The Impact of Education on Income in Adulthood

# Introduction

In recent years, researchers have become increasingly interested in examining the connection between wealth and education. Considering the substantial shifts in the global economy, where the demand for highly qualified employees has increased while the demand for less skilled labour has decreased [1], this is especially crucial. As a result, an individual's livelihood and lasting financial well-being are now greatly influenced by their level of education. Numerous studies have demonstrated that people with higher education levels typically have higher earning potential than people with lower education levels [2] [1] [3]. This connection has been noted in several nations and areas, and it is especially significant in industrialized economies where access to education is widespread. Additionally, the type and level of education can have a significant impact on earnings, with higher levels of education generally associated with higher salaries and better career prospects.

However, there are many variables that can affect the complex relationship between education and income. For instance, depending on the industry or vocation, the effect of education on income may vary, with certain professions or industries having better wage potential than others. A person's age, gender, and ethnicity are only a few examples of the demographic factors that might affect how education affects income. Additionally, it's not always easy to determine how education and income are related. While more education is typically linked to higher salaries, other elements like work experience and soft skills can also have a significant impact on a person's earning potential. Moreover, some studies suggest that there may be diminishing returns to education, where the impact of additional education on earnings decreases as the level of education increases.

Despite these difficulties, research on how education affects income is still important. Policymakers, educators, and people looking to increase their economic mobility can all benefit from having a better understanding of the variables that affect the link between education and income. Examining the complexities of this connection enables one to spot chances for bettering educational policy, expanding educational access, and enhancing the long-term financial security of people and society as a whole.

The goal of this study is to analyze the link between education and income in adulthood and to pinpoint the variables that influence it. The research will examine the association between people's income and education level by analyzing data from a sample of Americans. The study will examine a number of variables, including gender, race, age, and employment, that may have an influence on how schooling affects income. The purpose of the study is to shed light on the elements that might influence the link between education and income for policymakers and educators. This information may be utilized to create more efficient educational policies and initiatives that will boost Americans' chances of earning more money and advancing their careers. The initiative intends to add to the ongoing conversation about education and income and offer insightful information about the variables that might influence this relationship. The results of this study can have significant ramifications for decision-makers, educators, and those looking to further their educational and employment prospects.

## Research question 1

* What is the effect of education on income, and what factors contribute to this effect?

### Aim

* To examine the relationship between income and education across a range of demographic variables, and to pinpoint the major variables that influence this relationship.

### Objectives:

* To investigate the relationship between education and income.
* To determine how different levels of education impact income.
* To identify the factors that contribute to the relationship between education and income.
* To analyze the effect of demographic variables such as gender, race, and occupation on the relationship between education and income.

## Research question 2

* Are there any gender differences in the effect of education on income?

### Aim

* To determine whether there are any gender differences in the relationship between education and income.

### Objectives

* To compare the effect of education on income for males and females.
* To analyze any gender-based differences in the relationship between education and income.
* To investigate the potential impact of other demographic variables on the relationship between education and income for each gender.

## Research question 3

* What other factors besides education impact income?

### Aim

* To identify other factors besides education that may impact income.

### Objectives

* To explore the impact of demographic variables such as age, race, and occupation on income.
* To investigate the potential impact of other factors such as work experience, industry, and location on income.
* To determine which of these variables have a significant impact on income, and how they interact with education.

# Literature Review

Education has long been recognized as one of the most significant determinants affecting adult income. Higher levels of education are consistently linked to higher income levels, according to numerous studies that have looked at this relationship [2]. However, the precise nature of this relationship and the underlying mechanisms remain poorly understood. One reason for the link between education and income is that education equips people with the abilities to undertake tasks requiring a greater degree of competence, which are typically paid higher wages [4]. This explanation suggests that the effect of education on income is largely mediated by the type of job an individual is able to obtain.

According to other studies [5], even after accounting for the sort of work a person holds, education may still have a more significant influence on income. For instance, better education may be linked to greater productivity and efficiency, which might result in higher compensation even for professions that don't need a lot of skill. Additionally, having more education may suggest to employers that a person is more dedicated, dependable, and hardworking, which could result in higher pay even for entry-level positions [3].

