
Utrecht housing prices

Coursera assignment





Introduction

High housing prices:

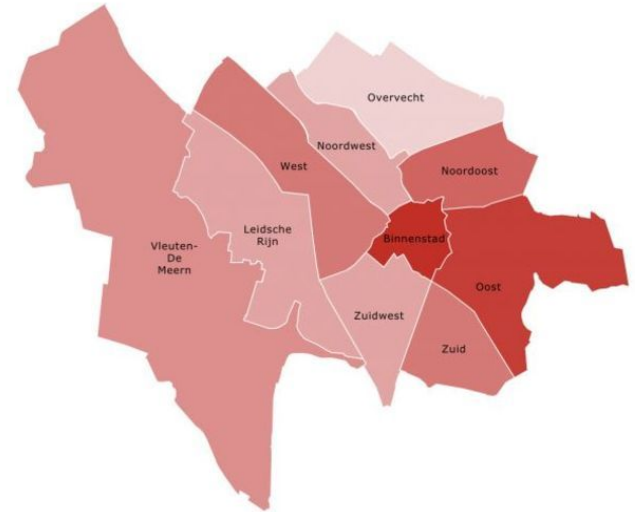
- Why is it so high?
- What are the factors influencing this?

Study:

- Clusters
- Determinants for price

Data

- General Neighbourhood data from CBS
- House prices data in euros
- Foursquare API to determine venue is in neighbourhood



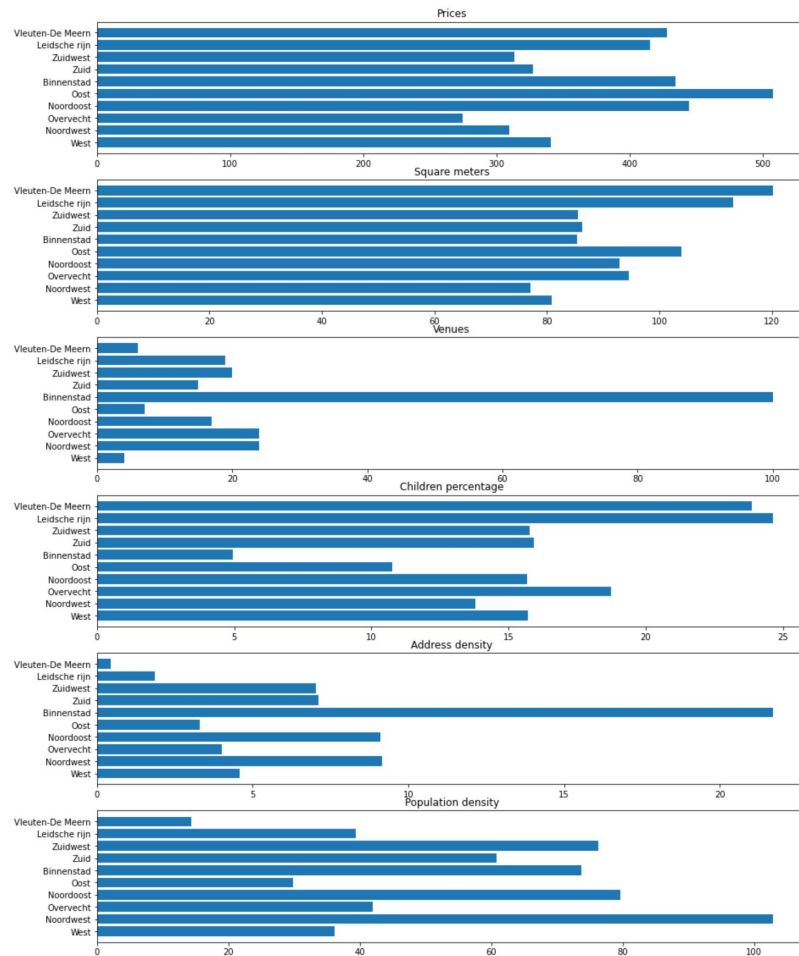
Methodology

Analyses:

- K-means cluster analysis
- Linear regression analysis

Preprocessing:

- Develop more meaningful variables (see tables)
- Min-max scaling





Results

- Venue and address_density have negative collinearity
- Children negatively related to housing prices
- Population density slightly negatively related to prices
- Floor area positively related to housing prices

	0	Coefs
0	Venue	-538.558745
1	children_percentage	-203.656263
2	address_density	637.866379
3	population_density	-52.330406
4	square_meters_2019	338.110482

Conclusion

- Clear clusters can be found
- Results provide insight, but more research and data is needed
- This study can be used as starting point for generation of further insights

