

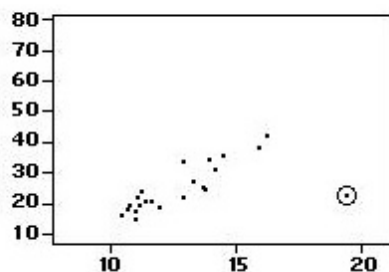
1. Which of the following is true of the slope of the least-squares regression line?

- A. It has the same sign as the correlation.
  - B. The square of the slope equals the fraction of the variation in the response variable that is explained by the explanatory variable.
  - C. It is unitless.
- 

2. Foresters use regression to predict the volume of timber in a tree using easily measured quantities such as diameter. Let  $y$  be the volume of timber in cubic feet and  $x$  be the diameter in feet (measured at 3 feet above ground level). One set of data gives  $\hat{y} = -30 + 60x$ . The predicted volume for a tree of diameter=18 inches is

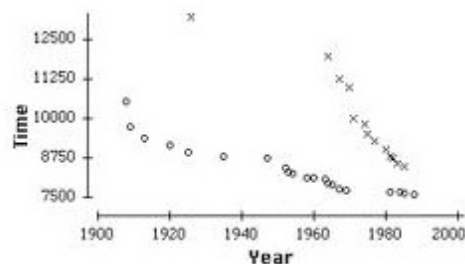
- A. 1050 cubic feet.
  - B. 90 cubic feet.
  - C. 60 cubic feet.
- 

3. In the graph below the circled point is



- A. an influential point. Deleting it should reduce the correlation and improve the fit.
- B. not an influential point because its  $y$  value is not unusually large or small.
- C. an influential point. Deleting it should increase the estimate of the slope.

4. The world record time (in seconds) in the marathon is plotted versus the year in which the record was set for men and women separately below. The plotting symbol o is used for men and x for women. The data includes only records set between 1908 and 1988.



Which of the following statements would be a valid conclusion from this plot?

- A. The world record times for women show a greater rate of improvement (decrease more rapidly) than the world record times for men for the period represented in the plot.
- B. We can expect the world record time for women to be lower than that for men sometime before the year 2010.
- C. By the year 2010, the world record time for men will reach a plateau beyond which no improvements will be possible.

---

5. A researcher wants to study the effect of regular exercise on cholesterol levels. The researcher compares the cholesterol level of fifty people who belong to a local gym and exercise regularly with the cholesterol level of fifty people from the community who do not exercise regularly. The cholesterol levels of the members of the gym were substantially lower. The researcher can conclude that

- A. belonging to a gym reduces cholesterol level.
- B. exercising regularly reduces cholesterol level.
- C. members of a local gym who exercise regularly have lower cholesterol levels than those in the community who do not exercise regularly.

---

6. Which of the following measures only the strength of a relationship?

- A. The correlation coefficient,  $r$ .
- B. The slope,  $b$ .
- C. The coefficient of determination,  $r^2$ .

7. A student has obtained the following computer output from a regression examining the relationship between BTUs input to a furnace and the BTU output.

Predictor	Coef	SE Coef	T	P
Constant	0.1868	0.2153	0.87	0.388
BTU.In	0.91109	0.01915	47.57	0.000

S = 0.5580      R-Sq = 96.3%

From this we conclude

- A. For every BTU input to a furnace, the output increases by 0.911, on average.
- B. For every 0.187 BTU input to a furnace, the output increases by 0.911, on average.
- C. For every 0.911 BTU input to a furnace, the output increases by 1, on average.

---

8. Statistician William Hammack examined the relationship between the number of public schools in each county of Florida and the crime rate for the county. The data showed a very strong linear relationship with  $r = 0.970$ . From this we conclude

- A. Public schools cause crime. We need more private schools.
- B. There is probably a lurking variable at work.
- C. This is totally meaningless.

---

9. Suppose you want to know how salaries (in thousands of dollars per year) are related to years of experience for your chosen career. You collect data for a random sample of 100 people with this type of job who have had from 0 to 20 years of experience. The relationship is linear and you determine the regression equation to be:  $\text{salary} = 45.59 + .798 \text{ years}$ . Predict the salary for an individual who has been working 15 years.

- A. \$57.56
- B. \$57,560
- C. \$45,590

10. For a sample of students, their total SAT scores and college grade point averages (GPA) are:

SAT	GPA
1070	3.50
1040	2.66
1120	3.54
1260	3.83
1100	2.60
1020	2.70
960	2.00

The least squares regression line equation for this data is

- A.  $\hat{y} = -3.65 + 0.008x$
- B.  $\hat{y} = 729.46 + 118.62x$
- C.  $\hat{y} = -3.27 + 0.006x$