Received: September 2025, Accepted: November 2025, Published: December 2025
Digital Object Identifier: https://doi.org/10.34302/CJEE/MEIZ3377

A DATA-DRIVEN ANALYSIS OF REMOTE WORK SALARIES AND SATISFACTION

Ioan Alexandru **VLAD**, Dacian **TEPFENHART**
Technical University of Cluj-Napoca, Romania
al3xro@gmail.com, dacianXpaul@gmail.com

Keywords: remote work, salary analysis, data science, job satisfaction

Abstract: *This paper provides a data-driven analysis of remote work salaries, leveraging a real-world dataset processed with Python. The study investigates how industry, experience, employment type, and remote flexibility impact salary and job satisfaction. Insights are visualized using statistical plots and support a broader understanding of global salary trends in remote work environments.*

1. INTRODUCTION

The rapid adoption of remote work has significantly reshaped global labor markets, influencing both salary structures and employee satisfaction. Advances in digital technologies and increased workplace flexibility have enabled organizations to recruit talent globally, while employees benefit from improved work-life balance. Understanding how factors such as industry, experience level, and employment type affect compensation in remote environments has become increasingly important. This paper presents a data-driven analysis of remote work salaries and job satisfaction using a real-world dataset, aiming to identify key trends and relationships that characterize contemporary remote work dynamics.

2. METHODOLOGY – DATA SOURCES AND TOOLS

The present analysis is based on a dataset containing remote work salary data, processed and visualized using Python. The references used provided both a socio-economic context and methodological foundation for the analysis.

Context:

- Sources [1], [2], and [3] provide insight into remote work trends, global salary expectations, and worker preferences, which helped frame the broader relevance of the dataset.
- Economic relevance: The OECD report [4] highlights how the COVID-19 pandemic influenced remote work, justifying the need for salary structure analysis in this new normal.
- Technical tools: Python libraries such as Pandas [5], Matplotlib [6], and Scikit-learn [7] were used for data cleaning, transformation, statistical summary generation, and data visualization. These tools are recognized and validated in the field of data science.

3. CODE FOR ANALYSIS IN GOOGLE COLAB:

```
# Remote Salary Analysis - Google Colab Notebook
import pandas as pd import seaborn as sns import
matplotlib.pyplot as plt # Load dataset df =
pd.read_csv("/content/Work_From_Anywhere_Salary_Data.
csv")
# Preview print(df.head()) print(df.info())
# Descriptive statistics print(df.describe())
# Check for missing values print(df.isnull().sum())
# Salary distribution plt.figure(figsize=(10, 6)) sns.histplot(df['Salary (Annual)'], kde=True, bins=30,
color='skyblue') plt.title('Annual Salary Distribution') plt.xlabel('Salary (Annual)') plt.ylabel('Frequency')
plt.show()
# Boxplot - Salary by Experience Level plt.figure(figsize=(10, 6))
sns.boxplot(data=df, x='Experience Level', y='Salary (Annual)', palette='
    Set2') plt.title('Salary by Experience Level') plt.show()
# Average salary per industry industry_salary = df.groupby('Industry')['Salary (Annual)'].mean().
    sort_values()
plt.figure(figsize=(10, 8)) industry_salary.plot(kind='barh', color='teal')
plt.title('Average Salary by Industry') plt.xlabel('Average Salary (Annual)')
plt.ylabel('Industry') plt.show()
# Comparison: Remote vs Onsite plt.figure(figsize=(8, 6))
sns.boxplot(data=df, x='Remote Flexibility', y='Salary (Annual)', palette
    ='coolwarm') plt.title('Salary by Remote Flexibility') plt.show()
# Average job satisfaction by Tech Stack
tech_stack_satisfaction = df.groupby('Tech Stack')['Job Satisfaction
    Score (1-10)'].mean().sort_values() plt.figure(figsize=(10, 8))
tech_stack_satisfaction.plot(kind='barh', color='orange')
plt.title('Average Job Satisfaction by Tech Stack') plt.xlabel('Average
Satisfaction Score') plt.ylabel('Tech Stack') plt.show()
# Employment types distribution employment_type_counts = df['Employment Type'].value_counts()
plt.figure(figsize=(6, 6)) employment_type_counts.plot(kind='pie', autopct='%1.1f%%', startangle
```

```
=140) plt.title('Distribution of Employment Types') plt.ylabel('')
plt.show()
# Correlation matrix for numeric variables plt.figure(figsize=(10, 6))
sns.heatmap(df.corr(numeric_only=True), annot=True, cmap='coolwarm', fmt
    ='.2f') plt.title('Correlation Matrix') plt.show()
```

