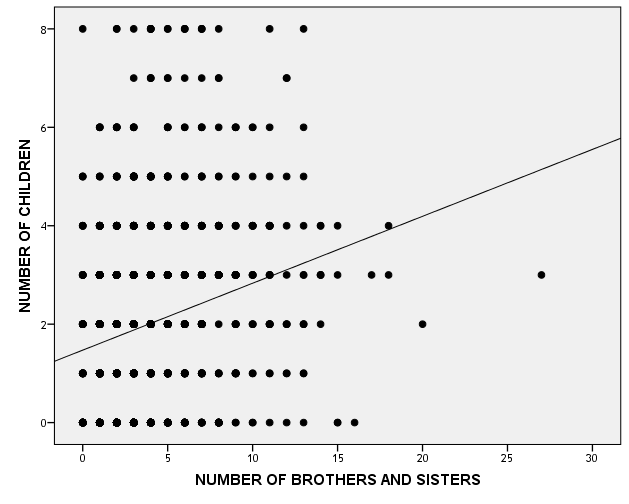
**CHAPTER 11 SPSS PROBLEMS SOLUTIONS**

1.

a. As the number of siblings a respondent has increases, the number of children they have also increases. There is quite a bit of scatter about the regression line, so there isn’t a very strong relationship between the two variables. A linear relationship appears to be a reasonable fit to the relationship. This is a positive relationship.



b. The output from SPSS is omitted. The intercept is 1.473 and the slope is .136. The coefficient of determination is .055 and the correlation coefficient is .234.

c. Predicted number of children for an individual with 3 siblings:

*Ŷ* = 1.473 + .136(3) = 1.88

d. For someone with 0 siblings, 1.473.

e. SPSS can calculate the predicted value for each respondent by clicking on the “Save” button in the “Regression dialog” box and clicking on the “Unstandardized” choice in the “Predicted values” box.

2.

a. The slope of the male equation is .131 and the slope of the female equation is .138. The intercept of the male equation is 1.391; of the female equation, 1.541. The values of the coefficient of determination for males and females are .043 and .065, respectively. Note that the intercept for females is higher than for males. The equation for females has a larger coefficient of determination; however both coefficients of determination suggest weak relationships between the number of siblings a respondent has and his or her number of children.

b. For females:

*Ŷ* = 1.541 + .138(2) = 1.82

*Ŷ* = 1.541 + .138(6) = 2.37

For males:

*Ŷ* = 1.391 + .131(2) = 1.65

*Ŷ* = 1.391 + .131(6) = 2.18

For both predicted values of 2 and 6 siblings, the values for females are greater.

3.

a. The slope of the white equation is .120 and the slope of the black equation is .132. The intercept of the white equation is 1.536; of the black equation, 1.625. The values of the coefficient of determination for whites and blacks are .039 and .062, respectively. Note that the slope for blacks is higher than for whites. The same can be said for the intercept.

The equation for blacks has a larger coefficient of determination; however both coefficients of determination suggest weak relationships between the number of

siblings a respondent has and the number of his or her children.

b. For whites:

*Ŷ* = 1.536 + .120(1) = 1.66

*Ŷ* = 1.536 + .120(4) = 2.02

*Ŷ* = 1.536 + .120(7) = 2.38

For blacks:

*Ŷ* = 1.625 + .132(1) = 1.76

*Ŷ* = 1.625 + .132(4) = 2.15

*Ŷ* = 1.625 + .132(7) = 2.55

4.

a. The slope of the married equation is .099 and the slope of the divorced equation is .112. The intercept of the married equation is 1.969; of the divorced equation, 1.675. The values of the coefficient of determination are very similar (.038 for married respondents and .041 for divorced respondents). Note that the slope for divorced respondents is slightly higher than for married respondents, yet the married respondents have a higher intercept than the divorced respondents. The equation for divorced respondents has a slightly larger coefficient of determination; however both coefficients of determination suggest a weak relationship between the number of siblings a respondent has and the number of his or her children.

b. For married:

*Ŷ* = 1.969 + .099(1) = 2.07

*Ŷ* = 1.969 + .099(4) = 2.37

*Ŷ* = 1.969 + .099(7) = 2.66

For divorced:

*Ŷ* = 1.675 + .112(1) = 1.79

*Ŷ* = 1.675 + .112(4) = 2.12

*Ŷ* = 1.675 + .112(7) = 2.46

c. Overall, on the basis of the coefficients of determination from both analyses, neither equation predicts number of children all that well. The coefficient of determination (*r*2) for both married and divorced respondents is very similar and small (.038 and .041).

