FIND THE RIGHT PLACE FOR A ORGANIC SUPERMARKET IN MUNICH

A INTRODUCTION:

A.1 DESCRIPTION & DISSCUSION OF THE BACKGROUND

With 1.5 million inhabitants, Munich is only the third largest city in Germany, but the one with the highest population density, with 4777 inhabitants per square kilometer. As a resident of this city, I decided to use Munich in my project. The city is divided into 55 neighborhoods in total in an area of 310 km².[1]

As you can see from these data, Munich is a city with a high population density. With my data analysis I would like to support someone who wants to open an organic supermarket in Munich. With the analysis of the venues in Munich I would like to find districts in which cultural facilities or restaurants are less available but sports facilities, playgrounds and food shopping facilities are more represented. However, the number of organic supermarkets should be low in these areas to have as little competitors as possible. Since a supermarket needs a relatively large area to operate, I use the rent index to find areas where rents are not as high. Also, it is expected that in areas with low rents, people are more likely to live and therefore need to buy food.

A.2 DATA DESCRIPTION

- I found the postal codes of Germany from Suche-Postleitzahl.org [2]. The .geojson file has coordinates of the postal codes of Germany. I have to clean the data and reduce it to the city of Munich.
- For the neighborhoods of Munich I fount the rent index at [3]. At Wikipedia the assignment of the postal codes to the neighborhoods i could find. [4]. To show the rent index on a choroplethmap of Munich I have to combine the postal codes with the rent index.
- With geopy/geocoder I will manage to get the central coordinates of the neighborhoods. [5]
- I will use Foursquare API to get the most common venues of given neighborhoods of Munich [6].

- By using kMeans I want to cluster the neighborhoods to find ones where the people live and cultural facilities are less available.
- If numbers of organic groceries are not available at Foursquare i will get them from googlemaps API or a phonebook. [7]
- Finally I want to give an overview on the data by a map and graphics on Clusters, Rent index and counts of organic supermarkets in the neighborhoods.

B. METHODS

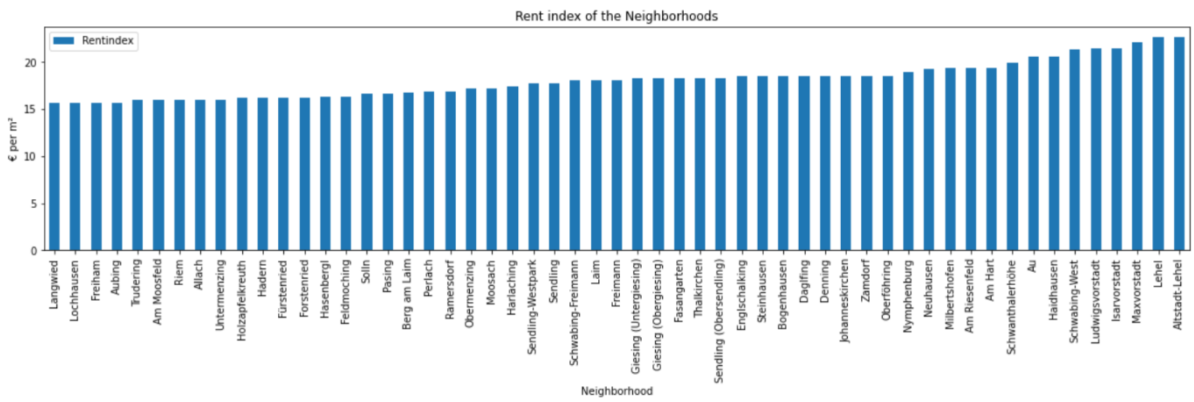
B.1 NEIGHBORHOODS WITH GEOCOORDINATES

For my study I started with an overview on the Munich neighborhoods on Wikipedia. I scraped the Names of the neighborhoods of [1] with Beautifulsoup. With geopy/geocoder I managed to get the geocoordinates of the centers of each neighborhood. At a few of the neighborhoods I had to correct the names to get the coordinates. I found 55 Neighborhoods and saved it to the dataframe for further evaluation.

