Capstone Project - The Battle of Neighborhoods

Clustering the Neighborhoods of Amsterdam, The Netherlands

by Kinga A. Lőrincz

1. Business problem

An expat housing company would like to be able to quickly recommend neighborhoods in Amsterdam to their customers based on their preferences for different venues.

Expats moving to a foreign country for a longer period often prefer to move to neighborhoods that remind them to their own neighborhood. Additionally, they might want their neighborhood to fit to their lifestyle: e.g. in general, young families prefer quieter neighborhoods with lots of parks and playgrounds, health conscious customers prefer neighborhoods with lost parks and lots gyms to choose from. Young professionals are looking for areas with lots of restaurants in the neighborhood.

2. Data

To explore the neighborhoods of Amsterdam and obtain the most common venue categories in each neighborhood I have used the Foursquare location data. I have used the Foursquare API called *explore* to obtain the 100 most common venues in each neighborhood in a radius of 1000 m from the center of the neighborhood.

Amsterdam has 8 districts (boroughs) and in total 481 neighborhoods. These neighborhoods are very close to each other, so I decided to use the aggregated neighborhoods in our analysis. There are 99 aggregated neighborhoods called quarters.

The geo-coordinates (latitude and longitude) of the neighborhoods of Amsterdam are freely available on the web: https://maps.amsterdam.nl/open_geodata/?LANG=en It can be downloaded in different formats, e.g. as a *csv* file or *json* file.

3. Methodology

3.1 Data cleaning

I have imported the data describing neighborhoods from the *csv* file into the pandas dataframe, `amsterdam_data`. The dataframe has 99 rows and 10 columns. The columns contain the features of the neighborhood: an object number (index), the code of the neighborhood, the name of the neighborhood, the code of the district, the area in m^2, WKT_LNG_LAT and WKT_LAT_LNG giving the vector geometry of the neighborhoods on the map, the longitude and latitude of the
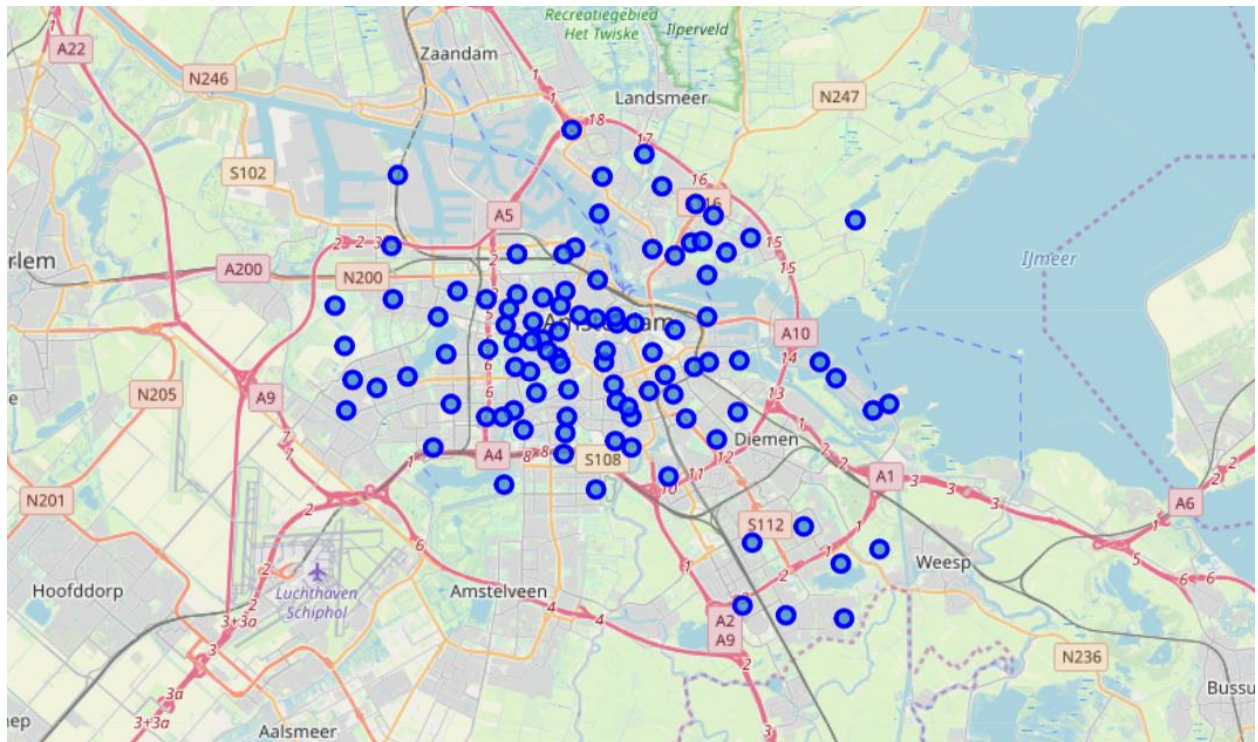
neighborhood and an unnamed:9 column which contains NaN values. The column names are in Dutch.