5.

a. The output from SPSS is omitted. The intercept is 10.158 and the slope is .318. The coefficient of determination is .187 and the correlation coefficient is .433. For every additional year that respondent’s father spends in school, we can expect an increase in the respondent’s level of education of .318 years. When a respondent’s father has zero years of education, we can expect the respondent to have nearly 10.158 years of school. When using a respondent’s father’s education to predict a respondent’s education, we improve our prediction by approximately 18.7 percent, indicating a moderate positive relationship (*r* = .433).

b. The output from SPSS is omitted. The intercept is 9.519 and the slopes are .223 and .159 for a father’s education (PAEDUC) and a mother’s education (MAEDUC), respectively. The coefficient of determination is .201 and the correlation coefficient is .448. When both a respondent’s father and mother have zero years of education, we expect a respondent to have 9.519 years of education, or about 9.5 years of school. Holding a respondent’s father’s education constant, for every additional year that respondent’s mother spends in school, we can expect an increase in the respondent’s level of education of .159 years, or about two months.

Holding a respondent’s mother’s education constant, for every additional year that respondent’s father spends in school, we can expect an increase in the respondent’s level of education of .223 years, or nearly three months. The value of R-Square is .201; thus, 20 percent of the variation in a respondent’s level of education can be explained by taking into account the respondent’s father’s and mother’s level of education.

c. Our prediction did improve by taking into account a respondent’s mother’s level of education. However, that said, our prediction only slightly improved. We only improved our prediction by 1.4 percent ().

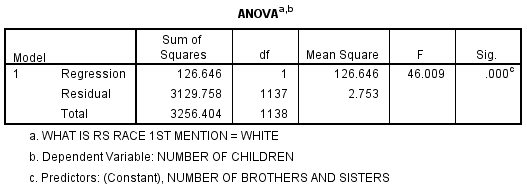
d. Father w/ 12 years of education:

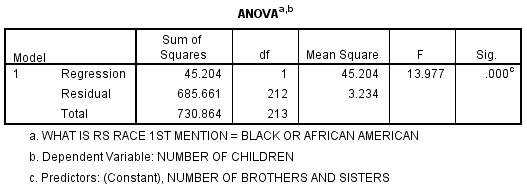
*Ŷ* = 10.518 + .318(12) = 14.33

Father and mother, each w/ 12 years of education:

*Ŷ* = 9.519 + .223(12) + .159(12) = 14.10

6. Based on the *F*-statistics for Whites (46.009) and Blacks (13.977) as well as the corresponding *p*-values (.000 for both Whites and Blacks), we can reject the null hypotheses that *r*2 = 0 in both populations. In other words, the number of siblings one has does affect the number of children he or she has. This holds true for both White and Black respondents.

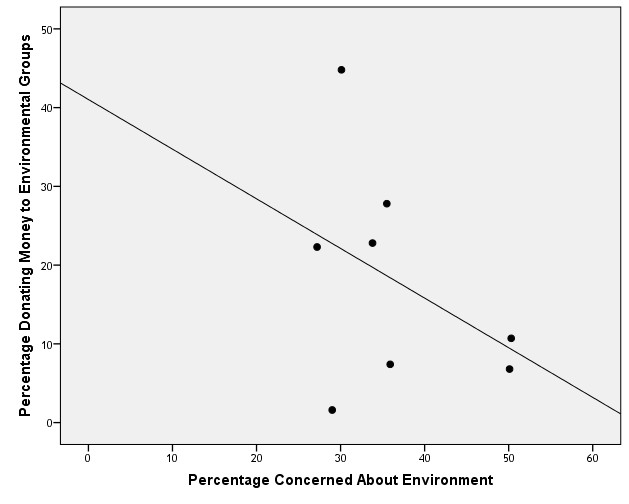




**CHAPTER 11 EXERCISE SOLUTIONS**

1.

a. On the scatterplot below, the regression line has been plotted to make it easier to see the relationship between the two variables.