Numerous other elements have also been found to be significant predictors of income in adulthood, in addition to education. Age is one such variable. Even after accounting for schooling and other criteria, research has repeatedly demonstrated that older workers tend to receive greater incomes than younger workers [6]. Because they have more work experience and are therefore more productive and efficient, older workers may also be in positions of greater responsibility and authority, which may also contribute to this. The occupation is a significant predictor of income. Even after accounting for schooling and other criteria, research repeatedly demonstrates that some professions are connected with greater wages than others [7]. For instance, jobs in the healthcare, technology, and finance industries often offer higher wages than jobs in the retail, hotel, and agricultural industries. Additionally recognized as significant predictors of income are race and gender. Even after accounting for education, occupation, and other criteria, research repeatedly demonstrates that women and people of colour tend to earn less than males and white people [8]. Disparities in access to school and other opportunities, as well as prejudice and discrimination in the labour market, may contribute to this.

According to the literature, education is a major predictor of income in adulthood, but other elements including age, occupation, race, and gender also have a significant impact. Future research is required to better understand the mechanisms underlying the relationship between education and income. This research will also help identify effective policies for expanding access to education and reducing income inequality.

# Methodology

The purpose of this study is to look at how education and other characteristics affect adult income. In order to establish the significance of the association between the covariates and income, a binary logistic regression model will be applied. When the goal of the study is to model binary outcomes as a function of predictor variables, logistic regression is a suitable approach. The outcome variable in this situation is whether a person earns more than $50,000 or less per year.

The logistic regression model will take the following form:

Where,

𝑙𝑜𝑔𝑖𝑡 (𝑃i) = natural log transformation of the odds of earning above 50,000 dollars (𝑃i) to earning below 50,000 dollars (1 − 𝑃i)

𝛽0 = Bias

𝛽k = k parameter estimates associated with k independent variable, X

These factors will be estimated, and then they will be used to calculate the likelihood that a person will earn more than $50,000 depending on their education level. The model will allow us to investigate how the specified parameters affect the log chances of earning more than $50,000.

# Data used for Empirical Application

The 1994 US Census Database, which has been taken from Kaggle, is the dataset utilized in this study. The dataset includes a sample of roughly 3000 people who were between the ages of 18 and 25 in 1994 and were drawn from a nationally representative sample. The dataset's rows each represent a unique person, and each row contains details on that person's education level, age, income, gender, ethnicity, and other demographic characteristics. The dataset offers a special chance to investigate how education and income relate to a variety of demographic factors. It enables researchers to look at how income is affected by education across a range of age groups, genders, ethnicities, and vocations. By using this dataset, researchers can analyze how different levels of education impact income and whether there are any gender or racial differences in the effect of education on income.

## Data Preparation and Summary Statistics

Age, Working Class, Education, Marital Status, Occupation, Relationship, Race, Sex, Hours per Week, Country, and Salary are among the 11 fields in the dataset utilized in this study. These variables include details about the people in the dataset, including demographic data like age, gender, and race as well as details on their employment circumstances, education, and pay. The dataset consists of 3020 tuples, or observations, each of which represents a single person. This dataset may be used for a variety of analysis, such as looking at the link between income and education or identifying trends in the distribution of the sample's demographics.

The Age column has a mean of 21.79 years with a range of 18 to 25 years, and a median of 22 years. The Hours.per.week column has a mean of 36.27 with a range of 1 to 99 hours, and a median of 40 hours. The remaining columns are categorical and have no numerical summary statistics available. However, it has to be noted that these categorical are later on converted to numerical columns using ordinal encoding. Table 1 summarizes the findings

.

*Table 1 Summary*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Column | Mean | Median | Minimum | Maximum |
| Age | 21.79 | 22 | 18 | 25 |
| Working.class | Categorical Column | | | |
| Education | Categorical Column | | | |
| Marital.status | Categorical Column | | | |
| Occupation | Categorical Column | | | |
| Relationship | Categorical Column | | | |
| Race | Categorical Column | | | |
| Sex | Categorical Column | | | |
| Hours.per.week | 36.27 | 40 | 1 | 99 |
| Country | Categorical Column | | | |
| Salary | Categorical Column | | | |

The table 2 shows the number of unique values for each field in the dataset. The Working.class field has 9 unique values, Education has 11, Marital.status has 5, Occupation has 15, Relationship has 6, Race has 5, Sex has 2, Country has only 1, and Salary has 2 unique values. The Country field has only one unique value, which means that this field will not contribute any information to the logistic regression model since it does not have any variability in the data. Therefore, the Country field needs to be dropped from the dataset before performing logistic regression.

