The Battle of Neighborhoods

for Applied Science Data Project

by Handi Rizkinugraha

Introduction

Aachen is one of the top 10 city in Germany with the most number of college students. Aachen also has a lot of international students who has just come to Germany. Therefore they need to look for pace to rent. The purpose of this case study is to help making some guidance for basic data about renting in Aachen, before they survey further for details based on their own needs and preferences. The conclusion will summarize about three things:

- How much is the usual cost rent in Aachen
- What number of rooms will be more efficient cost-wise
- What are the cost rate for each district around Aachen and what are the facilities around those districts

Data

The data on apartments: size, number of rooms, address, and coldprice (price with deposit and water & electricity cost excluded) is collected by web scraping a local website with apartment listings (immonet.de). We clean up the values and calculate the coldprice/m^2 by dividing the price by the area column. We also calculate the coldprice/room by dividing the price by the rooms column. The data is pre-processed and we get our first dataframe:

	PostalCode	District	Coldprice	Area	Rooms	Coldprice/m2	Coldprice/room
0	52062	Aachen	590.00	47.58	1.0	12.40	590.00
1	52062	Aachen	545.00	64.00	2.0	8.52	272.50
2	52062	Aachen	503.50	53.00	2.0	9.50	251.75
3	52062	Aachen	1399.48	101.87	3.0	13.74	466.49
4	52062	Aachen	1180.00	115.00	4.0	10.26	295.00

It is further cleaned by removing NA values and outliers, which results in a dataset of 344 apartments. From this dataset we sort them according their postal codes, so we can figure out the coordinates and then plot the data into folium map.

Using Foursquare we collect the closest venues (supermarket, restaurant, park, etc.) and select the top 10 venues for each district. After the data collection we can run k-means clustering to cluster the districts and plot them into the map.

Methodology

First, we describe our first database to know the apartment rent rate in Aachen.

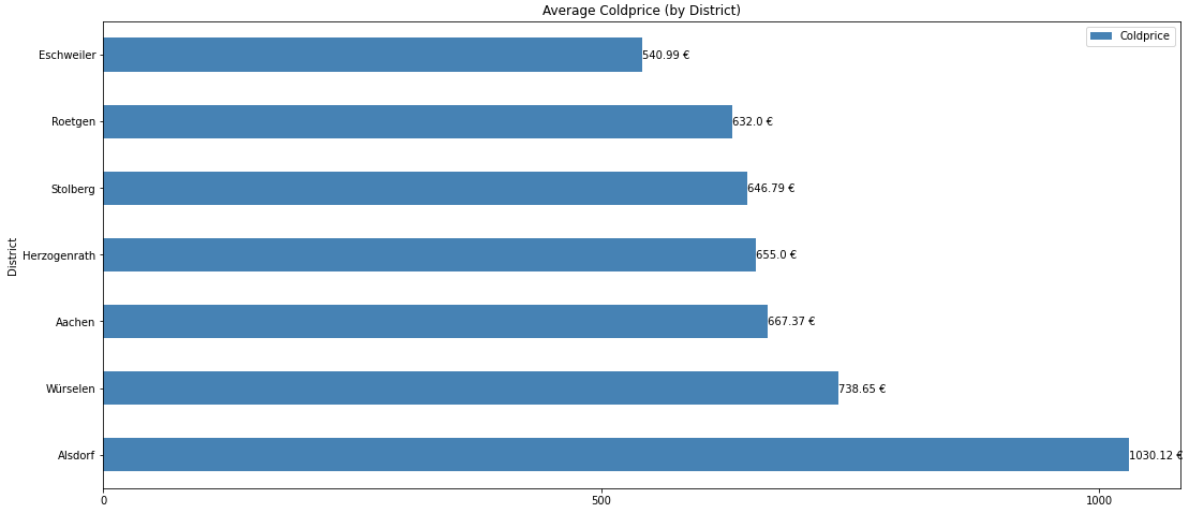
	Coldprice	Area	Rooms	Coldprice/m2	Coldprice/room
count	344.000000	344.000000	344.000000	344.000000	344.000000
mean	674.991831	65.393547	2.292151	11.077674	319.257529
std	318.449874	29.563325	0.974737	3.568296	140.339787
min	199.000000	13.000000	1.000000	4.760000	100.000000
25%	449.000000	45.000000	2.000000	8.515000	224.442500
50%	600.000000	64.000000	2.000000	10.070000	291.775000
75%	816.285000	83.250000	3.000000	12.420000	374.140000
max	2250.000000	200.000000	5.000000	26.940000	1399.000000

