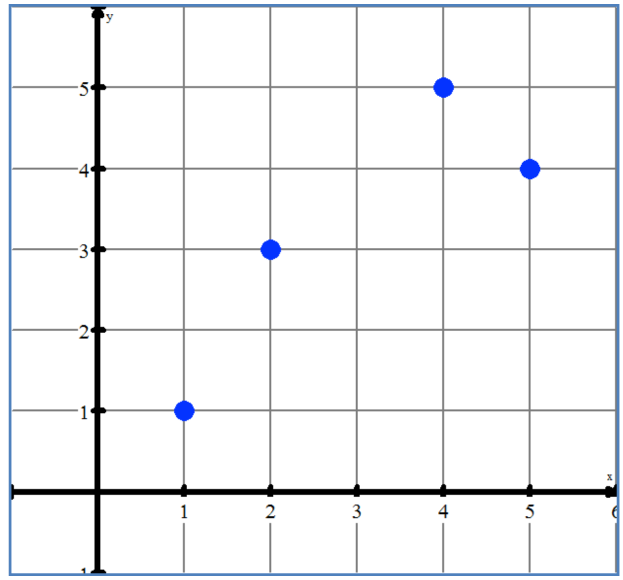
(A) **<u>EXPLORATION #1</u>** ➔ Data Set #1 ➔ (1,1); (2,3); (4,5); (5,4)

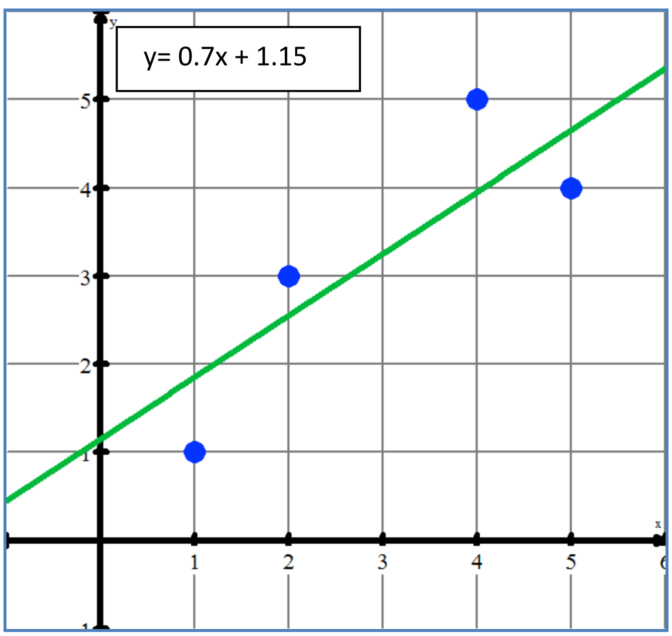


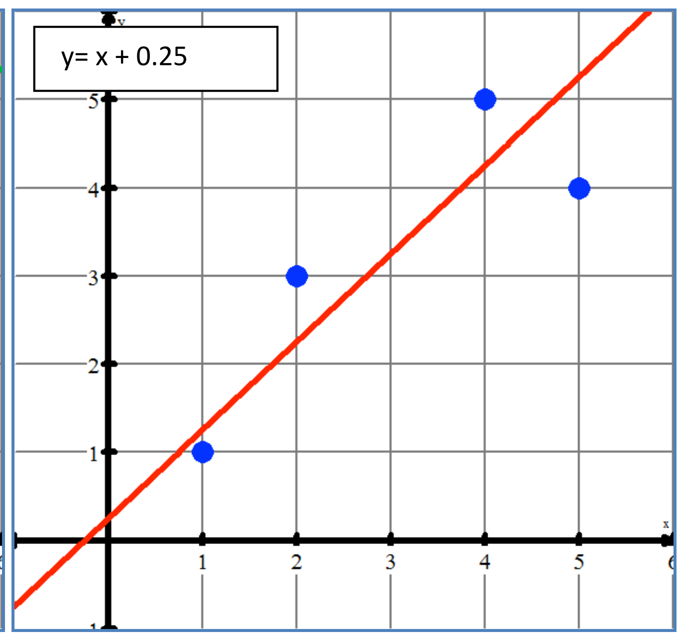
a. Calculate the mean point

b. Draw the line of best fit AS YOU PERCEIVE IT!!!

c. Determine the equation of the line you drew.

d. Student #1 drew the LoBF & decides that the line of best fit has an equation of y = 0.7x + 1.15. But Student #2 drew the LoBF & decides that the equation of the the line should be y = x + ¼. (see graph #2 & graph #3)





e. Q ➔ Is there some way ***<u>descriptive way</u>*** we can determine whose line is a "better" fit for the data?

f. Q ➔ Is there some ***<u>analytical/numeric way</u>*** we can determine whose line is a "better" fit for the data?

