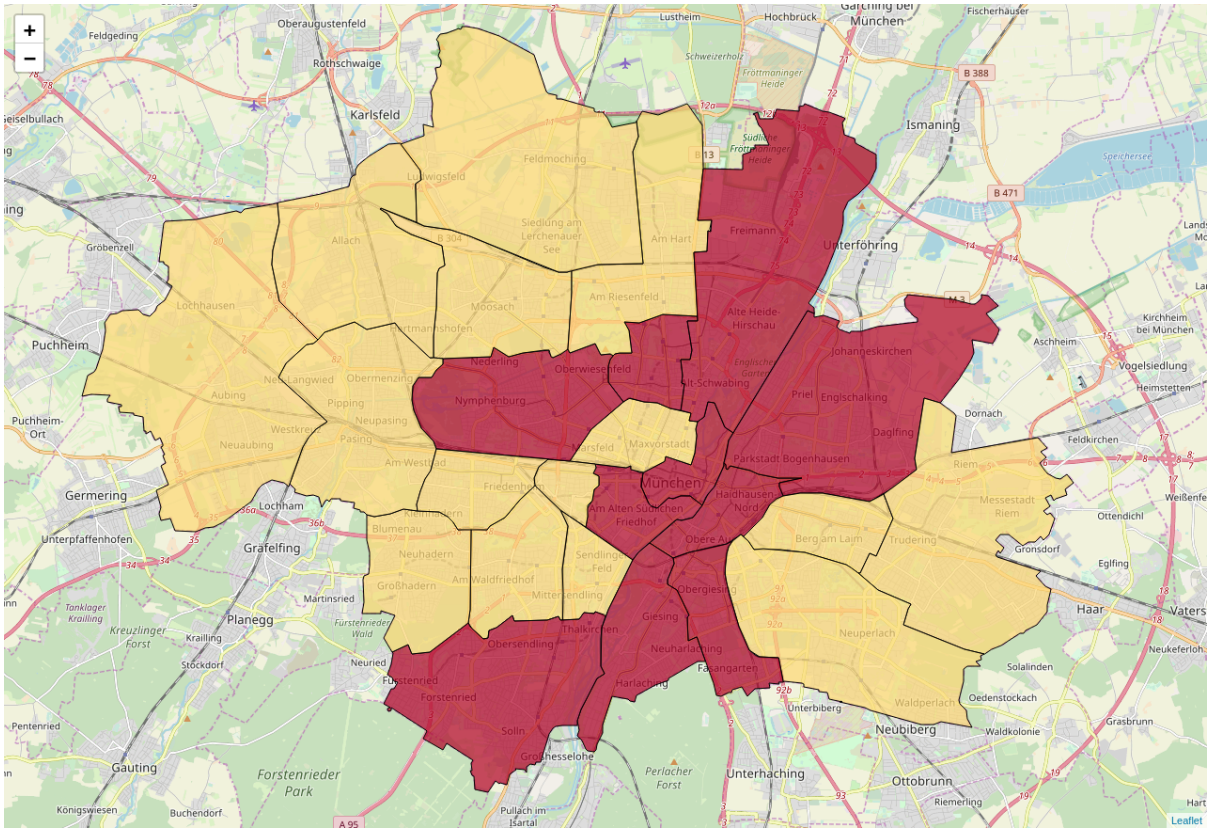
Capstone Project

The Battle of the Neighborhoods

Applied Data Science Capstone by Coursera/IBM

– Regina Castra – February 2020 –



Agenda

- Introduction
- Data / Sources
- Methodology
- Analysis
- Results
- Conclusion

Introduction

- The website www.moving-to-munich.com lists the “Best Neighborhoods in Munich”:

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

- Precise reasons for this selection are not given
- **Is it possible to come to a similar selection based on venues in each borough using machine learning?**