	Neighborhood	Borough	BoroughNo	Settlement	Latitude	Longitude
0	Allach	Allach-Untermenzing	23	Allach, Gerberau	48.195994	11.457013
1	Altstadt-Lehel	Altstadt-Lehel	01	Angerviertel, Graggenauviertel, Hackenviertel,...	48.137828	11.574582
2	Am Hart	Milbertshofen-Am Hart	11	Am Hart, Harthof (Ostteil), Nordhaide	48.195925	11.571815
3	Am Moosfeld	Trudering-Riem	15	Am Moosfeld	48.133867	11.666309
4	Am Riesenfeld	Milbertshofen-Am Hart	11	Studentenviertel Oberwiesenfeld, Am Oberwiesen...	48.176673	11.552154
5	Au	Au-Haidhausen	05	Obere Au, Untere Au, Falkenau, Neudeck	48.127919	11.592657
6	Aubing	Aubing-Lochhausen-Langwied	22	(Alt-)Aubing, Aubing-Ost, Dorniersiedlung, Kol...	48.158437	11.414066
7	Berg am Laim	Berg am Laim	14	Berg am Laim, Baumkirchen, Josephsburg, Werksv...	48.123483	11.633451
8	Bogenhausen	Bogenhausen	13	(Alt-)Bogenhausen, Arabellapark, Brunnthal, Ga...	48.154782	11.633484
9	Daglfing	Bogenhausen	13	Daglfing, Dianapark, Kolonie Daglfing, Kolonie...	48.149637	11.649305

B.2 RENTINDEX FOR MUNICH

For the neighborhoods I found the rent index as a graphical presentation at [3]. I had to transfer it to a Csv-File and import it into the data frame. I created a bar chart of the rent index.



B.3 VISUALISATION OF THE RENTINDEX WITH A CHOROPLETHMAP

Only JSON Data of the postal code for Germany were available for free. I used a GEOJSON file from [2] and reduced it to the postal codes of Munich. To show the rent index for each postal code, I extracted the postal codes of Munich from [1] and assigned the rent index to each postal code. I created a choroplethmap to visualize the rents in Munich.

B.4 VENUES OF MUNICH

With the geodata of the Neighborhoods I used the Foursquare API to get the Venues in the Radius of 1 km of the center of the Neighborhoods. With the app I got 2543 Venues.

To reduce the categories I summarized them as follows:

- Restaurants: all Venues with Restaurant in the Name and Osteria
- physical Health: 'Golf Course','Soccer Field','Gym Pool','Bowling Alley','Gym','Skate Park','Yoga Studio','Climbing Gym','Gym / Fitness Center','Water Park', 'Dance Studio','Escape Room','Sports Club','Water Park','Indoor Play Area','Golf Course','Pool','Spa','Pool','Bathing Area','Playground','Skating Rink', 'Athletics & Sports'
- Theater: 'Theater', 'Cinema', 'Opera House', 'Concert Hall', 'General Entertainment', 'Indie Theater', 'Indie Movie Theater', 'Movie Theater'
- Bar: 'Cocktail Bar', 'Wine Bar','Lounge', 'Juice Bar','Nightclub','Pub','Brewery', 'Gastropub','Deli / Bodega'
- Market: for all venues containing the word market
- Shop: for all shops
- Coffee: for coffee and café
- Museum: for all musem

With this procedure I managed to reduce the unique venues from 259 to 152.

By grouping the markets, only Farmers Markets, Fish Markets, Markets and Supermarkets were grouped. No organic markets are available. There are named organic groceries.

