

The Battle of Neighborhoods

Applied Data Science Capstone Project

by Maximilian

1) Introduction

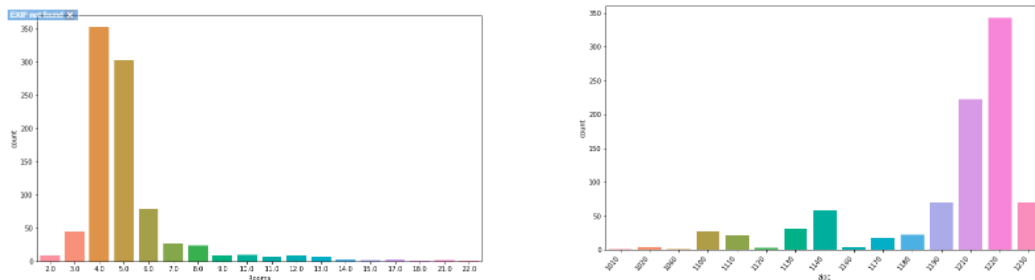
this project is about the fact that I have to move to Vienna professionally. My wife would like to buy a house with me, which should also contain her psychological practice. Since she does not yet have a client base in Vienna, it makes sense that the practice is located near popular and frequently visited areas.

It's important for us to know how:

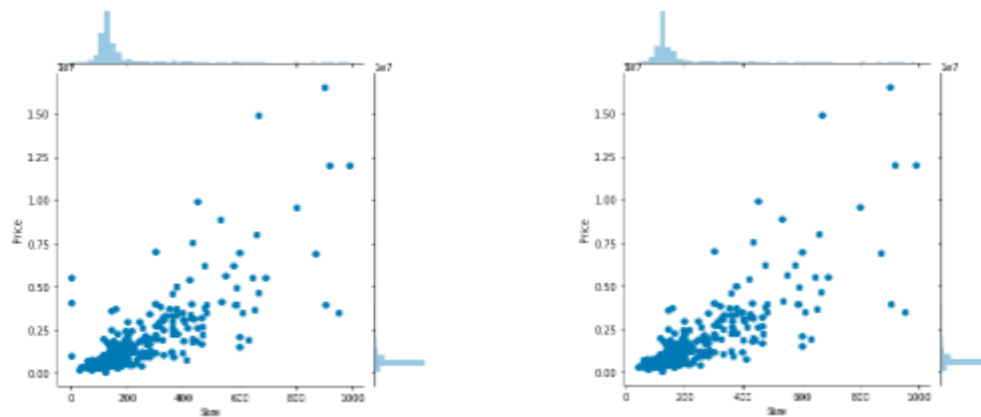
- the neighborhood is,
- to get a good price/m²,
- how many rooms the apartment has,
- how many square meters the apartment has. and also
- that there are many restaurants nearby, as this speaks for a visited area and the patients do not have to drive far.

2) Data collecting:

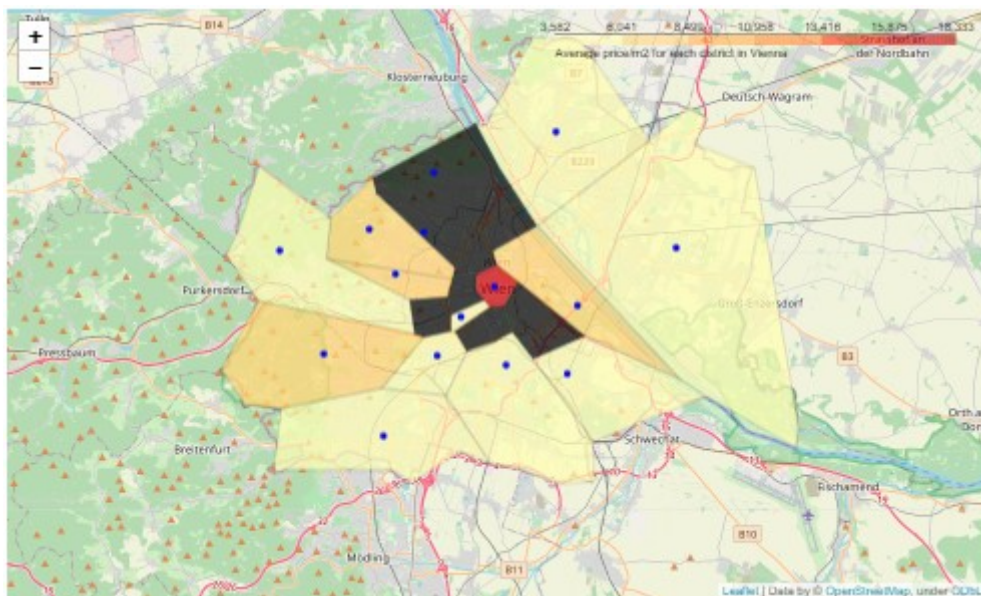
For the collection of real data a real housing search site was used. Here the pages were searched for all houses and implausible information/offers were removed. It is very pleasing in Vienna that the postcodes correspond to the areas, for this reason the collected data could be subsequently merged with the urban areas. Also characteristic values were generated, such as price/m² to compare the individual houses on offer. Histograms were used to interpret the data as shown in Figure 1.



