Capstone Project

The Battle of the Neighborhoods

Applied Data Science Capstone by Coursera/IBM

– Regina Castra – February 2020 –

Table of Contents

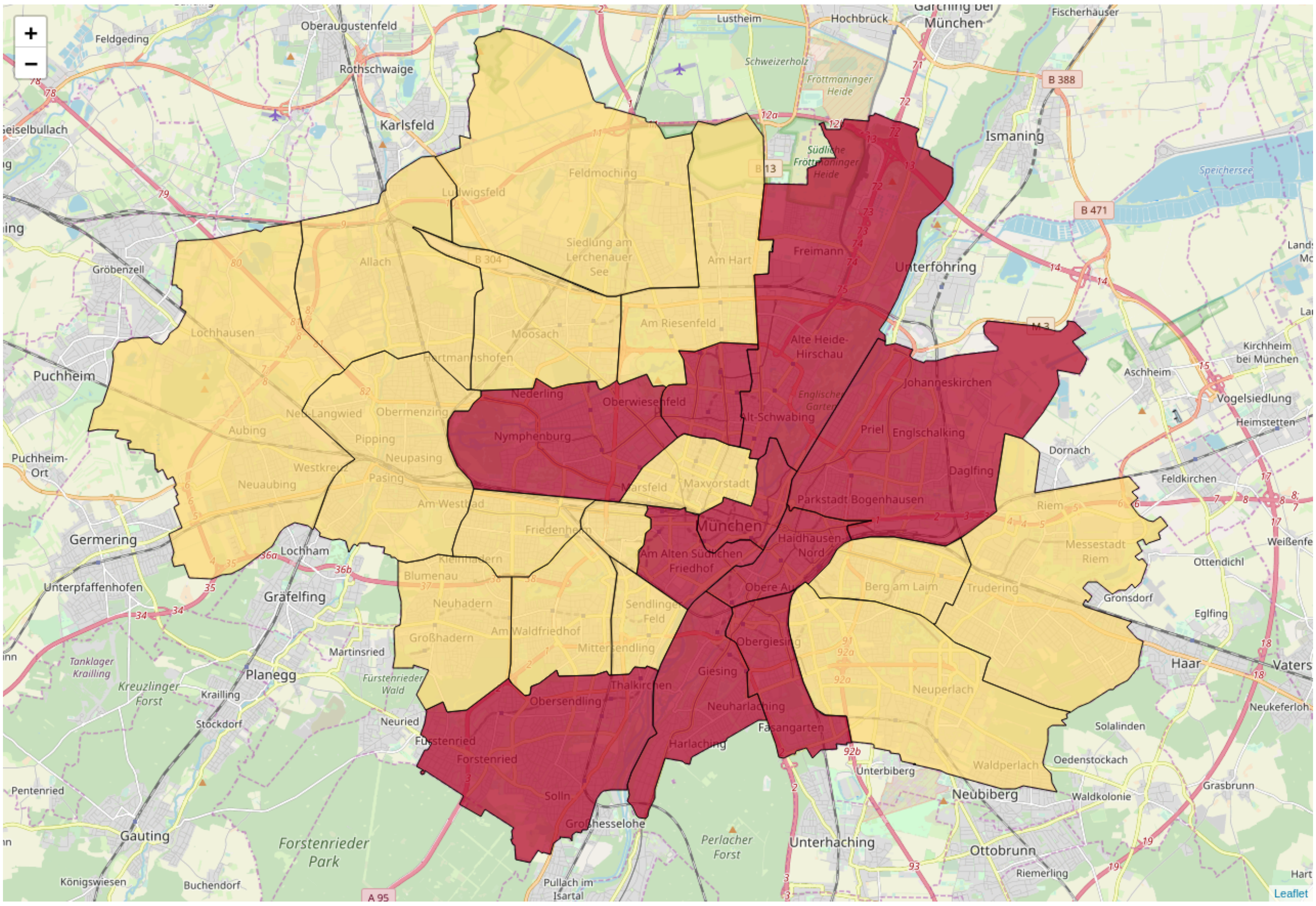

Fig. 1: Boroughs of Munich. Best Neighborhoods in Munich (highlighted in red) according to website www.moving-to-munich.com

1 Introduction

Munich is the capital and most populous city of Bavaria. With a population of around 1.5 million it is the third-largest city in Germany, according to Wikipedia. There is the website www.moving-to-munich.com which lists the Best Neighborhoods in Munich:

- Altstadt
- Au
- Bogenhausen
- Giesing (Ober- and Untergiesing)
- Haidhausen
- Isarvorstadt
- Lehel
- Neuhausen
- Schwabing
- Thalkirchen

These are roughly 10 of a total of 25 districts in Munich. Although the website describes the individual neighborhoods, it does not give precise reasons for the selection of these districts compared to the rest.

The idea of this project is to check if someone can make a similar selection based on the venues in each district. In addition, it might be possible to find neighborhoods that may have similar characteristics. These candidates would potentially have a comparable lifestyle, but probably lower rents.

If you want to move to Munich then this information might be of interest for you. Or if you are an investor looking for real estate assets to invest in then you are probably interested to find boroughs which are currently not high rated but have the potential to develop like the Best Neighborhoods in Munich have done as these “hidden” boroughs already show similar venues or venue distributions.

2 Data

The investigation is based on the following sources:

- Based on the website www.moving-to-munich.com a list of the Best Neighborhoods in Munich (https://www.moving-to-munich.com/best-neighborhoods-in-munich/) is created and modified by hand to represent the *official* names of Munich's boroughs.
- All 25 *official* names of Munich's boroughs are retrieved from a Wikipedia (https://de.wikipedia.org/wiki/Stadtbezirke_M%C3%BCnchens) page.
- For visualisation purpose the borders of Munich's boroughs are obtained from the website www.arcgis.com. It provides the *vector geometries* of Munich Districts and Subdistricts for free download and use. By using mapshaper.org the data is transformed into a suitable GeoJSON file format. The border of each borough is stored as a polygon which is used to determine each borough's center and extent.
- The venue data of Munich's boroughs is retrieved by using foursquare.com. Several radii are used for obtaining some kind of robust venue list for each borough. The lists of venues for all boroughs are used and a k-means clustering method is applied to group similar neighborhoods.

The above steps should make it possible to identify neighborhoods with a similar lifestyle like the Best Neighborhoods in Munich have.

2.1 Best Neighborhoods in Munich

According to the website www.moving-to-munich.com the Best Neighborhoods in Munich would be:

- Altstadt
- Au
- Bogenhausen
- Giesing (Ober- and Untergiesing)
- Haidhausen
- Isarvorstadt
- Lehel
- Neuhausen
- Schwabing
- Thalkirchen

The *official* names of the boroughs are a little bit different. A list of the *best* boroughs is created by hand using the *official* names:

```
Best Neighborhoods:                          Official Borough Names:
* Altstadt                               -->   Altstadt-Lehel
* Au                                     -->   Au-Haidhausen
* Bogenhausen                             =    Bogenhausen
* Giesing (Ober- and Untergiesing)       -->   Obergiesing-Fasangarten
                                         -->   Untergiesing-Harlaching
* Haidhausen                             -->   Au-Haidhausen (see above)
* Isarvorstadt                           -->   Ludwigsvorstadt-Isarvorstadt
* Lehel                                  -->   Altstadt-Lehel (see above)
* Neuhausen                              -->   Neuhausen-Nymphenburg
* Schwabing                              -->   Schwabing-West
                                         -->   Schwabing-Freimann
* Thalkirchen                            -->   Thalkirchen-Obersendling-
                                                 Forstenried-Fürstenried-Solln
```

2.2 Retrieve Boroughs of Munich

The *official* names of Munich's boroughs are retrieved from Wikipedia.org.
The data is modified, i.e. english column names are assigned, the data types are set correctly and an index is added. The columns of the resulting table are of following types:

The *official* names of Munich's boroughs are retrieved from from Wikipedia.org.
The data is modified, i.e. English column names are assigned, the data types are set correctly and an index is added. The columns of the resulting table are of the following type:

```
Borough                       object
Area(km**2)                  float64
Residents                      int64
Density(resident/km**2)        int64
Foreigners(%)                float64
bestBorough                    int64
```

2.3 Retrieve Borders of Munich's Boroughs

For visualisation purpose the borders of Munich's boroughs are required. The following procedure was applied to obtain a *GeoJSON* file describing these borders.

1. The website www.arcgis.com provides the *vector geometries* of Munich Districts and Subdistricts for free download and use.
2. The downloaded file *Munich_districts25_subdistricts105.lpk* is a 7-zip archive which contains (beside ohters) the following files:

```
commondata/data0/Munich_25_Bezirke_Dissolved.dbf
commondata/data0/Munich_25_Bezirke_Dissolved.prj
commondata/data0/Munich_25_Bezirke_Dissolved.shp
commondata/data0/Munich_25_Bezirke_Dissolved.shx
```

3. These files were uploaded to https://mapshaper.org/ using the options *detect line intersections* and *snap vertices*.
4. By exporting the data from https://mapshaper.org/ the file *Munich_25_Bezirke_Dissolved.json* is obtained.
5. There are two minor errors which are corrected by hand and saved again in file *Munich_25_Bezirke_Dissolved.json* :

```
"Tudering-Riem" --> "Trudering-Riem"
"Obergiesing"   --> "Obergiesing-Fasangarten"
```

2.4 Center and Extent of Munich's Boroughs

When analysing the file *Munich_25_Bezirke_Dissolved.json*, it turns out that the geometry of the boroughs is described by one polygon for each borough.
For each polygon its centroid (coord_lat, coord_lng) and an estimation of an radius (radius(m)) which describes roughly the extent of each borough, is calculated.

Table 1 shows the resulting list. The Best Neighborhoods in Munich are marked with a “1” in column “bestBorough”.

The 25 *official* boroughs of Munich (see Wikipedia) and the Best Neighborhoods in Munich according to the website www.moving-to-munich.com are visualized on a map in Fig. 2.

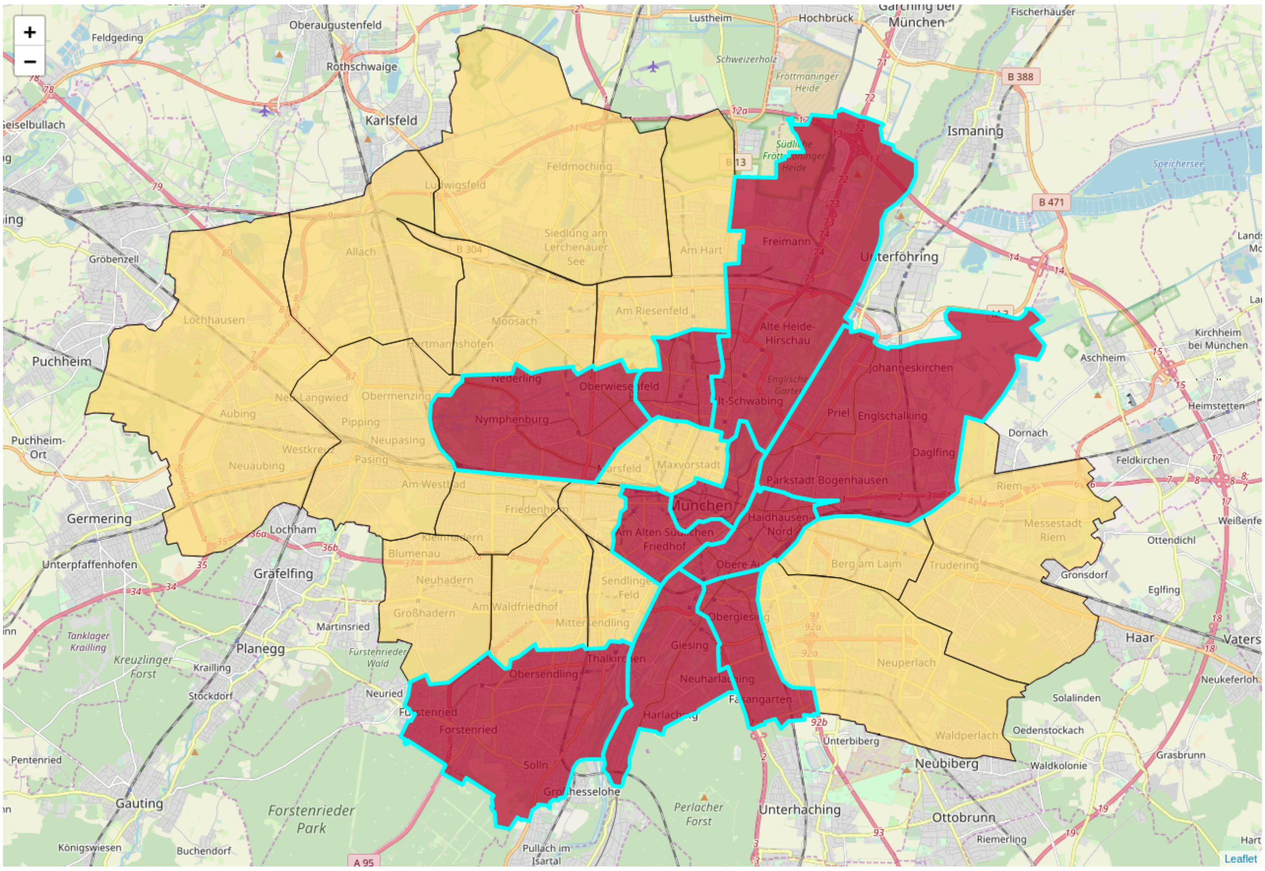

Fig. 2: [Best Neighborhoods in Munich]() are shown as red filled polygons with cyan borders