Data / Sources

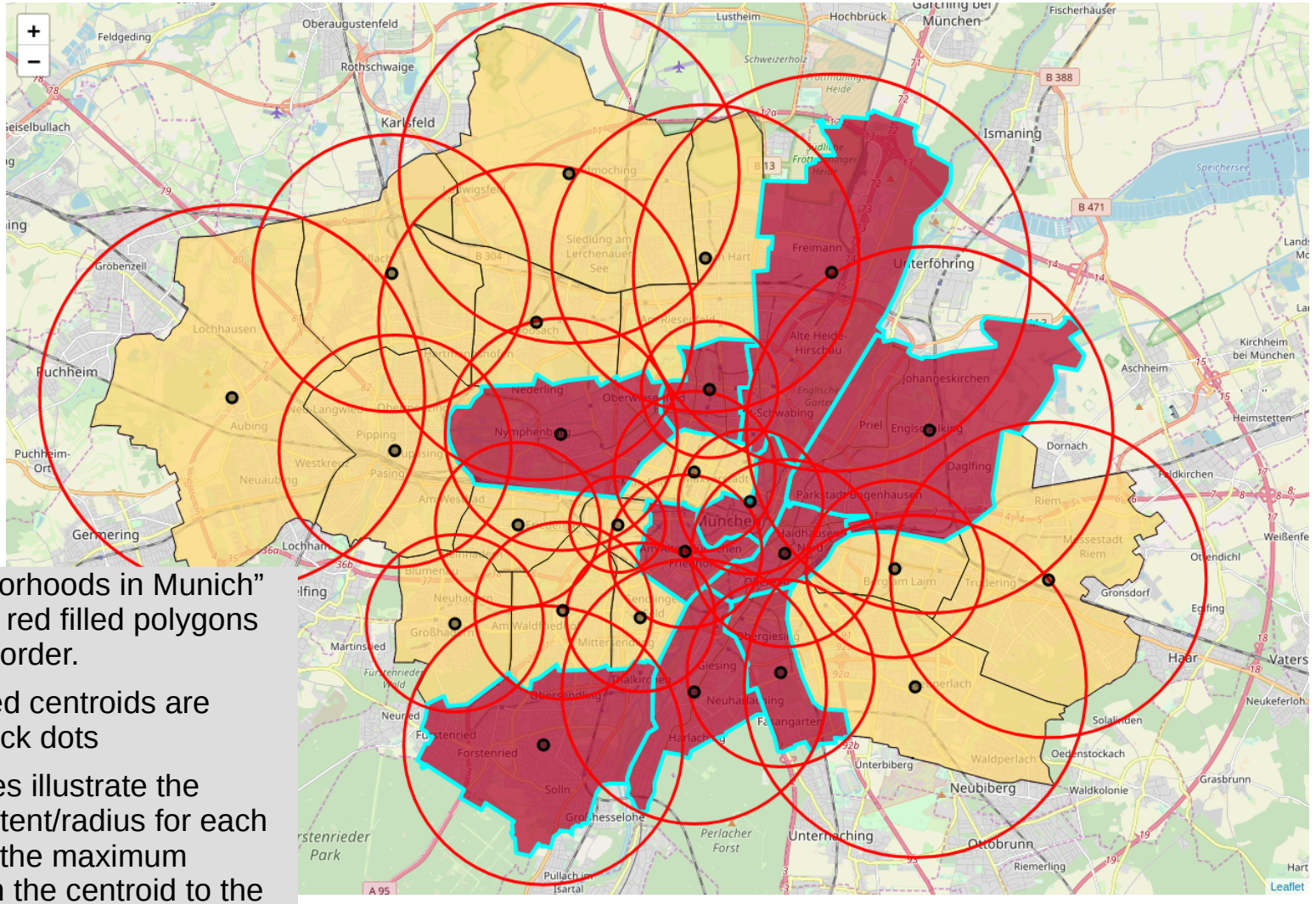
- List of the "Best Neighborhoods in Munich": Moving-to-Munich.com
  - https://www.moving-to-munich.com/best-neighborhoods-in-munich/
- Official names of Munich's boroughs: Wikipedia.org
  - https://de.wikipedia.org/wiki/Stadtbezirke_M%C3%BCnchens
- The border of Munich’s boroughs: Arcis.com
  - https://www.arcgis.com/home/item.html?id=369c18dfc10d457d9d1afb28adcc537b
  - mapshaper.org used to transform data into GeoJSON
- Venue data of Munich's boroughs: Foursquare.com
  - https://foursquare.com

Methodology

1) Retrieve venues for each borough using Foursquare-API.
2) Calculate mean values of the occurrence of individual venue categories for each borough
3) Use a k-means clustering algorithm to cluster the best boroughs
4) Do steps 1) to 3) for several radii around each borough's centroid
5) For each radius of 4) sum up how many times a borough was classified to be one of the best boroughs
6) Compare the result of 5) with the "Best Neighborhoods in Munich"

