Where You Should Open Your Next Restaurant in Zurich

Maria Kliesch

5/4/2020

Contents

1 Introduction/Business Problem

1.1 Background

According to the statistical office of the City of Zurich, there are currently around 2200 restaurants listed in the City of Zurich. With around 400'000 inhabitants, this means that there is one restaurant for every 200 citizens, or put differently, this is comparable to having five restaurants in a small village of 1000 inhabitants. Unsurprisingly, it is estimated that around 62% of restaurants in Zurich, and particularly small ones, are in the red (NZZ).

1.2 Problem

In order to be able to cover personnel and rental costs, stakeholders planning to open a restaurant in Zurich will have to make an educated decision on suitable locations that 1) attract a sufficient number of guests and 2) have affordable rental charges.

1.3 Aim

The aim of this project is to characterise the different neighborhoods in Zurich based on whether or not they may be a promising location for a new restaurant to be opened. The following aspects will be taken into account in the analysis:

- The number of already existing restaurants in the vicinity, which should be as low as possible in order to minimize the number of competitors.
- The number of "friendly" businesses in the vicinity, such as shopping facilities or bars, which increase the number of potential customers in the area.
- The average rental prices in the area, which should be as low as possible.

I will use machine learning to distinguish areas that meet the above-mentioned criteria from those that do not, and will provide a list of the most promising three areas as well as their scoring in terms of the selection criteria for stakeholders to make a final decision.

2 Data

2.1 Data Sources

The following data sources will be used for this analysis:

- There are 12 districts in Zurich (“Kreis”). Neighborhoods will be defined as sub-districts, of which there are 34 in Zurich (2-4 per district). The list of neighborhoods is provided by the statistical office of the City of Zurich (see here).
- The location data on restaurants in Zurich will be extracted via the Foursquare.com API, which, among others, provides information on the type of venues found in places all over the world, including restaurants, shopping facilities and entertainment venues.
- Information on the average rental costs for commercial spaces per district are provided by the statistical office of the City of Zurich (see here). The numbers refer to Swiss Francs (CHF) per m2 .
- Geospatial information (i.e. coordinates) of each neighborhood will be retrieved via reverse geocoding using the geopy package.

2.2 Data Preprocessing

First, a data sheet was created manually from the Statistical Office’s information on commercial space rent per district, using information from the most recent year for which a number was available. The data frame was then read into Python. It included the names of all districts and neighborhoods as well as the rent per m^2.

In a second step, the geopy-package was used to add geolocation information to the dataframe and provide longitude and latitude values of each neighborhood. To verify the correctness of the result, all districts were plotted on a map of Zurich:

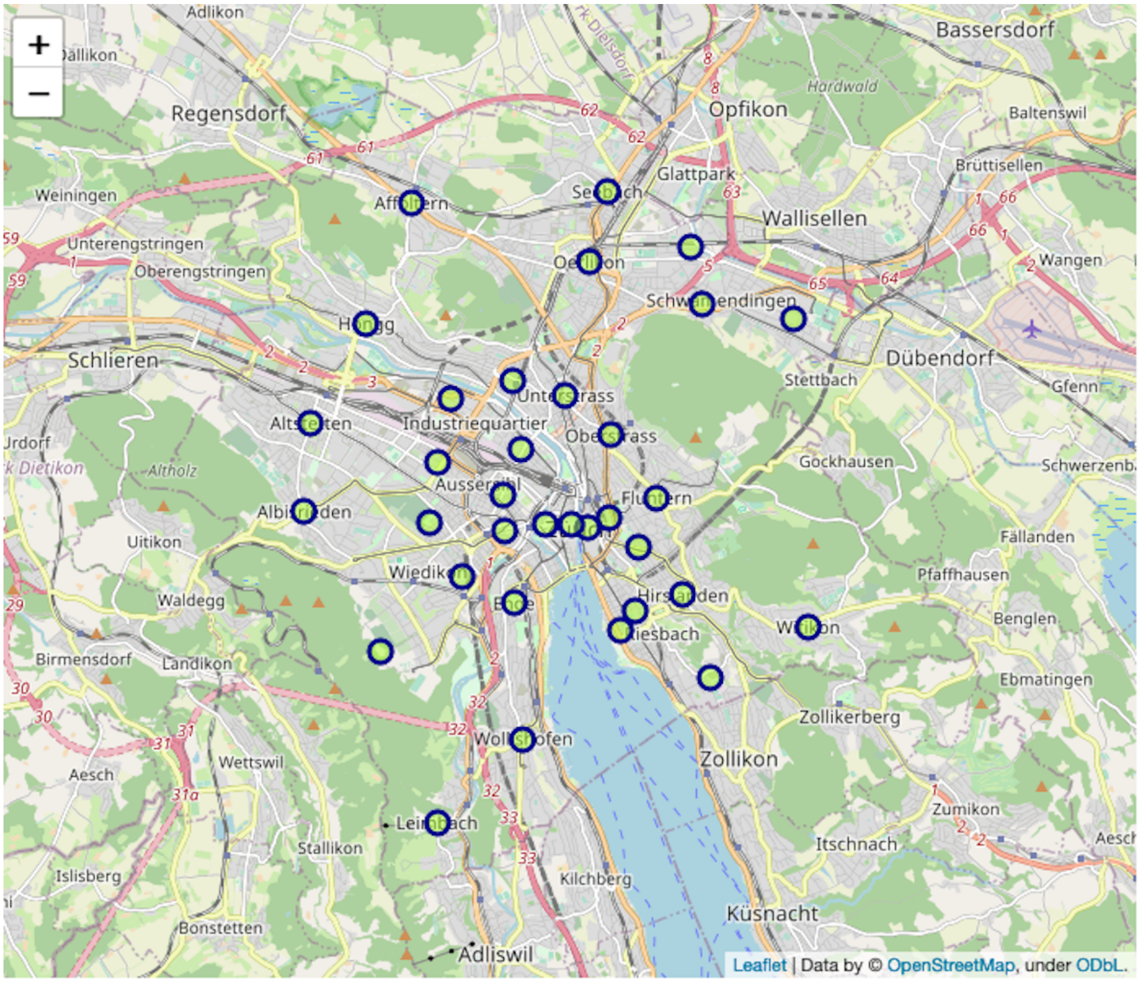

Fig.1: Overview of neighborhood centroids in Zurich.

To get information on all the venues that are registered on Foursquare, I performed an API request to search for venues within a radius of 1000m from the central coordinates of a given neighborhood. The thereby extracted information was concatenated with the previous dataframe.

This dataframe had to be cleaned up, because 1) not all venues that are registered on Foursquare have an assigned category and 2) some venues appeared twice, because they were within the 1000m radius of more than one neighborhood. Since category information was essential for the ensuing analysis, all venue entries without categories were removed from the dataframe. As a final step, all duplicate venues were removed by finding out which neighborhood a duplicated venue is closest to and removing it from all other neighborhoods.

	District	Neighborhood	Neighborhood Rent	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Kreis 1	Rathaus	540.0	47.372649	8.544311	Oliver Twist Pub Zürich	47.372280	8.544270	Sports Bar
1	Kreis 1	Rathaus	540.0	47.372649	8.544311	Café Henrici	47.372516	8.543686	Café
2	Kreis 1	Rathaus	540.0	47.372649	8.544311	Restaurant 1001	47.372974	8.543783	Falafel Restaurant
3	Kreis 1	Rathaus	540.0	47.372649	8.544311	Zürich	47.373158	8.544117	Motel
4	Kreis 1	Rathaus	540.0	47.372649	8.544311	Raclette Factory	47.372376	8.543813	Swiss Restaurant

Fig.2: First five rows of the final dataframe.

3 Methodology

3.1 Exploratory Analyses

In a first step, I examined the individual parameters that are going to decide whether or not a neighborhood is suitable as a new restaurant spot or not. To do so, I investigated how many restaurants there are in Zurich and how they are spread across the city.

In a second step, I compared neighborhoods on the basis of their average rental costs and checked whether there is a relationship between the rental price and the popularity of neighborhoods for restaurants.

Third, I examined the number of friendly businesses in each neighborhood and compared the number of friendly businesses in each neighborhood with the number of friendly businesses.

3.2 Main Analyses

The main analysis consisted of two parts:

First, I ranked each neighborhood by their number of competitors, friendly businesses and average rent and calculated an overall (mean) score that describes how favourable the conditions in a given neighborhood are. This overall score can then be used directly as a measure of whether (1) or not (0) a given neighborhood is overall suitable.

In a second step, I clustered the neighborhoods of Zurich in order to find similarities between them based on the three criteria. This approach is more nuanced than considering the overall mean, as neighborhoods can be very favourable regarding two criteria but unfavourable regarding the third. The cluster analysis can tell us which neighborhoods are similar and how they load on each of the assessment criteria.

Wherever possible, visualization of a given analysis is provided.

