Algebra I
5.6 & 5.7 – Line of Best Fit Notes

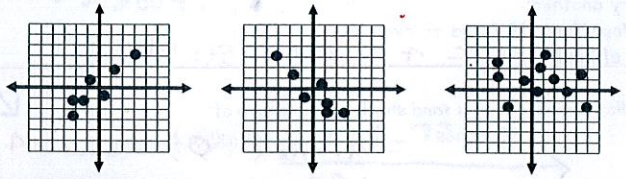
Name _____
Date _____ Block ___

So far, we have learned how to write equations of lines given various pieces of information (slope, points, graph, etc.).

We can also collect data, plot that data and (possibly) come up with the line that best fits that data. We can then use that line to make predictions.

Key Vocabulary
Scatter plot: ➢ a graph used to plot data correlation ➢ used to determine whether there is a relationship between paired data (two variables) ➢ can show possible trend in the data
Correlation: ➢ the relationship between paired data (two variables) ➢ can be Negative, positive or None
Line of Best Fit: ➢ the line that most closely follows a trend in data

Correlations:

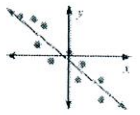


positive — as x increases, y increases

Negative — as x increases, y decreases

None — x and y have no apparent relationship

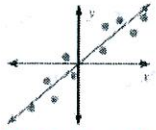
Describe the correlation of data:

a. 

Negative

b. 

None

c. 

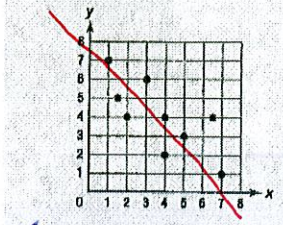
positive

Finding the Line of Best Fit:

1. If given a scatter plot and multiple choice answers:
   a. Use what you know!
      - slope (positive, negative)
      - y-intercept
   b. Eliminate answer choices you know are wrong.
   c. Enter the equations into "y =" and check the graph and the values in the table. Pick the equation that most closely represents the data.

Examples

Which is the best equation of a line of best fit for this scatterplot?



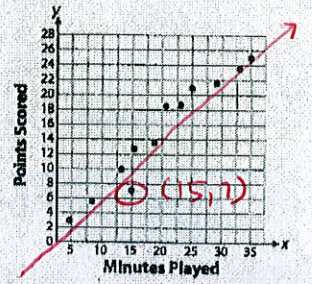
A. $y = x + 8$
B. $y = -x + 8$
C. $y = 8x$
D. $y = -8x$

Slope appears to be Negative
so we can eliminate A, C

Estimate the y-intercept: 8

Enter the equations into y= and check graph/table.

The coach for a girls' basketball team recorded the number of minutes each girl played and the number of points she scored. This is shown in the scatterplot.



Which equation most closely defines the line of best fit for the data?

F $y = \frac{3}{4}x - \frac{1}{3}$ — $\frac{3}{4}(15) - \frac{1}{3} = 10.92$

G $y = \frac{1}{2}x - 1$ — $\frac{1}{2}(15) - 1 = 6.5$

H $y = -\frac{1}{2}x + \frac{3}{4}$

J $y = -x + \frac{2}{3}$

Slope appears to be positive
so eliminate H, J

Estimate the y-intercept: Negative

Enter in y = and check graph/table.

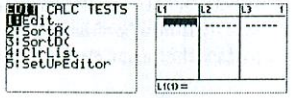
2. If given a table of data:

a. Perform a **LINEAR REGRESSION** using your calculator:

HOW TO PERFORM A LINEAR REGRESSION

Enter Your Data List

- › STAT
- › 1: EDIT
- › Type x values in L1
- › Type y values in L2



Calculate the Linear Regression

- › STAT
- › CALC
- › 4: LinReg(ax + b) … "a" is actually m (slope)
- › substitute the "a" and "b" values into y = ax+ b to get your equation

Examples

x	y
5	1
10	2.5
15	4
20	6
25	7
30	8.5
35	11
40	12.5

+s ↓ + increase ↓

$\frac{12.5-1}{40-5} = \frac{11.5}{35} \approx .33$

Which equation defines the line of best fit for the data in the table?

