

CAPSTONE PROJECT – STUTTGART APARTMENTS

BY AXEL SPAHR



PROBLEM DEFINITION

Rents in Stuttgart were rising a lot over the last years. It is therefore often difficult to find an apartment that fits your needs within your given budget limit.

So with this Project we can help people to see, in which districts of Stuttgart they should start looking:

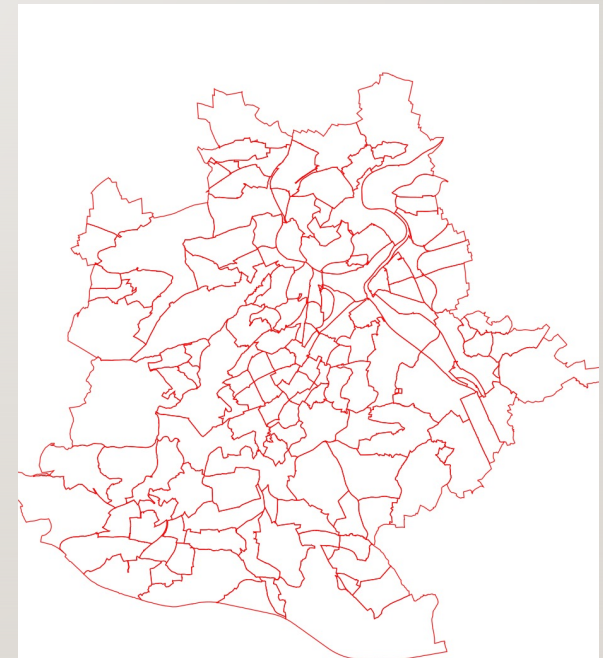
- If the rent demanded by the owner is too high for the specific District
- Which district is the cheapest district
- Which area is the best regarding the venues

This project helps people who want to move to Stuttgart or people who already live in Stuttgart but are looking for a new apartment.



DATA

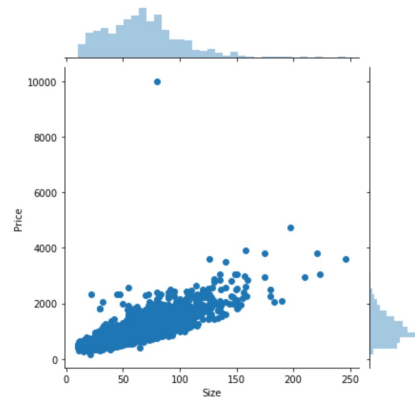
The data on apartments (size, number of rooms, address, and price) is collected by scraping a local website with apartment listings (Immoscout24.de). For the coordinates of the districts we use geopy and take the Data from the official website of the city stuttgart (https://www.stuttgart.de/medien/ibs/KLGL_DXF.zip) that I labeled and converted myself. To Collect the Venues I use Foursquare. After the Data collection process we run k-means clustering to cluster the districts into residential and commercial areas and visualize the data on a choropleth map.



METHODOLOGY

Plot Price to check for Outliers

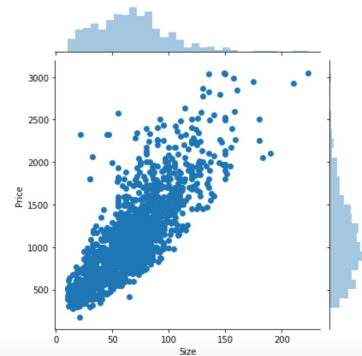
```
In [22]: sns.jointplot(x='Size', y='Price', data=rental_properties[['Size', 'Price']])
Out[22]: <seaborn.axisgrid.JointGrid at 0x7f91d342b510>
```



Identification of outliers

Remove outliers

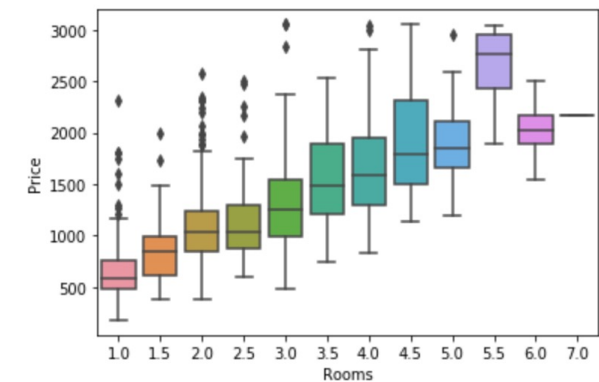
```
In [23]: rental_properties = rental_properties[rental_properties['Price'] <= 3200]
In [24]: rental_properties = rental_properties[rental_properties['Price/m2'] <= 200 ]
In [25]: sns.jointplot(x='Size', y='Price', data=rental_properties[['Size', 'Price']])
Out[25]: <seaborn.axisgrid.JointGrid at 0x7f91d2d299d0>
```



