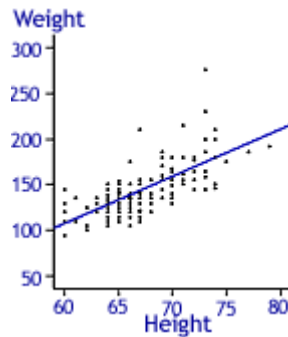


Each question is worth 10 pts. For multiple choice questions circle your answer. For other questions show your work.
 NOTE: Your quiz may have had choices in a different order, and there were two versions of numbers given in Q4.

Questions 1 to 4: The scatterplot below is from CyberStats, Unit D1, Warnings. The data are heights (x) and weights (y) for 126 college students. The equation for the regression line was given in Unit D1 as $\hat{y} = -200 + 5x$.



- Which of the following warnings were the scatterplot and regression line used to illustrate in Unit D1?
 - Males and females should not be used in the same analysis.
 - There is an outlier (a person whose height and weight are about 74 inches and 275 pounds, respectively) that will make the results meaningless.
 - It is dangerous to extrapolate beyond the range of the data, as illustrated by the fact that a child 30 inches tall would be predicted to weigh -50 pounds.***
 - Regression should not be used when the relationship is non-linear.
- Which one of the following is NOT a legitimate interpretation of the slope of 5?
 - Growing one inch will cause a student to gain 5 pounds.***
 - A student who is 1 inch taller than another student is predicted to be 5 pounds heavier.
 - Students who are 71 inches tall are predicted to have an average weight that is 5 pounds more than students who are 70 inches tall.
 - The average increase in weight for a one-inch increase in height is 5 pounds.
- What is the predicted weight for a college student who is 70 inches tall? Show your work.

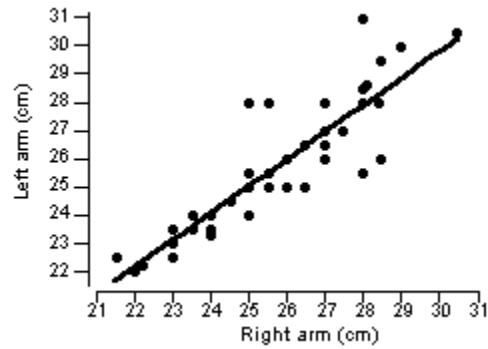
$$\hat{y} = -200 + 5x = -200 + 5(70) = -200 + 350 = 150 \text{ pounds}$$

- One student in the data set was 70 inches tall and weighed 130 pounds. Find the residual for this student, *and* explain in words what it means about the student's weight relative to what would be predicted. (You already found the predicted weight for this student in Question 3.)

Residual = observed value - predicted value = 130 - 150 (found in question 3) = -20 pounds.
In the other version of the exam the student weighed 135 pounds, so the residual was -15 pounds.
Interpretation: This student weighs 20 (or 15) pounds less than the predicted weight for students who are 70-inches tall. You could also say that this student weighs 20 (15) pounds less than the average weight the regression line specifies for students who are 70 inches tall.

- Suppose that in a relationship between x and y , it is found that $r^2 = 0.81$ (or 81%). Which of the following is true?
 - As x increases by one unit, y increases by 0.81.
 - As x increases by one unit, y increases by 0.9.
 - If you increase someone's x value, this will cause an increase in their y value, because the value of r^2 is positive and close to 1.
 - There is either a strong positive correlation or a strong negative correlation between x and y , but we can't tell which it is.***

Questions 6 and 7: The following scatterplot shows the relationship between the left and right forearm lengths (cm) for 55 college students along with the regression line, where y = left forearm length and x = right forearm length.



6. One of the following four choices is the equation for the regression line in the plot. Which one is it?
- $\hat{y} = 1.22 + 0.95x$ (Note that the slope is clearly positive.)
 - $\hat{y} = 1.22 - 0.95x$
 - $\hat{y} = 21.0 + 0.95x$ (Note that 21.0 is about where the line goes through the y -axis when $x = 21$, not $x = 0$.)
 - $\hat{y} = 21.0 - 0.95x$
7. One of the following four choices is the value of the correlation for this situation. Which one is it?
- 0.88
 - 0.00
 - 0.88** (Note that the slope is clearly positive, not negative, so the correlation is as well.)
 - 1.00
8. There is a strong correlation between verbal SAT scores and college GPAs. Which of the following is the most likely reason for this relationship?
- Having high SAT scores gives people the confidence to do well in college, thus causing higher GPAs.
 - The high (low) SAT scores and high (low) GPAs both result from a common cause.** (See p. 177)
 - Both SAT scores and GPA values change over time.
 - There are a few outliers – students with very high SAT scores and GPA values – that inflate the correlation.
9. A regression equation for y = grade point average (GPA) and x = number of classes skipped in a typical week was found based on self-reported data for a sample of $n = 1,673$ at a large northeastern university. The resulting equation is $\text{GPA} = 3.25 - 0.0614 \times \text{Skipped classes}$. One interpretation of the *intercept* is:
- A student who had a GPA of 3.25 would be predicted to skip no classes.
 - A student who skipped no classes would be predicted to have a GPA of 3.25.**
 - A student who skipped one more class in a typical week than another student would be predicted to have a GPA that was 0.0614 lower than the other student.
 - A student who skipped one more class in a typical week than another student would be predicted to have a GPA that was 0.0614 higher than the other student.
10. What impact can an outlier have on a correlation, compared to the same data set without the outlier?
- An outlier that is consistent with the trend of the rest of the data set will inflate the correlation, whereas an outlier that is not consistent with the rest of the data set can deflate the correlation.**
 - An outlier that is consistent with the trend of the rest of the data set will *deflate* the correlation, whereas an outlier that is not consistent with the rest of the data set can *inflate* the correlation.
 - Outliers always *inflate* the correlation.
 - Outliers always *deflate* the correlation.