*Table 2 Number of Unique Values per Field*

|  |  |
| --- | --- |
| Fields | Number of unique values |
| Working.class | 9 |
| Education | 11 |
| Marital.status | 5 |
| Occupation | 15 |
| Relationship | 6 |
| Race | 5 |
| Sex | 2 |
| Country | 1 |
| Salary | 2 |

To further investigate the distribution of the Salary variable, a count plot and a pie chart were generated as shown in Figure 1. The resulting visualizations indicated that the dataset is highly skewed towards individuals with salaries below 5000 USD per year. It has been decided that the data will not be balanced since the main objective is to solely explore the relationship between the different fields of the dataset. Therefore, it is not necessary to address the bias in the data.

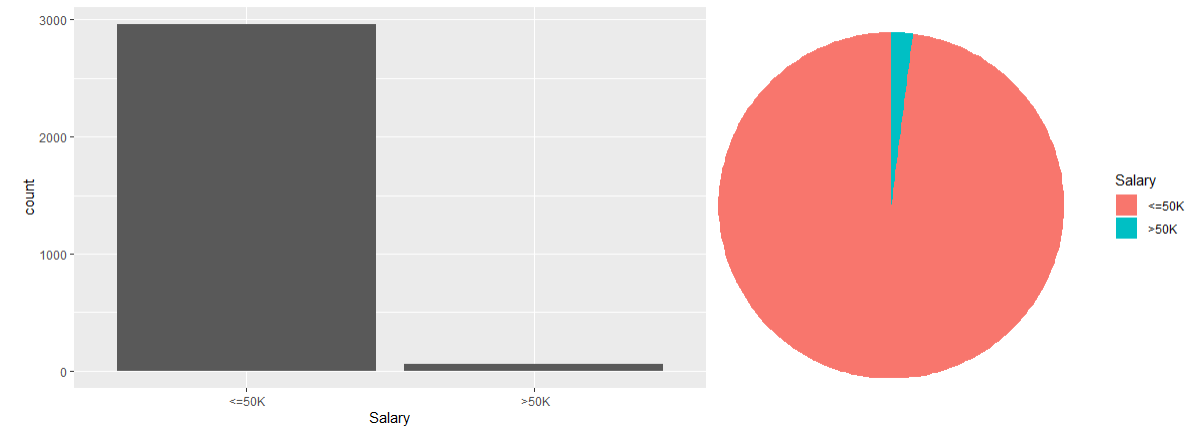


Figure Salary Distribution

Figure 2 explores the gender-wise salary distribution of the participants. From the bar chart, it can be seen that there is a higher probability that an individual with a higher salary is male.



Figure Gender Wise Salary Distribution

The correlation between the dependent and independent variables was calculated to determine the strength and direction of the relationship between them. A heatmap, shown in Figure 3, was used to visualize the correlation between the different independent variables and the dependent variable, while the bars in the figure 4 indicate the strength and direction of the relationship between the independent variables and the salary. The numerical values of the correlation of the independent variables with the target are shown in Table 3.



