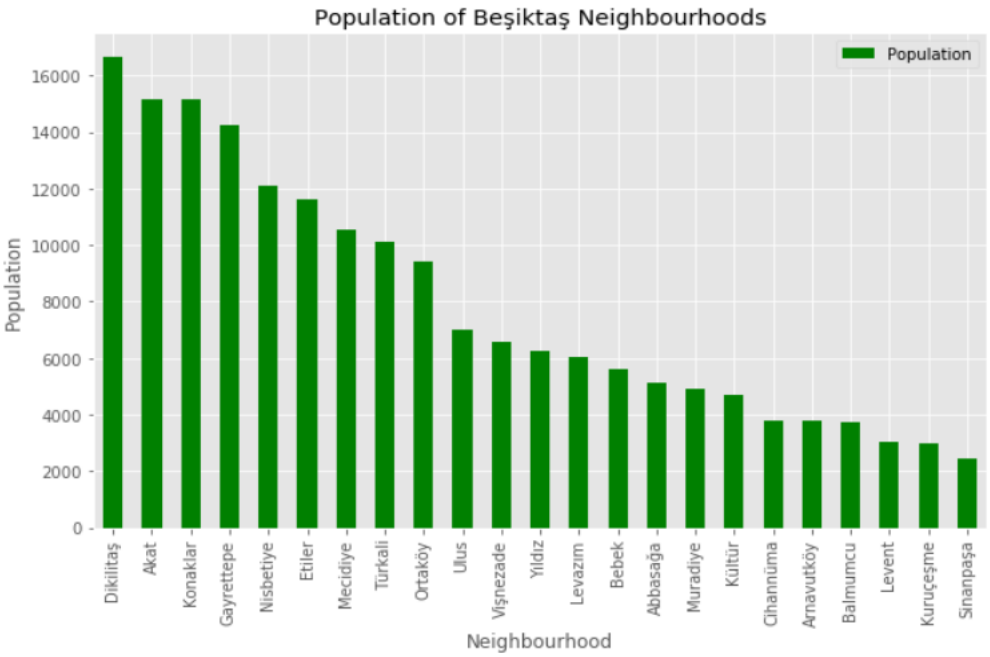
Number of Restaurants

Akat has the highest number of restaurants amongst all neighbourhoods.



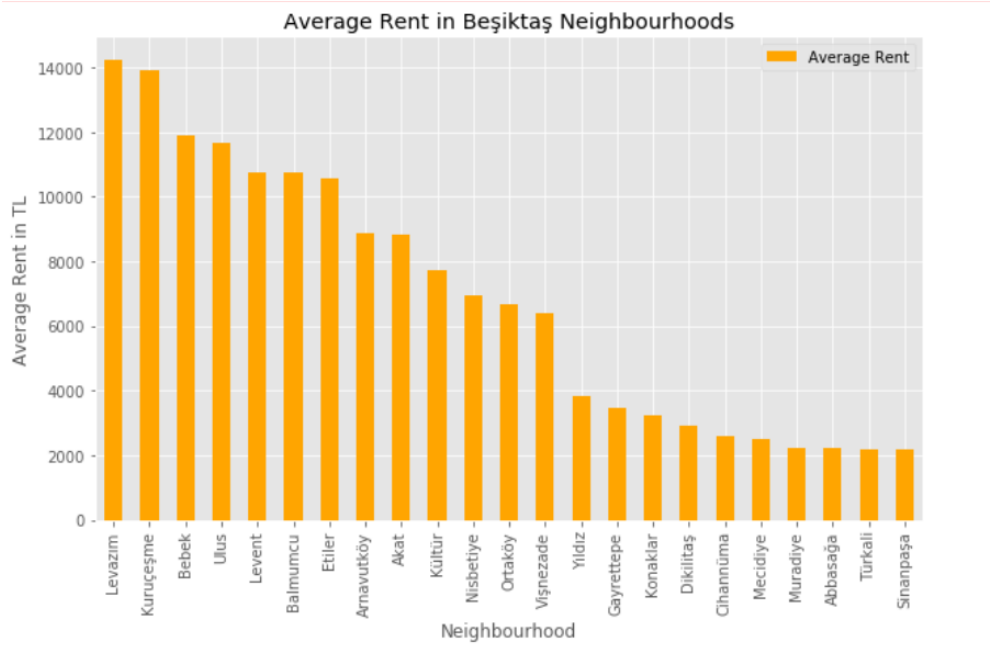
Population

It is important to have a critical mass in the vicinity of the neighbourhood. There are 8 neighbourhoods exceeding the required 10k population threshold.



