

1. 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 of the transformed data has the equation:

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

The regression model of the original data is:

- a. $\widehat{\text{growth}} = 0.7 + 1.93(\text{water})$
- b. $\widehat{\text{growth}} = 5.01 + 1.93(\text{water})$
- c. $\widehat{\text{growth}} = (5.01)(1.93)^{\text{water}}$
- d. $\widehat{\text{growth}} = 5.01(\text{water})^{1.93}$
- e. none of these

$$\begin{aligned} 10^{\text{Log}(\widehat{\text{growth}})} &= 10^{0.7 + 1.93 \text{Log Water}} \\ \widehat{\text{growth}} &= 10^{0.7} \cdot 10^{1.93 \log \text{water}} \\ \widehat{\text{growth}} &= 5.01 (\text{water})^{1.93} \end{aligned}$$

2. 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

$$\hat{y} = 7.908 + 4.098(12)$$

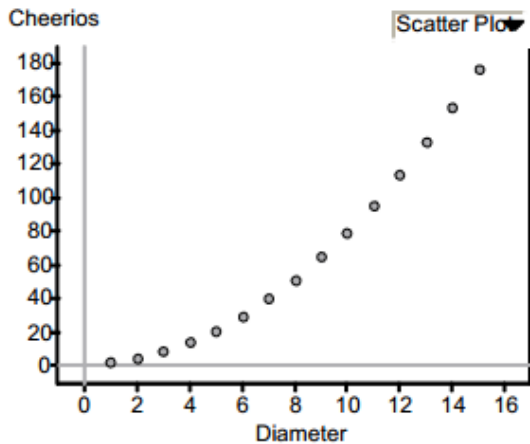
$$\hat{y} = 57.08$$

$$\text{Resid} = \text{Act} - \text{Pred}$$

$$\text{Resid} = 60 - 57 = 3 \text{ Not Too large}$$

3. The residual value of (\bar{x}, \bar{y}) in a linear regression is
- a. negative
 - b. 0
 - c. positive
 - d. dependent on the value of r
 - e. the value cannot be determined

Students in a statistics class drew circles of varying diameters and counted how many Cheerios[®] could be placed in the circle. The scatterplot shows the results.



The students wanted to determine an appropriate equation for the relationship between diameter and the number of Cheerios[®]. The students decided to transform the data to make it appear more linear before computing a least-squares regression line. Which of the following transformations would be reasonable for them to try?

Looks Quad or Exp Growth so try \sqrt{y} or $\log y$

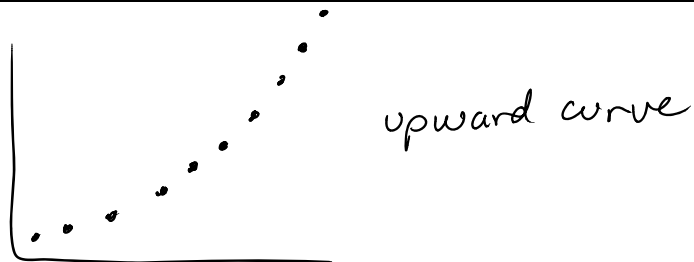
- I. Take the square root of the number of Cheerios[®].
- II. Cube the number of Cheerios[®].
- III. Take the log of the number of Cheerios[®].
- IV. Take the log of the diameter.

4. (a) I and II (b) I and III (c) II and III (d) II and IV (e) III and IV

5. Complete a linear regression analysis for the following age and income data as indicated

Age (years)	20	25	30	35	40	45	50	55	60
Income (\$1,000)	18.5	23.6	29.8	38.5	49	64.1	78.5	102.0	130.8

- a. Sketch a scatterplot of the data.



- b. Determine whether a linear model is appropriate. Explain.

The scatterplot has a curve, so a linear model is NOT appropriate

- c. If a linear model is not appropriate, write the non-linear regression equation found by re-expressing the data.

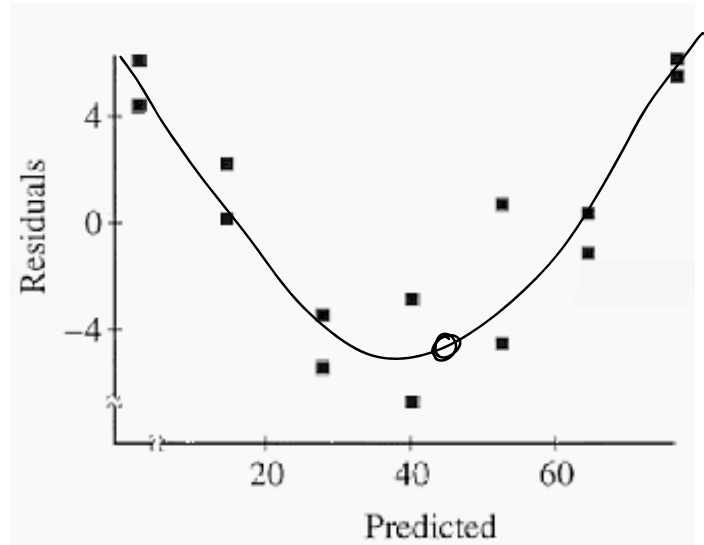
$$\widehat{\log(\text{Income})} = .8418 + .021(\text{age})$$

$$\widehat{\text{Income}} = 6.947(1.05)^{\text{age}}$$

6. In a study of the application of a certain type of weed killer, 14 fields containing large numbers of weeds were treated. The weed killer was prepared at seven different strengths by adding 1, 1.5, 2, 2.5, 3, 3.5, or 4 teaspoons to a gallon of water. Two randomly selected fields were treated with each strength of weed killer. After a few days, the percentage of weeds killed on each field was measured. The computer output obtained from fitting a least squares regression line to the data is shown below. A plot of the residuals is provided as well.

Dependent variable is: percent killed
 R squared = 97.2% R squared (adjusted) = 96.9%
 s = 4.505 with 14 - 2 = 12 degrees of freedom

Variable	Coefficient	s.e. of Coeff	t-ratio	Prob
Constant	-20.5893	3.242	-6.35	≤ 0.0001
No. Teaspoons	24.3929	1.204	20.3	≤ 0.0001



- a. What is the equation of the least squares regression line given by this analysis? Define any variables used in this equation.
- percent weeds killed = 24.3929(No of Teaspoons) - 20.5893*

- b. If someone uses this equation to predict the percentage of weeds killed when 2.6 teaspoons of weed killer are used, which of the following would you expect?

- The prediction will be too large.
- The prediction will be too small.
- A prediction cannot be made based on the information given on the computer output.

percent weeds killed = 24.3929(2.6) - 20.5893
= 42.8

Explain your reasoning.

predicted = 42.8

Residual from Graph is about -4

Resid = Act - Pred

-4 = Act - 42.8

Actual = 38.8

Prediction was too large w/ a negative residual.