Where to open a restaurant in Helsinki

Muzi Chen

March 9, 2021

Introduction

Helsinki is the capital city of Finland with a population of 657,674. Together with the cities of Espoo, Vantaa, and Kauniainen, and surrounding commuter towns, Helsinki forms the Greater Helsinki metropolitan area (Uusimaa), which has a population of over 1.5 million. This area is the country's most important center for politics, education, finance, culture, and research. The urbanization and development of the Uusimaa area has brought great opportunities for the tertiary sectory business, including catering.

Imagining someone is seeking for a suitable place in Helsinki to open a restaurant, he or she must be interested in how restaurants are situated in this city, the most popular types of restaurants and the surrounding venues of the restaurants, i.e. other venues in the same neighborhood.

My project will provide an analysis of the 60 neighborhoods in Helsinki area, mainly focusing on the restaurants and other popular venues in each neighborhood. The neighborhoods will finally be segmented into several clusters based on the venue situation. For those seeking for a suitable location to open a restaurant, this analysis may provide some clues.

Data

The data that will be used in this project include:

-Subdivision (neighborhoods) of Helsinki, collected from wikipedia page [1].

-The center coordinates of each neighborhood, collected from LatLong.net [2] and Google Maps [3].

-Venues in each neighborhood, collected from Foursquare API [4].

Methodology

1. Data acquisition and pre-processing

1.1 Get the Helsinki neighborhood (subdivision) data from Wikipedia page

This set of data was collected by webscraping, mainly using python library BeautifulSoup. There are in total 60 neighborhoods in Helsinki area.

1.2 Get the center coordinates of each neighborhood

The center coordinates of each neighborhood were mainly collected from LatLong.net, where batch search is possible. By uploading the neighborhoods name + “Helsinki” + “Finland” and making a query, I got the coordinates of each neighborhood. However, the same coordinates were given to several neighborhood, which is not what we want. Therefore, for those neighborhoods, I searched manually in Google Maps and got the correct coordinates for them.

1.3 Get the venue data by utilising Foursquare API

To get the venue data through Foursquare API, firstly the Foursquare API credentials should be defined. Then the API request RUL can be created, and request can be made (in this project, the limit of

returned venues for each neighborhood was set to 100). Data selected from the response includes the venues' name, latitude, longitude and category. In total, 1233 venues' information was collected, and they belong to 241 unique categories. An important step in the venue data pre-processing is one-hot encoding, so that we can have one venue category in one column. It is also important to use the groupby function to group the venue data by neighborhood name so that we have one row for one neighborhood.

2. Exploratory Data Analysis

2.1 Helsinki neighborhoods

As shown in Figure 1 below, a map was created by utilizing the folium library. Latitude and longitude values of the each neighborhood was used to define the location of the neighborhood in the map. From the map we can see that in the central region of Helsinki, the area of the each neighborhood is quite small compared with other regions. Therefore, when I made Foursquare API request, radius was set to 400 meters to ensure that different neighborhoods will not share the same venue.