	Borough	Area(km**2)	Residents	Density(resident/km**2)	Foreigners(%)	bestBorough	coord_lat	coord_lng	radius(m)
0	Altstadt-Lehel	3.15	21100	6708	26.1	1	48.141273	11.583178	1804.8
1	Ludwigsvorstadt-Isarvorstadt	4.40	51644	11734	28.4	1	48.130152	11.561218	1920.2
2	Maxvorstadt	4.30	51402	11960	25.4	0	48.148069	11.564298	2001.2
3	Schwabing-West	4.36	68527	15706	22.7	1	48.166539	11.569307	1700.6
4	Au-Haidhausen	4.22	61356	14541	23.5	1	48.129727	11.594758	2338.5
5	Sendling	3.94	40983	10405	26.9	0	48.115282	11.546057	1854.2
6	Sendling-Westpark	7.81	59643	7632	28.9	0	48.116790	11.520021	2223.8
7	Schwanthalerhöhe	2.07	29743	14367	33.5	0	48.136161	11.538246	1190.7
8	Neuhausen-Nymphenburg	12.91	98814	7651	24.3	1	48.156452	11.519305	2904.3
9	Moosach	11.09	54223	4888	31.5	0	48.181714	11.510923	3933.0
10	Milbertshofen-Am Hart	13.42	75094	5597	40.8	0	48.195994	11.567868	3832.0
11	Schwabing-Freimann	25.67	77936	3036	29.3	1	48.192805	11.610558	4974.5
12	Bogenhausen	23.71	87950	3709	24.4	1	48.157458	11.643526	4591.5
13	Berg am Laim	6.31	46098	7300	31.9	0	48.126164	11.631715	2205.1
14	Trudering-Riem	22.45	73206	3261	23.3	0	48.123764	11.683573	3946.5
15	Ramersdorf-Perlach	19.90	116327	5847	33.9	0	48.099738	11.638702	4277.9
16	Obergiesing-Fasangarten	5.72	54256	9485	31.1	1	48.102744	11.593269	2349.6
17	Untergiesing-Harlaching	8.06	53184	6601	24.1	1	48.098586	11.564134	3303.9
18	Thalkirchen-Obersendling-Forstenried-Fürstenri...	17.76	96714	5445	27.4	1	48.086525	11.513238	3492.3
19	Hadern	9.22	49898	5410	27.3	0	48.113956	11.483451	2206.6
20	Pasing-Obermenzing	16.50	74625	4523	22.9	0	48.152697	11.463274	2898.3
21	Aubing-Lochhausen-Langwied	34.06	47813	1404	28.4	0	48.164726	11.408340	4814.5
22	Allach-Untermenzing	15.45	33355	2159	24.2	0	48.192671	11.462006	3449.9
23	Feldmoching-Hasenbergl	28.94	61774	2135	32.4	0	48.215176	11.521979	4254.7
24	Laim	5.29	56546	10698	28.5	0	48.136082	11.504656	2074.8

Table 1: List of official names of Munich boroughs, supplemented by geodata

To get an impression what the data looks like, Fig. 3 shows a map with the *centroids* and the *extent* of each borough. The minimum radius is 1190.7m. The maximum radius is 4974.5m. Due to the large spread it is clear that the radius for each individual borough must be taken into account when requesting data from [Foursquare](https://foursquare.com).

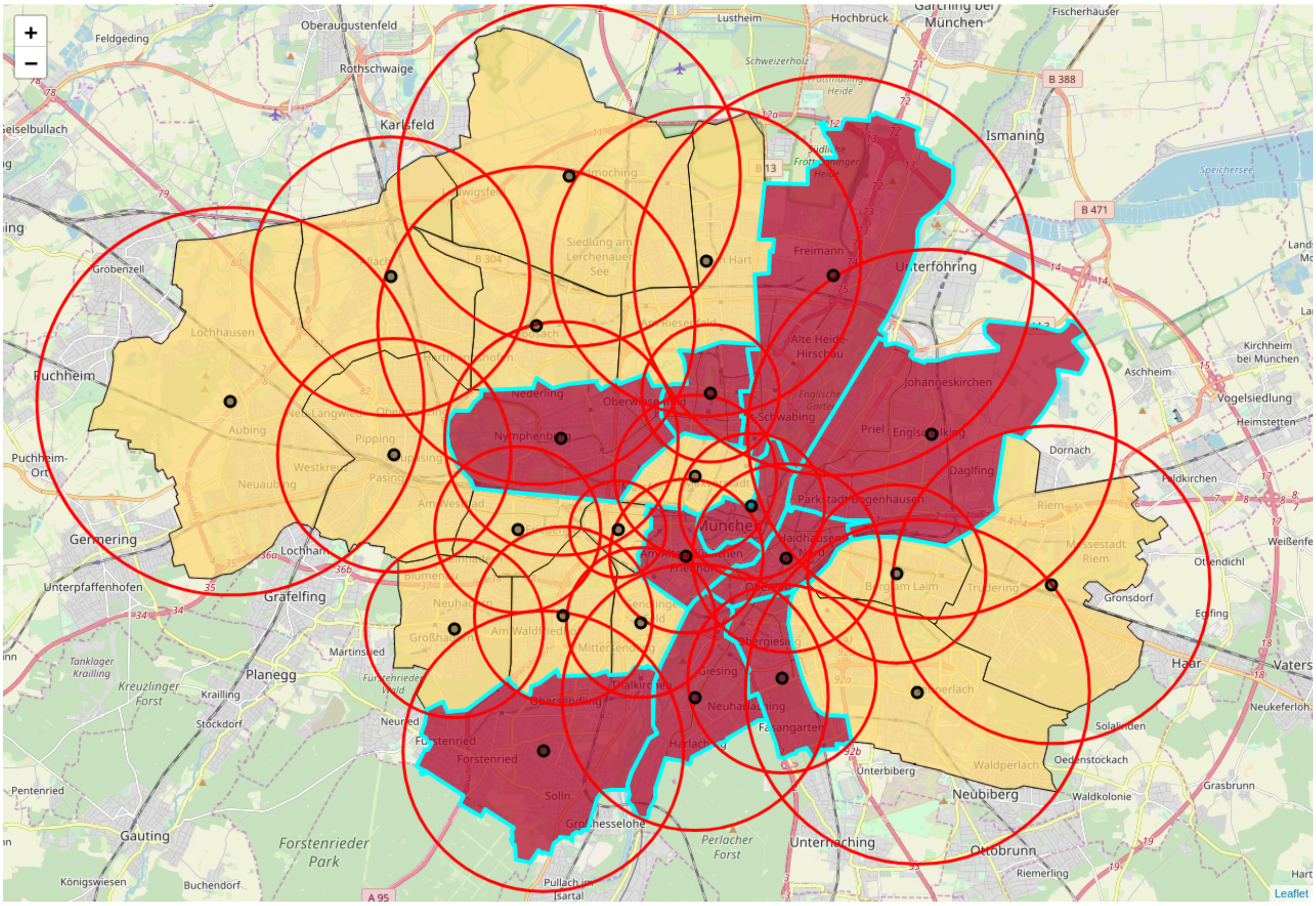

Fig. 3: Boroughs of Munich. [Best Neighborhoods in Munich]() are drawn as red filled polygons with a cyan border. The calculated centroids are shown as black dots and the red circles illustrate the calculated extent (radius(m)) for each borough, i.e. the maximum distance from the centroid to the border of a borough.

2.5 Retrieving venue data using [Foursquare](https://foursquare.com)

The venue data of Munich's boroughs is retrieved by using foursquare.com. For each borough, requests are carried out using the individual *centroid* location (latitudes, longitudes) and the individually calculated *radius(m)* to obtain the venues in each borough. The *radius(m)* is scaled by a list of *radiusFactors* to retrieve several lists of venues in different distances around the centroid of each borough.
The used radiusFactors are: 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5.

By applying this procedure, a list with more than 25000 entries is created, which contains the following categories:

```
Borough               object
Borough_Latitude     float64
Borough_Longitude    float64
radiusFactor         float64
Venue                 object
Venue_Latitude       float64
Venue_Longitude      float64
Venue_Category        object
```

3 Methodology and Analysis

This chapter acts as a linking element between the data wrangling and the following analysis sections and is therefore the main component of this report.

The idea of this investigation is quite simple:

1) Retrieve venues for each borough using Foursquare-API.
2) Do some data manipulation to calculate mean values of the occurrence of individual venue categories for each borough.
3) Use a *k-means* clustering algorithm and force all boroughs into two clusters: cluster one contains the *best* boroughs. This machine learning algorithm is used as it is the most suitable for this kind of categorization.
4) Do steps 1 to 3 for several radii around each borough's centroid, i.e. 50% to 150% of each borough's extent/radius.
5) For the several radii of 4) sum up all how many times a borough was classified as *best* borough.
6) Compare the result of 5) with the list of the website Best Neighborhoods in Munich

Is it possible to confirm the selection of the Best Neighborhoods in Munich by this simple analysis?

Are there boroughs with a similar “lifestyle” like the Best Neighborhoods in Munich but not listed?

The following analysis sections will show...

3.1 Venues of each Borough

Fig. 4 shows for all boroughs the venues that lie within a given radius. Here the smallest radius was applied by scaling the given *radius(m)* with the smallest available *radiusFactor* of 0.5.

It is easy to see that even with a *radiusFactor* of 0.5, there is still an overlap between the areas of influence in some boroughs. This means that there are boroughs that are also influenced by the venues of their neighboring boroughs.

Table 2 lists the boroughs of Munich and how many venues are located inside a certain distance to the centroid of each borough. Distance is calculated by scaling the extent of each borough (*radius(m)*) with a certain *radiusFactor* (*rFactor_x.x*).

The maximum number of hits seems to be limited to 100. In some boroughs, the number of hits increases as the `radiusFactor` increases. However, there are also sporadic irregular fluctuations in the quantities.

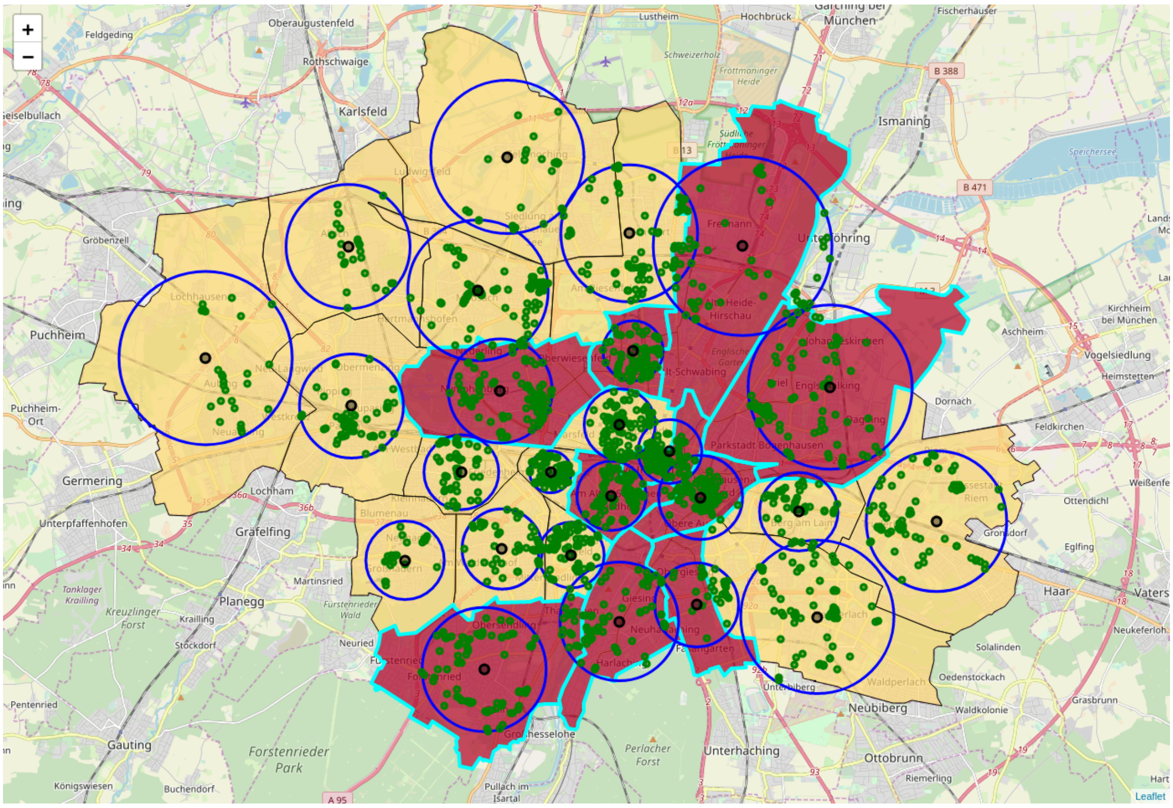

Fig. 4: Boroughs of Munich. Best Neighborhoods in Munich are drawn as red filled polygons with a cyan border. The green dots represent the venues retrieved from Foursquare in the catchment area around the centroid (black dots) of each borough. Blue circles illustrate the radius of each borough with a radiusFactor of 0.5 applied to radius(m) of each borough.