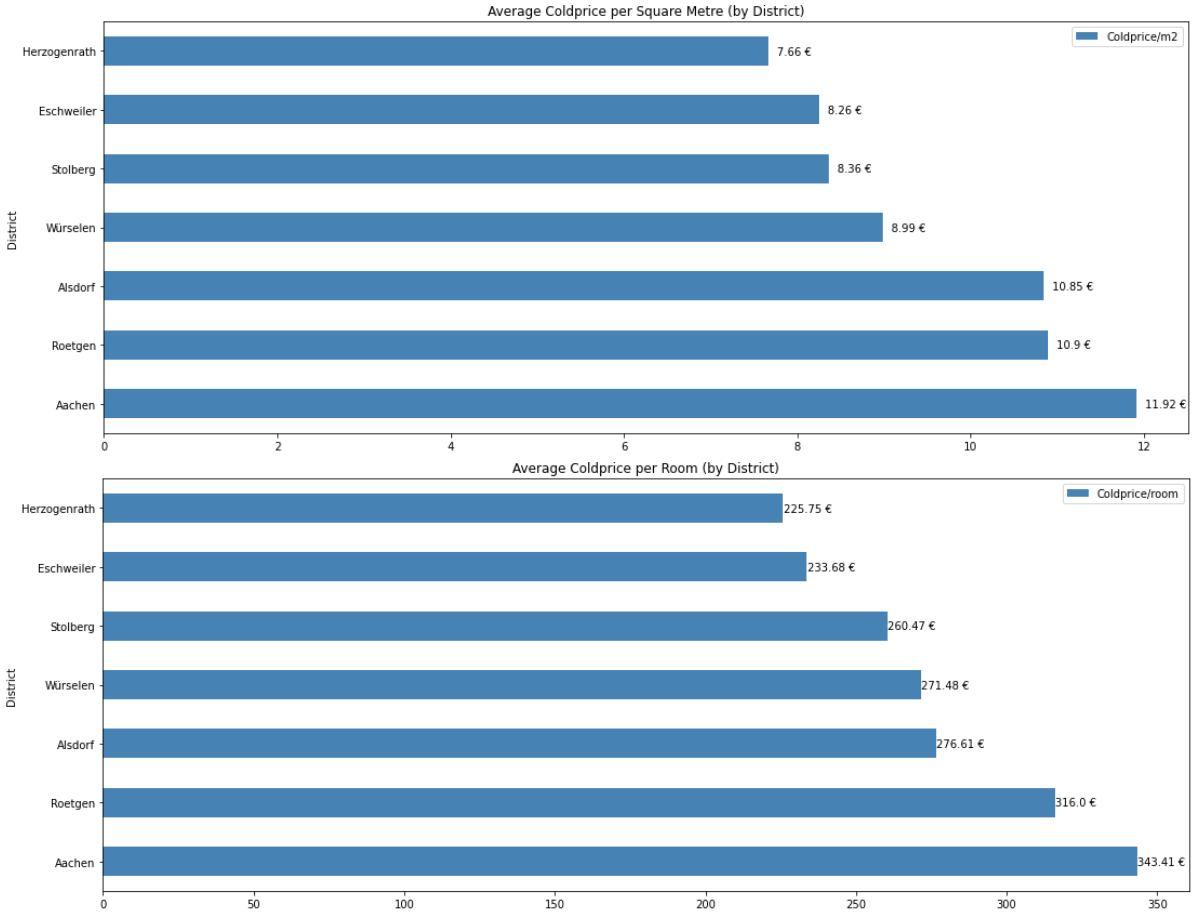
From here, we can see that the mean price for the properties around Aachen (per July 2021):

- Coldprice: 674.99 €
- Coldprice/m^2: 11.07 €
- Coldprice/room: 319.25 €

These numbers can be the first guidelines for student to know the rent cost rate around Aachen. We will break this down one by one. We then will break this down for each district to see which district has the cheapest rate.

