Apartment Rentals in Vienna: Prices & Venues Analysis

Ekaterina Bubenko

May 8th 2020

1. Introduction

1.1. Background Discussion

The Economist Intelligence Unit's (EIU) "Global Livability Report"[1] ranks the best places to live in the world, scoring 140 cities in five categories:

1. Stability
2. Healthcare
3. Culture & Environment
4. Education
5. Infrastructure

In the article called "These Are the 10 Best Places to Live in the World"[2] updated on April 23, 2020 the author, Cailin Morton, uses the EIU ranking to identify top 10 places to live in the world.

According to the article #1 city to live in is Vienna (Austria): "*Maintaining its position in the number one spot for the second year in a row, Vienna provides the perfect blend of adventure, affordability, ease of living, and safety. The Austrian capital draws crowds with its Baroque architecture and Danube cruises, but it's just as exciting as a center of what's new in the culinary and art worlds of Europe. And due to a city government that puts a lot of resources behind infrastructure and housing, the cost of living here is far lower than comparable European cities.*"

Overall rating: 99.1

1. Stability: 100
2. Healthcare: 100
3. Culture & Environment: 96.3
4. Education: 100
5. Infrastructure: 100

---

[1] https://www.eiu.com/n/
[2] https://www.cntraveler.com/gallery/the-most-livable-cities-in-the-world

1.2. Problem Description & Target Audience

Since Vienna is considered to be the best place to live many people are interested in moving to Vienna either permanently or temporarily. So in this project we will help people who want to be looking for renting apartment in Vienna. They will be able to see:

- How districts compare in prices for rent (where is more expensive, where is less)
- Compare districts by their type (either commercial or residential)

Additionally, for those who already live in Vienna they will be able to see:

- If they're paying above or below the average market price for their apartment
- If there are similar districts to theirs with lower rent prices

2. Data Acquisition, Cleaning & Preparation

2.1. Data Sources & Acquisition

To consider the problem we used the following data:

- To collect data about apartments for rent in Vienna (size, number of rooms, address, and price) we used **BeautifulSoup** library to scrap data from a local website with apartment listings willhaben.at;
- To obtain coordinates for districts we used **geopy** to get location data;
- To gather information regarding venues (supermarket, restaurant, park, etc.) in districts we used **Foursquare API**;
- To build choropleth maps we used JSON file for geo-data for districts in Vienna available on https://github.com/ginseng666/GeoJSON-TopoJSON-Austria

2.2. Data Cleaning & Preparation

Once we scraped a local website with apartment listings willhaben.at we created a dataframe to contain the data:

	PostalCode	District	Size	Rooms	Price
0	None	None	82	3	980,-
1	1100	Favoriten	110	3	990,56
2	1010	Innere Stadt	237	4	4.506,82
3	1040	Wieden	257	7	3.799,25
4	1100	Favoriten	63	2	754,70

The number apartments in the dataframe was 9,834.

As you can wee in order to use the data we needed to clean it and ensure that datatypes are correct for all of the columns. Once we did that we added a column that calculated relative price (Price/m2) in order to have a fair comparison later on. The resulted dataframe:

	PostalCode	District	Size	Rooms	Price	Price/m2
0	None	None	82.0	3.0	980.00	11.95
1	1100	Favoriten	110.0	3.0	990.56	9.01
2	1010	Innere Stadt	237.0	4.0	4506.82	19.02
3	1040	Wieden	257.0	7.0	3799.25	14.78
4	1100	Favoriten	63.0	2.0	754.70	11.98

Then we needed to check the dataframe for missing values and identify possible outliers. Below are descriptive statistics for the dataframe:

	Size	Rooms	Price	Price/m2
count	9812.000000	9124.000000	9800.000000	9794.000000
mean	82.042805	2.691144	1332.817451	16.905968
std	44.523381	3.436967	2956.034704	44.929006
min	1.000000	1.000000	30.000000	0.670000
25%	52.000000	2.000000	799.000000	13.410000
50%	70.000000	2.000000	999.000000	15.500000
75%	99.000000	3.000000	1524.867500	17.700000
max	664.000000	312.000000	280000.000000	4179.100000

As we can, some data is missing. We removed rows with no data for all columns except for the column 'Room' since it has too many missing values. The updated number apartments in the dataframe was 9,791.

Then we checked how many apartments per number of rooms were in the dataframe and saw that the majority of apartments have 1-5 rooms. Therefore, we removed apartments that have more than that (as they are also considered luxury) from the dataset and the updated number apartments in the dataframe was 8,967.

