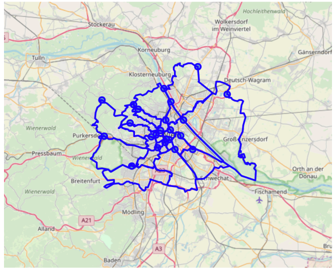


Vienna

- Where to settle for entertainment and recreational purposes as a student or young urban professional

Joern Grimmer

Coursera – Applied Data Science – Capstone Project
January 2020

Agenda

Abstract

Where to settle for entertainment & recreational purposes as a student or young urban professional

- According to a recent study, Vienna is the most attractive city worth living for the second year in a row.
  In our project we will try to find an optimal location so settle in Vienna, when your major interest is in recreational & sport facilities, restaurants and night life entertainment close to a university faculty of your choice
- The data of Vienna and its universities, faculties and colleges was taken from www.data.gv.at[1], a central catalogue of Austrian public data sources .The datasets included the geospatial data of Vienna‘s districts, universitites, faculties and colleges. Venue data was taken from the Foursquare API.
- Our investigation is focused around clustering the districts near the university campus of choice - in our case “Akademie der Bildenden Künste” – in order to provide the most feasible cluster to a student’s preference with respect to the assumed scope of interest.
- We make use of the k-Means Clustering method. The objective is to cluster districts on their venues of recreational and entertainment venues.
- Overall, the results demonstrate that most of Vienna inner districts have a similar dense offering of recreational venues, namely restaurants and cafes. K_means clustering provide indication of clusters although the evidence for distinctive clusters is limited as the elbow method demonstrates.

1 https:// www.data.gv.at

Motivation

Why this study?

- According to a recent study, Vienna is the most attractive city worth living for the second year in a row. (Source: Economist Intelligence Unit, Global Livability Ranking 2019). This study ranks 140 cities for their urban quality of life based on assessments of stability, healthcare, culture and environment, education and infrastructure. No surprise, Vienna is attractive for students and young urban professionals to start their studies or working life
- In our project we will try to find an optimal location so settle in Vienna, when your major interest is in recreational & sport facilities, restaurants and night life entertainment close to the university faculty of your choice
- This report is targeted at students & young urban professionals planning to settle in Vienna. We assume this target group is interested in general to settle close to spots of their daily live to minimize distances between every day lcoations.

Data set used in our project

Dataset was taken from www.data.gv.at and Foursquare

- The data of Vienna and its universities, faculties and colleges was taken from www.data.gv.at[1], a central catalogue of Austrian public data sources dedicated to provide the metadata of otherwise decentralized Austrian administrative data catalogues. The datasets included the geospatial data of Vienna‘s districts, universitites, faculties and colleges. Venue data was taken from the Foursquare API
- The data for Vienna‘s districts was taken from https://www.data.gv.at/katalog/dataset/stadt-wien_bezirksgrenzenwien/ (BEZIRKSGRENZEOGD.csv)
- The data for Vienna‘s universities, faculties and colleges was taken frrom https://www.data.gv.at/katalog/dataset/stadtwien_universittenundfachhochschulenstandortewien/ (UNIVERSITAETOGD.csv)
- Venue data for selected districts of Vienna was sourced from the Foursquare API
- The data sources used in our project will be described on the next slides in more detail

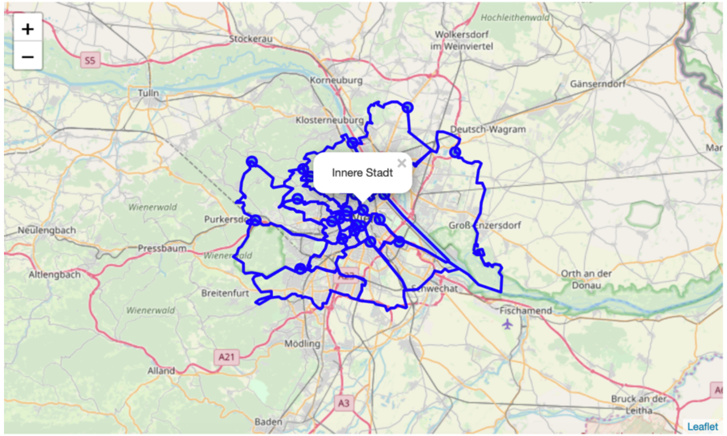
Shape of dataset (I)

