Exploration of Tübingen's rental market

Aline Breitinger [*1] **Okan S. Coskun** [*2] **Lili Even** [*3] **Lukas M. B. Seehuber** [*4]

Abstract

Navigating Tübingen's rental market for a suitable residence within budget is challenging. This report analyses the market using the *VALUE Marktdaten* market data, exploring geographic and temporal price variations while validating their general upward trend. Linear regression and statistical validation, combined with visualizations, uncover district-wise trends in the city and how students influence monthly market patterns. The goal is to provide insights about Tübingen's dynamic rental landscape.

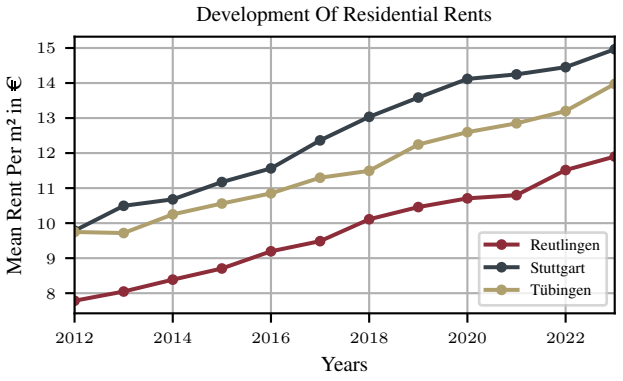

Figure 1. Comparison of the mean cold rent per m^2 in Tübingen to the rent in Reutlingen and Stuttgart.

1. Introduction

The cost of renting in Germany has increased throughout the years (Flemisch et al., 2023). This development can also be noticed regionally. Figure 1, shows a steady rise in renting costs in Tübingen and it's two major neighbouring cities. Because of Tübingen's student dominated population (Dr. Rijkhoek, 2023), its market situation is especially interesting and the main focus of this analysis.

This report explores the market using the *VALUE Marktdaten* dataset. To show an overall rent increase in Tübingen, we look at the price development over time in the districts (Section 4.1). Section 4.2 highlights the variance of apartment offers throughout the months of a year. Additionally, Section 4.3 reveals the market distribution and pricing trends for apartments based on their number of rooms.

[*]Equal contribution [1]Matrikelnummer 5451128, aline.breitinger@student.uni-tuebingen.de, MSc Computer Science [2]Matrikelnummer 6369042, okan.coskun@student.uni-tuebingen.de, MSc Computer Science [3]Matrikelnummer 5463873, lili.even@student.uni-tuebingen.de, MSc Computer Science [4]Matrikelnummer 6451084, lukas.seehuber@student.uni-tuebingen.de, MSc Computer Science.

Project report for the "Data Literacy" course at the University of Tübingen, Winter 2023/24 (Module ML4201). Style template based on the ICML style files 2023.

2. Data

The market data is provided by *VALUE Marktdaten* (Tzeggai et al., 2024a). After establishing contact with the company, a formal agreement was signed to secure access to the dataset encompassing real estate and rental offers from Tübingen, Reutlingen, and Stuttgart.

Stuttgart, Baden-Württemberg's state capital, has the highest amount of residents with approximately 610,000 people (Landeshauptstadt Stuttgart, 2023). Reutlingen has 117,000 residents (Stadt Reutlingen, 2023). The university city of Tübingen has a population of 92,800 residents (Universitätsstadt Tübingen, 2024a). According to (Dr. Rijkhoek, 2023) around 30% of Tübingen's population consists of students. The 20 to 35 age group is therefore more prominent in Tübingen compared to the whole state (Universitätsstadt Tübingen, 2023).

The Tübingen dataset includes data from January 1, 2012 to December 31, 2023, with 16,160 entries — 5,205 for buying and 10,955 for renting. Out of the rental offers, flats dominate with 10,668 entries, forming the focus of our report.

The provided dataset, curated by *VALUE Marktdaten* through a comprehensive quality control process, comprises 86 features related to the real estate market. Detailed information about all features can be found on their website (Tzeggai et al., 2024b).

