Coursera Capstone – "The Battle of the Neighbourhoods"

Expanding into Europe – Zurich District Analysis

Table of Contents

1 Problem Statement

An Analytics Consulting firm based in New York, is looking to expand into Europe by opening a small office in Zurich, Switzerland. The City of Zürich is divided into 12 districts and 34 quarters. The Zürich Metropolitan Area has some 1.9 million inhabitants and 900,000 employees spread across 238 municipalities in eight cantons[1].

The firm has already noted Zurich's advantages in relation to its solid transport links. Zürich central station is one of Europe's main railway intersections, with over 1,900 connections every day. Intercity trains leave every half-hour to various hubs in Switzerland and Europe, and the city is only a matter of hours by train from Paris (4.5 hours), Milan (3.5 hours), Munich (5 hours) and Hamburg (7 hours). Air journey times from the international airport to these destinations are around one hour[2].

This project aims to use Foursquare location data to explore and cluster Zurich's 12 districts to assist the firm's partners in deciding which districts to shortlist based on surrounding venues and prospective rental costs. The firm has detailed its priority criteria for its current and forthcoming employees and clients:

- transport accessibility
- proximity to clients
- local hospitality and hotels
- shopping
- dining possibilities

---

[1] https://www.stadt-zuerich.ch/portal/en/index/portraet_der_stadt_zuerich/zahlen_u_fakten.html
[2] https://www.stadt-zuerich.ch/portal/en/index/portraet_der_stadt_zuerich/zahlen_u_fakten.html

2 Data Description

To present the firm's partners with the project's resulting clustering analysis, the data that will be used is a CSV file prepared from multiple sources (Zurich_districts.csv) which includes geolocation data as well as location and venue information from FourSquare:
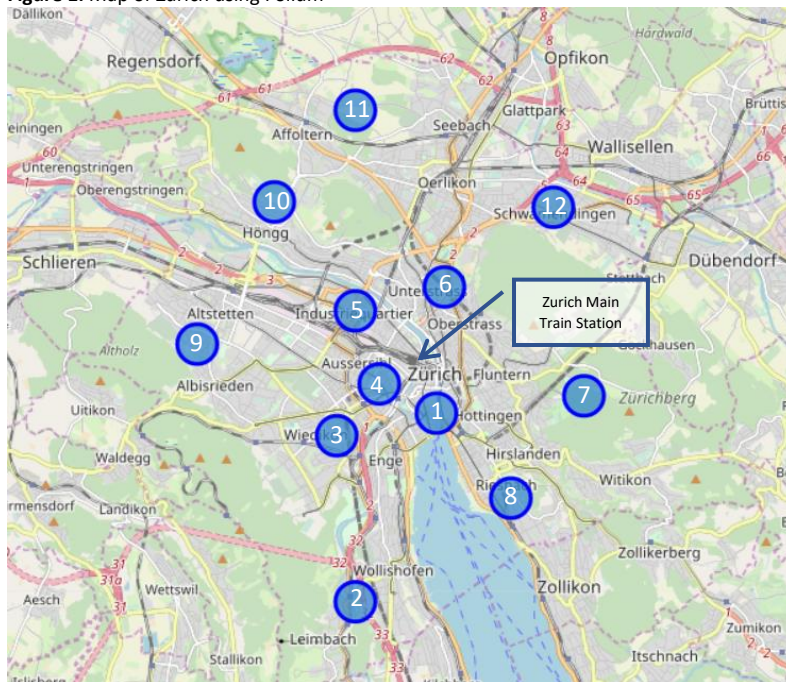
1. Data about Zurich and its 12 districts including their respective latitudes and longitudes (these will be scraped and cleansed accordingly) as shown in table 1. Each district in Zurich is officially numbered from 1 to 12[3].

Table 1: Zurich_districts.csv

No.	Latitude	Longitude	MedianRent
1	47.370	8.542	525
2	47.337	8.521	450
3	47.366	8.516	250
4	47.375	8.527	320
5	47.388	8.521	350

2. Data from Swiss real estate company JLL to append median rents (CHF/m^2 per annum) in all 12 districts for further analysis in clustering (these will be appended in the CSV file)[4].