4. DATASET PREVIEW AND STRUCTURE

To understand the structure and context of the analysis, Table 1 presents the first five entries of the dataset. Each row represents a job position in a remote or flexible working environment, and each column captures a specific characteristic of the job or employee:

Table 1. Entries of the dataset

	Company	Job Title	Industry	Location	Employment Type
1	Microsoft	Data Analyst	Media	Austin	Part-time
2	Apple	Data Scientist	Retail	San Francisco	Part-time
3	Amazon	Software Engineer	Healthcare	San Francisco	Full-time
4	Tesla	Data Analyst	Retail	Austin	Contract
5	Adobe	DevOps Engineer	Healthcare	New York	Contract

This snapshot highlights the diversity in employer type, industry, geographical location, and employment contract. Notably:

Industry Variety: The dataset includes entries from Media, Retail, and Healthcare industries, allowing for cross-sectoral comparisons.

Employment Types: Full-time, part-time, and contract-based employment are all represented, offering insight into different compensation structures and job stability.

Geographic Distribution: The jobs are based in prominent U.S. tech hubs like San Francisco, Austin, and New York, which are known for high employment demand and competitive salaries.

Job Roles: Positions vary from Data Analysts to Software and DevOps Engineers, covering a wide range of technical expertise. This initial view, summarized in Table 2, supports the generalizability of the dataset and establishes a solid foundation for further exploratory and statistical analysis.

Table 2. Dataset structure and job characteristics

	Experience Level	Remote Flexibility	Salary (annual)	Currency
1.	Mid	Remote	155200.11	AUD
2.	Lead	Remote	106365.54	INR
3.	Lead	Remote	91026.49	INR
4.	Mid	Onsite	41824.38	EUR
5.	Senior	Remote	143929.78	USD

This subset illustrates how experience level and remote flexibility interact with salary levels and global currencies:

Experience Impact: Lead and Senior professionals command higher salaries on average than Mid-level employees, consistent with broader labor market trends.

Remote Flexibility: Positions marked as fully remote tend to be better compensated than onsite roles, even when accounting for differences in geographic location and currency.

... geographic and contractual flexibility. The salary distribution presented in *fig. 1* further highlights these variations across the dataset.

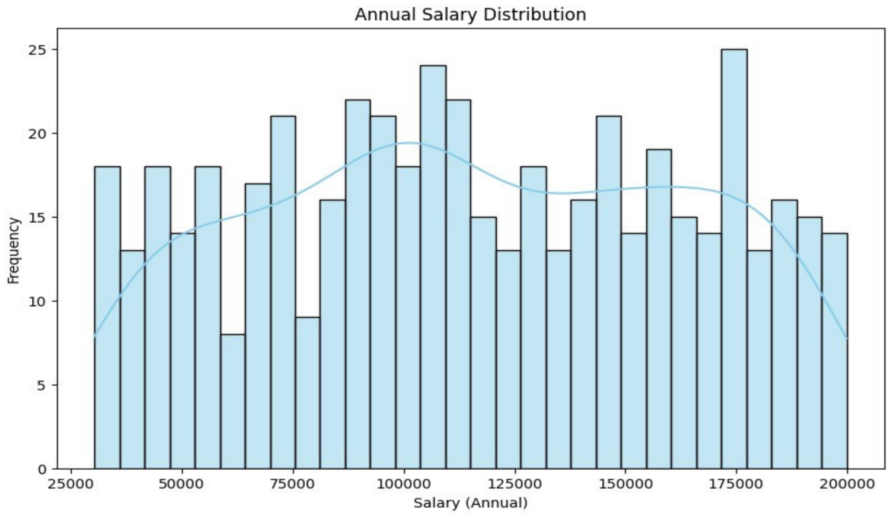

Fig. 1. Annual Salary Distribution (Histogram)

