AP Stats
Chap 8 Activity Name ____ Pd __________

Graduating Classes

The table at the right shows the number of students who graduated from a high school during the latter half of the last century. Data was not available for all of the years, especially those long ago.

Year	Number Graduated
1950	203
1953	211
1957	288
1960	319
1962	381
1965	446
1968	439
1970	521
1972	509
1976	527
1980	476
1984	413
1987	399
1989	362
1992	379
1994	413
1996	437
1998	426
2000	451

1. Using Excel (to create the spreadsheet) and Fathom (to make the display), create a scatterplot of the data. Print this out – one copy for each of you – and attach it to this sheet when handing it in.

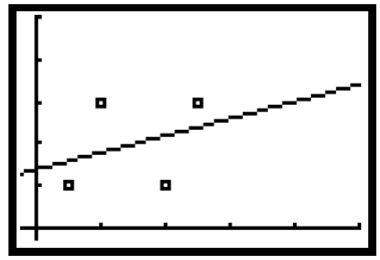
2. Describe the trends you see in the data.

3. These data do not show the size of the graduating class in 1969. Create an appropriate model and use it to estimate the size of that class. Explain what years you used and why.

4. If we wanted to project the graduating sizes of this school through the year 2005, what model would you use? What aspect of these data makes you cautious about your projections? Why?

AP Stats
Chap 8 Activity Name ____ Pd __________

The Wandering Point



The scatterplot shows the four points (1 , 2) , (2 , 6) , (4 , 2) , and (5 , 6) plotted in a 10-by-10 graphing window.

1. Find the correlation and the equation of the line of best fit.

2. Now investigate the influence of one more point on the correlation and the slope. Try each of these points as the fifth point **(one of them at a time)** and record the new correlation and slope. Also note whether the new point has a small or large residual.

Fifth point	Description	Correlation	Slope of the regression line	Size of the residual
None	the original four points			N/A
(3,4)	right in the center of the given points			
(8,6)	also on the line, but far from the other points			
(10,7)	only close to the line, but much farther away			
(3,8)	above the center of the original cluster			
(1,7)	nearby, but not consistent with the apparent pattern			
(8,9)	farther away, and also not consistent			
(10,0)	farther and stranger…			

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 be looking 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?