Finding the best place for a Coffee House

in Vienna, Austria

Applied Data Science Capstone

via IBM/Coursera

Final Report by John Oxales

June 10, 2021

Table of contents

---

1. Introduction: Business Problem

1.1 Background

Vienna's very first coffee house opened in 1683. Even if Vienna was not the pioneer in coffee house culture, it has - over the centuries - established a coffee house tradition like no other city in the world. Coffee and coffee houses are at their best in Vienna!

1.2 Problem

Since so many coffee houses can be already found in any district in Vienna it would be good to know where it would be still reasonable to establish a new coffee house business. It can be assumed that high frequented places like sightseeing spots or malls already have some

coffee houses in their vicinity. So an approach would be to find the neighborhood which is still near an area of interest but not overcrowded with or, even better, still lacking a coffee house.

1.3 Interest

These questions could be of high interest for any new investor or entrepreneur who wants to open one or maybe even more coffee houses in Austria's capital. This project can also be part of business analytics for larger companies who consider expanding their business. According to the findings, after the following thorough data analysis, those stakeholders can decide where they might open their new coffee house.

For this project I am going to focus on neighborhoods near Vienna's famous historical site: St. Stephen's Cathedral. This place has a very high frequency of tourists and passers-by.

2. Data

2.1 Data sources

The following criteria should be considered to answer the main questions:

- find proper neighborhood candidates near St. Stephen's Cathedral
- distance of those neighborhoods to that area of interest
- number of existing coffee shops in those neighborhoods

Following APIs were used:

Data source:	Purpose:
Foursquare API: "https://developer.foursquare.com/"	Get venues, categories i.e. coffee houses
Google Maps API reverse geocoding: "https://developers.google.com/maps/documentation/javascript/examples/geocoding-reverse"	Get centers of candidate areas that will be generated algorithmically and find the approximate addresses of centers of those areas

Google Maps API geocoding:"https://developers.google.com/maps/"	Get coordinates of St. Stephen's Cathedral and its surrounding

2.2 Data acquisition of proper neighborhoods

Started with obtaining necessary coordinates via Google Maps Geocoding specifically St. Stephen's Cathedral. Latitude and Longitude coordinates of potential neighborhoods were needed for the centroids. First approach was to use grid search. So, a layer of grid cells came into play. 10x10 kilometers centered around St. Stephen's should do the job in the beginning.

Next step was to create a grid of area candidates, equally spaced, centered around St. Stephen's Cathedral and within ~6km from it. Neighborhoods was defined as circular areas with a radius of 300 meters, so the neighborhood centers were 600 meters apart.

To accurately calculate distances a grid of locations in a cartesian 2D coordinate system was needed, which allowed the calculation of distances in meters (not in latitude/longitude degrees). Then it projected those coordinates back to latitude/longitude degrees to be shown on the Folium map. Further step was to write functions to convert between the WGS84 spherical coordinate system (latitude/longitude degrees) and UTM Cartesian coordinate system (X/Y coordinates in meters).

After that a hexagonal grid of cells was created (offset every other row, and adjusted vertical row spacing so that every cell center was equally distant from all its neighbors).

