

Where to live in Vienna?

-----Comparison between Vienna neighborhoods

1. Introduction

1.1 Background

Each year, the international consulting firm Mercer carries out a study to assess the quality of life in 231 cities around the world. In 2019 the Austrian federal capital was awarded the most livable city in the world for the tenth time in a row.

Although Vienna is the most livable city in the world, people still would like to know more about which district in the city would suit them more, before moving to Vienna. For example, there are so many factors affect people to make decisions to start a new life in a city, such as house price, school rating, crime rating, neighborhood analysis etc. That calls for a search algorithm which normally gives the requested features.

1.2 Aim

This project is to help stakeholders or anyone who wants to explore Vienna, in order to make a better decision which district is the best area to start a new life.

There are in total 23 districts in Vienna. The rental cost is quite different from one district to another. For example, the 1st district locates in the city center, which is of course the most convenient place to live. However, the renting price must be taken into consideration. According to the information we scraped from internet, the 1st district is the most expensive district.

So the purpose of this project is to help stakeholders to find a district to rent an apartment, with a relatively cheaper rental cost and relative convenient living facilities such as cafes, restaurants, bakeries, gyms etc.

2. Data Source

2.1 Vienna rental price:

In vienna, there are 23 districts in total. In order to find a balance between cheaper rental price and more convenient living facilities, we have to know the rental cost of each district. The link below provides the rental price of apartment with difference sizes. In this project, we only consider the average rental price (AVERAGE/M²). However, the algorithm used in this project also can be applied to small (<50 M²)/medium (50-129 M²)/large (>130M²) apartments.

<https://www.virtualvienna.net/moving-to-vienna/accommodation/real-estate-rental-prices/>

2.2 Vienna neighborhood data:

Vienna neighborhood data including the coordinates of each neighborhood can be download via

<https://data.wien.gv.at/daten/geo?service=WFS&request=GetFeature&version=1.1.0&typeName=ogdwien:GEONAMENOGD&srsName=EPSG:4326&outputFormat=csv>

2.3 Four-square API

Four-square API would be used in this project as the prime data gathering source, which provides the ability to perform location search, location sharing and details about a business. Photos, tips and reviews.

3. Methodology

3.1 Data wrangling

Libraries Pandas and numpy are used for data wrangling.

3.1.1 dataset of Vienna rental price

- the postcode of each Vienna district needs to be added into the dataframe in order to read the data more clearly.

- In this project, we only consider the average rental price (AVERAGE/M²). So the other columns with the price of different apartment size are dropped.

3.1.2 dataset of Vienna neighborhood

- columns ‘SHAPE’, ‘FEATURENAME’, ‘LOCALITYLIST’ will be kept from the original data.
- ‘SHAPE’ contains the coordinates of each neighborhood. So the latitude and longitude need to be extracted from ‘SHAPE’ and make two new columns ‘Latitude’ and ‘Longitude’.
- ‘LOCALITYLIST’ will be renamed to ‘Postcode’ and convert to the same format as the ‘Postcode’ column in Vienna rental price dataset.
- Some neighborhood are shared with more than 1 district, thus these neighborhoods has more than one postcode. These data will be dropped. We only consider the neighborhood uniquely belongs to one district.

3.1.3 Merge Vienna rental price dataset and Vienna neighborhood dataset based on the shared column ‘Postcode’.

3.1.4 There are more than 4000 rows after merging the two datasets. For simplifying, 5 data points (neighborhoods) in each district are randomly selected to make a new dataframe.