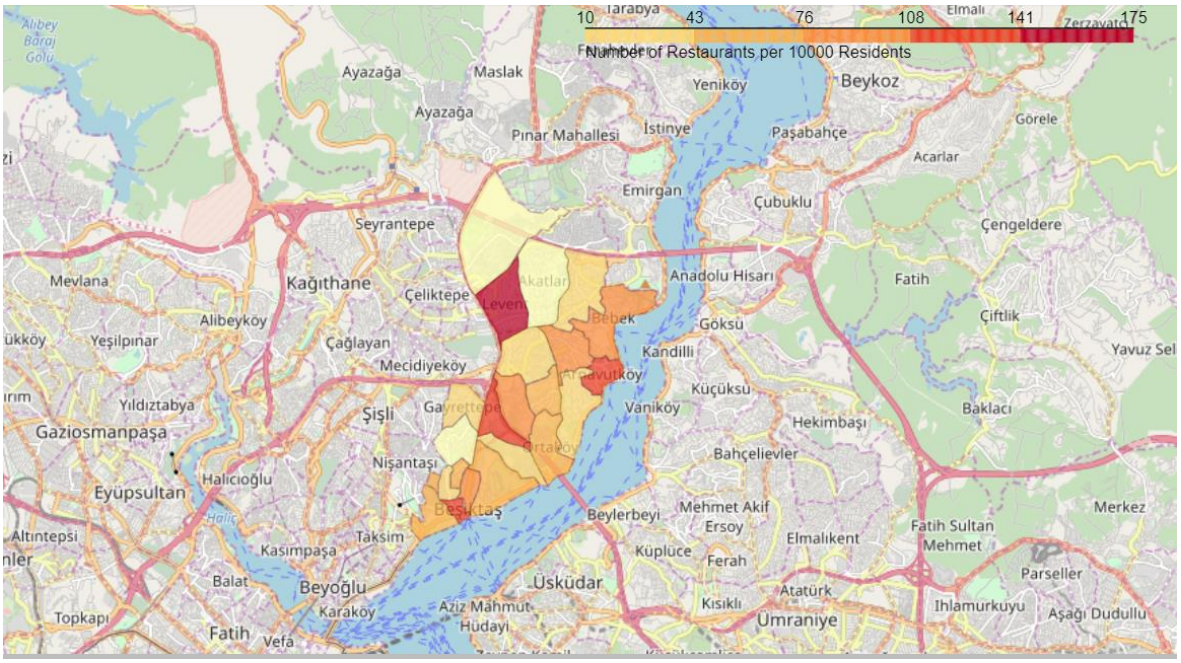
Average House Rent

Average house rent is an indication of household income, hence the higher the rent, the posher is the neighbourhood. Levazım, Kuruçeşme and Bebek are the top three neighbourhoods as far average rent is concerned.



Restaurant Density

The restaurant density is calculated as the number of restaurants per 10,000 residents. The choropleth map clearly shows the variations between neighbourhoods in terms of restaurant density. Dikilitaş, Türkali and Konaklar are the most advantageous neighbourhoods.

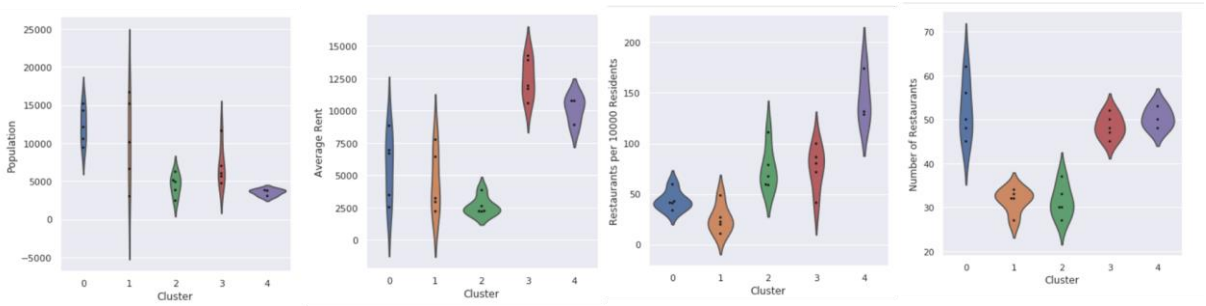


Clustering Analysis

A k-means clustering analysis was performed in order to explore the hidden similarities between the neighbourhoods. At data preparation stage, all categorical variables were dropped from the data set, and the remaining numerical variables were normalized with the standard deviation. The analysis yielded 5 clusters.

