

# German Job Market 2025 Analysis

1. Project Goal and Objectives .....	1
2. Data Sources .....	1
Data Set 1:.....	1
Data Set 2:.....	2
3. Research Questions .....	2
German Job Market (Dataset 1): .....	2
Data Science Deep Dive (Dataset 1 → Dataset 2):.....	3
4. Cleaning Methods and Data limitations .....	3
Data Limitation: .....	3
Cleaning Methods: .....	3
5. Descriptive Statistics and Correlation Metrics.....	4
Results 1: Job Market in Germany (dataset1) .....	4
Demand & Competitiveness: .....	4
Salary variances:.....	5
Figure 2: Salary Variances:.....	5
Qualifications: .....	6
Flexibility & Pay:.....	6
Figure 4: Flexibility & Pay: .....	6
Seasonality & Forecasting.....	7
Figure 5: Seasonality & Forecasting: .....	7
Results 2: Data Science Deep Dive (Dataset2) .....	8
Data Science Salary variances: .....	8
Results 3: Job Market Prediction models .....	8
Regression-focused: .....	9
Median Salary Regression Conclusion of four tested variables: .....	9
Regression Salary Median vs. Number of Openings.....	11
Classification-focused: .....	11

## 1. Project Goal and Objectives

**Project Goal:** To provide a strategic analysis of the German job market for 2025, identifying key trends in salary, demand, and hiring patterns. Special look at the German employment trend from 2010 to 2025.

**Project choice reason:** Personal interest in this job market analysis and future demand, also i found the data set is sufficient to practice data analysis and predictive models.

### Objectives (Key Deliverables):

1. Identify the highest-demand industries, job titles, and locations in the overall German job market.
2. Analyse salary distributions across different experience levels, industries, and regions (States/Cities).
3. Determine the impact of seasonality (quarter/month) and various factors (e.g., remote availability, education) on hiring job openings.
4. Provide actionable insights and data-driven recommendations for job seekers.

## 2. Data Sources

### Data Set 1:

- **Germany Job Market 2025 Dataset** by Kundan Sagar Bedmutha on Kaggle
- [!\[\]\(c6a8736a601a632e2c96605cf66055ed\_img.jpg\) Germany Job Market 2025 Dataset](#)

This dataset provides a **structured snapshot of the German job market for 2025**, including:

- Job titles, industries, salary ranges, experience requirements
- Geographic distribution (cities/states)
- Hiring intensity, demand index, remote work availability
- Seasonality (quarter/month) and education requirements

### Data Set 2:

German employment/unemployment numbers 2010-2025.

Source: Labour force survey (microcensus). Code: 13231-0001.

© Statistisches Bundesamt (Destatis), 2025 | created: 2025-12-10 / 15:34:14

**Data Content:** employment/unemployment numbers , Unemployment rate: Unemployed as a proportion of the economically active population (persons in employment + unemployed) of comparable age.

### 3. Research Questions

#### German Job Market (Dataset 1):

1. Demand & Competitiveness: Which industries, job titles, and cities show the highest demand and competitiveness?
2. Salary variances: How do median salaries vary across industry, job titles, cities based on experience levels?
3. Qualifications: What are the most common education and skills requirements for high-demand jobs?
4. Flexibility & Pay: What job titles offer remote availability? And how does remote availability affect salary offers? (remote vs on-site jobs)]
5. Seasonality & Forecasting: Which months/quarters show peak hiring activity, and what trends can be forecasted?  
→ [Regression/Time Series Forecasting: predicting number\_of\_openings over time]

#### German employment t (Dataset 1 → Dataset 2):

6. How is the German Employment is changing from 2010 to 2025. What aspects influence this change. What predictive model can be used for future predictions?

### 4. Cleaning Methods and Data limitations

#### Data Limitation:

**Dataset 1** Source is synthetic data designed by Kundan Sagar Bedmutha on Kaggle for the purpose of machine learning training.