4 Results

4.1 Exploratory Analyses

The exploratory analysis of already existing restaurants in Zurich showed that only 472 of them were registered on Foursquare (or marked as such by category). As per the statistical office of the City of Zurich, there should be more than 2000, but not all of them are registered on Foursquare. This is problematic, because it means that our analysis likely will only include the most popular/ most visited ones, so in a real-case scenario, we would change to a different location information provider that includes a more comprehensive list of venues in Zurich (e.g. Yelp or Trip-Advisor).

From the restaurants found on Foursquare, I created a map in order to visualize how restaurants are spread across the city.

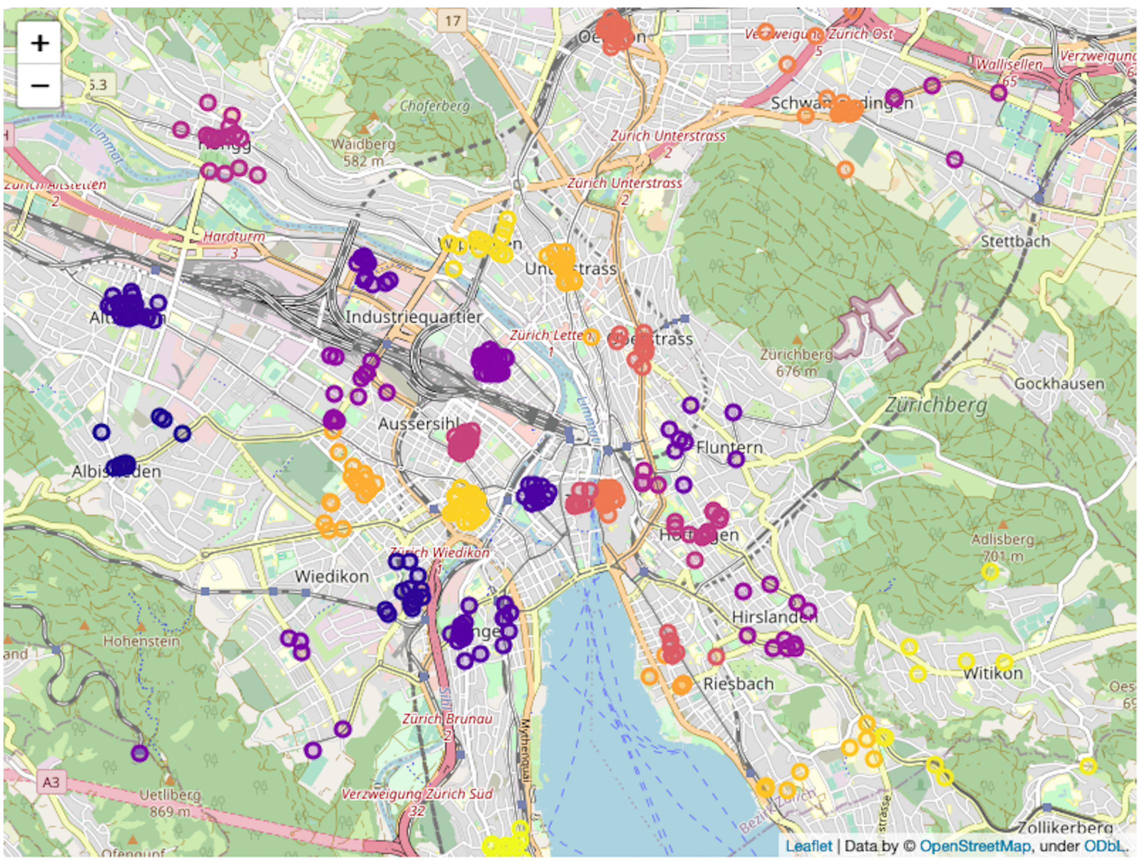

Fig.3: Restaurants in Zurich, color-coded by neighborhood.

As Figure 3 shows, especially in the city center, restaurants are clustered densely around specific areas, so opening another restaurant within these areas would mean a high number of competitors. From visual inspection, neighborhoods with particularly dense distributions are City, Affoltern, Gewerbeschule, Werd, Rathaus, Langstrasse and Unterstrass. In contrast, there are also less dense neighborhoods, such as Fluntern, Hirslanden or Hard, where, even though there are quite a few restaurants in the neighborhood, there might still be space for another one due to the large distance between existing ones.

4.1.1 Number of Restaurants per Neighborhood

To examine which neighborhoods have a comparably low number of competitors, the following table shows neighborhoods sorted by number of restaurants.