Figure Correlation Heatmap

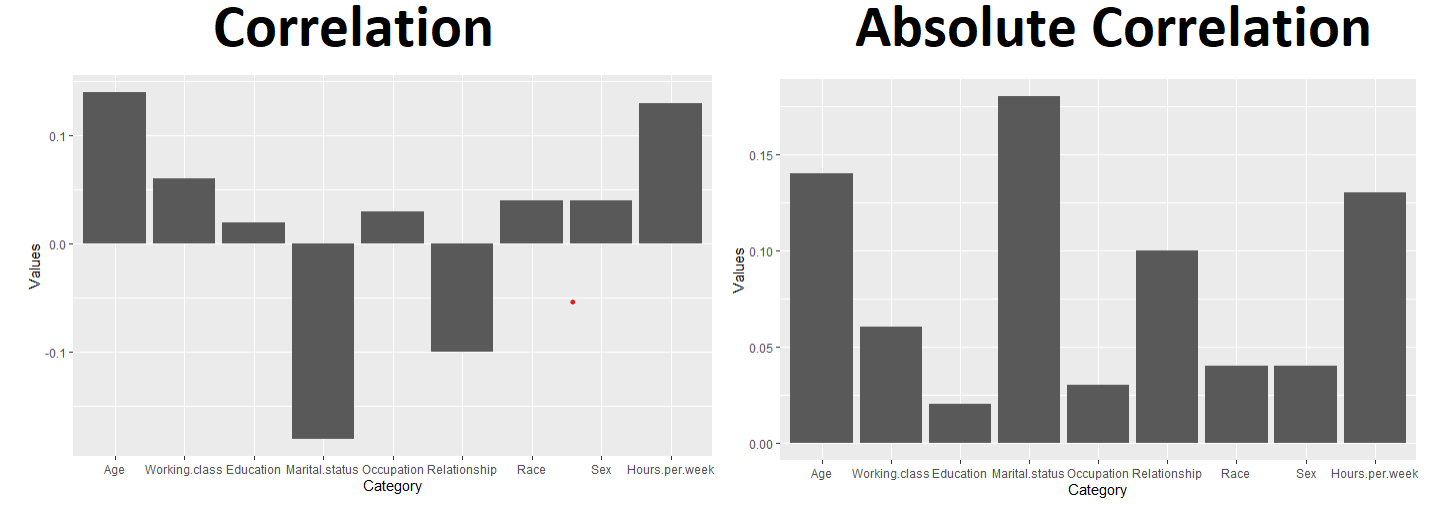


Figure Correlation and Absolute Correlation of Features with Salary

*Table 3 Correlation of Independent variables with Salary*

|  |  |
| --- | --- |
| Field | Correlation with Salary |
| Age | 0.14 |
| Hours.per.week | 0.13 |
| Working.class | 0.06 |
| Race | 0.04 |
| Sex | 0.04 |
| Occupation | 0.03 |
| Education | 0.02 |
| Relationship | -0.1 |
| Marital.status | -0.18 |

The results of the correlation analysis between the independent variables and the dependent variable (Salary) are presented in the table 3. The table shows the correlation coefficient values for each independent variable with Salary. It can be observed that Age and Hours.per.week have a moderately positive correlation with Salary, with correlation coefficients of 0.14 and 0.13, respectively. Working.class, Race, Sex, and Occupation have weak positive correlations with Salary, with correlation coefficients ranging from 0.02 to 0.06. On the other hand, Relationship and Marital.status have negative correlations with Salary, with correlation coefficients of -0.1 and -0.18, respectively.

Considering the results, it can be argued that Age and Hours.per.week are the independent variables that are the most strongly related to Salary. This is because they have moderately positive correlation coefficients, which indicate that an increase in either of these variables is associated with a higher Salary. On the other hand, while Relationship and Marital.status have the strongest negative correlations with Salary, it does not necessarily mean that they are the best predictors of Salary. Nonetheless, based on the results of this correlation analysis, it can be suggested that Age and Hours.per.week are good candidates for inclusion in a regression model to predict Salary.

# Results and Discussion

*Table 4 Regression Results*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | z | Pr(>|z|) |
| (Intercept) | -18.84326 | 3.19565 | -5.897 | 3.71E-09 |
| Age | 0.51697 | 0.10859 | 4.761 | 1.93E-06 |
| Working.class | 0.4486 | 0.14031 | 3.197 | 0.00139 |
| Education | -0.02755 | 0.05985 | -0.46 | 0.64527 |
| Marital.status | -0.55756 | 0.10168 | -5.483 | 4.17E-08 |
| Occupation | 0.02395 | 0.03372 | 0.71 | 0.47745 |
| Relationship | -0.13582 | 0.11301 | -1.202 | 0.22942 |
| Race | 0.28982 | 0.27919 | 1.038 | 0.29924 |
| Sex | -0.01226 | 0.35885 | -0.034 | 0.97274 |
| Hours.per.week | 0.04815 | 0.01071 | 4.497 | 6.90E-06 |

The logistic regression analysis revealed several significant independent variables that are associated with salary. The results are summarized in Table 4.

The intercept coefficient is significant at p < 0.001, which suggests that even if all of the independent variables are zero, there is still a baseline probability of earning a salary that is different from zero. This implies that there may be other factors not included in the model that affect salary.

The coefficient for Age is positive and significant (p < 0.001), indicating that as an individual's age increases, the likelihood of having a higher salary also increases. This finding suggests that experience and expertise gained with age may contribute to higher salaries. Similarly, the coefficient for Hours.per.week is positive and significant (p < 0.001), suggesting that as the number of hours worked per week increases, the likelihood of having a higher salary also increases. This finding implies that individuals who work longer hours are compensated with higher salaries. The Working.class variable is also significant (p = 0.00139) with a positive coefficient, which implies that individuals in a particular working class are more likely to earn higher salaries. This result suggests that the type of job an individual has is an important predictor of salary, and certain working classes may have higher salaries than others.