3. FourSquare API will be queried using the "explore" endpoint and will provide the surrounding venues of a given coordinates including restaurants, hotels and coffee shops, which will then be converted into counts to apply one hot encoding to cluster the districts.

As shown in figure 1, Folium was used to visualize Zurich and its 12 districts, using the latitude and longitude data from (Zurich_district.csv). Please note the proximity of the districts to the main Zurich train station, which serves as a direct link to Zurich airport as well as numerous international destinations.

Figure 1: Map of Zurich using Folium

---

[3] https://en.wikipedia.org/wiki/Subdivisions_of_Z%C3%BCrich
[4] http://www.jll.ch/switzerland/en-gb/Research/jll-switzerland-office-market-2019.pdf

3 Methodology

3.1 Exploratory data analysis

Foursquare API has been used to explore and segment Zurich's districts, setting the limit as 100 venues and the radius 500 meters for each Zurich district from their given latitude and longitude. As shown in table 2, is a head of the list Venues name, category, latitude and longitude information from Foursquare API is queried for *District 1*.

Table 2: Zurich_districts.csv

	name	categories	lat	lng
0	Sprüngli	Dessert Shop	47.369398	8.539486
1	Frauenbad Stadthausquai	Pool	47.368412	8.542296
2	Ban Song Thai	Thai Restaurant	47.369467	8.544068
3	Zeughauskeller	Swiss Restaurant	47.370250	8.539828
4	Café Schober	Café	47.371400	8.544149

Figure 2: Number of venues for each district
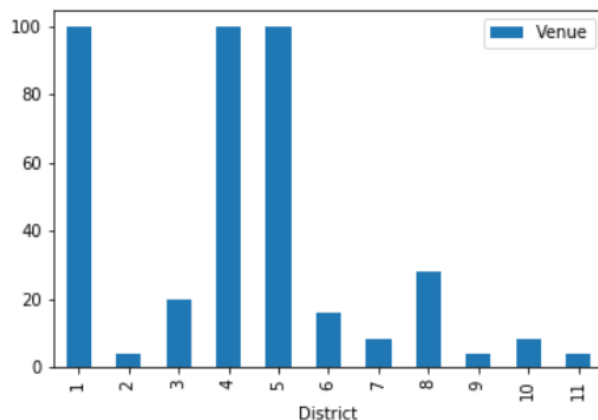
In plotting the number of venues against the specific districts, we can see the more popular districts by venue concentration. Looking at figure 2, *Districts 1, 4, and 5*, suggest levels high commercial activity. This also coincides with proximity to Zurich's main train station, which also suggests a higher level of coveted office space. As we continue at add more factors into the analysis, the chart serves well to keep in mind ways to label the resulting clusters from subsequent sections.