Analysis

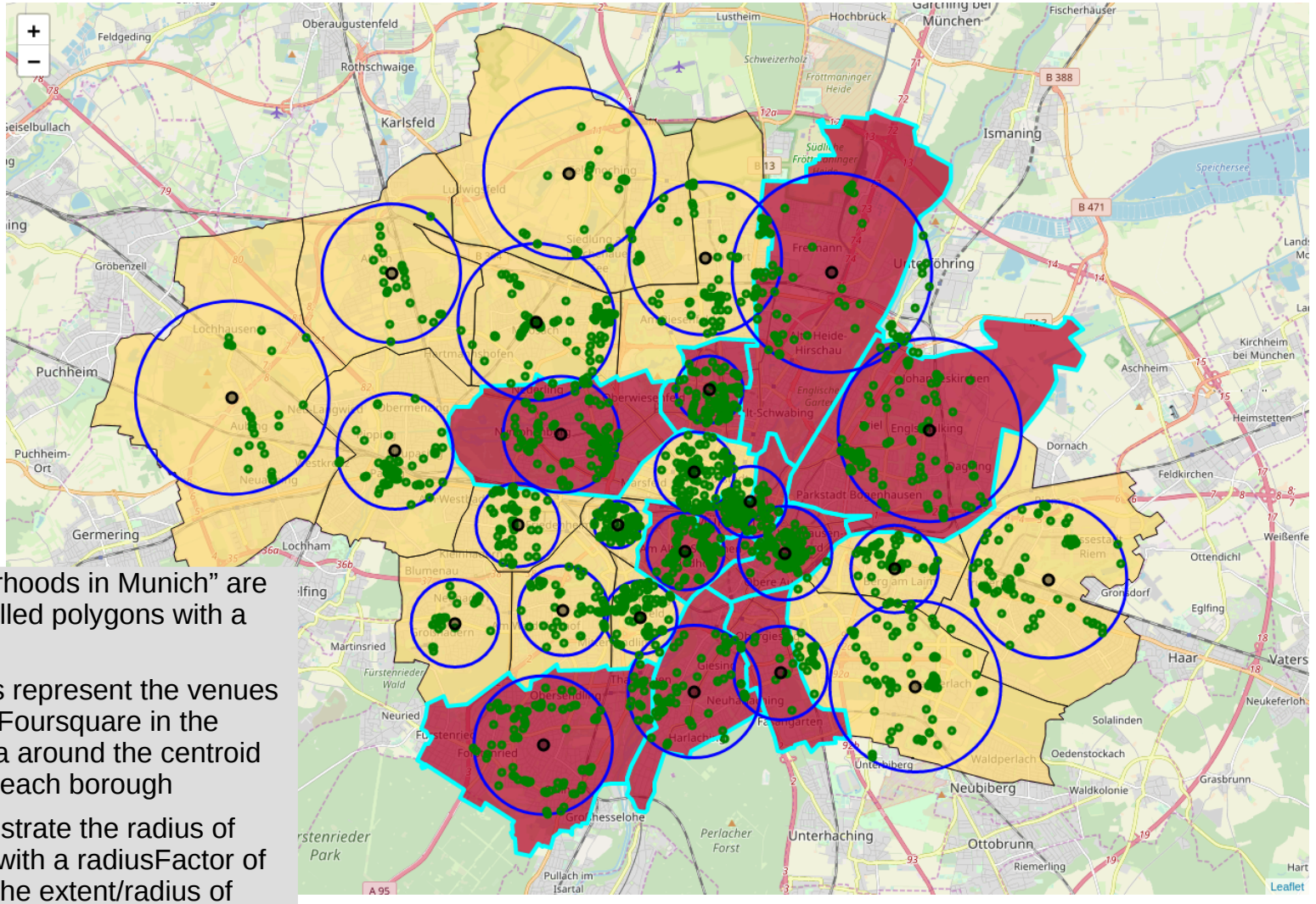
Boroughs and their individual extent (radius)



- “Best Neighborhoods in Munich” are drawn as red filled polygons with a cyan border.
- The calculated centroids are shown as black dots
- The red circles illustrate the calculated extent/radius for each borough, i.e. the maximum distance from the centroid to the border of a borough.

