Study Guide: Linear regression

1. A baseball coach graphs some data and finds the line of best fit. The equation for the line of best fit is $y = 0.32x - 20.51$, where x is the number of times at bat and y is the number of hits.

How many hits should he expect from a player who is at bat 175 times?

A) 35 hits ⟵ *(circled)*
B) 49 hits
C) 609 hits
D) 62 hits

$y = .32(175) - 20.51$

---

3. Chang wants to know if he is improving his skill on the cello. He created a scatter plot and drew a line of best fit.

(2, 8)

(5, 1.5)

Number of Mistakes Played / Time Practicing (hours)
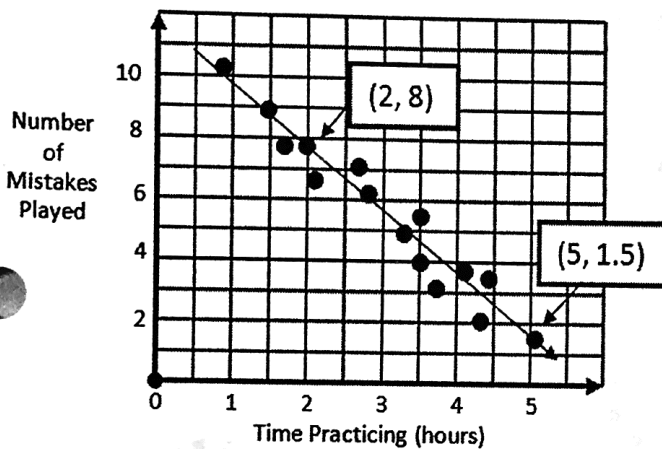
If he uses the point (2, 8) and (5, 1.5) from his line, which equation would best represent the line of best fit?

A) y = -2.17x + 12.3 *(circled)*
B) y = 2.17x + 3.77 ✗
C) y = -0.46x + 9 ✗
D) y = -2.17x - 9.35

$\frac{8 - 1.5}{2 - 5} = \frac{6.5}{-3} = -2.17$

$b = 12.3$

---

5. Jared collected data on the ages and heights of a random sample of elementary school students. If he plots the data on a scatter plot, what relationship will he most likely see between age and height?

A) A negative correlation
B) No correlation
C) A positive correlation *(circled)*
D) A constant correlation

---

2. Determine the correlation (positive, negative, or no correlation) for each statement. Explain your reasoning for each.

A) The number of people at your party and the number of sodas you have in the refrigerator.

Negative because the more people at the party, the less sodas you should have in the refrigerator.

B) The amount your drive to school and the amount you have to pay for gas.

Positive the more you drive to school, the more gas you use so the more you pay for gas. gas.

C) The size your dog and the number of times you take it to the vet.

None   or   Positive

your dog's size does not affect dog visits | If your dog is larger (obese) you might have to go to the vet more.

