

# Hacking the German Apartment Hunt

How I used Python to glean helpful insights for renters from scraped apartment listings.

By Jacy Marmaduke



Image credit: Roman Kraft on Unsplash

# About the project

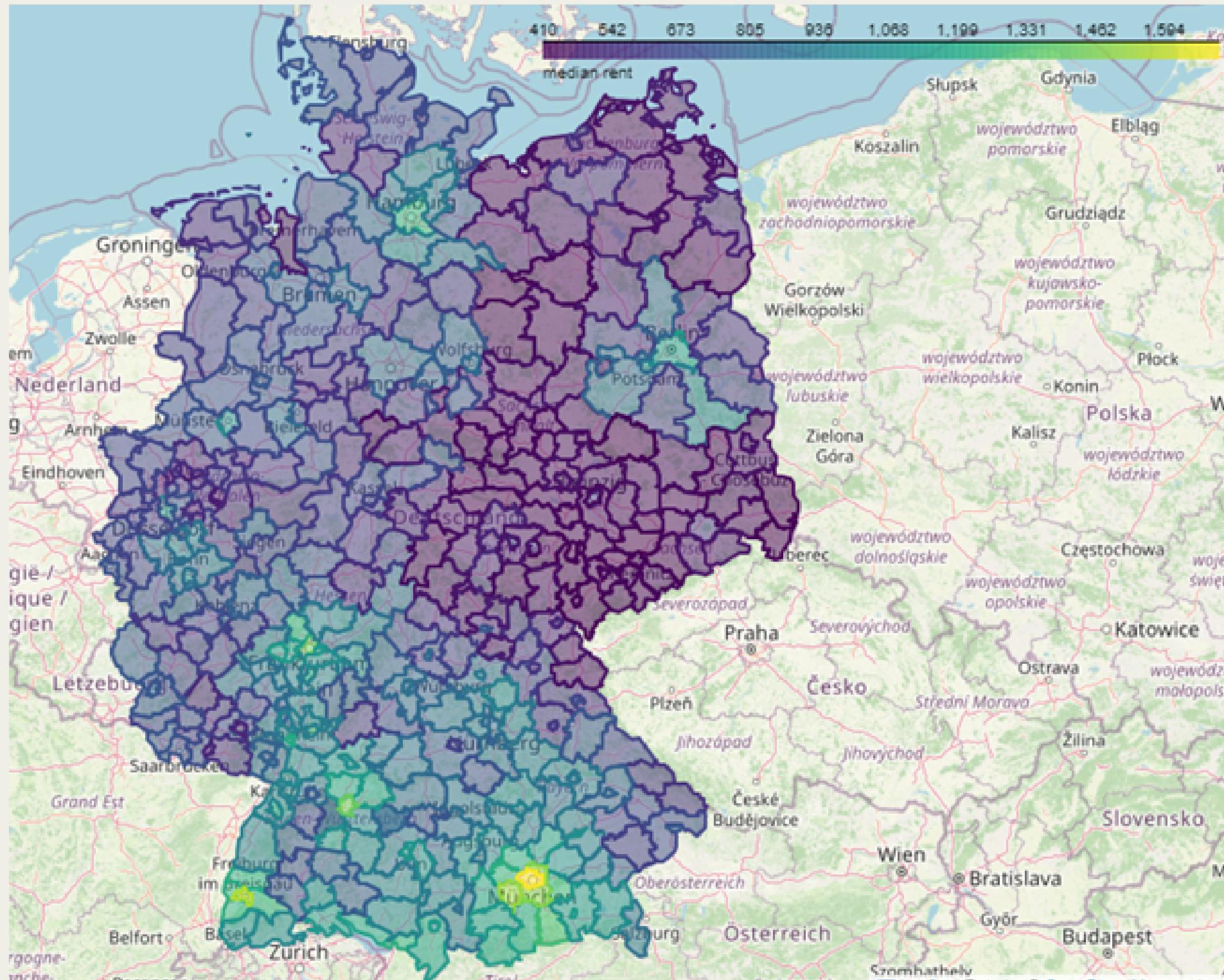


Image: Interactive Geopandas map of median rents by district

**The big question:** What factors have the biggest influence on rent in Germany?

**Objective:** Use German apartment listings and population data to produce meaningful insights about apartment hunting for renters new to the German housing market.

**Context and purpose:** I designed and executed this project as part of the Career Foundry Data Analytics course to demonstrate my skills in Python and Tableau.

