Table of Contents

[2) Definition of Variables 2](#_Toc463002344)

[5) Data 3](#_Toc463002345)

[6) Write the regression equation: 4](#_Toc463002346)

[6.1 Interpretation????? 4](#_Toc463002347)

[7) Identify and interpret the t tests for each of the coefficients (one separate paragraph for each variable, in numerical order): 5](#_Toc463002348)

[7.1 interpret the t tests - Education 5](#_Toc463002349)

[7.2 interpret the t tests- Gender 6](#_Toc463002350)

[7.3 interpret the t tests - experience 7](#_Toc463002351)

[7.4 interpret the t tests - marital 8](#_Toc463002352)

[7.5 interpret the t tests - children 9](#_Toc463002353)

[**7.6 interpret the t tests - Age** 10](#_Toc463002354)

[8) Identify and interpret the adjusted R2 (one paragraph):< Define "adjusted R2."< 11](#_Toc463002355)

[What does the value of the adjusted R2 reveal about the model? 62% related, closer than 1 is best…. 11](#_Toc463002356)

[If the adjusted R2 is low, how has the choice of independent variables created this result? 11](#_Toc463002357)

[9) Analyze multicollinearity of the independent variables (one paragraph): 12](#_Toc463002358)

[A) Generate the correlation matrix 12](#_Toc463002359)

[B) Define multicollinearity 12](#_Toc463002360)

[C) Are any of the independent variables highly correlated with each other? If so, identify the variables and explain why they are correlated? 12](#_Toc463002361)

[D) State the implications of multicollinearity (if found) for the model. 12](#_Toc463002362)

[10) Identify and interpret the F test (one paragraph): 14](#_Toc463002363)

[A) Using the p-value approach, perform f test 14](#_Toc463002364)

[B) is the null hypothesis for the F test rejected or not rejected? 14](#_Toc463002365)

[C) Why or why not? 14](#_Toc463002366)

[D) Interpret the implications of these findings for the model. 14](#_Toc463002367)

# 2) Definition of Variables

Y1 Annual Wage – the current annual wage is the dependent variable (DV) in this study. The annual wage is assumed influenced by other factors as outlined in the following independent variables (IV) to follow.

X2 Education –completed years of education

X3 Gender – gender (Female 0, Male 1)

X4 Experience - # of years of work experience

X5 Marital Status – Current marital status (Single 0, Married 1)

X6 Parental Status - Parental status (No Children 0, Children 1)

X7 Age – Age represented in full years

# 5) Data

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Y1 Wage** | **X2 Education** | **X3 Gender** | **X4 Experience** | **X5 Marital** | **X6 Children** | **X7 Age** |
| $ 17,000.00 | 8 | 0 | 1 | 1 | 1 | 19 |
| 22,000.00 | 12 | 1 | 2 | 0 | 0 | 20 |
| 27,000.00 | 12 | 1 | 2 | 0 | 1 | 20 |
| 20,000.00 | 12 | 0 | 3 | 0 | 1 | 21 |
| 44,000.00 | 12 | 1 | 3 | 0 | 0 | 21 |
| 48,000.00 | 13 | 0 | 3 | 1 | 0 | 21 |
| 20,000.00 | 12 | 1 | 4 | 1 | 0 | 22 |
| 36,000.00 | 18 | 0 | 5 | 0 | 1 | 23 |
| 30,000.00 | 16 | 1 | 6 | 0 | 1 | 24 |
| 28,000.00 | 12 | 0 | 7 | 0 | 1 | 25 |
| 47,000.00 | 16 | 1 | 10 | 0 | 1 | 28 |
| 50,000.00 | 16 | 0 | 12 | 1 | 1 | 30 |
| 65,000.00 | 16 | 1 | 12 | 1 | 1 | 30 |
| 75,000.00 | 18 | 1 | 5 | 1 | 1 | 30 |
| 70,000.00 | 16 | 0 | 14 | 1 | 1 | 32 |
| 42,000.00 | 13 | 0 | 17 | 0 | 1 | 35 |
| 61,000.00 | 14 | 0 | 18 | 1 | 0 | 36 |
| 80,000.00 | 17 | 0 | 18 | 1 | 1 | 36 |
| 100,000.00 | 18 | 0 | 18 | 1 | 1 | 36 |
| 37,000.00 | 18 | 0 | 19 | 1 | 1 | 37 |
| 33,000.00 | 12 | 1 | 20 | 1 | 0 | 38 |
| 32,000.00 | 14 | 0 | 22 | 1 | 0 | 36 |
| 38,000.00 | 12 | 0 | 24 | 0 | 1 | 36 |
| 50,000.00 | 11 | 1 | 25 | 0 | 0 | 43 |
| 34,000.00 | 17 | 1 | 26 | 1 | 1 | 44 |
| 30,500.00 | 16 | 1 | 27 | 1 | 0 | 45 |
| 50,500.00 | 12 | 0 | 33 | 1 | 1 | 51 |
| 85,000.00 | 11 | 1 | 33 | 0 | 1 | 51 |
| 45,000.00 | 12 | 0 | 43 | 0 | 0 | 61 |
| 32,000.00 | 12 | 1 | 45 | 0 | 1 | 63 |
| 38,000.00 | 16 | 1 | 45 | 1 | 0 | 61 |
| 49,000.00 | 12 | 1 | 45 | 1 | 0 | 57 |
| 55,000.00 | 16 | 0 | 45 | 1 | 1 | 61 |
| 145,000.00 | 18 | 1 | 45 | 1 | 1 | 63 |
| $ 40,000.00 | 14 | 0 | 47 | 0 | 1 | 61 |

