DataFest 2025
39 Insights

Annika Hörügel, Silas Kaatz, Samuel Roscher, Yanchu Zhang

# Research Question and Approach

Data

- ▶ ImmoScout cross-sectional ad data on advertised aparments and houses for sale
- ▶ Municipal information by municipality and 100x100m grid
- ▶ Election data from Bundestagswahl 2021
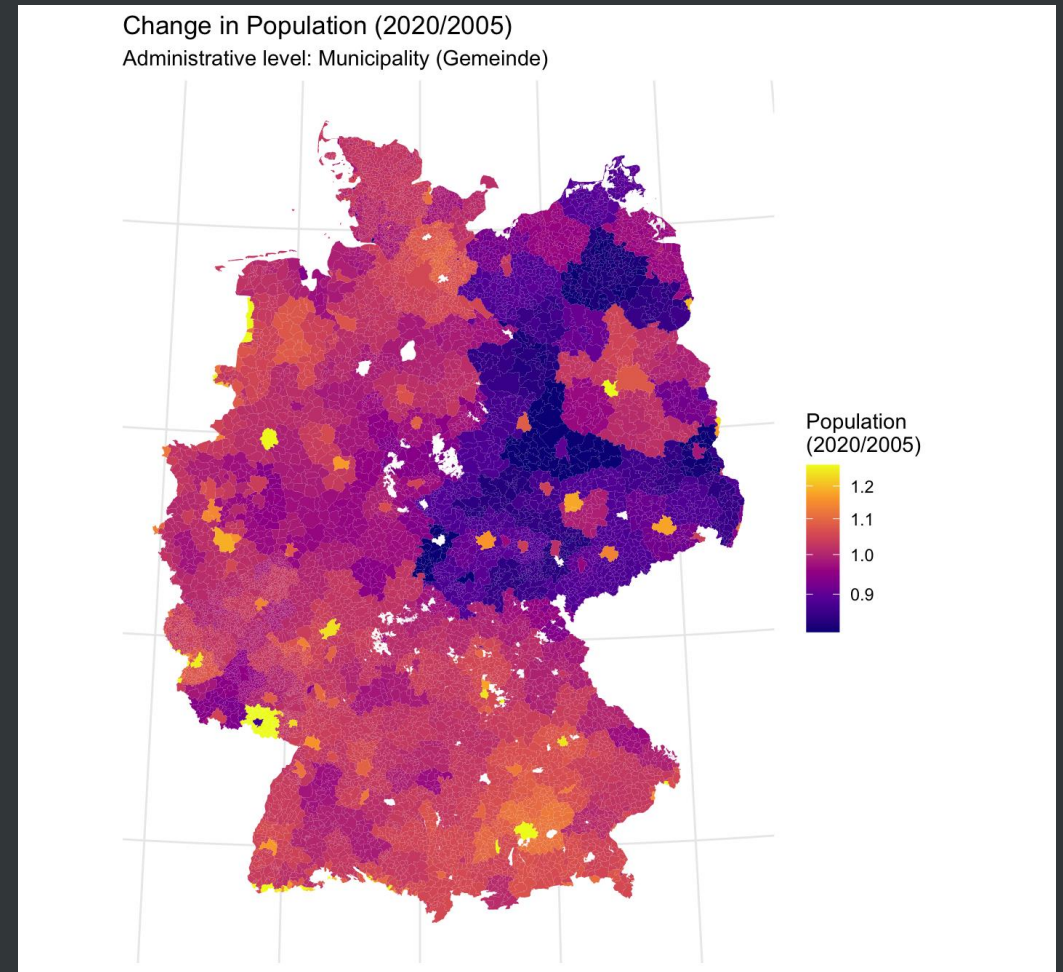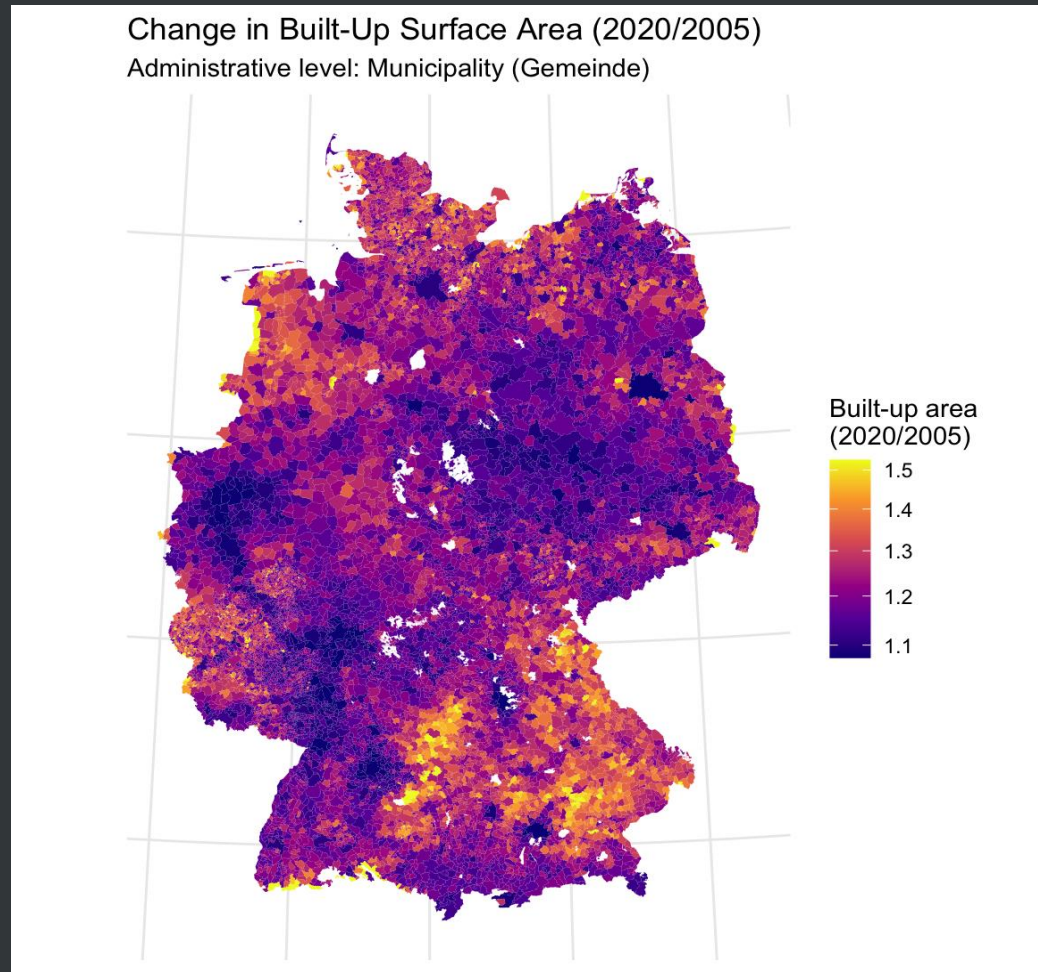
Research Question