**Dataset 2** no limitation to be mentioned for the purpose of this analysis.

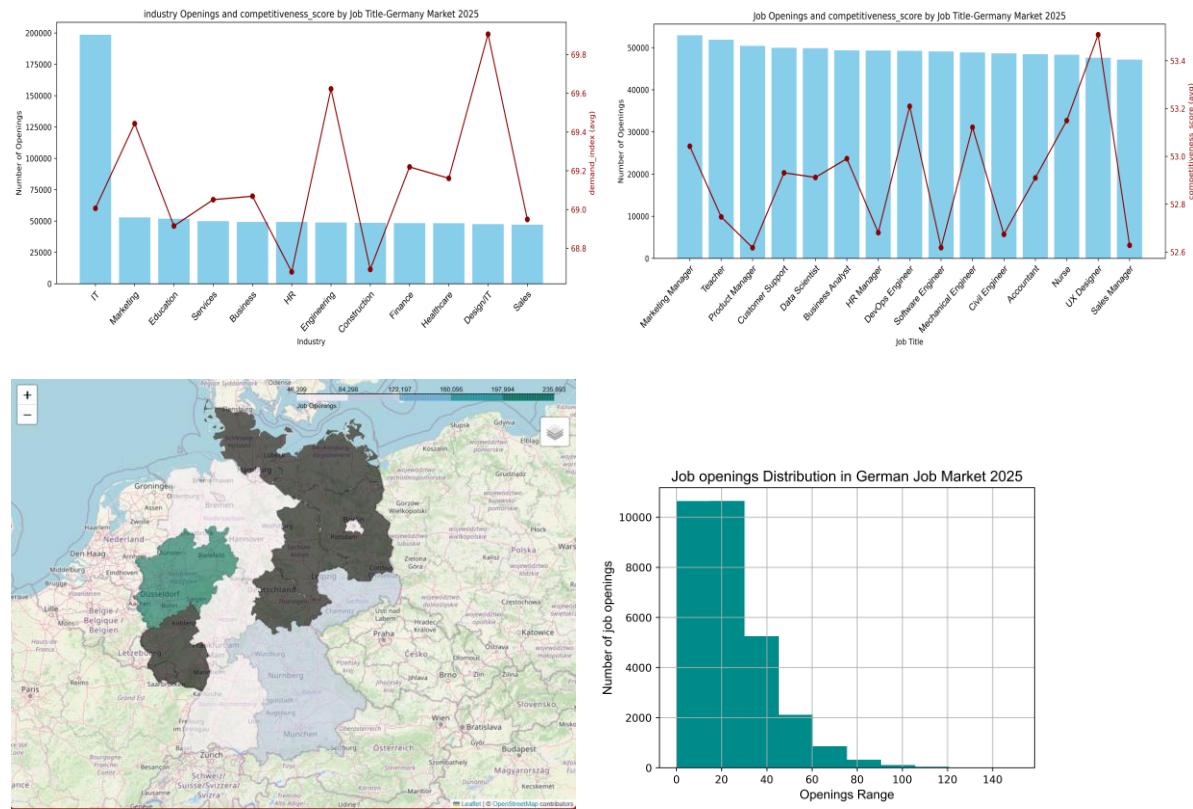
## Cleaning Methods:

	<b>German Job Market 2025 dataset 1</b>	<b>Data Science Global Market Dataset 2</b>
original no. of rows	30000	90
after cleaning no. of rows	30000	
removed columns	34-42 not needed and were removed	none
duplicates	none	1
Missing values	none	3
columns names change	none	none
data type change	none	None
New columns		Employment percentage calculated from total