b. The scatterplot shows there is a general linear relationship between the two variables. There is not a lot of scatter about the straight line describing the relationship. As the percentage of respondents concerned about the environment increases, the percentage of respondents donating money to environmental groups decreases.

c. The Pearson correlation coefficient between the two variables is –.40. This is consistent with the scatterplot indicated a negative relationship between being concerned about the environment and actually donating money to environmental groups.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | ***(1)*** | ***(2)*** | ***(3)*** | ***(4)*** | ***(5)*** | ***(6)*** | ***(7)*** |
|  | ***Percentage Concerned*** | ***Percentage Donating*** |  |  |  |  |  |
| ***State*** | **X** | **Y** |  |  |  |  |  |
| United States | 33.8 | 22.8 | -2.69 | 7.22 | 4.77 | 22.80 | -12.83 |
| Austria | 35.5 | 27.8 | -0.99 | 0.98 | 9.77 | 95.55 | -9.67 |
| Netherlands | 30.1 | 44.8 | -6.39 | 40.80 | 26.77 | 716.90 | -171.06 |
| Slovenia | 50.3 | 10.7 | 13.81 | 190.79 | -7.33 | 53.66 | -101.23 |
| Russia | 29.0 | 1.6 | -7.49 | 56.06 | -16.43 | 269.78 | 123.06 |
| Philippines | 50.1 | 6.8 | 13.61 | 185.30 | -11.23 | 126.00 | -152.84 |
| Spain | 35.9 | 7.4 | -0.59 | 0.35 | -10.63 | 112.89 | 6.27 |
| Denmark | 27.2 | 22.3 | -9.29 | 86.26 | 4.27 | 18.28 | -39.67 |
|  | ∑*X* = 291.9 | ∑*Y* = 144.2 | -0.02a | 567.76 | 0.04a | 1,415.85 | -357.97 |
| a. Answers may differ slightly due to rounding. | | | | | | | |

2.

The scatterplot of number of children by education indicates a negative relationship between the two variables. The negative coefficient (.-155) in the “Coefficients” table confirms this. That is, as a person has more schooling, the fewer children they have. Specifically, for each additional year of schooling a person has, they have .155 fewer children. The value of the intercept (4.06) indicates that a person with 0 years of schooling will have about 4 children. The value for *r* (.279) indicates a moderate relationship between level of education and number of children. The value of *r*2 is .078. In other words, by only knowing level of education, we can predict 7.8% of the variance in the number of children a person has.

3.

a. The correlation coefficient is -0.45.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | ***(1)*** | ***(2)*** | ***(3)*** | ***(4)*** | ***(5)*** | ***(6)*** | ***(7)*** |
| ***Country*** | ***GNP per Capita*** | ***Percentage Willing to Pay*** |  |  |  |  |  |
| ***State*** | **X** | **Y** |  |  |  |  |  |
| United States | 29.24 | 44.9 | 2.72 | 7.40 | -1.64 | 2.69 | -4.46 |
| Ireland | 18.71 | 53.3 | -7.81 | 61.00 | 6.76 | 45.70 | -52.80 |
| Netherlands | 24.78 | 61.2 | -1.74 | 3.03 | 14.66 | 214.92 | -25.51 |
| Norway | 34.31 | 40.7 | 7.79 | 60.68 | -5.84 | 34.11 | -45.49 |
| Sweden | 25.58 | 32.6 | -0.94 | 0.88 | -13.94 | 194.32 | 13.10 |
|  | ∑*X* = 132.62 | ∑*Y* = 232.7 | -0.02a | 132.99 | 0.04a | 491.74 | -115.16 |
| a. Answers may differ slightly due to rounding. | | | | | | | |

b. A correlation coefficient of -0.45 means that relatively high values of GNP are moderately negatively assoicated with low values of percentage of residents willing to pay higher prices to protect the environment.

4.

From Figure 11.16, we know that the *F* ratio is 77.757. The *p* value is .000 (technically *p*<.001). Since SPSS reports the exact *p* value, we are able to say that the relationship between education and television viewing is significant at the .001 level.