Borough	rFactor_0.5	rFactor_0.6	rFactor_0.7	rFactor_0.8	rFactor_0.9	rFactor_1.0	rFactor_1.1	rFactor_1.2	rFactor_1.3	rFactor_1.4	rFactor_1.5
Allach-Untermenzing	24	26	31	44	48	63	100	73	88	100	100
Altstadt-Lehel	100	100	100	100	100	100	100	100	100	100	100
Au-Haidhausen	100	100	100	100	100	100	100	100	100	100	100
Aubing-Lochhausen-Langwied	25	30	42	61	92	100	100	100	100	100	100
Berg am Laim	51	62	61	89	98	100	100	100	100	100	100
Bogenhausen	100	100	99	100	100	100	100	100	100	100	100
Feldmoching-Hasenbergl	24	30	41	59	52	92	100	100	100	100	100
Hadern	25	26	36	39	49	47	66	79	100	100	100
Laim	59	54	81	96	100	100	100	100	100	100	100
Ludwigsvorstadt-Isarvorstadt	100	100	100	100	100	100	100	100	100	100	100
Maxvorstadt	100	100	100	100	100	100	100	100	100	100	100
Milbertshofen-Am Hart	87	100	68	98	100	100	100	100	100	100	100
Moosach	93	100	100	100	100	100	100	100	100	100	100
Neuhausen-Nymphenburg	100	100	100	100	100	100	100	100	100	100	100
Obergiesing-Fasangarten	40	46	76	92	100	100	100	100	100	100	100
Pasing-Obermenzing	65	100	100	100	100	78	100	100	100	100	100
Ramersdorf-Perlach	88	91	76	92	100	100	100	100	100	100	100
Schwabing-Freimann	55	87	100	100	100	100	100	100	100	100	100
Schwabing-West	96	45	63	94	100	100	100	100	100	100	100
Schwanthalerhöhe	72	83	96	100	100	100	100	100	100	100	100
Sendling	87	100	100	100	100	100	100	100	100	100	100
Sendling-Westpark	51	70	77	100	100	80	100	100	100	100	100
Thalkirchen-Obersendling-Forstenried-Fürstenried-Solln	83	70	100	100	100	100	100	100	100	100	100
Trudering-Riem	90	100	95	100	100	100	100	100	100	100	100
Untergiesing-Harlaching	73	98	100	100	100	100	100	100	100	100	100

Table 2: List of Munich boroughs and how many venues are located inside a certain distance to the centroid of each borough. Distance is calculated by scaling the extent of each borough (*radius(m)*) with a certain *radiusFactor* (*rFactor_x.x*)

The map in Fig. 5 shows the boroughs of Munich with the borough of *Hadern* as an example to explain details on the approach of using different radii.

The *radiusFactors* (0.5, 1.0, 1.5) are visualized accordingly by circles in three different colors (blue, red, green). The retrieved venues of each radius are marked in the corresponding color.

It turns out that some venues of a smaller circle are not included in the group of a larger circle. There might be two reasons for that kind of behavior:

- The number of results is limited to 100. It is not known which venues will be provided if more are available within a certain radius.
- [Foursquare](#) *finds venues that a typical user is likely to checkin to at the provided location*, see [API documentation](#). Again, the selection method is unknown.

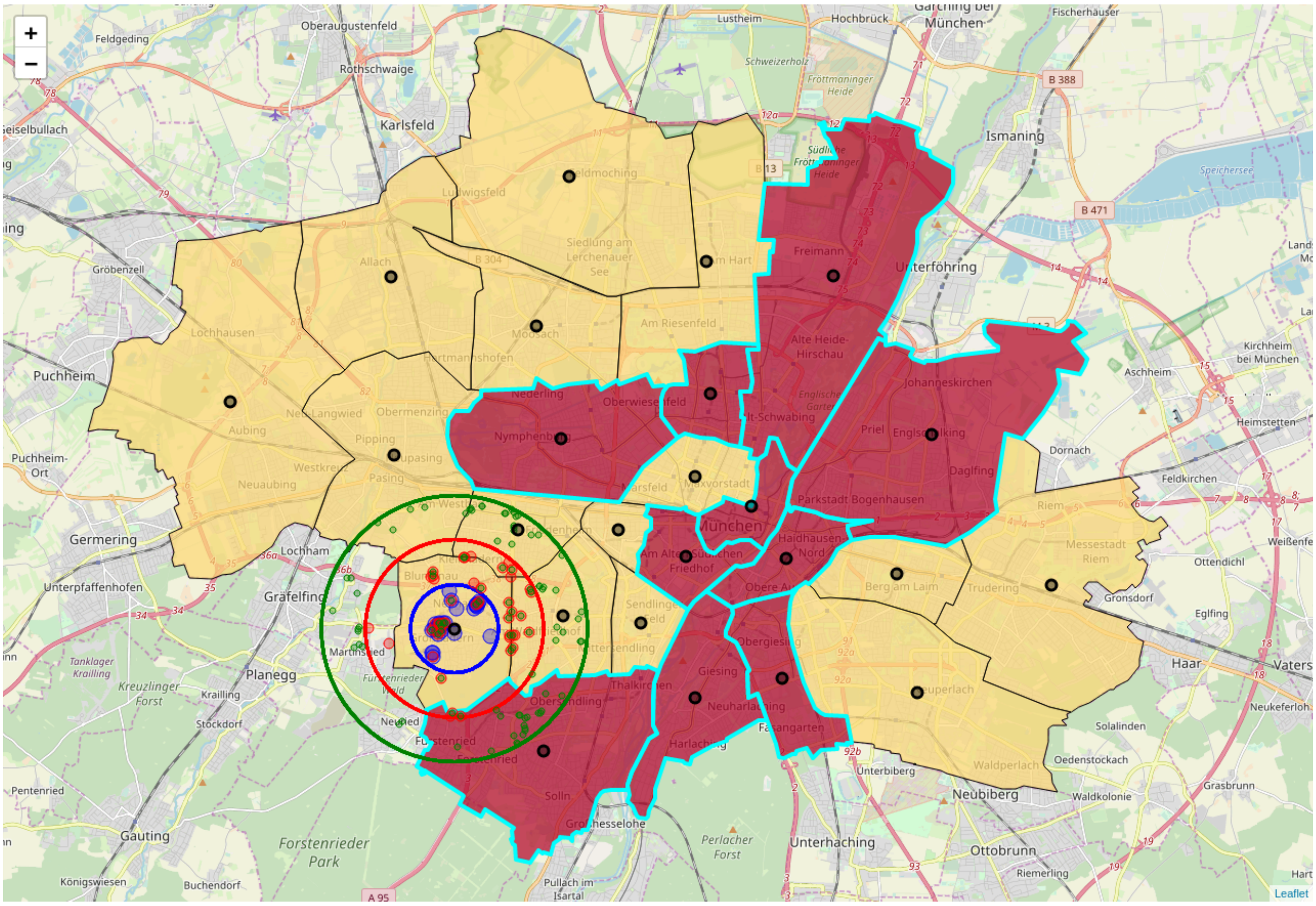

Fig. 5: Boroughs of Munich. [Best Neighborhoods in Munich](#) are drawn as red filled polygons with a cyan border. The calculated centroids are shown as black dots. The borough “Hadern” is taken as example to illustrate the method of retrieving venues (colored dots) from [Foursquare](#) for three different radiusFactors (0.5=blue, 1.0=red, 1.5=green)

3.2 Categories of Venues

The retrieved venues are assigned to 291 unique categories:

```
['Trattoria/Osteria' 'German Restaurant' 'Garden' 'Hotel' 'Steakhouse'
 'Pastry Shop' 'Opera House' 'Mediterranean Restaurant' 'Plaza'
 'Cocktail Bar' 'Palace' 'Historic Site' 'Concert Hall' 'Surf Spot'
 'Italian Restaurant' 'Theater' 'Boutique' 'Convenience Store'
 'Art Museum' 'Café' 'Snack Place' 'Restaurant' 'Vietnamese Restaurant'
 'Gourmet Shop' 'Bavarian Restaurant' 'Department Store'
 'Performing Arts Venue' 'American Restaurant' 'Pizza Place'
 'Shopping Mall' 'French Restaurant' 'Waterfall' 'Austrian Restaurant'
 'Camera Store' 'Candy Store' 'Clothing Store' 'Bar' 'Fountain' 'Wine Bar'
 'Museum' 'Monument / Landmark' 'Outdoor Sculpture' 'Beach' 'Coffee Shop'
 'Design Studio' 'Art Gallery' 'Irish Pub' 'Sporting Goods Shop' 'Church'
 'Japanese Restaurant' 'Tree' 'Greek Restaurant' 'Asian Restaurant'
 'Ice Cream Shop' 'Xinjiang Restaurant' 'Spanish Restaurant'
 'Vegetarian / Vegan Restaurant' 'Drugstore' 'Nightclub'
 'Falafel Restaurant' 'Creperie' 'Rock Club' 'Sandwich Place' 'Brewery'
 'Indian Restaurant' 'Movie Theater' 'Supermarket' 'Bistro'
 'Afghan Restaurant' 'Burger Joint' 'Indie Movie Theater'
 'Seafood Restaurant' 'Noodle House' 'Taverna' 'Middle Eastern Restaurant'
 'Tea Room' 'Chinese Restaurant' 'Dim Sum Restaurant' 'Salon / Barbershop'
 'Sushi Restaurant' 'Jazz Club' 'Portuguese Restaurant' 'Beer Bar' 'Pub'
 'Bakery' 'Event Space' 'Roof Deck' 'Burrito Place' 'History Museum'
 'Peruvian Restaurant' 'Martial Arts Dojo' 'Ramen Restaurant' 'Park'
 'Arcade' 'Gym / Fitness Center' 'Israeli Restaurant' 'Thai Restaurant'
 'Diner' 'Sausage Shop' 'Food & Drink Shop' 'Botanical Garden'
 'Doner Restaurant' 'Hotel Bar' 'Modern European Restaurant'
 'Arts & Crafts Store' 'Gastropub' 'Tapas Restaurant' 'Beer Store'
 'Recreation Center' 'Ethiopian Restaurant' 'Mexican Restaurant' 'Pool'
 'Hill' 'Spa' 'Post Office' 'Tram Station' 'Metro Station' 'Pharmacy'
 'Bus Stop' 'Jewish Restaurant' 'Turkish Restaurant' 'Cultural Center'
 'Organic Grocery' 'Music Venue' 'Beer Garden' 'Comedy Club'
 'Science Museum' 'Dance Studio' 'Climbing Gym' 'Planetarium' 'Fair'
 'Bosnian Restaurant' 'Gym' 'Deli / Bodega' 'Grocery Store' 'Food Court'
 'Wine Shop' 'Market' 'Laundry Service' 'Soccer Field'
 'Athletics & Sports' 'Butcher' 'Bank' 'Optical Shop' 'Mobile Phone Shop'
 'Bus Line' 'Print Shop' 'Gas Station' 'Light Rail Station'
 'Basketball Stadium' 'Pet Store' 'Lake' 'IT Services' 'Scenic Lookout'
 'Tennis Court' 'Shipping Store' 'Bagel Shop' 'Yoga Studio'
 'Modern Greek Restaurant' 'Salad Place' 'Lounge' 'Juice Bar' 'Canal'
 'Kebab Restaurant' 'Fast Food Restaurant' 'Soup Place'
 'English Restaurant' 'Hawaiian Restaurant' 'Szechuan Restaurant'
 'Electronics Store' 'Stadium' 'Food' 'Garden Center' 'Music Store'
 'Motel' 'Bed & Breakfast' 'Hostel' 'Outlet Store' 'Fried Chicken Joint'
 'Hardware Store' 'Donut Shop' 'Bookstore' 'Furniture / Home Store'
 'Automotive Shop' 'Motorcycle Shop' 'Indoor Play Area'
 'Paper / Office Supplies Store' 'Rental Car Location'
 'Rock Climbing Spot' 'Food Truck' 'Sports Bar' 'Lottery Retailer'
 'Train Station' 'Liquor Store' 'Business Service' 'Malay Restaurant'
 'Bowling Alley' 'Farmers Market' 'Racetrack' 'Flower Shop' 'Water Park'
 'Tunnel' 'Dog Run' 'Big Box Store' 'Discount Store' 'Hookah Bar'
 'Eastern European Restaurant' 'Storage Facility' 'Toy / Game Store'
 'Hot Dog Joint' 'Golf Course' 'Auto Dealership' 'Track' 'Trail'
 'Tour Provider' 'Sports Club' 'Dessert Shop' 'Baby Store' 'Hockey Arena'
 "Women's Store" 'Halal Restaurant' 'Pie Shop' 'Office' 'Zoo'
 'Zoo Exhibit' 'BBQ Joint' 'Gym Pool' 'Aquarium' 'Soccer Stadium'
 'Campground' 'Skate Park' 'Theme Restaurant' 'Castle'
 'Schnitzel Restaurant' 'Nature Preserve' 'Breakfast Spot'
 'Health & Beauty Service' 'Intersection' 'Caucasian Restaurant'
 'Gymnastics Gym' 'Platform' 'Laundromat' 'Festival' 'Gift Shop'
 'Poke Place' 'Cretan Restaurant' 'Coworking Space'
 'Comfort Food Restaurant' 'River' 'Tibetan Restaurant' 'Boarding House'
 'Outdoor Supply Store' 'Accessories Store' 'Cafeteria' 'Flea Market'
 'Mountain' 'Fruit & Vegetable Store' 'Forest' 'Boat Rental'
 'Residential Building (Apartment / Condo)' 'Skating Rink' 'Smoke Shop'
 'Pide Place' 'Fish Market' 'Radio Station' 'Korean Restaurant'
 'Indie Theater' 'Auto Garage' "Men's Store" 'Comic Shop' 'Manti Place'
 'Cupcake Shop' 'African Restaurant' 'General Entertainment' 'Taco Place'
 'Lebanese Restaurant' 'Public Art' 'Frozen Yogurt Shop' 'Beach Bar'
 'Playground' 'Cheese Shop' 'Airport' 'Kitchen Supply Store' 'Circus'
 'Mongolian Restaurant' 'Heliport' 'Hobby Shop' 'Shoe Store'
 'Drive-in Theater' 'Convention Center' 'Farm']
```