## Tools and methods

### Python

**Methods:** Data merging, cleaning and wrangling, correlation-testing, multiple linear regression, k-means clustering, geographic analysis, visualization

**Libraries:** Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, Folium, Geopandas

### Tableau

Storyboard design, combination mapping, histograms, scatterplots and boxplots

# Meet the data

- A Kaggle user scraped over **257,000 German apartment listings** from ImmoScout24, Germany's top real estate website, and posted the raw data to Kaggle.
- I combined the ImmoScout24 data with **my own dataset** of German population variables from the Federal Statistical Office of Germany.
- **Merging the datasets** meant I could assess rent through the lens of local and regional population growth, population density and residential building stock.

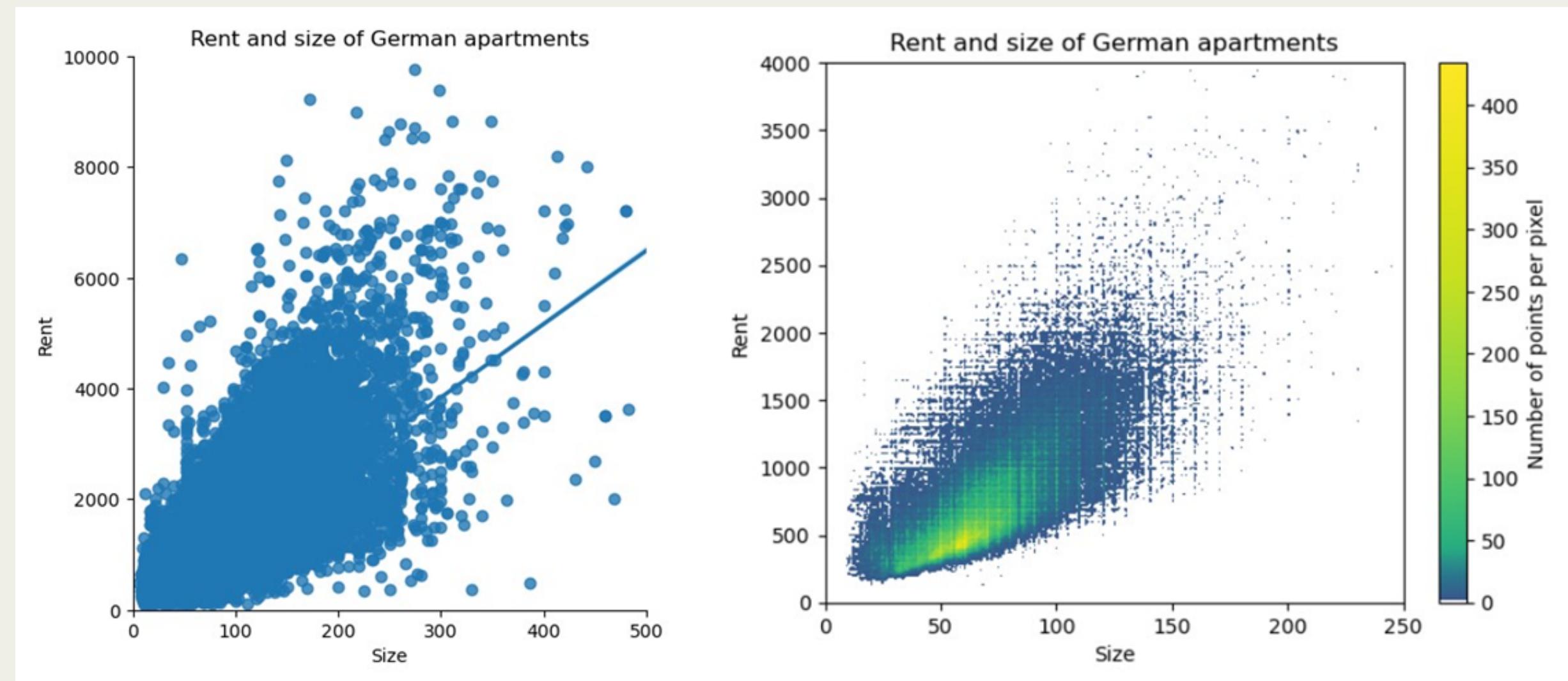


Credit: Paul Fiedler on Unsplash

# Challenges and solutions

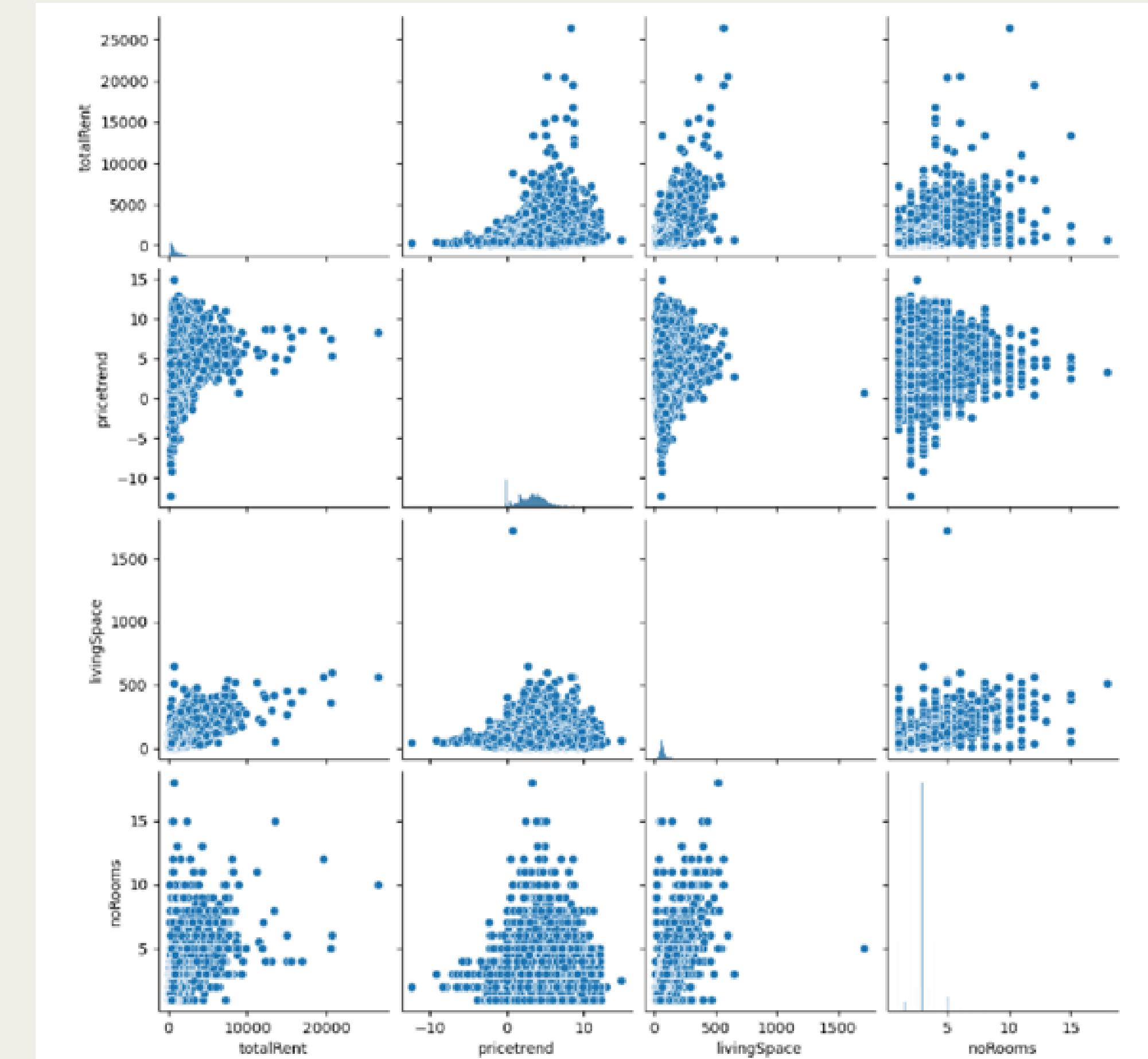
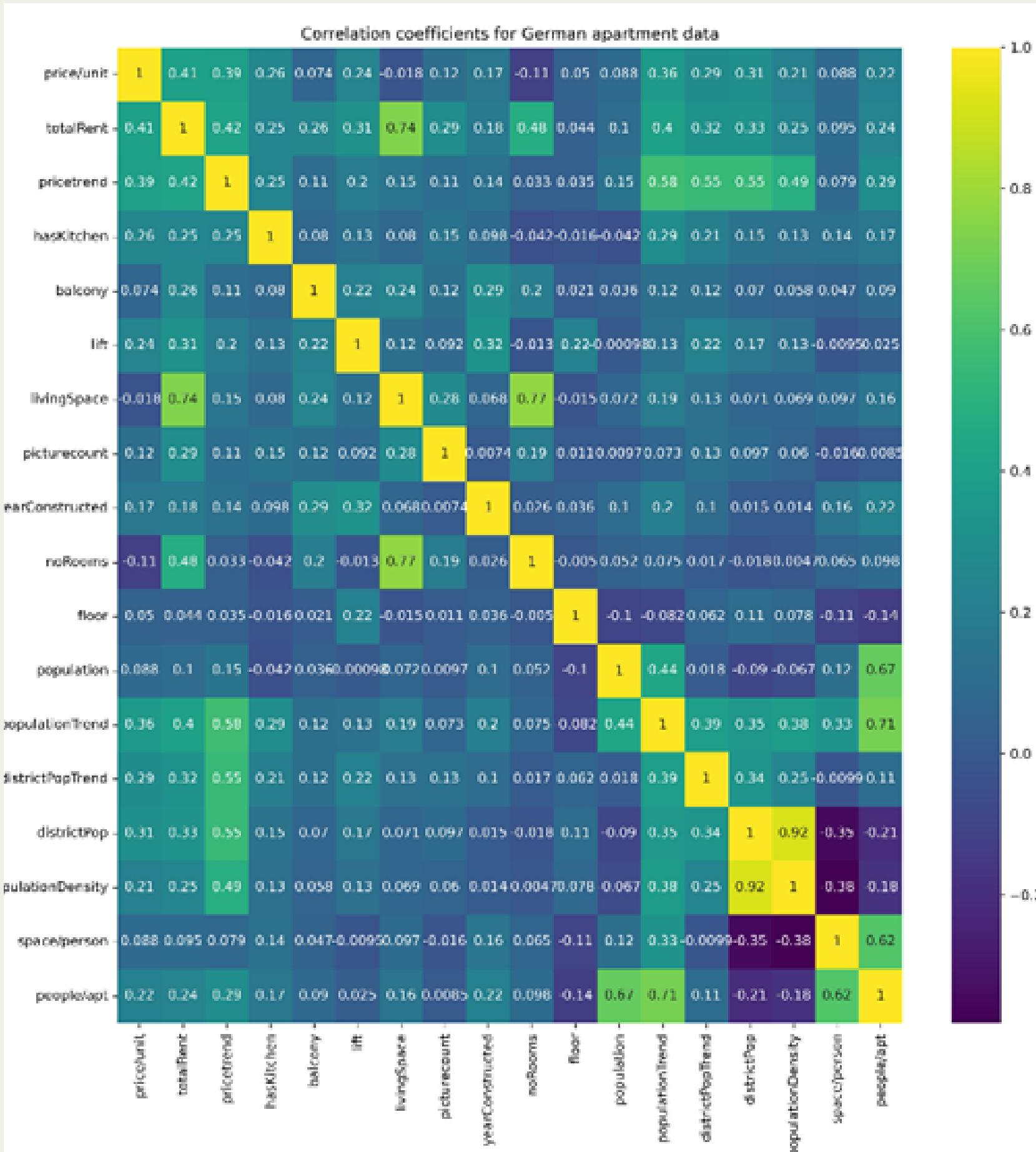
**Challenge:** The dataset was big and messy, with over 40 variables and so many points that scatterplots were tough to interpret.