	District	Coldprice	Area	Rooms	Coldprice/m2	Coldprice/room	DataEachDistrict
0	Aachen	667.373880	60.944520	2.114000	11.916040	343.406120	250
1	Alsdorf	1030.115625	92.080625	3.656250	10.848750	276.605625	16
2	Eschweiler	540.993000	66.295333	2.416667	8.259000	233.679667	30
3	Herzogenrath	655.000000	82.633333	2.916667	7.663333	225.750000	6
4	Roetgen	632.000000	58.000000	2.000000	10.900000	316.000000	1
5	Stolberg	646.791667	78.146250	2.604167	8.360417	260.473750	24
6	Würselen	738.651765	80.458235	2.764706	8.990000	271.477059	17

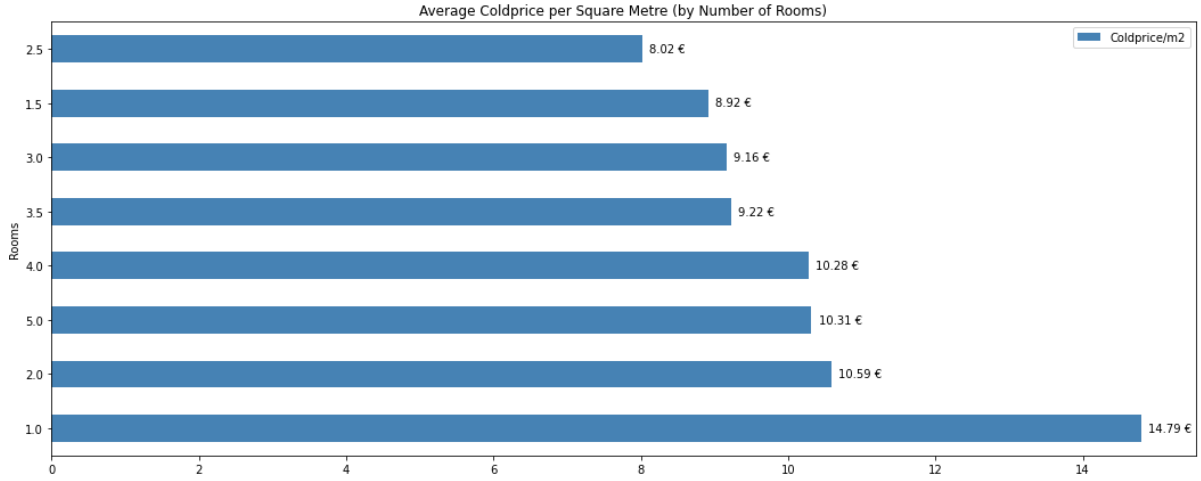


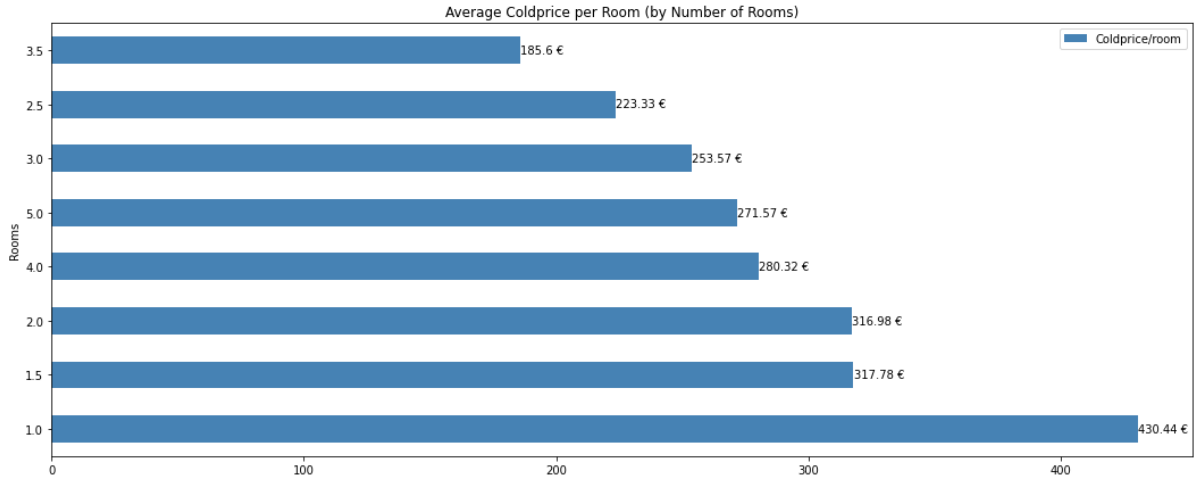


From the table and the graphics above, we can see that from cold price, cold price per square metre, and cold price per amount of room aspects, **Eschweiler** comes first as the overall cheapest district. Followed by **Stolberg** in second. The data also have quite good amount of samples for representation (around 10% from the total sample). So this analysis is quite valid.

We also will sort the “rental data” table by the number of rooms if students want to see what is usually the cheapest apartment in Aachen according its number of rooms.

	Rooms	Coldprice	Area	Coldprice/m2	Coldprice/room	Samples
0	1.0	430.442195	32.038049	14.791707	430.442195	82
1	1.5	476.666667	54.666667	8.916667	317.776667	3
2	2.0	633.961316	59.932982	10.586754	316.980702	114
3	2.5	558.317500	66.387500	8.015000	223.327500	4
4	3.0	760.723786	82.783592	9.160971	253.574563	103
5	3.5	649.595000	71.375000	9.225000	185.600000	2
6	4.0	1121.280312	106.802500	10.275625	280.320000	32
7	5.0	1357.840000	129.802500	10.312500	271.567500	4