3.3 Most common Venues for each Borough

To obtain a list of the most common venues for each borough a *one hot encoding* on the venue data is applied. The result is grouped by *Borough* using the `mean()`-aggregation method.

Table 3 gives an extract from the list of mean values of the occurrence of individual venue categories for each district of Munich using the *radiusFactor* of 1.0.

On the basis of this data the top 10 venues for each borough is calculated. Table 4 presents an extract from the list of the most common venues for each borough of Munich using the *radiusFactor* of 1.0

	Borough	Accessories Store	Afghan Restaurant	Airport	American Restaurant	Arcade	Art Gallery	Art Museum	Arts & Crafts Store	Asian Restaurant	...	Turkish Restaurant	Vegetarian / Vegan Restaurant	Vietnamese Restaurant	Waterfall	Wine Bar	Wine Shop	Xinjiang Restaurant	Yoga Studio	Zoo
0	Allach-Untermenzing	0.00	0.00	0.00000	0.000000	0.00	0.00	0.00	0.00	0.000000	...	0.015873	0.015873	0.015873	0.00	0.00	0.00	0.00	0.00	0.00
1	Altstadt-Lehel	0.00	0.00	0.00000	0.010000	0.00	0.01	0.01	0.00	0.000000	...	0.000000	0.000000	0.000000	0.01	0.01	0.00	0.00	0.01	0.00
2	Au-Haidhausen	0.00	0.01	0.00000	0.000000	0.00	0.01	0.01	0.00	0.000000	...	0.010000	0.000000	0.010000	0.00	0.02	0.00	0.00	0.01	0.00
3	Aubing-Lochhausen-Langwied	0.00	0.00	0.00000	0.000000	0.00	0.00	0.00	0.00	0.010000	...	0.000000	0.000000	0.010000	0.00	0.00	0.00	0.00	0.00	0.00
4	Berg am Laim	0.01	0.00	0.00000	0.000000	0.00	0.00	0.00	0.00	0.010000	...	0.000000	0.010000	0.010000	0.00	0.00	0.00	0.00	0.00	0.00
5	Bogenhausen	0.00	0.00	0.00000	0.000000	0.00	0.00	0.00	0.00	0.010000	...	0.000000	0.000000	0.020000	0.00	0.00	0.00	0.00	0.00	0.00
6	Feldmoching-Hasenbergl	0.00	0.00	0.01087	0.021739	0.00	0.00	0.00	0.00	0.010870	...	0.000000	0.000000	0.000000	0.00	0.00	0.00	0.00	0.00	0.00
7	Hadern	0.00	0.00	0.00000	0.000000	0.00	0.00	0.00	0.00	0.000000	...	0.000000	0.000000	0.000000	0.00	0.00	0.00	0.00	0.00	0.00
8	Laim	0.00	0.00	0.00000	0.020000	0.00	0.00	0.00	0.00	0.010000	...	0.000000	0.000000	0.000000	0.00	0.00	0.00	0.00	0.01	0.00
9	Ludwigsvorstadt-Isarvorstadt	0.00	0.01	0.00000	0.000000	0.00	0.01	0.00	0.01	0.010000	...	0.000000	0.010000	0.020000	0.00	0.01	0.00	0.00	0.00	0.00
10	Maxvorstadt	0.00	0.00	0.00000	0.000000	0.01	0.00	0.04	0.01	0.000000	...	0.000000	0.010000	0.020000	0.00	0.00	0.00	0.00	0.00	0.00
11	Milbertshofen-Am Hart	0.00	0.00	0.00000	0.000000	0.00	0.00	0.00	0.00	0.010000	...	0.000000	0.000000	0.000000	0.00	0.00	0.00	0.00	0.00	0.00

Table 3: Extract from the list of mean values of the occurrence of individual venue categories for each borough of Munich using the radiusFactor of 1.0

	Borough	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	Allach-Untermenzing	Supermarket	German Restaurant	Italian Restaurant	Hotel	Bus Stop	Drugstore	Garden Center	Beer Garden	Bavarian Restaurant	Intersection
1	Altstadt-Lehel	Café	Plaza	Hotel	Cocktail Bar	German Restaurant	Italian Restaurant	Bavarian Restaurant	Bookstore	Gourmet Shop	Boutique
2	Au-Haidhausen	German Restaurant	Café	Cocktail Bar	Plaza	Hotel	Italian Restaurant	Restaurant	Coffee Shop	Steakhouse	Concert Hall
3	Aubing-Lochhausen-Langwied	Supermarket	German Restaurant	Drugstore	Bakery	Greek Restaurant	Italian Restaurant	Hotel	Coffee Shop	Café	Ice Cream Shop
4	Berg am Laim	Supermarket	Hotel	Italian Restaurant	German Restaurant	Drugstore	Gym	Bakery	Greek Restaurant	Chinese Restaurant	Food & Drink Shop
5	Bogenhausen	German Restaurant	Italian Restaurant	Hotel	Gym / Fitness Center	Restaurant	Beer Garden	Bakery	Park	Supermarket	Gourmet Shop
6	Feldmoching-Hasenbergl	Supermarket	Hotel	Bakery	Drugstore	Greek Restaurant	Italian Restaurant	Lake	Electronics Store	Indian Restaurant	Intersection
7	Hadern	Supermarket	German Restaurant	Bakery	Bus Stop	Trattoria/Osteria	Drugstore	Italian Restaurant	Sushi Restaurant	Bank	Ice Cream Shop
8	Laim	Supermarket	Greek Restaurant	Drugstore	Italian Restaurant	Hotel	Gym / Fitness Center	Plaza	Beer Garden	German Restaurant	Bank
9	Ludwigsvorstadt-Isarvorstadt	Café	Hotel	German Restaurant	Ice Cream Shop	Bar	Italian Restaurant	Plaza	Cocktail Bar	Bavarian Restaurant	Coffee Shop
10	Maxvorstadt	Café	Plaza	Hotel	Art Museum	Italian Restaurant	Bar	German Restaurant	Ice Cream Shop	Steakhouse	Hotel Bar
11	Milbertshofen-Am Hart	Hotel	Café	Gym / Fitness Center	Greek Restaurant	Restaurant	Museum	Park	Lounge	Pool	Bakery

Table 4: Extract from the list of the most common venues for each borough of Munich using the radiusFactor of 1.0

3.4 Clustering Boroughs on base of their Venues

The mean values of the occurrence of individual venue categories for each district of Munich (see table 3 for example) are the input for a K-means clustering algorithm. In this case the algorithm is used to separate the boroughs into two clusters. The clustering is done for each *radiusFactor.*

In the Appendix the result for each *radiusFactor* is visualized on a map of Munich.

For all *radiusFactors* the number of occurrences of recommendations (cluster one) is summed up and declared as *score*. This *score* can be compared with the *BestBorough*-label recommendations from www.moving-to-munich.com, see table 5.

Fig. 6 uses a choropleth map to give a better overview: The different colors of the boroughs illustrate the score presented in table 5, i.e. how often a individual borough was included in the respective cluster for the best "boroughs" using the K-means clustering algorithm for several *radiusFactors*. The redder the color, the more often the respective borough was marked to be associated to the recommended boroughs.

	Borough	bestBorough	score	rFactor_0.5	rFactor_0.6	rFactor_0.7	rFactor_0.8	rFactor_0.9	rFactor_1.0	rFactor_1.1	rFactor_1.2	rFactor_1.3	rFactor_1.4	rFactor_1.5
0	Altstadt-Lehel	1	11	1	1	1	1	1	1	1	1	1	1	1
1	Ludwigsvorstadt-Isarvorstadt	1	11	1	1	1	1	1	1	1	1	1	1	1
2	Maxvorstadt	0	11	1	1	1	1	1	1	1	1	1	1	1
3	Schwabing-West	1	11	1	1	1	1	1	1	1	1	1	1	1
4	Au-Haidhausen	1	11	1	1	1	1	1	1	1	1	1	1	1
5	Sendling	0	11	1	1	1	1	1	1	1	1	1	1	1
6	Sendling-Westpark	0	6	0	0	0	0	0	1	1	1	1	1	1
7	Schwanthalerhöhe	0	11	1	1	1	1	1	1	1	1	1	1	1
8	Neuhausen-Nymphenburg	1	11	1	1	1	1	1	1	1	1	1	1	1
9	Moosach	0	5	0	0	0	0	1	1	1	1	1	0	0
10	Milbertshofen-Am Hart	0	7	0	0	1	1	1	1	1	1	1	0	0
11	Schwabing-Freimann	1	11	1	1	1	1	1	1	1	1	1	1	1
12	Bogenhausen	1	10	1	1	1	0	1	1	1	1	1	1	1
13	Berg am Laim	0	1	0	0	0	0	0	0	0	0	1	0	0
14	Trudering-Riem	0	1	1	0	0	0	0	0	0	0	0	0	0
15	Ramersdorf-Perlach	0	4	0	0	0	0	0	1	1	1	1	0	0
16	Obergiesing-Fasangarten	1	4	0	0	0	0	0	1	0	1	1	0	1
17	Untergiesing-Harlaching	1	10	1	1	1	1	0	1	1	1	1	1	1
18	Thalkirchen-Obersendling-Forstenried-Fürstenri...	1	4	0	0	0	0	0	1	1	1	1	0	0
19	Hadern	0	0	0	0	0	0	0	0	0	0	0	0	0
20	Pasing-Obermenzing	0	0	0	0	0	0	0	0	0	0	0	0	0
21	Aubing-Lochhausen-Langwied	0	0	0	0	0	0	0	0	0	0	0	0	0
22	Allach-Untermenzing	0	0	0	0	0	0	0	0	0	0	0	0	0
23	Feldmoching-Hasenbergl	0	0	0	0	0	0	0	0	0	0	0	0	0
24	Laim	0	3	0	0	0	0	0	0	0	0	1	1	1

Table 5: List of Munich's boroughs. The label bestBorough=1 indicates that this borough is on the list of Best Neighborhoods in Munich. The score is the sum of all occurrences of recommendations (marked with "1") over all radiusFactors (rFactor_x.x)