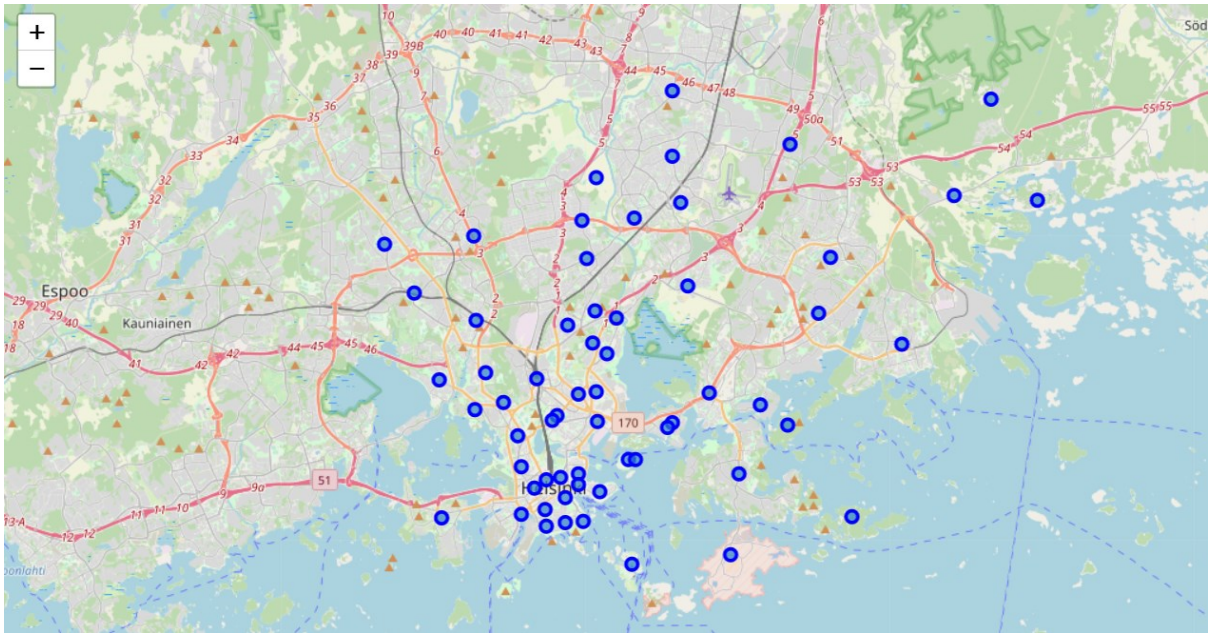

Figure 1. Helsinki neighborhood map

2.2 The total number of restaurants in each neighborhood

Since the goal of this project is to provide insights for choosing a good location for opening a restaurant, it is important to know how many restaurants there are in each neighborhood. To get the restaurant data, I filtered the venue data by selecting only columns whose names containing strings "Restaurant" "Pizza" "Sandwich" "Speakeasy" or "Food". From Figure 2 below we can see a big difference across neighborhoods regarding the number of restaurants. Kamppi, Kaartinkaupunki, Taka-Töölö, Aluemeri, Kruununhaka, Ulkosaaret, Punavuori, Vallila, Ullanlinna are neighborhoods which have a relatively high density of restaurants compared with other subdivisions.

