

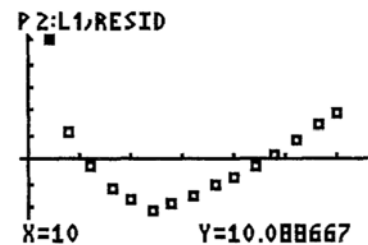
AP Stats
Chap 6-9
Practice Test

Name _____ Pd _____

SECTION I
Number of Questions – 29

Directions: Solve each of the following problems, using the available space (or extra paper) for scratchwork. Decide which is the best of the choices given and place that letter on the ScanTron sheet. No credit will be given for anything written on these pages for this part of the test. Do not spend too much time on any one problem.

1. 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 $\hat{y} = 10 + 0.9x$ where y is the final exam score and x is the score on the first test. Carla scored 95 on the first test. What is the predicted value of her score on the final exam?
A. 95 B. 85.5 C. 90 D. 95.5 E. none of these
2. Refer to the previous problem. On the final exam, Carla scored 98. What is the value of her residual?
A. 98 B. 2.5 C. -2.5 D. 0 E. none of these
3. A study of the fuel economy for various automobiles plotted the fuel consumption (in liters of gasoline used per 100 kilometers traveled) versus speed (in kilometers per hour). A Least Squares Regression Line (LSRL) was fit to the data. Here is the residual plot from this least squares fit. What does the pattern of the residuals tell you about the linear model?
A. The evidence is inconclusive.
B. The residual plot confirms the linearity of the fuel economy data.
C. The residual plot does not confirm the linearity of the data.
D. The residual plot clearly contradicts the linearity of the data.
E. None of the above.



4. All but one of the following statements contains a blunder. Which statement is correct?
- A. There is a correlation of 0.54 between the position a football player plays and his weight.
 - B. The correlation between planting rate and yield of corn was found to be $r = 0.23$.
 - C. The correlation between the gas mileage of a car and its weight is $r = 0.71$ mpg.
 - D. We found a high correlation ($r = 1.09$) between the height and age of children.
 - E. We found a correlation of $r = -0.63$ between gender and political party preference.
5. In the scatterplot in the previous question, if each x -value were decreased by one unit and the y -values remained the same, then the correlation r would
- A. decrease by 1 unit
 - B. decrease slightly
 - C. increase slightly
 - D. stay the same
 - E. you can't tell without knowing the data values
6. In regression, the residuals are which of the following?
- A. those factors unexplained by the data
 - B. the difference between the observed responses and the values predicted by the regression line
 - C. those data points which were recorded after the formal investigation was completed
 - D. possible models unexplored by the investigator
 - E. none of these
7. What does the square of the correlation measure?
- A. the slope of the LSRL
 - B. the intercept of the LSRL
 - C. the extent to which cause and effect is present in the data
 - D. the percent of the variation in the values of y that is explained by the variation in the values of x
 - E. the strength of the linear association between the variables
8. Which of the following statements are true?
- I. Correlation and regression require explanatory and response variables.
 - II. Scatterplots require that both variables be quantitative.
 - III. Every LSRL passes through $(\bar{x}, \bar{y})$
- A. I and II only
 - B. I and III only
 - C. II and III only
 - D. I, II, and III
 - E. none of these

9. A local community college announces the correlation between college entrance exam grades and scholastic achievement was found to be -1.08. On the basis of this, you would tell the college that
- A. the entrance exam is a good predictor of success.
 - B. the exam is a poor predictor of success.
 - C. students who do best on this exam will be poor students.
 - D. students at this school are underachieving.
 - E. the college should hire a new statistician.
10. A researcher finds that the correlation between the personality traits “greed” and “superciliousness” is -0.40. What percentage of the variation in greed can be explained by the relationship with superciliousness?
- A. 0% B. 16% C. 20% D. 40% E. 60%

11. Suppose the following information was collected, where x = diameter of tree trunk in inches, and y = tree height in feet.

x	4	2	8	6	10	6
y	8	4	18	22	30	8

- If the LSRL equation is $\hat{y} = -3.6 + 3.1x$, what is your estimate of the average height of all trees having a trunk diameter of 7 inches?
- A. 18.1 B. 19.1 C. 20.1 D. 21.1 E. 22.1

