Exploring Venues for a Swiss Restaurant in Zürich

Introduction

1. The city of Zürich

Zürich is the largest city in Switzerland with a population of over 428'700, an increase of 19'500 since year 2000. 1,4 million people live in the Zürich agglomeration. By the end of 2018, 32% of the city's population was made up of non-Swiss from a total of 172 different countries, all of whom help to make up Zürich's multi-cultural character. 6% of all foreigners who move to Switzerland come to the City of Zürich, emphasizing its central role as an economic hub. Germans make up the largest group with 33'579 living in the city, followed by Italians with 15'080 and Portuguese with 7'826. [4]

The City of Zürich is divided into 12 districts and 34 quarters. The agglomeration of Zürich is made up of the City of Zürich and 130 other municipalities. The lowest point of the city is on the banks of the river Limmat, at 392 metres above sea-level, while the highest – the peak of the Uetliberg mountain – is 871 metres. The water table of Lake Zürich is 405.94 metres above sea-level. [4]

2. Business Problem

Zürich is thus a very international city, with residents coming from all over the world to enjoy its relatively high and stable quality of life. In our study, we want to support our stakeholders in finding the optimal location within the city where to open a new Swiss restaurant. It is not so easy to decide on a venue for such a business: Swiss cuisine is not one of the most popular in the world, so we can expect that many tourists could be willing to try it out, but the large majority of the international population in the city may not be interested in going regularly to a Swiss restaurant. Therefore, we may presume that the majority of the existing Swiss restaurants are concentrated in the city centre and most touristic venues, to benefit from the influx of short term visitors. For our stakeholders, it may then be a good option to look for venues with fewer existing restaurants, so not to be hindered by excessive competition, and with a higher percentage of local Swiss residents, so as to attract a clientele of national residents that are accustomed to the Swiss cuisine and may build up a base of regular and affectionate clients.

In order to assess this problem, we will then be working with the 34 Zürich neighbourhoods and use the “Forsquare API” to build a venue data analysis and subsequently to cluster the various locations in order to identify the most promising places for a Swiss gastronomic business. Within this cluster, we will then be looking for the places with a lower concentrations of existing Swiss restaurants and possibly the lowest percentage of non-Swiss residents, in order to identify the most promising location for our stakeholders’ new business.

Data

In order to approach and analyse our problem, we work with the following data sources:

- List of the 34 Zürich neighbourhoods with their respective total population, number of Swiss residents and percentage of foreign inhabitants, as found in [2].
- Python Geopy, used to get the coordinates for each of the Zürich neighbourhoods, in terms of latitude and longitude. [3]
- Foursquare API, which is instrumental to derive and populate the most common venues and location for each Zürich neighbourhood. [1]

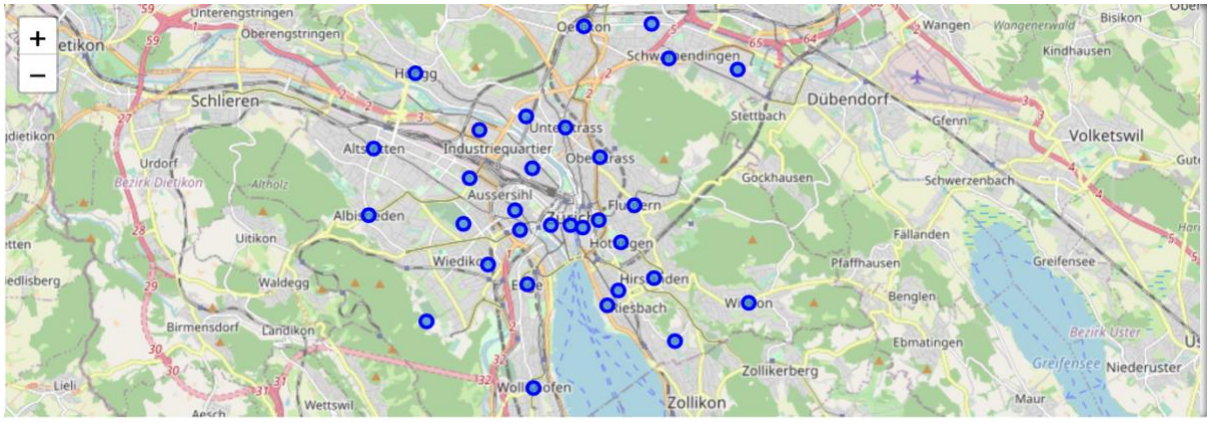
Methodology

1. Inferential Statistical Analysis

In this section, we describe the methodology we followed to generate and analyse the data for the venue location and distribution in the Zürich neighbourhoods. First of all, we assemble the dataset containing the list of 34 neighbourhoods and for each of them the total population, the Swiss population, the foreign inhabitants and their percentage, together with the coordinates (longitude and latitude). Our final dataset after proper cleaning takes the following form (as an example we report the first six neighbourhoods):