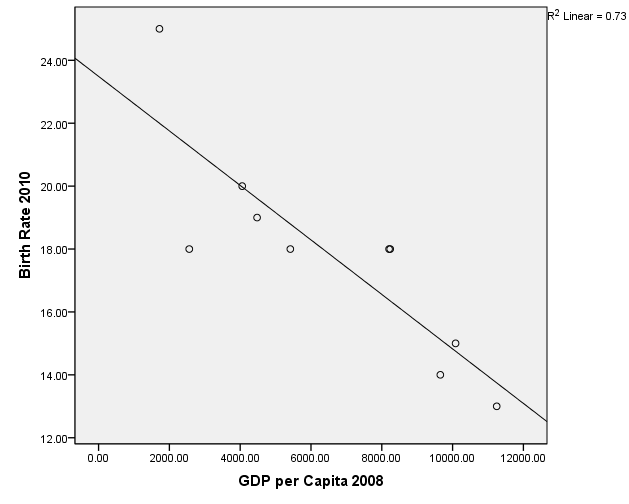
5. Though somewhat difficult to visually determine, there is evidence within the scatterplot that indicates a negative relationship between education and hours/day spent watching TV. Looking to the regression equation output, this is confirmed by the negative coefficient (-.231). This means that for each year of education a person has, the number of hours they spend watching TV per day decreases by 0.231 hours. The value of the intercept (6.130) indicates that a person with 0 years of education spends 6.130 hrs/day watching television. The value for *r* (.268) suggests a moderate relatioship between education and hours/day spent watching television. The value of *r*2 is .072; thus, 7.2 percent of the variation in hours/day spent watching TV can be explained by taking into account a person’s level of education.

6.

The scatterplot identifies a positive relationship between respondent level of education and mother’s level of education (as mother’s level of education increases, so does their children’s). The postive coefficient (.326) in the “Coefficients” table supports this: for each year of education a respondent’s mother has, their level of education increases by .326 years (approximately 4 months). The value of the intercept (9.988) indicates that a person whose mother has 0 years of education is predicted to complete about 10 years of schooling. The value for *r* (.383) suggests a moderate relatioship between mother’s level of education and respondent level of education. The value of *r*2 is .146. In other words, by only knowing mother’s level of education, we can predict 14.6% of the variance in respondent level of education.

7.

a. The scatterplot for GDP per capita and birth rate can be summarized with a straight line or linear relationship.



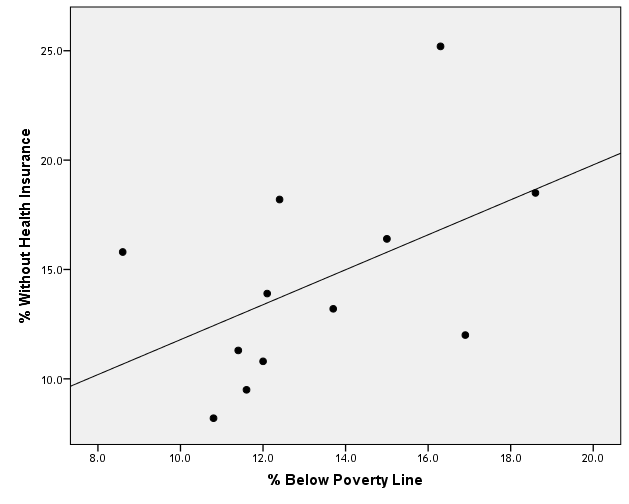
b. Students have to be careful and not calculate the coefficient of determination first and then take the square root to get *r* because they won’t then get the negative sign for *r*.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | ***(1)*** | ***(2)*** | ***(3)*** | ***(4)*** | ***(5)*** | ***(6)*** | ***(7)*** |
|  | ***GDP Per Capita*** | ***Birth Rate*** |  |  |  |  |  |
| ***Country*** | **X** | **Y** |  |  |  |  |  |
| Argentina | 8,236 | 18 | 1,670.5 | 2,790,570.25 | 0.2 | 0.04 | 334.1 |
| Bolivia | 1,720 | 25 | -4,845.5 | 23,478,870.25 | 7.2 | 51.84 | -34,887.6 |
| Brazil | 8,205 | 18 | 1,639.5 | 2,687,960.25 | 0.2 | 0.04 | 327.9 |
| Chile | 10,084 | 15 | 3,518.5 | 12,379,842.25 | -2.8 | 7.84 | -9,851.8 |
| Colombia | 5,416 | 18 | -1,149.5 | 1,321,350.25 | 0.2 | 0.04 | -229.9 |
| Ecuador | 4,056 | 20 | -2,509.5 | 6,297,590.25 | 2.2 | 4.84 | -5,520.9 |
| Paraguay | 2,561 | 18 | -4,004.5 | 16,036,020.25 | 0.2 | 0.04 | -800.9 |
| Peru | 4,477 | 19 | -2,088.5 | 4,361,832.25 | 1.2 | 1.44 | -2,506.2 |
| Uruguay | 9,654 | 14 | 3,088.5 | 9,538,832.25 | -3.8 | 14.44 | -11,736.3 |
| Venezuela | 11,246 | 13 | 4,680.5 | 21,907,080.25 | -4.8 | 23.04 | -22,466.4 |
|  | ∑*X*=65,655 | ∑*Y* = 178 | 0.0a | 100,799,948.50 | 0.0a | 103.60 | -87,338 |
| a. Answers may differ slightly due to rounding. | | | | | | | |

