The Battle of Neighborhoods - Stockholm

Introduction
Stockholm is a growing city and there are a lot of people in Sweden that is not from Stockholm that is, at this time, moving into Stockholm. A lot of people are therefore looking for apartments to buy, but is not familiar with the different neighbourhoods. This project will, based on data, give people that are new to Stockholm a description and comparison of different neighbourhoods with available apartments in Stockholm, so they easily can find a fit for them.

Data
Foursquare.com
Data will be retrieved through API from Foursquare and be delivered as a CSV file. The data will contain location data about Stockholm.

Hemnet.se
Data will be downloaded from Hemnet and be delivered as a CSV file. The data will contain data about available apartments in Stockholm.

I will combine the location data and the data of apartments to build a dataframe that contains important information about Stockholm and its different places for living. The data will for an example contain information about different venues near by, most common venues and the dresses to these apartments. The project aims to cluster Stockholm into smaller neighbourhoods(segments) from the dataframe to evaluate and define their key characteristics. In this way the project will define the different neighbourhoods from important properties when people buying an apartment in Stockholm.

Methodology
Methodology section which represents the main component of the report where you discuss and describe any exploratory data analysis that you did, any inferential statistical testing that you performed, if any, and what machine learnings were used and why.

Exploratory Data analys - Data Visualisation
To make sure the data was correct with the right geocodes for the addresses I used data visualisation to quickly get an overview if all geocodes are in Stockholm. This I did with the folium-package witch allowed me to plot markers(the addresses) on a map over Stockholm.

I also visualised the top of my data and data frames to see that it was structures in the correct way.

Other exploratory data analysis approached I used was showing a summary of the data collected from foursquare with function *count()* to make sure I received data in the correct way and also checking the types of the values and columns along the way to see if I built the data frames correct, this by using pandas *dtype* function.

Hot encoding
I used hot encoding to process my categorical variables and convert them to a form that is better used for my unsupervised machine learning model k-means. This was applied on the data frame where location and apartments data was joint and grouped by the adress.

KMeans Machine learning model

I used the machine learning algorithm K-means to cluster the different addresses of apartments in Stockholm based on the near by venues. K-means is a unsupervised learning and suited my end-goal to be able to find groups in the data which has not been explicitly labeled.

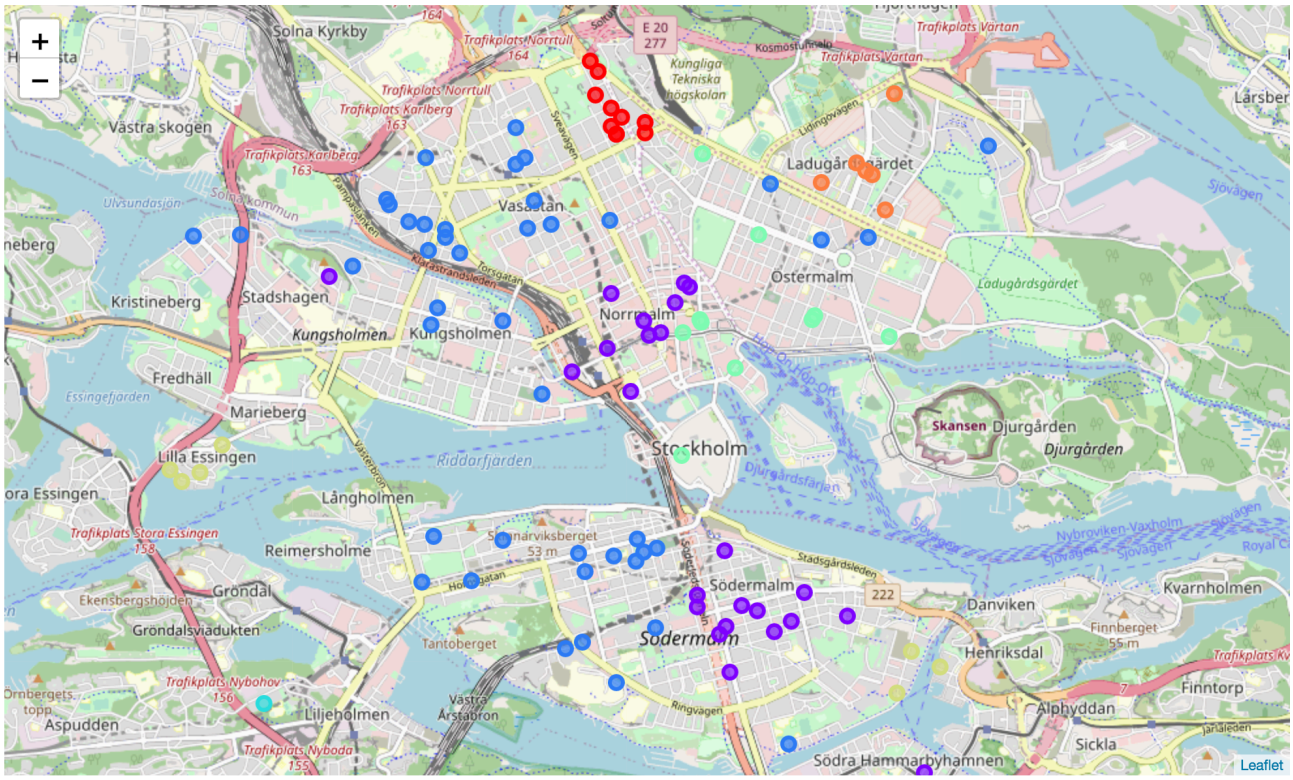
This will enable my goal to find different segments of addresses within Stockholm to later on be able to describe their characteristics and analyse their differences.