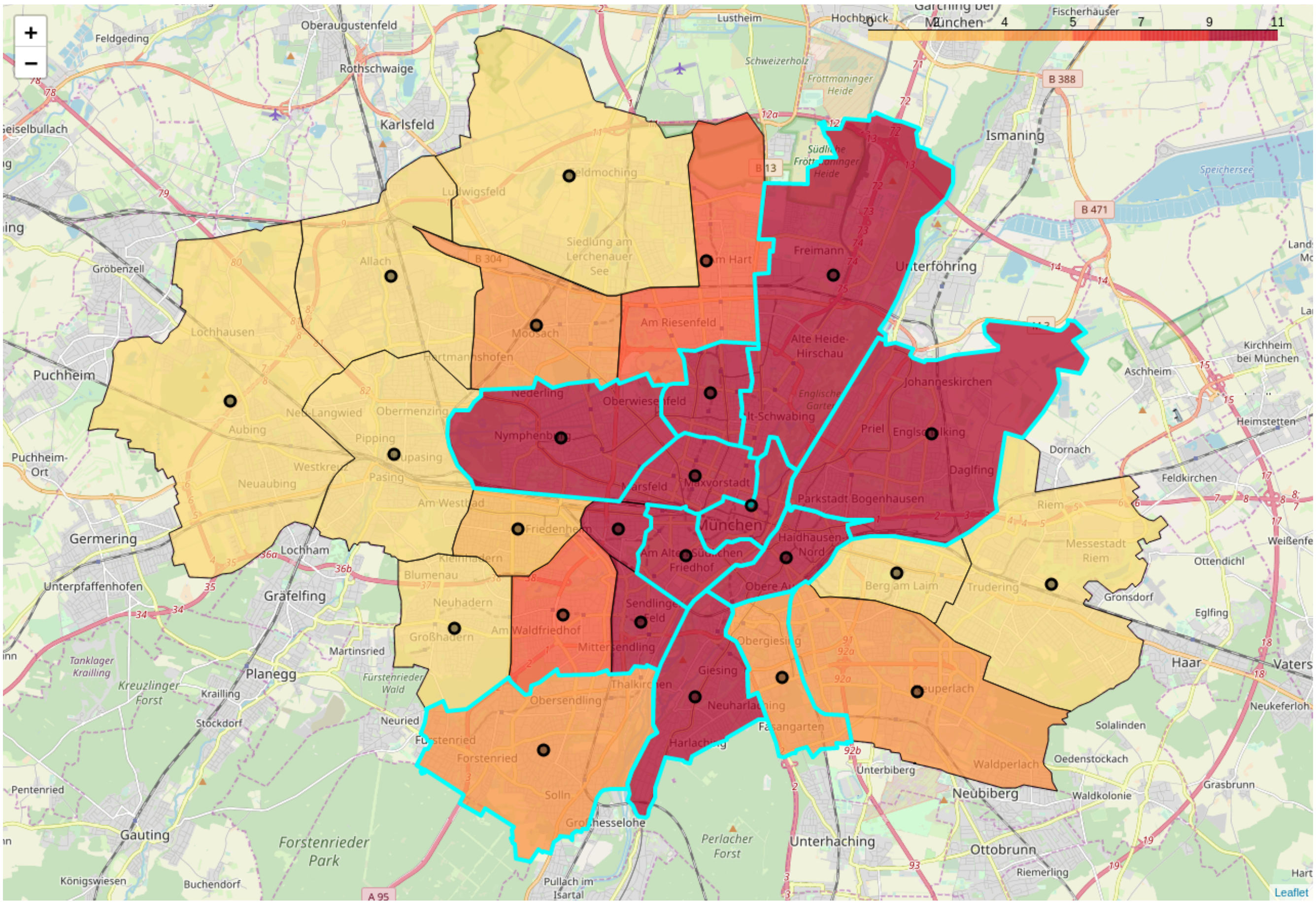

Fig. 6: Boroughs of Munich. [Best Neighborhoods in Munich](#) are drawn as polygons with a cyan border. The calculated centroids are shown as black dots. The different colors of the boroughs illustrate the score presented in table 5, i.e. how often a individual borough was included in the respective cluster for the best "boroughs" using the method of several radiusFactors.

4 Results

The map of Fig. 6 and the data of table 5 show:

- Most of the Best Neighborhoods in Munich are also recommended correctly by the clustering algorithm with the highest score of 11.
- Two of the Best Neighborhoods in Munich show almost a maximum score of 10:
    - Bogenhausen
    - Untergiesing-Harlaching
- Two other boroughs from the list of Best Neighborhoods in Munich were only able to achieve a score of 4 and are therefore not highly recommended by the clustering algorithm:
    - Obergiesing-Fasangarten
    - Thalkirchen-Obersendling-Forstenried-Fürstenried-Solln
- There are three boroughs that are not on the list of the Best Neighborhoods in Munich, but which are recommended by the clustering algorithm with the highest score of 11:
    - Maxvorstadt
    - Sendling
    - Schwanthalerhöhe

5 Discussion

In summary, the list of the Best Neighborhoods in Munich from the website www.moving-to-munich.com can be confirmed very well with the help of the *k-means* clustering algorithm.
The clustering algorithm gives an indication of which boroughs should be considered in a more differentiated way.
It also recommends three boroughs that are not on the original list of the Best Neighborhoods in Munich. So if you want to move to Munich these candidates would potentially have a comparable lifestyle, i.e. similar venues but probably lower rents (which is not proven yet because of lack of data):

- Maxvorstadt
- Sendling
- Schwanthalerhöhe

The following reservations and recommendations can be expressed about the above analyses:

- The website www.moving-to-munich.com presents a list of the Best Neighborhoods in Munich. Partly these are only parts of the official boroughs. This affects the exact comparability. Recommendation for a more in-depth analysis is given, for example by dividing the the boroughs in smaller sub-districts.
- The venue data retrieved by Foursquare is affected by the API behaviour: The number of results is limited to 100. It is not known which venues will be provided if more are available within a certain radius. Furthermore Foursquare *finds venues that a typical user is likely to checkin to at the provided location, at the current moment in time*, i.e. the exact selection method is unknown and depends on the time of the request. Here the recommendation would be to build up a more thorough venue list, for example by requesting all venues of Munich by placing a dens grid of locations with a small radius for which the venues are requested. This would avoid the limitation to 100 venues per request and the selection algorithm would be in your own hands.
- The method of clustering into two groups is quite simple. In order to obtain more confidence in the result, further statistical methods should be applied: How well are the individual boroughs comparable with each other? Do other data such as crime rates, house prices, rents, social composition of the population show similar relationships?

Last but not least:
By having a look at the map it is noticeable that the Best Neighborhoods in Munich are located at or close to the Isar river. Perhaps this is the feature that makes the boroughs particularly interesting.

6 Conclusion

The selection of the Best Neighborhoods in Munich was confirmed by the venue information provided by Foursquare.
In addition, the clustering algorithm used found further boroughs that may have similar characteristics. These would also be worth a look if you are planning to move to Munich or if you are an investor looking for real estate assets to invest in.

Due to the simplicity of the clustering algorithm used, the question could be raised whether the analysis was successful by accident.
Therefore, this project ends with the recommendation to conduct an in-depth investigation on the basis of further data not yet considered.

7 Appendix

On the following pages the results of the K-means clustering algorithm for each *radiusFactor* is visualized: the Best Neighborhoods in Munich are drawn as polygons with a cyan border. By k-means clustering recommended boroughs based on the specified *radiusFactor* are shown as red areas,

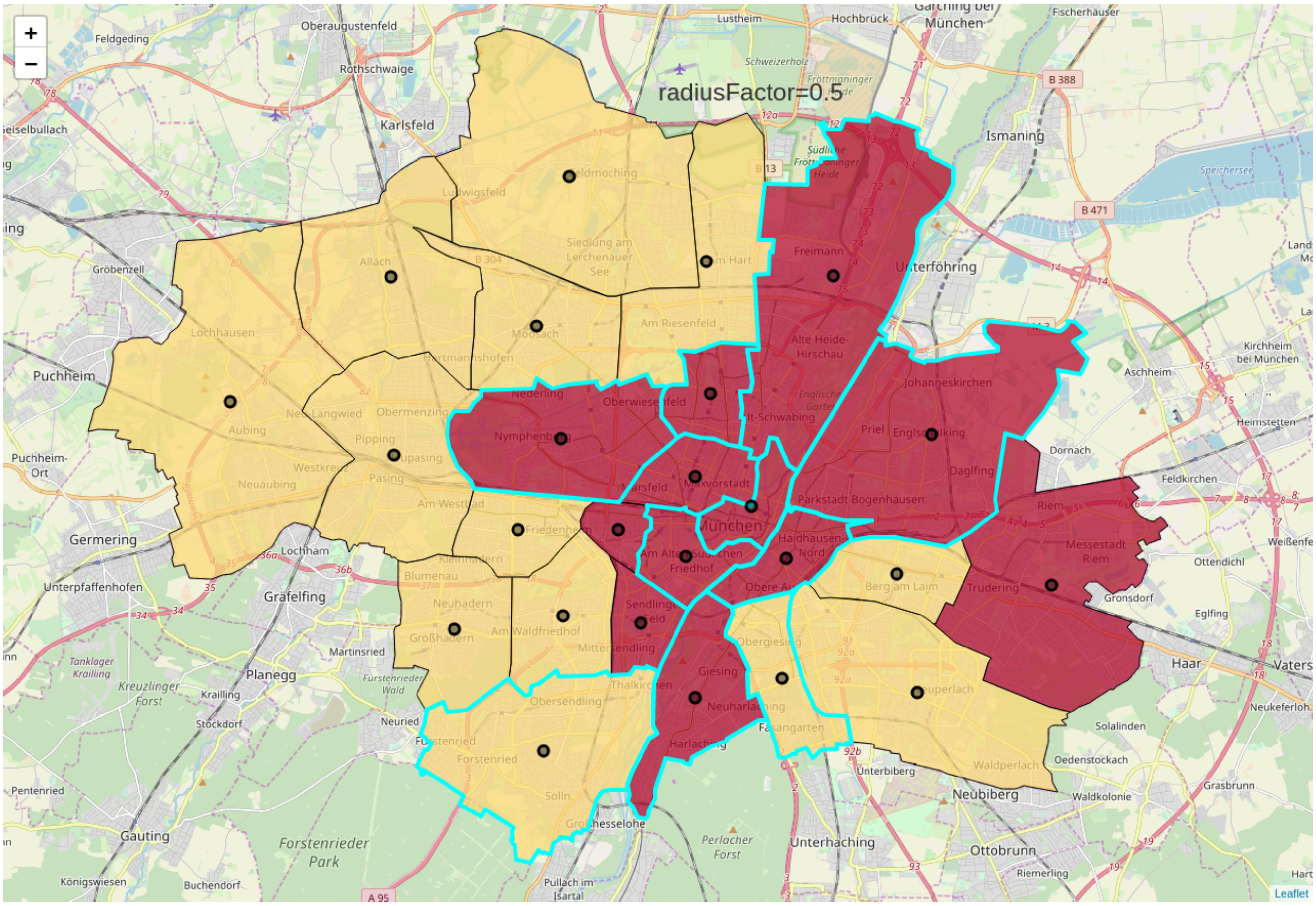

Fig. 7: Red areas show recommended boroughs for radiusFactor=0.5 by using K-means clustering. Best Neighborhoods in Munich are drawn as polygons with a cyan border.