---

4. The graph below shows the relationship between the distance in miles a delivery truck traveled and the number of hours each delivery took.
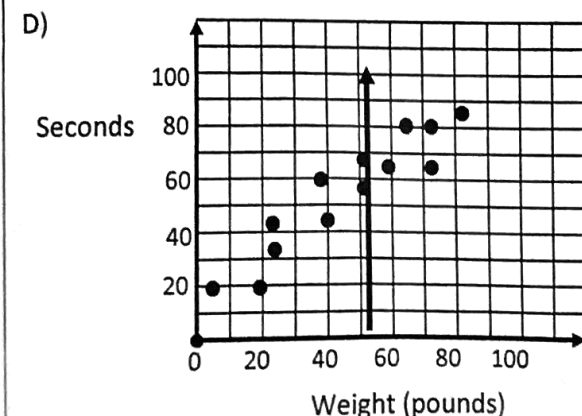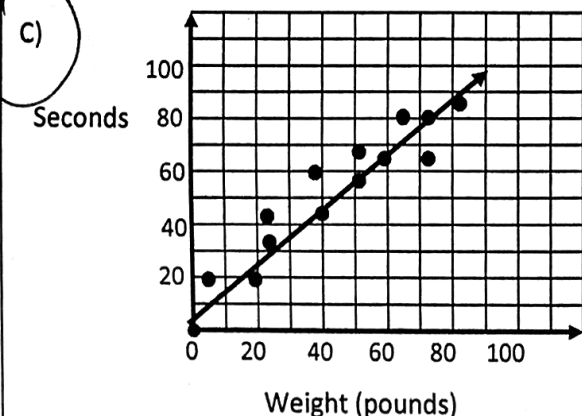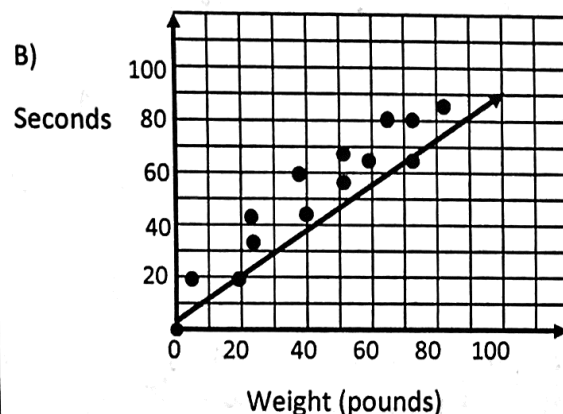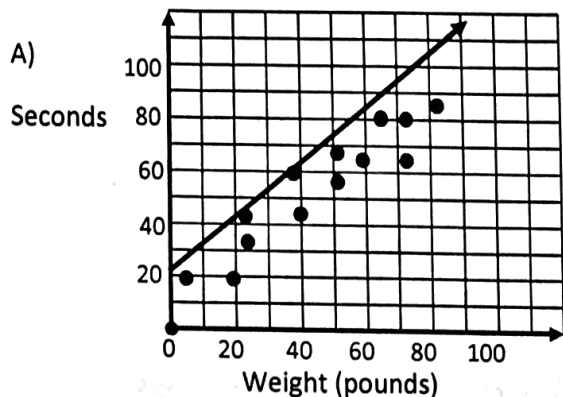

Number of hours / Distance (miles)
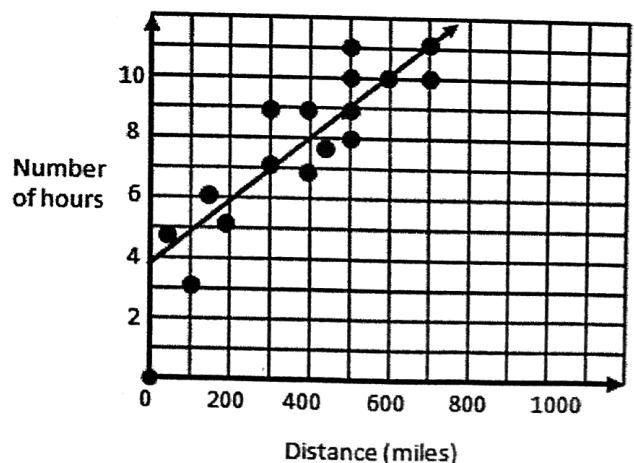
Which of the two given points would be the best to use to calculate the line of best fit?

A) (500, 11) and (700, 11) ✗
B) (300, 9) and (400, 7)
C) (400, 9) and (500, 11) ✗
D) (300, 7) and (600, 10) *(circled)*

6. The graph shows the weights of dogs and the time it took the same dogs to complete an agility course in seconds.

Which shows the line of best fit for the data?

A)

Seconds (y-axis), Weight (pounds) (x-axis)

B)

Seconds (y-axis), Weight (pounds) (x-axis)

C) ⭕

Seconds (y-axis), Weight (pounds) (x-axis)

D)

Seconds (y-axis), Weight (pounds) (x-axis)

7. A high school principal wants to predict the number of students will drop out of school so he can get funding for support services.

D) Year	The number of students who drop out of high school
2004	217
2005	202
2006	199
2007	185
2008	180
2009	163

He determined the equation that represents this data as $y = -10x + 216$, where x represents the years since 2004, and y is the number of students who drop out.