## 5. Descriptive Statistics and Correlations

### Results 1: Job Market in Germany

**Demand & Competitiveness:** Which industries, job titles, and cities show the highest demand and competitiveness?



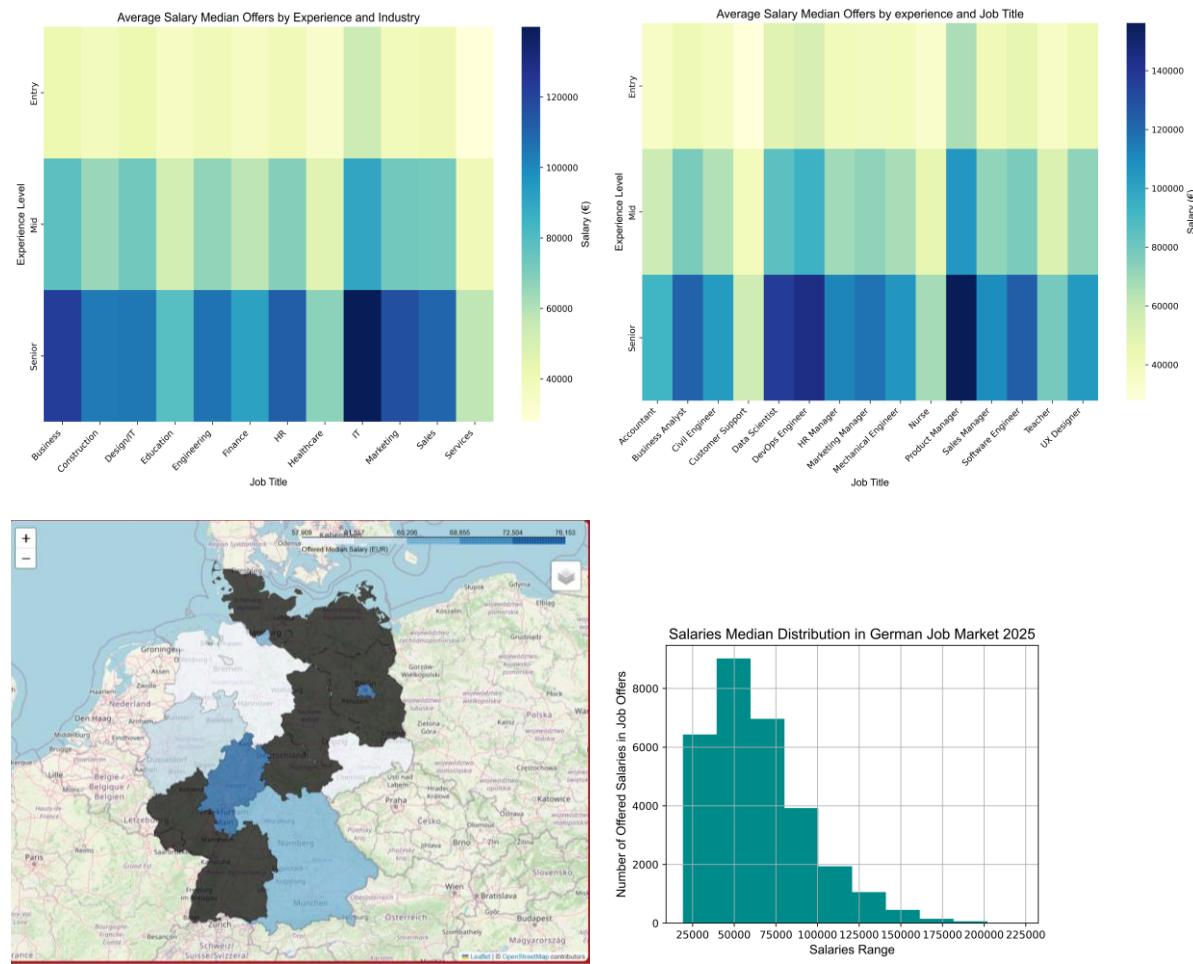
**Figure 1: Demand & Competitiveness:** **A.** Bar chart showing the number of job openings per industry in 2025 vs. the demand index line. **B.** Bar chart showing the number of job openings per job title in 2025 vs. the competitiveness score line. **C.** Geospatial distribution of the job openings among the German cities. **D.** Median Openings histogram showing left skewed distribution.

### Results Description: Demand & Competitiveness

Number of job openings across industry reveals that industries such as IT and Engineering are leading in job openings, accompanied by a high demand index. Job titles like UX Designer and DEV OP Engineer show the highest competitiveness scores, reflecting a

strong talent competition for these in-demand positions. While no difference in the number of job openings between job titles. a geospatial view of the number of job openings across the German states, illustrates that job openings are concentrated in North Rhine-Westphalia followed by Bavaria. **Subfigure D** shows a histogram of median openings, indicating a left-skewed distribution, where most job openings are clustered at the lower end, with fewer industries offering a very high number of openings.

