Project Report

# Introduction

This report examines the relationship between the sex and the number of hours worked per week using linear regression analysis. Additionally, we check how accounting for educational level affects the correlation. The goal is to examine patterns in working hours based on gender and evaluate the impact of other factors in enhancing the model.

# Linear Regression Analysis

We performed a linear regression with sex as the independent variable and hours-per-week as the dependent variable. To perform this, gender was encoded as a binary variable (0 for Female and 1 for Male). The results are as follows:

* Intercept: 36.45, average hours worked by females per week
* Slope: 6.03, males work an average of 6.03 more hours per week than females
* R²: 0.06, 6% of the variability in hours-per-week is explained by gender.

These results indicate that in general males work significantly more hours per week than females. However, the low R² value suggests that sex alone is a weak predictor of working hours, indicating that there are other factors that play a significant role.

In another linear regression analysis, we added education-num as a control variable, and the following are the updated results:

* Intercept: 29.4, average hours worked by females per week
* Slope (sex): 5.97, even after controlling for educational level, males work on average 5.97 more hours than females per week
* Slope (educational-num): 0.7, meaning for each additional year of education, individuals work for an additional 0.7 hours per week, regardless of gender.

Adding education-num to the model slightly reduced the gender effect but did not eliminate it, suggesting that differences in education slightly affect the working hours. However, even after controlling for education, males still work more hours per week than females. This means that education level has an effect on hours-per-week, but it is also affected by other factors.

# Conclusion

On average, men work almost 6 more hours than women per week, even after adding the education-num as a controlling variable. This means that education level is a significant predictor of working hours but does not fully explain the gender disparity.