Capstone Project - The Battle of Neighborhoods

1. Introduction and problem

This project is a part of IBM data science, which we are going to explore the data of Helsinki, Finland, where I am staying currently. Helsinki is the capital and most populous city in Finland and has not only its own unique vibe of a charming Scandinavian city but also a contemporary feel.

In 2017, according to , the number of overnight stays in Helsinki was about 5.3 million for both domestics and international visitors. Helsinki's main attractions are urban culture with the diversity in events and foods. In this final project, we are going to focus on exploring Helsinki in different categories in all areas to observe what people have been enjoyed when they come to Helsinki using Foursquare.

2. Data preparation

2.1 Getting Helsinki Neighborhoods Information

Since we are going to clustering the data, having the neighborhood information helps us in grouping the data into each district in Helsinki. Using Google and Beautifulsoup4 library in Python, a data-frame contains the name of 51 subdivisions in the capital of Finland is created.

	Neighborhood		Neighborhood
0	Kamppi		
1	Suomenlinna	26	Sompasaari
2	Munkkiniemi	27	Meilahti
3	Herttoniemi	28	Kumpula
4	Vuosaari	29	Tapanila
5	Itäkeskus	30	Kruununhaka
6	Hernesaari	31	Haaga
7	Ullanlinna	32	Arabianranta
8	Kallio	33	Eira
9	Vallila	34	Sörnäinen
10	Töölö	35	Maunula
11	Kalasatama	36	Pukinmäki
12	Katajanokka	37	Mellunmäki
13	Malmi	38	Jollas
14	Kontula	39	Kluuvi
15	Kuninkaantammi	40	Pihlajamäki
16	Pitäjänmäki	41	Jakomäki
17	Oulunkylä	42	Tapulikaupunki
18	Kannelmäki	43	Rastila
19	Myllypuro	44	Puotila
20	Malminkartano	45	Munkkivuori
21	Käpylä	46	Merihaka
22	Viikki	47	Pikku Huopalahti
23	Konala	48	Pakila
24	Ruoholahti	49	Kaisaniemi
25	Roihuvuori	50	Siltamäki

Using Geopy Client, a python library that helps us getting the latitudes and longitudes of all districts. Combing with the list of subdivisions above, we can have a completed data frame as below:

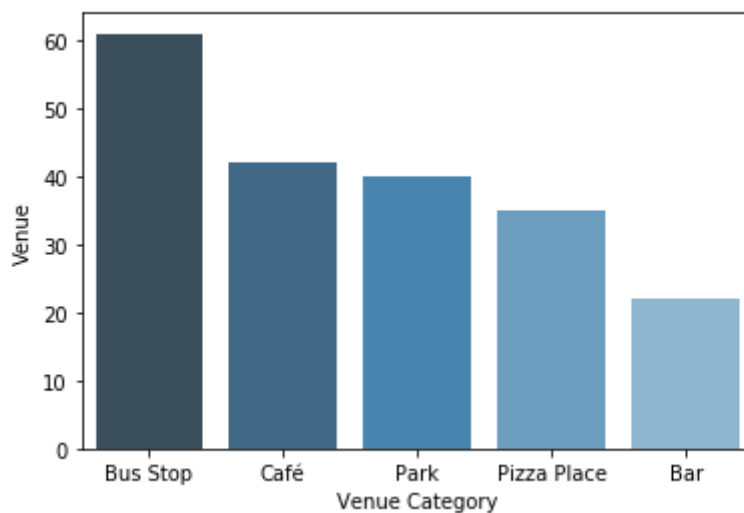|	Neighborhood	Latitude	Longitude
0	Kamppi	60.168535	24.930494
1	Suomenlinna	60.145706	24.988860
2	Munkkiniemi	60.199094	24.874358
3	Herttoniemi	60.195525	25.029063
4	Vuosaari	60.208348	25.143554
5	Itäkeskus	60.208652	25.084760
6	Hernesaari	60.148235	25.096343
7	Ullanlinna	60.158715	24.949404
8	Kallio	60.183982	24.952451
9	Vallila	60.196167	24.956710
10	Töölö	13.744717	-1.964599
11	Kalasatama	60.185295	24.980187
12	Katajanokka	60.166975	24.968151
13	Malmi	60.250761	25.008574
14	Kontula	60.238786	25.080050
15	Kuninkaantammi	60.265462	24.892970
16	Pitäjänmäki	60.222711	24.864356
17	Oulunkylä	60.229443	24.963583
18	Kannelmäki	60.243608	24.883289
19	Myllypuro	60.224010	25.066383
20	Malminkartano	60.249340	24.859959
21	Käpylä	60.215684	24.952786
22	Viikki	60.225892	25.022264
23	Konala	60.238550	24.846065
24	Ruoholahti	60.162925	24.911497
25	Roihuvuori	60.201457	25.061163
26	Sompasaari	60.180063	24.976118
27	Meilahti	60.191348	24.902664
28	Kumpula	60.204001	24.958108
29	Tapanila	60.522998	26.894329
30	Kruununhaka	60.172870	24.954733
31	Haaga	60.218979	24.897647
32	Arabianranta	60.208579	24.979143
33	Eira	60.156191	24.938375
34	Sörnäinen	60.183885	24.964409
35	Maunula	62.985124	21.993138
36	Pukinmäki	60.245247	24.989694
37	Mellunmäki	60.237838	25.115417
38	Jollas	60.163001	25.074154
39	Kluuvi	60.170778	24.947329
40	Pihlajamäki	62.626649	23.294745
41	Jakomäki	60.259171	25.078444
42	Tapulikaupunki	60.276210	25.025881
43	Rastila	60.210445	25.123320
44	Puotila	60.210204	25.103034
45	Munkkivuori	60.207101	24.871208
46	Merihaka	60.178833	24.961806
47	Pikku Huopalahti	60.202954	24.894140
48	Pakila	60.243545	24.939963
49	Kaisaniemi	60.173182	24.946400
50	Siltamäki	60.274750	24.988819

2.2 Collect Venues Locations using Foursquare:

Using Foursquare API, we can get a very large database of nearby venues for each neighborhood in JSON format. Here in this table, we are able to choose 100 popular locations maximum for each district within a 1 km radius.
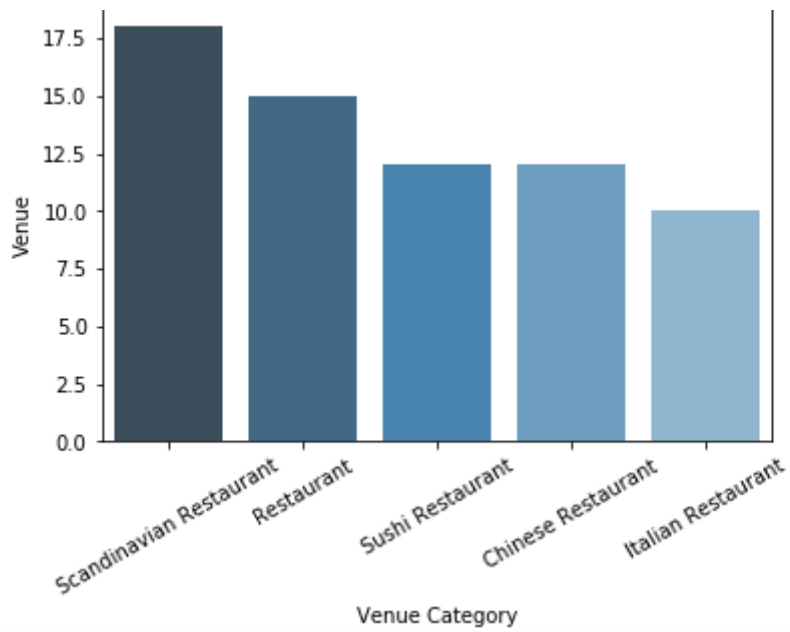
	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Kamppi	60.168535	24.930494	Cafe Rouge	60.168711	24.933027	Middle Eastern Restaurant
3	Kamppi	60.168535	24.930494	Pobre	60.169500	24.933484	Filipino Restaurant
12	Kamppi	60.168535	24.930494	Singapore Hot Wok	60.169163	24.933890	Malay Restaurant
14	Kamppi	60.168535	24.930494	Sushibar + Wine	60.167187	24.932170	Sushi Restaurant
15	Kamppi	60.168535	24.930494	Hoku	60.169490	24.933543	Hawaiian Restaurant
...	...	...	...	...	...	...	...
1162	Kaisaniemi	60.173182	24.946400	Momotoko	60.169834	24.946957	Ramen Restaurant
1169	Kaisaniemi	60.173182	24.946400	Döner Harju City	60.170082	24.941644	Doner Restaurant
1173	Kaisaniemi	60.173182	24.946400	Kiinalainen Ravintola Family	60.174245	24.950121	Chinese Restaurant
1178	Kaisaniemi	60.173182	24.946400	Fazer Food & Co Oodi	60.173648	24.937799	Restaurant
1195	Kaisaniemi	60.173182	24.946400	Ravintola Kaisaniemi	60.176665	24.941657	Restaurant