Analysis

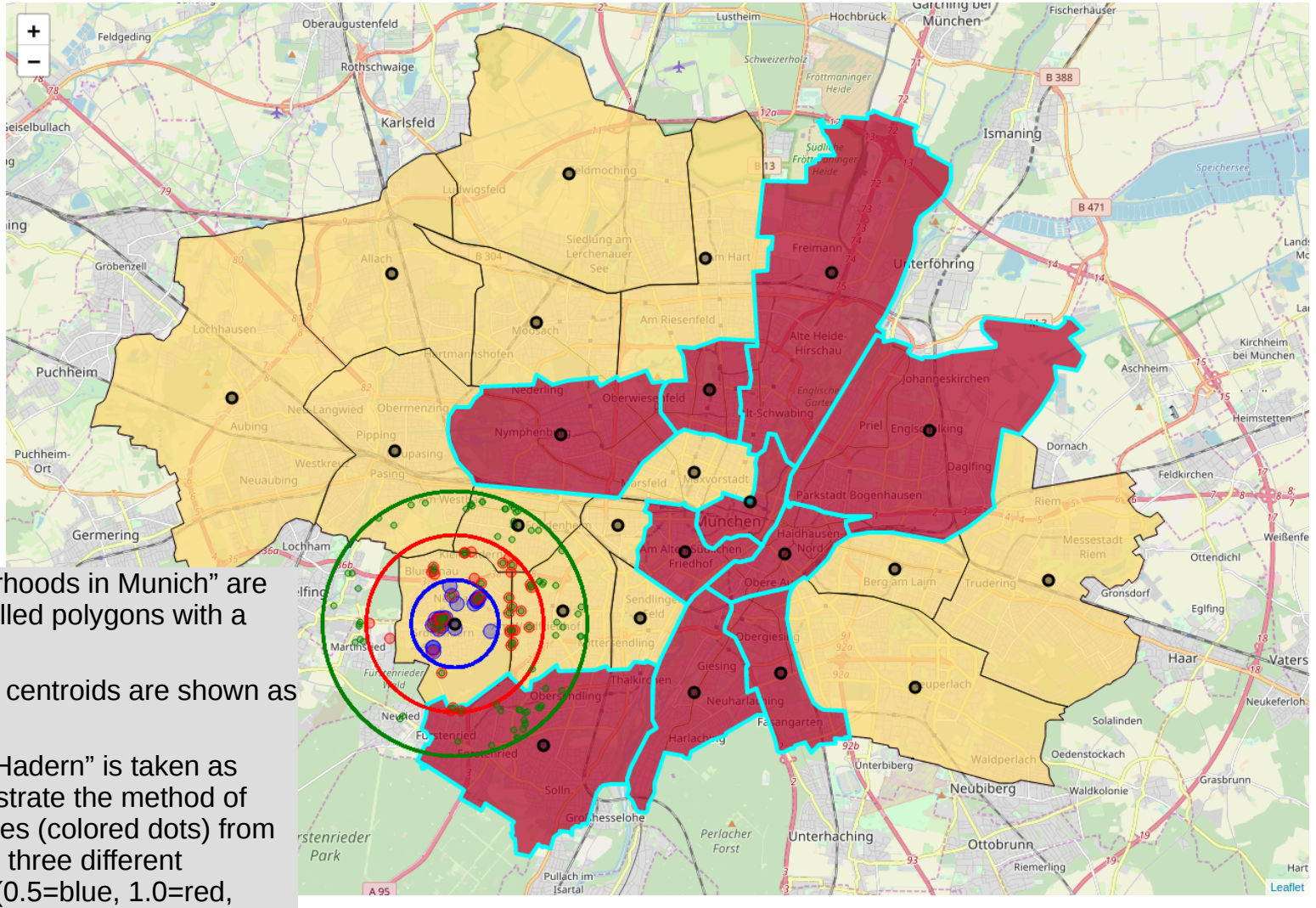
Venues retrieved from Foursquare



- “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 the extent/radius of each borough

Analysis

Venues retrieved from Foursquare



- “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)