2.3 Visualize data on map

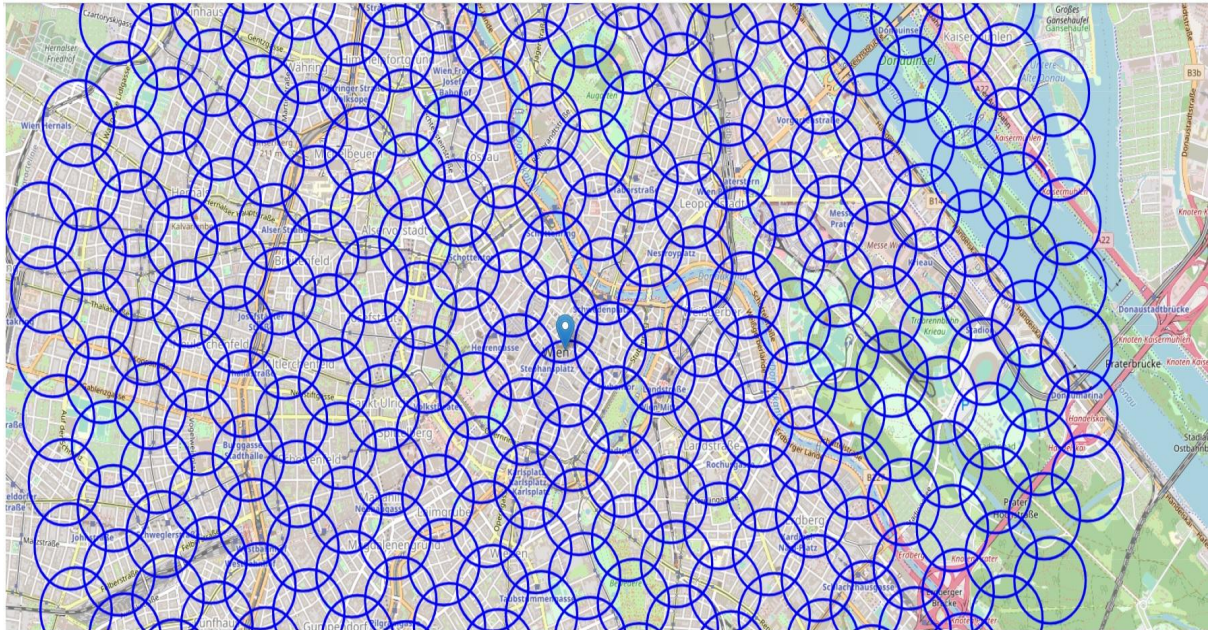Displaying city center location and candidate neighborhood centers:

Figure 1. Vienna Map with Grid

2.4 Create Data Frame out of addresses

After all addresses with their correlating distances were obtained a data frame was generated.

	Address	Latitude	Longitude	X	Y	Distance from center
0	Lienfeldergasse 37, 1160 Wien	48.216291	16.314260	5.608245e+06	9.533059e+06	5992.495307
1	Thaliastraße 122, 1160 Wien	48.212275	16.314919	5.608845e+06	9.533059e+06	5840.376700
2	Zagorskigasse 5a, 1160 Wien	48.208259	16.315578	5.609445e+06	9.533059e+06	5747.173218
3	Possingergasse 15, 1150 Wien	48.204243	16.316237	5.610045e+06	9.533059e+06	5715.767665
4	Schanzstraße 14, 1150 Wien	48.200227	16.316896	5.610645e+06	9.533059e+06	5747.173218
5	Märzstraße 99, 1150 Wien	48.196212	16.317554	5.611245e+06	9.533059e+06	5840.376700
6	Felberstraße 25, 1150 Wien	48.192197	16.318212	5.611845e+06	9.533059e+06	5992.495307
7	Hernalser Hauptstraße 160, 1170 Wien	48.222697	16.318478	5.607345e+06	9.533579e+06	5855.766389
8	Effingergasse 6, 1160 Wien	48.218681	16.319136	5.607945e+06	9.533579e+06	5604.462508
9	Degengasse 17, 1160 Wien	48.214664	16.319794	5.608545e+06	9.533579e+06	5408.326913