	Neighbourhood	Total	Swiss	Foreigner	Foreigner Percentage	Latitude	Longitude
0	Rathaus	3317	2296	1021	30.8	47.372649	8.544311
1	Hochschulen	675	464	211	31.3	47.373846	8.548613
2	Lindenhof	1010	715	295	29.2	47.372996	8.540799
3	City	829	567	262	31.6	47.372943	8.535346
4	Wollishofen	19757	13816	5941	30.1	47.342427	8.530708
5	Leimbach	6140	4113	2027	33.0	47.330511	8.512539

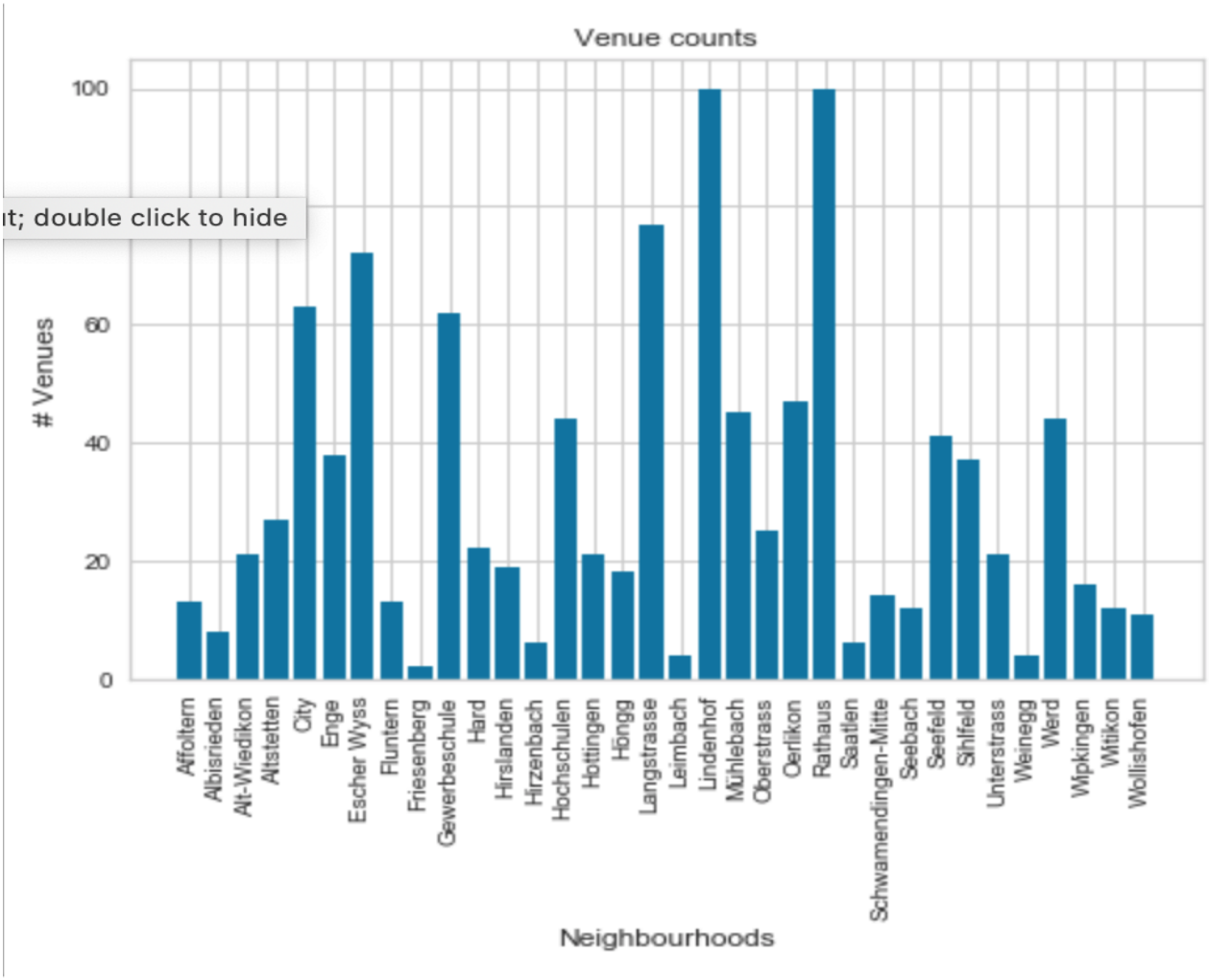
We can inspect the distribution of the neighbourhoods around Zürich using the folium library; we get the following map:



With the help of Foursquare API we can now explore the venues associated to each neighbourhood, with a choice of radius of 500 m for each quartier. We get the distribution of venues, their type and coordinates like in the following table for the 'Rathaus' neighbourhood:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Rathaus	47.372649	8.544311	Fitnesspark Münstergasse	47.370888	8.544999	Gym / Fitness Center
1	Rathaus	47.372649	8.544311	Café Schober	47.371400	8.544149	Café
2	Rathaus	47.372649	8.544311	Schwarzenbach Kolonialwaren	47.371444	8.544091	Gourmet Shop
3	Rathaus	47.372649	8.544311	Neumarkt 17 AG	47.372868	8.546121	Furniture / Home Store
4	Rathaus	47.372649	8.544311	Louis Take Away	47.373037	8.543776	Diner

In the plot below, we can also see how the venue distribution compares for each neighbourhood; we show in fact the count of venues for each quartier:



We see that the more central neighbourhoods, Rathaus and Lindenhof, have the majority of venue counts, as we could have expected, whereas the most peripheral locations have way fewer venues, like Friesenberg, Leimbach or Weinegg.

Another very interesting analysis that we perform is to identify the most common venues for each neighbourhood. This is also instrumental for the machine learning algorithm we are going to discuss in the next section and for tackling our problem of finding the best location for a new Swiss restaurant. Therefore, we use the results from Foursquare API calculation to determine the top ten venues for each neighbourhood. See for example the result for the first five neighbourhoods in the table below:

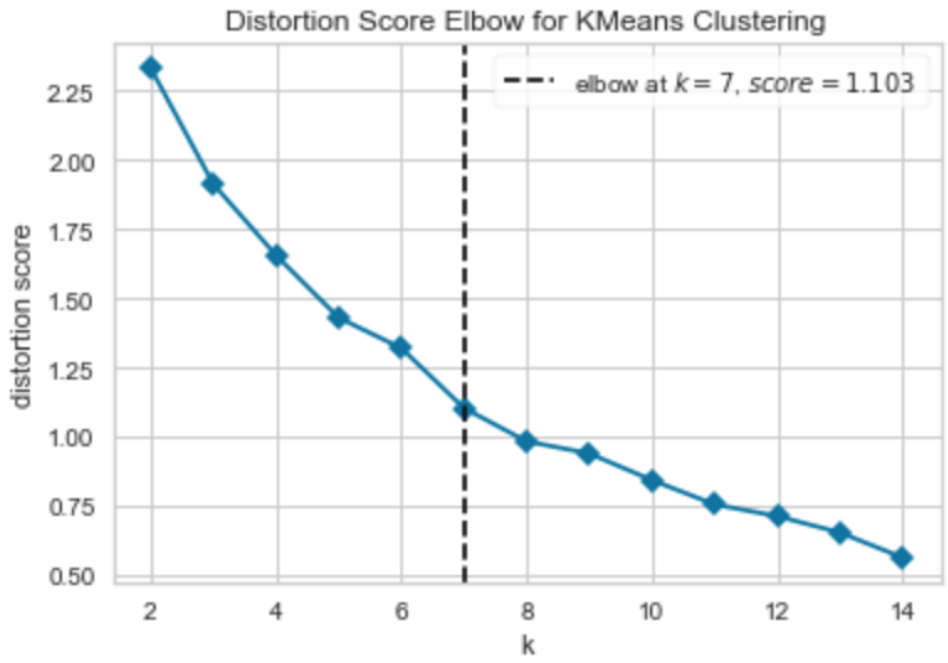
	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Affoltern	Bus Station	Supermarket	Athletics & Sports	Light Rail Station	Train Station	Gym / Fitness Center	Diner	Grocery Store	Italian Restaurant	Department Store
1	Albisrieden	Pizza Place	Trattoria/Osteria	Supermarket	Swiss Restaurant	Scenic Lookout	Bus Station	Grocery Store	Bakery	Diner	Dessert Shop
2	Alt-Wiedikon	Restaurant	Italian Restaurant	Playground	Bakery	Lounge	Farmers Market	Burrito Place	Light Rail Station	Supermarket	Beer Garden
3	Altstetten	Supermarket	Mediterranean Restaurant	Bus Station	Bakery	Plaza	Doner Restaurant	Mexican Restaurant	Café	Food	Burger Joint
4	City	Bar	Italian Restaurant	Department Store	Restaurant	Vegetarian / Vegan Restaurant	Plaza	Swiss Restaurant	Cocktail Bar	Café	Hotel

We can already infer some interesting conclusions from this table: central neighbourhoods like ‘City’ have already a high concentration of gastronomic venues due to their location and the presence of more tourists. More peripheral places like ‘Affoltern’ do have instead a higher concentration of centres for leisure activities, like fitness centres, and more utility stores or stations.