Neighborhood	Number of Restaurants
Hochschulen	4
Hirzenbach	5
Seefeld	5
Affoltern	6
Friesenberg	6
Muehlebach	8
Fluntern	8
Weinegg	9
Lindenhof	10
Witikon	10
Saatlen	10
Hirslanden	11
Escher Wyss	11
Albisrieden	11
Oberstrass	11
Schwamendingen-Mitte	12
Leimbach	12
City	13
Hottingen	13
Alt-Wiedikon	13
Hard	13
Wipkingen	14
Wollishofen	15
Hoengg	15
Oerlikon	17
Unterstrass	17
Seebach	18
Sihlfeld	20
Enge	21
Werd	23
Langstrasse	25
Rathaus	26
Altstetten	27
Gewerbeschule	33

Table 1: Number of Restaurants per Neighborhood

If stakeholders were simply looking for neighborhoods with few competitors, this overview would already suffice for a business decision. However, it would also make sense to take into account the number of friendly businesses as well as the rent per m^2 to gain a more comprehensive characterization of each neighborhood.

4.1.2 Rental Costs per Neighborhood

Information on rental costs for empty commercial space in Zurich is only available per district and not per neighborhood. Therefore, the following map is color-coded by rental costs (per m^2) per district.

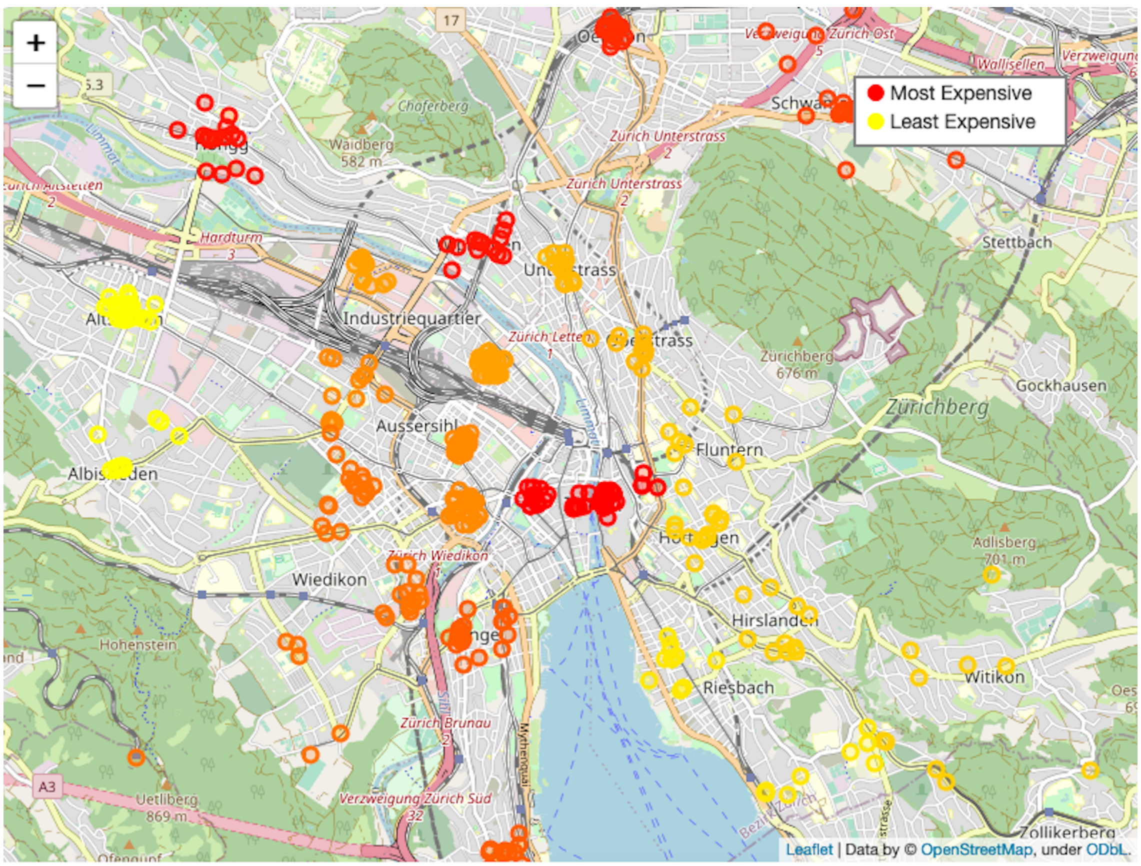

Fig.4: Restaurants in Zurich, color-coded by district rent.