Figure 2. Data Frame Addresses

2.5 Data acquisition of venues via Foursquare API

With the help of the Foursquare API all necessary venues containing coffee houses and similar could be gathered. So, a closer look at collected coffee houses in areas of interest was possible.
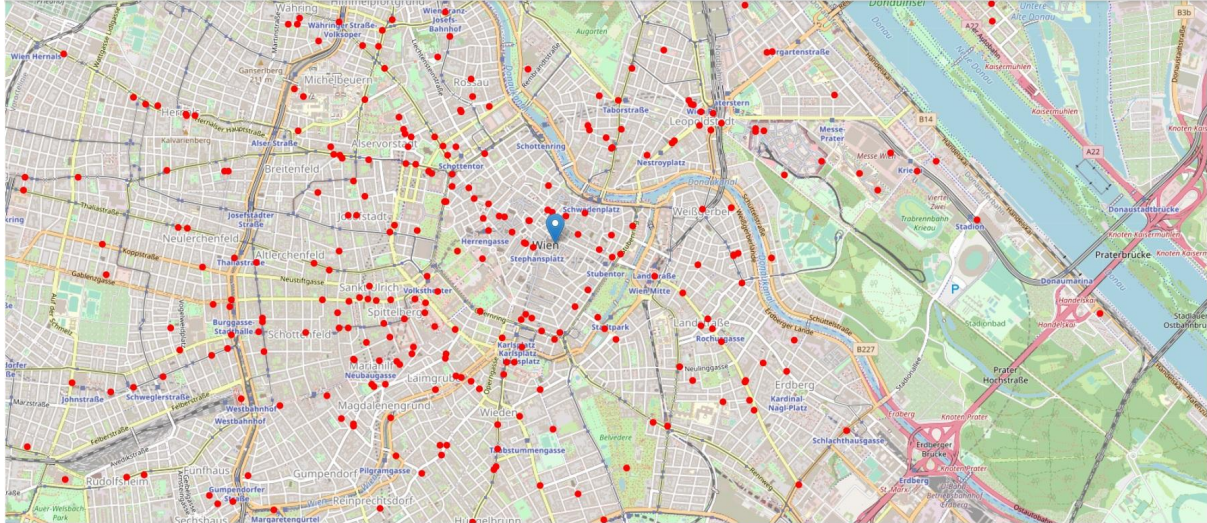

Figure 3. Map of Coffee Houses

3. Methodology

In order to find those promising neighborhoods with a low density of coffee houses (or no coffee shops at all) a heatmap in the beginning was the tool of choice. To keep it simple the area of inspection was limited within ~6km around St. Stephen's.

Next step included the use of clusters i.e. k-means clustering which was visualized on a map.

Started the exploratory analysis with counting coffee houses in every area candidate.

Next step was to calculate the distance to the nearest coffee house from every area candidate center (not only those within 300m - especially the distance to the closest one, regardless of how distant it was).

	Address	Latitude	Longitude	X	Y	Distance from center	Cafes in area	Distance to Coffee House
0	Lienfeldergasse 37, 1160 Wien	48.216291	16.314260	5.608245e+06	9.533059e+06	5992.495307	0	422.791942
1	Thaliastraße 122, 1160 Wien	48.212275	16.314919	5.608845e+06	9.533059e+06	5840.376700	2	151.701745
2	Zagorskigasse 5a, 1160 Wien	48.208259	16.315578	5.609445e+06	9.533059e+06	5747.173218	0	664.202150
3	Possingergasse 15, 1150 Wien	48.204243	16.316237	5.610045e+06	9.533059e+06	5715.767665	0	707.645372
4	Schanzstraße 14, 1150 Wien	48.200227	16.316896	5.610645e+06	9.533059e+06	5747.173218	1	386.037775
5	Märzstraße 99, 1150 Wien	48.196212	16.317554	5.611245e+06	9.533059e+06	5840.376700	1	298.772433
6	Felberstraße 25, 1150 Wien	48.192197	16.318212	5.611845e+06	9.533059e+06	5992.495307	1	217.043997
7	Hernalser Hauptstraße 160, 1170 Wien	48.222697	16.318478	5.607345e+06	9.533579e+06	5855.766389	1	54.630640
8	Effingergasse 6, 1160 Wien	48.218681	16.319136	5.607945e+06	9.533579e+06	5604.462508	0	609.972800
9	Degengasse 17, 1160 Wien	48.214664	16.319794	5.608545e+06	9.533579e+06	5408.326913	1	180.687513

Figure 4. Data Frame with Distances to nearby Coffee Houses

Observed that on average the next coffee house could be found within ~400m from every area center candidate. Next step showed those results i.e. the density on a heatmap.

