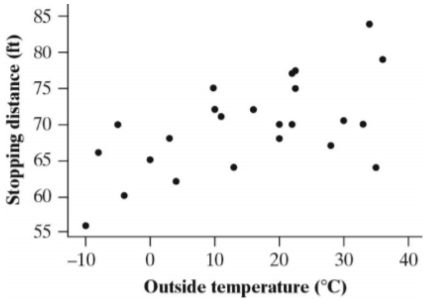
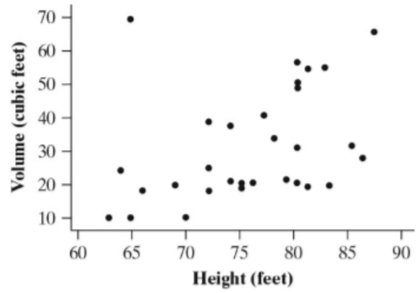
NAME____Period_____________Date_______________

C H A P T E R   3   P R A C T I C E   T E S T :   Describing Relationships



Graph 1



Graph 2

```
Predictor      Coef   SE Coef     T       P
Constant     10.658     7.636  1.40   0.193
Height      0.08237   0.04651  1.77   0.107

S = 1.223   R-Sq = 23.9%   R-Sq(adj) = 16.3%
```

Graph 3

1. A researcher wishes to determine whether the rate of water flow (in liters per second) over an experimental soil bed can be used to predict the amount of soil washed away (in kilograms). In this study, the response and explanatory variables are

2. Consider the scatterplot in graph 1 which describes the relationship between stopping distance (in feet) and air temperature (in degrees Celsius) for a certain 2,000-pound car traveling 40 mph. Estimate *r*, the correlation coefficient

3. The height (in feet) and volume (in cubic feet) of usable lumber of 32 cherry trees are measured by a researcher is shown in graph 2 above. The goal is to determine if volume of usable lumber can be estimated from the height of a tree. If the data point (65, 70) were removed from this study, how would the value of the correlation r change?

4. A researcher wishes to study how the average weight y (in kilograms) of children changes during the first year of life. He plots these averages versus the age x (in months) and decides to fit a least-squares regression line to the data with x as the explanatory variable and y as the response variable. He computes the following quantities.
r = correlation between x and y = 0.8
$\bar{x}$ = mean of the values of x = 6.2, $\bar{y}$ = mean of the values of y = 6.4
s_x = standard deviation of the values of x = 3.6, s_y = standard deviation of the values of y = 1.2
What is the equation of the least-squares regression line?

5. The manager of a water park collects data on the daily high temperature and the number of customers entering the park for 15 summer weekdays. Based on these data he produces a least squares regression equation to predict y = the number of customers based on x = daily high temperature. The equation is $\hat{y} = -3110 + 51.2x$. On one day, the high temperature was 93¤F and 1700 customers entered the park. Which of the following is the residual for this observation?

6. The correlation between number of service calls and the number of copy machines leased is 0.86. About what percent of the variability in the number of service calls is explained by the linear relation between number of service calls and number of machines?

7. Can you predict a person's shoe size from their height? The computer output in graph 3 (above) provides information on a regression analysis of these variables for 12 randomly selected high school students in England. Both height and foot length are measured in centimeters. What is the correct interpretation of s = 1.223?

8. A researcher finds that the correlation between the personality traits "greed" and "superciliousness" is –.40. What percentage of the variation in greed can be explained by the relationship with superciliousness?

9. 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 y = –3.6 + 3.1x, what is your estimate of the average height of all trees having a trunk diameter of 7 inches?

NAME____Period_____________Date_______________

10. An LSRL predicts the height of a tree in its first 10 weeks outside the nursery to be $\hat{y} = 3.2 + 1.8x$ where $\hat{y}$ is the predicted height in inches and x is the number weeks outside the nursery.

(a) Interpret the meanings of the *y*-intercept (*a*) and the slope (*b*) in terms of the data provided.
(b) Estimate the height of a tree that has been out of the nursery 6 weeks.
(c) The actual height of a tree that has been out of the nursery for 6 weeks was found to be 12.5 inches. Calculate the residual for this tree.

11. An LSRL between train departure time after 5pm and travel time found an r^2 of 0.66. If the relationship has negative correlation:
(a) What is the value of *r*?
(b) Explain the value of r^2 in context.
(c) What quantity is minimized in this LSRL?

12. The paper "Feeding of Predaceous Fishes on Out-Migrating Juvenile Salmonids in John Day Reservoir, Columbia River" gave the following data. The paper plotted the maximum size of salmonids consumed by northern squawfish against the length of the squawfish, both in millimeters. A linear regression yeilded the following computer output and a Residual Plot with no clear pattern.

