S.ID.B.6: Regression 2

1 The table below shows the minimum hourly wage, in U.S. dollars, for selected years since 1955.

Years Since 1955 (x)	0	5	10	15	20	25	30	35	40	45	50
Minimum Wage (y)	.75	1.00	1.25	1.45	2.00	3.10	3.35	3.80	4.25	5.15	5.15

Write the linear regression equation for this set of data, rounding all values to *three decimal places*. State the strength and direction indicated by the correlation coefficient.

2 The table below shows the attendance at a museum in select years from 2007 to 2013.

Attendance at Museum					
Year	2007	2008	2009	2011	2013
Attendance (millions)	8.3	8.5	8.5	8.8	9.3

State the linear regression equation represented by the data table when $x = 0$ is used to represent the year 2007 and y is used to represent the attendance. Round all values to the *nearest hundredth*. State the correlation coefficient to the *nearest hundredth* and determine whether the data suggest a strong or weak association.

3 The table below shows the number of SAT prep classes five students attended and the scores they received on the test.

Number of Prep Classes Attended (x)	3	1	6	7	6
Math SAT Score (y)	500	410	620	720	500

State the linear regression equation for this data set, rounding all values to the *nearest hundredth*. State the correlation coefficient, rounded to the *nearest hundredth*. State what this correlation coefficient indicates about the linear fit of the data.

4 The table below shows the number of math classes missed during a school year for nine students, and their final exam scores.

Number of Classes Missed (x)	2	10	3	22	15	2	20	18	9
Final Exam Score (y)	99	72	90	35	60	80	40	43	75

Write the linear regression equation for this data set. Round all values to the *nearest hundredth*. State the correlation coefficient for your linear regression. Round your answer to the *nearest hundredth*. State what the correlation coefficient indicates about the linear fit of the data.

5 A software company kept a record of their annual budget for advertising and their profit for each of the last eight years. These data are shown in the table below.

Annual Advertising Budget (in thousands, $) ($x$)	**Profit** (in millions, $) ($y$)
10	2.2
13	2.4
14	3.2
16	4.6
19	5.7
24	6.9
24	7.9
28	9.3

Write the linear regression equation for this set of data. State, to the *nearest hundredth*, the correlation coefficient of these linear data. State what this correlation coefficient indicates about the linear fit of the data.

6 Erica, the manager at Stellarbeans, collected data on the daily high temperature and revenue from coffee sales. Data from nine days this past fall are shown in the table below.

	Day 1	**Day 2**	**Day 3**	**Day 4**	**Day 5**	**Day 6**	**Day 7**	**Day 8**	**Day 9**
High Temperature, t	54	50	62	67	70	58	52	46	48
Coffee Sales, f(t)	$2900	$3080	$2500	$2380	$2200	$2700	$3000	$3620	$3720

State the linear regression function, $f(t)$, that estimates the day's coffee sales with a high temperature of t. Round all values to the *nearest integer*. State the correlation coefficient, r, of the data to the *nearest hundredth*. Does r indicate a strong linear relationship between the variables? Explain your reasoning.

7 The data given in the table below show some of the results of a study comparing the height of a certain breed of dog, based upon its mass.

Mass (kg)	4.5	5	4	3.5	5.5	5	5	4	4	6	3.5	5.5
Height (cm)	41	40	35	38	43	44	37	39	42	44	31	30

Write the linear regression equation for these data, where x is the mass and y is the height. Round all values to the *nearest tenth*. State the value of the correlation coefficient to the *nearest tenth*, and explain what it indicates.

8 The percentage of students scoring 85 or better on a mathematics final exam and an English final exam during a recent school year for seven schools is shown in the table below.

Percentage of Students Scoring 85 or Better	
Mathematics, x	**English, y**
27	46
12	28
13	45
10	34
30	56
45	67
20	42

Write the linear regression equation for these data, rounding all values to the *nearest hundredth*. State the correlation coefficient of the linear regression equation, to the *nearest hundredth*. Explain the meaning of this value in the context of these data.