Use this equation to help him predict the number of students who will drop out in 2012?

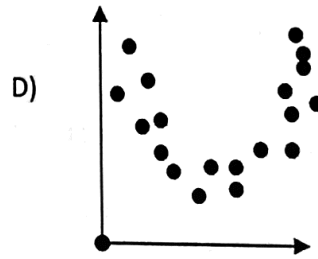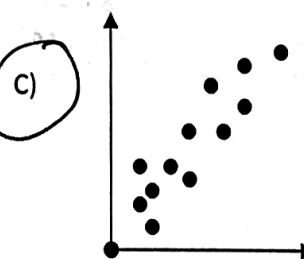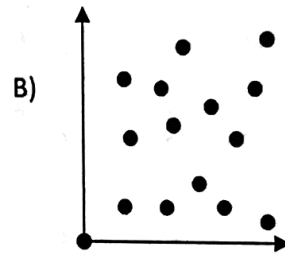A) -20, 2996 students
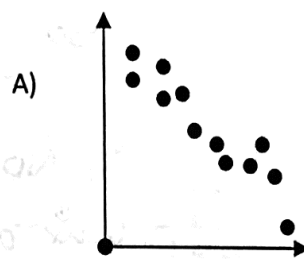B) 21 students
C) 136 students ⭕
D) 156 students

$y = -10(8) + 216$

$y = -80 + 216$

$y = 136$

8. Which graph represents a positive correlation?

A)

B)

C) ⭕

D)

9. The scatter plot below shows the average yearly consumption of bottled water by people in the United States starting in 1990.

Gallons of Bottled Water per Person (y-axis), Year (x-axis)
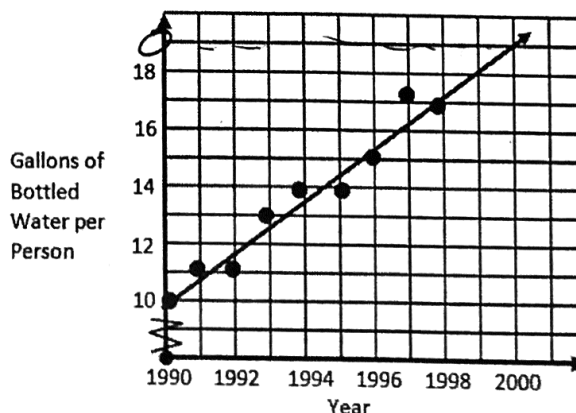
Using the line of best fit, predict the average consumption of bottled water in the year 2000?
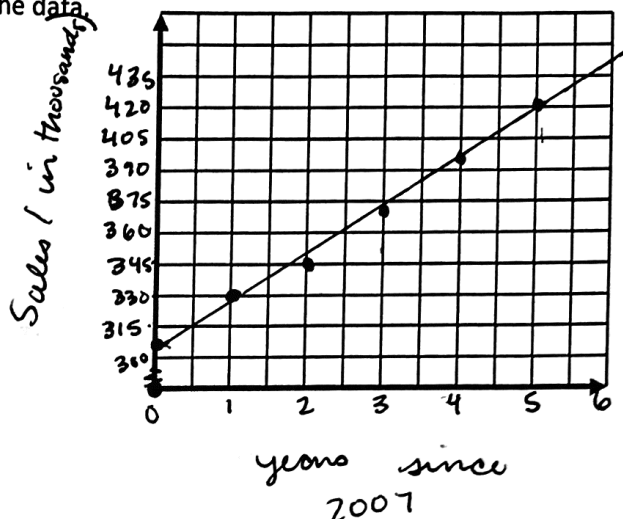
A) 20 gallons   B) 18 gallons   C) 20 gallons   D) 19 gallons ⭕

...able below shows the sales for a flower company for the years 2007 through 2012. Answer the given questions about
...e on your answer sheet.

Graph the data on the scatter plot and draw a line of best fit for the data.

FLOWER SALES

Year	Sales (in thousands)
2007	$305
2008	$330
2009	$345
2010	$370
2011	$395
2012	$420

B) Write an equation for the line of best fit for this data. Let x represent the years since 2007 and y represent the sales, in thousands of dollars.