	Neighborhood
Venue Category	
Farmers Market	7
Fish Market	3
Market	3
Supermarket	122

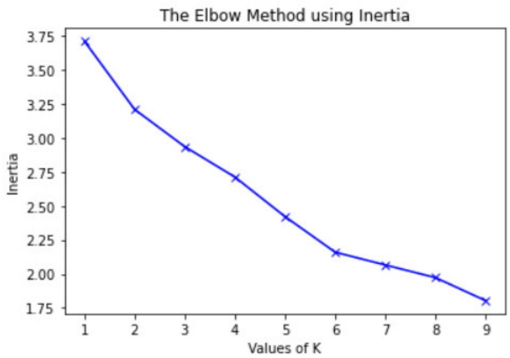
With onehot encoding, grouping and sorting the 10 most common venues for each neighborhood could be found.

As the Restaurants were mostly the first common venue, they do not contribute to the distinctiveness of the neighborhoods; therefore I dropped the category Restaurant.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Allach	48.195994	11.457013	Bäckerei Schuhmair	48.197175	11.459016	Bakery
3	Allach	48.195994	11.457013	Sport Bittl	48.191447	11.466553	Sporting Goods Shop
4	Allach	48.195994	11.457013	Westside Hotel	48.201045	11.458564	Hotel
7	Allach	48.195994	11.457013	dm-drogerie markt	48.194118	11.465640	Drugstore
8	Allach	48.195994	11.457013	Rossmann	48.193301	11.466388	Drugstore

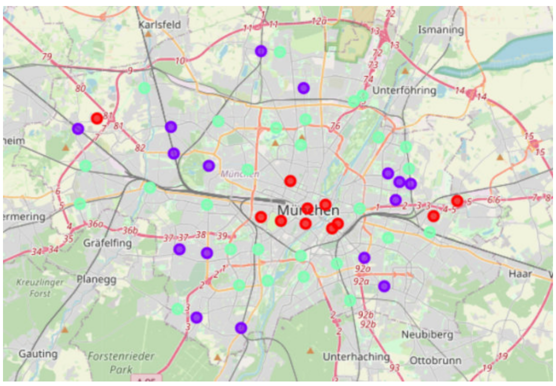
B.5 CLUSTERING THE NEIGHBORHOODS

Using the elbow method to find the optimal numbers of clusters, no real elbow could be detected.



The calculation with 3 Clusters resulted in groups with more than 10 neighborhoods each.

With kMeans 3 Clusters of the Neighborhoods were calculated, according to the most common venues. The cluster numbers I saved to the neighborhood data.



B.6 NUMBER OF ORGANIC GROCERIES

As by the Foursquare API only a few organic groceries were found, I searched for the organic groceries at the phone book [7]. With the restriction to München, Biomarkt and Naturkost I scraped with Beautifulsoup the content of the web-Page. After dublicate cleaning I found 74 organic groceries and saved the numbers per neighborhood to the Data.

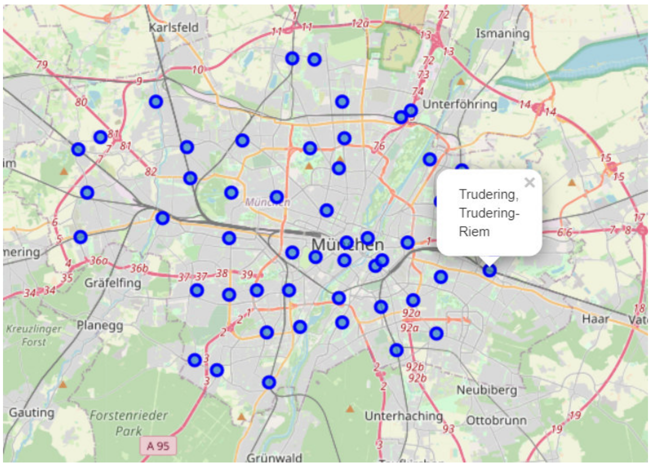
B.7 HEATMAP OF CLUSTERS, NUMBER OF ORGANIC GROCERIES AND RENTINDEX.

By creating a heatmap the addiction of clusters, number of organic groceries and rent index can be evaluated and used for the decision to find an optimal place for a new organic grocery.