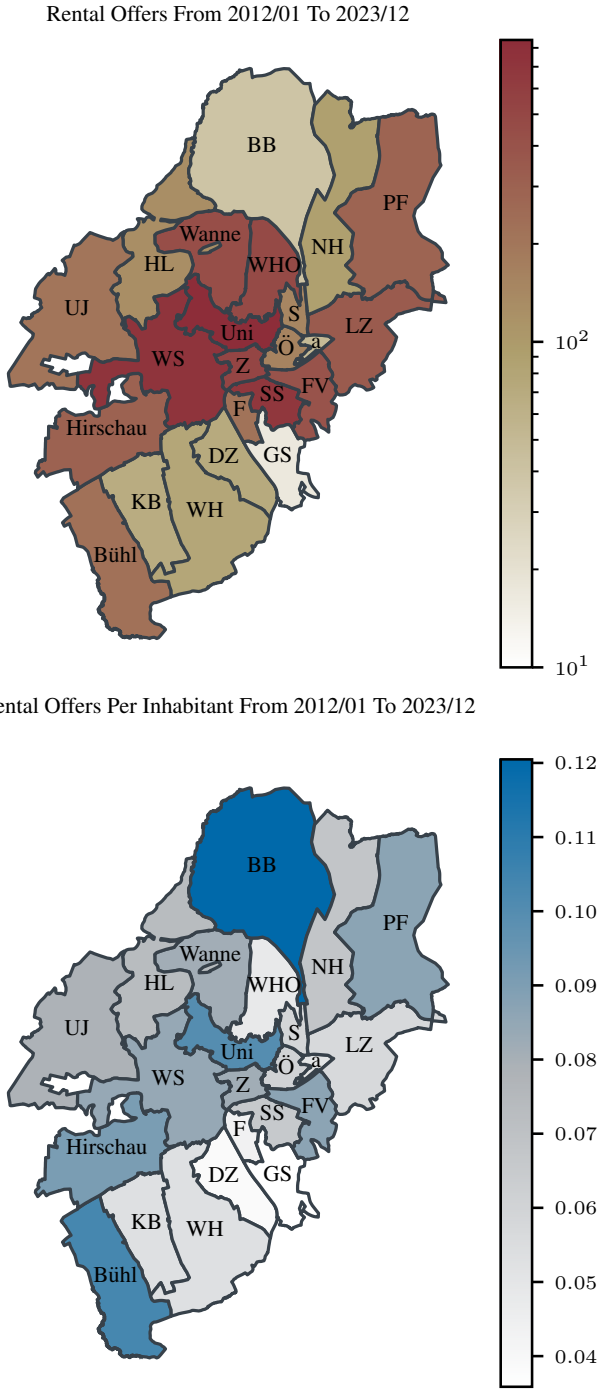

Figure 2. (Top) Distribution of rental apartment offers over Tübingen's municipal area (Universitätsstadt Tübingen, 2024b). *(Bottom)* Rental offers in context of the mean local population from 2016 to 2022 given by the (Universitätsstadt Tübingen, 2024a). District abbreviations: Aeule: **a**, Bebenhausen: **BB**, De-Zentrum: **DZ**, Feuerhägle/Mühlenviertel: **F**, Au/Unterer Wert/Französiches Viertel: **FV**, Gartenstadt: **GS**, Hagelloch: **HL**, Kilchberg: **KL**, Lustnau-Zentrum/Herrlesberg/Stäudach: **LZ**, Neuhalde: **NH**, Österberg/Gartenstraße: **Ö**, Pfrondorf: **PF**, Denzenberg/Sand: **S**, Südstadt: **SS**, Unterjesingen: **UJ**, Universität: **Uni**. Schönblick/Waldhäuser Ost: **WHO**, Weilheim: **WH**, Weststadt: **WS**, Zentrum: **Z**.

Since the data is based on rental offers, the amount of information provided for each dataset entry depends on the published details of an offer. Not every feature is relevant for our analysis, hence we focus on specific characteristics that are considered interesting and have sufficient data. These characteristics are mentioned in Section 3 when used.

To ensure a more accurate rental trend, we select the cold rent per m^2 feature as our primary metric. It excludes additional costs like electricity or gas, which could be influenced by external and unknown factors.

Due to privacy concerns, precise listing addresses are not available to us. In place of the address, we use the district location feature. This is the most precise location information for an offer that we can extract. Offers not assigned to a district are excluded from this analysis.

Figure 2 reveals that the provided dataset has the most offers in the city center. Some districts have very few offers, making it difficult to provide accurate statements about them. This is further validated by the bottom heatmap (Figure 2), which shows an uneven proportion of offers per inhabitant, especially in districts with a lower number of total offers.

3. Methods

Having established that Tübingen follows the overall trend of rising rents, we compare the individual rental price trends of Tübingen's 23 districts using linear regression.