In order to acquire coordinates for districts for further visualization we identified unique values for 'Postal Codes' (i.e. districts) in the dataframe and then used geopy to collect geolocation data:

|	PostalCode	District	Latitude	Longitude
0	1010	Innere Stadt	48.209023	16.369851
1	1020	Leopoldstadt	48.200638	16.426895
2	1030	Landstraße	48.193644	16.396286
3	1040	Wieden	48.208354	16.372504
4	1050	Margareten	48.188073	16.353386
5	1060	Mariahilf	48.195475	16.347023
6	1070	Neubau	48.202264	16.349123
7	1080	Josefstadt	48.210852	16.347360
8	1090	Alsergrund	48.225073	16.358398
9	1100	Favoriten	48.152994	16.382822
10	1110	Simmering	48.163109	16.458009
11	1120	Meidling	48.171965	16.328687
12	1130	Hietzing	48.178541	16.252986
13	1140	Penzing	48.225662	16.222790
14	1150	Rudolfsheim-Fünfhaus	48.195475	16.326301
15	1160	Penzing	48.225662	16.222790
16	1160	Ottakring	48.214955	16.302153
17	1170	Hernals	48.235403	16.284214
18	1180	Währing	48.234115	16.321606
19	1190	Döbling	48.261251	16.328471
20	1200	Brigittenau	48.243822	16.378147
21	1210	Floridsdorf	48.279815	16.412135
22	1220	Donaustadt	48.226966	16.494484
23	1230	Liesing	48.141106	16.293912

3. Methodology

3.1. Exploratory Data Analysis

To understand distribution of apartments by the number of rooms following plot:
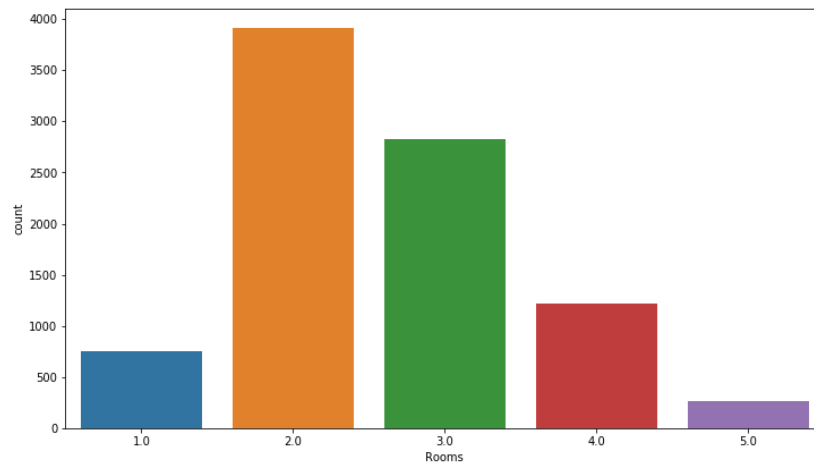

We saw that the majority of apartments in the listings have 2-3 rooms while the least frequent are listings for 5-room apartments. Below you can see the plot with distribution of apartments by postal codes (i.e. districts):
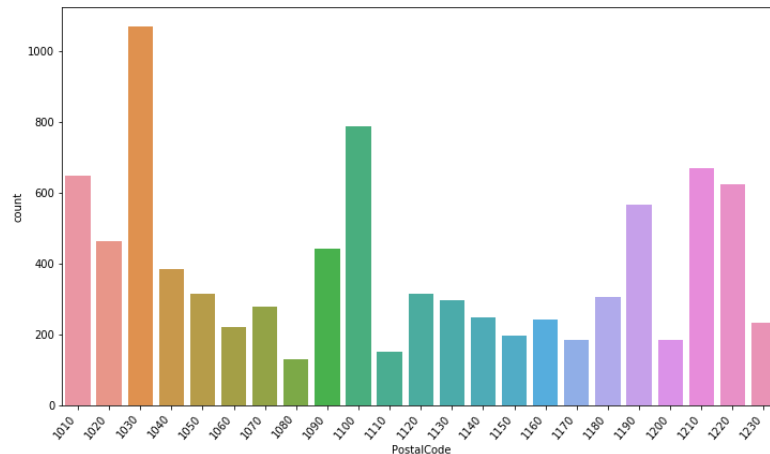
The plot showed that listings for 1030 (Landstraße) district are the most common followed by 1100 (Favoriten) while the least frequent listings are for 1080 (Josefstadt) district.

In order to identify outliers for apartment price and size used scatter plots. We saw that we have some outliers, but we did not have many of them (only 3 records). After removed them the scatter plot significantly improved with the new dataframe size of 8,864 apartments:
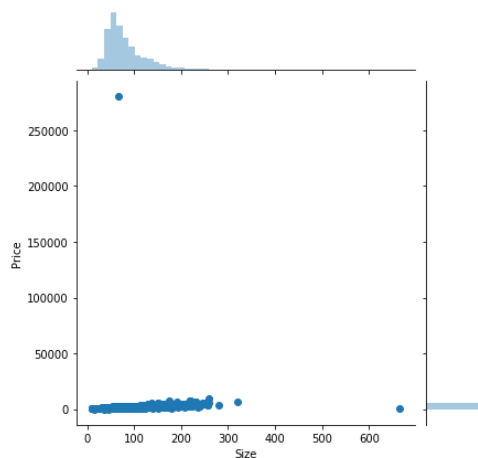