First, I have removed all the features I did not need for this analysis and kept only the name of the neighborhood and the geo-coordinates (longitude and latitude) and translated the remaining column names to English. Next, I have fixed the problem with the latitude and longitude columns: changed the decimal separator from comma to dot and change the datatype to float. I show the first 5 rows of the `amsterdam_data` dataframe below.

	Neighborhood	LNG	LAT
0	Amstel III/Bullewijk	4.950592	52.298741
1	Bijlmer Centrum (D,F,H)	4.954527	52.315304
2	Gein	4.993931	52.295510
3	Kinkerbuurt	4.865445	52.368083
4	Overtoomse Sluis	4.859836	52.359725

3.2. Visualize the neighborhoods

The following picture shows the neighborhoods superimposed on the map of Amsterdam. I have used the `Folium` library to generate the map. I have obtained the latitude and longitude values of Amsterdam using the `geopy` library.

Each blue dot on the map corresponds to one neighborhood.

3.3. Foursquare location data

I have defined and used the function `getNearbyVenues,` which loops through all the neighborhoods from the `amsterdam_data` dataframe and calls the Foursquare API *explore* to obtain the most common venues in each neighborhood. I chose to explore a radius of 1000 m from the center of the neighborhood to obtain a relevant number of venues in each neighborhood.

The resulting venues are stored in the `amsterdam_venues` dataframe. In addition to the venue name, I also store the venue latitude and longitude and venue category. In total there were 6557 venues found. I show the first 5 rows of the `amsterdam_venues` dataframe in the following table.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Amstel III/Bullewijk	52.298741	4.950592	Hotel Lowell	52.304526	4.950671	Hotel
1	Amstel III/Bullewijk	52.298741	4.950592	De Proefzaak	52.303585	4.950890	Café
2	Amstel III/Bullewijk	52.298741	4.950592	Fitness365	52.298421	4.945692	Gym
3	Amstel III/Bullewijk	52.298741	4.950592	Museum Vrolik	52.293968	4.960014	Science Museum
4	Amstel III/Bullewijk	52.298741	4.950592	Joy	52.305268	4.950310	Hotel

Grouping the venues by neighborhoods I have observed that some of the neighborhoods have a very small number of venues found in them. I have decided to remove the neighborhoods that have less than 10 venues found in them. In this instance there were 4 such neighborhoods.

Neighborhood	Venue	Neighborhood	Venue	Neighborhood	Venue
Amstel III/Bullewijk	46	De Weteringschans	100	IJburg West	29
Apollobuurt	100	Driemond	12	IJburg Zuid	42
Banne Buiksloot	19	Elzenhagen	40	IJplein/Vogelbuurt	55
Betondorp	29	Erasmuspark	100	IJselbuurt	100
Bijlmer Centrum (D,F,H)	95	Frankendael	100	Indische Buurt Oost	51
Bijlmer Oost (E,G,K)	17	Frederik Hendrikbuurt	100	Indische Buurt West	100
Buikslotermeer	37	Gein	12	Jordaan	100
Buitenveldert-Oost	77	Geuzenbuurt	100	Kadoelen	18
Buitenveldert-West	21	Geuzenveld	22	Kinkerbuurt	100
Burgwallen-Nieuwe Zijde	100	Grachtengordel-West	100	Landlust	49
Burgwallen-Oude Zijde	100	Grachtengordel-Zuid	100	Lutkemeer/Ookmeer	13
Centrale Markt	100	Haarlemmerbuurt	100	Middelveldsche Akerploder	24
Chassébuurt	100	Helmersbuurt	100	Middenmeer	35
Da Costabuurt	100	Holendrecht/Reigersbos	20	Museumkwartier	100
Dapperbuurt	100	Hoofddorppleinbuurt	100	Nellestein	16
De Kolenkit	69	Hoofdweg e.o.	100	Nieuwe Pijp	100
De Punt	37	Houthavens	91	Nieuwendammerdijk/Buiksloterdijk	51
		IJburg Oost	25	Nieuwmarkt/Lastage	100