Visual inspection shows that the three most central neighborhoods (belonging to district 1) are also the most expensive ones in terms of rental costs. It is possible that neighborhoods that are attractive for customers and therefore frequently visited could afford to charge more for the rent, so that neighborhoods with an already high number of restaurants or other venues would be more expensive. I tested this with a regression and did not find any correlation between the number of venues/restaurants and the rental costs.

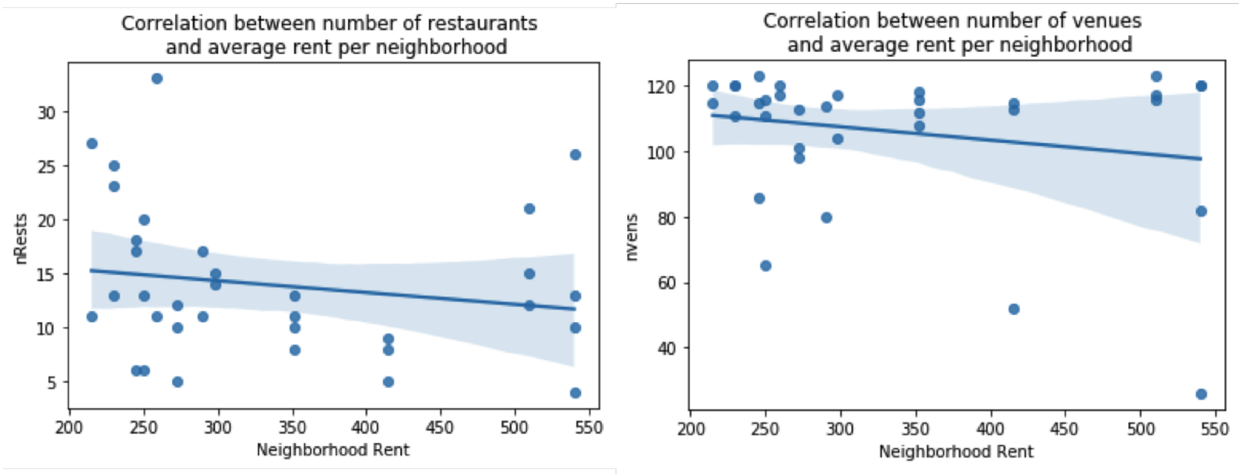

Fig.4: Correlation between number of restaurants per neighborhood and rent (left) and between number of overall venues and rent (right).

4.1.3 Friendly Businesses per Neighborhood

As a final step, I examined the ratio of competitors to friendly businesses in a neighborhood. As mentioned above, friendly businesses are those that attract potential customer but do not compete for the same market, meaning that restaurant owners and owners of friendly businesses can mutually benefit from one another. The competitor-to-friend ratio was calculated by counting all venues in a given neighborhood and determining how many of them are restaurants or friendly businesses.