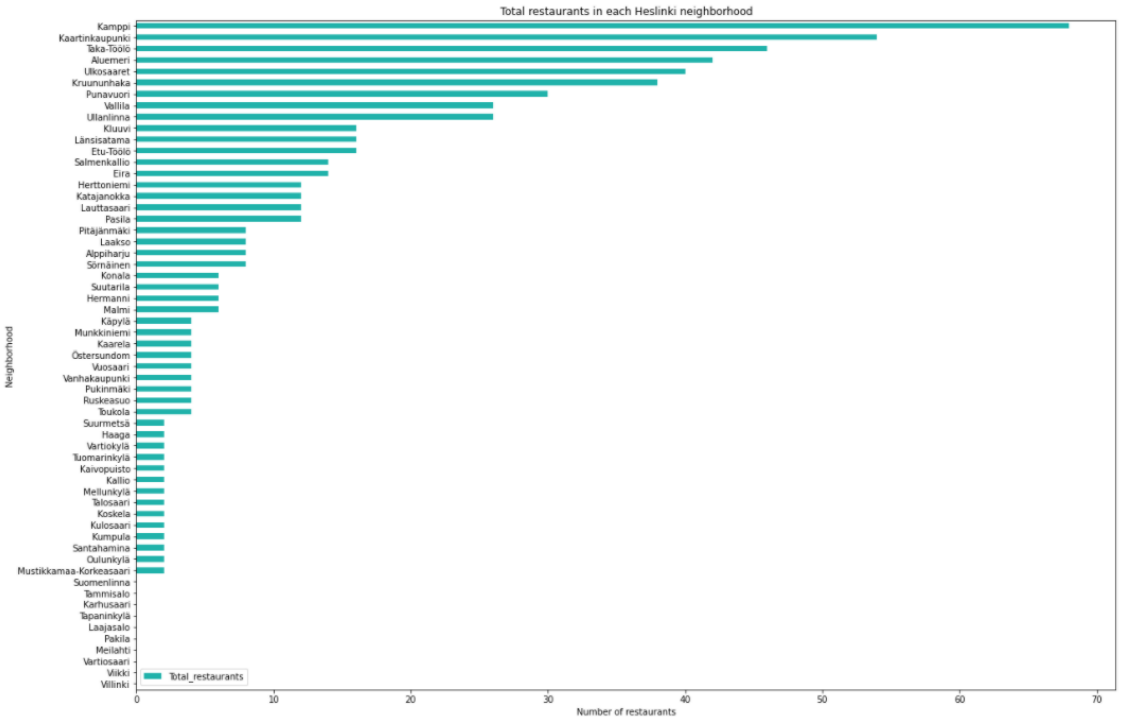

Figure 2. Number of restaurants in each neighborhood

Figure 3 below shows the distribution of the neighborhoods based on the total number of restaurants in the neighborhood. Around 40 neighborhoods have less than 10 restaurants. Around 10 neighborhoods have 10-20 restaurants. And 9 neighborhoods have more than 20 restaurants.

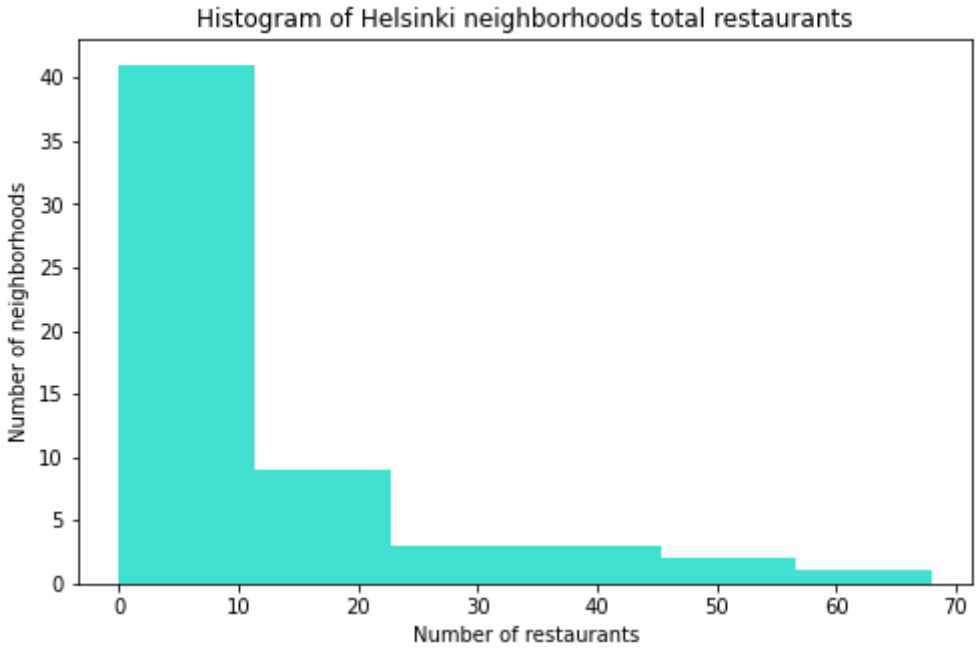

Figure 3. Distribution of neighborhoods based on number of restaurants

2.3 The number of each restaurant category in Helsinki

It might also be useful to know which types of restaurant are the most popular in Helsinki. Therefore, I calculated the number of each restaurant category (Figure 4). It seems that pizza place (51 in total) and Scandinavian restaurant (43) are the most common/popular restaurant types, followed by sushi restaurant (18), Chinese restaurant (14) and modern European restaurant (12).