3.2 Visualization

- Library seaborn is used for data visualization.
- Geopy library is used to get the latitude and longitude values of Vienna city.
- Folium library is used to visualize the neighborhoods of Vienna

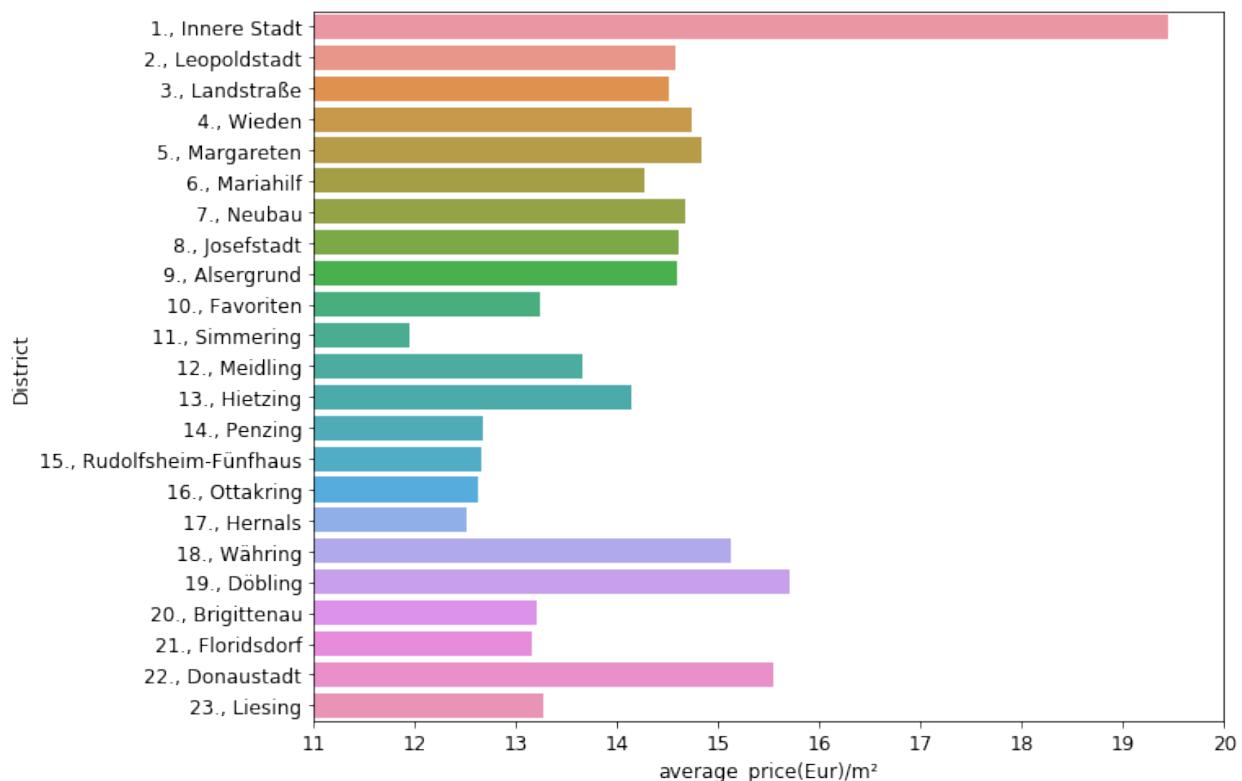
3.3 Clustering

- Neighborhoods need to be clustered in this project. However, neighborhoods are categorical data. So one-hot encoding on categorical variables is being used to convert categorical variables to numerical variables.
- Library Scikit learn is used for k means clustering.