Neighborhood		Neighborhood		Neighborhood	
Noordelijke IJ-oevers Oost	43	Schinkelbuurt	100	Tuindorp Oostzaan	23
Noordelijke IJ-oevers West	33	Sloter-/Riekerpolder	36	Van Galenbuurt	100
Omval/Overamstel	47	Sloterdijk	42	Van Lennepbuurt	100
Oostelijk Havengebied	34	Slotermeer-Noordoost	30	Volewijck	63
Oostelijke Eilanden/Kadijken	100	Slotermeer-Zuidwest	51	Vondelbuurt	100
Oosterparkbuurt	100	Slotervaart Noord	36	Waterlandpleinbuurt	21
Oostzanerwerf	17	Slotervaart Zuid	40	Weesperbuurt/Plantage	100
Osdorp-Midden	71	Spaarndammer- en Zeeheldenbuurt	93	Weesperzijde	100
Osdorp-Oost	52			Westindische Buurt	100
Oude Pijp	100	Staatsliedenbuurt	100	Westlandgracht	72
Overtoomse Sluis	100	Stadionbuurt	93	Willemspark	100
Overtoomse Veld	70	Transvaalbuurt	100	Zeeburgereiland/Nieuwe Diep	27
Prinses Irenebuurt e.o.	87	Tuindorp Buiksloot	63	Zuid Pijp	100
Rijnbuurt	98	Tuindorp Nieuwendam	40	Zuidas	97

In the remaining 95 neighborhoods there are 315 unique venue categories.