# 6) Write the regression equation:

Equation | Y1 Wage = -66,521 + 2,649 X2 Education + 1,131 X3 Gender – 2,350 X4 Experience + 8,325 X5 Marital + 8,926 X6 Children + 3,006 X7 Age

Wage = Education (x2) + Gender (x3) – Experience (x4) + Marital (x5) + Children (x6) + Age (x7)

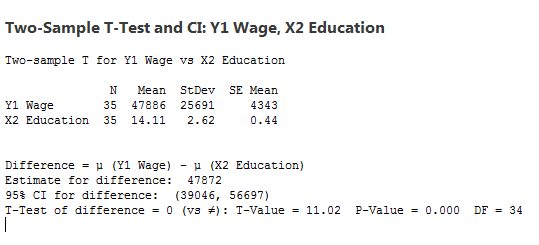
Y1 Wage = -66521 + 2649 X2 Education + 1131 X3 Gender - 2350 X4 Experience + 8325 X5 Marital+ 8926 X6 Children + 3006 X7 Age

# 6.1 Interpretation?????

1. Provide a brief definition of each independent variable and how it impacts the equation
   1. X2 Education
   2. X3 Gender
   3. X4 Experience
   4. X5 Marital
   5. X6 Children
   6. X7 Age

*???The regression summary above suggests that individuals will earn 2,649 for each year of education obtained, females will make 1,131 more than males, individuals will earn 8,325 more per year of education, married individuals will earn 8,926 more than single individuals, and each year of age will increase wages by 3,006.*??? NOT SURE IF THIS IS RIGHT?

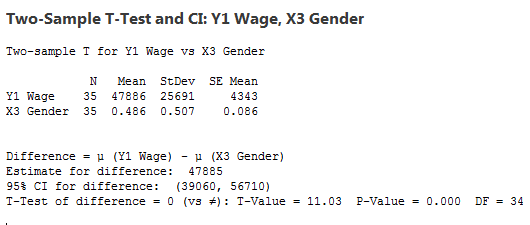
# 7) Identify and interpret the t tests for each of the coefficients (one separate paragraph for each variable, in numerical order):



## 7.1 interpret the t tests - Education

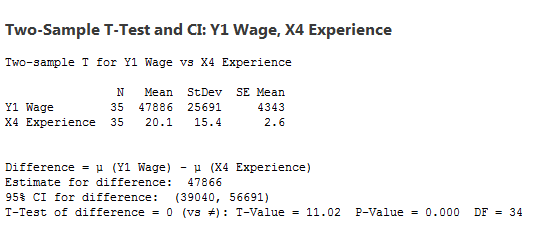
1. Are the signs of the coefficients as expected? Expected to income to increase with wages

1. If not, why not?
2. For each of the coefficients, interpret the numerical value. – listed in screenshots
3. Using the p-value approach, is the null hypothesis for the t test rejected or not rejected for each coefficient? Why or why not?
4. Interpret the implications of these findings for the variable.
5. Identify the variable with the greatest significance.