2

Vienna: 23 districts including district code and geospatial data (Longitude/ Latitude)

	index	Borough	BEZNR	DISTRICT_CODE	STATAUSTRIA_BEZ_CODE	STATAUSTRIA_GEM_CODE	Longitude	Latitude
0	3	Innere Stadt	1	1010	901	90101	16.372641	48.216617
1	6	Leopoldstadt	2	1020	902	90201	16.403453	48.231919
2	1	Landstraße	3	1030	903	90301	16.396617	48.207387
3	19	Wieden	4	1040	904	90401	16.369165	48.200713
4	18	Margareten	5	1050	905	90501	16.359449	48.196617
5	20	Mariahilf	6	1060	906	90601	16.363064	48.201827
6	0	Neubau	7	1070	907	90701	16.338725	48.208537
7	2	Josefstadt	8	1080	908	90801	16.349147	48.215158
8	5	Alsergrund	9	1090	909	90901	16.361652	48.231918
9	16	Favoriten	10	1100	910	91001	16.383819	48.185157
10	15	Simmering	11	1110	911	91101	16.425986	48.185575
11	17	Meidling	12	1120	912	91201	16.341743	48.188466
12	21	Hietzing	13	1130	913	91301	16.214234	48.206523
13	10	Penzing	14	1140	914	91401	16.209138	48.264112
14	22	Rudolfsheim-Fünfhaus	15	1150	915	91501	16.327324	48.205005
15	4	Ottakring	16	1160	916	91601	16.276206	48.227037
16	8	Hernals	17	1170	917	91701	16.285159	48.256800
17	7	Währing	18	1180	918	91801	16.295017	48.249609
18	11	Döbling	19	1190	919	91901	16.356813	48.282287
19	9	Brigittenau	20	1200	920	92001	16.373612	48.261269
20	13	Floridsdorf	21	1210	921	92101	16.437762	48.316811
21	12	Donaustadt	22	1220	922	92201	16.507839	48.273446
22	14	Liesing	23	1230	923	92301	16.280553	48.159055

Table : Vienna‘s districts



1 Source: https://www.data.gv.at/katalog/dataset/stadt-wien_bezirksgrenzenwien/ (BEZIRKSGRENZEOGD.csv)

Shape of dataset (II)

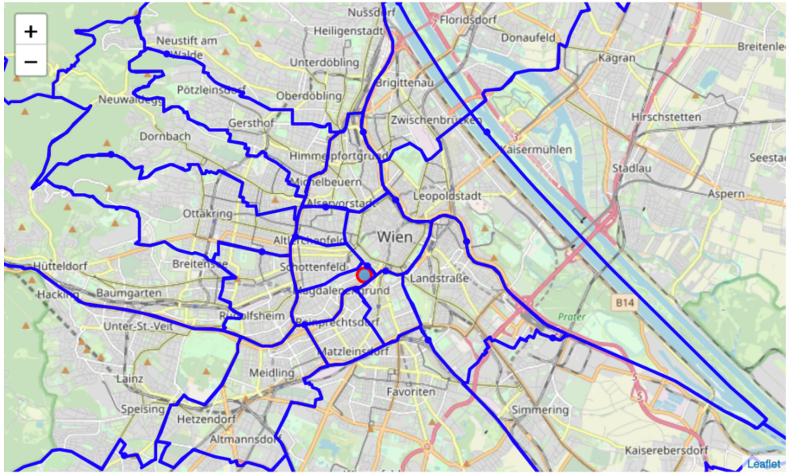
2

Vienna: 158 universities/ faculties or colleges and geospatial data (Longitude/ Latitude)

	NAME	Longitude	Latitude
0	FH Technikum Wien	16.377856	48.239443
1	FH Technikum Wien	16.426908	48.269503
2	FH Campus Wien	16.382288	48.157733
3	Fachhochschule des bfi Wien	16.403446	48.219132
4	Fachhochschule des bfi Wien	16.426908	48.269503
5	FHWien-Studiengänge der Wirtschaftskammer Wien	16.349201	48.226579
6	Lauder Business School	16.352469	48.242701
7	FH Technikum Wien	16.355891	48.200143
8	Technische Universität Wien	16.363088	48.200171
9	Akademie der bildenden Künste Wien	16.361984	48.199806
10	Universität Wien	16.348993	48.233515

Table : Universities and Colleges of Vienna (First 10 entries only)

Faculty of choice for further analysis



1 Source: https://www.data.gv.at/katalog/dataset/ stadtwien_universittenundfachhochschulenstandortewien/ (UNIVERSITAETOGD.csv)

Data Preparation and Cleaning

2

A three-step approach to prepare data

- Drop columns not required
- Sorting colums (districts) in ascendig order
- Extracting a latitude/ longitude datapoint from a polygone shape

Research Question & Method

Unsupervised machine learning – Clustering with k-Means algorithm

Are there clusters of similar but distinguishing recreational venues close to a university location of choice?