3.4. Analyze the neighborhoods

To analyze the neighborhoods, I have created a new dataframe `amsterdam_grouped` consisting of as many rows as the number of neighborhoods (95) and as many columns (features) as unique venue categories. I have grouped the venues by neighborhoods by taking the mean frequency of occurrence of each venue category. The first 5 rows and a small sample of features (10) of the `amsterdam_grouped` dataframe are shown in the table below.
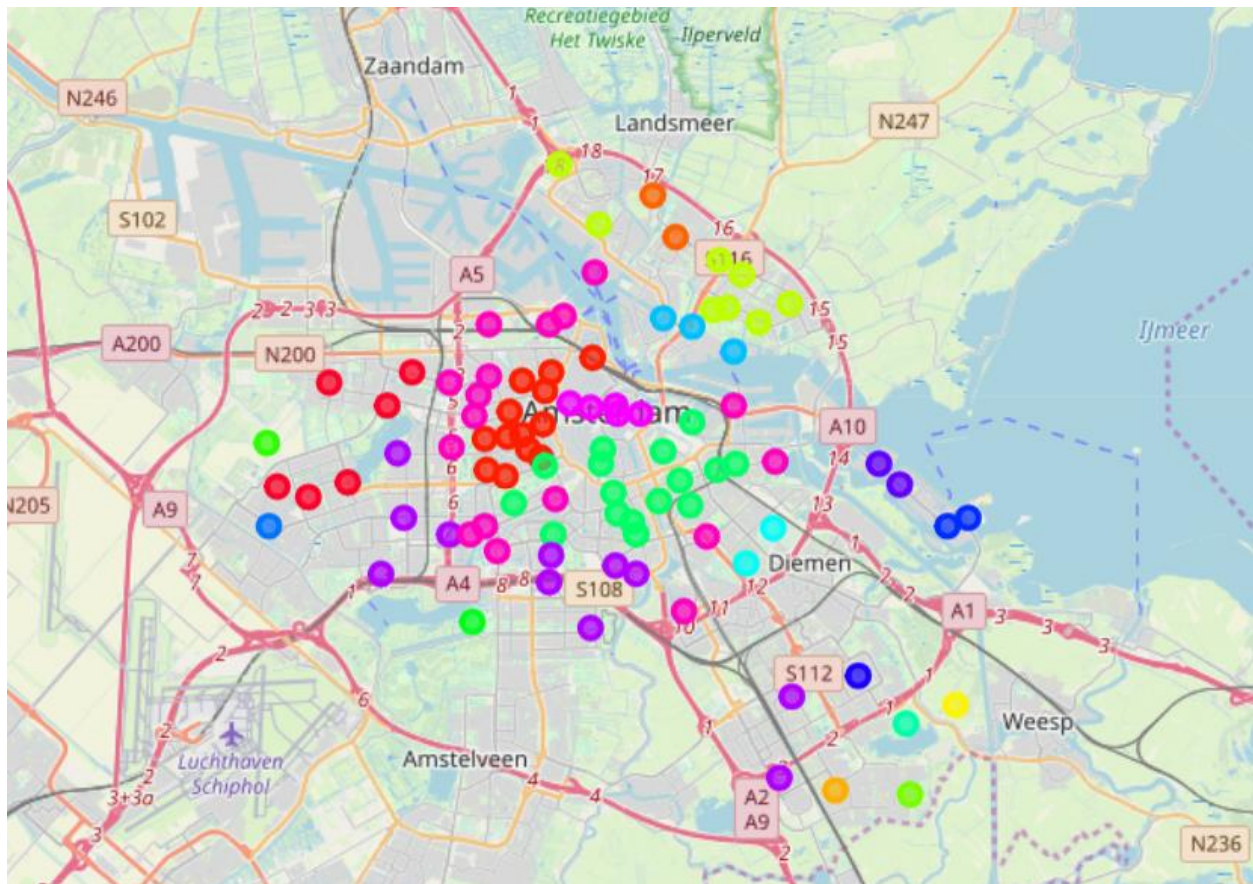
	Neighborhood	Asian Restaurant	Athletics & Sports	Australian Restaurant	Auto Workshop	BBQ Joint	Baby Store	Bagel Shop	Bakery	Bar	Baseball Field
0	Amstel III/Bullewijk	0.021739	0.000000	0.0	0.0	0.0	0.0	0.00	0.000000	0.000000	0.0
1	Apollobuurt	0.010000	0.000000	0.0	0.0	0.0	0.0	0.02	0.070000	0.030000	0.0
2	Banne Buiksloot	0.000000	0.000000	0.0	0.0	0.0	0.0	0.00	0.000000	0.000000	0.0
3	Betondorp	0.000000	0.034483	0.0	0.0	0.0	0.0	0.00	0.034483	0.000000	0.0
4	Bijlmer Centrum (D,F,H)	0.021053	0.000000	0.0	0.0	0.0	0.0	0.00	0.031579	0.021053	0.0

Next, I have used the `Kmeans` machine learning algorithm to cluster the neighborhoods. I chose to set the number of clusters to 20 to have a finer classification of the neighborhoods. Finally, I have constructed the `amsterdam_merged` dataframe which contains the name of the

neighborhoods, the geocoordinates (latitude and longitude), the cluster labels and the 10 most common venues.

4. Results

First, we visualize the clusters of neighborhoods superimposed on the map of Amsterdam. The map is generated using the `Folium` library. Here the neighborhoods are color coded according to the cluster labels. Neighborhoods represented by the same color belong to the same cluster.

I have also displayed the neighborhoods belonging to the same cluster together with the 7 most common venue categories in the tables below. In the notebook you can see the 10 most common categories.

Cluster 0 - Contains 17 neighborhoods. It is represented with ⬤ on the map.