- ▶ We decided to look at real estate prices as our target variable to understand the most impactful features that explain the variance in prices.
- ▶ We ran two different regressions, once for house sales ads and once for apartment sales ad.
- ▶ *H1: Unconventional features explain price variance more than expected.*
- ▶ *H2: House and apartment sales have similar coefficients for the same variables.*
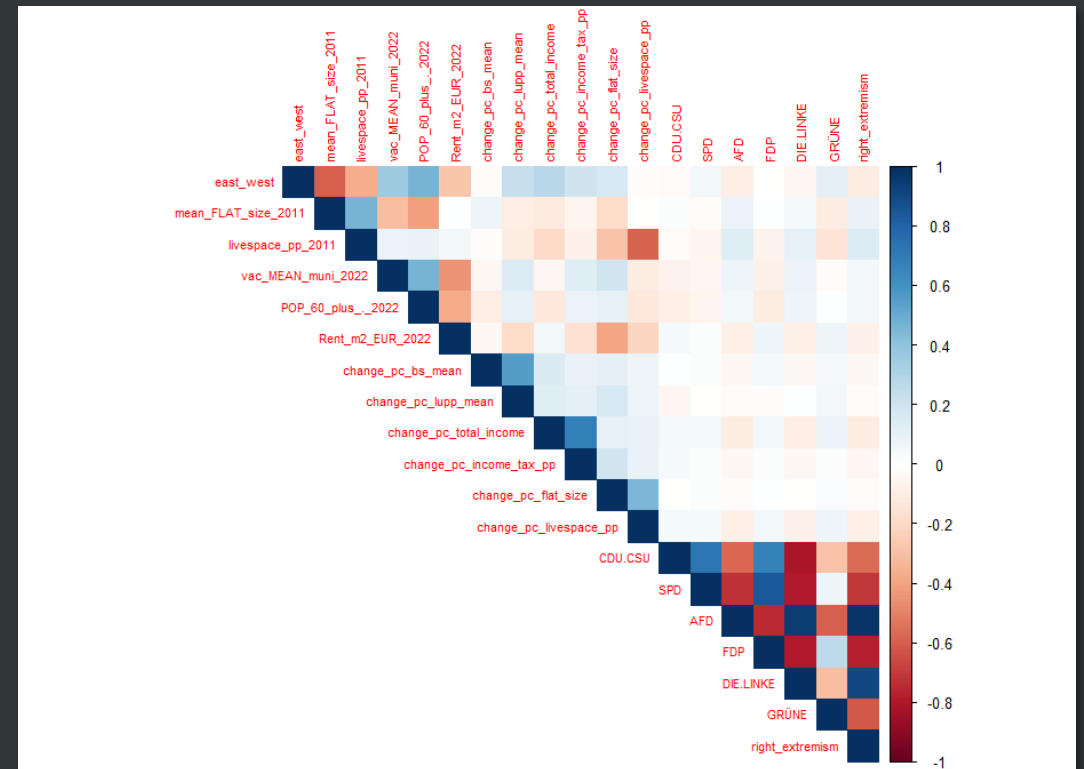
Correlation plot of voting results and municipal data