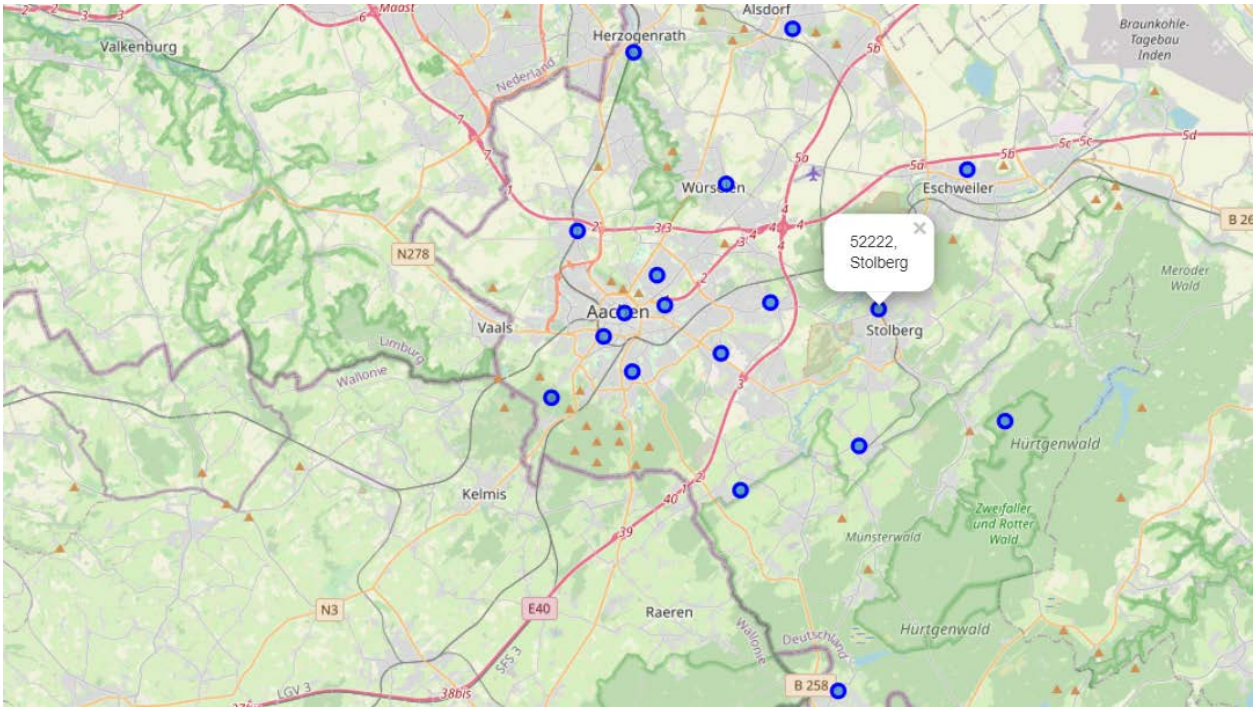




If we look at the graphs, the 3,5 and 2,5 rooms apartment seems to come as the cheapest option. But there is too few sample numbers from them (<5), so we are not quite sure whether this is valid analysis or not.

Meanwhile, the **3 room apartments** come close to them, and also with much more number of samples (around 30% of the total samples). So we can say that this is the cheapest number of apartment rooms option around Aachen.

Using the apartment data and the district coordinates we can visualize each district and their distance to the centre of Aachen. We can produce the following folium map:



As we have analyzed from the dataframe above, we have mentioned that **Eschweiler** and **Stolberg** are the two cheapest apartment place around Aachen. From this map, we can see that **Eschweiler** district has the shorter distance to the centre of Aachen, around 1 hour using public transport. Meanwhile, **Stolberg** takes **even further** 30 minutes more using public transport.

We then analyze each district facilities using Foursquare. We search for the most common venues (supermarket, restaurant, park, etc.) and select the top 10 venues for each district. After the data collection we can run k-means clustering to cluster the districts, in this case we will use k = 2 value.

```
cluster1
```

	District	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
12	Roetgen	Supermarket	Spa	Gas Station	Italian Restaurant	Fast Food Restaurant
13	Stolberg	Supermarket	Drugstore	Train Station	Big Box Store	Museum
15	Stolberg	Supermarket	Restaurant	Bank	Zoo	Fried Chicken Joint
17	Alsdorf	Supermarket	Gas Station	Drugstore	Zoo	Bus Station

```
cluster2
```