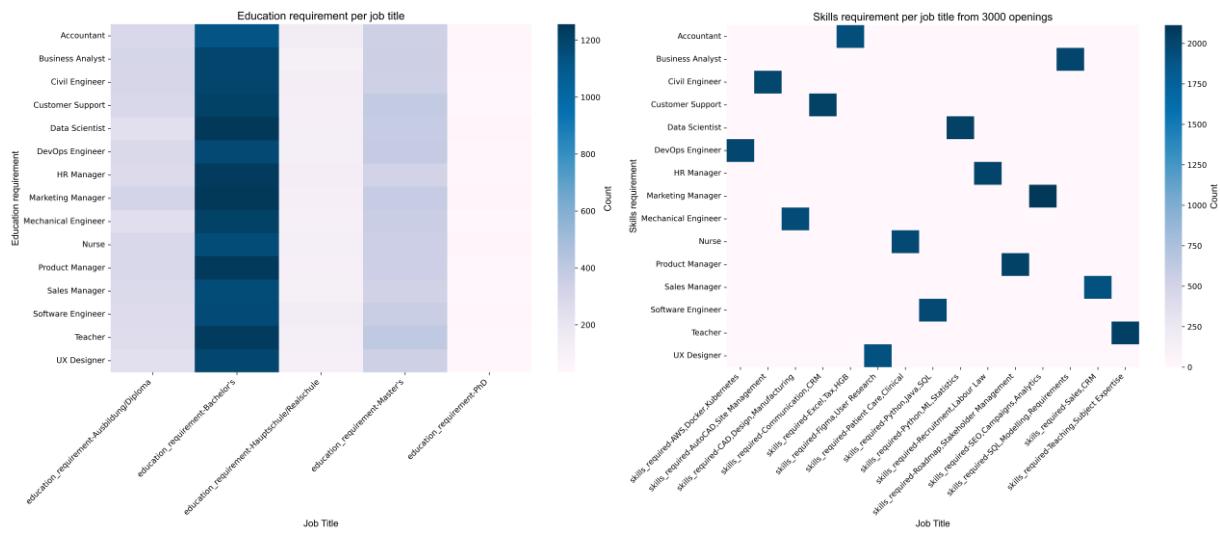
## Salary variances: How do median salaries vary across industry, job titles, cities based on experience levels?



**Figure 2: Salary Variances:** **A.** Heatmap showing the offered Average Salary Median across industries sorted by Experience Level. **B.** Heatmap showing the offered Average Salary Median across job titles sorted by Experience Level. **C.** Geospatial distribution of the offered salaries among the German cities. **D.** Median Salaries histogram showing left skewed distribution.

**Results description: Salary Variances:** Median salary analysis across industries shows a clear upward trend with increasing experience level, reaching the highest values at the senior level. Among industries, IT offers the highest salaries, followed by the business sector. Across job titles, the same pattern is observed, with Product Manager, Development Operations Engineer, and Data Scientist roles commanding the highest median salaries. The geospatial distribution highlights that Berlin and Hessen provide the most competitive salary packages, followed by Bavaria and North Rhine-Westphalia, reflecting the concentration of high-paying opportunities in major economic hubs. Subfigure D shows a left-skewed histogram of median salaries, indicating that while a few industries and regions offer very high salaries, most positions cluster around moderate ranges.