Removing those outliers

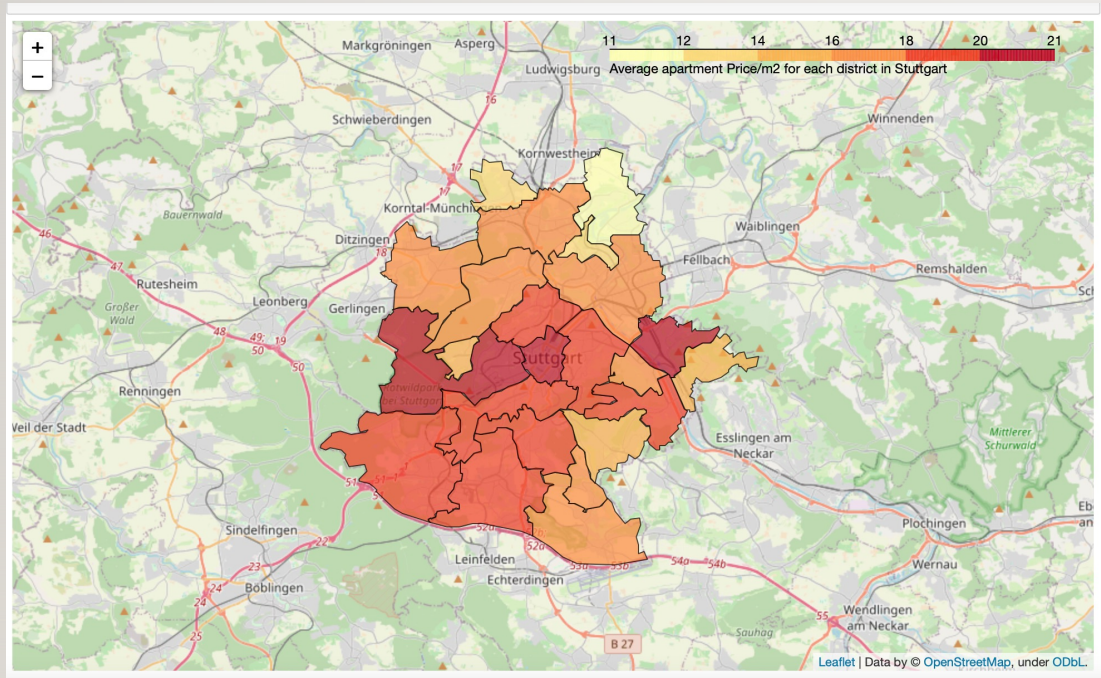
```
In [28]: sns.boxplot(x='Rooms', y='Price', data=rental_properties)
```

```
Out[28]: <matplotlib.axes._subplots.AxesSubplot at 0x7f91d325ddd0>
```