$(0, 305) \quad (5, 420)$

$\dfrac{115}{5} = 23$

$y = 23x + 305$

C) According to your equation, in what year will the sales reach about $500 (in thousands)? Use mathematics to explain how you determined your answer.
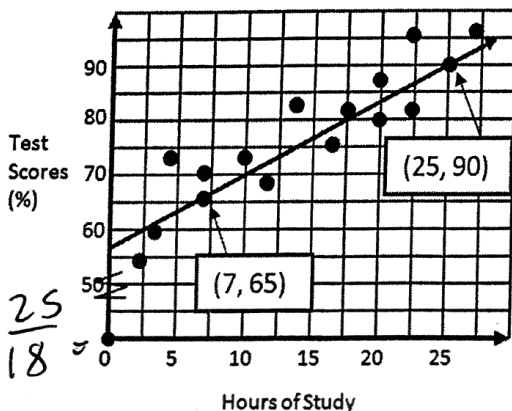
$500 = 23x + 305$
$-305 \qquad\qquad -305$

$195 = 23x$

$8.48 = x$

$\boxed{2016}$

---

11. Mr. Van made a graph to represent the time his students spent studying for their test and their actual test score.
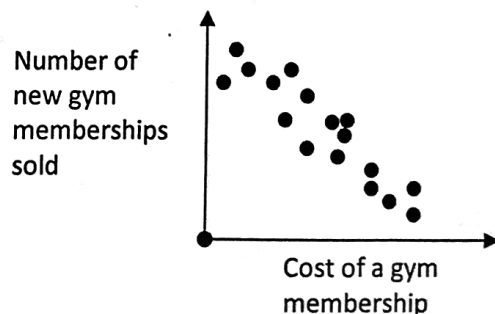Which is the correct equation for the line of best fit?

A) y = 1.4x + 55
B) y = 1.4x − 84
C) y = 0.72x + 60
D) y = 0.72 + 56

Test Scores (%)
Hours of Study

(25, 90)
(7, 65)

$\dfrac{90 - 65}{25 - 7} = \dfrac{25}{18}$

1.4

12. Which relationship is shown by this scatter plot?

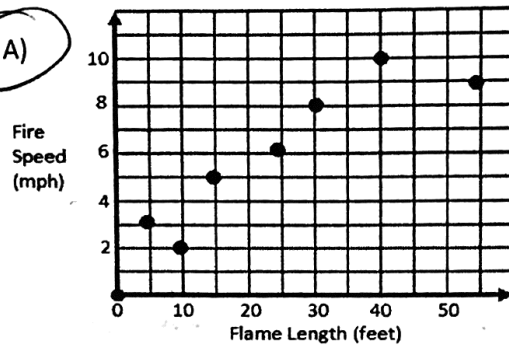Number of new gym memberships sold

Cost of a gym membership

A) As the cost goes down, the number sold goes down.
B) As the cost goes up, the number sold goes down.
C) As the cost does down, the number sold remains the same.
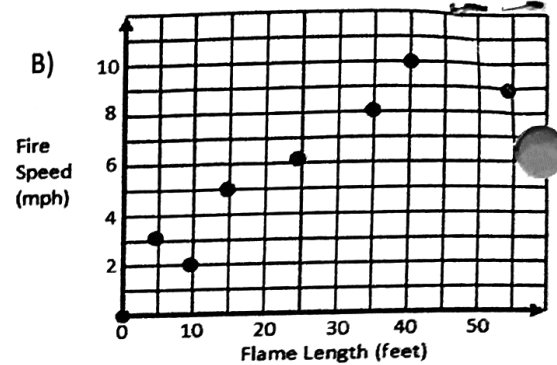D) There is no relationship between cost and the number sold.