- Unsupervised machine learning: K-Means algorithm
  - Selection of ten districts closest to university of choice. Distance of districts taken as they appear closest on map as a matter of simplification, no calcuation of distances from the Polygon midpoint of the districts
  - Start with 5 clusters
  - Run algorithm 10 times independently with different random centroids to choose final model as the one with the lowest SSE
  - Maximum number of iterations: 300
  - Elbow method to verify the optimal number of clusters

Initial K-Means clustering with 5 clusters

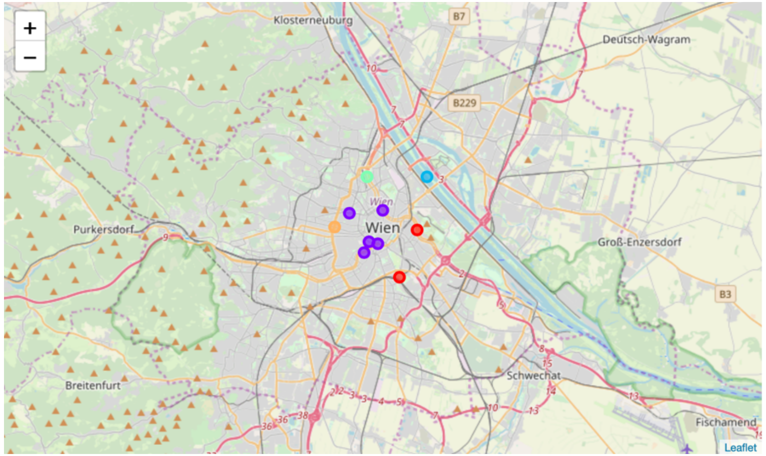
4

Five clusters identified, however clusters show overlaps in the top three most common venues

		Most Common Venue		
Borough	Cluster Label	1st	2nd	3rd
Landstraße	0	Art Museum	Bus Stop	Hotel
Favoriten	0	Hotel	Cafe	Restaurant
Innere Stadt	1	Restaurant	Cafe	Hotel
Wieden	1	Hotel	Cafe	Plaza
Margareten	1	Cafe	Bar	Asian Restaurant
Mariahilf	1	Cafe	Museum	Plaza
Josefstadt	1	Hotel	Cafe	Restaurant
Leopoldstadt	2	Beach Bar	Latin American Restaurant	Mediterranean Restaurant
Alsergrund	3	Hotel	Austrian Restaurant	Nightclub
Neubau	4	Bakery	Turkish Restaurant	Fast Food Restaurant

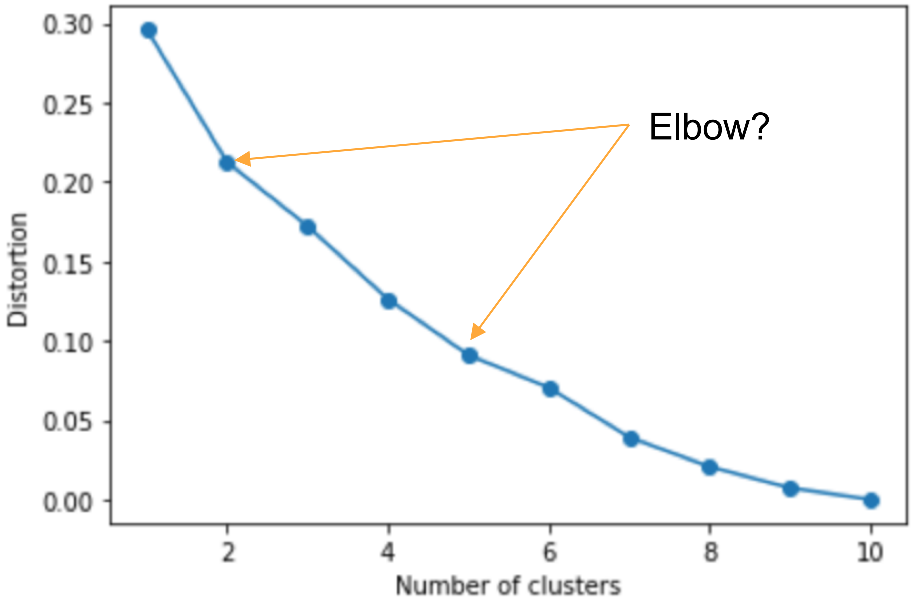
Cluster 0: „Museum & Hotel“
Cluster 1: „Cafe & Restaurant“
Cluster 2: „Maritim & Mediterrenean restaurants“
Cluster 3: „Domestic Food & Nightclub“
Cluster 4: „Bakery & Fast Food“

