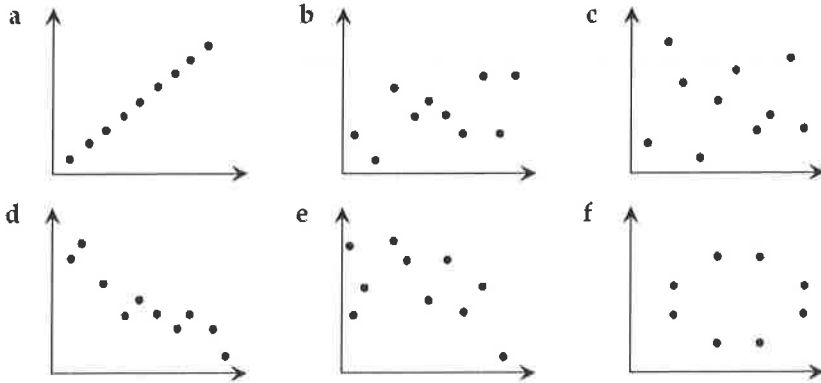


Matching: Match each graph with the best description of its goodness of fit.

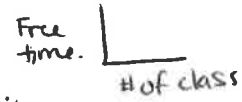


- D 1. Strong, negative
- F, C 2. No correlation
- A 3. Strong, positive
- E 4. Weak, negative
- B 5. Weak, positive

6. Determine whether a scatterplot for the following situations would have a positive, negative, or no correlation.

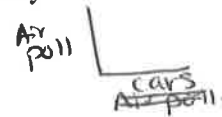
A) The amount of free time you have and the number of classes you take

negative.



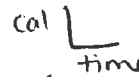
B) The air pollution levels for a city and the number of cars registered in that city

positive.



C) Number of calories burned and the time spent exercising

positive



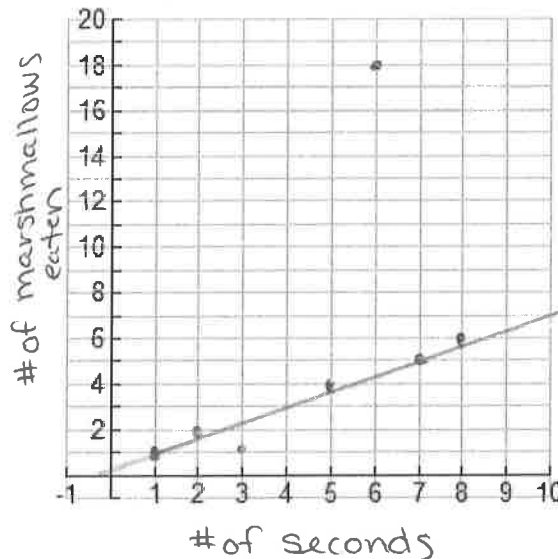
D) Length of a baby at birth and the month in which the baby was born.

None

7. Draw Scatter Plots and Best-Fitting Lines

X = number of seconds	Y = number of marshmallows eaten
1	1
2	2
3	2
5	4
6	18
7	5
8	6

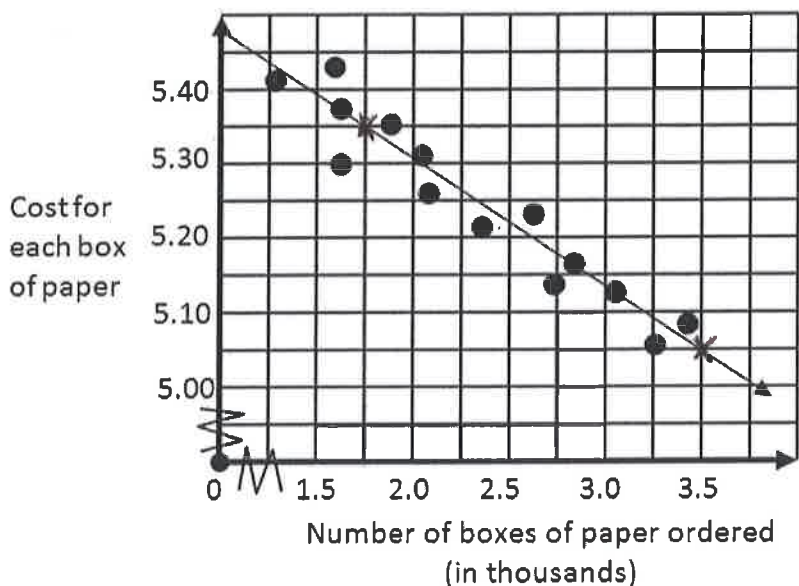
A) Graph the data on the grid below, then draw a best-fit line.



B) Based on your line of best fit how many marshmallows were eaten in 4 seconds? 3 Did you interpolate or extrapolate to find this value? \_\_\_\_\_

C) Based on your line of best fit, how many marshmallows were eaten in 10 seconds? 7 Did you interpolate or extrapolate to find this value? \_\_\_\_\_

8. At the Happy Paper company the more boxes of paper you order the cheaper the price you have to pay for each box of paper. Below are the prices charged per box of paper to different companies ordering various quantities of paper.