Scatter plot before the removal of outliers
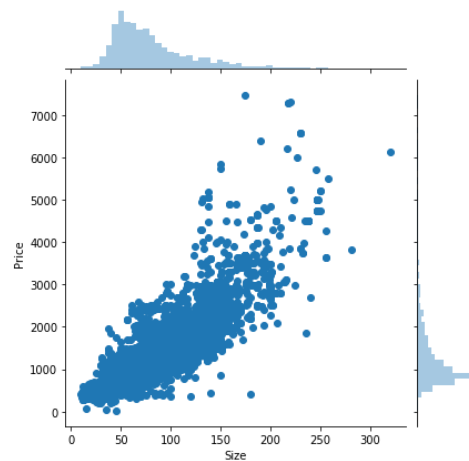
Scatter plot after the removal of outliers

To further improve our understanding of the data we wanted to check if our intuition about the relationships between the price of apartment and number of rooms is correct. We expected to see direct proportion between the absolute apartment price and the number of rooms, but inverse proportion between the relative prices. For that we created box plots:
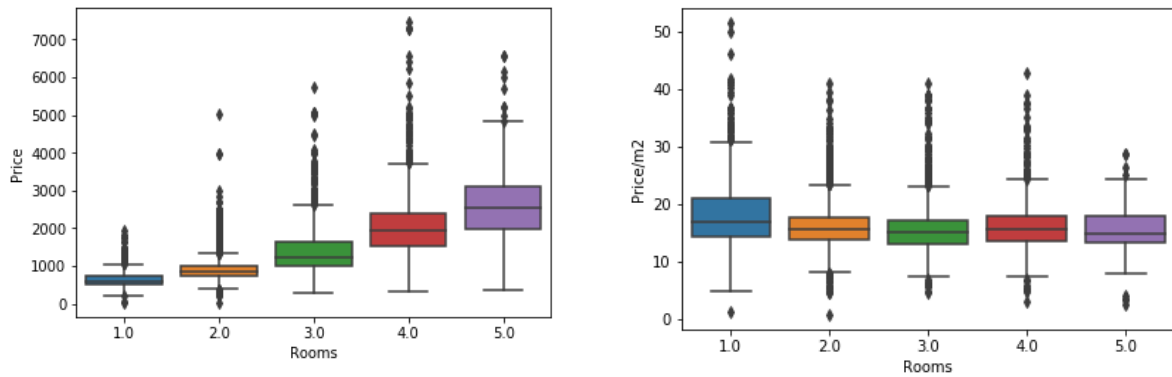
The plot on the left confirmed our intuition regarding the absolute price, however the plot on the right showed that the inverse proportion is not as strong as we expected. Yet indeed 1-room apartments are more expensive per m^2 than apartments with more rooms.

We were also curious to see price distributions per postal codes (i.e. districts) and learned that 1010 (Innere Stadt) was the most expensive district followed by 1190 (Döbling) then by 1090 (Alsergrund) and that there were lots of outliers for luxurious apartments in the majority of districts:
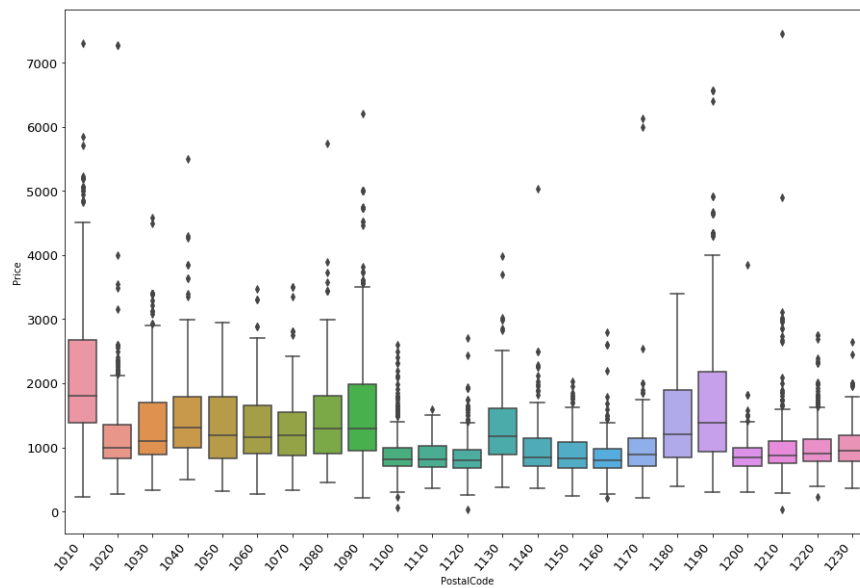
We then checked whether the data for geocoordinates we acquired was good using map visualization:
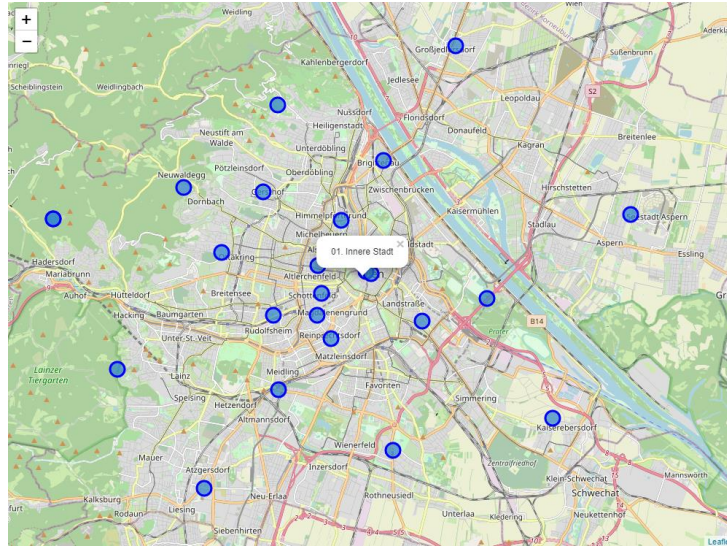
Once we saw that coordinates are ok we went ahead with preparing average metrics for the further analysis of Vienna's districts:

	District	Size	Rooms	Price	Price/m2
0	Alsergrund	95.761905	2.759637	1610.149297	17.065805
1	Brigittenau	66.064516	2.263441	887.505054	13.963871
2	Donaustadt	63.229167	2.375000	984.550513	16.150128
3	Döbling	97.458554	2.941799	1695.753810	17.019965
4	Favoriten	59.193916	2.302915	885.269037	15.521825
5	Floridsdorf	67.536622	2.500747	969.630807	14.663259
6	Hernals	69.837838	2.410811	1002.018378	14.532162
7	Hietzing	87.791946	2.721477	1315.067248	15.283490
8	Innere Stadt	108.176198	2.806801	2129.883076	20.018594
9	Josefstadt	94.046154	2.800000	1479.580154	15.762077
10	Landstraße	80.458879	2.689720	1312.083336	16.712196
11	Leopoldstadt	76.550538	2.516129	1200.218344	16.128409
12	Liesing	72.431034	2.629310	1016.273836	14.212155
13	Margareten	90.980952	2.831746	1392.930063	15.496857
14	Mariahilf	86.677273	2.709091	1352.071500	15.899273
15	Meidling	61.584127	2.247619	862.789810	14.639492
16	Neubau	82.118705	2.525180	1259.179388	15.838597
17	Ottakring	61.780083	2.248963	855.672490	14.516017
18	Penzing	71.888000	2.516000	971.758080	13.947560
19	Rudolfsheim-Fünfhaus	63.604061	2.340102	904.361066	14.829543
20	Simmering	68.697368	2.486842	875.892500	13.319934
21	Wieden	95.932642	2.769430	1520.269378	16.000415
22	Währing	90.058632	2.859935	1375.726156	15.361205

3.2.   Average Apartment Price & Size per District

We noticed that regarding the absolute apartment prices majority of districts were very comparable having average price up to 1,000 EUR.:
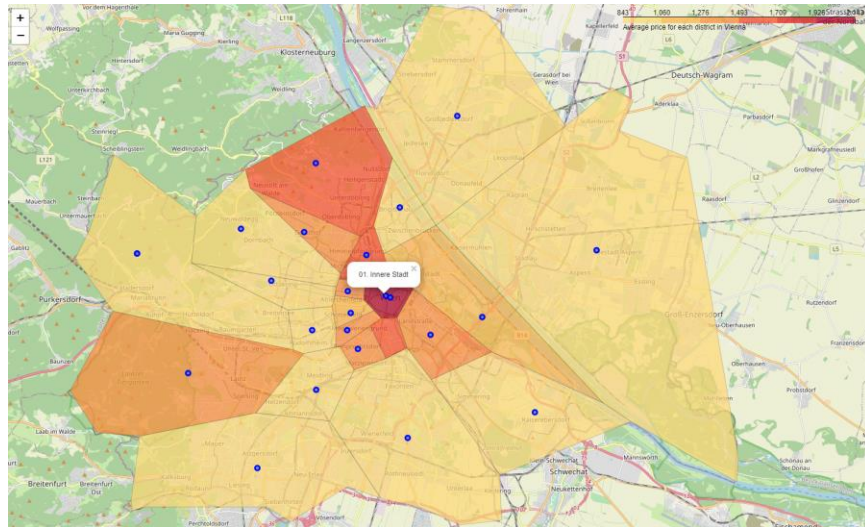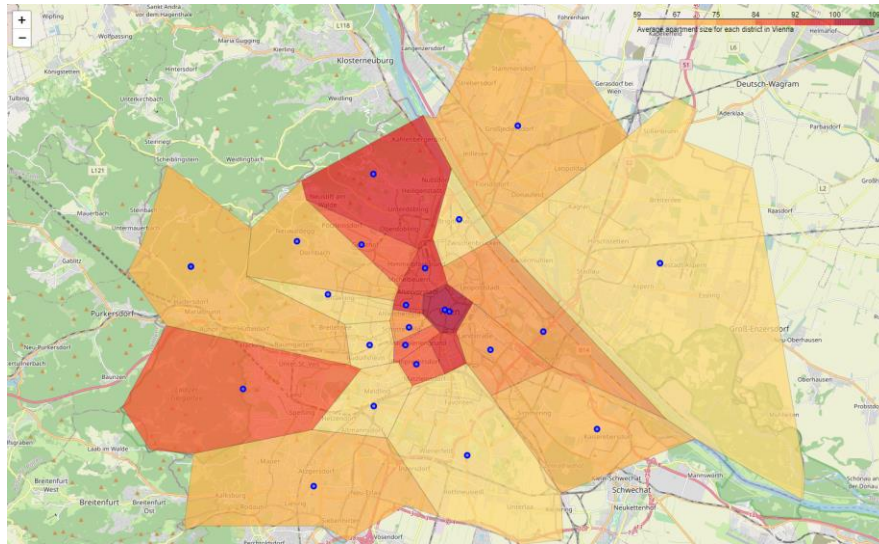