Results

Results section where you discuss the results.

Vizualisation of clusters in Stockholm

fig 1. Visualisation of the different segments and addresses in Stockholm



All clusters are relatively gathered together in the similar location, only the cluster 3 which is represented by blue is more spread out across Stockholm. Cluster 2, represented by purple is also spread out but still in most cases placed in very central places.

The different Segments of apartment addresses in Stockholm

	adress	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
1	Birger Jarlsgatan 113c	Sushi Restaurant	Squash Court	Playground	Scandinavian Restaurant	Miscellaneous Shop	Italian Restaurant	Plaza	Electronics Store	Pizza Place	Sporting Goods Shop
10	Surbrunnsgatan 7	Scandinavian Restaurant	Café	Italian Restaurant	Sushi Restaurant	Gym / Fitness Center	Coffee Shop	Bakery	Salad Place	Mediterranean Restaurant	Spanish Restaurant
18	Birger Jarlsgatan 97	Scandinavian Restaurant	Italian Restaurant	Café	Park	Pizza Place	Salad Place	Mediterranean Restaurant	French Restaurant	Spanish Restaurant	Sporting Goods Shop
23	Roslagsgatan 15A	Scandinavian Restaurant	Italian Restaurant	Middle Eastern Restaurant	Burger Joint	Vegetarian / Vegan Restaurant	Café	French Restaurant	Pizza Place	Salad Place	Breakfast Spot
28	Frejgatan 19	Scandinavian Restaurant	Italian Restaurant	Mediterranean Restaurant	Chinese Restaurant	Park	French Restaurant	Spanish Restaurant	Sporting Goods Shop	Café	Salad Place

fig 2. table of cluster 1, see all results in Capstone project code, results section

	adress	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
7	John Bergs Plan 3	Hotel	Café	Gym / Fitness Center	Metro Station	Bakery	Plaza	Office	Castle	Scandinavian Restaurant	Scenic Lookout
26	Bondegatan 27	Café	Clothing Store	Scandinavian Restaurant	Thai Restaurant	Coffee Shop	Indian Restaurant	Record Shop	Italian Restaurant	Deli / Bodega	Beer Bar
61	Korphoppsgatan 20	Café	Gym / Fitness Center	Scandinavian Restaurant	Hotel	Bakery	Sushi Restaurant	Indian Restaurant	Candy Store	Furniture / Home Store	Light Rail Station
78	sveavagen 2	Hotel	Scandinavian Restaurant	Clothing Store	Coffee Shop	Café	Gym / Fitness Center	Asian Restaurant	Cocktail Bar	Burger Joint	Department Store
81	kungsgatan 12	Scandinavian Restaurant	Hotel	Italian Restaurant	Movie Theater	Asian Restaurant	Burger Joint	Coffee Shop	Clothing Store	Café	Bakery