The slopes and intercepts of the district-wise trends are calculated using ordinary least squares, minimizing the sum of squared residuals Q as described by Stahel (2018):

$$Q(a,b) = \sum_i (y_i - (a + bx_i))^2.$$

The t-test assesses the significance of the slope in our linear regression model. Let the significance level of the calculated slope be $\alpha = 0.01$. We are prioritizing the prevention of a type 1 error, because the slope should not be accepted as meaningful when it is not. The test is defined by H_0: $b = 0$ and H_1: $b \neq 0$, with b as the slope of the linear regression. The test statistic T is given by

$$T = \frac{\hat{b} - 0}{se \div \sqrt{\sum (x - \bar{x})^2}},$$

where the residual standard error se is defined as

$$se = \sqrt{\frac{1}{n-2} \sum (y_i - \hat{y_i})^2}.$$

The distribution of the test statistic under H_0 is defined as $T \sim t_{n-2}$ with $n-2$ degrees of freedom. After calculating T, its p-value is derived to check if it is smaller than α. The test is two-sided to ensure that both negative and positive

trends are considered (Hartmann et al., 2023). For calculating the above formulae, we use the Python package *scipy* (The SciPy community, 2024).

Further exploration focuses on the market demand during the year, discovering Tübingen related patterns. This is done by analyzing all of Tübingen's offers based on their month of publication and removal. After validating the homogeneous distribution of our data over the time span, the average over all 12 years is taken to avoid year specific anomalies. This yields information about the added and removed offers on the market in each month.

Section 4.3 examines price trends segmented by the number of rooms in an apartment. This is compared to Stuttgart and Reutlingen, considering the different market shares of these apartment types.

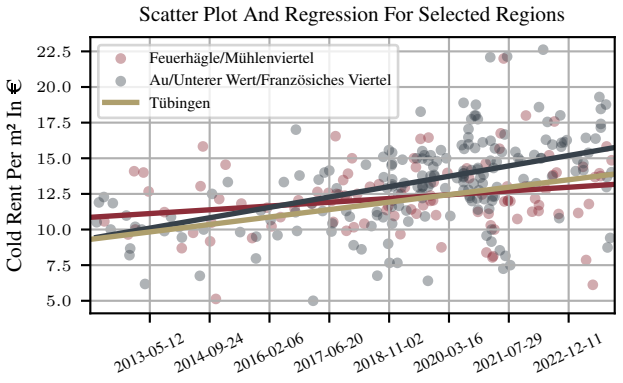

Figure 3. Development of basic rent per m^2 over the years. The regression lines belong to the district with the smallest and the largest increase, as well as to the overall increase in Tübingen. The scatterplot visualizes every second offer in the respective district.

District	R^2	p-Value	Price
Unterjesingen	0.288	2.6e-16	0.331
Hirschau	0.290	6.9e-24	0.377
Südstadt	0.209	7.4e-36	0.416
Pfrondorf	0.272	7.1e-22	0.327
Zentrum	0.172	1.2e-22	0.388
Weststadt	0.167	4.3e-30	0.318
Bühl	0.305	2.9e-18	0.449
Wanne	0.237	2.0e-27	0.364
Lustnau-Zentrum/ Herrlesberg/Stäudach	0.212	9.8e-19	0.427
Hagelloch	0.327	8.1e-13	0.426
Universität	0.136	6.4e-28	0.310
Österberg/Gartenstraße	0.202	8.6e-09	0.339
Feuerhägle/Mühlenviertel	0.057	5.0e-04	0.193
Weilheim	0.291	5.9e-07	0.320
Kilchberg	0.477	6.9e-12	0.377
Bebenhausen	0.420	2.0e-06	0.377
Schönblick/Waldhäuser Ost	0.184	5.1e-22	0.339
Denzenberg/Sand	0.149	2.4e-06	0.356
Au/Unterer Wert/ Französisches Viertel	0.264	4.5e-27	0.537
De-Zentrum	0.268	7.1e-06	0.512
Aeule	0.080	5.2e-02	0.318
Neuhalde	0.139	4.3e-04	0.473
Gartenstadt	0.185	9.6e-02	0.552

Table 1. Results of the linear regression model for the districts of Tübingen. The slope is converted to the price increase per year of the cold rent per m^2 by multiplying it with 365 days, neglegting leap years for simplicity.

4. Results

In the subsequent sections, we detail the outcomes of the above mentioned analyses and explorations. Additional information and insights are available on our dedicated Github repository.

