AP Statistics Name_____
Unit 02 – Bivariate Data Period_____
Homework #5 – Transforming Data Review

For each problem 1-3, choose the best-fit answer.

1. If (12, 60) is an influential point for the regression line $\hat{y} = 7.908 + 4.098x$, then which of the following must be true?
   a) removal of (12, 60) will improve r
   b) removal of (12, 60) will not affect r
   c) removal of (12, 60) will change the value of the slope of the regression line
   d) (12, 60) has a large residual
   e) none of these

2. Suppose a data set is transformed using $(x, y) \rightarrow (x, \log y)$ and a least squares linear regression procedure is performed on the transformed data. If the residual plot of this regression shows a curved pattern, which of the following is an appropriate conclusion?
   a) A quadratic model should be used with the original data
   b) A square root transformation should be applied to the transformed data
   c) The correlation coefficient of the set of transformed data is zero
   d) The exponential transformation is not appropriate
   e) none of these are appropriate

3. After data are collected from an agricultural experiment, suppose a transformation is performed on the bivariate set (inches of water, total plant growth). If the linear regression for the transformed data has the equation:

$$\mathbf{\log(\widehat{growth}) = 0.7 + 1.93 \log(water)}$$

The regression model of the original data is:
   a) $\widehat{growth} = 0.7 + 1.93(water)$
   b) $\widehat{growth} = 5.01 + 1.93(water)$
   c) $\widehat{growth} = (5.01)(1.93^{water})$
   d) $\widehat{growth} = 5.01(water^{1.93})$
   e) none of these

4. Earthquakes are among the most damaging kinds of natural disasters. The size of an earthquake is generally reported as a rating on the Richter scale-usually a number between 1 and 9. That Richter scale rating indicates the energy released by the shaking of the ground and the height of the shock waves recorded on seismographs.

   The data in the following table show Richter scale ratings and amounts of energy released for six earthquakes.

Earthquake Location	Richter Scale Rating	Energy (in sextillion ergs)
San Francisco, CA, 1906	8.25	1500
Yugoslavia, 1963	6.0	0.63
Alaska, 1964	8.6	5000
Peru, 1970	7.8	320
Italy, 1976	6.5	3.5
Loma Prieta, CA, 1989	7.1	28

a) Use your calculator to make a scatter plot for this data. Explain which model would be a good model to use and why.

b) Write a linear regression model for this data.

c) How confident would you be with predicting the Energy if the Richter Scale Rating was 6.3? Why?

d) Use the model to estimate energy released by earthquakes listed in the following chart.

Earthquake Location	Richter Scale Rating	Energy (in sextillion ergs)
Quetta, India, 1906	7.5	
Kwanto, Japan, 1923	8.2	
Chillan, Chile, 1939	7.75	
Agadir, Morocco, 1960	5.9	
Iran, 1968	7.4	
Tangshan, China, 1976	7.6	
Northridge, CA, 1994	6.7	
Kobe, Japan, 1995	7.2	

5. The femur is a large bone found in the leg or hind limb of an animal. Scientists use the circumference of an animal's femur to estimate the animal's weight. The bale at the right shows the femur circumference C (in millimeters) and the weight W (in kilograms) of several animals.

a) Use your calculator to make a scatter plot for this data. Explain which model would be a good model to use and why.

b) Write a linear regression model for this data.

c) Find the inverse equation for the regression model.

Animal	C (mm)	W (kg)
Meadow mouse	5.5	0.047
Guinea pig	15	0.385
Otter	28	9.68
Cheetah	68.7	38
Warthog	72	90.5
Nyala	97	134.5
Grizzly bear	106.5	256
Kudu	135	301
Giraffe	173	710

▶ Source: Zoological Society of London

d) Predict the weight of an animal whose femur has a circumference of 21 mm.