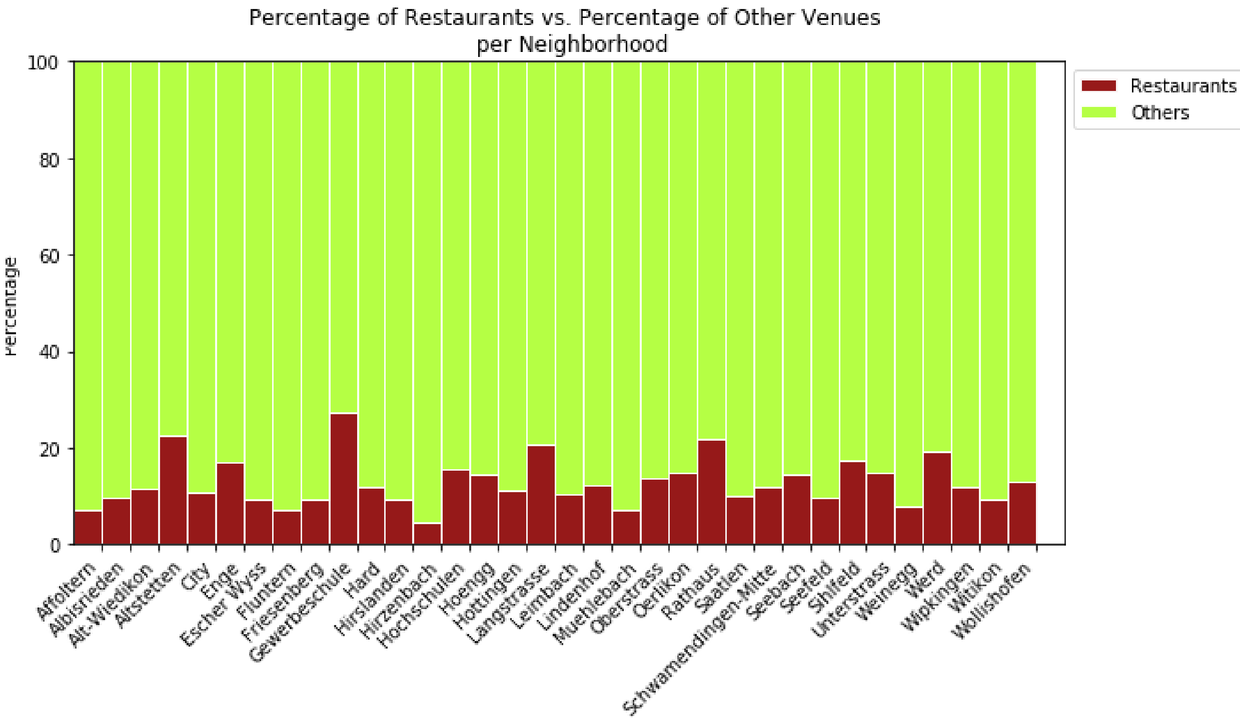

Fig.5: Competitor-to-friend ratio per neighborhood.

Figure 5 shows that neighborhood Gewerbeschule has the highest number of restaurants compared to friendly businesses, while Hirzenbach has the lowest ratio of competitors to friendly businesses. Again, stakeholders could make a decision based on this ratio alone, but the combination of all three criteria (number of competitors, number of friendly businesses and rent per m^2) is likely to yield a more objective measure of suitability.

4.2 Overall Score for Neighborhoods

Based on the three criteria 1) number of competitors, 2) number of friendly businesses and 3) rent per m^2, we can rank the neighborhoods from 0 (bad) to 1 (good). This results in three scores per neighborhood, which can then be combined into one overall score for suitability by calculating the mean over them.

Sorting neighborhoods by overall suitability, the following ranking emerges:

District	Neighborhood	Rent_score	Rest_score	Friend_score	Overall_score
Kreis 9	Albisrieden	1.00	0.68	0.82	0.83
Kreis 12	Hirzenbach	0.55	0.95	1.00	0.83
Kreis 5	Escher Wyss	0.64	0.68	0.91	0.74
Kreis 4	Hard	0.91	0.58	0.64	0.71
Kreis 11	Seebach	0.82	0.37	0.86	0.68
Kreis 11	Affoltern	0.82	0.89	0.23	0.65
Kreis 3	Alt-Wiedikon	0.73	0.58	0.64	0.65
Kreis 7	Fluntern	0.27	0.84	0.82	0.64
Kreis 8	Weinegg	0.18	0.79	0.91	0.63
Kreis 11	Oerlikon	0.82	0.42	0.64	0.63
Kreis 8	Muehlebach	0.18	0.84	0.86	0.63
Kreis 7	Hirslanden	0.27	0.68	0.86	0.60
Kreis 3	Friesenberg	0.73	0.89	0.09	0.57
Kreis 7	Hottingen	0.27	0.58	0.86	0.57
Kreis 4	Werd	0.91	0.21	0.59	0.57
Kreis 10	Wipkingen	0.36	0.53	0.77	0.55
Kreis 7	Witikon	0.27	0.74	0.64	0.55
Kreis 12	Saatlen	0.55	0.74	0.32	0.54
Kreis 3	Sihlfeld	0.73	0.32	0.55	0.53
Kreis 2	Leimbach	0.09	0.63	0.86	0.53
Kreis 4	Langstrasse	0.91	0.16	0.50	0.52
Kreis 12	Schwamendingen-Mitte	0.55	0.63	0.36	0.51
Kreis 1	City	0.00	0.58	0.95	0.51
Kreis 9	Altstetten	1.00	0.05	0.41	0.49
Kreis 6	Unterstrass	0.45	0.42	0.59	0.49
Kreis 6	Oberstrass	0.45	0.68	0.14	0.42
Kreis 2	Wollishofen	0.09	0.47	0.68	0.41
Kreis 10	Hoengg	0.36	0.47	0.36	0.40
Kreis 8	Seefeld	0.18	0.95	0.05	0.39
Kreis 2	Enge	0.09	0.26	0.73	0.36
Kreis 1	Hochschulen	0.00	1.00	0.00	0.33
Kreis 1	Lindenhof	0.00	0.74	0.18	0.31
Kreis 5	Gewerbeschule	0.64	0.00	0.27	0.30
Kreis 1	Rathaus	0.00	0.11	0.45	0.19

Table 2: Neighborhoods sorted by overall suitability.