12. The following is / are resistant to drastic, unusual data:
- A. least squares regression line
 - B. correlation coefficient
 - C. both the least squares line and the correlation coefficient
 - D. neither the least squares line nor the correlation coefficient
 - E. it depends
13. A study found correlation $r = 0.61$ between the sex of a worker and his or her income. You conclude that
- A. women earn more than men on average.
 - B. women earn less than men on average.
 - C. an arithmetic mistake was made; this is not a possible value of r .
 - D. this is nonsense because r makes no sense here.
 - E. there is a moderately strong, positive, linear relationship between sex and income.

14. If **Data Set A** of (x, y) data has correlation coefficient $r = 0.65$, and a second **Data Set B** has correlation $r = -0.65$, then
- the points in **A** exhibit a stronger linear association than **B**.
 - the points in **B** exhibit a stronger linear association than **A**.
 - neither **A** nor **B** has a strong linear association.
 - you can't tell which data set has a stronger linear association without seeing the data or seeing the scatterplots.
 - A** and **B** are actually the same data set.

15. The effect of removing the right-most point (near the positive x -axis) in the scatterplot shown would be that

- the slope of the LSRL will increase; r will increase.
- the slope of the LSRL will increase; r will decrease.
- the slope of the LSRL will decrease; r will increase.
- the slope of the LSRL will decrease; r will decrease.
- there will be no change.



16. If removing an observation from a data set would have a marked change on the position of the LSRL fit to the data, what is the point called?

- | | | |
|----------------------|-----------------------|-------------------------|
| A. robust | B. a residual | |
| C. a response | D. influential | E. none of these |

17. Which of the following statements are correct?

- Two variables that are strongly associated will have a correlation near 1.
- Regression requires an explanatory-response relationship, while correlation does not.
- Even though the correlation between two variables may be high, in order to use the LSRL to predict, there needs to be an explanatory-response relationship between x and y .

- | | | |
|--------------------------|--|---------------------------|
| A. I and II only | B. I and III only | C. II and III only |
| D. I, II, and III | E. none of these gives the complete set of true responses | |

18. Suppose the correlation between two variables x and y is due to the fact that both are responding to changes in some unobserved third variable. What is this due to?

- | | |
|--|--|
| A. cause and effect between x and y | B. the effect of a lurking variable |
| C. extrapolation | D. common sense |
| E. none of these | |

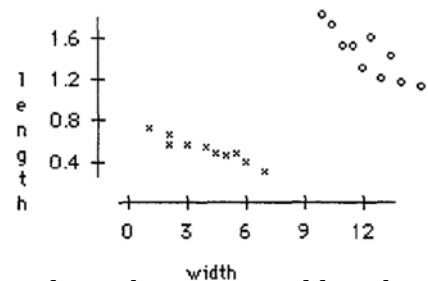
19. Suppose a straight line is fit to data having response variable y and explanatory variable x . Predicting values of y for values of x outside the range of the observed data is called

- | | |
|-------------------------|---------------------|
| A. correlation | B. causation |
| C. extrapolation | D. sampling |
| E. none of these | |

20. If the correlation between body weight and annual income were high and positive, we could conclude that

- A. high incomes cause people to eat more food.
- B. low incomes cause people to eat less food.
- C. high-income people tend to spend a greater proportion of their income on food than low-income people, on average.
- D. high-income people tend to be heavier than low-income people, on average.
- E. high incomes cause people to gain weight.