The salary distribution shown in *fig. 2* follows a slightly right-skewed curve, suggesting that while most employees earn within a moderate range, a minority earn substantially higher or lower salaries. These outliers may be attributed to senior roles, niche technical skills, or geographic salary differentials.

Fig. 2. Salary by Experience Level (Boxplot)

This boxplot compares annual salaries across four categories of experience level: Junior, Mid, Senior, and Lead. A positive correlation between experience and compensation is evident. While junior employees have more clustered earnings, higher-level professionals display a broader salary range, which may reflect diverse negotiation power, specializations, or company size.

The average annual salaries grouped by industry, illustrated in *fig. 3*, show that sectors such as Technology, Finance, and Healthcare offer higher average salaries, while Retail and Media tend to fall below the overall average. These findings highlight how industry type plays a crucial role in shaping remote salary structures.

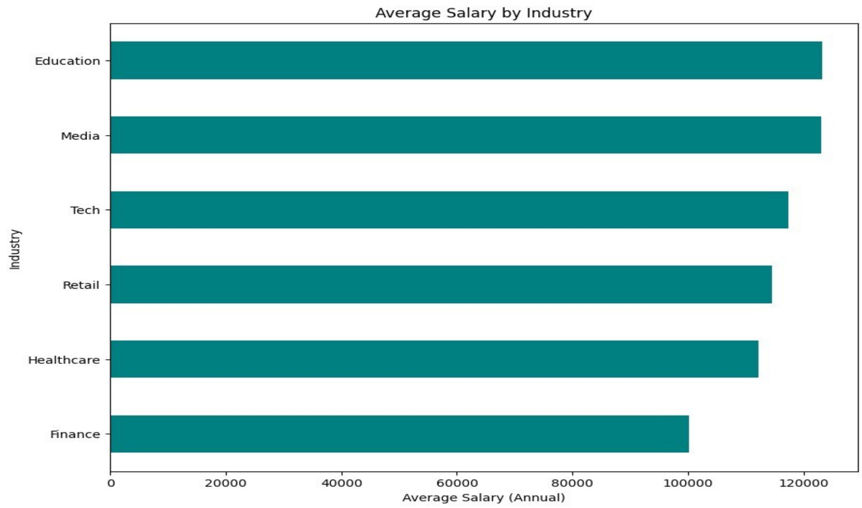

Fig. 3. Average Salary by Industry (Horizontal Bar Plot)

Salary variation by work arrangement type—Fully Remote, Hybrid, or Onsite—is illustrated in *fig. 4*. Fully remote roles demonstrate the highest median salary and wider distribution, which may result from global hiring practices and increased demand for flexible work. Onsite positions, by contrast, show more constrained salary ranges.