When it comes to the average relative price (price per m^2) per district we got more interesting picture:
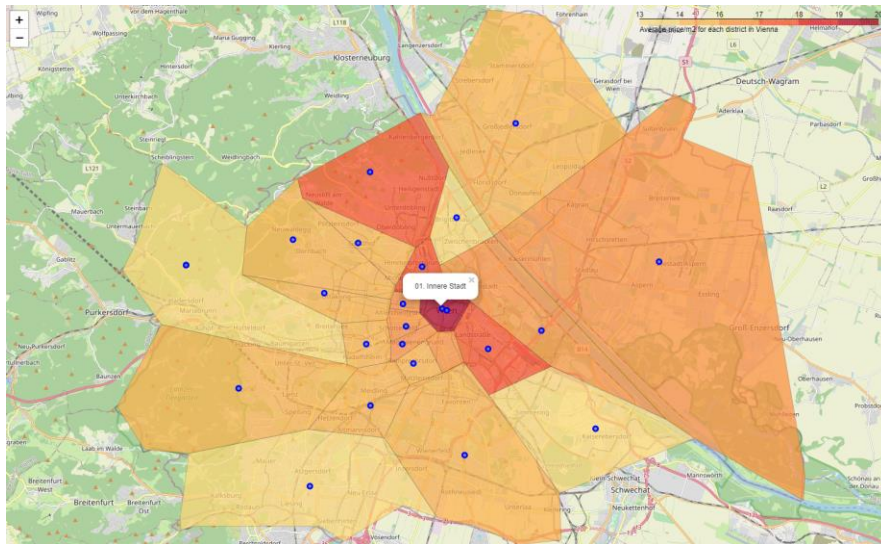
We saw that there are 4 districts that are more expensive that others (18-20 EUR per m^2) and 4 that are less expensive than others (under 14 EUR per m^2).

In terms of the average apartment size per district, the map also showed curious results:

We noticed that bigger apartments were mainly concentrated in the city center (84-109 m^2), yet Hietzing district on the west side of the city also has quite large apartments (avg. 84-92 m^2).

3.3.   Most Common Venues in Districts

We utilized Foursquare API to explore and segment districts and set the limit as 100 venues and the radius of 500 meters for each district. We collected 1,659 venues for 211 unique venue categories:
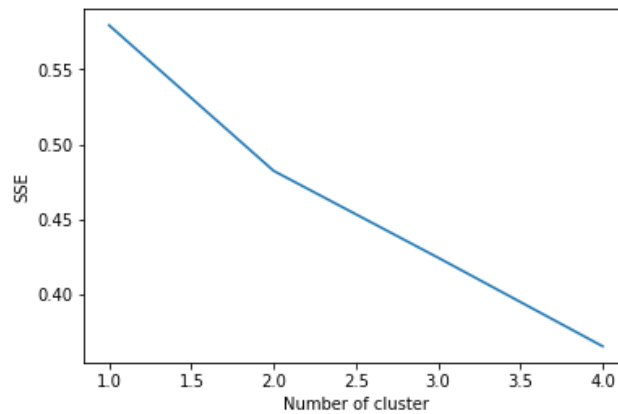
	District	District Latitude	District Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Innere Stadt	48.209023	16.369851	Stephansplatz	48.208299	16.371880	Plaza
1	Innere Stadt	48.209023	16.369851	Stephansdom	48.208626	16.372672	Church
2	Innere Stadt	48.209023	16.369851	Michaelerplatz	48.208007	16.366625	Plaza
3	Innere Stadt	48.209023	16.369851	Graben	48.208915	16.369379	Pedestrian Plaza
4	Innere Stadt	48.209023	16.369851	Park Hyatt Vienna	48.210691	16.367788	Hotel

