Name _____ Date _____

AP Statistics Period _____

Unit 3 HW 3

1) As reported in *The New York Times* (September 21, 1994, page C10), a study at the University of Toronto determined that, for every 10 grams of saturated fat consumed per day, a woman's risk of developing ovarian cancer rises 20%. What is the meaning of the slope of the appropriate regression line?

(A) Taking in 10 grams of fat results in a 20% increased risk of developing ovarian cancer.
(B) Consuming 0 grams of fat per day results in a zero increase in the risk of developing ovarian cancer.
(C) Consuming 50 grams of fat doubles the risk of developing ovarian cancer.
(D) Increased intake of fat causes higher rates of developing ovarian cancer.
(E) A woman's risk of developing ovarian cancer rises 2% for every gram of fat consumed per day.

2) As reported in the *Journal of the American Medical Association* (June 13, 1990, page 3031), for a study of ten nonagenarians (subjects were age 90 ± 1), the following tabulation shows a measure of strength (heaviest weight subject could lift using knee extensors) versus a measure of functional mobility (time taken to walk 6 meters). Note that the functional mobility is greater with lower walk times.

Strength (kg):	7.5	6	11.5	10.5	9.5	18	4	12	9	3
Walk time (s):	18	46	8	25	25	7	22	12	10	48

What is the sign of the slope of the regression line and what does it signify?

(A) The sign is positive, signifying a direct cause-and-effect relationship between strength and functional mobility.
(B) The sign is positive, signifying that the greater the strength, the greater the functional mobility.
(C) The sign is negative, signifying that the relationship between strength and functional mobility is weak.
(D) The sign is negative, signifying that the greater the strength, the greater the functional mobility.
(E) The slope is close to zero, signifying that the relationship between strength and functional mobility is weak.

3) Given a set of ordered pairs (x, y) so that $s_x = 1.6$, $s_y = 0.75$, $r = 0.55$. What is the slope of the least-square regression line for these data?

(a) 1.82
(b) 1.17
(c) 2.18
(d) 0.26
(e) 0.78

The following table shows the salaries of 10 employees (in $1000's) along with the purchase price of the cars they drive (in $1000's).

Salary	27	62	42	18	35	20	55	25	27	26
Car Price	21	39	28	11	24	13	36	18	20	16

4) Using the information in the table, what would be the expected value of a car driven by an employee with a salary of $30,000?

A. $20,300
B. $20,500
C. $20,700
D. $20,900
E. $21,100

5) Using the information in the table, what would be the expected value of the salary for an employee who drives a car valued at $30,000?

A. $43,550
B. $44,400
C. $45,650
D. $47,800
E. $49,850

6) A study of stopping distances found that the least squares regression line for predicting mileage (in miles per gallon) from the weight of the vehicle (in hundreds of pounds) was MPG = 32.50 – 0.45(Weight). The mean weight for the vehicles in the study was 2980 pounds. What was the mean MPG in the study?

a. 19.09
b. 15.27
c. –1308.5
d. 18.65
e. 20.33

7) Suppose the LSRL for predicting Weight (in pounds) from Height (in inches) is given by *Weight* = –115 + 3.6 (*Height*). Which of the following statements is correct?

I. A person who is 61 inches tall will weigh 104.6 pounds.
II. For each additional inch of Height, Weight will increase on average by 3.6 pounds.
III. There is a strong positive linear relationship between Height and Weight.

(a) I only
(b) II only
(c) III only
(d) II and III only
(e) I and II only

8) A simple random sample of 35 world-ranked chess players provides the following statistics:

Number of hours of study per day: $\bar{x} = 6.2$, $s_x = 1.3$
Yearly winnings: $\bar{y} = \$208{,}000$, $s_y = \$42{,}000$
Correlation $r = .15$

Based on this data, what is the resulting linear regression equation?

(A) $\widehat{\text{Winnings}} = 178{,}000 + 4850$ Hours
(B) $\widehat{\text{Winnings}} = 169{,}000 + 6300$ Hours
(C) $\widehat{\text{Winnings}} = 14{,}550 + 31{,}200$ Hours
(D) $\widehat{\text{Winnings}} = 7750 + 32{,}300$ Hours
(E) $\widehat{\text{Winnings}} = -52{,}400 + 42{,}000$ Hours