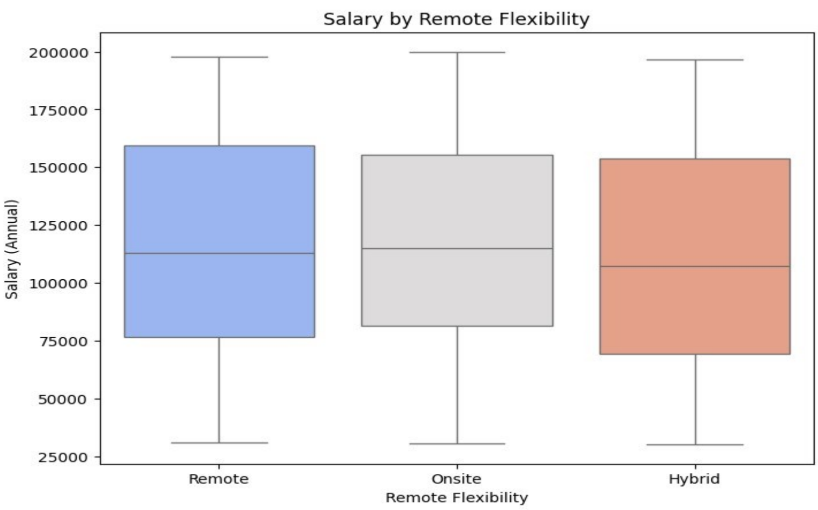

Fig. 4. Salary by Remote Flexibility (Boxplot)

Average job satisfaction scores by technological stack are illustrated in *fig. 5*. Employees using stacks such as Python, Go, and Kubernetes report significantly higher satisfaction levels, while roles based on older or more standardized stacks like JavaScript or .NET show lower satisfaction averages, potentially due to project maturity or limited innovation.

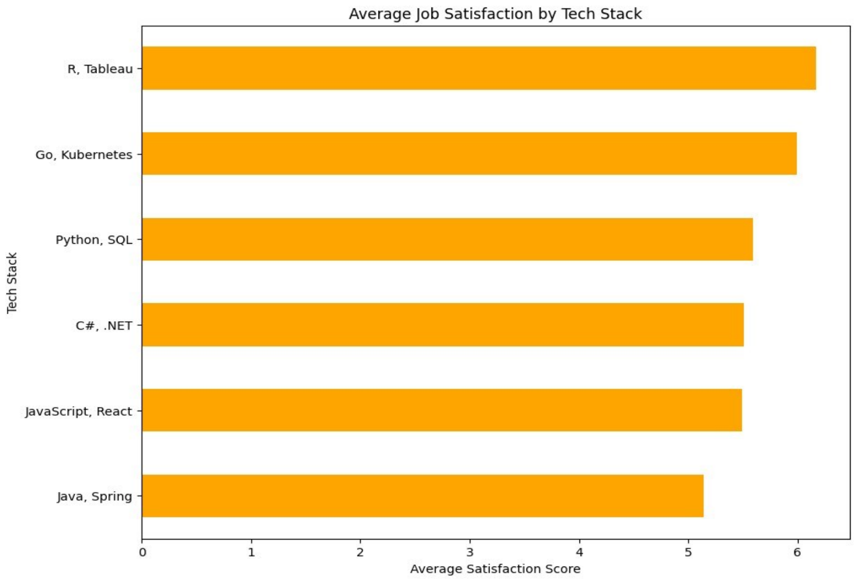

Fig. 5. Job Satisfaction by Tech Stack (Bar Plot)

The distribution of employment contracts in the dataset is illustrated in *fig. 6*. Full-time positions represent the majority, reflecting companies' tendency to maintain long-term employment relationships even in remote settings. Part-time and contract roles are also present but account for smaller segments of the workforce.

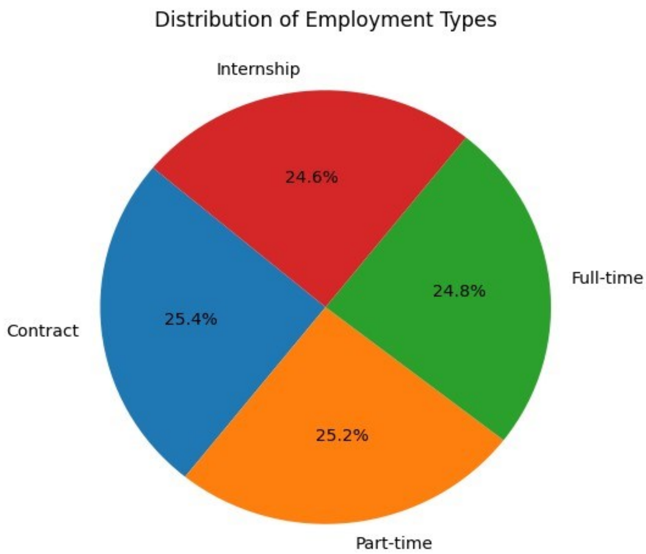

Fig. 6. Employment Type Distribution (Pie Chart)

The correlations among key numerical variables—including annual salary, years of experience, job satisfaction score, and time since last promotion—are illustrated in *fig. 7*. As expected, years of experience positively correlates with salary, while job satisfaction shows weak correlation with both salary and promotion recency, suggesting that factors such as work-life balance and team culture may have a stronger influence on satisfaction.