fig 3. table of cluster 2, see all results in Capstone project code, results section

	adress	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
4	Parkgatan 14	Indian Restaurant	Thai Restaurant	Bakery	Sushi Restaurant	Gym / Fitness Center	Steakhouse	Burger Joint	Middle Eastern Restaurant	Park	Italian Restaurant
5	Kalmgatan 44	Scandinavian Restaurant	Gym / Fitness Center	Bar	Automotive Shop	Supermarket	Gas Station	Restaurant	Bakery	Thai Restaurant	Furniture / Home Store
9	Lindhagensgatan 149	Scandinavian Restaurant	Gym / Fitness Center	Middle Eastern Restaurant	Park	Indian Restaurant	Poke Place	Doner Restaurant	Plaza	Supermarket	Steakhouse
11	Birkagatan 14	Ice Cream Shop	Pizza Place	Indian Restaurant	Sushi Restaurant	French Restaurant	Cocktail Bar	Coffee Shop	Café	Bakery	Thai Restaurant
13	Hornsgatan 150 B	Indian Restaurant	Bar	Italian Restaurant	Middle Eastern Restaurant	Sushi Restaurant	Playground	Bakery	Gym / Fitness Center	Organic Grocery	Pub

fig 4. table of cluster 3, see all results in Capstone project code, results section

	adress	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	Nybohovsbacken 83	Stadium	Department Store	Café	Athletics & Sports	Electronics Store	Yoga Studio	Falafel Restaurant	Fountain	Food Truck	Food Court

fig 5. table of cluster 4, see all results in Capstone project code, results section

	adress	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
6	Storgatan 59	Scandinavian Restaurant	Restaurant	Park	Bus Station	Museum	Tram Station	Café	Fast Food Restaurant	Liquor Store	Furniture / Home Store
16	Danderydsgatan 30	Scandinavian Restaurant	Café	Sushi Restaurant	Light Rail Station	Italian Restaurant	Chinese Restaurant	Greek Restaurant	Stadium	Medical Center	General College & University
20	Grev Turegatan 54	Scandinavian Restaurant	Italian Restaurant	Deli / Bodega	Bar	Bakery	Park	Grocery Store	Pizza Place	French Restaurant	Yoga Studio
21	Styrmansgatan 25	Scandinavian Restaurant	Italian Restaurant	Grocery Store	Café	Sushi Restaurant	Yoga Studio	Thai Restaurant	French Restaurant	History Museum	Restaurant
39	Styrmansgatan 21	Scandinavian Restaurant	Italian Restaurant	Sushi Restaurant	Grocery Store	Café	Thai Restaurant	History Museum	French Restaurant	Yoga Studio	Tram Station

fig 6. table of cluster 5, see all results in Capstone project code, results section

	adress	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
2	Poliargatan 24	Bus Stop	Park	Pizza Place	Italian Restaurant	Beer Garden	Grocery Store	Asian Restaurant	Harbor / Marina	Café	Pier
8	Essinge Brogata 3	Grocery Store	Bus Stop	Bakery	Scandinavian Restaurant	Restaurant	Gym / Fitness Center	Salad Place	Beach	Thai Restaurant	Indian Restaurant
12	Primusgatan 58	Bus Stop	Pizza Place	Thai Restaurant	Harbor / Marina	Grocery Store	Restaurant	Park	Pier	Beach	Café
24	Primusgatan 78	Thai Restaurant	Park	Pizza Place	Harbor / Marina	Restaurant	Grocery Store	Beach	Bus Stop	Pier	Deli / Bodega
47	Luxgatan 23	Thai Restaurant	Park	Pizza Place	Harbor / Marina	Restaurant	Grocery Store	Bus Stop	Pier	Deli / Bodega	Café

fig 7. table of cluster 6 , see all results in Capstone project code, results section