**Solution:** Me and copy-editing go way back – I was a high school newspaper geek – so smoking out and correcting errors with Python was a labor of love. I used density scatterplots to convey relationships more clearly in the crowded dataset.

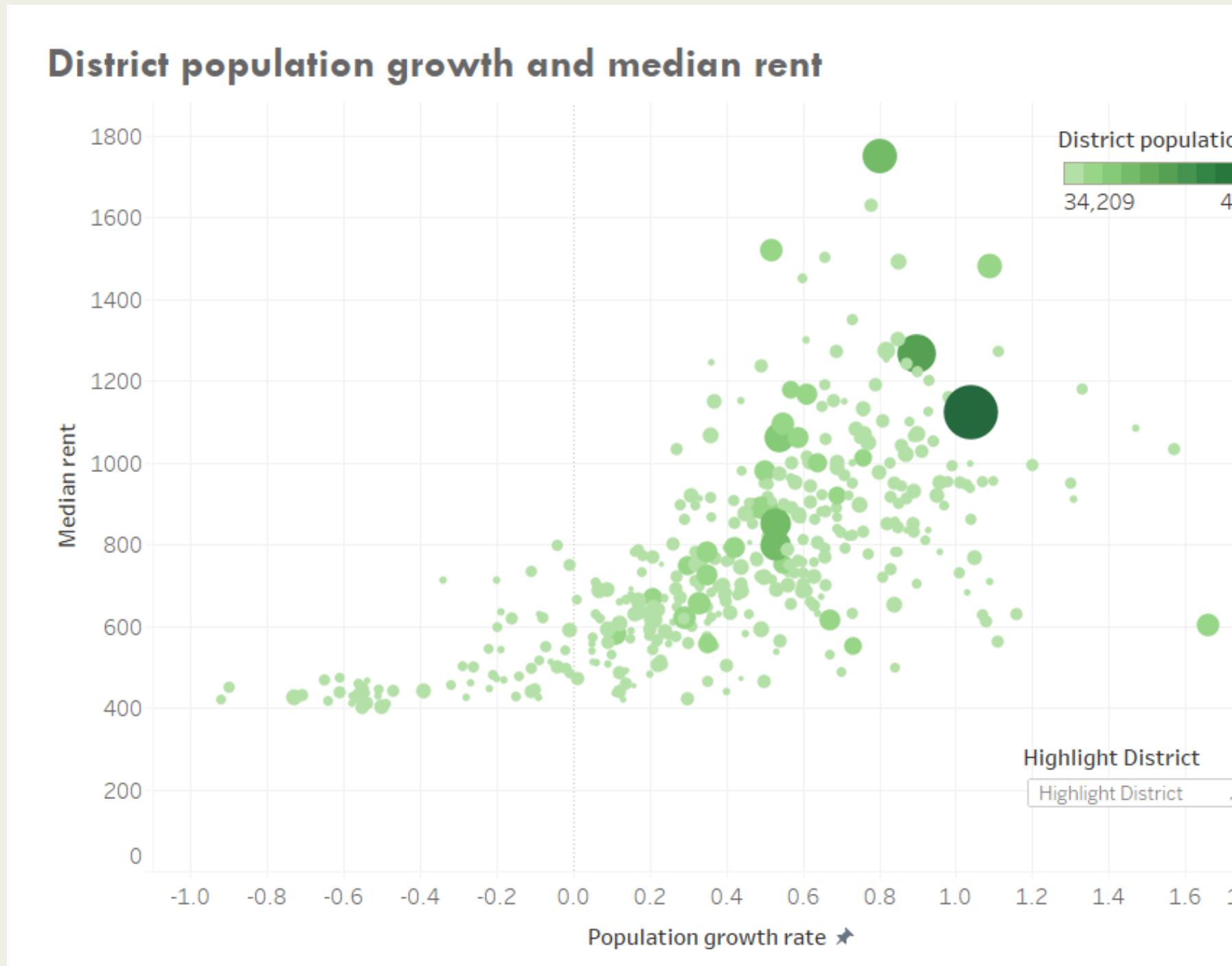


What a difference graphing method can make! These charts display the same cleaned data.