In table 3 below, the districts are listed with the top 10 venues by mean occurrence. *Districts 1 and 5* are highlighted, 1 and 5 at this stage to earmark the only districts with hotels in the top 10 which satisfies the "local hospitality and hotels" criteria. Even prior to clustering, it is evident that as the radius from Zurich main train station increases we begin to see a different dynamic of venues, as shown by district 11 which is although home to important clients has "Farms" and "Forests" as the most common venue which is unlikely to suit the firm's criteria for an office.
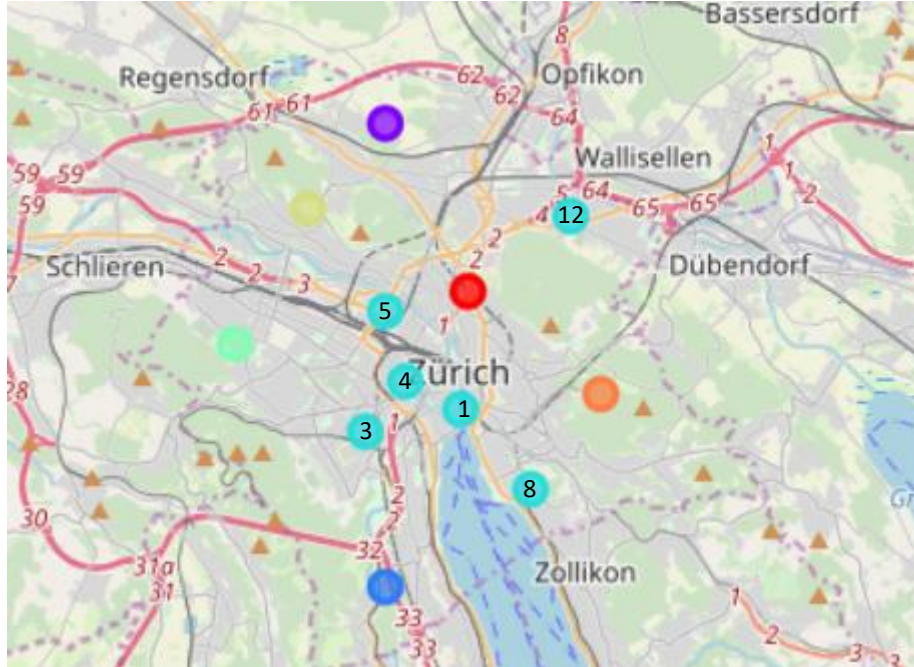
Table 3: Top 10 most common venues for each district

District	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
1	Swiss Restaurant	Café	Plaza	Bar	Boutique	French Restaurant	Cocktail Bar	Hotel	Italian Restaurant	Restaurant
2	Light Rail Station	Convenience Store	Bus Station	Yoga Studio	Gay Bar	Electronics Store	Falafel Restaurant	Farm	Fast Food Restaurant	Food
3	Restaurant	Italian Restaurant	Light Rail Station	Playground	Diner	Food	Bus Station	Burrito Place	Lounge	Beer Garden
4	Bar	Italian Restaurant	Restaurant	Swiss Restaurant	Café	Thai Restaurant	Coffee Shop	Chinese Restaurant	Nightclub	Pub
5	Nightclub	Bar	Restaurant	Café	Hotel	Bakery	Tram Station	Italian Restaurant	Pizza Place	Mexican Restaurant
6	Italian Restaurant	Bakery	Bistro	Indian Restaurant	Grocery Store	Middle Eastern Restaurant	Food & Drink Shop	Park	Falafel Restaurant	Coffee Shop
7	Bar	Swiss Restaurant	Music Venue	Pool	Tram Station	Golf Course	Skating Rink	Scenic Lookout	Food & Drink Shop	Food
8	Swiss Restaurant	Restaurant	Italian Restaurant	Museum	Tram Station	Bakery	Beach	Café	Mexican Restaurant	Coffee Shop
9	Middle Eastern Restaurant	Soccer Field	Moving Target	Bus Station	Yoga Studio	Furniture / Home Store	Electronics Store	Falafel Restaurant	Farm	Fast Food Restaurant
10	Bus Station	Grocery Store	Wine Bar	Steakhouse	Soccer Field	Other Great Outdoors	Food & Drink Shop	Food Truck	Food Court	Yoga Studio
11	Farm	Swiss Restaurant	Forest	Bus Station	Yoga Studio	Gay Bar	Eastern European Restaurant	Electronics Store	Falafel Restaurant	Fast Food Restaurant
12	Arts & Crafts Store	Bus Station	Restaurant	Thai Restaurant	Shopping Mall	Café	Supermarket	Swiss Restaurant	Plaza	Tram Station

3.2 Clustering

The subsequent step in the project is to evaluate which district is most suited using the criteria described above. The unsupervised learning K-means algorithm has been used to cluster the districts. Although an "elbow curve" was generated to choose the appropriate cluster number, it's evident from the labelled "Cluster 3" below that other districts stand alone in dissimilarity.

Figure 3: Zurich District Clustering (Cluster 3 has been labelled with respective district numbers)

'Cluster 3' below provides a clear distinction in terms of dining possibilities with plenty of diverse restaurants, shopping choices with plazas and shopping malls, as well as transport accessibility with proximity to Zurich main station as well as more flexible tram and bus stations.