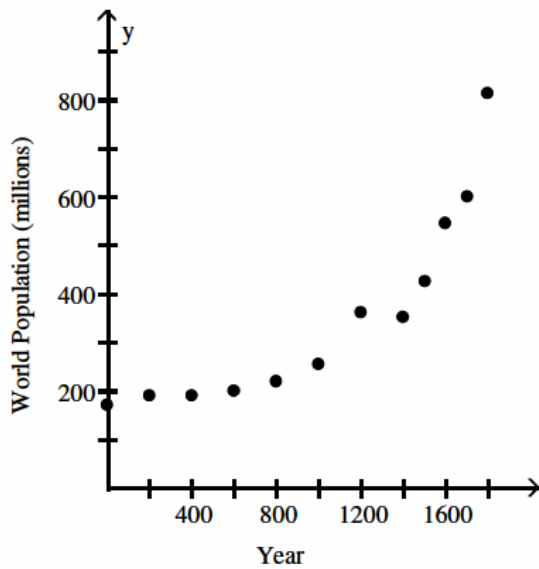
21. A study examined the relationship between the sepal length and sepal width for two varieties of an exotic tropical plant. **Variety A** and **Variety B** are represented by xs and os, respectively, in the following plot:



Which of the following statements is false?

- A. Considering **Variety A** alone, there is a negative correlation between sepal length and sepal width.
- B. Considering **Variety B** alone, the least squares regression line for predicting sepal length from sepal width has a negative slope.
- C. Considering both varieties together, there is a positive correlation between sepal length and sepal width.
- D. Considering each variety separately, there is a positive correlation between sepal length and sepal width.
- E. Considering both varieties together, the least squares regression line for predicting sepal length from sepal width has a positive slope.

22. The scatterplot below displays world populations (in millions) for years 0-1800. Where the population is an estimate, the lower estimate is given. For what range of years is a linear model appropriate?

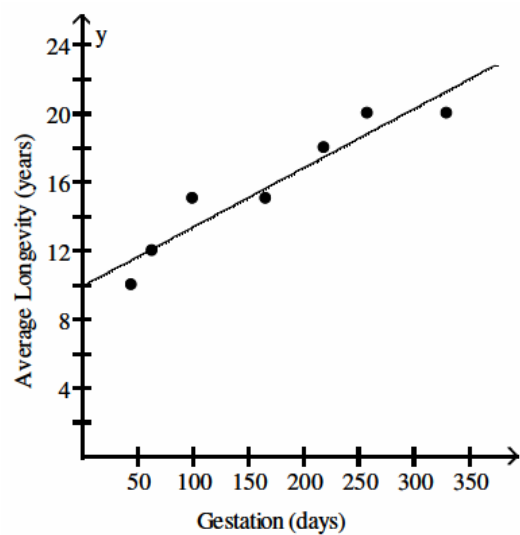


- A. A linear model should not be used for any part of the data.
- B. A linear model should be used for each pair of adjacent points.
- C. A single linear model is appropriate for the entire data set.
- D. One linear model is appropriate for the years 0 through 600 and another linear model for the years 600 through 1800.
- E. One linear model is appropriate for the years 0 through 1000 and another linear model for the years 1400 through 1800.

23. The table below shows the gestation (in days) and average longevity (in years) for a number of different mammals:

	<i>Gestation (days)</i>	<i>Average Longevity (years)</i>
Black Bear	219	18
Domestic Cat	63	12
Monkey	166	15
Lion	100	15
Horse	330	20
Gorilla	258	20
Gray Squirrel	44	10

The scatterplot and line of best fit are shown here:



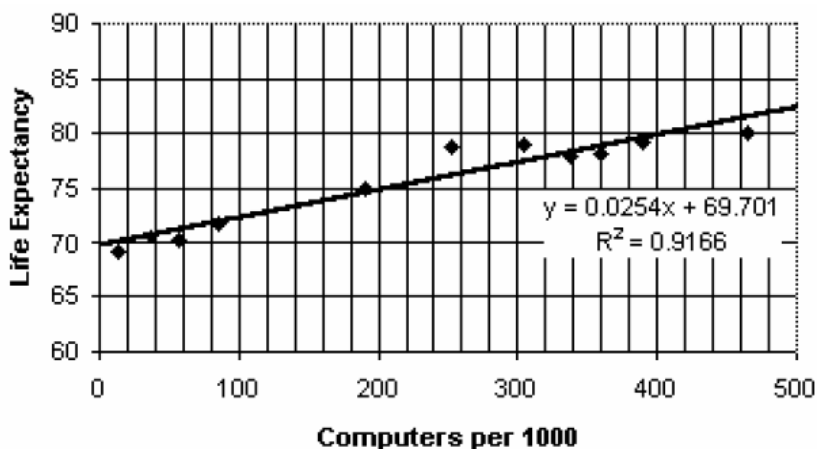
The regression analysis of this data yields the following values:

Variable	Coefficient
constant	9.90
gestation	0.0345

Use this information to predict the average longevity (to the nearest tenth of a year) of an African elephant whose gestation is 660 days.