(B) Calculating "Residuals"

a. Working With Student #______. and the equation f(x) = __________________________.

x	y	f(x) =	Residual = y – f(x)	Square of residual
1	1			
2	3			
4	5			
5	4			
$\bar{x}$ =	$\bar{y}$ =			$\sum$

b. Working With Student #______. and the equation f(x) = __________________________.

x	y	f(x) =	Residual = y – f(x)	Square of residual
1	1			
2	3			
4	5			
5	4			
$\bar{x}$ =	$\bar{y}$ =			$\sum$

c. CONCLUSION → Which line "fits" better? Why?

d. Now let's go to the following animation and "play" with a line of best fit on a data set: at http://hspm.sph.sc.edu/courses/J716/demos/LeastSquares/LeastSquaresDemo.html

Closing Question → How do I determine which line minimizes the squares of the residuals???

(C) Determining the Eqn of the Least Squares Regression Line & Regression Coefficient

Recall some of our formulas from Lesson 67:

$s_{xy} = \frac{\sum(x-\bar{x})(y-\bar{y})}{n} = \frac{\sum xy}{n} - \bar{x}\bar{y}$ and $s_x = \sqrt{\frac{\sum(x-\bar{x})^2}{n}} = \sqrt{\frac{\sum x^2}{n} - \bar{x}^2}$ and

$s_y = \sqrt{\frac{\sum(y-\bar{y})^2}{n}} = \sqrt{\frac{\sum y^2}{n} - \bar{y}^2}$ and $r = \frac{s_{xy}}{s_x s_y}$ as well as other formulas for r ➔

$$r = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{\sum x^2 - \frac{(\sum x)^2}{n}}\sqrt{\sum y^2 - \frac{(\sum y)^2}{n}}} = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2(y-\bar{y})^2}} = \frac{\sum xy - n\bar{x}\bar{y}}{\sqrt{\sum x^2 - n\bar{x}^2}\sqrt{\sum y^2 - n\bar{y}^2}},$$

so let's apply these calculations here:

x^2	X	y	y^2	xy
	1	1		
	2	3		
	4	5		
	5	4		
$\sum$	$\sum$ so $\bar{x}$ =	$\sum$ so $\bar{y}$ =	$\sum$	$\sum$

So then s_x = ________, s_y = ________, s_{xy} = ________, and r = __________.

Then we add one more formula that will help us write an equation of the least square regression line ➔

$y - \bar{y} = \frac{s_{xy}}{(s_x)^2}(x - \bar{x})$ ➔ so our equation in our example should be ➔

Now test it out on the TI-84 ➔

(D) Further Examples ➔ HH Textbook 18C, p589, Q3

Data Set:

Spray Concentration (mL/L)	3	5	6	8	9	11
Yield of Tomatoes (per bush)	67	90	103	120	124	150

x^2	x	y	y^2	xy
	3	67		
	5	90		
	6	103		
	8	120		
	9	124		
	11	150		
$\sum$	$\sum$ so $\bar{x}$ =	$\sum$ so $\bar{y}$ =	$\sum$	$\sum$

So then s_x = ________, s_y = ________, s_{xy} = ________, and r = __________.

the least square regression line ➔ $y - \bar{y} = \frac{s_{xy}}{(s_x)^2}(x - \bar{x})$ ➔ so our equation in our example should be ➔

(E) Further Examples ➔ HH Textbook 18C, p589, Q5

Data Set:

Frost Free days	75	100	125	150	175	200
Rate of Reaction	44.6	42.1	39.4	37.0	34.1	31.2

x^2	x	y	y^2	xy
	75	44.6		
	100	42.1		
	125	39.4		
	150	37		
	175	34.1		
	200	31.2		
$\sum$	$\sum$ so $\bar{x}$ =	$\sum$ so $\bar{y}$ =	$\sum$	$\sum$

So then s_x = ________, s_y = ________, s_{xy} = ________, and r = __________.

the least square regression line ➔ $y - \bar{y} = \frac{s_{xy}}{(s_x)^2}(x - \bar{x})$ ➔ so our equation in our example should be ➔

Now test it out on the TI-84 ➔