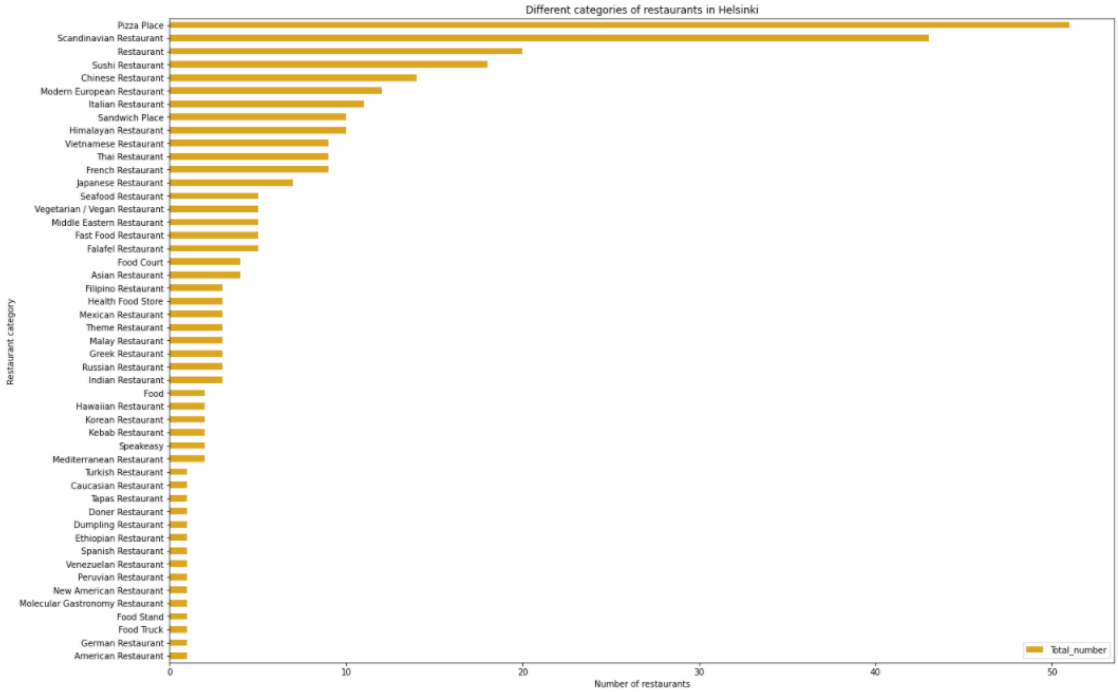

Figure 4. The number of restaurants in each category

2.4 The top 10 popular venues in Helsinki

Besides restaurants, we might also want to know the situation of other venues in Helsinki. Therefore, I calculated the number of the top 10 popular venues (see Figure 5). It seems that in Helsinki area, coffee places (in total 88 Café & Coffee Shop), bus stops, restaurants (mostly pizza places and Scandinavian Restaurant), parks, grocery stores, fitness center, hotel and boat/ferry are the most common venues.

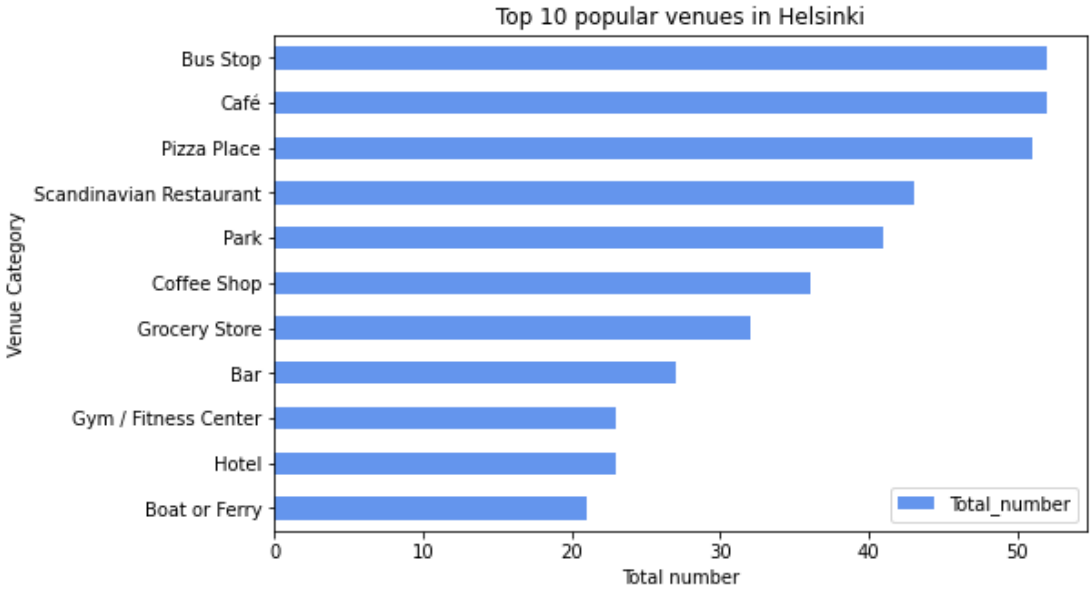

Figure 5. Top 10 popular venues in Helsinki

3. Machine learning Methods – K-means clustering algorithm

Now that we have the venues data, it might be helpful if we can divide the neighborhoods to several groups based on the venue features. Then we might be able to see which group has the best potential for catering business. Clustering would be a suitable machine learning method for this purpose. K-