c. The relationship between GDP per capita and birth rate is strong and is negative, with GDP explaining about 73.1% of the variation in birth rate. As a country becomes wealthier, couples have fewer children.

8.

a. A straight line does seem to fit the data, as shown in the scatterplot.



b. The equation, , supports the assertion that a straight line best fits these data. In fact, *b* is 0.8, which indicates that for a 1% increase in those living below the poverty line, there is a corresponding 0.8% increase in the percentage of people without health insurance. Though the straight line fits the data reasonably well, Texas has a much higher rate of uninsured than would be predicted from its poverty rate.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | ***(1)*** | ***(2)*** | ***(3)*** | ***(4)*** | ***(5)*** | ***(6)*** | ***(7)*** |
|  | ***% Below Poverty*** | ***% Without Health Insurance*** |  |  |  |  |  |
| ***State*** | **X** | **Y** |  |  |  |  |  |
| Alabama | 16.9 | 12.0 | 3.62 | 13.08 | -2.42 | 5.84 | -8.74 |
| California | 12.4 | 18.2 | -0.88 | 0.78 | 3.78 | 14.31 | -3.34 |
| Idaho | 12.1 | 13.9 | -1.18 | 1.40 | -0.52 | 0.27 | 0.61 |
| Louisiana | 18.6 | 18.5 | 5.32 | 28.27 | 4.08 | 16.67 | 21.71 |
| New Jersey | 8.6 | 15.8 | -4.68 | 21.93 | 1.38 | 1.91 | -6.48 |
| New York | 13.7 | 13.2 | 0.42 | 0.17 | -1.22 | 1.48 | -0.51 |
| Pennsylvania | 11.6 | 9.5 | -1.68 | 2.83 | -4.92 | 24.17 | 8.28 |
| Rhode Island | 12.0 | 10.8 | -1.28 | 1.65 | -3.62 | 13.08 | 4.64 |
| South Carolina | 15.0 | 16.4 | 1.72 | 2.95 | 1.98 | 3.93 | 3.40 |
| Texas | 16.3 | 25.2 | 3.02 | 9.10 | 10.78 | 116.28 | 32.53 |
| Washington | 11.4 | 11.3 | -1.88 | 3.55 | -3.12 | 9.71 | 5.87 |
| Wisconsin | 10.8 | 8.2 | -2.48 | 6.17 | -6.22 | 38.65 | 15.44 |
|  | ∑*X* = 159.4 | ∑*Y* = 173.0 | 0.0a | 91.88 | 0.0a | 246.32 | 73.41 |
| a. Answers may differ slightly due to rounding. | | | | | | | |

c. 5 = 3.796 + 0.8*X*

*X* = 1.51%. So about 1.51% need to be living below the poverty level to have only 5% without health insurance.

d. You cannot go outside the scope of your data. That said, it is interesting to see how closely poverty and lack of health insurance are related. Although we cannot generalize statistically, this example does help us see just how closely the two are related. Thus, as we go about our research, this may be a consideration to keep in mind for future studies.