F $y = \frac{1}{3}x - 10$ H $y = \frac{2}{3}x - 1$

G $y = \frac{1}{3}x - 1$ J $y = \frac{2}{3}x - 10$

Age of car (a)	Value of car (v)
0	12,500
1	9,200
2	7,850
4	6,100
8	3,425

+ ↓ ↓ − * Negative Slope

Which equation most closely defines the line of best fit for the data?

A $v = a + 12{,}500$

B $v = 11{,}000a - 12{,}500$

C $v = -1000a + 8{,}000$

D $v = -1000a + 11{,}000$

Your Turn – Find the line of best fit

1. The table below gives the number of hours spent studying for a science exam (x) and the final exam grade (y).

$y = ax + b$

L_1	x	2	5	1	0	4	2	3
L_2	y	77	92	70	63	90	75	84

a: 6.09 b: 63.93 Line of best fit: $y = 6.09x + 63.93$

2. The table below shows the lengths and corresponding ideal weights of sand sharks.

x (in.)	60	62	64	66	68	70	72
y (lbs.)	105	114	124	131	139	149	158

a: 4.36 b: −156.14 Line of best fit: $y = 4.36x - 156.14$

Making Predictions

Once we have found our line of best fit, we can use that line to make predictions.

Using Your Turn #1 (hours spent studying vs. final exam grade):

a) Line of best fit: $y = 6.09x + 63.93$

b) Predict the final exam grade for a student who studies:

If I study for 0 hrs, I make ____

i. 2.5 hours $= 6.09(2.5) + 63.93$

Grade pt increase per 1 hr of studying

$= 79.16$ ← if I study 2.5 hrs, I make ____

ii. 6 hours

$= 100.47$

c) If a student earned a 98 on the exam, how many hours did he/she study?

$98 = 6.09x + 63.93$ $x \approx 5.60$ hrs

$34.07 = 6.09x$

Let's try another:

Using Your Turn #2 (sand sharks):

a) Line of best fit: $y = 4.36x - 156.14$

b) Predict the weight of a sand shark with a length of:

a zero in shark weighs

i. 80 inches $= 4.36(80) - 156.14$

A shark gain 4 lbs per 1 in of growth

$= 193.66$ ← an 80 in shark weighs

ii. 105 inches

$= 301.66$

c) If a sand shark weighs 250 pounds, what is its length?

$250 = 4.36x - 156.14$

$406.14 = 4.36x$

$93.15 = x$ in

Your Turn

Police investigating traffic accidents often estimate the speed of a vehicle by measuring the length of the tire skid distance. The following table gives the average tire skid distance for an automobile with good tires on dry pavement.

Tire Skid Distance (in feet)	Speed (in miles per hour)
25	28
54	35
89	45
132	55
184	65
244	75
313	85

1. Find the linear regression equation: $y = .2004x + 25.77$

2. Estimate the speed of a vehicle with the following tire skid mark distances:
   (are you finding x or y?)
   a) 150 feet

   $= .2004(150) + 25.77$

   $= 55.83$ mph

   b) 200 feet

   $= .2004(200) + 25.77$

   $= 65.85$ mph

3. Find the tire skid distance of a car travelling at the following speeds:
   (are you finding x or y?)
   a) 52 miles per hour

   $52 = .2004(x) + 25.77$ $\quad x = 130.89$ ft

   $26.23 = .2004x$

   b) 105 miles per hour

   $105 = .2004x + 25.77$

   $79.23 = .2004x$

   $395.36 = x$ ft

Example 2

Listed below are the yearly total enrollment (number of students) figures in Loudoun County Public Schools for the last 10 years.

	Year	Enrollment
0	2005	47,361
1	2006	50,478
2	2007	54,047
3	2008	57,009
4	2009	60,096
5	2010	63,220
6	2011	65,668
7	2012	68,289
8	2013	70,858
9	2014	73,461

1. What is the linear regression equation for the data above?

   $y = 2896.19x + 48015.85$

2. Using your linear regression equation, what is the predicted enrollment for:

   a. 2015 10 years $= 76,977.75$

   b. 2018 13 years $= 85,666.32$

3. In what year can we expect the LCPS enrollment to surpass 95,000 students? (Hint: is "year" the x or the y?).

   $2896.19x + 48015.85 > 95000$

   $2896.19x > 46984.81$

   $x > 16.2$ years

   17 years

   2022

Correlation Coefficient r

Negative

-1 -.5

Negative Relationship Strong

Weak Moderate

None

0

Positive

.5 1

Strong positive relationship

Weak Moderate