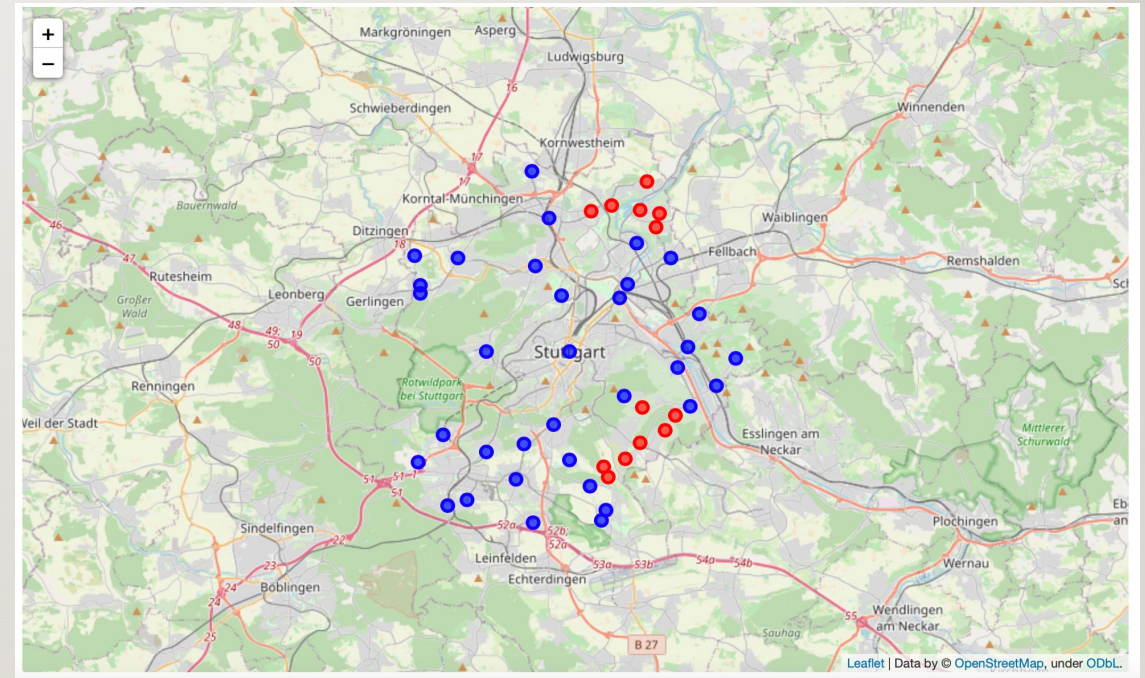


Plotting dependencies

METHODOLOGY



Average apartment Price/m² for each district in Stuttgart



Clustering the city districts

RESULTS

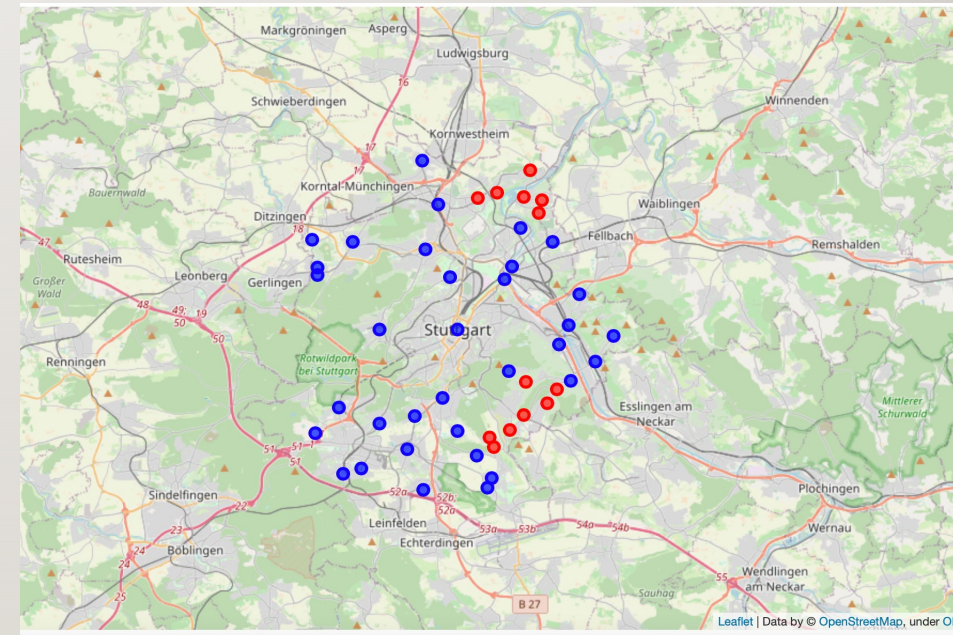
- By analyzing the clusters we can see that the blue cluster is more central and more commercial/touristic with restaurants, theaters, zoos, etc. while the red cluster is more decentral and residential with metro stations, supermarkets, and soccer fields.

```
In [55]: stuttgart_grouped.insert(0, 'Cluster Labels', kmeans.labels_)
```

```
In [56]: art_merged.drop(['District', 'Latitude', 'Longitude'], axis='columns', inplace=False).groupby('Cluster Labels').max()
```

Out[56]:

	Price/m2	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
Cluster Labels											
0	21.418207	Supermarket	Turkish Restaurant	Zoo Exhibit	Zoo Exhibit	Supermarket	Turkish Restaurant	Zoo Exhibit	Theater	Theater	Taverna
1	20.386505	Supermarket	Supermarket	Supermarket	Train Station	Trail	Soccer Field	Spa	Soccer Field	Taverna	Shopping Mall



CONCLUSION

- Based on our analysis we can determine that the districts **West, Mitte and Untertürkheim** are the most expensive district to live in, however by clustering we determined that there are several more similar districts where the price/m2 is significantly lower. Therefore, if someone wants to rent an apartment but cannot afford to live in those the district, they could look for apartments in **Botnang** or **Wangen** which are also very central but have lower prices for renting apartments and are also comparable in venues. However the cheapest district in **Stuttgart** overall is **Mühlhausen**, even tho it's a little decentral.

