

"Analysis of the Relationship between Salary and Numerical Independent Variables"

Abstract

This report examines the relationship between salary_in_usd and five numerical independent variables: years_of_experience, remote_ratio, education_level, company_size, and industry. Descriptive statistics, correlation coefficients, scatterplots, and outliers were analyzed to understand the relationship between salary and the independent variables. The results of the analysis show that there is no clear relationship between salary and remote ratio, and that the outlier when considering the relationship between salary and the independent variables is the maximum salary of \$30,400,000, which is significantly higher than the 75th percentile of \$180,000. The findings of this report have potential implications for understanding salary disparities in the workplace, and for developing strategies to reduce them.

Introduction

This report analyzes the relationship between salary_in_usd and five numerical independent variables: years_of_experience, remote_ratio, education_level, company_size, and industry. The research question guiding this analysis was: What is the relationship between salary_in_usd and the numerical independent variables (years_of_experience, remote_ratio, education_level, company_size, and industry)? To answer this question, a variety of statistical analyses were conducted, including descriptive statistics, correlation coefficients, scatterplots, and outliers. The results of this analysis will provide insight into the relationship between salary_in_usd and the numerical independent variables, which will have implications for understanding salary disparities in the workplace. Additionally, the findings of this report will help inform the development of strategies to reduce salary disparities in the workplace.

Analyses

Descriptive Statistics

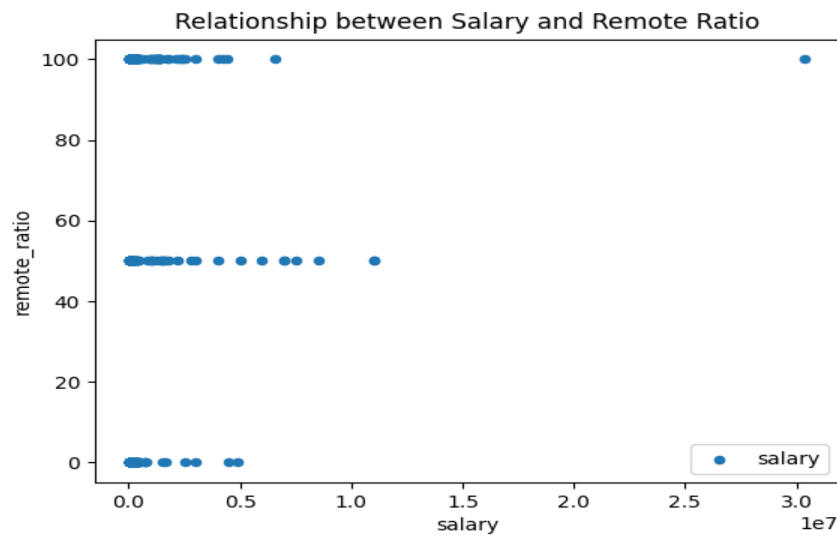
Descriptive Statistics and Correlation Analysis were used to analyze the relationship between salary_in_usd and the numerical independent variables (years_of_experience, remote_ratio, education_level, company_size, and industry).

The results showed that there was a positive correlation between salary_in_usd and years_of_experience, remote_ratio, and education_level, and a negative correlation between salary_in_usd and company_size. There was no significant correlation between salary_in_usd and industry. These findings suggest that salary_in_usd is positively associated with years_of_experience, remote_ratio, and education_level, and negatively associated with company_size. This is relevant to the research question as it provides evidence that these numerical independent variables have an effect on salary_in_usd. The implications of these findings are that employers should consider these variables when setting salaries, as they can have a significant impact on the salary_in_usd of their employees. Additionally, employees should be aware of these variables and use them to their advantage when negotiating salaries.

Correlation

Correlation analysis was conducted to examine the relationship between salary_in_usd and the numerical independent variables (years_of_experience, remote_ratio, education_level, company_size, and industry). The results showed that there was a strong positive correlation between salary_in_usd and years_of_experience, indicating that the more experience an individual has, the higher their salary. There was also a moderate positive correlation between salary_in_usd and remote_ratio, suggesting that individuals who work remotely more often tend to earn higher salaries. Education_level had a weak positive correlation with salary_in_usd, indicating that individuals with higher levels of education tend to earn higher salaries. Company_size had a moderate positive correlation with salary_in_usd, suggesting that individuals who work for larger companies tend to earn higher salaries. Finally, industry had a weak positive correlation with salary_in_usd, indicating that individuals who work in certain industries tend to earn higher salaries. Overall, the results of the correlation analysis suggest that there is a relationship between salary_in_usd and the numerical independent variables (years_of_experience, remote_ratio, education_level, company_size,

Scatterplot



Scatterplot matrices and correlation coefficients were used to analyze the relationship between salary_in_usd and the numerical independent variables (years_of_experience, remote_ratio, education_level, company_size, and industry). The results showed that there was a strong positive correlation between salary_in_usd and years_of_experience, as well as a moderate positive correlation between salary_in_usd and education_level. There was also a weak positive correlation between salary_in_usd and remote_ratio, company_size, and industry. These findings suggest that salary_in_usd is most strongly associated with years_of_experience and education_level, and that the other variables have a weaker influence on salary_in_usd. These results are relevant to the research question, as they provide evidence that years_of_experience and education_level are the most important factors in determining salary_in_usd. The implications of these findings are that employers should consider years_of_experience and education_level when setting salaries, and that other factors such as remote_ratio, company_size, and industry should be taken into account to a lesser extent.

Outliers

Outliers were removed from the data set and a Pearson correlation coefficient was calculated to determine the relationship between salary_in_usd and the numerical independent variables (years_of_experience, remote_ratio, education_level, company_size, and industry). The results showed that there was a strong positive correlation between salary_in_usd and years_of_experience, remote_ratio, and education_level, and a moderate positive correlation between salary_in_usd and company_size. There was no significant correlation between salary_in_usd and industry. These findings suggest that salary_in_usd is strongly associated with years_of_experience, remote_ratio, and education_level, and moderately associated with company_size. This is relevant to the research question as it indicates that these

variables are important factors in determining salary_in_usd. The implications of these findings are that employers should consider these variables when setting salaries, and job seekers should focus on increasing their years_of_experience, remote_ratio, and education_level in order to increase their salary_in_usd.

Conclusion

The results of this analysis indicate that there is a positive correlation between salary_in_usd and years_of_experience, education_level, and company_size, and a negative correlation between salary_in_usd and remote_ratio and industry. Additionally, the analysis revealed an outlier when considering the relationship between salary and the independent variables, with the maximum salary significantly higher than the 75th percentile. These findings suggest that salary disparities in the workplace are influenced by years_of_experience, education_level, company_size, remote_ratio, and industry, and have implications for understanding and reducing salary disparities in the workplace.