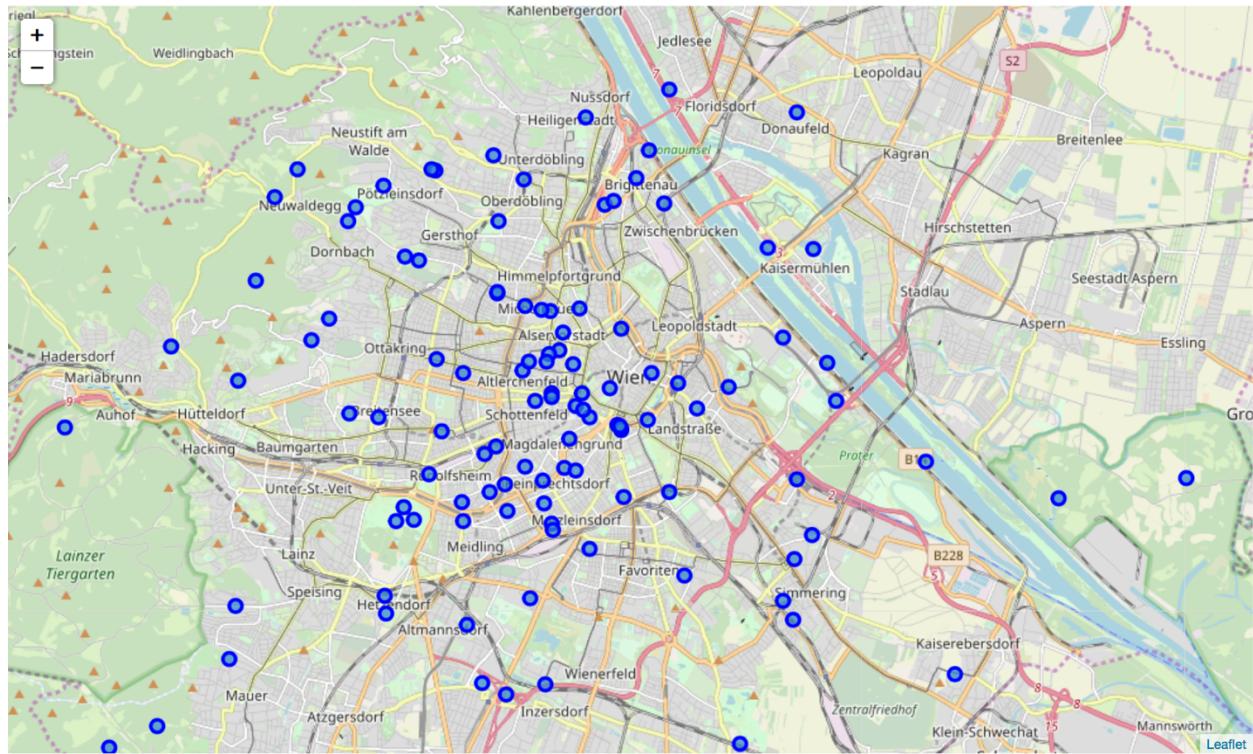
4. Results

4.1 Renting price is difference from district to district

Vienna rental price has a huge difference among districts. As shown in the barplot below, the first district Innere Stadt has the highest rental price with almost 20 euro per M². However, the lowest rental price is in the 11th district Simmering, which only cost around 12 euro per M². This makes sense since Innere Stadt has more convenient living facilities.



4.2 Five representative neighborhoods of each district visualization



4.3 The Top10 venues of each neighborhood

The file of ‘top10 venues of each neighborhood’ can be seen in the jupyter notebook. Here only one screen shot is presented.

	PostalCode	District	average/m ²	Neighborhood	Longitude	Latitude	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
0	1010	1., Innere Stadt	19.44	Hofburg	16.364944	48.206558	8	Café	Plaza	Restaurant	Hotel	Museum	Boutique
1	1010	1., Innere Stadt	19.44	Grünanlage Dr.-Karl-Lueger-Platz	16.380016	48.207450	8	Austrian Restaurant	Restaurant	Hotel	Italian Restaurant	Café	
2	1010	1., Innere Stadt	19.44	U-Bahn Station Stubentor	16.379991	48.206918	8	Austrian Restaurant	Restaurant	Italian Restaurant	Hotel	Bar	
3	1010	1., Innere Stadt	19.44	Weltmuseum Wien	16.363530	48.204799	8	Plaza	Restaurant	Art Museum	Café	Hotel	Museum
4	1010	1., Innere Stadt	19.44	Kartensammlung	16.367352	48.206001	8	Plaza	Café	Restaurant	Hotel	Italian Restaurant	Historic
5	1020	2., Leopoldstadt	14.57	Sportplatz Cricket	16.416773	48.206231	5	Hot Dog Joint	Other Event	Bus Stop	Soccer Stadium	Soccer Field	Racecourse