3. Visualization and Data Exploration

After collecting all the needed information, it's time to visualize and explore the data. First of all, we can see the top 5 venue categories in Helsinki from the chart below:

There are 228 unique venue categories in total and surprisingly as the surprising result, **Bus stop** is on top of the list. People might find it beautiful and attractive. Comes in the top 5 are **Cafe**, **Park**, **Pizza Place,** and **Bar.** Additionally, when it comes to food categories, it's clear that **Scandinavian** restaurant takes the win alongside **Chinese**, **Sushi**, and **Italian** restaurants.

Overall, A data frame for venue contributions grouped by each neighborhood is created. (noted that each one can query 100 venues maximum)

	Neighborhood	Venue
0	Arabianranta	22
1	Eira	33
2	Haaga	14
3	Herttoniemi	37
4	Itäkeskus	33
5	Jakomäki	19
6	Jollas	4
7	Kaisaniemi	81
8	Kalasatama	33
9	Kallio	87
10	Kamppi	100
11	Kannelmäki	7
12	Katajanokka	34
13	Kluuvi	93
14	Konala	18
15	Kontula	26
16	Kruununhaka	78
17	Kumpula	6
18	Kuninkaantammi	4
19	Käpylä	13
20	Malmi	22
21	Malminkartano	10
22	Meilahti	22
23	Mellunmäki	13
24	Merihaka	11
25	Munkkiniemi	18
26	Munkkivuori	9
27	Myllypuro	15
28	Oulunkylä	15
29	Pakila	4
30	Pikku Huopalahti	8
31	Pitäjänmäki	9
32	Pukinmäki	15
33	Puotila	14
34	Rastila	4
35	Roihuvuori	14
36	Ruoholahti	19
37	Siltamäki	14
38	Sompasaari	21
39	Suomenlinna	29
40	Sörnäinen	29
41	Tapanila	1
42	Tapulikaupunki	12
43	Ullanlinna	56
44	Vallila	44
45	Viikki	9
46	Vuosaari	31