C. RESULTS

C.1 55 NEIGHBORHOODS OF MUNICH:

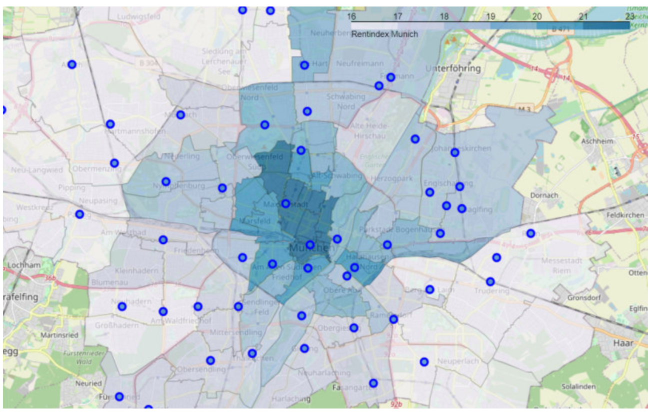
I could identify 55 Neighborhoods in Munich. Here I show the middle of each neighborhood.



C.2 RENT INDEX

The rent index of Munich is very different between 16 € and 22 € per squaremeter. The index was available for the neighborhoods. For visualization the index was transformed to the postal code as only GEOJSON Data of the postal codes are available for Munich.

C.3 VISUALIZATION OF RENTS



Giving an overview the rents were visualized with the centers of the neighborhoods.

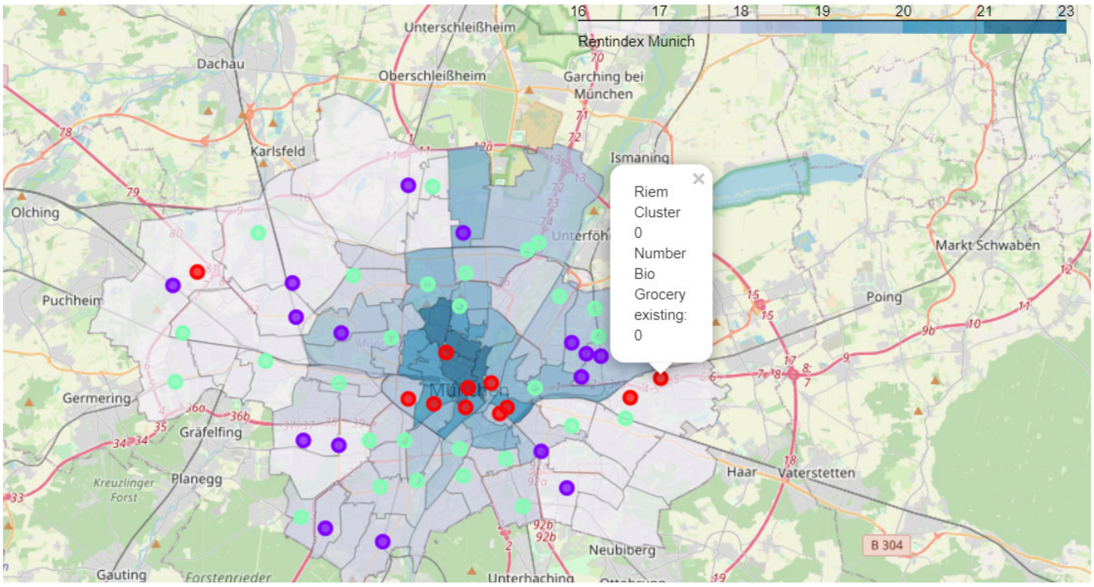
C.4, C.5 VENUES FOR CLUSTERING

The reduction of the venue categories led to a meaningful clustering. Optimal clustering with tree clusters resulted in the following clusters with the most common venues.

Clusters	1st common venue	2nd common venue	3rd common venue
No. 0 red	Coffee, Hotel	Bar, Shop, Physical Health	Plaza, Bar, Coffee
No 1 purple	Bus stop, Market, Physical Health, Coffee, Hotel	Bus stop, Market, Bakery	Market, Shop, Hotel
No 2 green	Physical Health, Bakery	Market, Physical Health, Coffee, Shop	Drugstore, Physical Health

C.6 ORGANIC GROCERIES

The clusters were visualized together with the rent index on a choroplethmap. For each neighborhood the number of organic groceries is visible by clicking on the symbol.