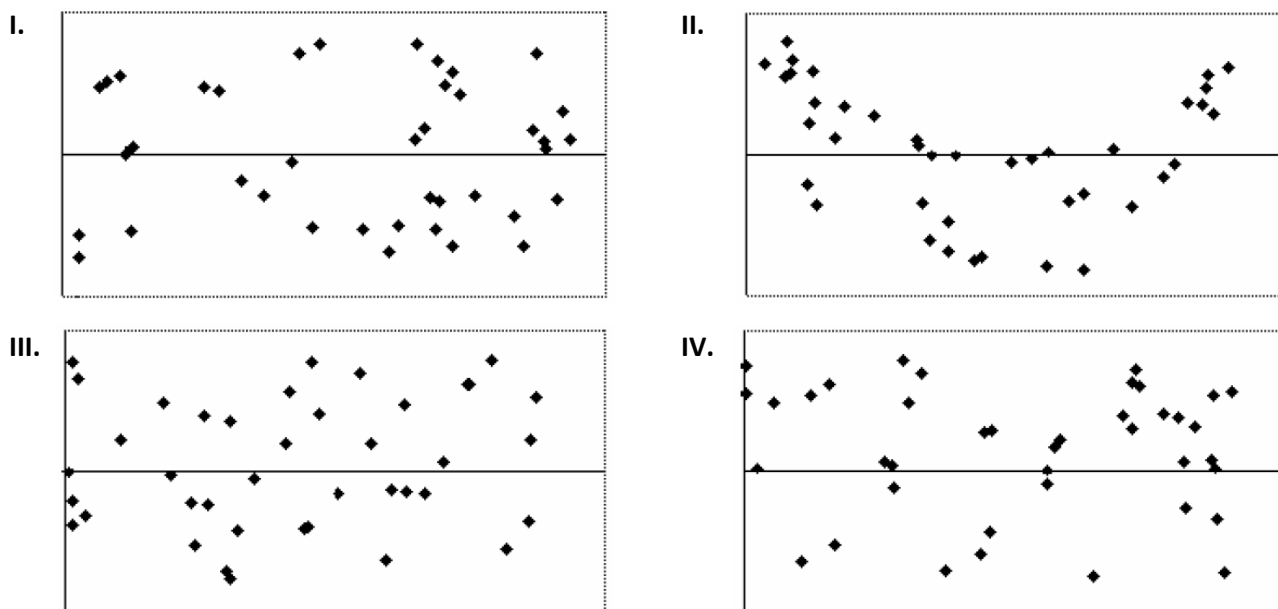
- A. 22.8 years
- B. 38.7 years
- C. 41.9 years
- D. 32.7 years
- E. 51.2 years

24. The figure below examines the association between life expectancy and computer ownership for several countries. Also shown are the equation and R^2 value from a linear regression analysis of the data. What is the best conclusion to draw from the figure?



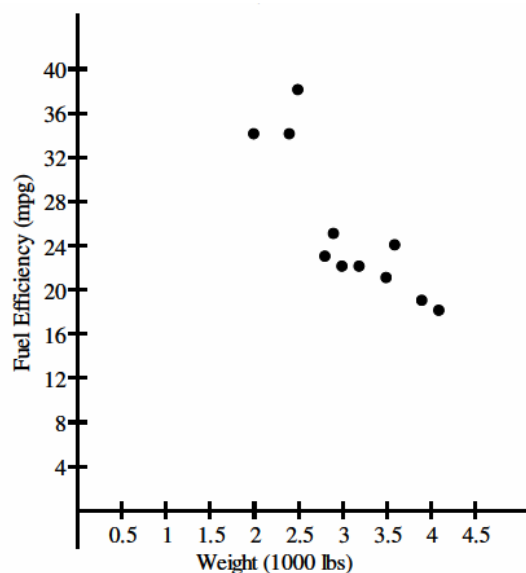
- A. Although the association is strong, computer ownership probably does not promote longevity. Instead, national per capital wealth is probably a lurking variable that drives both life expectancy and computer ownership.
- B. Clearly, there must be some as-yet unknown health benefit associated with using computers.
- C. Persons who live longer buy more computers over the course of their longer lifetimes.
- D. Exposure to the radiation from computer monitors is causing a clear decline in life expectancy.
- E. Computer ownership promotes health and long life, probably due to the greater access that computer owners have to health information on the Internet.

