Statistics 13V NAME:_____
Sample Quiz 3 Last six digits of Student ID#:_____

Inserted answers in italics; multiple choice answers in **bold.** *Your test may have questions and answers in a different order. Different numerical values are also used; answers are given for both versions with each question.*

1. If X = height measured in inches and Y = height measured in centimeters for 126 people (as shown in the Self-assess quiz in Unit A-8) what is the correlation between X and Y? Explain.

   The correlation is 1.0. If height in either measurement is known, height in the other one is known exactly. This is a deterministic relationship.

2. A scatter plot of number of teachers and number of people with college degrees for cities in California reveals a positive association. The most likely explanation for this positive association is:
   A. Teachers encourage people to get college degrees, so an increase in the number of teachers is causing an increase in the number of people with college degrees.
   B. Larger cities tend to have both more teachers and more people with college degrees, so the association is explained by a third variable, the size of the city.
   C. Teaching is a common profession for people with college degrees, so an increase in the number of people with college degrees causes an increase in the number of teachers.
   D. Cities with higher incomes tend to have more teachers and more people going to college, so income is a confounding variable, making causation between number of teachers and number of people with college degrees difficult to prove.

Questions 3 to 5: A group of adults aged 20 to 80 were tested to see how far away they could first hear an ambulance coming towards them. An equation describing the relationship between distance (in feet) and age was found to be: $Distance = 600 - 3 \times Age$

3. Estimate the distance for an individual who is 30 years old.
   Distance = 600 - 3(30) = 600 - 90 = 510 feet.

   Version 2 asked for age=40. Distance = 600 - 3(40) = 600 - 120 = 480 feet.

4. Does the intercept of 600 have a useful interpretation in this example? Explain.

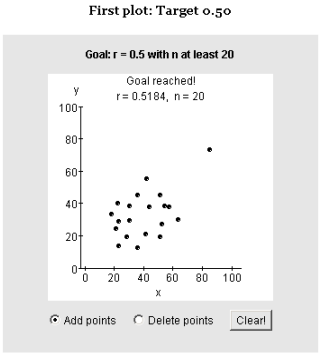
   No. It would be the distance at which someone with age = 0 could hear the ambulance. Age = 0 is far beyond the range of the data, so the intercept has no useful interpretation.

5. Mary is 30 years old and could first hear the ambulance at a distance of 500 feet. Find the residual for Mary and explain what it means in terms of Mary's hearing relative to others.
   Residual = observed value - predicted value = 500 - 510 = -10.

   In Version 2 Mary was 40 and her distance was 470, so residual = 470 - 480 = -10.
   Note that in both cases, the "predicted" distance was found already in Question 3.

   Interpretation: Mary's hearing is not quite as good as would be expected for someone her age. She first heard the ambulance when it was 10 feet closer than would be predicted for someone her age.

6. Shown below is a partially completed attempt at one of the Interactivities in Unit A8, Basics Practice 3. Add *one* more point on the plot that would help achieve the goal of r=0.5 from the current status of r = 0.2077. In the space to the right of the plot, explain why you selected that location to place your point.
   The new picture has an additional point added in the upper right corner. The idea is to add an outlier that "lines up" with the rest of the data, to increase the correlation. (See page 129.)



---

7. The value of a correlation is reported by a researcher to be $r = -0.5$. Which of the following statements is correct?
   A. The x-variable explains 25% of the variability in the *y*-variable.
   B. The x-variable explains −25% of the variability in the *y*-variable.
   C. The x-variable explains 50% of the variability in the *y*-variable.
   D. The x-variable explains −50% of the variability in the *y*-variable.

8. What is the effect of an outlier on the value of a correlation coefficient?
   A. An outlier will always decrease a correlation coefficient.
   B. An outlier will always increase a correlation coefficient.
   C. An outlier might either decrease or increase a correlation coefficient, depending on where it is in relation to the other points.
   D. An outlier will have no effect on a correlation coefficient.

9. One use of a regression line is
   A. to determine if any x-values are outliers.
   B. to determine if any y-values are outliers.
   C. to determine if a change in x causes a change in y.
   D. to estimate the change in y for a one-unit change in x.

10. Past data has shown that the regression line relating the final exam score and the midterm exam score for students who take statistics from a certain professor is:

$$\textit{final exam} = 50 + 0.5 \times \textit{midterm}$$

    One interpretation of the slope is
    A. a student who scored 0 on the midterm would be predicted to score 50 on the final exam.
    B. a student who scored 0 on the final exam would be predicted to score 50 on the midterm exam.
    C. a student who scored 10 points higher than another student on the midterm would be predicted to score 5 points higher than the other student on the final exam.
    D. students only receive half as much credit (.5) for a correct answer on the final exam compared to a correct answer on the midterm exam.