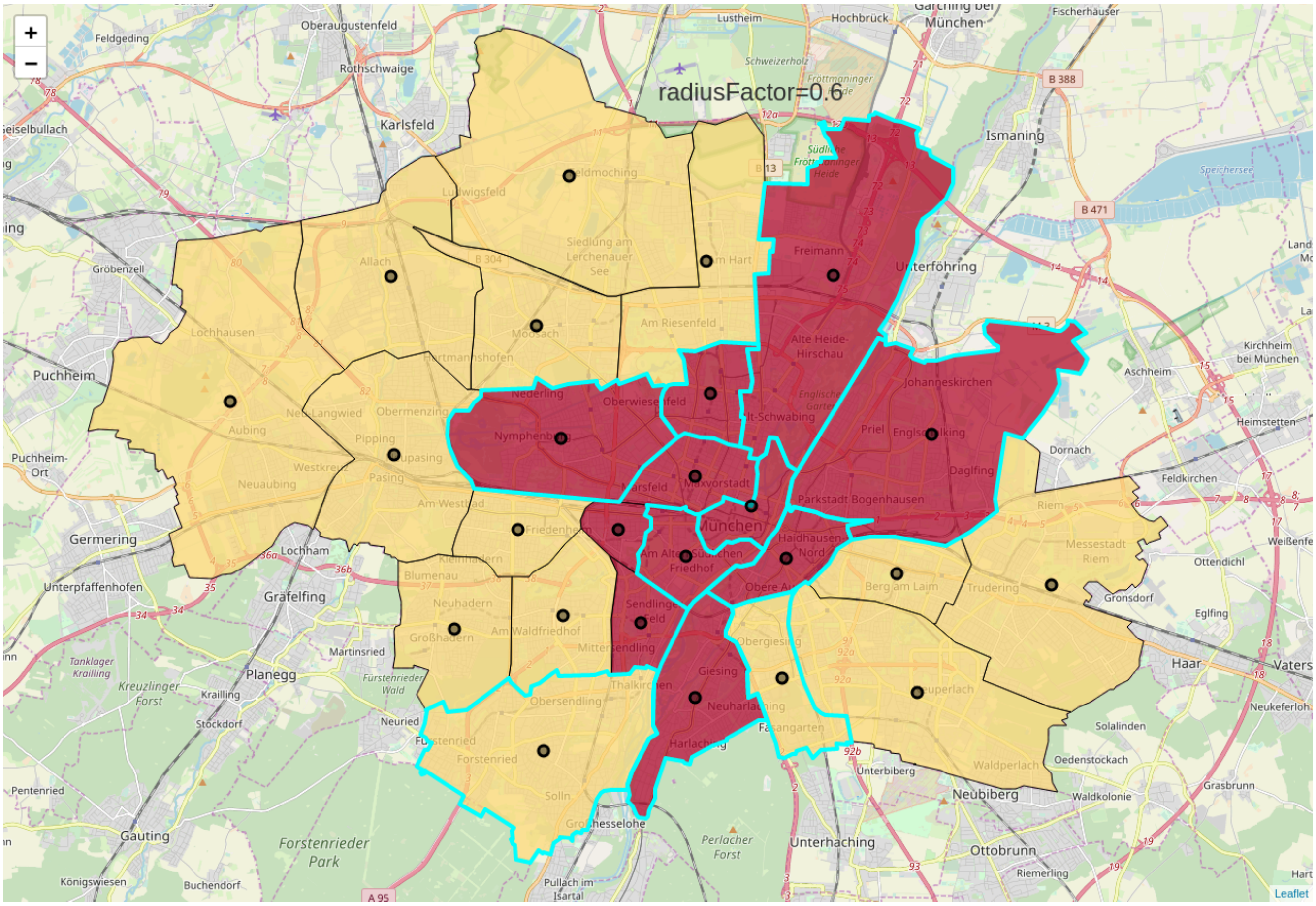

Fig. 8: Red areas show recommended boroughs for radiusFactor=0.6 by using <u>K-means clustering</u>. <u>Best Neighborhoods in Munich</u> are drawn as polygons with a cyan border.

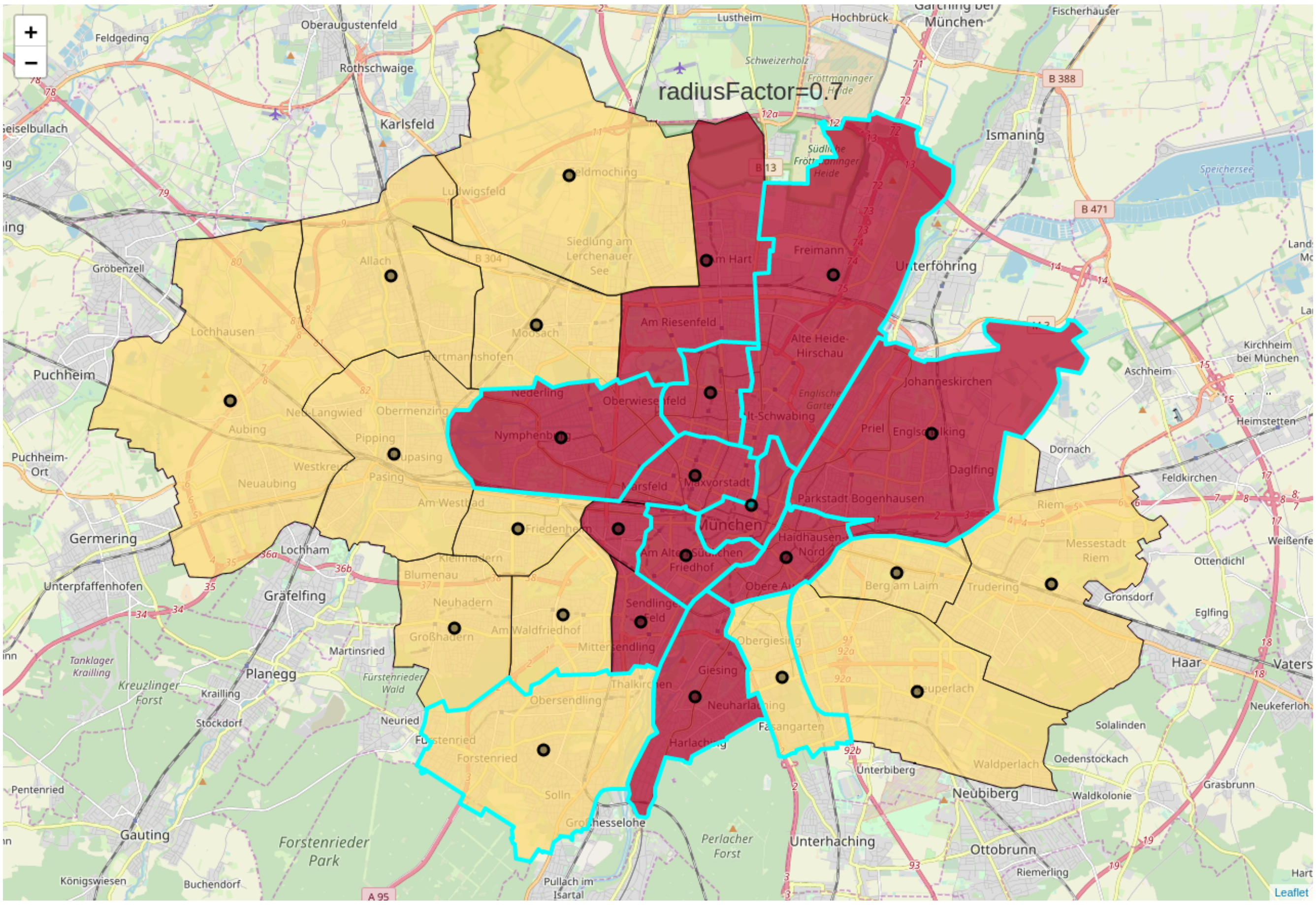

Fig. 9: Red areas show recommended boroughs for radiusFactor=0.7 by using <u>K-means clustering</u>. <u>Best Neighborhoods in Munich</u> are drawn as polygons with a cyan border.

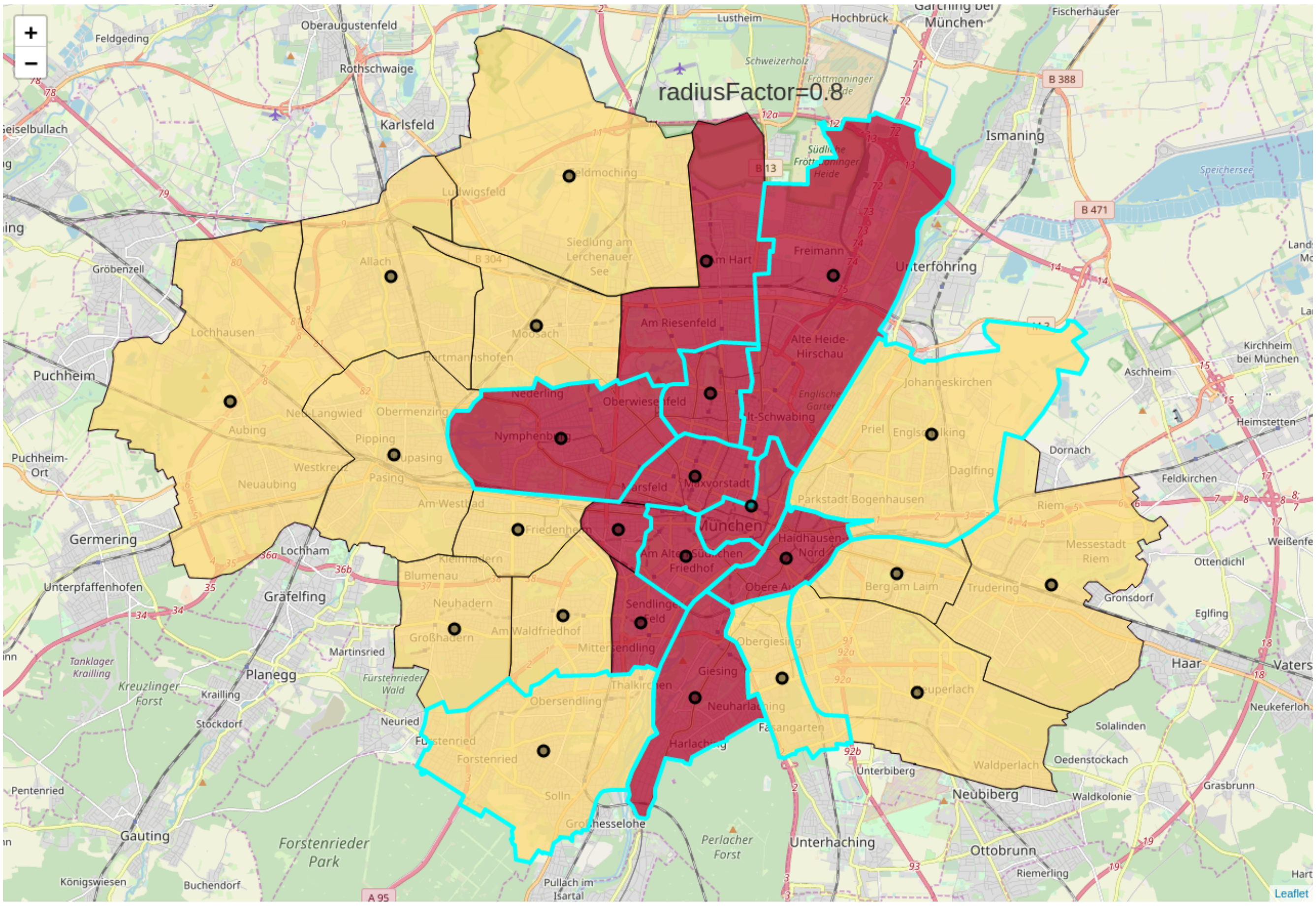

Fig. 10: Red areas show recommended boroughs for radiusFactor=0.8 by using [K-means clustering](). [Best Neighborhoods in Munich]() are drawn as polygons with a cyan border.

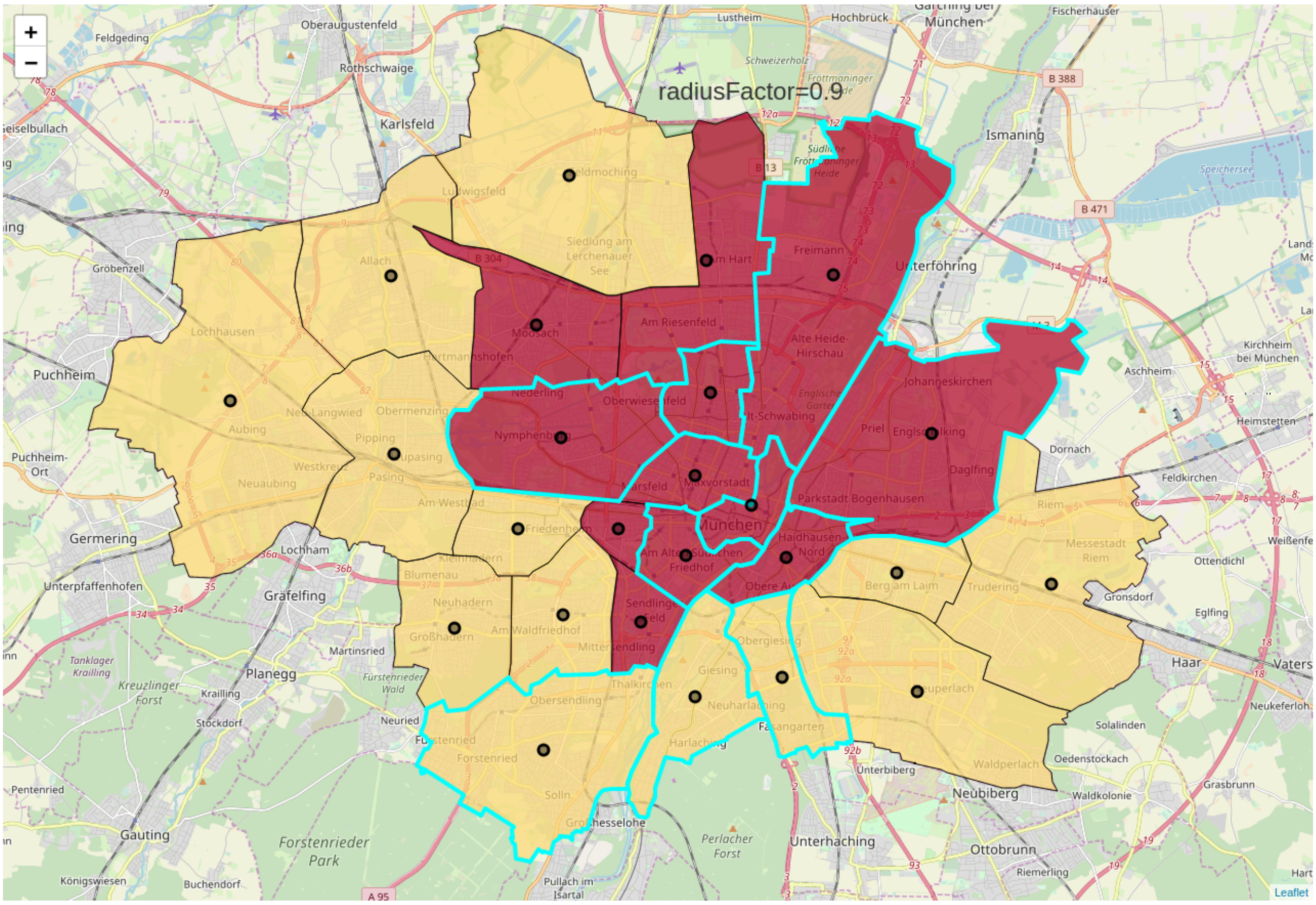

Fig. 11: Red areas show recommended boroughs for radiusFactor=0.9 by using [K-means clustering](). [Best Neighborhoods in Munich]() are drawn as polygons with a cyan border.