9.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | ***(1)*** | ***(2)*** | ***(3)*** | ***(4)*** | ***(5)*** | ***(6)*** | ***(7)*** |
|  | ***GNP Per Capita*** | ***% Willing to Pay Higher Taxes*** |  |  |  |  |  |
| ***Country*** | **X** | **Y** |  |  |  |  |  |
| Canada | 19.71 | 24.0 | 1.51 | 2.28 | -3.81 | 14.52 | -5.75 |
| Chile | 4.99 | 29.1 | -13.21 | 174.50 | 1.29 | 1.66 | -17.04 |
| Finland | 24.28 | 12.0 | 6.08 | 36.97 | -15.81 | 249.96 | -96.12 |
| Ireland | 18.71 | 34.3 | .51 | .26 | 6.49 | 42.12 | 3.31 |
| Japan | 32.35 | 37.2 | 14.15 | 200.22 | 9.39 | 88.17 | 132.87 |
| Latvia | 2.42 | 17.3 | -15.78 | 249.01 | -10.51 | 110.46 | 165.85 |
| Mexico | 3.84 | 34.7 | -14.36 | 206.21 | 6.89 | 47.47 | -98.94 |
| Netherlands | 24.78 | 51.9 | 6.58 | 43.30 | 24.09 | 580.33 | 158.51 |
| Norway | 14.60 | 22.8 | -3.60 | 12.96 | -5.01 | 25.10 | 18.04 |
| Portugal | 34.31 | 17.1 | 16.11 | 259.53 | -10.71 | 114.70 | -172.54 |
| Russia | 10.67 | 29.9 | -7.53 | 56.70 | 2.09 | 4.37 | -15.74 |
| Spain | 2.66 | 22.2 | -15.54 | 241.49 | -5.61 | 31.47 | 87.18 |
| Sweden | 14.10 | 19.5 | -4.10 | 16.81 | -8.31 | 69.06 | 34.07 |
| Switzerland | 25.58 | 33.5 | 7.38 | 54.46 | 5.69 | 32.38 | 41.99 |
| United States | 39.98 | 31.6 | 21.78 | 474.37 | 3.79 | 14.36 | 82.55 |
|  | ∑*X*=272.98 | ∑*Y*=417.10 | 0.0a | 2,029.08 | 0.0a | 1,426.13 | 318.23 |
| a. Answers may differ slightly due to rounding. | | | | | | | |

a. The equation is, *Ŷ* = 24.95 + .157*X*, indicating a positive relationship between GNP per capita and willingness to pay more in taxes.

\* Answers may differ due to rounding; however, the exact value of these column totals, properly calculated will always be equal to zero.

b. *Ŷ* = 24.95 + .157(3) = 25.42% are willing to pay more in taxes.

*Ŷ* = 24.95 + .157(30) = 29.66% are willing to pay more in taxes.

10.

a. False. Both *b* and *r* will always have the same sign because both tell us the direction of the relationship.

b. Both *a* and *b* refere to changes in the dependent variable.

c. The coefficient of determination, *r*2, is a PRE measure. PRE stands for proportional reduction of error. This means that *r*2 indicates the extent to which prediction error is reduced when we take into account the independent variable in our predictions.

d. False. The regression equation serves many functions, most notably to make predictions. Whether a regression equation models a causal relationship is a matter of meeting the causal requirements as they were discussed in Chapter 1.

Regression equations are commonly used to model relationships wherein researchers look for how changes in one or more variables (referred to as the independent variables) correspond with changes in another variable (referred to as the dependent variable).

11. a. The Pearson’s correlation coefficient of .221 indicates a weak, positive relationship between one’s number of siblings and the number of children they wish to have.

b. The y-intercept of 2.301 means that an only child will be predicted to think that 2.301 children is the ideal number to have. The slope of .066 indicates that, for each additional brother or sister a respondent has, their ideal number of children is predicted to rise by .066 kids.

12. a. Education has a greater impact on men’s internet habits. Their r2 of .048 indicates that education can explain 4.8% of the variation in internet use for men. In comparison, women’s r2 of .003 indicates that education can explain only .3% of the variation in their internet usage.

b. For those who have received zero years a formal education – an amount that few, if any, respondents reported – men can be predicted to spend -4.301 hours per week online (an impossible amount) and women can be predicted to spend 5.469 hours a week online. Each additional year of education is predicted to cause a .978 hour rise in males’ internet usage and only a .248 hour rise in females’. Generally speaking, uneducated females spend more time online than their male counterparts, but education increases males’ time spent online more quickly than females’.

c. Males = -4.301+.978(16) = 11.347 hours

Females = 5.469 + .248(16) = 9.437 hours