IBM Capstone Project - The Battle of Neighborhoods

Finding the Best Place for our Customer to Rent a House in Hamburg

Problem Description

- Find the best place for our customer to rent a house in Hamburg
- Conditions: Family-friendly and close to his new job location
- Hamburg has 104 quarters, 5.3 million inhabitants, is really diverse and is home to many companies and sports clubs
 - This leads to high rental prices and quarters that might be diverse
- Our aim: Cluster the relevant quarters of Hamburg, find a cluster which stands for family friendliness with parks and playgrounds and select the quarter which has the lowest rental prices in the family-friendly cluster.

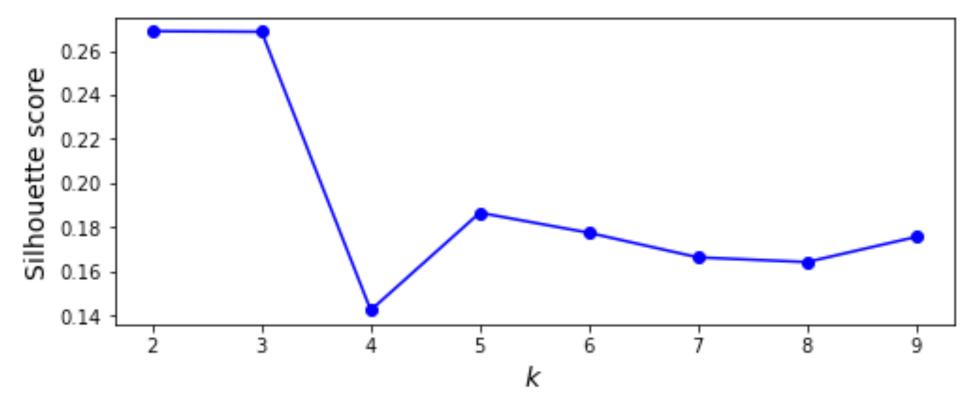
Data

- Data set with the quarters scraped from Wikipedia: https://de.wikipedia.org/wiki/Liste_der_Bezirke_und_Stadtteile_Hamburgs.
- Location (latitude and longitude coordinates) from GeoPy, a Python client for geocoding web services.
- Data set with rental prices of the quarters scraped from Hamburg Portal: https://www.hamburgportal.de/immobilien/mietwohnungen/mietenspiegel/

	Quarter	Borough	Latitude	Longitude	Price per m^2 in EURO
0	Neustadt	Hamburg-Mitte	53.549881	9.979048	10,69
1	Finkenwerder	Hamburg-Mitte	53.530882	9.858523	7,85
2	Altona-Altstadt	Altona	53.549660	9.945352	10,60
3	Sternschanze	Altona	53.561768	9.963282	11,73
4	Altona-Nord	Altona	53.561400	9.944720	10,61