4.1. Regional price developments over time

Across all observed areas, a consistent upward trend in rent is visible in Table 1. In the Feuerhägle/Mühlenviertel district, the rent experiences a modest increase of 0.19 Euro per m^2 annually. Even more substantially, Au/Unterer Wert/Französisches Viertel stands out with a higher annual increase of 0.54 Euro per m^2. In the remaining areas, the rent increase falls between these two extremes.

To put these values into perspective, Figure 3 visualizes their deviation from the overall trend in Tübingen. The t-test is employed to verify the robustness of our findings. As mentioned early in the paper, there are regions with insufficient data. Namely Auele and Gartenstadt, which show a significantly higher p-value compared to the remaining districts. For them, the H_0 cannot be rejected, so we cannot justify statements about their rental price trends.

4.2. Offered apartments during the year

By segmenting the offers as shown in Figure 4, it can be seen that January and August have more offers added than removed. The opposite is true for April, October, and December. Using the same method for Stuttgart and Reutlingen, a different pattern is observed. In Tübingen, there are way more offers removed than added in April, which cannot be seen in Reutlingen. For Stuttgart, the highest market activity concentrates at the beginning and end of the year. For all cities, December is the month with the least activity.

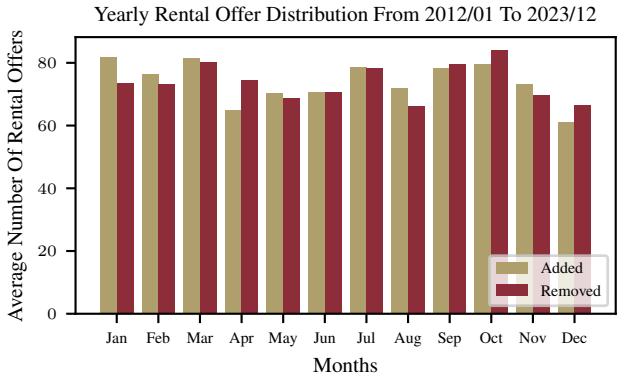

Figure 4. The average amount of offers added and removed over the last 12 years visualized per month. A listings can be in both bars of the same month if it is published and withdrawn in the same month. The median duration of an offer on the market is 11 days.

4.3. Rent trend by number of rooms

Comparing the rental market shares of different apartment types in Tübingen, Stuttgart and Reutlingen, the following can be observed. The share of one-room apartments in Tübingen is 21.9%. In contrast to that, the share in Stuttgart and Reutlingen is roughly 14%. This difference is balanced out by a share of 24% for three-room apartments in Tübingen compared to 35% in Reutlingen and Stuttgart. Interestingly, the trends shown in Figure 5 are however similar to those in the other two cities.

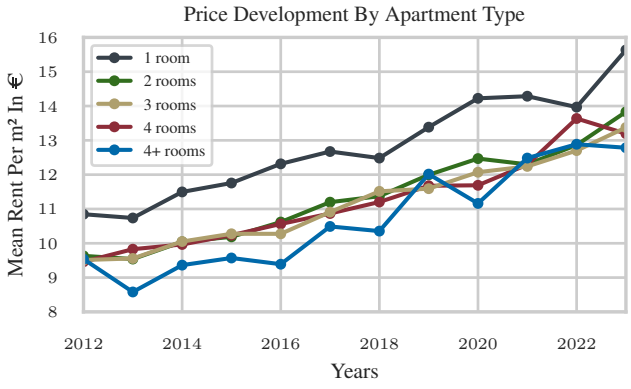

Figure 5. Rental prices segmented by the number of rooms per apartment. Comparing the market trends over time.

Throughout the 12-year period, apartments with only one room had the highest rent per m^2, making them proportionally the most expensive. Currently, the price per m^2 is lower the more rooms there are in an apartment.

5. Discussion & Conclusion

Focusing on Tübingen's districts, Figure 2 shows a higher density of offers in the city centre. Motivated by this finding, we tracked the development of each region individually to gain deeper insights into Tübingen's market. The Figure 3 shows significant price deviations, which contribute to a relatively low R^2 value. Nevertheless, the aim was to provide an insight into the evolution of rents instead of modelling the exact prices, which we can certainly do by rejecting H_0 for all regions except Auele and Gartenstadt.

However, this deep dive results in information loss, as not every listing is equally detailed regarding its location. Our ratio metric gets more sensitive with decreasing population. This leads to outliers like Bebenhausen, which stands out for having the highest ratio between listings and population. An anomaly can be seen in Schönblick/Waldhäuser Ost. Here, the dataset lacks offers due to the high amount of residence halls, which are distributed directly by the *Studierendenwerk Tübingen-Hohenheim*.

