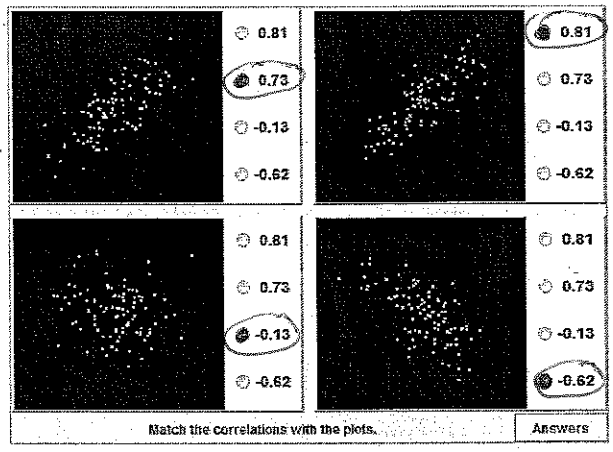
Statistics AP Practice – Chapter 3 Name KEY Date: _____

1. Which of the following have **no** unit of measure? Circle them.
   μ z s σ IQR $\bar{x}$ r s^2 median

2. Which of the following are **not** resistant to outliers or a few extreme values?
   μ z s σ IQR $\bar{x}$ r s^2 median

3. Match the correlations, -0.62, -0.13, 0.73, and 0.81 with the plots below.



 # 4 – 8 are true or false. If false, you must explain why and correct it.

4. If $r^2 = 0.49$, then the correlation r = 0.2401 False
   ±.7 depending on slope.

5. The least squares regression line of y on x makes the sum of the vertical distances of the data points from the line as small as possible. In symbols, the LSRL minimizes $\Sigma(y-\hat{y})$. False
   minimizes the $\Sigma(y-\hat{y})^2$
   (the squares of the error)

6. Extrapolation is the use of a regression line for prediction outside the range of values of the explanatory variable x used to obtain the line. Such predictions are often not accurate. True

7. A correlation, r, near +1 or near –1 indicates that there is a strong linear relationship. We can show that the explanatory variable causes the change in the response variable by obtaining an r near ±1. False
   NO! If the relationship is linear, then r measures the strength of it. AND – correlation does NOT IMPLY causation!!!

8. All density curves are symmetric and within one standard deviation of the mean there is 68% of the observations. False

only normally distributed density curves satisfy these properties

Questions # 9 - 15 are multiple-choice problems. Choose the most correct answer and circle it

D

9. Suppose the correlation between two variables is r = 0.12. What will be the new correlation if 0.25 is added to all the y variables, every value of the x variable is doubled, and the two variables are interchanged?
   a. 0.24 b. 0.49 c. 0.37 d. 0.12 e. –0.12

$r = \frac{1}{n-1}\sum\left(\frac{x-\bar{x}}{s_x}\right)\left(\frac{y-\bar{y}}{s_y}\right)$ r not affected by:
- change in units of x and y
- distinction between expl. & resp. variables

A

10. Data are obtained from a group of AP Statistics students by examining the number of letters in their full names and the scrabble value of their names. The resulting least squares regression line showed a correlation of 0.83. What percent of the variation in the scrabble value can be explained by looking at the linear relationship with the number of letters in the name?
    a. 69 % b. 83 % c. 31 % d. 91 % e. 17 %

$(.83)^2 = .6889$

E

11. Which one of the following statements is **true**?
    a. A lurking variable is one that has no effect on the outcomes of a study and whose influence was part of the investigation.
    b. The mean and standard deviation of the standard normal curve are $\bar{x} = 0$ and $s = 1$. $\mu = 0, \sigma = 1$
    c. The correlation coefficient and the slope of the LSR line sometimes have opposite signs.
    d. There is a very strong correlation between height and gender.
    e. The 10th percentile of the standard normal curve is $z \approx -1.28$.

C

12. In a statistics course a linear regression equation was computed to predict the final exam score from the score on the first test. The equation was $\hat{y} = 8 + 0.9x$ where y is the final exam score and x is the score on the first exam. James scored a 90 on the first test. On the final exam James scored an 86. What is the value of his residual?
    a. 4 b. –4 c. –3 d. 3 e. –1

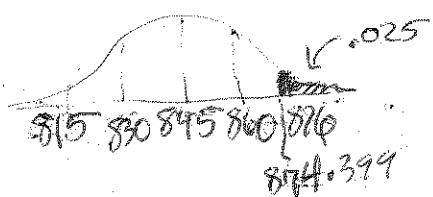
$\hat{y} = 8 + .9(90) = 89$

$res = 86 - 89 = -3$

C

13. If a store runs out of advertised material during a sale, customers become upset, and the store loses not only the sale but also goodwill. From past experience, a music store finds that a mean number of CD's sold in a sale is 845, the variance is 225, and a histogram of the demand is approximately Normal. The manager is willing to accept a 2.5% chance that a CD will be sold out. About how many CD's should the manager order for an upcoming sale?
    a. 1295 b. 1070 c. 875 d. 935 e. 860

$s = \sqrt{225} = 15$



14) Can we **predict the number of calories** in a fast-food hamburger from the amount of fat? Here are the fat and calorie contents of several brands of burgers. Find the regression line and the correlation for this data set.

A

Calories	410	580	570	640	680	660
Fat (g)	19	31	35	39	39	43

a. $\widehat{calories} = 210.33 + 11.06(fat)$, $r = 0.96$

b. $\widehat{calories} = 210.33 + 11.06(fat)$, $r^2 = 0.92$

c. $\widehat{fat} = -14.91 + 0.08(calories)$, $r = 0.96$

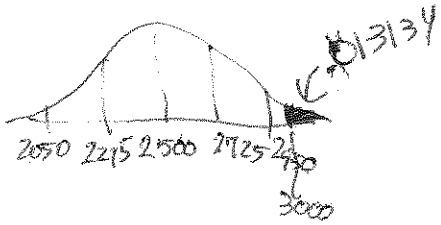
d. $\widehat{fat} = -14.91 + 0.08(calories)$, $r^2 = 0.92$

e. $\widehat{calories} = -14.91 + 0.08(fat)$, $r = 0.96$

15) The mean daily demand for bread at a popular bakery is 2500 loaves, with a standard deviation of 225 loaves. Every morning the bakery bakes 3000 loaves. What is the probability that today they will run out of bread? Assume that the mean daily demand for bread at this bakery is approximately normally distributed.

B

a. 0.0905 b. 0.0132 c. 0.1667 d. 0.8333 e. 0.9869



16) Fill in the missing information in the table below.

	$\bar{x}$	s_x	$\bar{y}$	s_y	r	$\hat{y} = a + bx$
a)	30	4	20	6	-0.2	$\hat{y} = 29 - .3x$
b)	$\frac{10}{3}$	0.8	40	15	.8	$\hat{y} = -10 + 15x$

$a = \bar{y} - b\bar{x}$

$b = r\frac{s_y}{s_x}$

a) $b = -.2\left(\frac{6}{4}\right) = -.2(1.5) = -.3$

$a = 20 + .3(30) = 29$

b) $15 = r\left(\frac{15}{.8}\right) \Rightarrow r = .8$

$-10 = 40 - 15(\bar{x}) \Rightarrow \bar{x} = \frac{10}{3}$