A) Using the line of best fit, if your company wants to only pay \$5.05 for each box of paper, how many boxes of paper should be ordered from Happy Paper company? **3,500 boxes**

B) Write an equation for the line of best fit by hand.

$(3.5, 5.05)$   $(1.75, 5.35)$

$y = -0.17x + 5.475$  or  $5.645$   
 (if using graph) (if plugin for x+y)

C) What does the y-intercept mean in the context of the problem? (fee?)  
 For 0 boxes of paper ordered, the cost per box is \$5.475

D) What does the slope mean in the context of the problem?  
 For every # of boxes (in thous) ordered, the price decreases by .174

E) Use your equation to predict the cost of each paper box if 4.5 boxes were ordered.  
 $-0.17(4.5) + 5.475 = \$4.71$  or  $\$4.88$

9. The table below shows the winning times for the Women's 100-Meter Freestyle swims in the Summer Olympics since 1972. Find the equation for best-fit line, letting x represent the number of years after 1972. (Round to nearest hundredth)

0	4	8	12	16	20	24	28	32	36	2012
1972	1976	1980	1984	1988	1992	1996	2000	2004	2008	2012
58.59	55.65	54.79	55.92	54.93	54.64	54.50	53.83	53.84	53.12	

Answer each of the following questions:

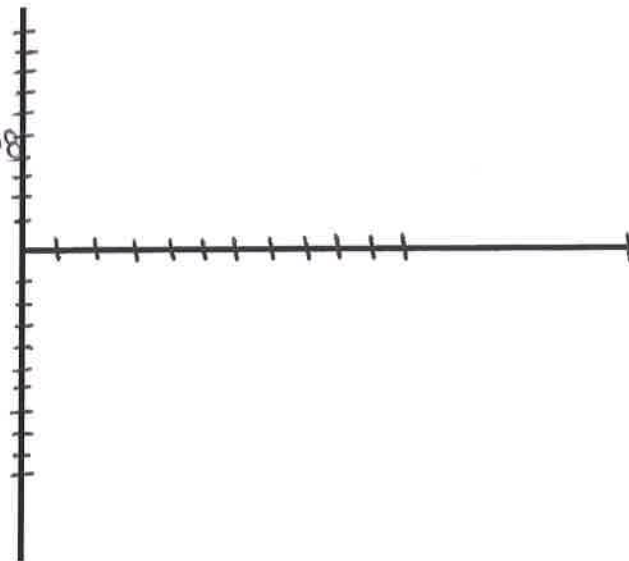
<p>a. What is the slope of the best-fit line?                      (Round to nearest hundredth)  <math>-0.108</math></p>	<p>b. What is the y-intercept of the best-fit line? Round to nearest whole  <math>56.92</math>  <math>57</math></p>	<p>c. What is the correlation coefficient of the best-fit line? (Round to nearest thousandth)  <math>r = -0.85859</math>  <math>r^2 = .73738</math>                      fairly strong negative correlation</p>	<p>d. Predict the winning time for the 2012 Olympic Games.  <math>52.68</math></p>
--	---	---	--

What is the equation of the best-fit line?  $y = -0.108x + 57$

10. Use the line-of-best-fit to find the predicted winning times (rounding to the nearest hundredth), then calculate the residuals. let  $x = \#$  of years after 1960

Line of best fit Equation:  $y = -0.04x + 23.13$

	Year	Time (seconds)	Predicted Time	Residuals
0	1960	24.13	23.13	1.00(41)
4	1964	23.05	22.98	.0753 $\approx$ .08
8	1968	22.58	22.82	-.243 $\approx$ -.24
12	1972	22.40	22.67	-.272 $\approx$ -.27
16	1976	22.37	22.52	-.151 $\approx$ -.15
20	1980	22.03	22.37	-.3398 $\approx$ -.34
24	1984	21.81	22.22	-.4086 $\approx$ -.41
28	1988	21.34	22.07	-.7273 $\approx$ -.73
32	1992	21.81	21.92	-.1061 $\approx$ -.11
36	1996	22.12	21.77	.3551 $\approx$ .36
40	2000	21.84	21.61	.2263 $\approx$ .23
44	2004	22.05	21.46	.5875 $\approx$ .59



Would you describe the residual plot as scattered and random or do you see a pattern? Do you think a linear model is best? yes.

11. Determine if the correlations are positive, negative, or none. Then determine if the statement is a casual relationship.

- A) A student's scores on an examination and students cumulative grade- point average (GPA) upon graduation?  
GPA | positive ; Not casual.
- B) A workers' <sup>SCORES</sup> education levels and wages he/she makes.  
(causal) positive ; not casual
- C) The average speed cars travel from Philadelphia to New York on the turnpike vs. the average amount of times it takes. Negative ; Casual yes.

12. Which example is the best example of "correlation does not imply causation?"

- A. As the availability of a smallpox vaccine increases, the number of smallpox deaths decreases (negative correlation)
- B. The more homework a student turns in, the higher their homework grade will be. (positive correlation)
- C. As the population of foxes increases in an area, the population of bunnies decreases (negative correlation)
- D. As the amount of ice cream sold per month increases, the amount of time spent in a pool increases (positive correlation)