As shown on the table, **Kamppi** and **Kluuvi** contribute the most venues with 100 and 93 venues respectively. The reason is that those areas are located in the center of Helsinki contain various famous places for tourists and domestic people.
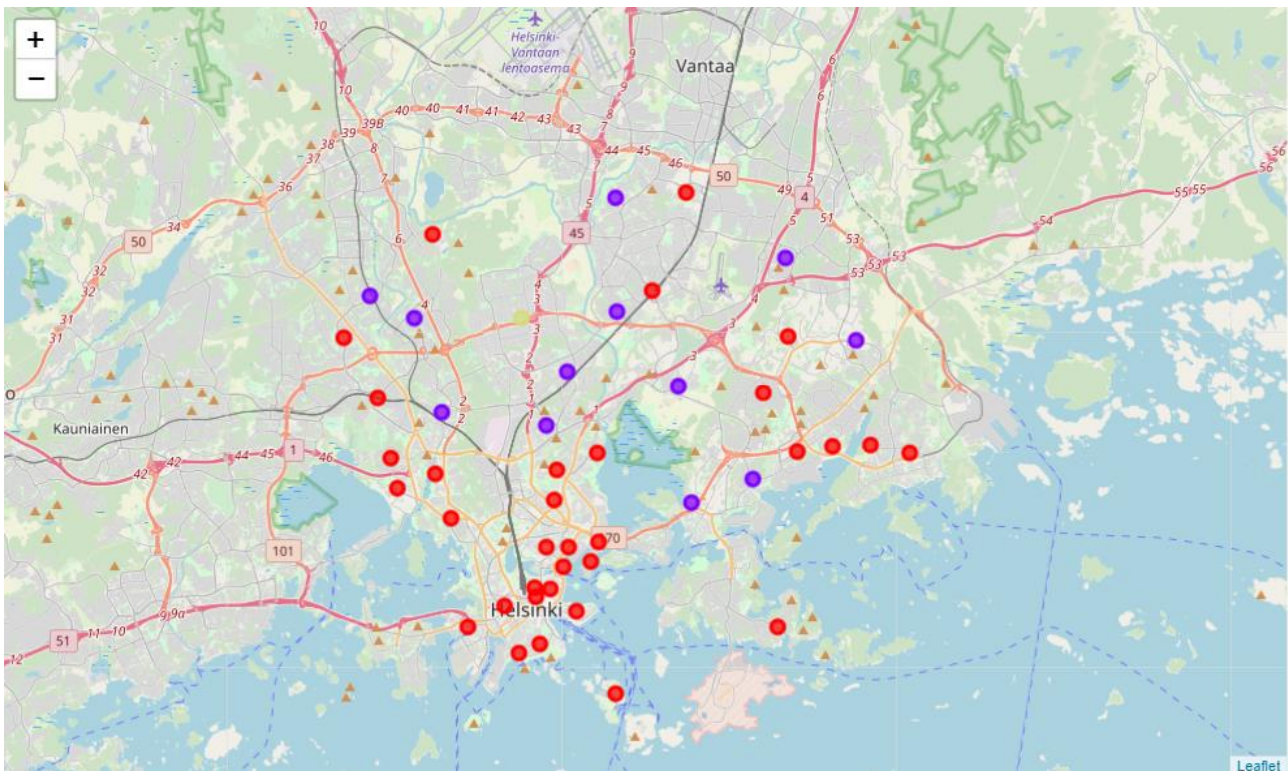
4. Clustering the neighborhoods

Finally, using K-means clustering, we try to cluster all districts based on the venue's information into 4 groups using the following snippet:

```
# set number of clusters
kclusters = 4


helsinki_grouped_clustering = helsinki_grouped.drop('Neighborhood', 1)

kmeans = KMeans(n_clusters=kclusters,random_state=0).fit(helsinki_grouped_clustering)
```

Let's visualize the clustered neighborhood in a map using Folium:

The following tables show the details from each cluster:

Cluster 1

	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
0	Kamppi	Scandinavian Restaurant	Beer Bar	Wine Bar	Art Museum	Chinese Restaurant	Café	Middle Eastern Restaurant	Sushi Restaurant	Outdoor Supply Store	Burger Joint
1	Suomenlinna	Café	History Museum	Historic Site	Park	Bar	Bistro	Grocery Store	Theater	Tourist Information Center	Restaurant
2	Munkkiniemi	Café	Disc Golf	Park	Museum	Gastropub	Pharmacy	Health Food Store	Vegetarian / Vegan Restaurant	Historic Site	Pizza Place
4	Vuosaari	Discount Store	Café	Bus Stop	Playground	Clothing Store	Grocery Store	Gym	Cafeteria	Shoe Repair	Scandinavian Restaurant
5	Itäkeskus	Coffee Shop	Sushi Restaurant	Clothing Store	Grocery Store	Gym / Fitness Center	Hardware Store	Chinese Restaurant	Outdoor Supply Store	Thai Restaurant	Liquor Store
7	Ullanlinna	Park	Grocery Store	Scandinavian Restaurant	Pizza Place	Coffee Shop	French Restaurant	Bistro	Ice Cream Shop	Farmers Market	Spanish Restaurant
8	Kallio	Bar	Café	Bakery	Beer Bar	Thai Restaurant	Pizza Place	Thrift / Vintage Store	Grocery Store	Park	Vegetarian / Vegan Restaurant
9	Vallila	Tram Station	Pizza Place	Park	Cafeteria	Bar	Chinese Restaurant	Sushi Restaurant	Garden	Dive Bar	Burrito Place
11	Kalasatama	Italian Restaurant	Sushi Restaurant	Coffee Shop	Supermarket	Gym / Fitness Center	Burger Joint	Indonesian Restaurant	Skate Park	Gastropub	Brewery
12	Katajanokka	Park	Hotel	Scandinavian Restaurant	Restaurant	Bar	Tram Station	Russian Restaurant	Plaza	Boat or Ferry	Gourmet Shop
13	Malmi	Coffee Shop	Gym / Fitness Center	Cultural Center	Beer Bar	Plaza	Pizza Place	Pharmacy	Restaurant	Fast Food Restaurant	Chinese Restaurant
14	Kontula	Bar	Supermarket	Grocery Store	Falafel Restaurant	Bus Stop	Nightclub	Skate Park	Sushi Restaurant	Gas Station	Middle Eastern Restaurant
15	Kuninkaantammi	Trail	Playground	Café	Ski Trail	Zoo	Fountain	Forest	Food Truck	Food Court	Food
16	Pitäjänmäki	Pizza Place	Train Station	Office	Cafeteria	Warehouse Store	Vietnamese Restaurant	Park	Athletics & Sports	Zoo	Flea Market
19	Myllypuro	Pizza Place	Grocery Store	ATM	Turkish Restaurant	Racetrack	Playground	Pharmacy	Shopping Mall	Café	Bus Stop
23	Konala	Automotive Shop	Cafeteria	Supermarket	Chinese Restaurant	Pizza Place	Paper / Office Supplies Store	Gym / Fitness Center	Convenience Store	Bakery	Bus Stop
24	Ruoholahti	Gym	Museum	Restaurant	Food Truck	Park	Beer Bar	Middle Eastern Restaurant	Himalayan Restaurant	Historic Site	Canal
26	Sompasaari	Art Gallery	Café	Zoo	Basketball Court	Pub	Pier	Restaurant	Park	Sauna / Steam Room	Scandinavian Restaurant
27	Meilahti	Bar	Café	Gym / Fitness Center	Tram Station	Disc Golf	Flower Shop	Supermarket	Scandinavian Restaurant	Tennis Court	Gym
28	Kumpula	Garden	Supermarket	Botanical Garden	Bus Stop	Café	Park	Furniture / Home Store	French Restaurant	Fountain	Forest
30	Kruununhaka	Boat or Ferry	Grocery Store	Theater	History Museum	Bar	Scandinavian Restaurant	Café	Coffee Shop	Indie Movie Theater	Modern European Restaurant
32	Arabianranta	Cafeteria	Park	Restaurant	Café	Plaza	Shopping Mall	Taxi Stand	Grocery Store	Gym	Comic Shop
33	Eira	Scandinavian Restaurant	Park	Bakery	French Restaurant	Café	Italian Restaurant	Boat or Ferry	Playground	Ice Cream Shop	Restaurant
34	Sörnäinen	Bar	Pizza Place	Burger Joint	Thai Restaurant	Recreation Center	Event Space	Beer Bar	Park	Bus Stop	Café
38	Jollas	Forest	Playground	Pizza Place	Historic Site	Zoo	Fast Food Restaurant	French Restaurant	Fountain	Food Truck	Food Court
39	Kluuvi	Café	Coffee Shop	Scandinavian Restaurant	Theater	Furniture / Home Store	Hotel	Park	Chinese Restaurant	Pizza Place	Clothing Store
42	Tapulikaupunki	Bar	Trail	Playground	Grocery Store	Outdoors & Recreation	Pizza Place	Convenience Store	Flea Market	Recreation Center	Pub
43	Rastila	Music Store	Café	Campground	Trail	Zoo	Furniture / Home Store	French Restaurant	Fountain	Forest	Food Truck
44	Puotila	Pub	Restaurant	Bus Stop	Café	Shopping Mall	Grocery Store	Beer Bar	Park	Pizza Place	Harbor / Marina
45	Munkkivuori	Supermarket	Pizza Place	Sushi Restaurant	Bus Stop	Café	Flower Shop	Gym / Fitness Center	Liquor Store	French Restaurant	Doner Restaurant
46	Merihaka	Beach Bar	Bar	Burger Joint	Grocery Store	Bistro	Health Food Store	Sauna / Steam Room	Speakeasy	Dance Studio	Waterfront
47	Pikku Huopalahti	Spa	Himalayan Restaurant	Garden	Taxi Stand	Convenience Store	Plaza	Park	Pizza Place	Fish & Chips Shop	Fish Market
49	Kaisaniemi	Clothing Store	Coffee Shop	Theater	Restaurant	Gym / Fitness Center	Bar	Chinese Restaurant	Burger Joint	Waterfront	Pizza Place