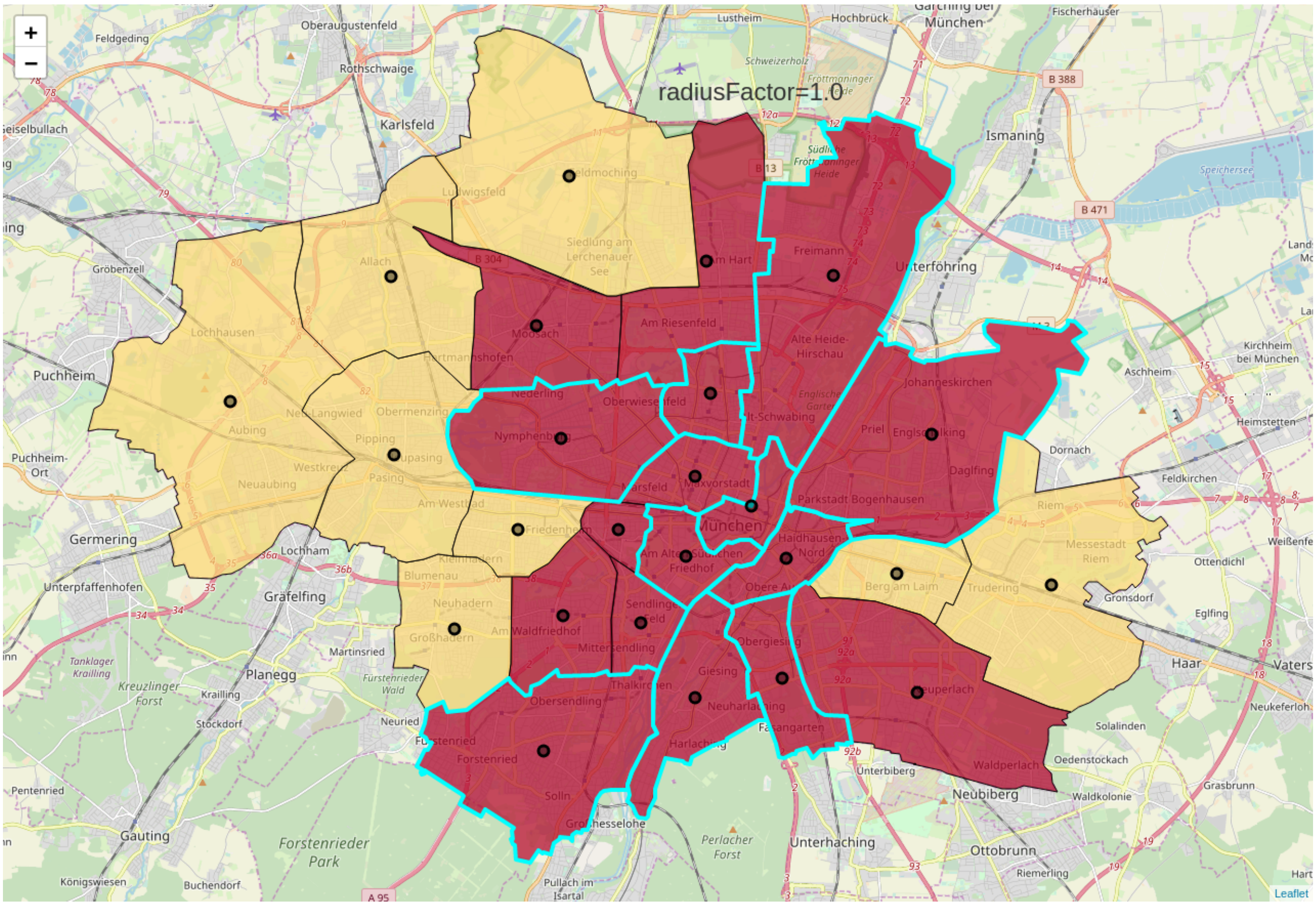

Fig. 12: Red areas show recommended boroughs for radiusFactor=1.0 by using K-means clustering. Best Neighborhoods in Munich are drawn as polygons with a cyan border.

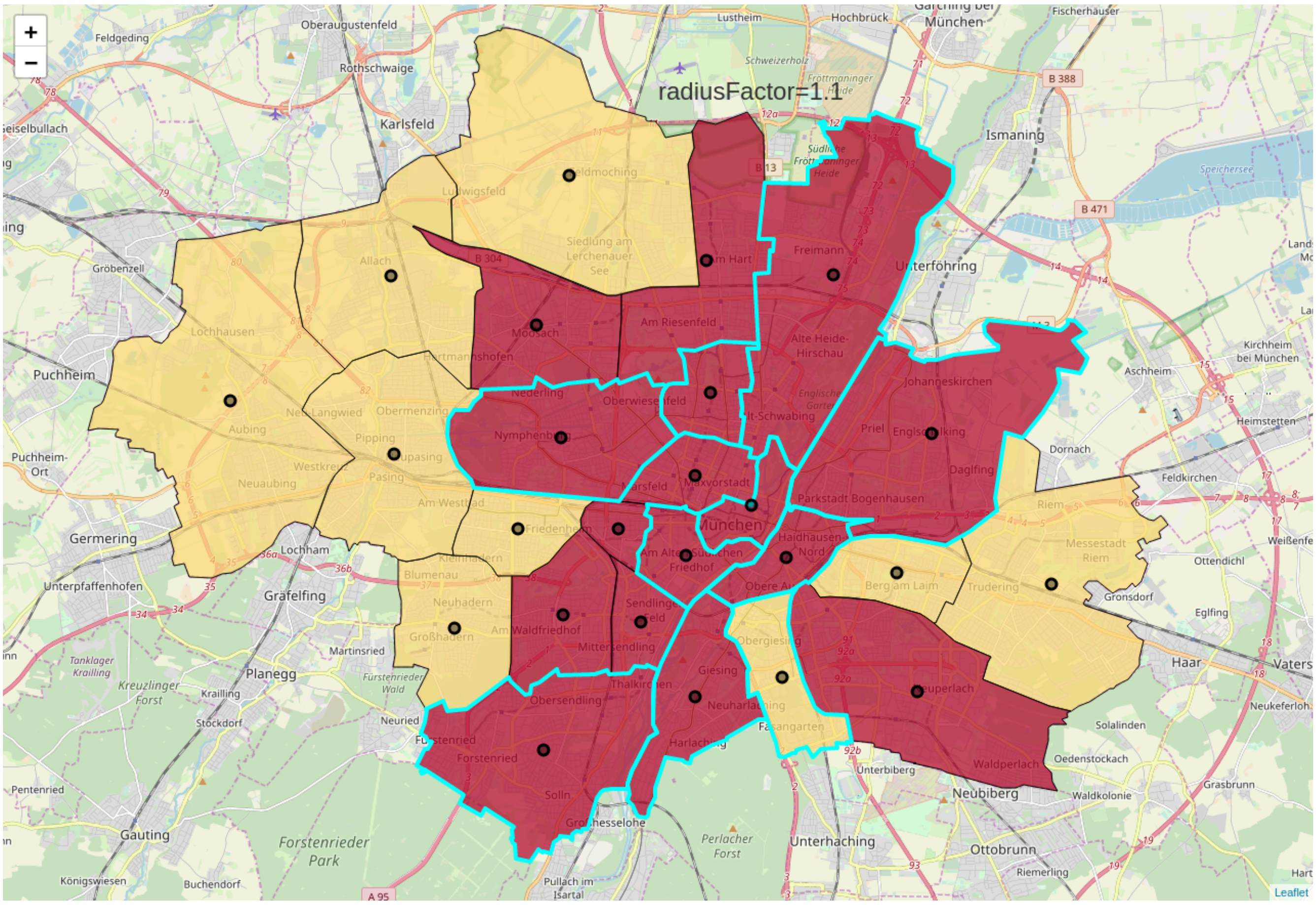

Fig. 13: Red areas show recommended boroughs for radiusFactor=1.1 by using K-means clustering. Best Neighborhoods in Munich are drawn as polygons with a cyan border.

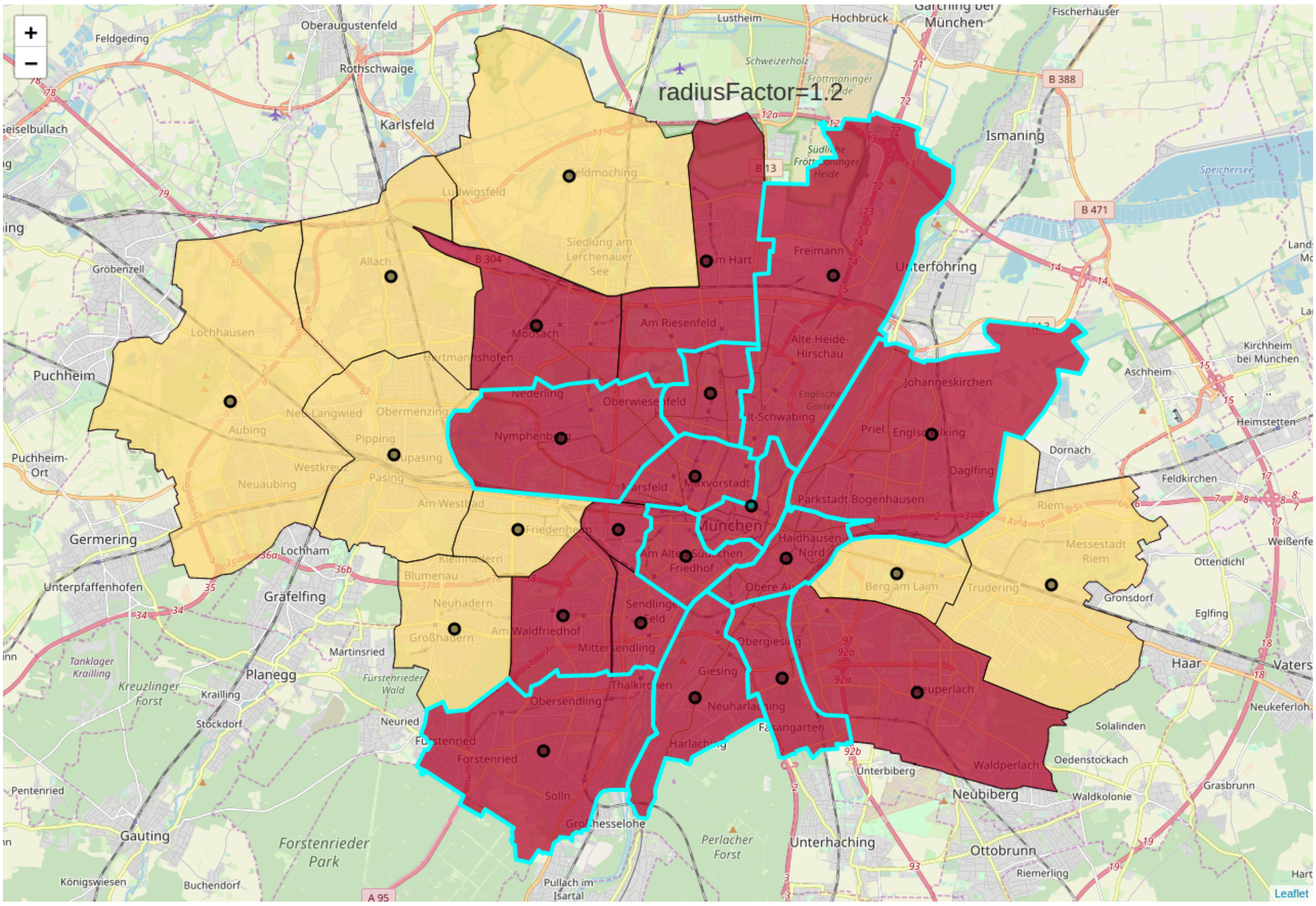

Fig. 14: Red areas show recommended boroughs for radiusFactor=1.2 by using K-means clustering. Best Neighborhoods in Munich are drawn as polygons with a cyan border.