The bar chart (Figure 4) reveals a notable shortage of housing options in April and October, which corresponds to the beginning of the semester. The low market activity in December, that all three cities have in common, can be explained by the holiday season.

Since Tübingen is a university city, there is a more frequent change of tenants due to the high amount of students. Frequent tenant turnover enables more frequent rent increases. These are unrestricted by the 'Capping limit' that limits rent increase to under 15% within a 3-year period. For a new renting contract, the increase of rent is only capped by the average rent in the district, regulated by rent control (Universitätsstadt Tübingen, 2024c).

Despite the dissimilar distribution of apartment types, as described in Section 4.3, we examine no significant difference in the development of cold rent between the cities. This suggests that the market is adjusted to supply and demand. Therefore, the persistent high demand, driven by students, exerts a significant influence on rental costs in Tübingen.

The last two findings combined deliver an explanation on the rising rental trend in Tübingen and why it is more expensive compared to its neighbour Reutlingen. Hence it is likely for Tübingen, characterized by its students and university, to continue its rising rental trend.

Contribution Statement

Aline Breitinger and Lili Even negotiated access to the data and conducted research. Okan Coskun and Lukas Seehuber produced visualizations and kept contact with *VALUE Marktdaten*. All authors jointly conducted the data exploration and wrote the text of the report.

References

Tamara Flemisch, Hendrik Lehmann, Lennart Tröbs, and Helena Wittlich. *Mieten steigen ortsweise um 100 Prozent, Kaufpreise um 200 Prozent*, November 2023. URL https://interaktiv.tagesspiegel.de/lab/mietsteigerungen-wohnungskrise-in-welchen-deutschen-staedten-die-mieten-und-kaufpreise-besonders-stark-steigen/.

Karl Guido Dr. Rijkhoek. *Universität Tübingen verzeichnet Rekordzahl bei Studierenden*. November 2023. URL https://uni-tuebingen.de/universitaet/aktuelles-und-publikationen/pressemitteilungen/newsfullview-pressemitteilungen/article/universitaet-tuebingen-verzeichnet-rekordzahl-bei-studierenden/.

Stefan Tzeggai, Andreas Vater, and Sebastian Hein. *Real estate market data for Germany - Value AG the valuation group*, 2024a. URL https://www.value-marktdaten.de/.

Statistisches Amt Landeshauptstadt Stuttgart. *Informationen zur Einwohnerentwicklung*, December 2023. URL https://www.domino1.stuttgart.de/web/komunis/komunissde.nsf/fc223e09e4cb691ac125723c003bfb31/4e6620005e228f76c12584d300483532/$FILE/bs701_.PDF.

Kommunale Statistikstelle Stadt Reutlingen. *Einwohner im Dezember 2023*, December 2023. URL https://www.reutlingen.de/de/Leben/Unsere-Stadt/Daten-Fakten/Einwohnerzahl.

Universitätsstadt Tübingen. *Bevölkerungszahlen: Entwicklung der Gesamtbevölkerung*, January 2024a. URL https://www.tuebingen.de/1370.html.

Universitätsstadt Tübingen. *Sozialbericht 2023*, September 2023. URL https://www.tuebingen.de/verwaltung/uploads/sozialbericht_2023.pdf.

Stefan Tzeggai, Andreas Vater, and Sebastian Hein. *VALUE Market Database*, 2024b. URL https://www.value-marktdaten.de/en/portfolio/market-database/.

Universitätsstadt Tübingen. *Geodaten*, January 2024b. URL https://www.tuebingen.de/33688.html.

Werner A Stahel. *Statistische Datenanalyse: Eine Einführung für Naturwissenschaftler*. Vieweg Sohn Verlag, 2018.

K. Hartmann, J. Krois, and A. Rudolph. *Statistics and Geodata Analysis using R (SOGA-R)*, 2023. URL https://www.geo.fu-berlin.de/en/v/soga-r/Basics-of-statistics/Hypothesis-Tests/Inferential-Methods-in-Regression-and-Correlation/Inferences-About-the-Slope/index.html.

The SciPy community. *SciPy v1.12.0 Manual: linregress*, January 2024. URL https://docs.scipy.org/doc/scipy-1.12.0/reference/generated/scipy.stats.linregress.html.

Universitätsstadt Tübingen. *Ist die Miete angemessen?*, 2024c. URL https://www.tuebingen.de/26317.html.