Figure 5. Heatmap displaying Coffee House densities

3.2 Heatmap interpretation

It seemed obvious that around St. Stephen's was a high density of coffee houses. This was not surprising since it is in the heart of Vienna and many other historical sites and shopping possibilities can be found here. Also within 1km, which can be defined as Vienna's downtown, showed some significance in coffee houses density.

So, shifting focus to another promising area within a 3km range was interesting. My personal decision resulted in the neighborhood called 'Stuwerviertel'. It is located about 2,5km north-eastern from St. Stephen's. It is a high emerging, modern area with a lot of newly built office complexes, apartments and the University for Business and Economics is also very near. Not forgetting to mention its close range to Vienna's famous theme park 'Praters'. All in all a very promising spot with potential coffee consuming customers and very low density of coffee houses.
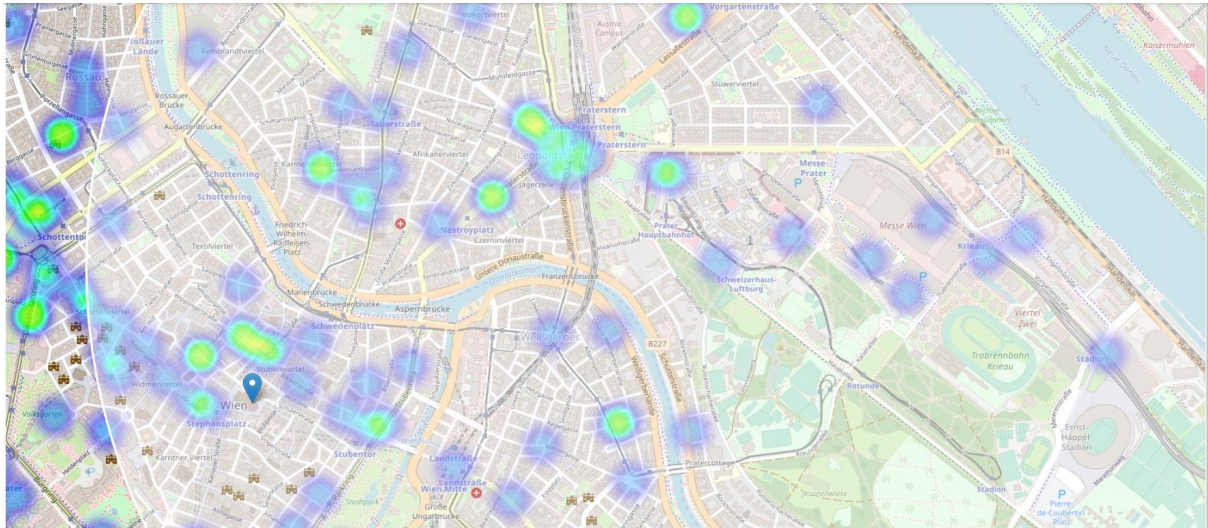
3.3 Focus: 'Stuwerviertel'

Figure 6. Stuwerviertel

After a new area of interest was chosen, a more narrow grid for 'Stuwerviertel' was created to find potential spots.Then distances (radius=250 meters) to closest coffee shops in vicinity were calculated.