Analysis

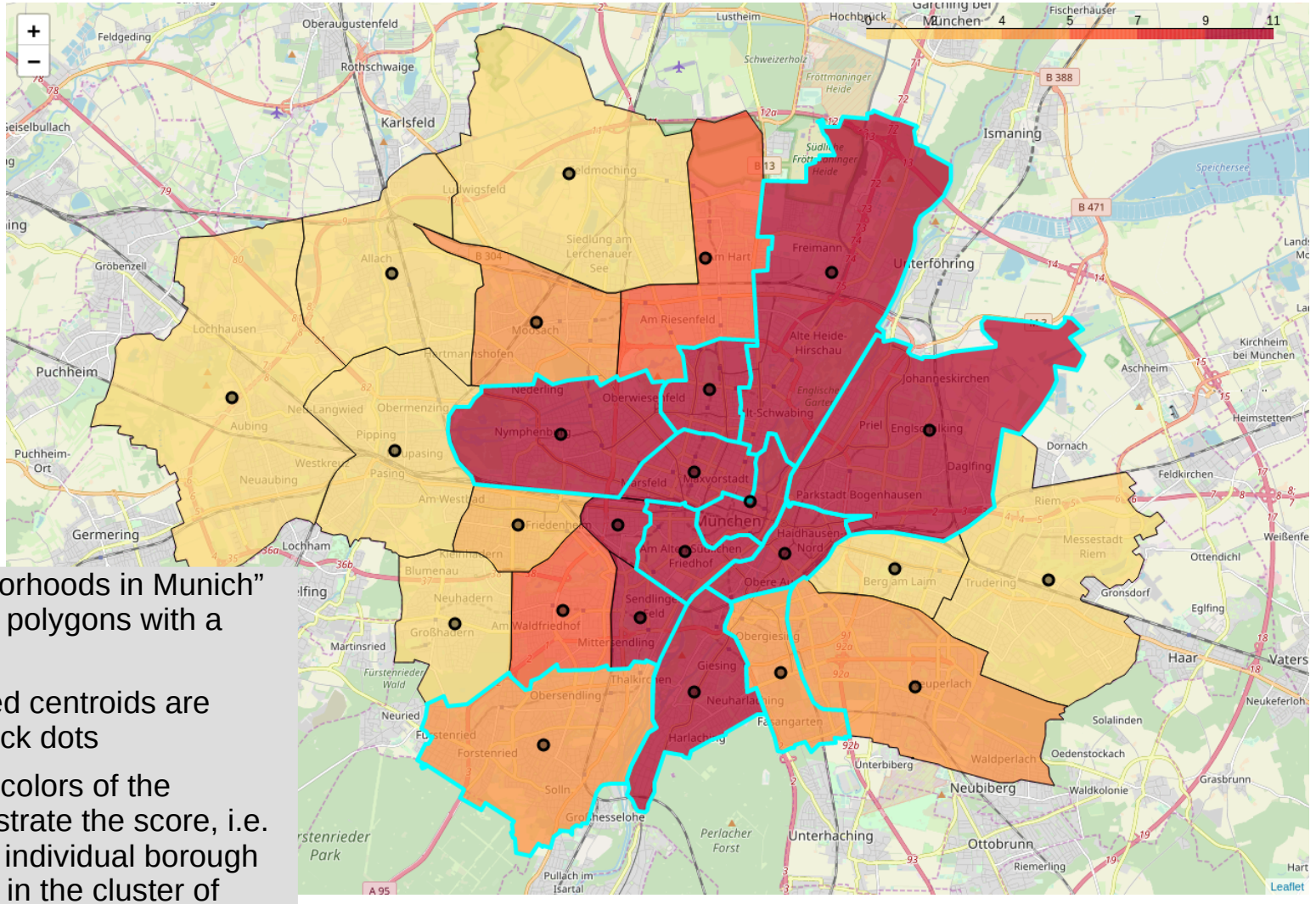
Boroughs recommended by K-means algorithm

	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

- The higher the *score* the stronger the recommendation

Analysis

Boroughs recommended by K-means algorithm



- “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, i.e. how often an individual borough was included in the cluster of “recommended boroughs"

Results

- Most of the Best Neighborhoods in Munich are 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
- 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
- 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

Conclusion

- The selection of the "Best Neighborhoods in Munich" was confirmed by the venue information provided by Foursquare
- The clustering algorithm found additional boroughs that may have similar characteristics. If you are planning to move to Munich these would be worth a look
- The clustering algorithm used is very simple. It is recommended that further in-depth investigations be carried out.