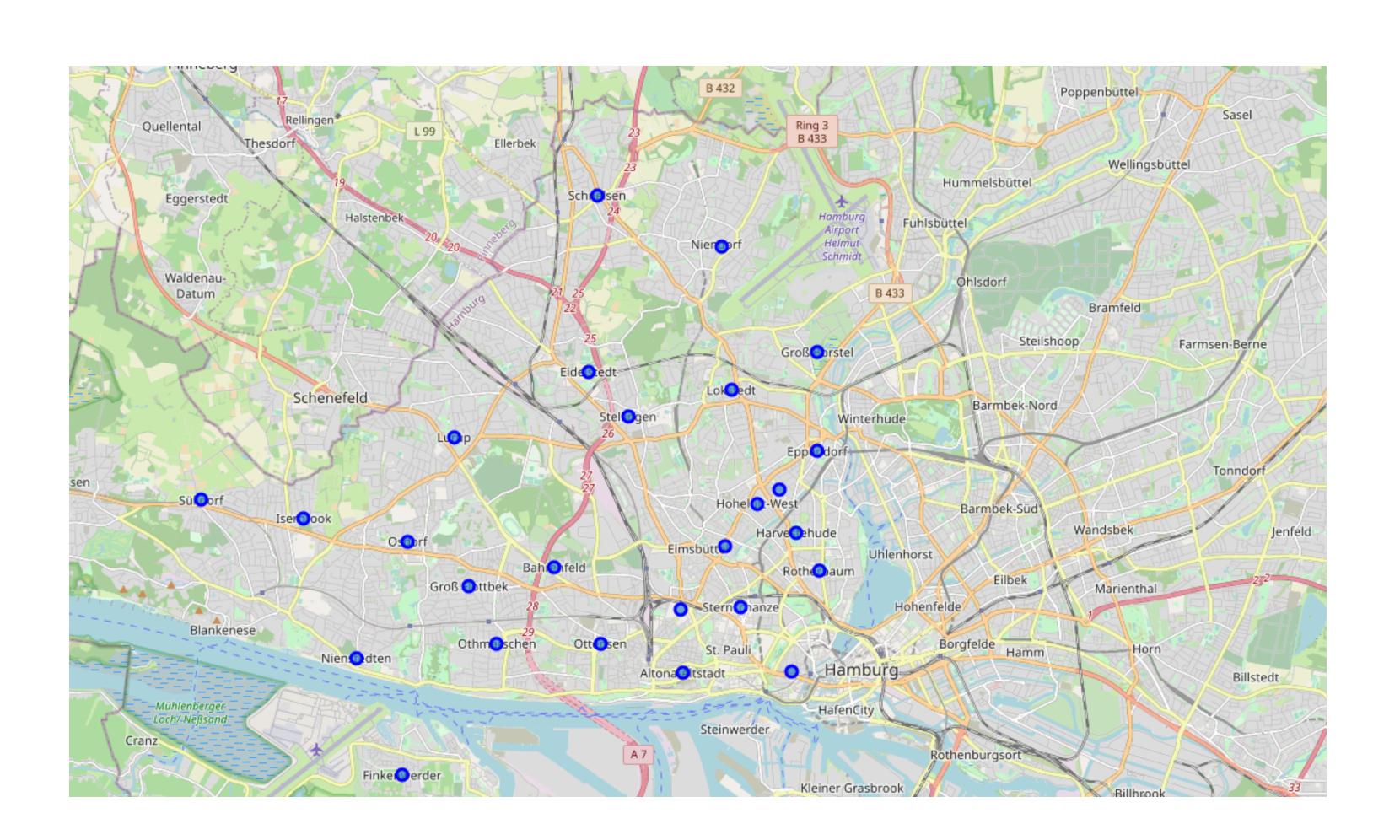
Methodology

- In this project we have unlabeled data therefore need an algorithm from the unsupervised learning section of machine learning. We chose the kmeans algorithm due its reliability and ease of use.
- The silhouette scoring tells which which k (number of clusters) to pick for the algorithm. Result: k=2 or k=3. We picked k=3, hoping to get 3 clusters that strongly differentiate from each other and to get one clean cluster with many family-friendly venues