	adress	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	Sandelsgatan 30	Bus Stop	Thai Restaurant	Café	Restaurant	Bakery	Convenience Store	Pizza Place	Burger Joint	Chocolate Shop	Clothing Store
14	Sandelsgatan 38	Thai Restaurant	Sushi Restaurant	Café	Convenience Store	Burger Joint	Bus Stop	Restaurant	Chocolate Shop	Clothing Store	Coffee Shop
41	Strindbergsgatan 53	Café	Coffee Shop	Park	Thai Restaurant	Bakery	Convenience Store	Clothing Store	Persian Restaurant	Bistro	Supermarket
44	Studentbacken 32	Bus Stop	Asian Restaurant	Sushi Restaurant	Thai Restaurant	Fast Food Restaurant	Burger Joint	Tennis Stadium	Pizza Place	Scandinavian Restaurant	Grocery Store
51	Sandelsgatan 42	Thai Restaurant	Park	Café	Sushi Restaurant	Persian Restaurant	Pizza Place	Coffee Shop	Clothing Store	Electronics Store	Chocolate Shop

fig 8. table of cluster 7, see all results in Capstone project code, results section

Cluster 1 is located close to the Royal institute of Technology and all addresses are located very close together. This cluster most common venue is Scandinavian restaurant followed by other resturants, cafés and hotels. Close by is both a park and theater, and different gyms and fitness clubs are also venues commonly used.

Cluster 2 and 3 are both located in the very central places of Stockholm. There are a lot of different resutrants, hotels and bars close by. Cluster 2 has more hotels as commonly visited places, but also arts, museum, marina, boat ferry, movie theatres.
Cluster 3 is close to everything but has more gyms, yoga studios and fitness clubs close by.

Cluster 4 is based on the outside of Stockholm and only one adress is in this segment. It is close to both a stadium, yoga studio, athletics and sports etc.

Cluster 5 - has a mix of venues to visits such as Scandinavian resturants, museums etc. But was characterises this cluster the most is the closeness to train station, railway station and tram station. A central place with easy access to communication.

Cluster 6 - has close venues such as parks, harbor / marina, piers, beach and less restaurants and hotels as the other clusters.

Cluster 7 - This segment has both thai resturants, sushi restaurants and other asian resturants close by.

Discussion

Discussion section where you discuss any observations you noted and any recommendations you can make based on the results.

Cluster 1 - Exploring swedish and other countries food

This cluster represents addresses close to the Royal Institute of Technology, so just for that it is a good place for students to live in. It seems to offer a lot of Scandinavian restaurants, but also a mixture of other foods. Definitely a place for a foodie.

Cluster 2 and 3 - close to everything

Cluster 2 represents addresses to live of for people who really like to live the city life, it has both for an example bars, movie theatre, boat ferry and restaurants. It is in my point of view, the perfect place for the Stockholm experience with both night life, arts and culture as well as the water. Also has a lot of close by hotels, which can be a result of a tourist place.

Cluster 3 also has a lot of different places close by as restaurants and bars but seems to fit a more active lifestyle with higher density of yoga, gyms and fitness centers.

Cluster 4 - the outsider

This cluster is only one adrdess that has different characteristics from the others. Its located a little more outside Stockholm than the others but is definitely a place for active people, with a stadium, gym and yoga center close by.

Cluster 5 - Easy access to communication

This cluster represents the addresses that is very close to a train, tram or subway station. This is a perfect place for someone who communicates to work or like to go outside Stockholm on the weekends. It is also very central and has different venues to visit.

Cluster 6 - By the water

This place is perfect for someone who loves the water and the nature. Since it is the segment with addresses that has close by venues such as marinas, piers, beaches and also parks. It dosen't have the same offer of different restaurants but seems to have some near by as well as cafés.

Cluster 7- Asian food

This cluster definitely has the asian restaurants in their heart. Here you can find thai, sushi and asian foods and also other restaurants and bars. Definitely something for people who likes the city life with a special place for the asian kitchen in their hearts.

Conclusion

Conclusion section where you conclude the report.

With the machine learning model this project enabled to group different addresses to find their characteristics in a simple way.

For further analysis I would like to in the future do the same segmentation but add information about the square meter prices of the different addresses. I believe this would give more value for the analysis and people moving into Stockholm. Do dive even deeper the is a potential to add more information about the apartments in general for an example add size to add more value to the user.