	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
5	Erasmuspark	Coffee Shop	Supermarket	Restaurant	Bar	Café	Fast Food Restaurant	Bakery
29	Indische Buurt Oost	Restaurant	Café	Coffee Shop	Supermarket	Bakery	Playground	Bar
36	Sloterdijk	Restaurant	Hotel	Music Venue	Farm	Bistro	Supermarket	Café
39	Houthavens	Restaurant	Coffee Shop	Indie Movie Theater	Italian Restaurant	Café	Pizza Place	French Restaurant
45	Schinkelbuurt	Restaurant	Bar	Café	Coffee Shop	Hotel	Gym / Fitness Center	Athletics & Sports
57	Spaarndammer- en Zeeheldenbuurt	Restaurant	Coffee Shop	Café	Music Venue	Theater	Indie Movie Theater	Snack Place
60	De Kolenkit	Turkish Restaurant	Park	Restaurant	Hotel	Fast Food Restaurant	Café	Coffee Shop
68	Hoofddorppleinbuurt	Restaurant	Coffee Shop	Café	Bar	Bakery	Gym / Fitness Center	Athletics & Sports
75	Landlust	Café	Park	Restaurant	Coffee Shop	Bakery	Sandwich Place	Hostel
77	Stadionbuurt	Restaurant	Bakery	Café	Bar	Coffee Shop	Snack Place	Athletics & Sports
79	Overtoomse Veld	Restaurant	Supermarket	Tram Station	Hotel	Bar	Coffee Shop	Ice Cream Shop
83	Frankendael	French Restaurant	Coffee Shop	Restaurant	Italian Restaurant	Bar	Soccer Field	Café
84	Oostelijk Havengebied	Restaurant	Café	Music Venue	Deli / Bodega	French Restaurant	Bar	Supermarket
85	Van Galenbuurt	Restaurant	Coffee Shop	Supermarket	Bar	Ice Cream Shop	Bakery	Fast Food Restaurant
91	Museumkwartier	Restaurant	Art Museum	Coffee Shop	French Restaurant	Bakery	Hotel	Men's Store
96	Omval/Overamstel	Restaurant	Hotel	Gym / Fitness Center	Supermarket	Athletics & Sports	Dance Studio	Diner
97	Noordelijke IJ-oevers West	Restaurant	Hotel	Steakhouse	Bar	Harbor / Marina	Furniture / Home Store	Flea Market

Cluster1 - Contains 6 neighborhoods. It is represented with ⬤ on the map.

	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
18	Osdorp-Oost	Turkish Restaurant	Supermarket	Snack Place	Clothing Store	Tram Station	Shopping Mall	Chinese Restaurant
19	De Punt	Supermarket	Fast Food Restaurant	Tram Station	Bakery	Turkish Restaurant	Clothing Store	Hotel
24	Slotermeer-Noordoost	Hotel	Turkish Restaurant	Park	Dessert Shop	Other Repair Shop	Gym	Supermarket
25	Geuzenveld	Bus Stop	Supermarket	Park	Tram Station	Italian Restaurant	Furniture / Home Store	Fast Food Restaurant
41	Slotermeer-Zuidwest	Turkish Restaurant	Tram Station	Park	Dessert Shop	Café	Plaza	Discount Store
42	Osdorp-Midden	Supermarket	Tram Station	Snack Place	Clothing Store	Turkish Restaurant	Bakery	Gym / Fitness Center

Cluster 2 - Contains 13 neighborhoods. It is represented with 🔴 on the map.

	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
3	Kinkerbuurt	Bar	Italian Restaurant	Coffee Shop	Restaurant	Yoga Studio	Grocery Store	Vegetarian / Vegan Restaurant
4	Overtoomse Sluis	Restaurant	Bar	Café	Bistro	Coffee Shop	Ethiopian Restaurant	Yoga Studio
6	Hoofdweg e.o.	Bar	Restaurant	Italian Restaurant	Coffee Shop	Pizza Place	Deli / Bodega	Bistro
17	Haarlemmerbuurt	Bar	Restaurant	Marijuana Dispensary	Italian Restaurant	Café	Bakery	Sandwich Place
21	Centrale Markt	Italian Restaurant	Coffee Shop	Bar	Café	Restaurant	Ice Cream Shop	Indonesian Restaurant
32	Da Costabuurt	Bar	Coffee Shop	Italian Restaurant	Pizza Place	Café	Restaurant	Yoga Studio
33	Helmersbuurt	Bar	Café	Coffee Shop	Hotel	Italian Restaurant	Hostel	Restaurant
34	Staatsliedenbuurt	Bar	Italian Restaurant	Coffee Shop	Café	Thai Restaurant	Bistro	Sandwich Place
35	Geuzenbuurt	Bar	Italian Restaurant	Coffee Shop	Restaurant	Café	Pizza Place	Hotel
38	Westindische Buurt	Restaurant	Bar	Ethiopian Restaurant	Café	Vegetarian / Vegan Restaurant	Bistro	Breakfast Spot
55	Frederik Hendrikbuurt	Bar	Italian Restaurant	Coffee Shop	Café	Yoga Studio	Ice Cream Shop	Thai Restaurant
58	Van Lennepbuurt	Coffee Shop	Café	Bar	Restaurant	Vegetarian / Vegan Restaurant	Italian Restaurant	Yoga Studio
93	Chassébuurt	Bar	Coffee Shop	Restaurant	Italian Restaurant	Yoga Studio	Café	Grocery Store