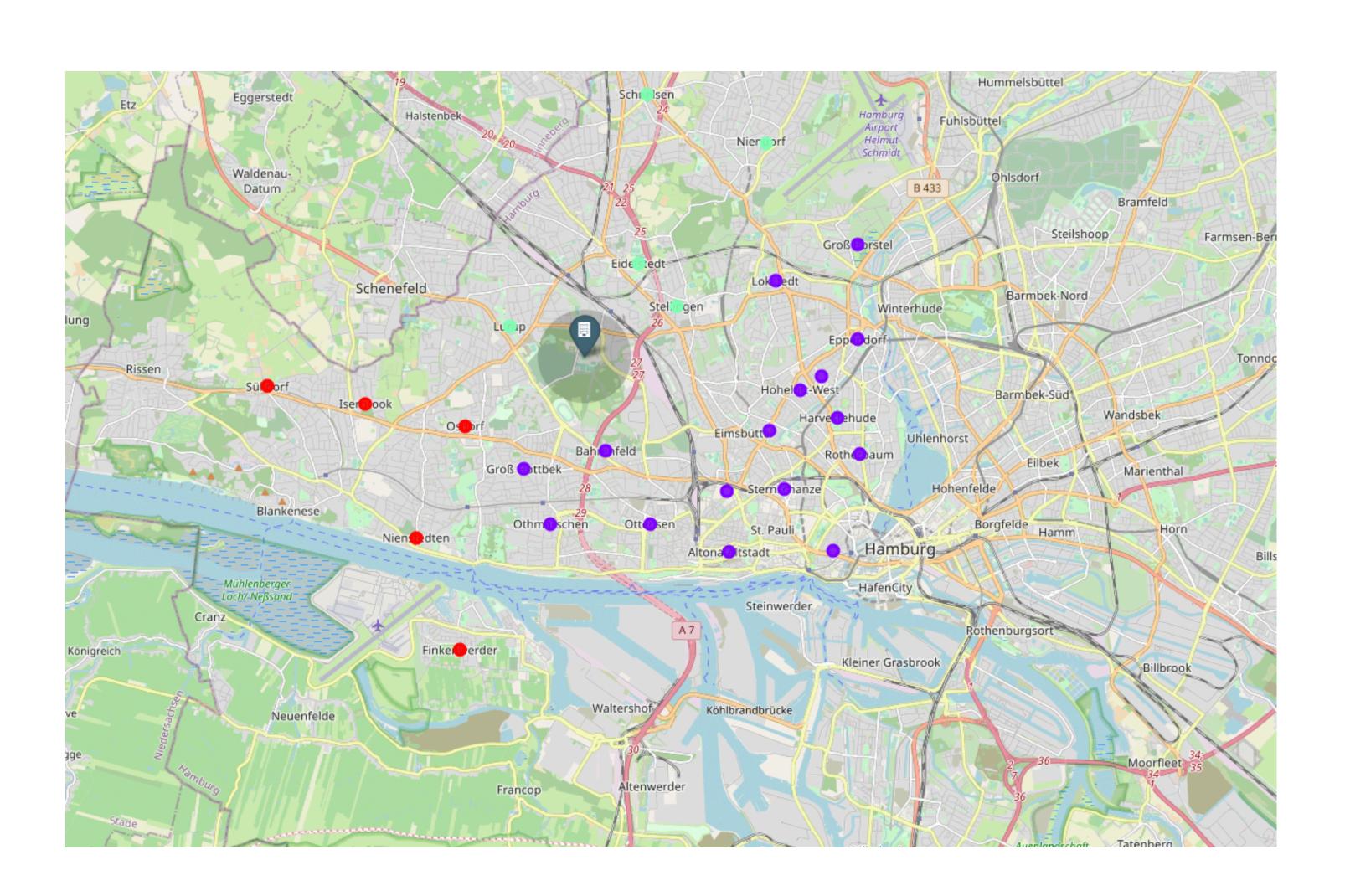
Results

• The quarters which are within a 7 km radius from our clients job location.



Results

 After the kmeans clustering we have 3 clusters (green, red and blue dots).



Results Cluster 1

- Seems not to be the cluster we were looking for.
- It has one quarter which could be interesting (i.e. Osdorf) but the other quarters don't show a clear tendency towards parks and/or playgrounds.

	Quarter	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	
0	Altona- Altstadt	Café	Park	Coffee Shop	Bakery	Seafood Restaurant	Pizza Place	
1	Altona-Nord	Café	Bakery	Coffee Shop	Park	Seafood Restaurant	Pizza Place	
2	Bahrenfeld	Café	Bakery	Park	Coffee Shop	Farmers Market	Ice Cream Shop	
3	Eidelstedt	delstedt Café Supermark		Bakery	Zoo Exhibit	Park	Ice Cream Shop	

Results Cluster 2

- We almost have in every quarter a park either as the most common venue or as the 2nd most common.
- Playgrounds seem to be underrepresented across all clusters.
- Groß Borstel is the way to go due to the lowest rental price