Final step was filtering out locations with no more than two coffee houses within a range of 250 meters:
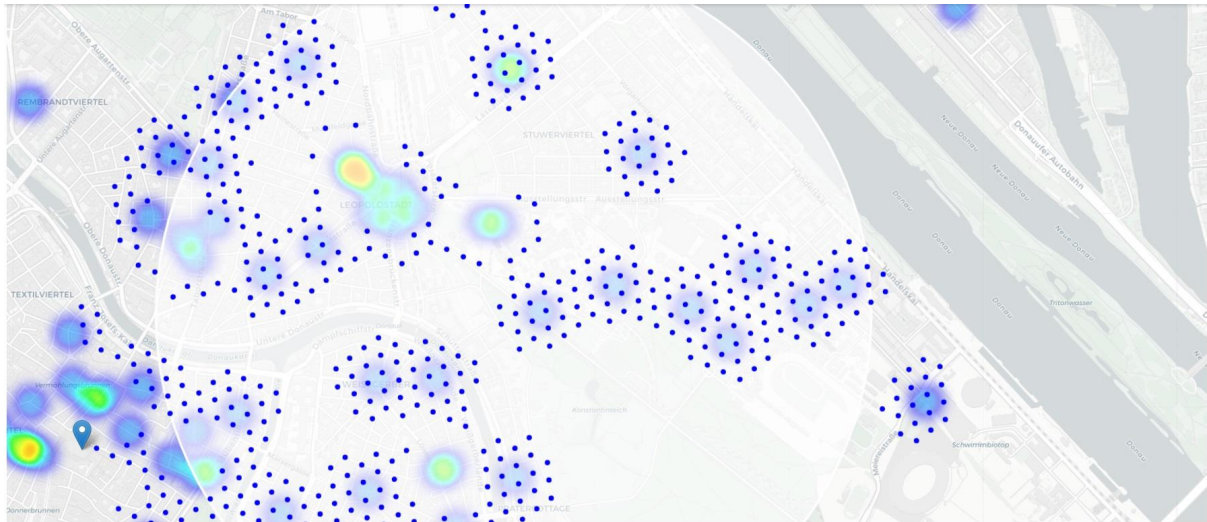
Figure 7. Map of Potential Spots in Stuwerviertel

3.4 Map interpretation

All blue markers show potential locations for a new coffee house. They meet the specified parameters:

- no more than 2 coffee houses
- within 250m

Calculated results:

- Locations with no more than two Cafes nearby: 2075
- Locations with no Cafes within 250m: 889
- Locations with both conditions met: 703

The following heatmap clarify those results:

Figure 8. Heatmap of Stuwerviertel

3.5 Cluster Results

With the use of a clustering method those promising areas were stressed out a bit more. After those zones were defined, their centers and addresses should lead to the final result of the analysis.
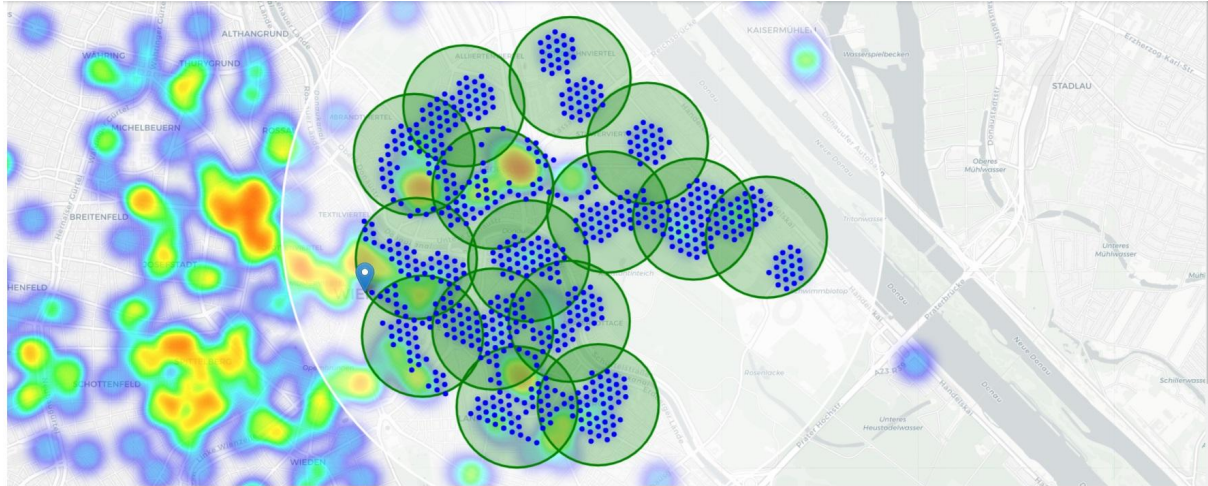

Figure 9. Clusters in Stuwerviertel

3.6 Cluster Interpretation

The clusters (green circles) represent groupings of the most promising locations.