9 The table below shows the number of hours ten students spent studying for a test and their scores.

Hours Spent Studying (x)	0	1	2	4	4	4	6	6	7	8
Test Scores (y)	35	40	46	65	67	70	82	88	82	95

Write the linear regression equation for this data set. Round all values to the *nearest hundredth*. State the correlation coefficient of this line, to the *nearest hundredth*. Explain what the correlation coefficient suggests in the context of the problem.

10 Stephen collected data from a travel website. The data included a hotel's distance from Times Square in Manhattan and the cost of a room for one weekend night in August. A table containing these data appears below.

Distance From Times Square (city blocks) (x)	0	0	1	1	3	4	7	11	14	19
Cost of a Room (dollars) (y)	293	263	244	224	185	170	219	153	136	111

Write the linear regression equation for this data set. Round all values to the *nearest hundredth*. State the correlation coefficient for this data set, to the *nearest hundredth*. Explain what the sign of the correlation coefficient suggests in the context of the problem.

11 The following table represents a sample of sale prices, in thousands of dollars, and number of new homes available at that price in 2017.

Sale Price, p (in thousands of dollars)	160	180	200	220	240	260	280
Number of New Homes Available f(p)	126	103	82	75	82	40	20

State the linear regression function, $f(p)$, that estimates the number of new homes available at a specific sale price, p. Round all values to the *nearest hundredth*. State the correlation coefficient of the data to the *nearest hundredth*. Explain what this means in the context of the problem.

12 Joey recorded his heart rate, in beats per minute (bpm), after doing different numbers of jumping jacks. His results are shown in the table below.

Number of Jumping Jacks **x**	**Heart Rate** (bpm) **y**
0	68
10	84
15	104
20	100
30	120

State the linear regression equation that estimates the heart rate per number of jumping jacks. State the correlation coefficient of the linear regression equation, rounded to the *nearest hundredth*. Explain what the correlation coefficient suggests in the context of this problem.

13 Suzanna collected information about a group of ponies and horses. She made a table showing the height, measured in hands (hh), and the weight, measured in pounds (lbs), of each pony and horse.

Height (hh) x	**Weight** (lbs) y
11	264
12	638
13	700
14	850
15	1000
16	1230
17	1495

Write the linear regression equation for this set of data. Round all values to the *nearest hundredth*. State the correlation coefficient for the linear regression. Round your answer to the *nearest hundredth*. Explain what the correlation coefficient indicates about the linear fit of the data in the context of the problem.

14 An insurance agent is looking at records to determine if there is a relationship between a driver's age and percentage of accidents caused by speeding. The table below shows his data.

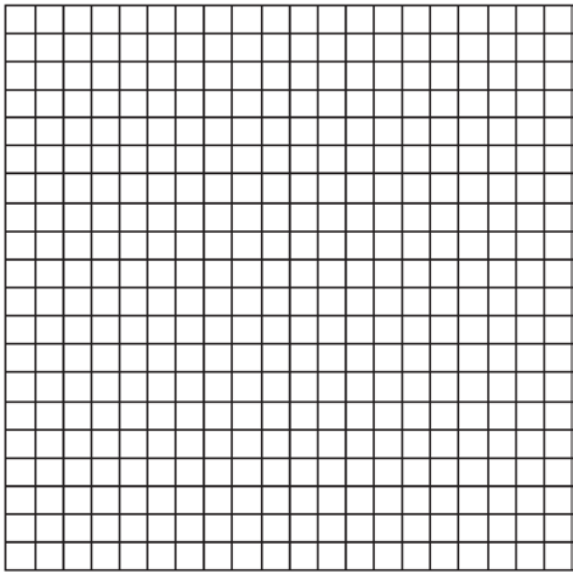
Age (x)	17	18	21	25	30	35	40	45	50	55	60	65
Percentage of Accidents Caused by Speeding (y)	49	49	48	38	31	33	24	25	16	10	5	6

State the linear regression equation that models the relationship between the driver's age, x, and the percentage of accidents caused by speeding, y. Round all values to the *nearest hundredth*. State the value of the correlation coefficient to the *nearest hundredth*. Explain what this means in the context of the problem.