Marital.status is negatively significant (p < 0.001), indicating that married individuals are less likely to earn higher salaries than unmarried individuals. This finding may suggest that married individuals may prioritize other factors, such as family responsibilities or work-life balance, over earning a higher salary. However, further research is needed to confirm this relationship and identify potential underlying mechanisms.

The coefficients for Education, Occupation, Relationship, Race, and Sex are not significant at the 0.05 level, indicating that there is no strong evidence to suggest that these variables are associated with salary. However, it is important to note that the lack of statistical significance does not necessarily mean that these variables do not have an impact on salary. Other factors such as sample size, measurement error, and model specification may influence the significance of these variables.

Based on the logistic regression results the initial research questions can be answered as follows:

*Research question 1: What is the effect of education on income, and what factors contribute to this effect?*

The coefficient for Education is not significant at the 0.05 level, indicating that there is no strong evidence to suggest that Education is associated with Salary. However, it is important to note that this coefficient is only one of several variables that could impact income. Other variables such as Age, Working.class, and Marital.status have significant coefficients, indicating that these variables are important predictors of Salary. For example, the coefficient for Working.class is significant with a positive coefficient, suggesting that individuals in a particular working class are more likely to earn higher salaries. Marital.status is negatively significant, indicating that married individuals are less likely to earn higher salaries than unmarried individuals.

*Research question 2: Are there any gender differences in the effect of education on income?*

The coefficient for Sex is not significant at the 0.05 level, indicating that there is no strong evidence to suggest that gender differences impact the relationship between education and income.

*Research question 3: What other factors besides education impact income?*

The coefficients for Age and Hours.per.week are significant with positive coefficients, suggesting that as Age and Hours.per.week increase, the likelihood of having a higher salary also increases. This highlights the importance of work experience and the number of hours worked in predicting Salary. Other variables such as Working.class, Marital.status, and Relationship also have significant coefficients, indicating that these variables are important predictors of Salary.

# Conclusion

This project aimed to investigate the relationship between education and income while considering various demographic variables such as age, gender, race, occupation, and marital status. The analysis was performed using logistic regression with salary as the dependent variable and the various demographic variables and education as independent variables. The results indicated that education alone is not a significant predictor of income, with an estimated coefficient of -0.02755 and a p-value of 0.64527. However, age and hours worked per week were significant predictors of income, with positive coefficients of 0.51697 and 0.04815, respectively. Additionally, the results suggested that certain demographic variables significantly impact income. Working class was found to have a positive relationship with income, while marital status was negatively associated with income. Race, sex, occupation, and relationship were found to have no significant effect on income. The study also investigated the gender differences in the relationship between education and income and found that there were no significant differences in the effect of education on income between males and females. The study identified that education is not the sole predictor of income, and several other demographic variables play a crucial role in determining income levels.

# References

|  |  |
| --- | --- |
| [1] | A. Bartel and F. Lichtenberg, "The comparative advantage of educated workers in implementing new technology.," *The Review of Economics and statistics,* 1987. |
| [2] | J. Morgan and M. David, "Education and income. T," *he Quarterly Journal of Economics,* 1963. |
| [3] | A. Shahabadi, M. Nemati and S. Hosseinidoust, "THE EFFECT OF EDUCATION ON INCOME INEQUALITY IN SELECTED ISLAMIC COUNTRIES.," *International Journal of Asia-Pacific Studies,* 2018. |
| [4] | G. W. J. G. R. C. P. H. B. W. J. a. B. E. Hawker, "The effect of education and income on need and willingness to undergo total joint arthroplasty.," *Arthritis & Rheumatism,* 2002. |
| [5] | A. Abdullah, H. Doucouliagos and E. Manning, "Does education reduce income inequality? A meta‐regression analysis.," *Journal of Economic Surveys,* 2015. |
| [6] | P. Hedström and S. Ringen, "Age and income in contemporary society: a research note.," *Journal of Social Policy, ,* 1987. |
| [7] | C. Mustard, S. Derksen, J. Berthelot, M. Wolfson and L. Roos, "Age-specific education and income gradients in morbidity and mortality in a Canadian province.," *Social science & medicine,* 1997. |
| [8] | D. Bobbitt-Zeher, "The gender income gap and the role of education.," *Sociology of education,* 2007. |