Cluster 2

	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
3	Herttoniemi	Bus Stop	Gym / Fitness Center	Pharmacy	Supermarket	Playground	Kebab Restaurant	Gastropub	Metro Station	Sandwich Place	Garden
17	Oulunkylä	Bus Stop	Pizza Place	Irish Pub	Recreation Center	Outdoor Sculpture	Grocery Store	Park	Chinese Restaurant	Fish & Chips Shop	Forest
18	Kannelmäki	Pizza Place	Bus Stop	Soccer Field	Karaoke Bar	Playground	Grocery Store	Zoo	Filipino Restaurant	Forest	Food Truck
20	Malminkartano	Bus Stop	Platform	Trail	Pet Store	Cafeteria	Soccer Field	Hotel	Pizza Place	Flea Market	Fish & Chips Shop
21	Käpylä	Bus Stop	Grocery Store	Scenic Lookout	Flower Shop	Scandinavian Restaurant	Coffee Shop	Kebab Restaurant	Café	Recreation Center	Gym Pool
22	Viikki	Bus Stop	Recreation Center	Italian Restaurant	Dog Run	Grocery Store	Zoo	Fish & Chips Shop	French Restaurant	Fountain	Forest
25	Roihuvuori	Bus Stop	Park	Café	Pizza Place	Convenience Store	Grocery Store	Shopping Mall	Harbor / Marina	Zoo	Fish & Chips Shop
31	Haaga	Bus Stop	Supermarket	Recreation Center	Shoe Repair	Bakery	Cafeteria	Pizza Place	Garden	Indian Restaurant	Sushi Restaurant
36	Pukinmäki	Bus Stop	Platform	Convenience Store	Karaoke Bar	Pet Store	Himalayan Restaurant	Playground	Grocery Store	IT Services	Dive Bar
37	Mellunmäki	Flea Market	Bus Stop	Grocery Store	Metro Station	Fast Food Restaurant	Gym / Fitness Center	Gym	Pub	Convenience Store	Taxi Stand
41	Jakomäki	Bus Stop	Soccer Field	Supermarket	Convenience Store	Shopping Mall	Beach	Pharmacy	Karaoke Bar	Pub	Plaza
50	Siltamäki	Bus Stop	Grocery Store	Dive Bar	Pharmacy	Pizza Place	Shopping Mall	Convenience Store	Waterfront	Recreation Center	Pool

Cluster 3

	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
29	Tapanila	Sports Club	Zoo	Garden	French Restaurant	Fountain	Forest	Food Truck	Food Court	Food	Flower Shop

Cluster 4

	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
48	Pakila	Bus Stop	Zoo	Gas Station	Furniture / Home Store	French Restaurant	Fountain	Forest	Food Truck	Food Court	Food

5. Conclusion and Summary

After using k-means clustering, we can specify that **Cluster 1** is a group of neighborhoods that has the common known for good restaurants and pubs while **Cluster 2** is mainly reviewed by local people. Meanwhile, **Cluster 3** and **Cluster 4** are classified when contains some unexpected venues categories such as **Gas Station**, **Garden**, etc.

By using data collected in Foursquare, we have been able to have a small project about analyzing Helsinki venue information. Even though the amount of available data from Foursquare for the Helsinki area is still quite limited, A small guidance for tourists is made in order to provide solutions for future individuals and businesses to improve.