Quarter	Price per m^2 in EURO	Cluster Labels	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	
Neustadt	10,69	1	Park	Café	Coffee Shop	Bakery	Seafood Restaurant	Plaza	Pizza Place	Ice Cream Shop	Pub	Nightclub	
Altona- Altstadt	10,60	1	Café	Park	Coffee Shop	Bakery	Seafood Restaurant	Pizza Place	Ice Cream Shop	Hotel	Nightclub	Scenic Lookout	
Sternschanze	11,73	1	Café	Bakery	Coffee Shop	Plaza	Seafood Restaurant	Pizza Place	Park	Wine Bar	Falafel Restaurant	Austrian Restaurant	
Altona-Nord	10,61	1	Café	Bakery	Coffee Shop	Park	Seafood Restaurant	Pizza Place	Ice Cream Shop	Nightclub	Bistro	Supermarket	
Ottensen	11,60	1	Café	Park	Bakery	Coffee Shop	Pizza Place	Ice Cream Shop	Seafood Restaurant	Supermarket	Beach Bar	Nightclub	
Bahrenfeld	10,09	1	Café	Bakery	Park	Coffee Shop	Farmers Market	Ice Cream Shop	Seafood Restaurant	Supermarket	Pizza Place	German Restaurant	
Groß Flottbek	11,38	1	Park	Bakery	Café	Seafood Restaurant	Ice Cream Shop	Supermarket	Falafel Restaurant	Restaurant	Harbor / Marina	Shopping Mall	
Othmarschen	11,81	1	Park	Bakery	Café	Ice Cream Shop	Seafood Restaurant	Farmers Market	Wine Bar	Pizza Place	Supermarket	Bistro	
Eimsbüttel	11,36	1	Bakery	Café	Coffee Shop	Park	Pizza Place	Wine Bar	Ice Cream Shop	French Restaurant	Plaza	Hotel	F
Rotherbaum	12,41	1	Coffee Shop	Café	Ice Cream Shop	Wine Bar	Restaurant	Park	Cocktail Bar	Plaza	Hotel	Farmers Market	
Harvestehude	12,72	1	Café	Bakery	Coffee Shop	Park	Plaza	Ice Cream Shop	Restaurant	Wine Bar	Farmers Market	Seafood Restaurant	
Hoheluft- West	11,68	1	Café	Bakery	Coffee Shop	Park	Wine Bar	Plaza	Ice Cream Shop	Farmers Market	German Restaurant	Supermarket	ı
Lokstedt	10,28	1	Café	Park	Bakery	Supermarket	Zoo Exhibit	Coffee Shop	Wine Bar	Farmers Market	Hotel	Beer Store	
Hoheluft-Ost	12,28	1	Café	Park	Bakery	Coffee Shop	Wine Bar	Farmers Market	Supermarket	Ice Cream Shop	French Restaurant	Falafel Restaurant	
Eppendorf	11,83	1	Café	Park	Bakery	Coffee Shop	Wine Bar	Farmers Market	Supermarket	Ice Cream Shop	Tapas Restaurant	Bistro	
Groß Borstel	9,59	1	Café	Park	Supermarket	Bakery	Coffee Shop	Italian Restaurant	Ice Cream Shop	Wine Bar	Hotel	Bistro	ı
	Neustadt Altona- Altstadt Sternschanze Altona-Nord Ottensen Bahrenfeld Groß Flottbek Othmarschen Eimsbüttel Rotherbaum Harvestehude Hoheluft- West Lokstedt Hoheluft-Ost Eppendorf	Quarter in EURONeustadt10,69Altona- Altstadt10,60Sternschanze11,73Altona-Nord10,61Ottensen11,60Bahrenfeld10,09Groß Flottbek11,38Othmarschen11,81Eimsbüttel11,36Rotherbaum12,41Harvestehude12,72Hoheluft- West11,68Lokstedt10,28Hoheluft-Ost12,28Eppendorf11,83	Quarter in EURO per m^2 Labels Cluster Labels Neustadt 10,69 1 Altona-Altstadt 10,60 1 Sternschanze 11,73 1 Altona-Nord 10,61 1 Ottensen 11,60 1 Bahrenfeld 10,09 1 Groß Flottbek 11,38 1 Othmarschen 11,81 1 Eimsbüttel 11,36 1 Rotherbaum 12,41 1 Harvestehude 12,72 1 Hoheluft-West 11,68 1 Lokstedt 10,28 1 Hoheluft-Ost 12,28 1 Eppendorf 11,83 1	Quarter in EUROCommon VenueNeustadt10,691ParkAltona-Altstadt10,601CaféSternschanze11,731CaféAltona-Nord10,611CaféOttensen11,601CaféBahrenfeld10,091CaféGroß Flottbek11,381ParkOthmarschen11,811ParkEimsbüttel11,361BakeryRotherbaum12,411Coffee ShopHarvestehude12,721CaféHoheluft-West11,681CaféLokstedt10,281CaféHoheluft-Ost12,281CaféEppendorf11,831Café	Quarter in EUROPer m^2 LabelsCommon VenueCommon VenueNeustadt10,691ParkCaféAltona- Altstadt10,601CaféParkSternschanze11,731CaféBakeryAltona-Nord10,611CaféBakeryOttensen11,601CaféParkBahrenfeld10,091CaféBakeryGroß Flottbek11,381ParkBakeryOthmarschen11,811ParkBakeryEimsbüttel11,361BakeryCaféRotherbaum12,411Coffee ShopCaféHarvestehude12,721CaféBakeryHoheluft- West11,681CaféBakeryLokstedt10,281CaféParkHoheluft-Ost12,281CaféParkEppendorf11,831CaféPark	Quarter in EUROper m^2 LabelsCommon VenueCommon VenueCommon VenueCommon VenueNeustadt10,691ParkCaféCoffee ShopAltona-Altstadt10,601CaféBakeryCoffee ShopSternschanze11,731CaféBakeryCoffee ShopAltona-Nord10,611CaféBakeryCoffee ShopOttensen11,601CaféParkBakeryBahrenfeld10,091CaféBakeryParkGroß Flottbek11,381ParkBakeryCaféOthmarschen11,811ParkBakeryCaféEimsbüttel11,361BakeryCaféCoffee ShopRotherbaum12,411Coffee ShopCaféCoffee ShopHarvestehude12,721CaféBakeryCoffee ShopHoheluft-West11,681CaféBakeryCoffee ShopLokstedt10,281CaféParkBakeryHoheluft-Ost12,281CaféParkBakeryEppendorf11,831CaféParkBakery	Quarter in EUROper m^2 LabelsCommon VenueCommon VenueCommon VenueCommon VenueCommon VenueNeustadt10,691ParkCaféCoffee ShopBakeryAltona-Altstadt10,601CaféParkCoffee ShopPlazaSternschanze11,731CaféBakeryCoffee ShopParkAltona-Nord10,611CaféBakeryCoffee ShopParkOttensen11,601CaféParkBakeryCoffee ShopBahrenfeld10,091CaféBakeryParkCoffee ShopGroß Flottbek11,381ParkBakeryCaféSeafood RestaurantOthmarschen11,811ParkBakeryCaféIce Cream ShopEimsbüttel11,361BakeryCaféCoffee ShopParkRotherbaum12,411CoffeeCaféCoffee ShopParkHarvestehude12,721CaféBakeryCoffee ShopParkHoheluft-West11,681CaféBakeryCoffee ShopParkLokstedt10,281CaféParkBakeryCoffee ShopEppendorf11,831CaféParkBakeryCoffee Shop	Quarter in EURO Clabels Labels Common Venue Common Restaurant Altona-Altstadt 10,60 1 Café Park Coffee Shop Palza Seafood Restaurant Altona-Nord 10,61 1 Café Bakery Coffee Shop Park Seafood Restaurant Ottensen 11,60 1 Café Park Bakery Coffee Shop Pizza Place Bahrenfeld 10,09 1 Café Bakery Café Seafood Restaurant Groß Flottbek 11,38 1 Park Bakery Café Café Seafood Restaurant Groß Flottbek 11,81 1 Park Bakery Café Café	Quarter in EURO Clasers (Labels) Common Venue Park Restaurant Plaza Seafood Restaurant Plaza Place Plaza Place Plaza Place Seafood Restaurant Plaza Place Plaza P	Quarter per ne IEURO Cluster Labels Common Venue Plaza Factor Seafood Restaurant Plaza Place Plaza Place Plaza Place Incommon Venue Plaza Place Plaza Plaze Plaza Plaze	Quarter per mic Unitables in EURO Common Venue Park Park Park Coffee Shop Park Seaturant Park Seaturant Coffee Shop Park Seaturant Park Seaturant Coffee Shop Shop	Quarter in EURO Camber in EURO Common Venue Venue	Quarter per mc Labels Common feature Common Venue Color Common Venue Co

