Influencing Regression

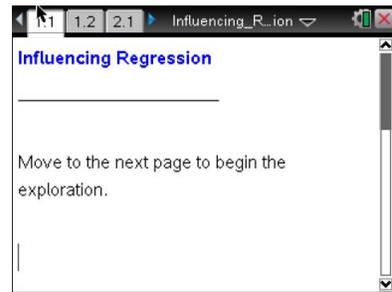## Student Activity

Name _____

Class _____

Open the TI-Nspire document *Influencing_Regression.tns.*

Does one point really make a difference? In this activity, you will consider whether a single point can change the correlation coefficient and the regression line.

Move to page 1.2.

1. a. Select the line on the scatter plot on Page 1.2. What is the equation of the least-squares regression line?

    b. The correlation coefficient for these data is 0.8944. Interpret this correlation coefficient in terms of the graph.

Move to page 2.1.

2. On this page, click the arrow to add the data point (3, 5) from the table to the plot.

    a. Fill in the table below with the new correlation coefficient, the equation of the least-squares regression line, the position of the line, and the change from the original plot with respect to the correlation coefficient.
        •   Select the arrow to replace the new point with the next ordered pair. Continue until you have completed the table.

Point added	Correlation coefficient r	Equation of the Least-Squares Regression Line $\hat{y} = a + bx$ and Position of Line	What changed from original?
none			
(3, 5)			
(5, 7)			
(3, 0)			
(4, 10)			
(10, 0)			
(10, 5)			
(10, 12)			

b. Did you add any points for which the regression equation did not change? Explain.

c. Did you add any points for which the correlation coefficient did not change? Explain.

d. Did you add any points for which the slope did not change?

Move to page 3.1.

The question below will repeat the process from question 2 with more extreme data values.

3. In the table below, **predict** how this new data point will change the slope of the original least-squares regression line, and its effect on the correlation coefficient (e.g., change sign, decrease a little, decrease a lot, no change, increase a little, increase a lot).
    - Check your prediction by using the arrow to add the point in your TI-Nspire document.
    - Record the new equation in the table below.
    - Repeat the process with the remaining points.

Point added	Change in slope	Change in correlation coefficient	Actual correlation coefficient	Equation of least-squares regression line
none				
(50, 0)				
(50, 5)				
(50, 53)				

Use the information in the tables from Questions 2 and 3 to answer the questions below.

4. Points that lie within the *x*-range of the data and change the *y*-intercept but do not change the slope of the least-squares regression line are called *outliers in regression*. Which points did you add that would be considered outliers?

5. Points that do not lie within the *x*-range of the data and change the slope of the regression line are called *influential points*. Which points did you add that would be considered influential?

Assessment

1. Where would you place a point to make the correlation coefficient as close to 0 as possible?

2. Where would you place a point to make the correlation coefficient as close to 1 as possible?

3. Where would you place a point to not change the correlation coefficient or the equation of the least-squares regression line?

Adapted from an idea from Matching Descriptions to Scatter Plots, (1996). *Activity Based Statistics,* by Richard Scheaffer, Mrudulla Gnanadesikan, Ann Watkins, & Jeffrey Witmer. Springer.