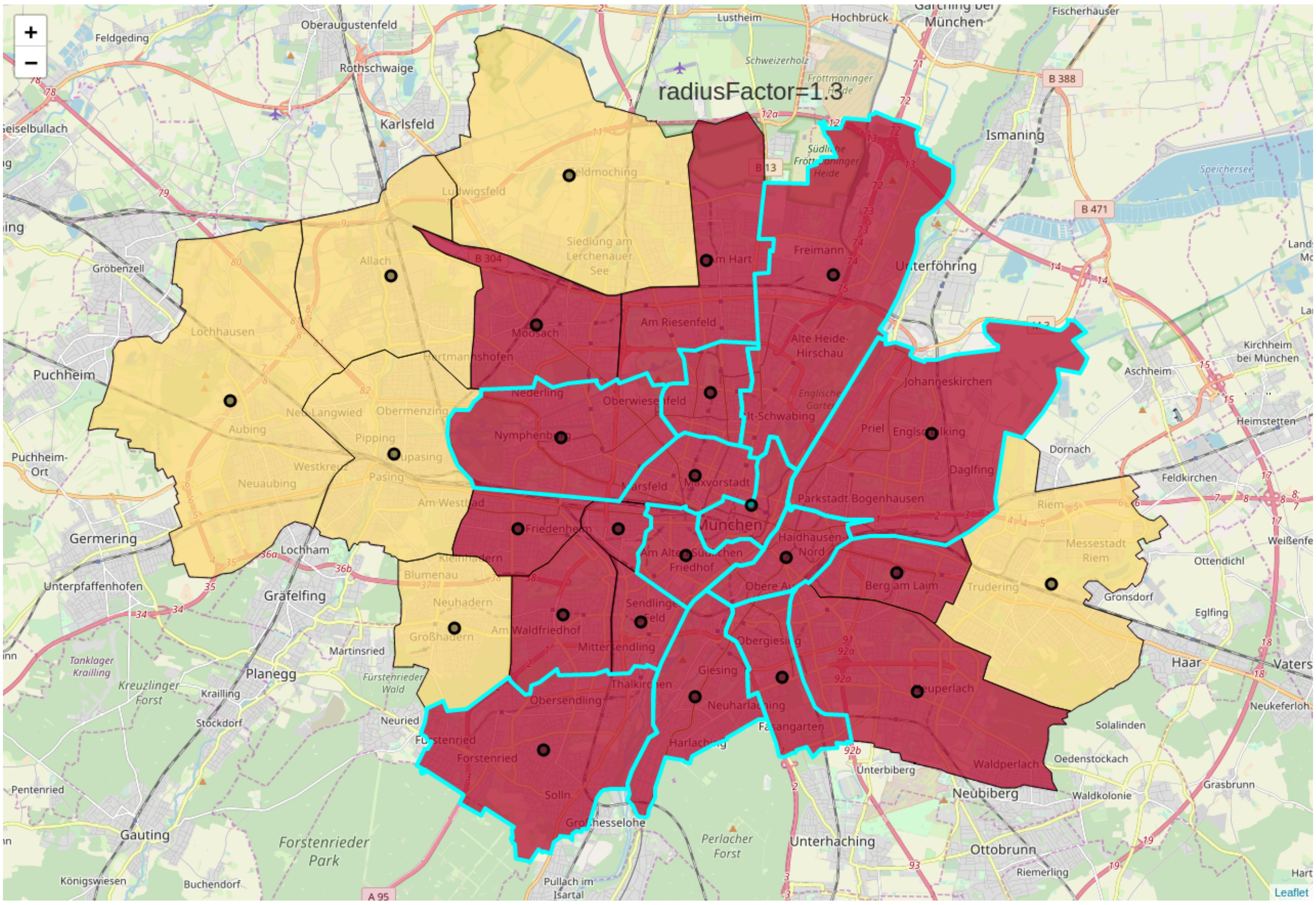

Fig. 15: Red areas show recommended boroughs for radiusFactor=1.3 by using K-means clustering. Best Neighborhoods in Munich are drawn as polygons with a cyan border.

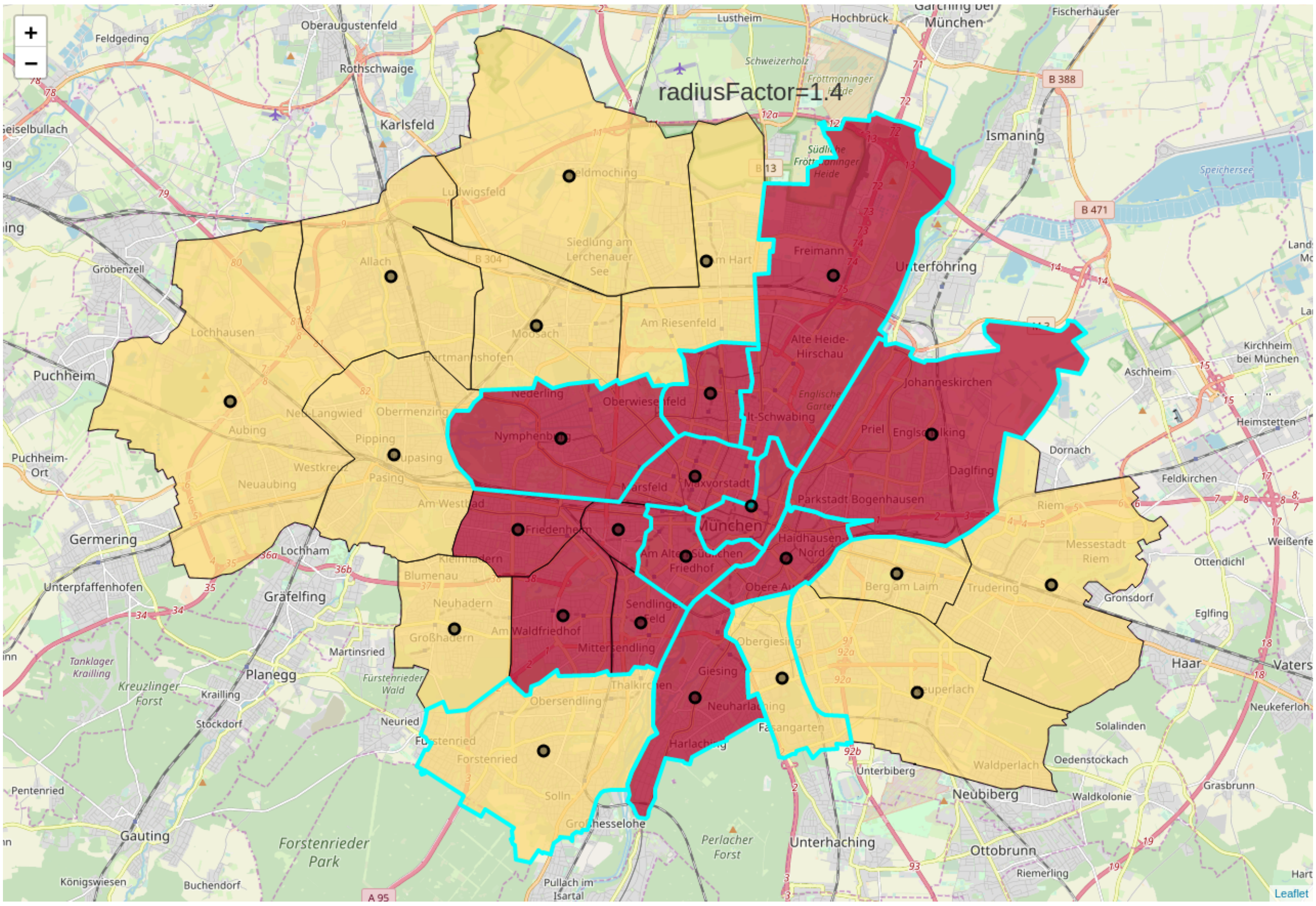

Fig. 16: Red areas show recommended boroughs for radiusFactor=1.4 by using [K-means clustering](). [Best Neighborhoods in Munich]() are drawn as polygons with a cyan border.

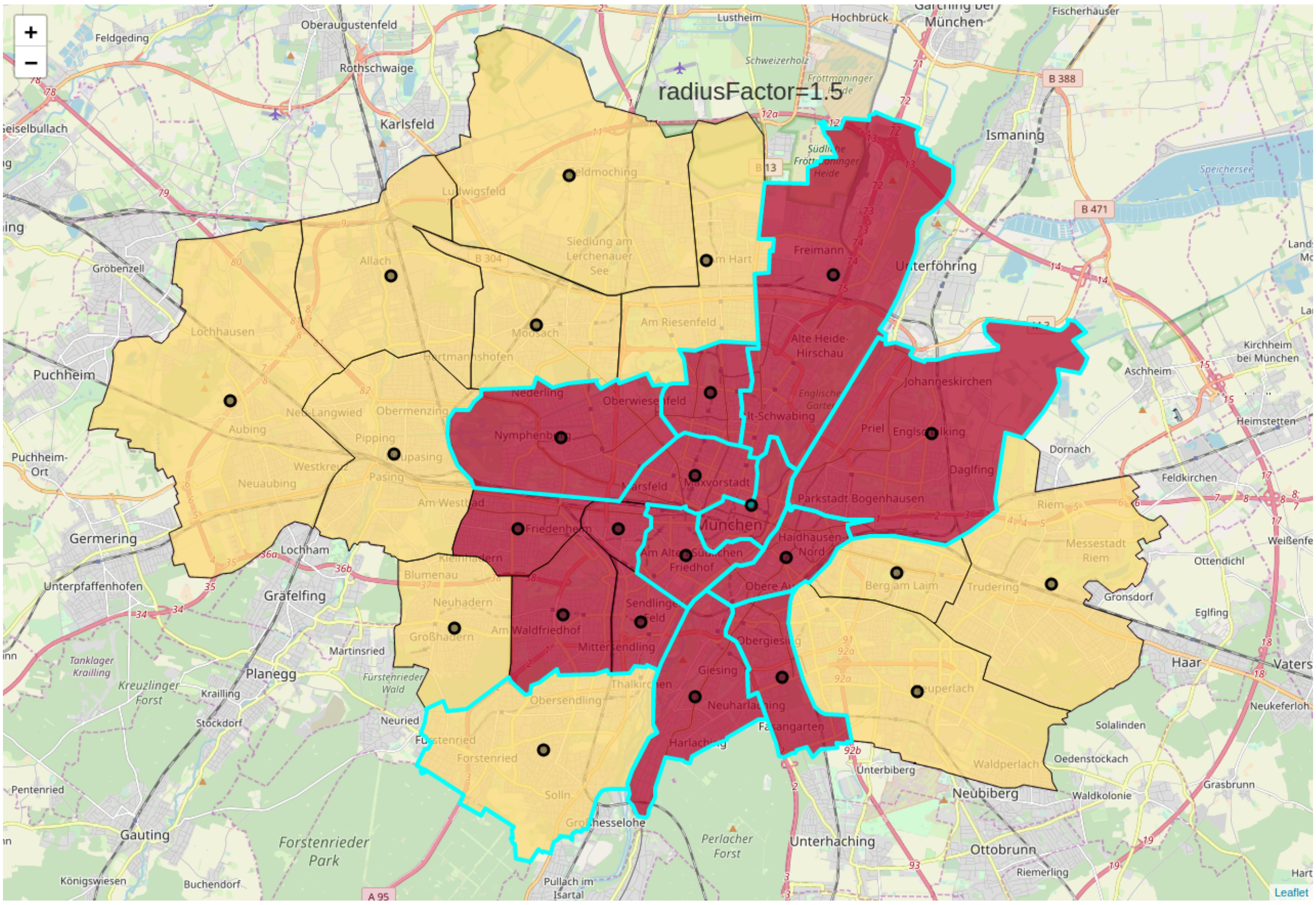

Fig. 17: Red areas show recommended boroughs for radiusFactor=1.5 by using [K-means clustering](). [Best Neighborhoods in Munich]() are drawn as polygons with a cyan border.