Table 4: "Cluster 3"

No.	MedianRent	Cluster Labels	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
1	525	3	Swiss Restaurant	Café	Plaza	Hotel	French Restaurant	Bar	Boutique	Cocktail Bar	Italian Restaurant	Restaurant
3	250	3	Restaurant	Italian Restaurant	Light Rail Station	Playground	Diner	Food	Bus Station	Burrito Place	Lounge	Beer Garden
4	320	3	Bar	Italian Restaurant	Restaurant	Swiss Restaurant	Café	Thai Restaurant	Coffee Shop	Chinese Restaurant	Nightclub	Pub
5	350	3	Nightclub	Bar	Restaurant	Café	Hotel	Bakery	Tram Station	Italian Restaurant	Pizza Place	Mexican Restaurant
8	410	3	Swiss Restaurant	Restaurant	Italian Restaurant	Museum	Tram Station	Bakery	Beach	Café	Mexican Restaurant	Coffee Shop
12	210	3	Arts & Crafts Store	Bus Station	Restaurant	Thai Restaurant	Shopping Mall	Café	Supermarket	Swiss Restaurant	Plaza	Tram Station

At this stage, it is appropriate to drop District 8 out of the final shortlist owing to distance from the centre and proximity to clients. *District 4* also known colloquially Chreis Cheib (animal cadaver), earned the name as the area historically contained pits for the deposition of dead animals, will also be dropped for strong similarity and immediacy to *District 3*. This leaves Districts 1, 3, 5 and 12 in the shortlist for the final recommendation stage.

4 Final recommendations
Analysing the final shortlist.

District 1 looks to be the most preferable option in terms of optimum location but comes at a steeper median rent to the three districts. Bahnhofstrasse, Zurich's main downtown street, is home to a plethora of dining and shopping options in the immediate vicinity. The district is also home to numerous renowned consultancy companies who are already on the books as clients for firm, while local hospitality is second to none of the shortlisted districts as aforementioned. As shown in figure 4, smaller offices in shopping streets can be an option in growing the business in Zurich in the first few years of the expansion strategy.

Figure 4: Nüschelerstrasse 1, *District 1*

District 3 lies to the south west of the city centre and is bordered by the river Sihl and parts of the Albis mountain range. The district is known to be largely residential as shown by a high concentration of "playgrounds" and transport options but is home to several authentic international restaurants. The district can easily be reached from the city centre in 5-10 minutes by public transport or around 40 minutes on foot, and could be an option. The district could serve the firm as a more scenic option away from the hustle and bustle of the city centre.

Figure 5: Giesshübelstrasse 40, *District 3*

District 5, also known as Industriequartier, is Zurich's trendy and rising district. As shown in previous sections, the district's most common venues are nightclubs and bars but has a strong selection of hotels and restaurants. The once abandoned industrial quarter has undergone a remarkable transformation in recent decades, helping revamp the image of Switzerland's largest city from conservative and sleepy to trendy and modern. *District 5* lies in the opposite direction, offering grittier vistas of the railroad tracks on one side and the Limmat River on the other, with an eclectic mix of businesses, flats, art galleries, hip restaurants and lovely parks randomly squeezed in between[5].

Figure 6: Hardturmstrasse 76, *District 5*

District 12's Stettbach's train station has a 20-minute train or 12-minute drive to the airport. The district is also in closer proximity to the firm's higher margin pharmaceutical clients, which are on average a 5-10-minute drive from an office situated in the district, which could provide a higher value for investment. The district also has the lowest median rent which furthers the firm's choice as a potential less risky option for the first stage of its European entrenchment.

---

[5] https://www.thelocal.ch/20140413/gritty-district-turns-into-zurichs-trendiest-spot

5 Conclusion

The project has provided the Analytics Consulting firm with a short-list of 4 clustered districts, which all fit the firm's office criteria but have distinct nuances based on the criteria specified. While *District 1* will be the most expensive investment, the requirement of a few employees and smaller office space in the first two years would provide the firm strong foundations with suppliers and potential clientele.