13. Which scatter plot best represents the data given in the table?

Flame Length	Fire Speed (mph)
10	2
40	10
15	5
5	3
55	9
30	8
25	6

A) (circled)

Fire Speed (mph) vs Flame Length (feet)

B)

Fire Speed (mph) vs Flame Length (feet)

C)

Fire Speed (mph) vs Flame Length (feet)
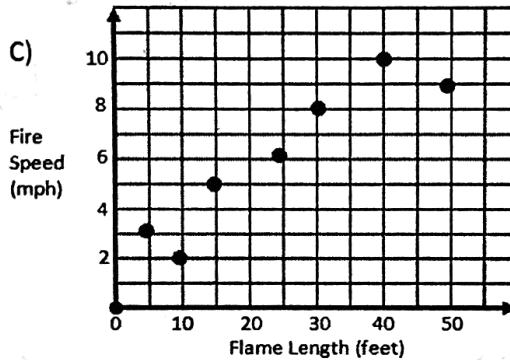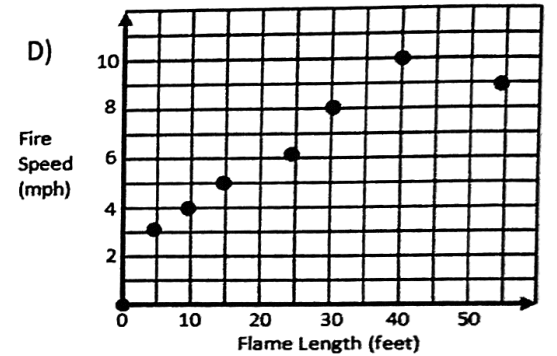
D)

Fire Speed (mph) vs Flame Length (feet)

---

14. Given in the table and scatter plot are the samplings of average annual temperatures collected at different elevations in the United States. Pick two points from the line of best fit and determine the equation for the line of best fit.

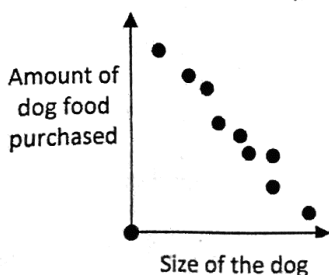| Elevation (meters)	Average Annual Temp (°C)
1,000	15
490	20
2,200	8
650	17
1,750	8
1,500	12.5
800	18
1,400	15
2,000	10
1,200	12

Average Annual Temp (C°) vs Elevation (m)
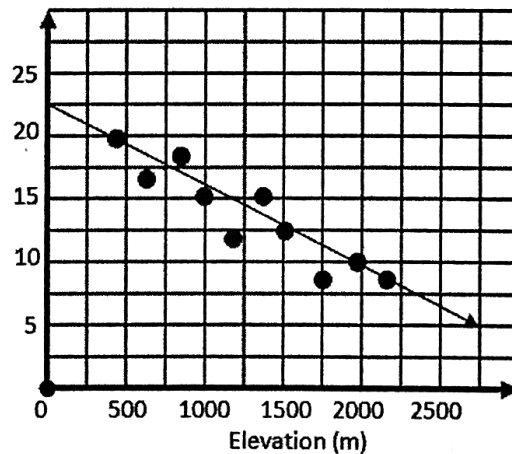
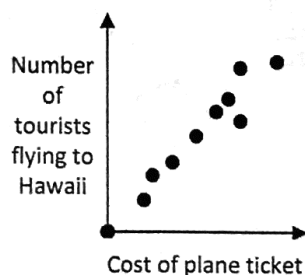$(0, 22.5)\quad (750, 18$

$$\frac{4.5}{-750} = -0.006$$

$$y = -0.006x + 22.5$$

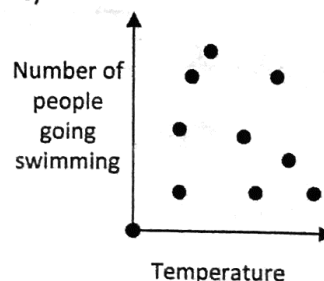**15.** Which graph represents the correlation of its given situation correctly?

A) Amount of dog food purchased vs Size of the dog

B) Number of tourists flying to Hawaii vs Cost of plane ticket

C) Number of people going swimming vs Temperature

D) (circled) Amount of clothing sold vs Number of customers at clothing store