Cluster 3 - Contains 2 neighborhoods. It is represented with 🟠 on the map.

	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
12	Kadoelen	Bus Stop	Soccer Field	Tennis Court	Park	Shopping Mall	Sports Club	Stables
70	Banne Buiksloot	Bus Stop	Soccer Field	Tennis Court	Grocery Store	Farm	Sports Club	Stables

Cluster 4 - Contains 1 neighborhood. It is represented with 🟡 on the map.

	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
72	Holendrecht/Reigersbos	Pizza Place	Supermarket	Chinese Restaurant	Shopping Mall	Grocery Store	Train Station	Snack Place

Cluster 5 - Contains 1 neighborhood. It is represented with ⬤ on the map.

	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
16	Driemond	Restaurant	Athletics & Sports	Plaza	Monument / Landmark	Convention Center	Hotel	Print Shop

Cluster 6 - Contains 8 neighborhoods. It is represented with ⬤ on the map.

	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
10	Tuindorp Nieuwendam	Bus Stop	Supermarket	Park	Harbor / Marina	Snack Place	Indonesian Restaurant	Electronics Store
11	Tuindorp Buiksloot	Bus Stop	Park	Clothing Store	Bakery	Supermarket	Harbor / Marina	Hotel
13	Elzenhagen	Bus Stop	Supermarket	Clothing Store	Electronics Store	Soccer Field	Fast Food Restaurant	Breakfast Spot
52	Nieuwendammerdijk/Buiksloterdijk	Bus Stop	Clothing Store	Supermarket	Bakery	Electronics Store	Sporting Goods Shop	Flea Market
53	Tuindorp Oostzaan	Bus Stop	Soccer Field	Chinese Restaurant	Shopping Mall	Shop & Service	Flea Market	Sandwich Place
67	Oostzanerwerf	Bus Stop	Supermarket	Soccer Field	Flower Shop	Park	Shop & Service	Dutch Restaurant
69	Waterlandpleinbuurt	Bus Stop	Athletics & Sports	Supermarket	Plaza	Sandwich Place	Seafood Restaurant	Park
80	Buikslotermeer	Bus Stop	Supermarket	Park	Clothing Store	Soccer Field	Shopping Mall	Market

Cluster 7 - Contains 1 neighborhood. It is represented with ⬤ on the map.

	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
2	Gein	Lake	Playground	Gym	Smoke Shop	Metro Station	Supermarket	Bakery

Cluster 8 - Contains 1 neighborhood. It is represented with ⬤ on the map.

	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
62	Lutkemeer/Ookmeer	Clothing Store	Snack Place	Toy / Game Store	Fast Food Restaurant	Sandwich Place	Pet Store	BBQ Joint

Cluster 9 - Contains 1 neighborhood. It is represented with ⬤ on the map

	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
48	Buitenveldert-West	Restaurant	Harbor / Marina	Tennis Court	Flower Shop	Spa	Garden	Dog Run

Cluster 10 - Contains 16 neighborhoods. It is represented with ● on the map.