2. Machine learning analysis

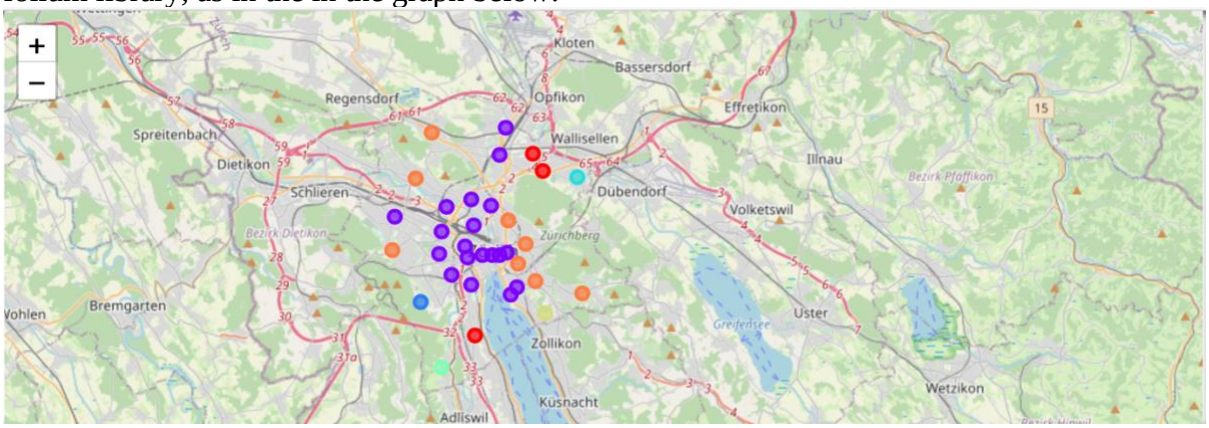
We are now in a good place to describe our machine learning algorithm and the corresponding analysis. We apply an unsupervised learning algorithm for clustering, K-Means. This is one of the simplest and most common clustering techniques and we aim at using it to identify groups of venues that share characteristics in common. By grouping the different neighbourhoods in clusters, we may be able to identify if there is any particular set of locations which are more likely to be associated with restaurants or gastronomic businesses and therefore better suited for our new venture.

We start by performing an analysis to find out the optimal number of clusters by using the elbow algorithm. We get the following result:



From the plot above, we see that in a range from 2 to 14, the optimal number of clusters is 7. We proceed then by clustering our neighbourhoods using the K-Means algorithm with 7

clusters. We get seven different groups of neighbourhoods which we can visualize using the folium library, as in the in the graph below.



Here, each different colour corresponds to a different cluster and we see that the majority of the neighbourhoods belongs to the violet cluster, mainly distributed along the city centre. Upon closer inspection, we can identify the following seven clusters:

- Supermarkets and restaurants (red blobs): the most common venues here are supermarkets and restaurants, distributed slightly out of the city centre;
- Bars and restaurants area (violet blobs): this is the main cluster for eating out and finding a hotel, as bars, restaurants and hotels are the most common venues distributed around the city centre;
- Fliesenberg (blue blob): known for its tennis court area;
- Hirzenbach (light blue blob): various stores and a flea market;
- Leimbach (light green blob): gas station;
- Weinegg (yellow blob): Yoga studio and farmers market;
- Tram and bus stations (orange blob).

After classifying our neighbourhoods by mean of our machine learning algorithm, we are now in a position to analyse the clusters in order to find an answer to our initial question: which cluster and neighbourhoods offer the best chances for a new Swiss restaurant?

Results

Looking at the clusters we got in the previous section, we can infer that the best location for a new Swiss restaurant would be in the second clusters, the bars and restaurants area. Many similar businesses are also located here and we can therefore expect to have the best chances to attract visitors. As noted in the Introduction, however, we want to focus on a more national clientele and avoid neighbourhoods which are already crowded with Swiss restaurants for tourists. Therefore, we search among the neighbourhoods in the second (violet) cluster those which fulfil the following conditions:

- there are no existing Swiss restaurants among the top ten venues;
- there is the lowest percentage of foreign inhabitants, so as to have better chances to attract Swiss customers who already know and grew up with this cuisine.