# Challenge: A lot of factors have small-to-moderate influences on rent, with non-linear relationships.



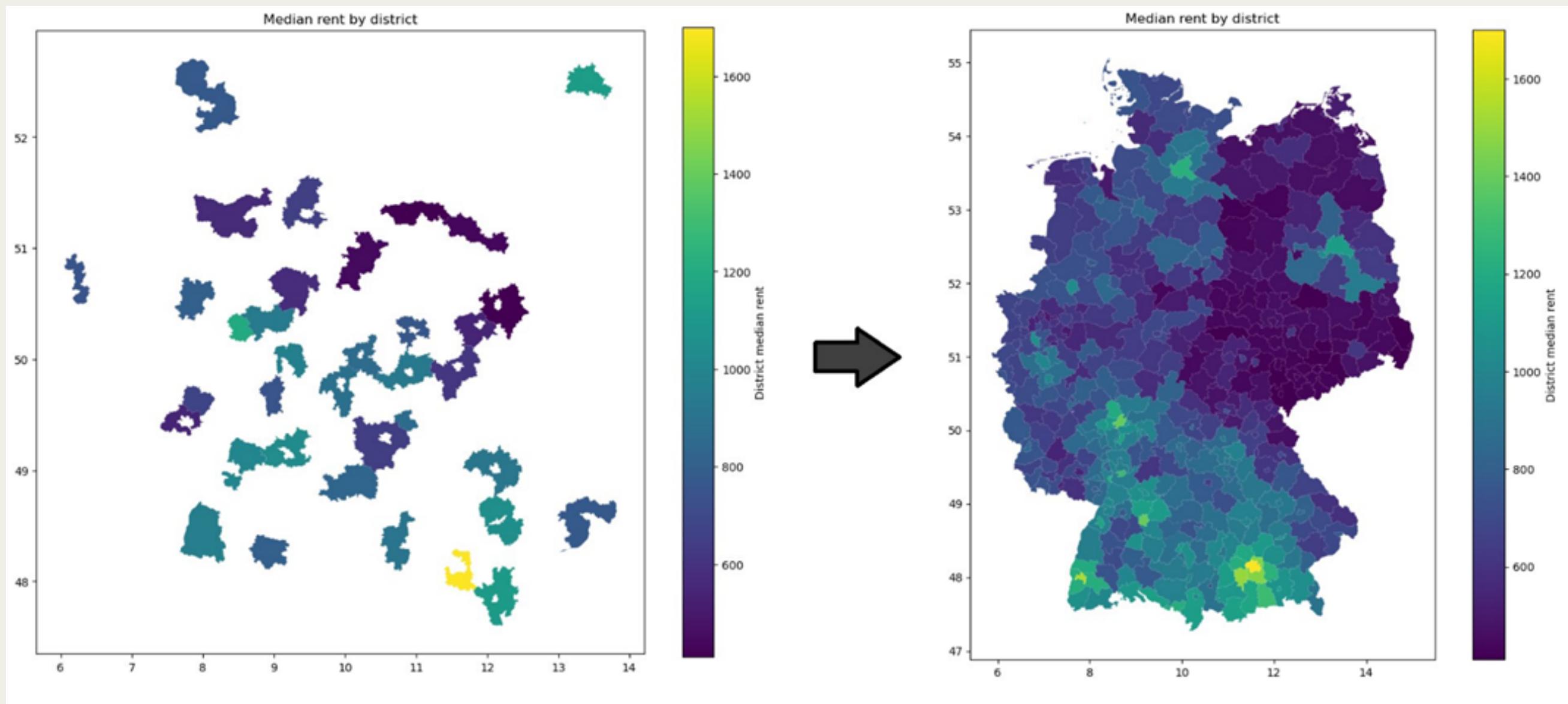
**Solution:** I highlighted the factors with the strongest correlations with rent – apartment size and district population growth – and used k-means clustering to explore rent variation in a different way.