## Qualifications: What are the most common education and skills requirements for high-demand jobs?

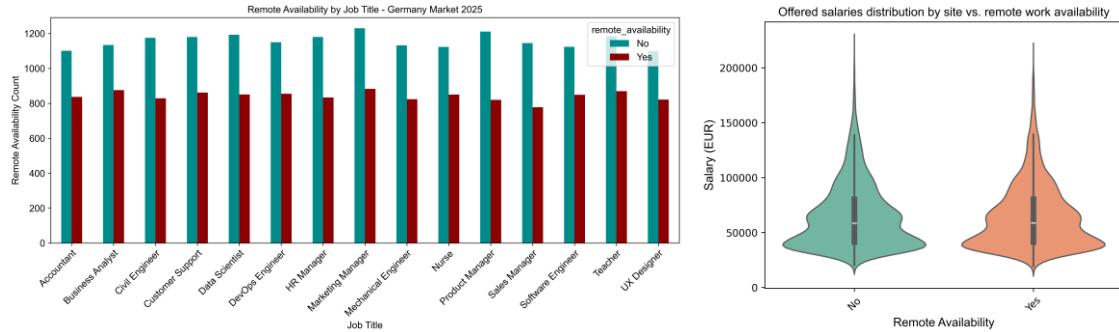


**Figure 3: Qualifications:** A. Heatmap showing the required education for each job title. B. Heatmap showing the most required skills package for each job title.

## Results description: Qualifications

The most required education level is Bachelor's followed by Masters among all job roles. While the required skills package are being unique to each job title.

**Flexibility & Pay:** What job titles offer remote availability? And how does remote availability affect salary offers?

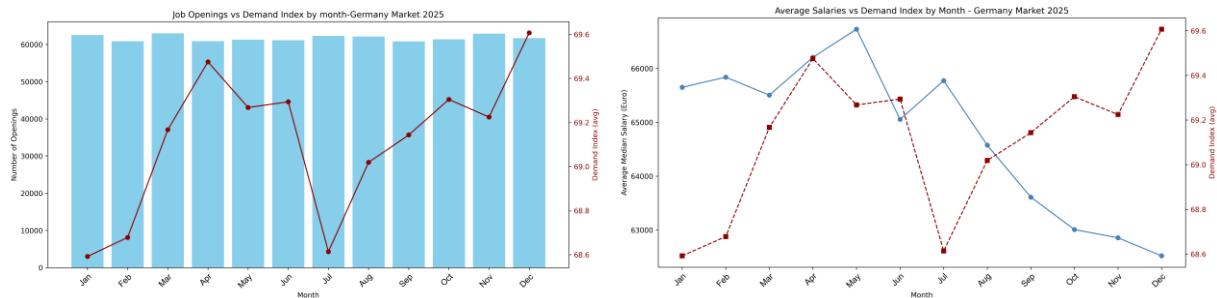


**Figure 4: Flexibility & Pay:** **A.** Bar chart showing the job titles that offer remote availability. **B.** Violin chart illustrating the Offered salary distribution by site vs. remote work availability.

### Results description: Flexibility & Pay:

The German job market shows more job offers with in/presence than remote, but this has no effect on the offered salary.

**Seasonality:** Which months/quarters show peak job openings, and what trends can be forecasted?



**Figure 5: Seasonality & Forecasting:** **A.** Bar chart showing the job titles that offer remote availability. **B.** Line chart illustrating the Offered Average monthly salaries over 2025 year.