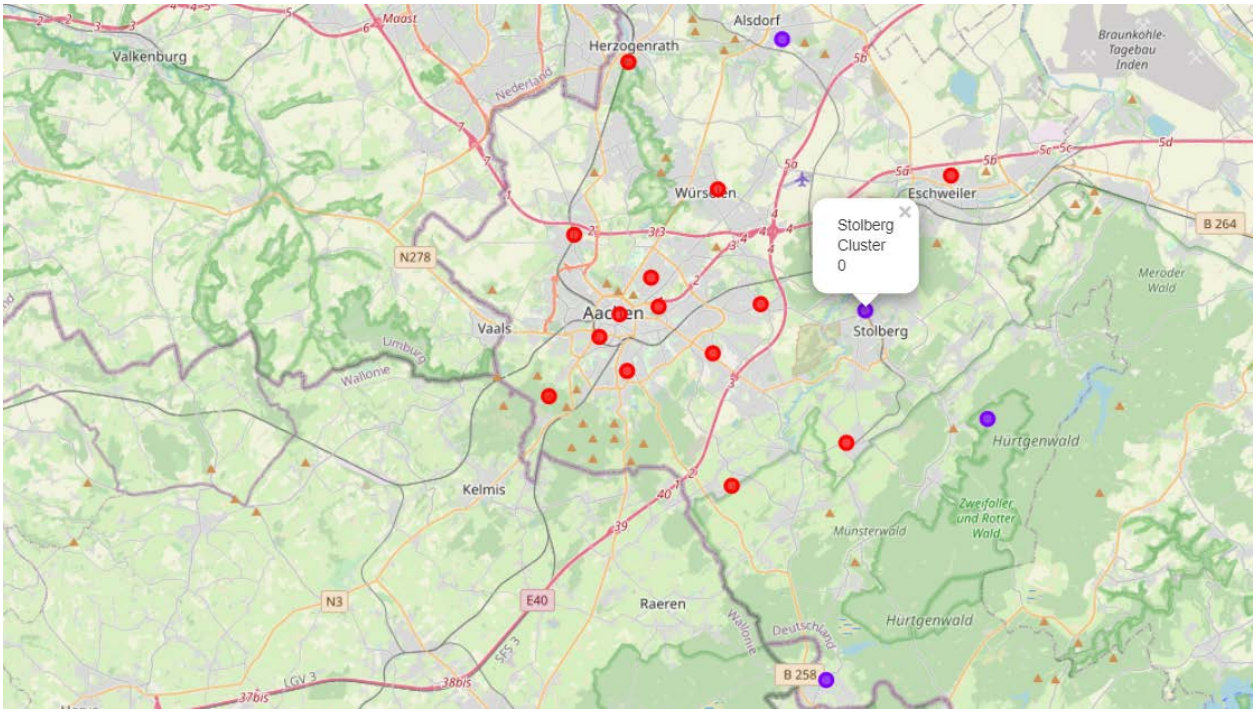
	District	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Aachen	Bar	German Restaurant	Bakery	Plaza	Supermarket
1	Aachen	Bar	German Restaurant	Park	Bakery	Café
2	Aachen	Bar	Café	German Restaurant	Bakery	Turkish Restaurant
3	Aachen	Bar	Bakery	Park	Plaza	Italian Restaurant
4	Aachen	Bar	Bakery	Plaza	Supermarket	Italian Restaurant
5	Aachen	Supermarket	Hotel	Liquor Store	Outlet Store	Gym
6	Aachen	Supermarket	Café	Restaurant	Pool	Scenic Lookout
7	Aachen	Supermarket	German Restaurant	Hotel	Bakery	Trail
8	Aachen	Supermarket	Bus Stop	Hotel	Turkish Restaurant	Bar
9	Aachen	Supermarket	Fast Food Restaurant	Bakery	Hotel	Bus Station
10	Herzogenrath	Supermarket	Snack Place	Train Station	Drugstore	Hotel
11	Würselen	Supermarket	Drugstore	Furniture / Home Store	Big Box Store	Hotel
14	Stolberg	German Restaurant	Restaurant	Pizza Place	Gastropub	Supermarket
16	Eschweiler	Supermarket	Drugstore	Electronics Store	Hotel	Thai Restaurant

From this data, we can see the only difference between inside the centre of Aachen and outside is that bar is the most common venue in the centre of Aachen.

Other than that, these districts have pretty much the same common facilities, such as supermarket, restaurant, and public transport. So all district provide quite enough facilities for daily needs.

Result

By analyzing the clusters we can see that cluster 1 is more residential since it contains mostly supermarket and drugstore, while cluster 2 is more commercial / touristic and contains many hotels, bars, and restaurants. We can visualize this cluster in a map.



Discussion & Conclusion

From this analysis, we can determine the average cold price for the properties around Aachen based on immonet.de as follows (per July 2021):

- Coldprice: 674.99 €
- Coldprice/m^2: 11.07 €
- Coldprice/room: 319.25 €

We also find by the price, **Eschweiler** comes first as the overall cheapest district. Followed by **Stolberg** in second. However, **Eschweiler** district takes around 1 hour to the centre of Aachen using public transport. Meanwhile, **Stolberg** takes **even further** 30 minutes more using public transport. All the district have pretty much the same facilities too, so there is no need to worry about daily needs in this radius.

Whereas by the number of room categories, we find that **3 room apartments** is the cheapest option.