As was to be expected from the exploratory analyses, the three most central neighborhoods are least suitable as a location for a new restaurant, but if we look at the individual scores, we find that the reasons for their unsuitability appear to be different. For instance, neighborhood Rathaus scores particularly poorly on rent and competition, but is sort of average in terms of friendly businesses, whereas Gewerbeschule, which scores almost as poorly, is actually above average in terms of rent, that is, reasonably affordable. Thus, it becomes clear that the overall score may not be nuanced enough to make an informed decision, but at the same time, it would be tedious to compare each of the neighborhoods one by one based on each of the criteria. A cluster analysis can help reduce the dimensions of the search and allow stakeholders to consider similar neighborhoods based on a weighting of the criteria that they consider most relevant.

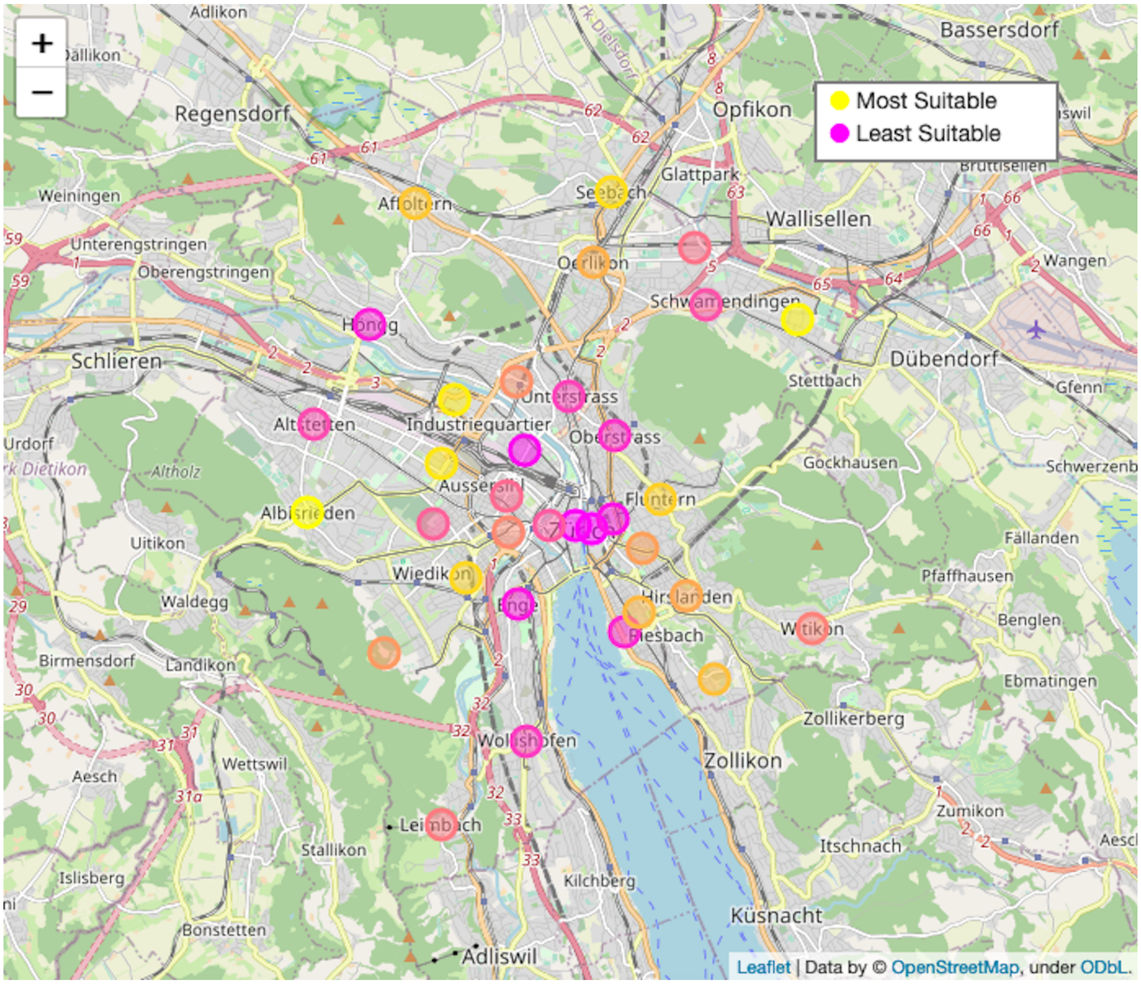

Fig.6: Overall suitability scoring of neighborhoods.