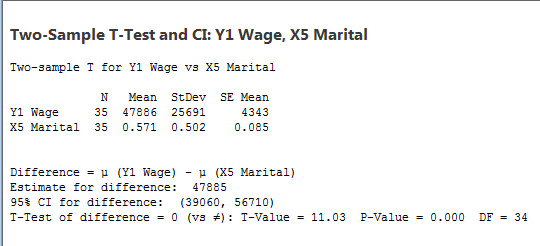
## 7.2 interpret the t tests- Gender

1. Are the signs of the coefficients as expected? Expected gender to be a factor (men make more)
2. If not, why not?
3. For each of the coefficients, interpret the numerical value. – listed in screenshots
4. Using the p-value approach, is the null hypothesis for the t test rejected or not rejected for each coefficient? Why or why not?
5. Interpret the implications of these findings for the variable.
6. Identify the variable with the greatest significance.



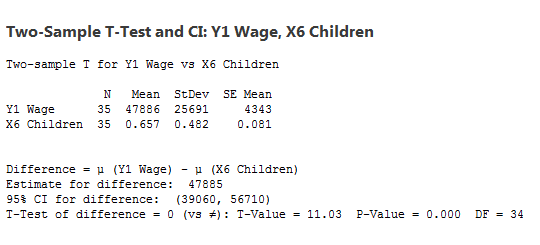
## 7.3 interpret the t tests - experience

1. Are the signs of the coefficients as expected? Expected more experience more money
2. If not, why not?
3. For each of the coefficients, interpret the numerical value. – listed in screenshots
4. Using the p-value approach, is the null hypothesis for the t test rejected or not rejected for each coefficient? Why or why not?
5. Interpret the implications of these findings for the variable.
6. Identify the variable with the greatest significance.



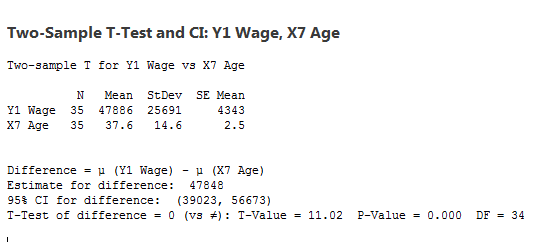
## 7.4 interpret the t tests - marital

1. Are the signs of the coefficients as expected? Expected married individuals to make more than single individuals
2. If not, why not?
3. For each of the coefficients, interpret the numerical value. – listed in screenshots
4. Using the p-value approach, is the null hypothesis for the t test rejected or not rejected for each coefficient? Why or why not?
5. Interpret the implications of these findings for the variable.
6. Identify the variable with the greatest significance.



## 7.5 interpret the t tests - children

1. Are the signs of the coefficients as expected? Expected individuals with children to earn more
2. If not, why not?
3. For each of the coefficients, interpret the numerical value. – listed in screenshots
4. Using the p-value approach, is the null hypothesis for the t test rejected or not rejected for each coefficient? Why or why not?
5. Interpret the implications of these findings for the variable.
6. Identify the variable with the greatest significance.



## **7.6 interpret the t tests - Age**

1. Are the signs of the coefficients as expected? Expected income to increase with age
2. If not, why not?
3. For each of the coefficients, interpret the numerical value. – listed in screenshots
4. Using the p-value approach, is the null hypothesis for the t test rejected or not rejected for each coefficient? Why or why not?
5. Interpret the implications of these findings for the variable.
6. Identify the variable with the greatest significance.