insert k-means viz

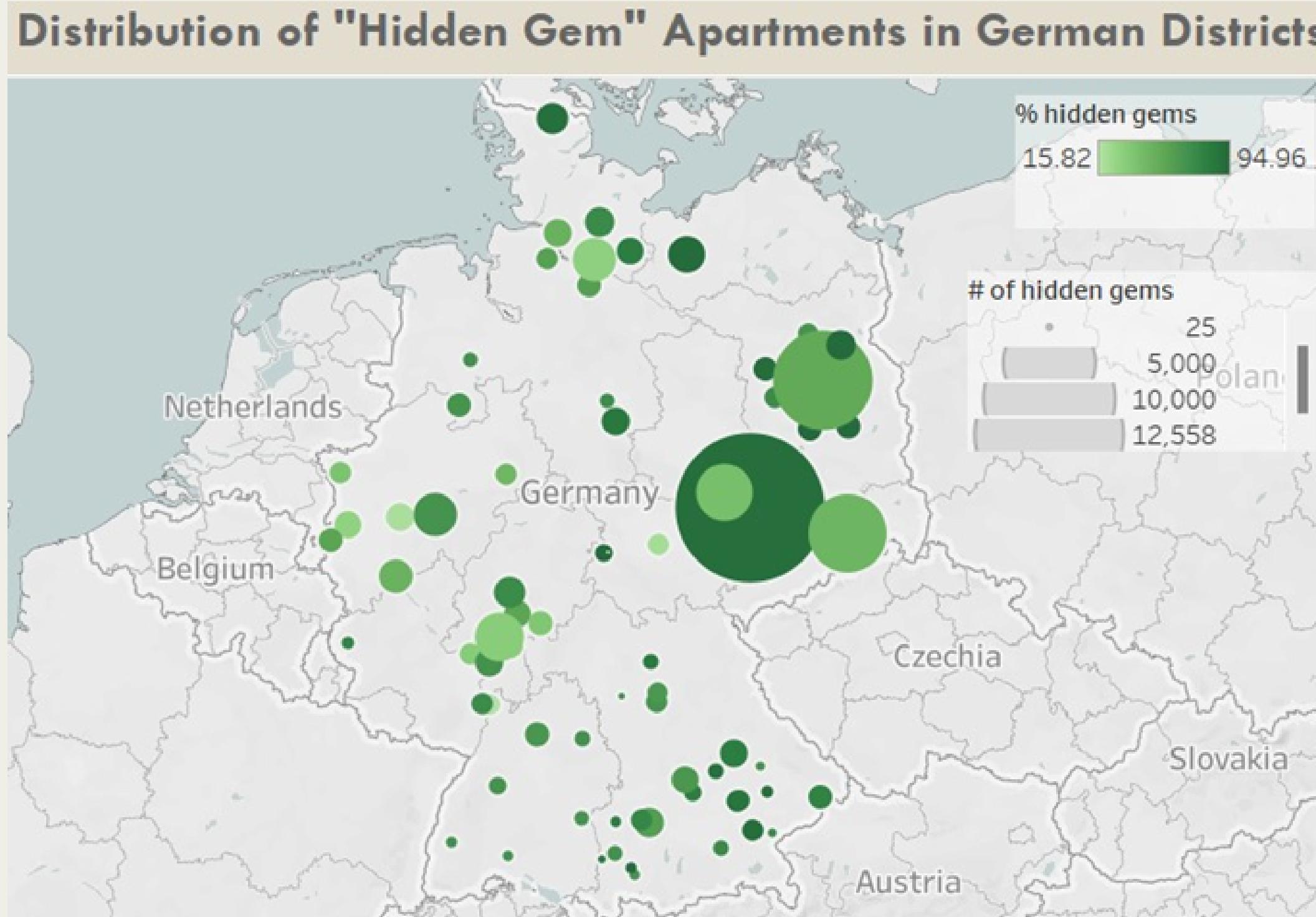
# Mapping the data

**Challenge:** Mapping median rents throughout Germany would explain another dimension of the data, but accurate shapefiles of the country's 400 districts are hard to come by.