15 Two different tests were designed to measure understanding of a topic. The two tests were given to ten students with the following results:

Test x	75	78	88	92	95	67	58	72	74	81
Test y	81	73	85	88	89	73	66	75	70	78

Construct a scatter plot for these scores, and then write an equation for the line of best fit (round slope and intercept to the *nearest hundredth*). Find the correlation coefficient. Predict the score, to the *nearest integer,* on test y for a student who scored 87 on test x.



S.ID.B.6: Regression 2
Answer Section

1 ANS:
$y = 0.098x + 0.402$ high, positive correlation

REF: 011736a2

2 ANS:
$y = 0.16x + 8.27$ $r = 0.97$, which suggests a strong association.

REF: 081536ai

3 ANS:
$y = 40.48x + 363.81$, 0.84, strong

REF: 012434ai

4 ANS:
$y = -2.81x + 97.55$, -0.97, strong

REF: 012334ai

5 ANS:
$y = 0.41x - 2.31$, 0.99, strong

REF: 082335ai

6 ANS:
$f(t) = -58t + 6182$ $r = -.94$ This indicates a strong linear relationship because r is close to -1.

REF: 011635ai

7 ANS:
$y = 1.9x + 29.8$ $r = 0.3$ This indicates a weak relationship between a dog's height and mass.

REF: 011934ai

8 ANS:
$y = 0.96x + 23.95$, 0.92, high, positive correlation between scores 85 or better on the math and English exams.

REF: 061836ai

9 ANS:
$y = 7.79x + 34.27$ $r = 0.98$ high, positive correlation between hours spent studying and test scores

REF: 061935ai

10 ANS:
$y = -7.76x + 246.34$, -0.88 As the distance from Times Square increases, the cost of a room decreases.

REF: 081935ai

11 ANS:
$f(p) = -.79p + 249.86$ $r = -.95$ There is a strong negative correlation as the higher the sales price, the fewer number of new homes available.

REF: 012035ai

12 ANS:
$y = 1.72x + 69.4$, 0.97, high, positive correlation between the number of jumping jacks and heart rate

REF: 062133ai

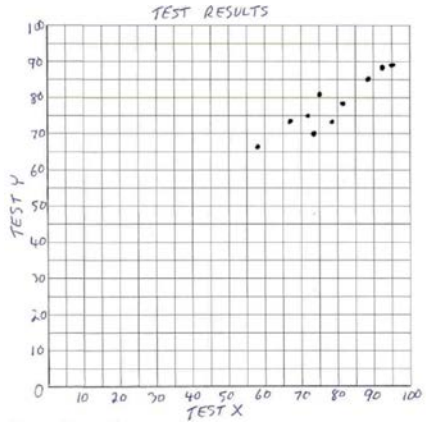
13 ANS:
$y = 184.89x - 1706.07$, 0.99, As the height of the horse increases, the weight of the horse increases.

REF: 062336ai

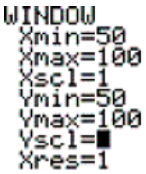
14 ANS:
$y = -0.96x + 64.74$, $r = -0.98$. There is a strong correlation between the driver's age and the percentage of accidents caused by speeding.

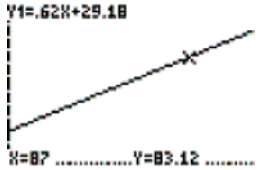
REF: 062235ai

15 ANS:



$y = 0.62x + 29.18$, $r = 0.92$, 83.





REF: 010234b