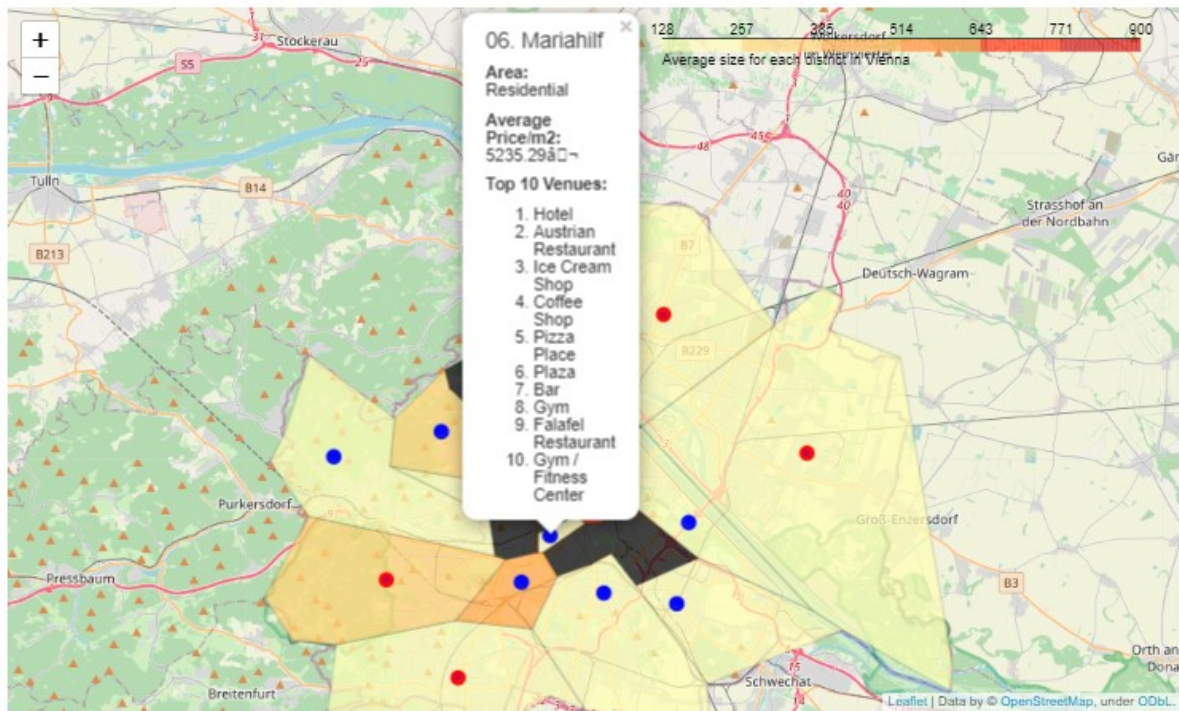
Seaborn was also used to identify outliers in the data. Non-logical data, such as houses with a square meter count of 0, were found.



As shown in Fig 3., Folium was used to visualize these data in a map, like the price per m^2



Afterwards Foursquare was used to explore the region and make clustering. With all the data collected, we can now create a map that shows the average price/ m^2 for each district and displays on the marker labels information about the type of area and the 10 best locations for each district.



As can be seen, a region similar to the centre is available in the district. We have bought a house in this region and are lucky to have had so much luck with the help of this course. Many of our neighbours say that this knowledge would have been more of an inside knowledge.