

Rent influencing factors – in Stuttgart

Applied Data Science Capstone Project on Coursera

Rents are getting more and more expensive

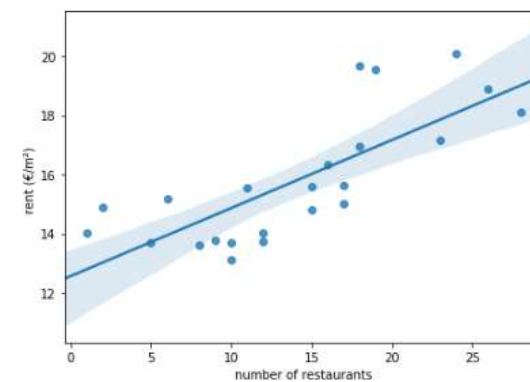
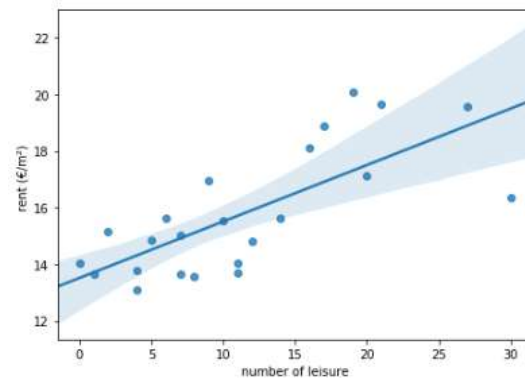
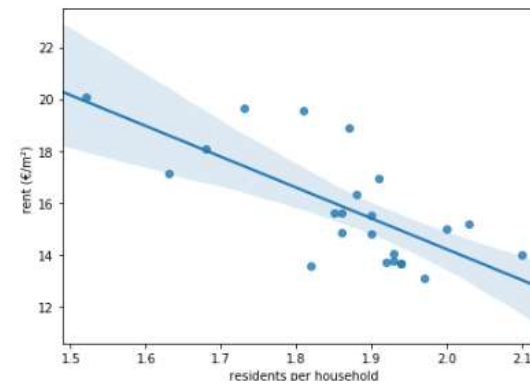
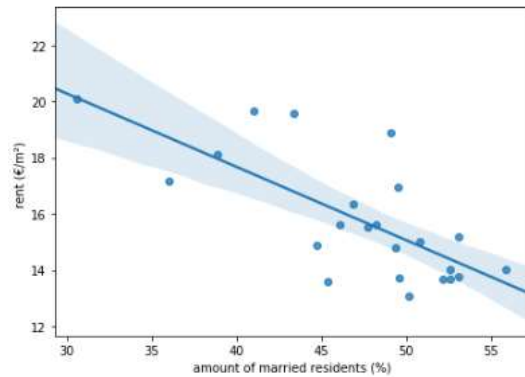
- Rents are influenced by many factors
- What factors make the rent rise?
- Can we build a model that can predict the average rent ?

Data acquisition and preperation

- Average rents in the boroughs from <https://www.wohnungsboerse.net/mietspiegel-Stuttgart/972>
- Coordinates of the boroughs from <https://www.suche-postleitzahl.org/stuttgart-plz-70173-70629.608e>
- General statistics about the boroughs from <https://statistik.stuttgart.de/statistiken/statistikatlas/atlas/atlas.html?indikator=i0&select=00>
- Data about the venues in the boroughs from foursquare
-> 12 features were extracted and prepared

Correlation between feature and rent

- Regression plots were used to determine relevant features:



Correlation between feature and rent

- The Pearson Correlation and the p-value validates our finding

feature	Pearson Correlation	p value
rent (€/m ²)	1.0	0.0
residents per km ²	0.312	0.14696753
sport and leisure area (m ² /resident)	0.08	0.71755271
Average age	-0.461	0.02680288
amount of non-german residents (%)	0.121	0.58231485
amount of married residents (%)	-0.723	9.702e-05
residents per household	-0.706	0.00016678
living area per flat	-0.152	0.48838061
number of restaurants	0.771	1.697e-05
number of sports	0.11	0.61821963
number of leisure	0.738	5.747e-05
number of transport	-0.333	0.11995345
number of shopping	-0.607	0.00214644

Building a model

- A multiple linear regression model was used because we have several features with a linear correlation to the rent
- The cross validation of 0.24 shows however, that the used features are not very informative

Conclusion

- We have found some features that show a correlation to the rent
- Other cities should be looked at to see if the same features occur
- More data and a wider set of features should be included in future researches