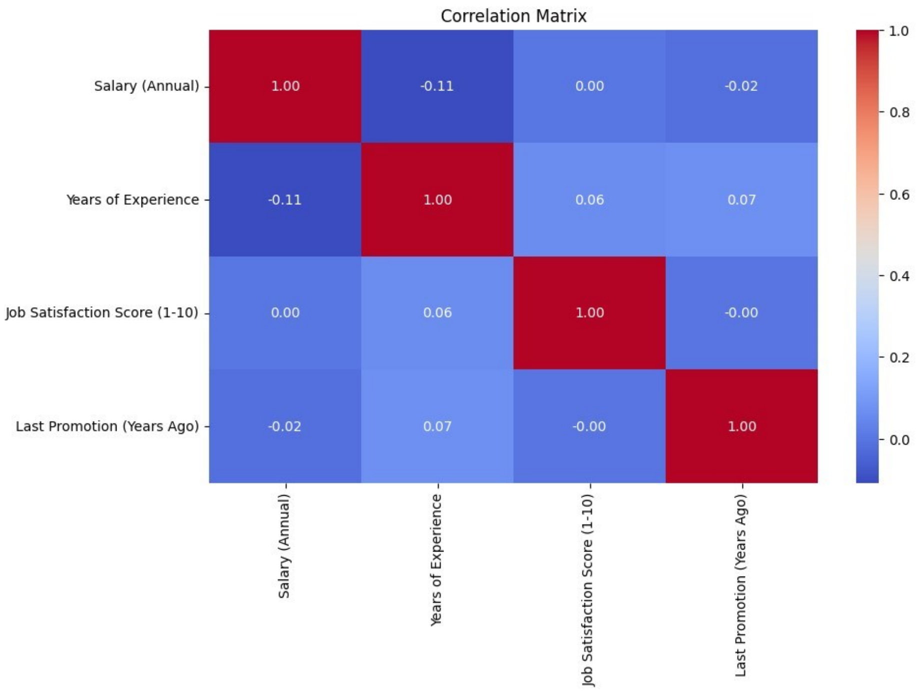

Fig. 7. Correlation Matrix (Heatmap)

5. CONCLUSIONS

In conclusion, the analysis highlights the industries that offer the highest salaries in remote work, the impact of experience and promotions on compensation, and the connection between work flexibility and employee satisfaction.

REFERENCES

[1] Globalization Partners, *Work from Anywhere: The Evolution of Global RemoteWork*, https://www.globalization-partners.com, 2022.

[2] Buffer, *State of Remote Work 2023*, https://buffer.com/stateof-remote-work, 2023..

[3] Glassdoor Economic Research, *Remote Work: What We've Learned So Far*, https://www.glassdoor.com/research, 2022.

[4] OECD, *Teleworking in the COVID-19 pandemic: Trends and prospects*, OECD Publishing, https://doi.org/10.1787/5fad3b8e-en, 2021.

[5] The Pandas Development Team, *pandas-dev/pandas: Powerful Python data analysis toolkit (v2.1)*, Zenodo, https://doi.org/10.5281/zenodo.3509134, 2023.

[6] J. D. Hunter, Matplotlib: *A 2D Graphics Environment*, Computing in Science& Engineering, vol. 9, no. 3, pp. 90–95, doi: 10.1109/MCSE.2007.55, 2007.

[7] F. Pedregosa et al., *Scikit-learn: Machine Learning in Python*, Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.