**Solution:** I found the most current district-level shapefile I could and meticulously updated the punctuation, capitalization and naming conventions in the rent data to match it. The result: a high-level map illustrating the contours of rent variation in Germany.

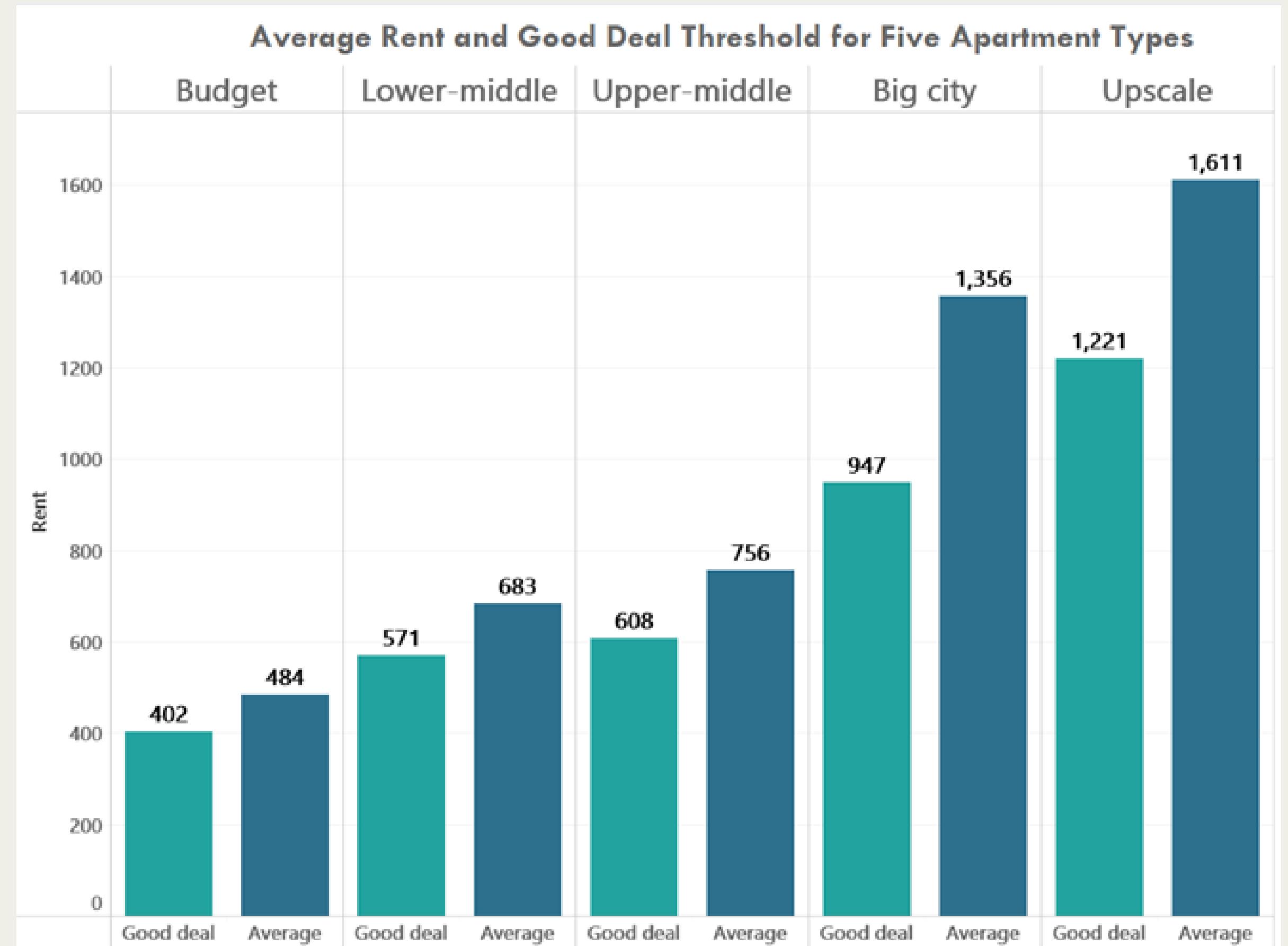
# Results: Turning insights into tools



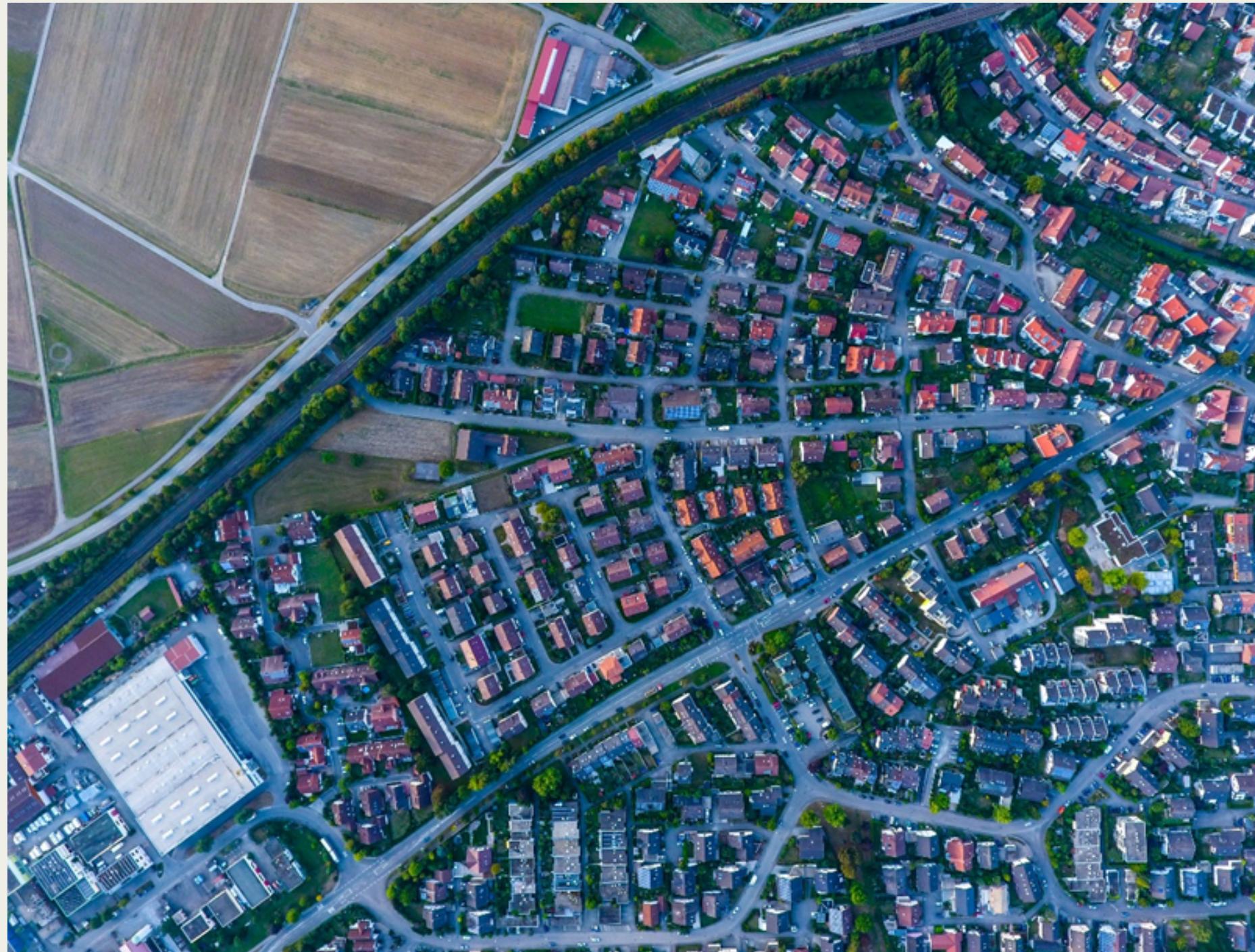
- The correlation between rent and local population growth ( $r = .61$ ) was second only to the correlation between rent and apartment size.
- To map exceptions to population growth correlation, I created the “hidden gem” statistic to measure unit affordability against local population growth.
- I used the statistic to show renters where to find good deals on rent in high-demand destinations.

# Results continued

- I used the rent statistics for each apartment personality to define a good deal for each type – half a standard deviation below the average rent.
- These parameters help apartment hunters target rents relative to their personal preferences.



# Next steps



- Scrape newer listings to reflect more recent rent trends.
- Refine the hidden gem variable to reflect attributes renters would want in an apartment and ignore population increases from district boundary changes.
- Define the apartment clusters in more detail so renters can take an “apartment personality quiz” and get personalized rent ranges and apartment-hunting tips.
- Map rents by PLZ for more granular geographic analysis.

Credit: Max Bottinger on Unsplash