# 8) Identify and interpret the adjusted R2 (one paragraph):< Define "adjusted R2."<

|  |  |
| --- | --- |
| *Regression Statistics* | |
| Multiple R | 0.621968084 |
| R Square | 0.386844298 |
| Adjusted R Square | 0.255453791 |
| Standard Error | 22168.07017 |
| Observations | 35 |

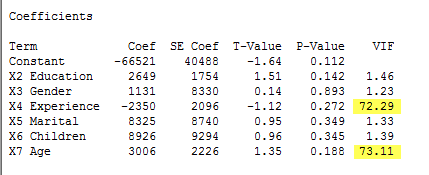
**Questions:**

## What does the value of the adjusted R2 reveal about the model? 62% related, closer than 1 is best….

## If the adjusted R2 is low, how has the choice of independent variables created this result?

# 9) Analyze multicollinearity of the independent variables (one paragraph):

## A) Generate the correlation matrix



## B) Define multicollinearity

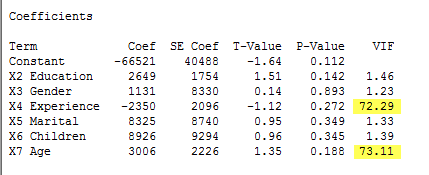
*Multi-collinear exists when the independent variables are correlated. It does not affect a multiple regression equation’s ability to find the dependent variable but it does evaluate the relationship between each independent variable and dependent variable. (Lind 461) The result of the correlation analysis reveals that multi-collinear does not exist between the independent variables. The VIF is not larger than 10 so it is satisfactory. If any of the independent variables were larger than the upper limit 10 then they should be removed from the analysis.*

## C) Are any of the independent variables highly correlated with each other? If so, identify the variables and explain why they are correlated?

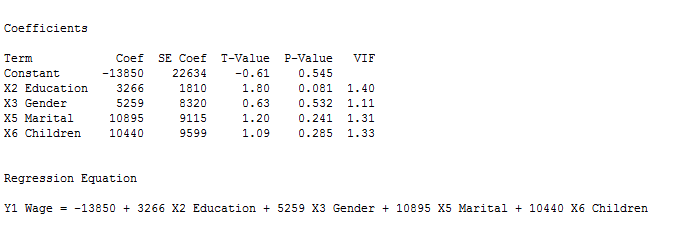
There are two independent variables that are highly correlated X4 Experience and X7 Age ?

## D) State the implications of multicollinearity (if found) for the model.

*The result of the correlation analysis reveals that multi-collinear does not exist between the independent variables. The VIF is not larger than 10 so it is satisfactory. If any of the independent variables were larger than the upper limit 10 then they should be removed from the analysis.*



Excluding Experience and Age independent variables from analysis the revised regression analysis makes more sense. This equation suggest education will increase by 3,266 by each year of education obtained, males will make 5,259 more than women, married individuals will make 10895 more than single individuals, and parents will earn 10,440 more than non children families.*??*



# 10) Identify and interpret the F test (one paragraph):

## Using the p-value approach, perform f test

## is the null hypothesis for the F test rejected or not rejected?

## Why or why not?

## Interpret the implications of these findings for the model.

**One-way ANOVA:**

**Method**

Null hypothesis All means are equal

Alternative hypothesis At least one mean is different

Significance level α = 0.05

Equal variances were assumed for the analysis.

**Factor Information**

Factor Levels Values

Factor 7 Y1 Wage, X2 Education, X3 Gender, X4 Experience, X5 Marital, X6 Children, X7 Age

**Analysis of Variance**

Source DF Adj SS Adj MS F-Value P-Value

Factor 6 68756062436 11459343739 121.53 0.000

Error 238 22441058450 94290162

Total 244 91197120886

**Model Summary**

S R-sq R-sq(adj) R-sq(pred)

9710.31 75.39% 74.77% 73.92%

**Means**

Factor N Mean StDev 95% CI

Y1 Wage 35 47886 25691 ( 44652, 51119)

X2 Education 35 14.114 2.621 ( -3219.300, 3247.528)

X3 Gender 35 0.4857 0.5071 (-3232.9283, 3233.8998)

X4 Experience 35 20.11 15.42 ( -3213.30, 3253.53)

X5 Marital 35 0.5714 0.5021 (-3232.8426, 3233.9855)

X6 Children 35 0.6571 0.4816 (-3232.7569, 3234.0712)

X7 Age 35 37.63 14.60 ( -3195.79, 3271.04)

Pooled StDev = 9710.31

**Regression Equation**

Y1 Wage = -66521 + 2649 X2 Education + 1131 X3 Gender - 2350 X4 Experience + 8325 X5 Marital + 8926 X6 Children + 3006 X7 Age