Then we identified the 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
0	Alsergrund	Restaurant	Café	Park	Tapas Restaurant	Cocktail Bar	Coffee Shop	Plaza	Italian Restaurant	Austrian Restaurant	
1	Brigittenau	Supermarket	Plaza	Park	Austrian Restaurant	Italian Restaurant	Restaurant	Kebab Restaurant	Ice Cream Shop	Gym / Fitness Center	Fast Food Restaurant
2	Donaustadt	Fast Food Restaurant	Metro Station	Restaurant	Indian Restaurant	Convenience Store	Soccer Field	Café	Bus Stop	Pizza Place	Breakfast Spot
3	Döbling	Café	Austrian Restaurant	Restaurant	Wine Bar	Italian Restaurant	Bus Stop	Scenic Lookout	Bakery	Mountain	Vineyard
4	Favoriten	Austrian Restaurant	Supermarket	Gym / Fitness Center	Park	Grocery Store	Soccer Field	Bus Stop	Café	Train Station	Fast Food Restaurant
5	Floridsdorf	Supermarket	Shopping Mall	Electronics Store	Fast Food Restaurant	Restaurant	Tram Station	Grocery Store	Gym / Fitness Center	Café	Furniture / Home Store
6	Hernals	Austrian Restaurant	Park	Gastropub	Bus Stop	Hotel	Gluten-free Restaurant	Comfort Food Restaurant	Scenic Lookout	Restaurant	Beer Garden
7	Hietzing	Park	Supermarket	Hotel	Austrian Restaurant	Bakery	Scenic Lookout	Historic Site	Asian Restaurant	Food	Pizza Place
8	Innere Stadt	Hotel	Plaza	Austrian Restaurant	Café	Ice Cream Shop	Restaurant	Park	Japanese Restaurant	Art Museum	Pedestrian Plaza
9	Josefstadt	Coffee Shop	Plaza	Hotel	Park	Tapas Restaurant	Ice Cream Shop	Italian Restaurant	Café	Pizza Place	Gym
10	Landstraße	Hotel	Restaurant	Austrian Restaurant	Concert Hall	Asian Restaurant	Italian Restaurant	Café	Grocery Store	Bakery	Palace
11	Leopoldstadt	Restaurant	Supermarket	Hotel	Café	Gym / Fitness Center	Plaza	Seafood Restaurant	Sports Club	Beer Garden	Bar
12	Liesing	Supermarket	Hotel	Italian Restaurant	Bus Stop	Pharmacy	Plaza	Bakery	Food	Breakfast Spot	Furniture / Home Store
13	Margareten	Austrian Restaurant	Hotel	Plaza	Ice Cream Shop	Coffee Shop	Pizza Place	Restaurant	Burger Joint	Organic Grocery	Japanese Restaurant
14	Mariahilf	Hotel	Ice Cream Shop	Austrian Restaurant	Plaza	Coffee Shop	Restaurant	Pizza Place	Thai Restaurant	Japanese Restaurant	Hostel
15	Meidling	Hotel	Monument / Landmark	Pool	Café	Fountain	Austrian Restaurant	History Museum	Gastropub	Gift Shop	Indie Theater
16	Neubau	Hotel	Plaza	Coffee Shop	Ice Cream Shop	Tapas Restaurant	Art Museum	Gym	Austrian Restaurant	Pizza Place	Park
17	Ottakring	Supermarket	Restaurant	Austrian Restaurant	Café	Wine Bar	Gastropub	Hotel	Italian Restaurant	Tram Station	BBQ Joint
18	Penzing	Restaurant	Trail	Supermarket	Café	Electronics Store	Greek Restaurant	Smoke Shop	Ski Area	Multiplex	Market
19	Rudolfsheim-Fünfhaus	Austrian Restaurant	Hotel	Restaurant	Café	History Museum	Ice Cream Shop	Coffee Shop	Persian Restaurant	Tapas Restaurant	Italian Restaurant
20	Simmering	Tram Station	Flower Shop	Fast Food Restaurant	Restaurant	Café	Asian Restaurant	Supermarket	Grocery Store	Bus Stop	Shopping Mall
21	Wieden	Plaza	Hotel	Austrian Restaurant	Ice Cream Shop	Café	Restaurant	Concert Hall	Park	Art Museum	French Restaurant
22	Währing	Supermarket	Café	Park	Austrian Restaurant	Ice Cream Shop	Restaurant	Gastropub	Italian Restaurant	Plaza	Bakery

3.4. K-means Clustering for Partitioning Districts

First, we created Elbow Curve to determine the optimum number of clusters and found that it was 2:

After that we ran the algorithm to assign cluster labels to districts:

	PostalCode	District	Latitude	Longitude	Price/m2	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
0	1010	Innere Stadt	48.209023	16.369851	20.018594	1	Hotel	Plaza	Austrian Restaurant	Café	Ice Cream Shop	Restaurant	Park	Japanese Restaurant	Art Museum	Pedestrian Plaza
1	1020	Leopoldstadt	48.200638	16.426895	16.128409	0	Restaurant	Supermarket	Hotel	Café	Gym / Fitness Center	Plaza	Seafood Restaurant	Sports Club	Beer Garden	Bar
2	1030	Landstraße	48.193644	16.396286	16.712196	1	Hotel	Restaurant	Austrian Restaurant	Concert Hall	Asian Restaurant	Italian Restaurant	Café	Grocery Store	Bakery	Palace
3	1040	Wieden	48.208354	16.372504	16.000415	1	Plaza	Hotel	Austrian Restaurant	Ice Cream Shop	Café	Restaurant	Concert Hall	Park	Art Museum	French Restaurant
4	1050	Margareten	48.188073	16.353386	15.496857	1	Austrian Restaurant	Hotel	Plaza	Ice Cream Shop	Coffee Shop	Pizza Place	Restaurant	Burger Joint	Organic Grocery	Japanese Restaurant
5	1060	Mariahilf	48.195475	16.347023	15.899273	1	Hotel	Ice Cream Shop	Austrian Restaurant	Plaza	Coffee Shop	Restaurant	Pizza Place	Thai Restaurant	Japanese Restaurant	Hostel
6	1070	Neubau	48.202264	16.349123	15.838597	1	Hotel	Plaza	Coffee Shop	Ice Cream Shop	Tapas Restaurant	Art Museum	Gym	Austrian Restaurant	Pizza Place	Park
7	1080	Josefstadt	48.210852	16.347360	15.762077	1	Coffee Shop	Plaza	Hotel	Park	Tapas Restaurant	Ice Cream Shop	Italian Restaurant	Café	Pizza Place	Gym
8	1090	Alsergrund	48.225073	16.358398	17.065805	1	Restaurant	Café	Park	Hotel	Tapas Restaurant	Cocktail Bar	Coffee Shop	Plaza	Italian Restaurant	Austrian Restaurant
9	1100	Favoriten	48.152994	16.382822	15.521825	0	Austrian Restaurant	Supermarket	Gym / Fitness Center	Park	Grocery Store	Soccer Field	Bus Stop	Café	Train Station	Fast Food Restaurant
10	1110	Simmering	48.163109	16.458009	13.319934	0	Tram Station	Flower Shop	Fast Food Restaurant	Restaurant	Café	Asian Restaurant	Supermarket	Grocery Store	Bus Stop	Shopping Mall
11	1120	Meidling	48.171965	16.328687	14.639492	0	Hotel	Monument / Landmark	Pool	Café	Fountain	Austrian Restaurant	History Museum	Gastropub	Gift Shop	Indie Theater
12	1130	Hietzing	48.178541	16.252986	15.283490	0	Park	Supermarket	Hotel	Austrian Restaurant	Bakery	Scenic Lookout	Historic Site	Asian Restaurant	Food	Pizza Place
13	1140	Penzing	48.225662	16.222790	13.947560	0	Restaurant	Trail	Supermarket	Café	Electronics Store	Greek Restaurant	Smoke Shop	Ski Area	Multiplex	Market
14	1160	Penzing	48.225662	16.222790	13.947560	0	Restaurant	Trail	Supermarket	Café	Electronics Store	Greek Restaurant	Smoke Shop	Ski Area	Multiplex	Market
15	1150	Rudolfsheim-Fünfhaus	48.195475	16.326301	14.829543	1	Austrian Restaurant	Hotel	Restaurant	Café	History Museum	Ice Cream Shop	Coffee Shop	Persian Restaurant	Tapas Restaurant	Italian Restaurant
16	1160	Ottakring	48.214955	16.302153	14.516017	0	Supermarket	Restaurant	Austrian Restaurant	Café	Wine Bar	Gastropub	Hotel	Italian Restaurant	Tram Station	BBQ Joint
17	1170	Hernals	48.235403	16.284214	14.532162	0	Austrian Restaurant	Park	Gastropub	Bus Stop	Hotel	Gluten-free Restaurant	Comfort Food Restaurant	Scenic Lookout	Restaurant	Beer Garden
18	1180	Währing	48.234115	16.321606	15.361205	0	Supermarket	Café	Park	Austrian Restaurant	Ice Cream Shop	Restaurant	Gastropub	Italian Restaurant	Plaza	Bakery
19	1190	Döbling	48.261251	16.328471	17.019965	0	Café	Austrian Restaurant	Restaurant	Wine Bar	Italian Restaurant	Bus Stop	Scenic Lookout	Bakery	Mountain	Vineyard
20	1200	Brigittenau	48.243822	16.378147	13.963871	0	Supermarket	Plaza	Park	Austrian Restaurant	Italian Restaurant	Restaurant	Kebab Restaurant	Ice Cream Shop	Gym / Fitness Center	Fast Food Restaurant
21	1210	Floridsdorf	48.279815	16.412135	14.663259	0	Supermarket	Shopping Mall	Electronics Store	Fast Food Restaurant	Restaurant	Tram Station	Grocery Store	Gym / Fitness Center	Café	Furniture / Home Store
22	1220	Donaustadt	48.226966	16.494484	16.150128	0	Fast Food Restaurant	Metro Station	Restaurant	Indian Restaurant	Convenience Store	Soccer Field	Café	Bus Stop	Pizza Place	Breakfast Spot
23	1230	Liesing	48.141106	16.293912	14.212155	0	Supermarket	Hotel	Italian Restaurant	Bus Stop	Pharmacy	Plaza	Bakery	Food	Breakfast Spot	Furniture / Home Store

We noticed that based on top-10 most common venues districts from the first cluster (cluster = 0) appear to be more residential since they have more venues like parks and supermarkets whereas districts from the second cluster (cluster = 1) appear to be more commercial / touristic since they have more hotels and restaurants.