We get the following table by excluding neighbourhoods in this cluster which already have Swiss restaurants and sorting the remaining by the percentage of foreign residents:

	Neighborhood	Total	Swiss	Foreigner	Foreigner Percentage	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
27	Wipkingen	16605	12127	4478	27.0	47.393495	8.528602	1	Grocery Store	Italian Restaurant	Train Station	Supermarket	Restaurant
15	Unterstrass	24493	17683	6810	27.8	47.391447	8.539572	1	Bakery	Middle Eastern Restaurant	Café	Italian Restaurant	Hotel
9	Sihlfeld	21921	15118	6803	31.0	47.373218	8.510820	1	Café	Italian Restaurant	Bar	Plaza	Supermarket
14	Escher Wyss	6198	4187	2011	32.4	47.390899	8.515360	1	Café	Hotel	Bar	Restaurant	Gym / Fitness Center
7	Alt-Wiedikon	17764	11707	6057	34.1	47.365562	8.517851	1	Restaurant	Italian Restaurant	Playground	Bakery	Lounge
22	Mühlebach	6414	4158	2256	35.2	47.360772	8.554362	1	Italian Restaurant	Café	Supermarket	Hotel	Salad Place
25	Altstetten	34098	21830	12268	36.0	47.387403	8.486061	1	Supermarket	Mediterranean Restaurant	Bus Station	Bakery	Plaza
29	Oerlikon	23548	14557	8991	38.2	47.410421	8.544585	1	Supermarket	Italian Restaurant	Tram Station	Hotel	Indian Restaurant
30	Seebach	26056	15869	10187	39.1	47.420438	8.548377	1	Hookah Bar	Tram Station	Korean Restaurant	Laser Tag	Eastern European Restaurant

We see from the result above that **Wipkingen** is a good candidate for our client: it lies in the "Restaurant cluster", it does not have yet many Swiss restaurants (this venue does not figure in the top ten) and the percentage of foreign inhabitants is relatively low, 27%, especially if compared with the most international neighbourhoods which can reach peaks of 43%-45% of foreign residents. Wipkingen is also quite centrally located and it occupies that part of the city which is most in development ("Kreis 5"), with modern new buildings and a growing population. We can therefore reinforce our suggestion also based on these considerations.

Discussion

Despite the promising conclusion that we got in the previous section, there are some limitations to our analysis and possible extensions that could potentially make our conclusions more robust.

First of all, we used the population data for 2020 only, meaning we analysed the concentration of foreign and Swiss residents based solely on this year. A more sophisticated approach could take into account the population data from the demographic register and build a predictive model on how the population is going to evolve in the future. Do we expect a growing immigration in the city? From which nationality groups and concentrated in which neighbourhoods? Answering this questions could help further analysing the ethnic diversity of Zürich and possibly identifying restaurant venues for different types of cuisines based on how diverse/international a neighbourhood is.

Another possible extension could be to consider a wider range of neighbourhoods and sub-neighbourhoods, by dividing each quartier into its subdivision. We could also consider the greater Zürich area, including the peripheral cities counting for more the 1 million people. We could then analyse a wider distribution of venues and possibly have a better statistical basis to conclude on where, within and around Zürich, a business could be more successful.

Conclusions

We conclude our study by noting that with a very simple approach based on the K-Means algorithm and the geographical distribution of the Zürich neighbourhoods we could infer the possible location of a new Swiss restaurant within the city. We can corroborate the result of the analysis by qualitative observations based on the recent urban development of the city and the increase in population in the Wipkingen area, which appears to be a lively and promising quartier where to start a new business. We highlighted limitations and possible extensions of our methodology, which can be explored in the future and be of interest for new business entrepreneurs.

Bibliography

1. (n.d.). Retrieved from https://developer.foursquare.com
2. Präsidialdepartment, S. Z. (n.d.). *Kreise und Quartiere*. Retrieved from https://www.stadt-zuerich.ch/prd/de/index/statistik/themen/bevoelkerung/bevoelkerungsentwicklung/kreise-und-quartiere.html
3. *Python Geopy*. (n.d.). Retrieved from https://geopy.readthedocs.io/en/stable/
4. *Stadt Zürich*. (n.d.). Retrieved from Facts & Figures : https://www.stadt-zuerich.ch/portal/en/index/portraet_der_stadt_zuerich/zahlen_u_fakten.html