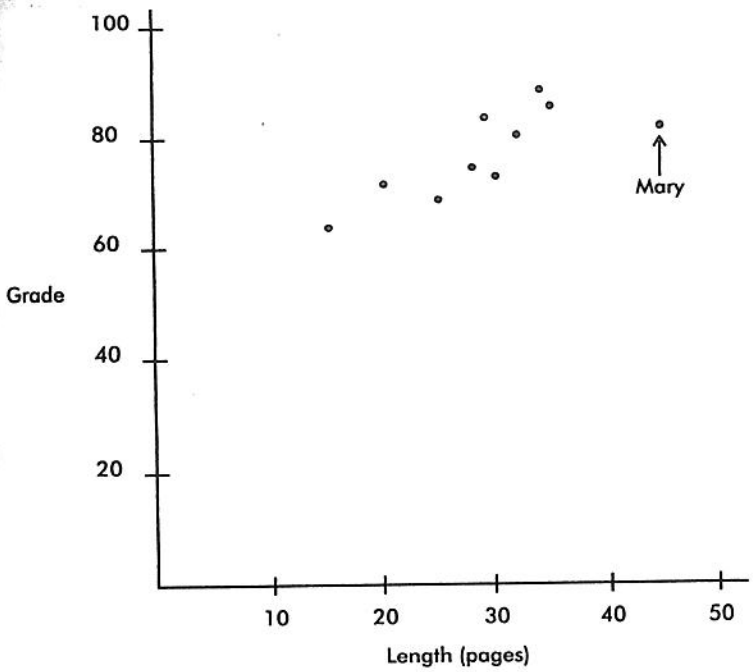
Free Response

1) The shoe sizes and the number of ties owned by ten corporate vice presidents are as follows.

Shoe size, x:	8	9.5	9	11	9	9.5	8.5	9	9	9.5
Number of ties, y:	10	10	8	15	12	13	16	7	12	4

a. Draw a scatterplot for these data.
b. Find the correlation r.
c. Can we find the best-fitting straight-line approximation to the above data? Does it make sense to use this equation to predict the number of ties owned by a corporate executive who wears size 10 shoes? Explain.

2) The following scatterplot shows the grades for research papers for a sociology professor's class plotted against the lengths of the papers (in pages).

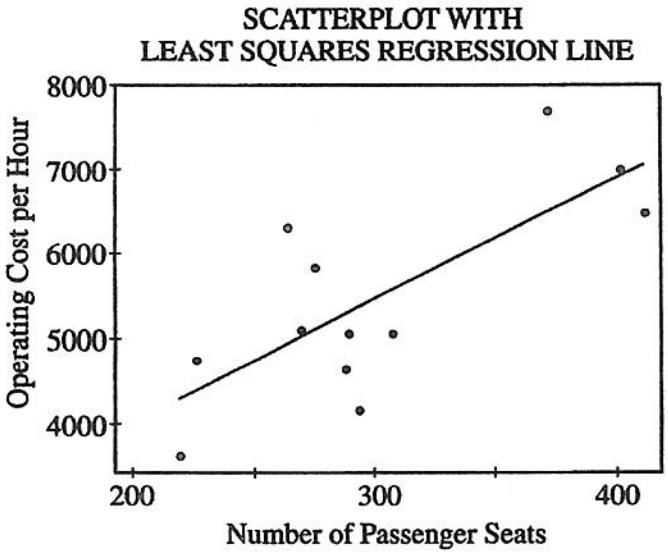


Mary turned in her paper late and was told by the professor that her grade would have been higher if she had turned it in on time. A computer printout fitting a straight line to the data (not including Mary's score) by the method of least squares gives

```
Grade = 46.51 + 1.106 Length
R-sq = 74.6%
```

(a) Find the correlation coefficient for the relationship between grade and length of paper based on these data (excluding Mary's paper).
(b) What is the slope of the regression line and what does it signify?
(c) How will the correlation coefficient change if Mary's paper is included? Explain your answer.
(d) How will the slope of the regression line change if Mary's paper is included? Explain your answer.
(e) What grade did Mary receive? Predict what she would have received if her paper had been on time.

3) Commercial airlines need to know the operating cost per hour of flight for each plane in their fleet. In a study of the relationship between operating cost per hour and number of passenger seats, investigators computed the regression of operating cost per hour on the number of passenger seats. The 12 sample aircraft used in the study included planes with as few as 216 passenger seats and planes with as many as 410 passenger seats. Operating cost per hour ranged between $3,600 and $7,800. Some computer output from a regression analysis of these data is shown below.



Predictor	Coef	StDev	T	P
Constant	1136	1226	0.93	0.376
Seats	14.673	4.027	3.64	0.005
S = 845.3	R-Sq = 57.0%		R-Sq (adj) = 52.7%	

(a) What is the equation of the least squares regression line that describes the relationship between operating cost per hour and number of passenger seats in the plane? Define any variables used in this equation.

(b) What is the value of the correlation coefficient for operating cost per hour and number of passenger seats in the plane? Interpret this correlation.

(c) Suppose that you want to describe the relationship between operating cost per hour and number of passenger seats in the plane for planes only in the range of 250 to 350 seats. Does the line shown in the scatterplot still provide the best description of the relationship for data in this range? Why or why not?