Predictor	Coef	SE Coef	T	P
Constant	89.09	16.83	5.29	0.00
Squawfish length	0.729	0.047	15.26	0.00
S = 12.56	R-Sq = 96.3%	R-Sq (adj) = 95.9%		

(a) What is the value of the slope of the least squares regression line? Interpret the slope in the context of this situation.
(b) What is the value of the y-intercept of the least squares regression line? Interpret the y-intercept in the context of this situation.
(c) What maximum salmonid size would you predict for a squawfish whose length is 375 mm?
(d) A 414 mm squawfish consumed a salmonid 219 mm in length. Find the value of the residual for this squawfish.

13. A certain psychologist counsels people who are getting divorced. A random sample of six of her patients provided the following data where
x = number of years of courtship before marriage, and
y = number of years of marriage before divorce.

x	3	0.5	2	1.5	5
y	9	6	14	10	20

(a) Construct a scatterplot of these points:
(b) Use your calculator to determine the least-squares regression line (LSRL). Write the equation, and plot this line on your graph. (Be sure to show what information you're using to plot the line.)
(c) Find the values of the slope and the y-intercept for the regression line in (b). Interpret these values in context.
(d) What is the correlation between x and y? Interpret this number.
(e) What is the value of r^2? Interpret this number.
(f) Show how the residual for the first data point in the table is calculated. Then construct a residual plot on the second axis. Does your residual plot confirm or refute the choice of a linear regression? Explain.

NAME____Period_____________Date_______________

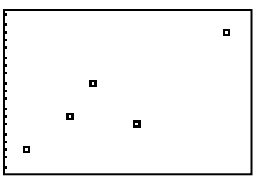
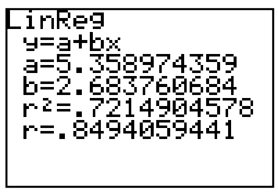
^)ANSWERS(T_T)ANSWERS(^_^)ANSWERS(T_T)ANSWERS(^_^)ANSWERS(T_T)ANSWERS(^_^)ANSWERS(T_T

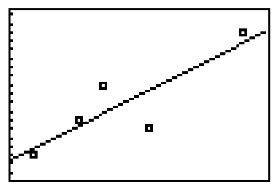
1. Explanatory: rate of water flow, Response: amount of soil washed away
2. About 0.6   3. *r* would increase since the point does not follow the pattern of the rest of the data
4. (predicted weight) = 4.75 + 0.27(age)   5. 48.4   6. About 74%
7. The typical error in predicting foot length from height based on this regression model is about 1.22 cm
8. trick question: r does not make sense here.   9. 18.1

10(a) The height of a tree that has just left the nursery is estimated to be 3.2 inches. For every additional week outside the nursery, a tree grows 1.8 inches on average.   (b) A tree that has been out for 6 weeks will be approximately 14 inches tall.   (c) Residual = actual - predicted; $\text{residual} = 12.5 - 14 = -1.5$.

11 (a) $r = \sqrt{.66} = 0.8124$ (note: the value of r is positive because the relationship has positive association.
(b) 66% of the variation in travel time can be explained by and LSRL relating travel time with departure time.
(c) The LSRL minimizes the sum of the squares of the residuals.

12. (a) b = 0.729 mm per mm; For every increase of 1 mm in squawfish length, the maximum salmonid consumed by that squawfish increases by approximately 0.729mm (b) a = 89.09 mm; the largest salmonid consumed by a squawfish that is 0 mm in length is predicted to be 89.09 mm. This is a nonsensical measurement since a squawfish that is 0 mm in length is non-existent. (c) (predicted maximum salmonid length) = 89.09 + 0.729 x (375) = 362.7 mm; a squawfish that is 375 mm in length is predicted to eat a maximum salmonid length of about 363 mm. (d) resid = actual - predicted = 219 - (89.09 + .729 x 414) = 219 - 390.9 = -171.9 mm

13. (a)  (b) 

 (e) 

(b) (predicted years before divorce) = 5.36 + 2.68 (years of courtship before marriage) (c) Slope is 2.68 years before divorce per years of courtship before marriage. For every increase of 1 year of courtship before marriage, the number of years before divorce increases by 2.68 years. Y-intercept is 5.36 years. A couple with 0 years of courtship is predicted to divorce in 5.36 years. (d) r = 0.849; there is a strong positive association between years of courtship before marriage and predicted years before divorce. (e) $r^2 = 0.7208$; 72% of the variation in number of years before divorce can be explained with an LSRL relating number of years before divorce with number of years of courtship (f) residual = Actual value – Predicted value = (9) – (5.36 + 2.68 x 3) = -4.4. Since the residual plot shows no structured pattern, our choice of a linear pattern is confirmed.