

_____ 1. A study is conducted to determine if one can predict the yield of a crop based on the amount of fertilizer used. The response variable in this study is:

- a. Yield of crop
- b. Amount of fertilizer
- c. The experimenter
- d. Amount of rainfall

The following information shows a season’s home run totals and batting averages for 10 MLB players from both the National League (N) and the American League (A). Use this information to answer questions 2-10.

<u>Player</u>	<u>Average(x)</u>	<u>Homeruns(y)</u>
A	280	32
N	290	34
N	341	46
N	320	48
A	220	25
A	245	20
N	270	30
N	310	52
A	260	48
A	300	38

2. Design a scatterplot below. Differentiate between American and National League players.

3. Calculate “r.”

r = _____

4. Is the association positive or negative? _____

_____ 5. The linear relationship would best be described as:

- a. Extremely strong
- b. Moderately strong
- c. Extremely weak
- d. Non-existent

6. Write the equation of the regression line using the form $y = a + bx$.

7. Draw the equation line on your scatterplot. Be as accurate as possible.

8. Using the regression equation, predict how many home runs a player would hit if he had an average of 380.

9. Suppose a player had an average of 100. Why is using the regression equation to predict his home runs an example of extrapolation?

10. What percentage of the variation is explained by the linear relationship?

_____ 11. Which of the following statements is correct concerning correlation?

- a. Changing the units of measurement of x or y does not change “r.”
- b. If “r” is negative, then the association between the variables must be weak.
- c. “r” must have the same units of measurement as the x variable, but not the y.
- d. If $r=1$, then no line fits the data on the scatterplot.

_____ 12. Which of the following statements is NOT true about the linear regression equation?

- a. No distinction is needed between explanatory and response variables.
- b. A close connection exists between slope and correlation.
- c. The line always passes through (\bar{X}, \bar{Y}) .
- d. The square of the correlation is the percentage of the variation explained by the linear relationship.

13. What is a lurking variable?

14. Suppose that a study done by a Denver Lutheran student shows a strong negative correlation between number of cigarettes smoked per week and GPA. The student then publishes a statement claiming that “smoking causes bad grades.” What problems might there be with this statement?

Bonus: Show which league, American or National, has a higher correlation between average and home runs.

Bonus: Explain why the assumption that a higher batting average means more home runs is generally not true in baseball.