### Results description-Seasonality:

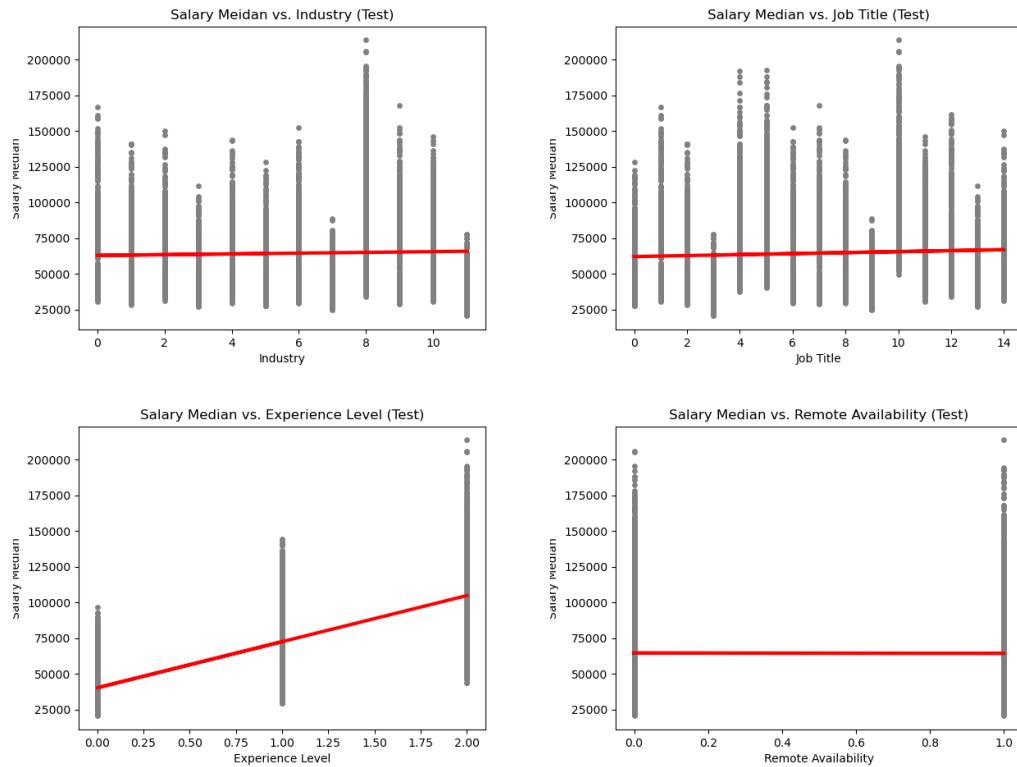
One year of data cannot show seasonality in general, but we see seasonal trend in the demand index that is very low in January, and July, but very high in December and April. While there is no significant difference of the job openings numbers between the months.

While we see decreasing trend of offered salary from August to December but this decrease is not significant and no enough information about seasonality.

## Results 2: Job Market Prediction models

### Regression-focused:

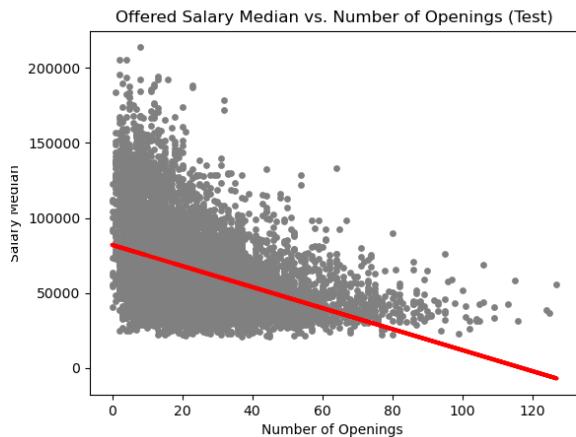
1. Can we predict the median salary of a job posting based on job title, experience level, industry, and remote availability?"



### Median Salary Regression Conclusion of four tested variables:

Among the four variables tested, **experience level** is the only strong predictor of salary, explaining nearly 57% of the variation and showing a clear positive relationship. In contrast, **industry** and **job title** have almost no explanatory power, with slopes close to zero and  $R^2$  values near zero. **Remote availability** even shows a weak inverse relationship with salaries and a negative  $R^2$ , making it an unreliable predictor. Overall, salary differences in the German data science job market are driven primarily by **experience**, while the other encoded factors contribute little to meaningful prediction.