4.3 Clustering Neighborhoods Based on Similarity

A cluster analysis groups neighborhoods based on how similarly they score on the three criteria. Since neighborhoods can often have contradictory scores on each of the criteria (a neighborhood can be expensive but feature loads of friendly businesses), the cluster analysis allows us to go into more detail on how each criterion should be weighted, while also narrowing down the number of items to consider.

Based on a k-means algorithm searching for 4 overall clusters, the following clusters were extracted:

- **Cluster 0**: The perfect spot where all three conditions are above average.
- **Cluster 1**: Neighborhoods with a high number of friendly businesses and little competition, but a fairly high rent.
- **Cluster 2**: Neighborhoods with low rent, some friendly businesses but loads of competition.
- **Cluster 3**: Neighborhoods with hardly any competitors, somewhat high rent and virtually no friendly businesses.

Cluster	Rent.Score	Competition.Score	Friends.Score
0	0.7814286	0.6085714	0.7828571
1	0.1938462	0.5746154	0.7700000
2	0.8380000	0.1480000	0.4440000
3	0.4044444	0.7766667	0.1833333

Table 3: Identified clusters and how each of them scores on the three criteria.

Fig.7: Neighborhood clusters in Zurich.

In contrast to the previous mean-based scores, the cluster means provide more nuanced information on how each criterion is met in a given cluster of neighborhoods. For example, both the neighborhoods City (left of Lindenhof) and Lindenhof were ranked as one of the least suitable ones previously, but we did not know why unless we compared each of the criteria between the two. Using clustering, we can see that neighborhood City belongs to the cluster with a particularly high rent, whereas Lindenhof forms part of the cluster that has almost no friendly businesses. Similar comparisons can be drawn for other neighborhoods.

4.4 Discussion

The decision to open a new restaurant, particularly in a city where 60% of restaurants are in the red, needs to be based on a number of careful considerations in order to ensure the success of its business. The Where is key in this process, as it will decide the number of customers that visit on a daily basis and whether that number is sufficient to pay for the rent, personnel costs and other expenses.

This analysis examined three main factors to characterise Zurich's neighborhood in terms of their suitability for new restaurants: 1) The number of competing food businesses in the neighborhood, 2) the number of other businesses that could attract potential customers and 3) the average rent for commercial space in that area.

Visual inspection of the distribution of restaurants across the city indicated that restaurants are particularly dense in the three central neighborhoods belonging to District 1. At the same time, commercial space is most expensive in these neighborhoods, so that although there may be a lot of potential customers from other businesses, high competition and high rents may reduce the potential for revenue.

When examining areas outside the city center, however, the story becomes a little more complex, as competition, friendly businesses and rental prices can go in opposite directions and cancel one another out. For neighborhoods that score high on all three criteria (Albisrieden, Hirzenbach and Escher Wyss) with overall scores of 0.8 out of 1.0, it is safe to assume that they did not score extremely low on any of the criteria. For lower ranking neighborhoods, however, the overall score is little informative and does not tell us whether a neighborhood scored lower on all three criteria or just on one or two.

The cluster analysis helps us identifying neighborhoods based on how similarly the score on each of the three criteria. For instance, neighborhoods in Cluster 3 feature a very low number of friendly businesses and may therefore score low on the average rating. However, a low number of businesses can mean that there are a lot of parks, hiking paths or other non-commercial spaces that attract people, so that these neighborhoods may be just as suitable as those that meet all three criteria.

Thus, while this analysis constitutes an objective characterisation of neighborhoods, it is only a starting point for stakeholders to reach their final decision, and a number of further considerations will have to be taken into account.

5 Conclusion

The aim of this project was to identify and characterize suitable neighborhoods in Zurich to open a new restaurant. Based on my findings, stakeholders can make an informed choice on where it makes sense to dig deeper, make onsite visits, look for available commercial space etc. and where it does not. In addition, the cluster analysis provides insight into characteristically similar neighborhood and allows stakeholders to decide for themselves how each of the criteria matters in their business decision.

A limiting factor of the current analysis is the use of Foursquare as location information provider, as it only included 1/4 of the restaurants in Zurich, so that findings are likely not representative. Alternative options of location providers would be Yelp, Trip Advisor or similar.

Finally, the reason why most of the restaurants in Zurich do not make any profit is often not even due to bad choice of location but due to personnel and other costs (NZZ). Therefore, the right choice of location does not guarantuee a successful business, but it certainly forms an important pillar of starting one.

6 Links

- The Jupyter Notebook including interactive code and maps can be found here.
- The HTML version of the analysis can be found here.