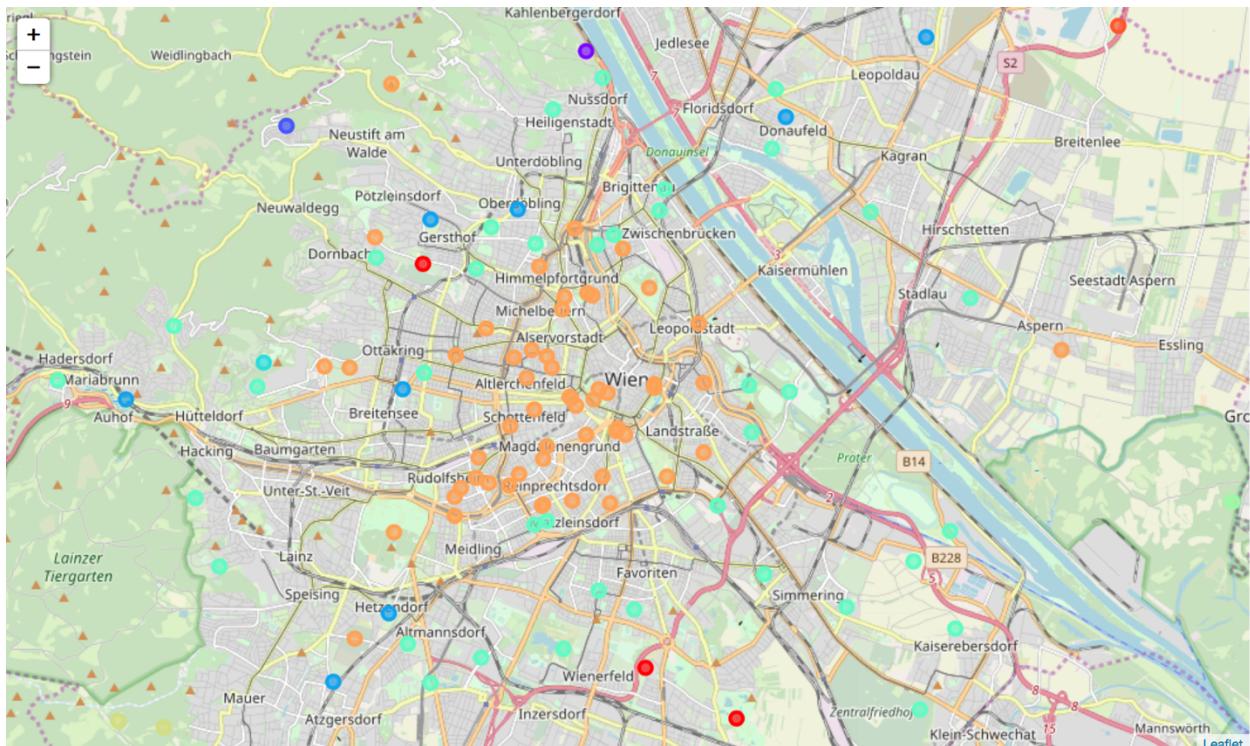
4.4 Clustering Vienna neighborhoods

In total 10 clusters are made from total Vienna neighborhoods.

Each cluster contains the a few neighborhoods shown in below:

Cluster Labels	Num_of_each_neighborhoods
0	3
1	1
2	1
3	10
4	1
5	38
6	1
7	2
8	55
9	1

4.5 Visualization of Vienna neighborhoods cluster



5.Discussions

In this project, we clustered the Vienna neighborhoods in order to find similar or more convenient living facility with relative cheaper price. Depends on which neighborhoods feature the stakeholders care about, they can easily focus on the cluster which they are interested in. Then it is more easier to make a decision according to rental price where to start a few life in Vienna.

6.Conclusions

A algorism of clustering Vienna neighborhood has been made in order to find a compromise between rental price and living convenience.