Plots
show the change in Population by grid and municipality as a quotient

Change in Built-Up Surface Area (2020/2005)
Administrative level: Municipality (Gemeinde)

Built-up area
(2020/2005)
1.5
1.4
1.3
1.2
1.1

Change in Population (2020/2005)
Administrative level: Municipality (Gemeinde)

Population
(2020/2005)
1.2
1.1
1.0
0.9

Regression Coefficients Comparison and Analysis

House data

```
Coefficients:

                           Estimate
(Intercept)               -3.306e+06
baujahr                    1.749e+03
wohnflaeche                2.060e+03
zimmeranzahl              -9.625e+03
laufzeittage             -3.522e+01
denkmalobjekt1             1.710e+05
einliegerwohnung1          1.655e+04
keller1                    3.297e+04
rollstuhlgerecht1        -1.108e+04
KFL_2021                 -8.867e+02
Total_income_EUR_2020      2.443e-01
drive_distance.km_2025   -1.526e+03
ratio_pop25_pop05          1.727e+05
ratio_bs20_bs05            3.430e+05
ratio_lupp20_lupp05      -4.358e+05
```

Apartment data

```
Coefficients:

                           Estimate
(Intercept)               -2.955e+06
baujahr                    1.238e+03
wohnflaeche                4.135e+03
zimmeranzahl              -1.413e+04
keller1                    1.969e+04
rollstuhlgerecht1          6.728e+04
KFL_2021                 -4.341e+02
Total_income_EUR_2020      7.460e-02
drive_distance.km_2025     9.427e+02
ratio_pop25_pop05          5.678e+05
ratio_bs20_bs05          -2.614e+05
ratio_lupp20_lupp05        1.391e+05
```