Homework Assignment #8 (due by 11:59pm CST on Sunday, June 27th)

Please study CH12 to finish this homework assignment. You can calculate the answers following steps in the textbook practice (Computing *a* and *b* for the prediction Equation on Page 406-407 and TESTING THE SIGNIFICANCE OF R^2 USING **ANOVA** on Page 419-421) and weekly materials in Blackboard. Also, please familiar with how to interpret the results using your own words following each definition (See interpretation on Page 410-411). Table 12.3 (P.412) provides a clear, sequential example of the steps to follow to calculate the regression coefficients. By following the steps in this table, you not only learn about the calculation of the regression coefficients, but also about the logic of how the regression equation is constructed. Of course, one additional caveat: continue to remind yourself about the issue of rounding error and should develop a general classroom rule to round to a specific decimal point (i.e., **2 decimal points**) for all calculations.

Problem set 1: Use the following data to answer questions 1-4 (30 points):
There is often thought to be a relationship between a person's educational attainment and the number of children he or she has. The hypothesis is that as one's educational level increases, he or she has fewer children. Investigate this conjecture with 20 cases drawn from the 2018 GSS file. The following table displays educational attainment, in years, and the number of children for each respondent.

Education	Children	**Education**	Children
16	1	**14**	2
12	3	**13**	2
16	2	**15**	1
8	6	**12**	2
15	1	**12**	3
14	2	**11**	3
18	1	**12**	3
12	4	**11**	4
24	0	**12**	4
12	3	**12**	4

1. Calculate the Pearson correlation coefficient for these two variables. Does its value support the hypothesized relationship (10 pts)?

EDUC	CHILDS	$(X - \overline{X})$	$(X - \overline{X})^2$	$(Y - \overline{Y})$	$(Y - \overline{Y})^2$	$(X - \overline{X})(Y - \overline{Y})$
16	1					
12	3					
16	2					
8	6					
15	1					
14	2					
18	1					
12	4					
24	0					
12	3					
14	2					
13	2					

15	1					
12	2					
12	3					
11	3					
12	3					
11	4					
12	4					
12	4					
Σ*X*=271	**Σ***Y*=51	**Σ**=	**Σ**=	**Σ**=	**Σ**=	**Σ**=

Mean X = **(Σ**X **/** N**)** =

Mean Y = **(Σ**Y / N) =

Variance (X) = S^2_X = Σ $(X - \overline{X})^2$ / N-1 =

Standard deviation (X) = Sx = square root (_____)=

Variance (Y) = S^2_Y = Σ $(Y - \overline{Y})^2$ / N-1 =

Standard deviation (Y) = Sy = square root (_____)=

Covariance (X,Y) = Sxy = Σ $(X - \overline{X})(Y - \overline{X})$ / N-1 =

r = Sxy / Sx * Sy =

*Answers may differ slightly due to rounding.

2. Calculate the least squares regression equation using education as a predictor variable. What is the value of the slope, b (5 pts)? What is the value of the intercept, a (3 pts)? What is the prediction equation (2 pts)?

3. What is the predicted number of children for a person with a college degree (18 years of education) (5 pts)?

4. Does any respondent actually have this number of children? If so, what is his or her level of education? If not, is this a problem or an indication that the regression equation you calculated is incorrect? Why or why not (5 pts)?

Problem set 2: use the following information to answer questions 5-9 (40 points total):
Let's examine the relationship between GNP per capita and the percentage of respondents willing to pay more in taxes.

5. In this chapter, we used Table 12.3 to illustrate how to calculate the slope and intercept in the regression table. Using table 12.3 as a model, create a similar table using the data below for GNP per capita and the percentage willing to pay higher taxes (15 pts).

GNP per Capita (*X*)	% Willing to Pay Higher Taxes (*Y*)	$(X - \overline{X})$	$(X - \overline{X})^2$	$(Y - \overline{Y})$	$(Y - \overline{Y})^2$	$(X - \overline{X})(Y - \overline{Y})$
23.70	24.50					
14.10	22.20					
20.33	22.20					
8.37	28.90					
22.28	14.00					
18.71	24.30					
32.35	37.20					
29.24	31.60					
3.84	34.70					
27.78	**24.80**					
14.60	31.10					
34.31	22.80					
10.67	15.10					
2.66	29.90					
14.10	22.20					
25.58	19.50					
39.98	33.50					
Σ*X*=342.60	**Σ***Y*=438.50	**Σ***X*=	**Σ**=	**Σ**=	**Σ**=	**Σ**=

Mean *X* = **Σ***X* / N =

Mean *Y* = **Σ***Y* / N =

Variance (X) = $S_X^2 = \frac{\sum (X - \overline{X})^2}{N - 1} =$

Standard deviation (*X*) = Sx = square root (_____)=

Variance (*Y*) = $S_Y^2 = \frac{\sum (Y - \overline{Y})^2}{N - 1} =$

Standard deviation (*Y*) = Sy = square root (_____)=

Covariance (*X,Y*) = $S_{YX} = \frac{\sum (X - \overline{X})(Y - \overline{Y})}{N - 1} =$

*Answers may differ due to rounding (2 decimal points); however the exact value of these column totals, properly calculated will always be equal to zero.

Sources: The World Bank Group, *Development Education Program Learning Model: Economics, GNP per Capita*, 2004. International Social Survey Programme, 2000.

6. From the table that you created in #5, calculate *a*, *b,* and write out the regression equation (i.e., prediction equation) (6 pts).

7. Calculate and interpret error type, E_2 (10 pts).

% Willing to Pay Higher Taxes (Y)	Predicted % Willing to Pay Higher Taxes ($\hat{Y}$)	$(Y- \hat{Y})$	$(Y- \hat{Y})^2$
24.50	26.01		
22.20	25.89		
22.20	25.89		
28.90	26.23		
14.00	25.48		
24.30	26.00		
37.20	26.64		
31.60	26.36		
34.70	26.52		
24.80	26.02		
31.10	26.34		
22.80	25.92		
15.10	25.54		
29.90	26.28		
22.20	25.89		
19.50	25.76		
33.50	26.46		
			Σ=

8. Using your answer from #7, calculate the PRE measure, r^2. Interpret (3 pts).

9. About what percentage of citizens are willing to pay higher taxes for a country with a GNP per capita of 8.0 (i.e., $8,000)? For a GNP per capita of 80.0 (i.e., $80,000) (6 pts)?

Problem set 3: use the following information to answer questions 10-12 (30 points total):

10. Calculate the Y-intercept, a, of the regression equation using the following information: (10 pts)
    Mean number of years of schooling: 16 (X)
    Mean monthly income: $3,580 (Y)
    Slope, b, of the regression equation: $210 (b, slope)

11. Consider the following linear regression prediction equation: Y = 10 + -1x. If x = 5, what is the predicted value of, Y, for this observation (10 pts)?

12. Consider the following linear regression prediction equation: Y = 10 + -1x. If the predicted value of, Y, for an observation is 5, calculate the residual for an observation where the observed value, Y, is equal to 7 (10 pts)?