Cluster	Neighbourhoods	Population (sum)	Average Rent (mean)	Restaurant Density (mean)
0	Akat, Gayrettepe, Mecidiye, Nisbetiye, Ortaköy	61520	5681	43.6
1	Dikilitaş, Konaklar, Kuruçeşme, Türkali, Vişnezade	51541	4501	25.6
2	Abbasağa, Cihannüma, Muradiye, Sinanpaşa, Yıldız	22482	2615	75.1
3	Bebek, Etiler, Kültür, Levazım, Ulus	34948	12461	75.8
4	Arnavutköy, Balmumcu, Levent	10583	10129	144.7

The resulting clusters were examined in terms of the distributions of key input variables.



Observations:

- In terms of **Population**, Cluster 0 is the most populous cluster whilst having a restaurant density (43.6) which is lower than Beşiktaş average (53,1).
- As far as **Average House Rent,** goes Cluster 3 is the clear leader, which is a strong indication for household income. Indeed, this cluster includes some of the most expensive neighbourhoods such as **Bebek, Etiler, and Ulus**.

- Looking at **restaurant density**, Cluster 1 is the most advantageous, whilst Cluster 4 is the least. In fact, **Cluster 4** has the highest restaurant density 144.7. **Levent**, a neighbourhood with many shopping malls is included in this cluster.
- **Cluster 1** has the lowest restaurant density, and it is mainly composed of residential neighbourhoods such as Dikilitaş and Türkali.

Results

Among the 39 boroughs of Istanbul, Beşiktaş is by far the wealthiest borough with an average household income of 10,560 TL. This is where we have the highest concentration of high income segment, therefore it is advisable that the restaurant chain opens its first venue in Beşiktaş.

The Foursquare analysis returned 969 restaurants which are labelled under 40 different restaurant categories. Most of Beşiktaş restaurants are situated along the Bosporus or on the main avenues and streets of the neighbourhoods.

The Akat neighbourhood has the highest number of restaurants, whereas Türkali has the lowest.

Beşiktaş' total population is 181,074, and with 969 restaurants counted in the study the average restaurant density per 10000 residents is calculated as 53,51. As far as the restaurant density goes, there are a handful of choices where the number per 10000 residents are less than the average. The most advantages areas are Dikilitaş and Konaklar.

Levazım, Kuruçeşme and Bebek are the top three neighbourhoods as far average house rent is concerned. There are 7 neighbourhoods which are above 10000 TL threshold, which can be classified as super high income.

Looking at population of neighbourhoods, only 8 out of 22 meet the minimum 10000 population criteria for the new restaurant reaching a critical mass in its vicinity. Examining the average rent as a proxy for household income, Levent and Kuruçeşme occupy the top two spots.

Discussion

In order to formulate a **recommendation** for our stakeholders, we need to examine the neighbourhoods' population, average rent and restaurant density all in combination. The below scatterplot deals with this issue in a three-dimensional manner, with restaurant density, population are plotted on x and y axes, respectively. The size of the bubbles represent the average house rent.

The upper half of the chart is the area where we can find the neighbourhoods which meet the minimum **population requirement** of 10,000. Luckily, in this area the restaurant density is lower than Beşiktaş average. At a **restaurant density** of only 40, which is quite favourable for a new market entrant, one can spot immediately three neighbourhoods sporting relatively large bubble sizes. They have the **highest average house rent**, which is an indication for high income segment concentration. In order of preference, the top three candidates are **Etiler, Akat and Nisbetiye**. Based on the criteria outlined in the methodology and the findings of the analyses, we can recommend that the best spots for the new restaurant are these three neighbourhoods.

Having said that the above recommendation can be further refined by collecting more data about the potential consumers. For example, a sample of neighbourhood residents can be contacted by an online survey to enquire about their preferences for the menu mix, price range and inner layout of the restaurant. Restaurants' foot traffic data can be obtained through third party location data providers.

The recommendation can be further be refined by adding other factors into the analysis. One of these factors is the number of potential customers who would travel from other boroughs and neighbourhoods to visit the restaurant. In this case, travel time to other major centres as well as access and promixity to public transport services can be considered as additional criteria.

Finally, online consumer feedback (likes, reviews, recommendations) about restaurants in particular locations can be collected and incorporated into this study.

Conclusion

This study deals with the business problem of finding the best location for the first venue of an international restaurant chain. It outlines the findings of explanatory data analyses, data visualisation and clustering analyses. The analyses were run in two steps. In the first step, we looked at Istanbul's boroughs, and made comparisons in terms of household income, which led to the selection of Beşiktaş borough. In the second step, the characteristics of Beşiktaş' 22 neighbourhoods were analysed. A k-means clustering analyses was run to identify the neighbourhoods that were most similar. Finally, neighbourhoods' population, restaurant density and average rent were analysed in combination to come up with a short list of three recommended locations for the stakeholders.

The work can further be improved by acquiring additional data such as a consumer surveys, restaurant reviews, and foot traffic data. With the additional data at hand, one can try to build a model to predict the foot traffic of a new restaurant at a particular location.