25. Which of the following scatterplots of residuals suggests that a linear model may not be applicable?



- A. II
- B. I
- C. III
- D. IV
- E. none of these

26. One of the important factors determining a car’s fuel efficiency is its weight. This relationship is examined for 11 cars, and the association is shown in the scatterplot below.



Dependent variable: Fuel Efficiency
R-squared = 84.7%

VARIABLE	COEFFICIENT
Intercept	47.1181
Weight	-7.34614

If a linear model is considered, the regression analysis is:

What does the slope say about this relationship?

- A. Gas mileage decreases an average of 0.7346 mpg for each thousand pounds of weight.
- B. Gas mileage decreases an average of 4.712 mpg for each thousand pounds of weight.
- C. Gas mileage decreases an average of 7.346 mpg for each thousand pounds of weight.
- D. Gas mileage increases an average of 4.712 mpg for each thousand pounds of weight.
- E. Gas mileage increases an average of 7.346 mpg for each thousand pounds of weight.

27. Consider the data listed in the following table:

X	Y
2	0.0714
4	0.0333
6	0.0217
8	0.0161
10	0.0128
12	0.0106
14	0.0091

Create an appropriate model. What re-expression of Y does this model involve?

- A. y^2
- B. $\log y$
- C. y^1
- D. $-\sqrt{\frac{1}{y}}$
- E. $\sqrt{y}$

28. Consider the data listed in the following table:

X	Y
2	-0.0417
4	-0.02
6	-0.0132
8	-0.0098
10	-0.0078
12	-0.0065
14	-0.0056

Create an appropriate model. Estimate the value of Y when X = 33. Round your answer to four decimal places.

- A. -0.0024 B. 3.1427 C. 4.2377 D. 2.1677 E. 0.1207

29. Dioxins are a class of long-lived and highly toxic pollutants. The topsoil concentration in parts per million (ppm) are shown in the table as a function of distance from the dump where they were collected.

Distance from dump (meters)	Dioxin concentration (ppm)
21	435
42	235
49	150
79	65
102	45
157	22
204	14

Re-express the dioxin concentration levels. Then determine the regression equation and coefficient of determination for the re-expressed data.

- A. $(\text{dioxin})^2 = 17.98 + 0.0829(\text{distance})$; $R^2 = 0.8830$
- B. $\log(\text{dioxin}) = -0.0079 + 2.619(\text{distance})$; $R^2 = 0.9308$
- C. $\sqrt{\text{dioxin}} = 1153.58 - 663.91(\text{distance})$; $R^2 = 0.4061$
- D. $-\frac{1}{\sqrt{\text{dioxin}}} = -0.0218 - 0.0012(\text{distance})$; $R^2 = 0.9971$
- E. $-\frac{1}{\text{dioxin}} = 0.3624 - 0.00497(\text{distance})$; $R^2 = 0.9971$

SECTION II

Questions 30-38

Directions: Show all of your work. Indicate clearly the methods you use, because you will be graded on the correctness of your methods as well as on the accuracy of your results and explanations.

Assembly Line.

Your new job at Panasonic is to do the final assembly of camcorders. As you learn how, you get faster. The company tells you that you will qualify for a raise if, after 13 weeks, your assembly time averages under 20 minutes. The data shows your average assembly time during each of your first 10 weeks.