## 2. Does the number of openings correlate with the offered median salaries?



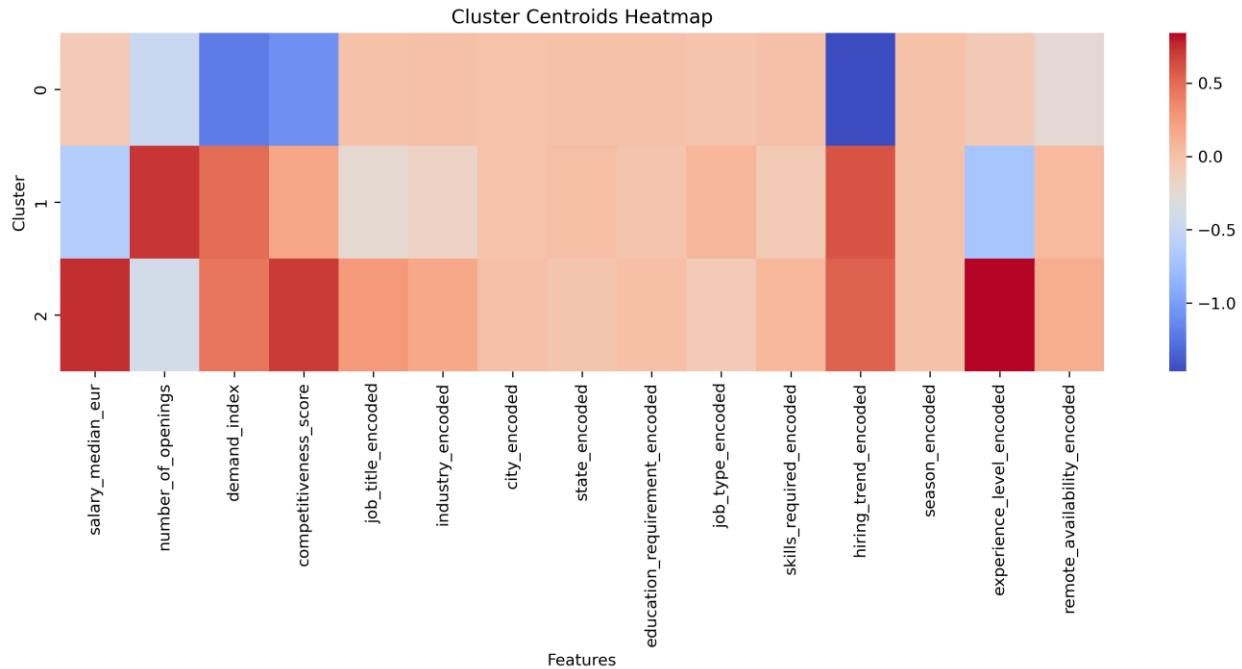
### Regression Salary Median vs. Number of Openings

The slope is negative suggesting a reverse relationship. -700.11 For every additional job opening, the model predicts that the median salary decreases by about €700. This suggests that when there are more positions available, salaries tend to be lower. R<sup>2</sup> Score: 0.164 (16.4%) The number of openings explains only about 16% of the variation in salaries i.e., not strong correlation. Most of the variation is due to other factors.

Note: removing the extreme values in the two variables has distorted any possible correlation, therefore, no regression was performed.

## 3. Correlation of Hiring index and competitiveness' score with number of openings and salaries median: Results show very weak correlation, i decided not to show it here as i did it for practicing and not for providing an answer to a logic question.

## Classification-focused: unsupervised learning method



## German Job Market 2025 – unsupervised Classification Analysis

In analyzing the German job market of 2025, a heatmap was created using **K-Means clustering**, which grouped the data into three distinct clusters. Each cluster highlights different characteristics of the labor market:

### Cluster 0

This cluster reflects a segment with **low hiring trends, low demand index, low competitiveness scores, and relatively fewer job openings**. It represents areas of the market where opportunities are limited and competition is weak.

### Cluster 1

Here we see jobs with a **relatively low median salary and lower experience requirements**, but at the same time, a **high number of job openings, strong hiring trends, and a high demand index**. This suggests that while compensation is modest, opportunities are abundant and demand is strong.

### Cluster 2

This cluster is characterized by **very high median salaries, high experience levels, and high competitiveness scores**, combined with **strong hiring trends and high demand index**. However, it shows a **relatively lower number of job openings**, indicating that these positions are lucrative but scarce and highly competitive.

