r^2 HW

1) Each of 100 laboratory rats has available both plain water and a mixture of water and caffeine in their cages. After 24 hours, two measures were recorded for each rat: the amount of caffeine the rat consumed, X, and the rat's blood pressure, Y. The correlation between X and Y was 0.428. Which of the following conclusions is justified on the basis of this study?

(A) The correlation between X and Y in the population of rats is also 0.428.
(B) If the rats stop drinking the water/caffeine mixture, this would cause a reduction in their blood pressure.
(C) About 18 percent of the variation in blood pressure can be explained by a linear relationship between blood pressure and caffeine consumed.
(D) Rats with lower blood pressure do not like the water/caffeine mixture as much as do rats with higher blood pressure.
(E) Since the correlation is not very high, the relationship between the amount of caffeine consumed and blood pressure is not linear.

2) A study was done on the relationship between high school grade point average (GPA) and scores on the SAT. The following 8 scores were from a random sample of students taking the exam:

X (GPA)	3.2	3.8	3.9	3.3	3.6	2.8	2.9	3.5
Y (SAT)	725	752	745	680	700	562	595	730

What percent of the variation in SAT scores is explained by the regression of SAT score on GPA?

a. 62.1%
b. 72.3%
c. 88.8%
d. 94.2%
e. 78.8%

3) A study found a correlation of $r = -0.58$ between hours per week spent watching television and hours per week spent exercising. That is, the more hours spent watching television, the less hours spent exercising per week. Which of the following statements is most accurate?

(a) About one-third of the variation in hours spent exercising can be explained by hours spent watching television.
(b) A person who watches less television will exercise more.
(c) For each hour spent watching television, the predicted decrease in hours spent exercising is 0.58 hrs.
(d) There is a cause-and-effect relationship between hours spent watching television and a decline in hours spent exercising.
(e) 58% of the hours spent exercising can be explained by the number of hours watching television.

4)) Which of the following statements about the correlation r are true?

   I. A correlation of .2 means that 20% of the points are highly correlated.
   II. The square of the correlation measures the proportion of the y-variance that is predictable from a knowledge of x.
   III. Perfect correlation, that is, when the points lie exactly on a straight line, results in $r = 0$.

(A) I only
(B) II only
(C) III only
(D) None of these statements is true.
(E) None of the above gives the complete set of true responses.

5) A study is conducted relating GPA to number of study hours per week, and the correlation is found to be .5. Which of the following are true statements?

   I. On the average, a 30% increase in study time per week results in a 15% increase in GPA.
   II. Fifty percent of a student's GPA can be explained by the number of study hours per week.
   III. Higher GPAs tend to be associated with higher numbers of study hours.

(A) I and II
(B) I and III
(C) II and III
(D) I, II, and III
(E) None of the above gives the complete set of true responses.

Free Response

1) One of the variables that is related to college success (as measured by *GPA*) is socio-economic status. In one study of the relationship, $r^2 = 0.45$. Explain what this means in the context of the problem.

2) The equation of the least-squares line for a set of points is $\hat{y} = -8 + 4.6x$. If 75% of the variation in y is explained by the variation in x, what is the ratio of the standard deviations of the x and y-values?

3) A study of right-handed people found that the regression equation for predicting left-hand strength (measured in kg) from right-hand strength is *left-hand strength* = 7.1 + 0.35 (*right-hand strength*).

   (a) What is the predicted left-hand strength for a right-handed person whose right-hand strength is 12 kg?
   (b) Interpret the intercept and the slope of the regression line in the context of the problem.

4) A study was conducted in a mid-size U.S. city to investigate the relationship between the number of homes built in a year and the mean percentage appreciation for that year. The data for a 5-year period are as follows:

Number	110	80	95	70	55
Percent appreciation	15.7	10	12.7	7.8	10.4

   (a) Obtain the LSRL for predicting appreciation from number of new homes built in a year.
   (b) The following year, 85 new homes are built. What is the predicted appreciation?
   (c) How strong is the linear relationship between number of new homes built and percentage appreciation? Explain.
   (d) Suppose you didn't know the number of new homes built in a given year. How would you predict appreciation?