Week	Time (min)
1	43
2	39
3	35
4	33
5	32
6	30
7	30
8	28
9	26
10	25

30. Which is the explanatory variable? 30. _____
31. What is the correlation between these variables? 31. _____
32. You want to predict whether or not you will qualify for that raise. Would it be appropriate to use a linear model? Defend your position.

Earning Power.

A college's job placement office collected data about students' GPAs and the salaries they earned in their first jobs after graduation.

GPA	3.4	2.6	3.1	2.4	3.9	2.6	3.8
Salary	51840	48900	50861	39675	47260	42590	62950

33. Write an equation of the model that can predict salary based on GPA. 33. _____
34. Do you think these predictions will be reliable? Defend your answer.
35. Your brother just graduated from that college with a GPA of 3.30. He tells you that based on this model, the residual for his pay is -\$1,880. What salary is he earning? 35. _____

Music and Grades.

A couple of years ago a local newspaper published research results claiming a positive association between the number of years high school children had taken instrumental music lessons and their performances in school as measured by their GPA.

36. A group of parents then went to the school board demanding more funding for music programs as a way to improve student chances for academic success in high school. As a statistician, do you agree or disagree with their reasoning? Briefly explain.

Consumer Price Index.

The consumer price index (CPI) is a measure of the relative cost of goods in the US for a particular year. The table here shows the CPI for various years beginning in 1940.

Year	CPI
1940	14.0
1950	24.1
1960	29.6
1970	38.8
1980	82.4
1990	130.7
2000	201.4

37a. Use an appropriate re-expression of the data to develop a model that can be used for predicting CPI values.

37a. _____

37b. Use your model to predict the CPI for 2010.

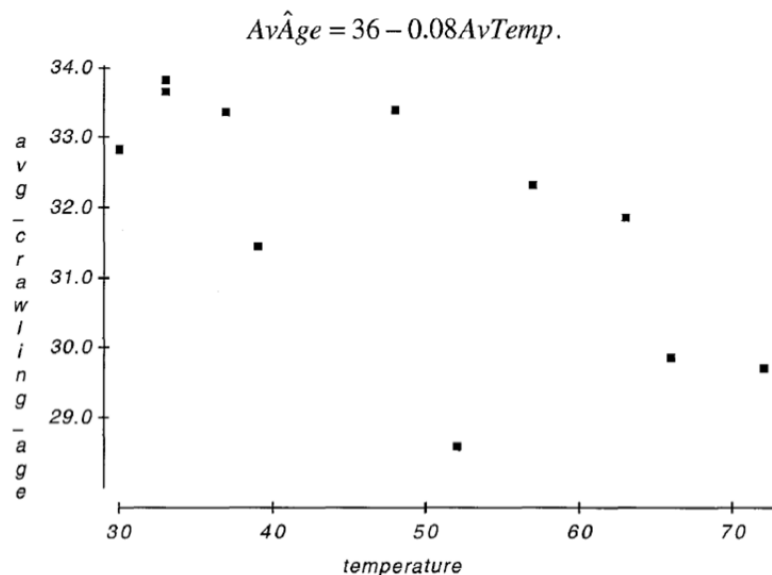
37b. _____

37c. State which rung of the Ladder of Powers you used for your re-expression and explain why you chose this rung.

37c. _____

Crawling.

Researchers at the University of Denver Infant Study Center investigated whether babies take longer to learn to crawl in cold months (when they are often bundled in clothes that restrict their movement) than in warmer months. The study sought an association between babies' first crawling age (in weeks) and the average temperature during the month they first try to crawl (about 6 months after birth). Between 1988 and 1991 parents reported the birth month and age at which their child was first able to creep or crawl a distance of four feet in one minute. Data were collected on 208 boys and 206 girls. The graph below plots average crawling ages (in weeks) against the mean temperatures when the babies were 6 months old. The researchers found a correlation of $r = -0.70$ and their scatterplot was



- 38a. Draw the line of best fit on the graph.
- 38b. Describe the association in context.
- 38c. Explain, in context, what the slope of the line means.
- 38d. Explain, in context, what the y-intercept of the line means.
- 38e. Explain, in context, what the R^2 means.
- 38f. In this context, what does a negative residual indicate?