Visualization: Clustered districts



No clear evidence on where the elbow is located

4



- The curve is flattening first at 2 slighlty, then a second time at 5, before it is steepening at 6 again

Second iteration K-Means clustering with 2 clusters

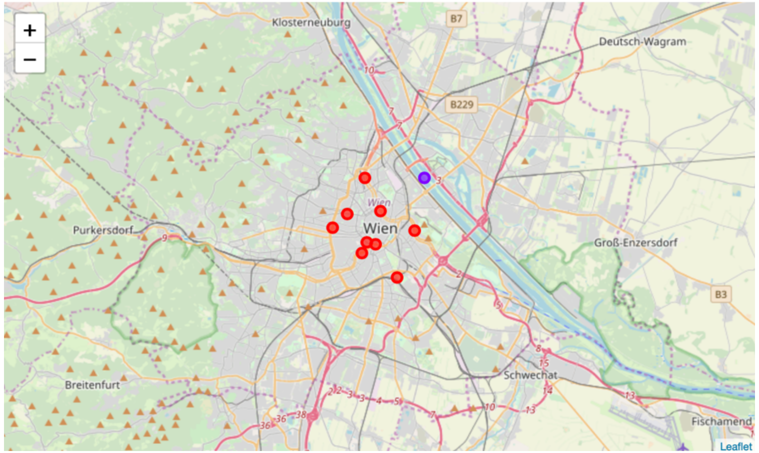
4

With two clusters identified, a Restaurant/Cafe/ Hotel cluster is separated from a Maritim/ Mediterrenean cluster restaurant cluster

		Most Common Venue		
Borough	Cluster Label	1st	2nd	3rd
Alsergrund	0	Hotel	Austrian Restaurant	Nightclub
Favoriten	0	Hotel	Cafe	Restaurant
Innere Stadt	0	Restaurant	Cafe	Hotel
Josefstadt	0	Hotel	Cafe	Restaurant
Landstraße	0	Art Museum	Bus Stop	Hotel
Margareten	0	Cafe	Bar	Asian Restaurant
Mariahilf	0	Cafe	Museum	Plaza
Neubau	0	Bakery	Turkish Restaurant	Fast Food Restaurant
Wieden	0	Hotel	Cafe	Plaza
Leopoldstadt	1	Beach Bar	Latin American Restaurant	Mediterranean Restaurant

Cluster 0: Restaurant/Cafe/Hotel
Cluster 1: Maritim/ Mediterrenean restaurants

Visualization: Clustered districts



Limitations

- The gesopatial shape of Vienna's districts is provided by a Polygon dataset of Longitude/ Latitude datapoints. We extract the first Longitude/ Latitude datapoint instead of calculating the mid-point of the Polygon for simplification. Thus, the selection of venues in a radius of 500 metres around this datapoint might bias our results
- Second, the Foursquare data limits the radius und number of venues. Again, this limitation might bias our clustering results
- While k-means is very good at identifying clusters with a spherical shape, we observe the distribution of our venue data seems not to provide such a shape within the given districts of Vienna
- Nevertheless, a comparison with an external qualitative study confirms our results on a high level[1]

1 https://austrianadaptation.com/where-to-live-vienna/ Title: Where to live in Vienna. Your epic guide to Vienna's Districts, In Living Abroad, Vienna, Vienna Local Experiences by Carly, November 2, 2017

Conclusions

5

- The project assignment was targeted to find an optimal location so settle in Vienna, when the major interest is in recreational & sport facilities, restaurants and night life entertainment close to the university faculty of your choice

- The popular kMeans clustering method delivered cluster results which were limited. The reason for this was found in the apparently even distribution of recreational venues across the districts rather than having a distinct sperical shaped distribution of specific recreational venues

- Nevertheless, the study can serve as a starting point for an additional analysis of other preferences, e.g. cultural diversity or distance to public transportation et.al., to find the most preferred spot in line with your preferences to settle near your university or faculty of choice in Vienna. This type of data is available on www.data.gv.at., too

References

6

- https:// www.data.gv.at
- https://de.foursquare.com
- https://austrianadaptation.com/where-to-live-vienna/ Title: Where to live in Vienna. Your epic guide to Vienna's Districts, In Living Abroad, Vienna, Vienna Local Experiences by Carly, November 2, 2017