means, which is the simplest unsupervised learning algorithm that solves clustering problem, was used for this project. To decide the optimal number of clusters, i.e. value of K, I applied the elbow method. Elbow method consists of plotting the explained variation as a function of the number of clusters, and picking the elbow of the curve as the number of clusters to use. Figure 6 below shows the results of the elbow method for my data. According to the figure, 3 is the best k value.

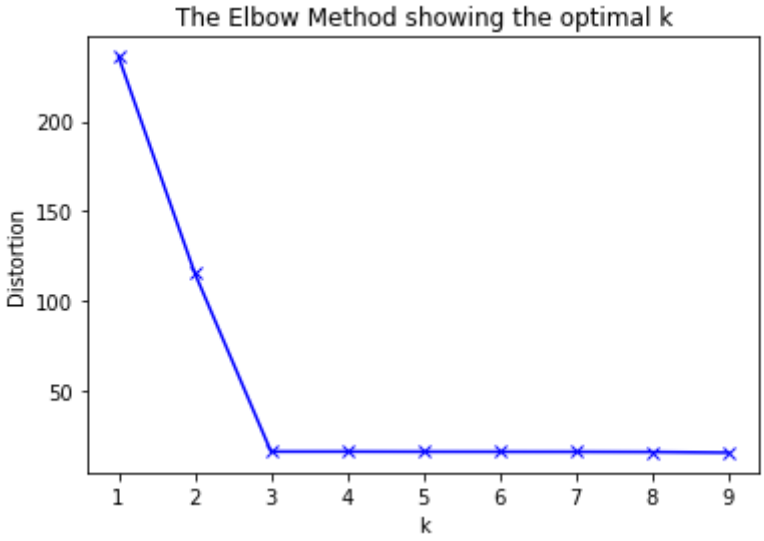

Figure 6. Result of the elbow method

Before clustering, I grouped rows of the venue data by neighborhood and by taking the mean of the frequency of occurrence of each category. Then I sorted the venues based on its frequency of each neighborhood. Next, I created a new dataframe which displays the top 10 venues of each neighborhood.

Results

The neighborhoods were segmented into three clusters as shown in the map below (Figure 7).