Addresses of those cluster centers will be a good starting point for further

exploration in those neighborhoods based on specified requirements.

Figure 10. Map of clusters without Heatmap

3.7 Reverse Geocoding

To find the addresses of those potential zones the coordinates had to be reverse geocoded which represents the final results.

3.8 End of Analysis

This concludes the analysis. 15 addresses representing centers of zones containing locations with a low number of Coffee Houses nearby, all zones being fairly close to St. Stephen's Cathedral (all less than 5km from this location). Although zones are shown on map with a radius of ~500 meters (green circles), their shape is actually very irregular and their centers/addresses should be considered only as a starting point for exploring area neighborhoods in search for potential coffee house locations. Most of the zones are located in 'Stuwerviertel' borough, which is identified to be interesting due to being popular with tourists, fairly close to St. Stephen's and well connected by public transport.
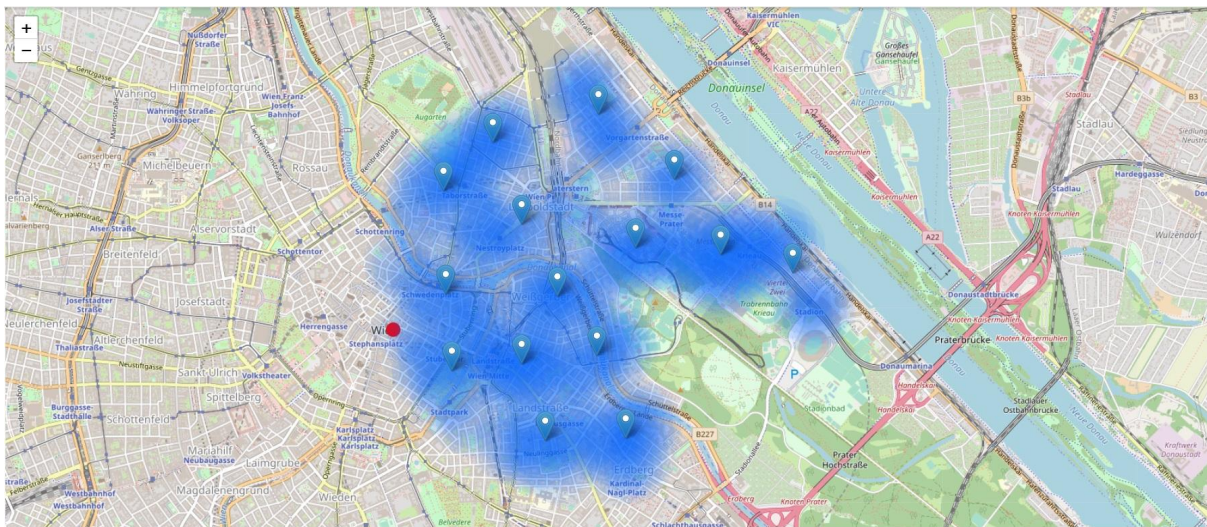

Figure 11. Map with Canditate Areas

4. Results and Discussion

Summarizing all the findings it led to focus on a new area of interest which is 'Stuwerviertel' and its surroundings. Started at Vienna's most centered and important historical site, St. Stephen's Cathedral to locate areas with high densities of coffee shops which could be excluded and identify spots with low density. One of those zones was finally 'Stuwerviertel'.

After calculating specified distances and finding nearby coffee houses it still meets defined requirements. With the help of clustering potential zones and addresses could be obtained.

5. Conclusion

It should be mentioned that there are other potential candidate areas. Shifting focus to that specific region 'Stuwerviertel' was a personal decision based on current observations and knowledge. A lot of factors speak for the attractiveness of that upcoming and trendy area which offers a lot of business possibilities in the end. For a more accurate analysis of course, different key indicators would have to be taken in consideration like rent-prices, available rent objects, tourist-frequency, number of offices and so on. This project delivers a first overview of potential locations to offer initial recommendations for the stakeholders.