AP Statistics - Class Activity: The Wandering Point   *Chapter 9*

1. The scatterplot shows the four points (1,2), (2, 6), (4, 2) and (5,6) plotted in a 10x10 graphing window. Find the correlation and the equation of the line of best fit.
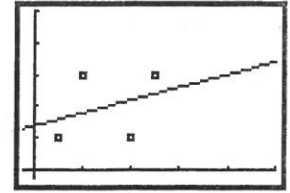
    $r =$ __.3162__ $\hat{y} =$ __2.8 + .4x__

2. Now investigate the influence of one more point on the correlation and the slope. Try each of these as the fifth point and record the new correlation and slope. Also note whether the new point has a small or large residual. (NOTE: Add each point to the original four, one at a time, see what happens, and then remove that point. There are never more than 5 points in the plot!)

Fifth point	Description	Correlation	Slope of the regression line	Size of the residual
None	the original four points	.3162	.4	N/A
(3,4)	right in the center of the given points	.3162	.4	0
(8,6)	also on the line, but far from the other points	.5	.4	0
(10,7)	only close to the line, but much farther away	.6157	.4228	small
(3,8)	above the center of the original cluster	.2357	.4	large
(1,7)	nearby, but not consistent with the apparent pattern	-.0457	-.06	medium?
(8,9)	farther away, and also not consistent	.7303	.8	small
(10,0)	farther and stranger…	.4888	-.3740	small

3. A point that dramatically changes the apparent slope of the regression line is called an *influential* point. You need to be able to spot potential influential points in a scatterplot. What should you look for?

4. Originally there were only four points here. Suppose instead that we had started with 50 points clustered in essentially the same region and displaying an association of roughly the same strength and direction. Would our fifty-first point still be as influential? Where would you locate one additional point so influential that it changed the line as dramatically as (10,0) did above?