Clusters description:

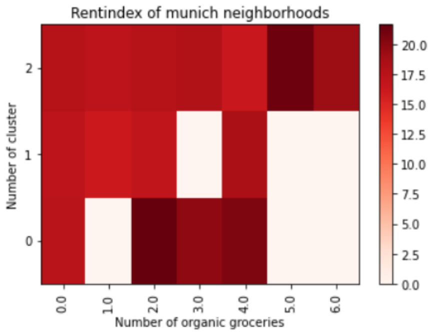
Cluster 0 in red is more present in the center of the city. Venues for nightlife, relaxation and holiday are more present and the rents are higher. Markets or Bakeries are not as present.

Cluster 1 in purple is more present in the outer parts of the city. Bus stop and markets seem to be very important and the rents are lower. It must be expected, that people with less money live in these neighborhoods.

Cluster 2 in green forms a ring around the center of the city. Very important venues are the one for physical health. Also present are markets and drugstores. The neighborhoods were people with middle income are living.

C.7 HEATMAP

For better understanding I prepared a heatmap to show the coherence of rent index and occurrence of organic groceries in the clusters.
Here you can see: in areas with lower rent are less organic groceries and vice versa in Cluster 0 and 2 in areas with high rents more organic groceries are found. Especially in Cluster 2 the Neighborhoods with the most organic groceries are found and have the highest rent.



Recommendation for the place of a new organic grocery:

A good location for the opening of a new organic grocery appears to be a neighborhood from cluster 2, since the rents there are in the middle range, the available venues are geared to daily needs and there are still a number of clusters without an organic grocery. The rent level in the middle range indicates that the residents have sufficient means to buy the more expensive goods from the organic groceries and on the other hand the rent for the shop itselfe will not be too big. Neighborhoods with middle rent and no organic groceries are for example *Neuhausen* and *Fasanengarten*.

E. DISCUSSION

As Munich is a very heterogeneous city with 55 neighborhoods and a complex city structure different approaches can be tried in clustering and classification studies. Moreover it is obvious that not every classification method can yield the same high quality.

I used the KMeans algorithm with data from the Foursquare API as a part of this clustering study. As the Foursquare-App is not common in Germany, I suspect that the entries are mostly not done by city residents of Munich. As vacationers focus on other things than residents, the used venues might be biased. To overcome this problem Googlemaps API might be used. Nevertheless the calculated 3 Clusters appear reasonable. As with the Elbow method no optimum for k was found, different k Values were tested and adjusted to 3, as no small clusters appeared.

For decision making it seemed to be important to know the number of existing organic groceries in the city and to take the rent index into account. These data was used to present by choropleth- and heat-map. Therefore I used coordinates of the postal codes of Munich.

I ended the study by visualizing the data and clustering information on a Munich map and give a recommendation of neighborhoods suitable for opening an organic grocery.

F. CONCLUSION

An analysis like this might be helpful for investors searching a good place for investment. Further aspects (for example, the population density or the ratio of offices to apartments) from different sources might be included in the analysis to sharpen the recommendation.

This analysis might also be a good approach for city planner and people who are responsible for making a city attractive to live in.

E. Wesner

SOURCES:

[1] https://de.wikipedia.org/wiki/Liste_der_Stadtteile_M%C3%BCnchens
[2] https://www.suche-postleitzahl.org/plz-karte-erstellen
[3] https://de.statista.com/statistik/daten/studie/260438/umfrage/mietpreise-in-muenchen-nach-bezirken/#professional
[4] https://www.suche-postleitzahl.org/muenchen-plz-80331-85540.52bb
[5] https://geopy.readthedocs.io/en/stable/
[6] https://foursquare.com/
[7] https://www.dasoertliche.de/