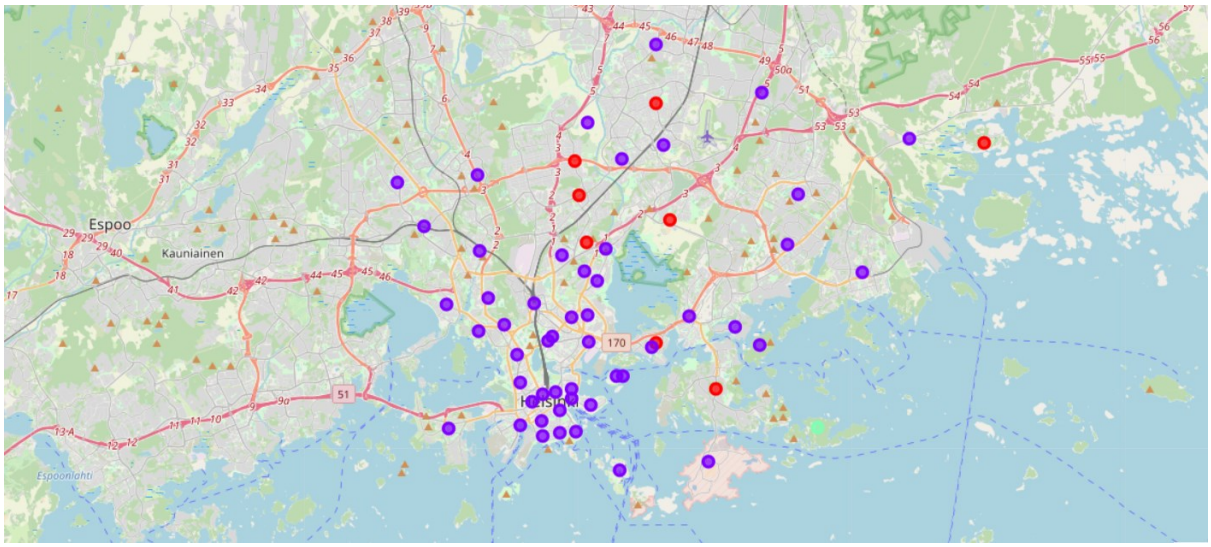

Figure 7. Neighborhood clusters map

- Cluster 1 – Red markers in the map

Table 1 below shows the eight neighborhoods which belong to the first cluster and their top 10 most common venues. Bus stop is the 1st most common venues in these neighborhoods. Restaurants only start to appear from the 3rd most common venues. I'll name this cluster as the "Transporation Cluster".

	Neighborhood	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	Cluster Labels
10	Kallio	Bus Stop	Music School	Flower Shop	Taxi Stand	Grocery Store	Chinese Restaurant	Park	Dog Run	Doner Restaurant	Filipino Restaurant	0
25	Koskela	Bus Stop	Café	Pizza Place	Grocery Store	Farm	Food Truck	Food Stand	Food Court	Food	Flower Shop	0
27	Oulunkylä	Bus Stop	Grocery Store	Soccer Field	Recreation Center	Skating Rink	Fast Food Restaurant	Curling Ice	Food	Flower Shop	Flea Market	0
33	Pakila	Bus Stop	Playground	Historic Site	Grocery Store	Film Studio	Farmers Market	Fast Food Restaurant	Filipino Restaurant	Zoo Exhibit	Falafel Restaurant	0
35	Viikki	Bus Stop	Cafeteria	Eye Doctor	Lake	Farm	Food Stand	Food Court	Food	Farmers Market	Food Truck	0
38	Tapaninkylä	Bus Stop	Gourmet Shop	Forest	Food Truck	Food Stand	Food Court	Food	Flower Shop	Flea Market	Fish & Chips Shop	0
48	Laajasalo	Bus Stop	Sauna / Steam Room	Falafel Restaurant	Food Truck	Food Stand	Food Court	Food	Flower Shop	Flea Market	Fish & Chips Shop	0
57	Karhusaari	Bus Stop	Park	Lounge	Falafel Restaurant	Food Truck	Food Stand	Food Court	Food	Flower Shop	Flea Market	0

Table 1. Cluster 1 - "Transporation Cluster"

- Cluster 2 – Purple markers in the map

Table 2 below shows part of the 51 neighborhoods which belong to the second cluster and their top 10 most common venues. Restaurants, café (or coffee shop) and bars (or pub) appear quite often among the top 3 most common venues in these neighborhoods. I'll call this cluster "Food & Drinks Cluster".