Results Cluster 3

• Cluster 3 looks like cluster 1 with supermarkets on top of the ranking.

	Quarter	Price per m^2 in EURO	Cluster Labels	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
9	Lurup	8,40	2	Supermarket	Bakery	Zoo Exhibit	Gym / Fitness Center	Drugstore	Café	Hotel	Park	Clothing Store	Shopping Mall
19	Niendorf	9,55	2	Supermarket	Hotel	Bakery	Café	Greek Restaurant	Park	Zoo Exhibit	Airport Service	Ice Cream Shop	Pool
20	Schnelsen	8,88	2	Zoo Exhibit	Supermarket	Ice Cream Shop	Furniture / Home Store	Indoor Play Area	Bakery	German Restaurant	Italian Restaurant	Hotel	Greek Restaurant
21	Eidelstedt	8,92	2	Café	Supermarket	Bakery	Zoo Exhibit	Park	Ice Cream Shop	German Restaurant	Italian Restaurant	Gym / Fitness Center	Wine Bar
22	Stellingen	9,81	2	Café	Bakery	Zoo Exhibit	Supermarket	Italian Restaurant	Ice Cream Shop	German Restaurant	French Restaurant	Coffee Shop	Park

Further Discussion

- Interesting to compare the results with other clustering algorithms. That is an idea for further analysis.
- Also the amount of categories (15) could be reduced in order to put the focus on e.g. 3 or 4 venue categories. That way, you could eventually make better distinctions and the algorithm would probably lead to even better results.