	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
7	De Weteringschans	Café	Coffee Shop	Art Museum	Steakhouse	Clothing Store	Hotel	French Restaurant
8	Oostelijke Eilanden/Kadijken	Zoo Exhibit	Restaurant	Bar	Café	Museum	Hotel	Gastropub
28	Transvaalbuurt	Café	Coffee Shop	Hotel	French Restaurant	Italian Restaurant	Restaurant	Bar
30	Nieuwe Pijp	Italian Restaurant	Japanese Restaurant	Pizza Place	Seafood Restaurant	Bar	Breakfast Spot	Salad Place
44	Oude Pijp	Italian Restaurant	Pizza Place	Hotel	Bar	Seafood Restaurant	French Restaurant	Breakfast Spot
46	Apollobuurt	Hotel	Restaurant	Bakery	French Restaurant	Plaza	Coffee Shop	Italian Restaurant
47	IJselbuurt	Italian Restaurant	Ice Cream Shop	Café	Restaurant	Hotel	Pizza Place	Breakfast Spot
49	Oosterparkbuurt	Hotel	Restaurant	Italian Restaurant	Café	Coffee Shop	Bar	Breakfast Spot
50	Indische Buurt West	Restaurant	Bar	Café	Coffee Shop	Turkish Restaurant	Vegetarian / Vegan Restaurant	Italian Restaurant
65	Dapperbuurt	Bar	Coffee Shop	Café	Turkish Restaurant	Italian Restaurant	French Restaurant	Vegetarian / Vegan Restaurant
71	Willemspark	Restaurant	Hotel	Bar	Coffee Shop	Breakfast Spot	Café	Bistro
73	Zuid Pijp	Italian Restaurant	Restaurant	Pizza Place	Breakfast Spot	Hotel	Japanese Restaurant	Coffee Shop
76	Grachtengordel-Zuid	Sandwich Place	Café	Restaurant	Clothing Store	Hotel	Italian Restaurant	Coffee Shop
78	Vondelbuurt	Coffee Shop	Bar	Café	Hotel	Restaurant	Art Museum	Music Venue
86	Weesperbuurt/Plantage	Zoo Exhibit	Bar	Hotel	Restaurant	French Restaurant	Coffee Shop	History Museum
94	Weesperzijde	Italian Restaurant	Hotel	Café	Coffee Shop	Yoga Studio	Restaurant	French Restaurant

Cluster 11 - Contains 1 neighborhood. It is represented with ● on the map.

	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
54	Nellestein	Restaurant	Lake	Playground	Hotel	Park	Stables	Bus Station

Cluster 12 - Contains 2 neighborhoods. It is represented with ● on the map.

	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
27	Betondorp	Soccer Field	Café	Stadium	Scandinavian Restaurant	Supermarket	Sports Bar	Flower Shop
92	Middenmeer	Café	Soccer Field	Fast Food Restaurant	Coffee Shop	Bus Stop	Train Station	Restaurant

Cluster 13 - Contains 3 neighborhoods. It is represented with ⬤ on the map.

	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
9	Volewijck	Café	Furniture / Home Store	Bakery	Park	Sporting Goods Shop	Market	Bus Stop
90	IJplein/Vogelbuurt	Café	Bus Stop	Park	Turkish Restaurant	Brewery	Restaurant	Market
95	Noordelijke IJ-oevers Oost	Bus Stop	Café	Music Venue	Bar	Restaurant	Harbor / Marina	Furniture / Home Store

Cluster 14 - Contains 1 neighborhood. It is represented with ⬤ on the map.

	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
74	Middelveldsche Akerploder	Tram Station	Trail	Hotel	Museum	Music Store	Spa	Drugstore

Cluster 15 - Contains 1 neighborhood. It is represented with ⬤ on the map.

	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
62	Lutkemeer/Ookmeer	Clothing Store	Discount Store	Turkish Restaurant	Café	Fast Food Restaurant	Electronics Store	Sandwich Place

Cluster 16 - Contains 1 neighborhood. It is represented with ⬤ on the map

	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
15	Bijlmer Oost (E,G,K)	Supermarket	Bar	Bakery	Metro Station	Bus Stop	Discount Store	Park

Cluster 17 - Contains 2 neighborhoods. It is represented with ⬤ on the map.

	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
26	IJburg West	Farmers Market	Italian Restaurant	Tram Station	Park	Shopping Mall	Asian Restaurant	Supermarket
51	Zeeburgereiland/Nieuwe Diep	Farmers Market	Harbor / Marina	Hotel	Beach	Supermarket	Bookstore	Soccer Field

Cluster 18 - Contains 11 neighborhoods. It is represented with ⬤ on the map.