	Neighborhood	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	Cl La
0	Kruununhaka	Bar	Grocery Store	Café	Boat or Ferry	Pizza Place	Scandinavian Restaurant	Modern European Restaurant	Theater	Organic Grocery	German Restaurant	1
1	Kluuvi	Café	Scandinavian Restaurant	Clothing Store	Burger Joint	Hotel	Coffee Shop	Bistro	Park	Furniture / Home Store	Bookstore	1
2	Kaartinkaupunki	Scandinavian Restaurant	Hotel	Cocktail Bar	Vegetarian / Vegan Restaurant	Music Venue	French Restaurant	Furniture / Home Store	Bistro	Park	Café	1
3	Kamppi	Wine Bar	Coffee Shop	Scandinavian Restaurant	Chinese Restaurant	Burger Joint	Sushi Restaurant	Gym / Fitness Center	Food Court	Salon / Barbershop	Rock Club	1
4	Punavuori	Bakery	Coffee Shop	Pizza Place	Hotel	Scandinavian Restaurant	Sushi Restaurant	Park	Café	Gay Bar	Furniture / Home Store	1
5	Eira	Bakery	Park	Harbor / Marina	Boat or Ferry	Scandinavian Restaurant	Café	Mexican Restaurant	Beach	Modern European Restaurant	Turkish Restaurant	1
6	Ullanlinna	Scandinavian Restaurant	Grocery Store	Park	Ice Cream Shop	Coffee Shop	Plaza	Pharmacy	Nightclub	Boat or Ferry	Restaurant	1
7	Katajanokka	Boat or Ferry	Hotel	Park	Scandinavian Restaurant	Tram Station	Pub	Harbor / Marina	Film Studio	Café	Plaza	1
8	Kaivopuisto	Boat or Ferry	Park	Harbor / Marina	Monument / Landmark	Tram Station	Nightclub	Taxi Stand	Grocery Store	Scandinavian Restaurant	Coffee Shop	1
9	Sörnäinen	Thai Restaurant	Circus School	Cocktail Bar	Café	Dance Studio	Pizza Place	Burger Joint	Martial Arts School	Vietnamese Restaurant	Organic Grocery	1
11	Alppiharju	Theme Park Ride / Attraction	Park	Bar	Greek Restaurant	Café	Thai Restaurant	Dog Run	Vietnamese Restaurant	History Museum	Theater	1
12	Etu-Töölö	Scandinavian Restaurant	Pub	Bookstore	Russian Restaurant	Plaza	Indie Theater	Asian Restaurant	Gym / Fitness Center	Restaurant	Bakery	1
13	Taka-Töölö	Coffee Shop	Sushi Restaurant	Music Venue	French Restaurant	Bowling Alley	Indian Restaurant	Himalayan Restaurant	Italian Restaurant	Japanese Restaurant	Pizza Place	1
14	Meilahti	Gym / Fitness Center	Disc Golf	Tennis Court	Zoo Exhibit	Falafel Restaurant	Food Stand	Food Court	Food	Flower Shop	Flea Market	1

Table 2. Cluster 2 – “Food & Drinks Cluster”

- Cluster 3 – Green markers in the map

There is only one neighborhood in the third cluster as shown in Table 3 below. The top 2 most common venues are memorial site and zoo exhibit. I’ll name this cluster “Memorial Site & Zoo Cluster”.

	Neighborhood	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	Cluster Labels
49	Villinki	Memorial Site	Zoo Exhibit	French Restaurant	Food Truck	Food Stand	Food Court	Food	Flower Shop	Flea Market	Fish & Chips Shop	2

Table 3. Cluster 3 – “Memorial Site & Zoo Cluster”

Discussion

Though venues data collected through Foursquare API does not reflect the full picture of venues in Helsinki, we can still have a rough overview of the restaurants situation in Helsinki from the above exploratory analysis and cluster results. Neighborhoods in the second cluster, i.e. “Food & Drinks Cluster”, have a higher density of restaurants compared with other neighborhoods. And where there are many restaurants, there are often many cafés and bars. If we pay attention in our daily life, we’ll notice that people are more likely to visit a restaurant if it is close to other restaurants. Because sometimes we might not know what to eat, and we’d probably go to a place where there are many

restaurants so that we can have a wider range of choices. As a brand-new restaurant, it is important to increase the chances that it appears in people's eyes. The simplest way is to place this restaurant among other restaurants. Competition also brings opportunities. Therefore, I would recommend people to open the restaurant in one of the "Food & Drinks Cluster" neighborhoods.

In addition, based on my analysis, pizza place, Scandinavian restaurant, sushi restaurant, Chinese restaurant and modern European restaurant are the top 5 most common types of restaurant in Helsinki. There would be enough demands in the market if the new restaurant is one of the above five categories. However, if someone plans to open an other type of restaurant, such as Indian restaurant, he or she might need to think more.

We might also need to consider the cost of opening the restaurant, including the rental cost of the place. This is something that we can improve in the future. We can get the housing price (per square meter) of the neighborhoods and add to our analysis. Housing prices in a way reflect the renting prices.

Conclusion

To open a brand-new restaurant, I would suggest choosing a location where there are many other restaurants. Therefore, the neighborhoods in the "Food & Drinks Cluster" should be considered first. We also need to consider if people will like the type of food our restaurant serves. It might be a safer choice to serve pizza, Scandinavian food, sushi, Chinese food and modern European food.

Reference

[1] Subdivisions of Helsinki, Wikipedia. https://en.wikipedia.org/wiki/Subdivisions_of_Helsinki

[2] LatLong.net. https://www.latlong.net/

[3] Google Maps. https://www.google.com/maps

[4] Foursquare developer page https://developer.foursquare.com/