4. Results

We combined our findings into one choropleth map below where:

- the intensity of the fill color for districts depend on the average relative price of apartment (i.e. price per m^2),

- color of markers for districts depend on the cluster it assigned to ('Residential' or 'Commercial / Touristic')

- labels contain the following information:
    - District Name
    - Type of Area
    - Average price/m2
    - Top 10 Venues
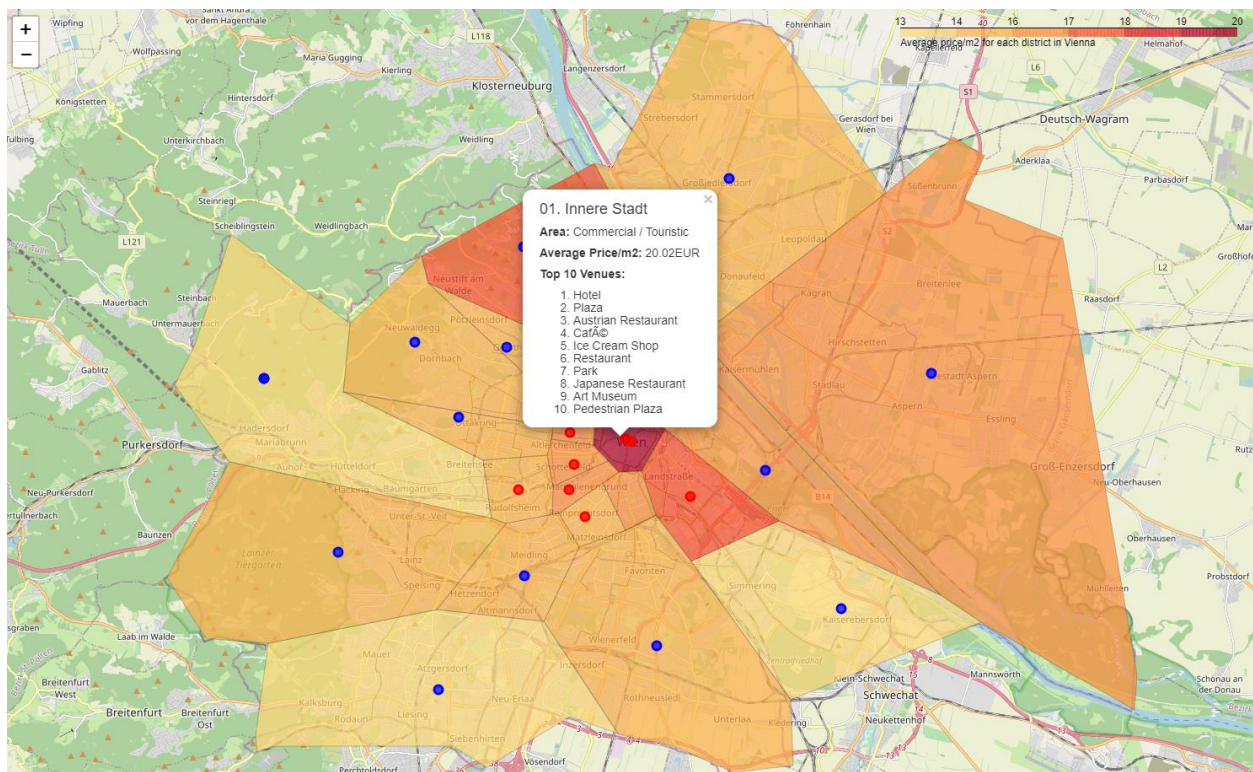
5. Discussion

We would like to point out that this project has a lot of room for improvement. For example, this analysis was based on the dataset for apartments that have 1-5 rooms with apartment size up to 400 m^2, however usually people have more specific needs when they're searching for the apartment. Also, we only used size, number of rooms, price and district information from apartment listings, yet people might want to have other criteria in the analysis to help them with decision-making process. Additionally, one can argue that

we clustering did not help to gain valuable insights since categorizing areas into "Commercial / Touristic" and "Residential" may be not sophisticated enough since usually people intuitively understand that districts in the inner circle of the city (city center) are usually considered "Commercial / Touristic" anyway.

6. Conclusion & Future Directions

Having an interactive map of the city that enables you to quickly assess which district in the city can be more preferable to you to live in is extremely useful. For this reason, people can make better decisions when considering to move to another place to leave (either from another city / country or within the same city).

Similar interactive maps can be created for multiple cities around the world to enable cross-city comparisons for those who consider several cities for relocation.

Tool can be further improved by increasing the level of sophistication for data inputs and by providing more information for apartments under the consideration