	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
0	Amstel III/Bullewijk	Hotel	Coffee Shop	Restaurant	Fast Food Restaurant	Food Court	Café	Furniture / Home Store
1	Bijlmer Centrum (D,F,H)	Coffee Shop	Office	Chinese Restaurant	Restaurant	Park	Bakery	Hotel
22	Westlandgracht	Supermarket	Coffee Shop	Café	Restaurant	Gym	Hotel	Gym / Fitness Center
23	Zuidas	Coffee Shop	Hotel	Supermarket	Restaurant	Gym	Italian Restaurant	Breakfast Spot
31	Rijnbuurt	Café	Italian Restaurant	Supermarket	Restaurant	Gym / Fitness Center	Hotel	Japanese Restaurant
43	Slotervaart Zuid	Coffee Shop	Tram Station	Supermarket	Plaza	Chinese Restaurant	Asian Restaurant	Convenience Store
61	Slotervaart Noord	Supermarket	Liquor Store	Pharmacy	Gym	Beach	Moroccan Restaurant	Pool
63	Prinses Irenebuurt e.o.	Hotel	Italian Restaurant	Restaurant	Plaza	Japanese Restaurant	Salad Place	Breakfast Spot
64	Buitenveldert-Oost	Hotel	Korean Restaurant	Gym	Gym / Fitness Center	Sandwich Place	Coffee Shop	Playground
82	Scheldebuurt	Japanese Restaurant	Restaurant	Italian Restaurant	Café	Hotel	Ice Cream Shop	Supermarket
88	Sloter-/Riekerpolder	Soccer Field	Hotel	Tram Station	Bus Stop	Café	Coffee Shop	Athletics & Sports

Cluster 19 - Contains 5 neighborhoods. It is represented with ⬤ on the map.

	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
14	Burgwallen-Oude Zijde	Hotel	Bar	French Restaurant	Marijuana Dispensary	Café	Coffee Shop	Thai Restaurant
37	Burgwallen-Nieuwe Zijde	Hotel	Bar	Marijuana Dispensary	Café	Cocktail Bar	French Restaurant	Restaurant
56	Grachtengordel-West	Hotel	Bar	Bookstore	Café	Sandwich Place	French Restaurant	Cheese Shop
59	Nieuwmarkt/Lastage	Bar	Hotel	Pub	Dessert Shop	Coffee Shop	Thai Restaurant	History Museum
87	Jordaan	Hotel	Bar	Restaurant	Pizza Place	Cheese Shop	Café	French Restaurant

5. Discussion

The clustering analysis gives a good classification of the neighborhoods of Amsterdam. We observe that some of the neighborhoods belonging to the same cluster are situated in each other's vicinity, others are situated further apart from each other.

Looking at the most common venue categories of each cluster it is possible to make recommendation to customers based on their preferences. For example, cluster 0 is dominated by the presence of restaurants, cafes, bars, and entertainment (theater, museum, music venues). These neighborhoods would be generally suitable for young professionals. The most common venue category of Cluster 1 is the Turkish restaurant, so I would certainly recommend it to customers with Turkish background. The most common venue categories for cluster 2 are bars and cafes, for cluster 3 sport areas, for cluster 4 hotels and bars, for cluster 6 bus stops, shops and sport venues, cluster 10 Italian restaurants, for cluster 13 café, Turkish restaurant and bus stop, for cluster 18 shops and restaurants and for cluster 19 hotels, bars and an abundance of marijuana dispensaries. These are the clusters with more than 2 neighborhoods in them.

There are also a couple of clusters with only one neighborhood in them. I have tried to change the number of clusters in the classification algorithm, but it does not help. These are the neighborhoods with a very limited amount of Foursquare location data.

Although I feel that the clustering of the neighborhoods of Amsterdam gives a lot of information to the housing company, this can be greatly enhanced with improved location data. The problems with the current location data are: the limited amount of venues given for each neighborhood (max 100 even if the limit in the *explore* API is increased) and the lack of registered venues on Foursquare for less touristic neighborhoods.

6. Conclusion

Using the freely available geocoordinates of Amsterdam and the Foursquare location data I have obtained a nice clustering of the neighborhoods of Amsterdam. These clusters can be used to recommend neighborhoods to the customers of the expat housing company. These recommendations can be made using the preferences of the customers for different venues. If these preferences are not known the clustering analysis can be used to recommend neighborhoods based on the general preference of customers with the same age range, family status, education, work etc.

This analysis can be improved using better location data and can be expended to other cities.