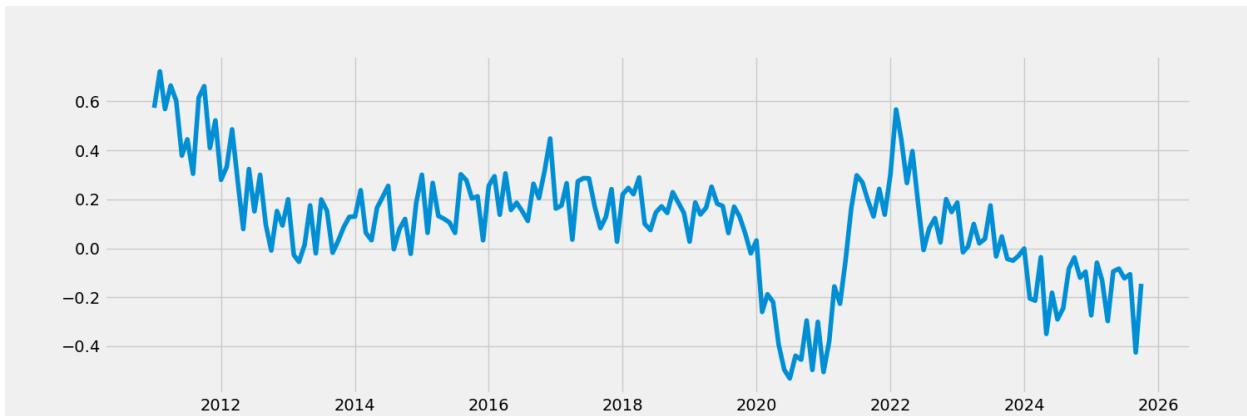
### Overall Insight

The analysis reveals that **job type, job title, industry, city, state, education level, and remote availability do not significantly influence the number of job openings across these clusters**. Instead, the clusters are primarily differentiated by salary levels, experience requirements, competitiveness, and demand dynamics.

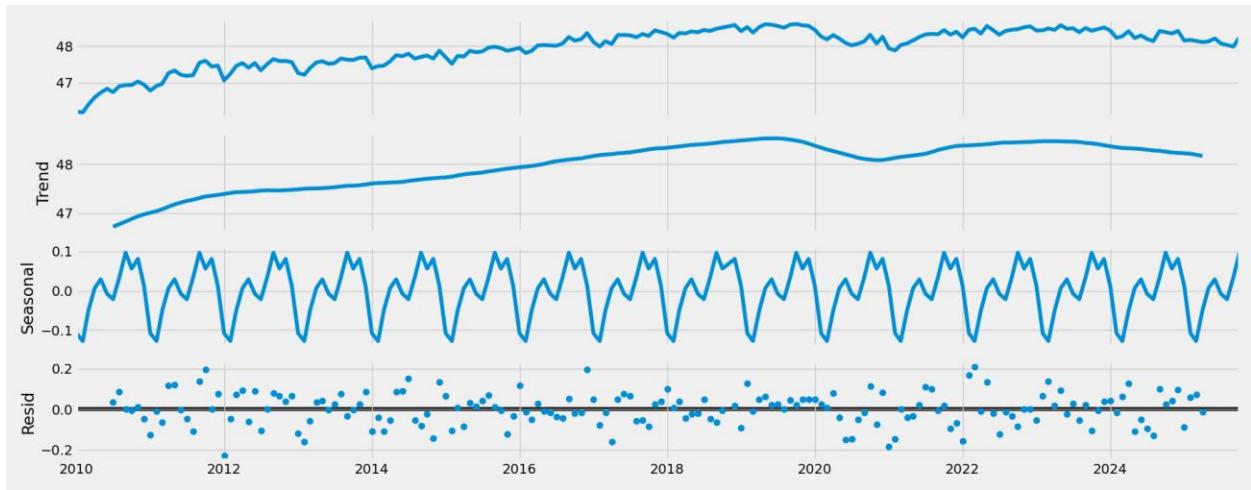
## Time series Analysis of German employment 2010-2025



German employment time series shows uprising trend over the years with one drop in 2021. Statistical exam shows non-stationary data.



German employment time series after 12 month lag seasonal differencing shows more stable trend over the years with the same drop in 2021. Stationarity has been reached significantly.



Decomposition of the German employment percentage before differencing. Statistical exam shows non-stationary data.