Revised 2017

AP Statistics - 3.1B Activity

Linear Regression Activity

Name: KEY (Version 2 of 2) Class: _____ Date: _____

Part 1: The purpose is to understand correlation coefficient

ACTIVITY: Guessing correlation http://istics.net/Correlations/ **Complete 3-4 Round**

There are 4 graphs, guess the value of the correlation coefficient.

Part 2: The purpose is to explore the impact of outliers on an LSRL equation.

Enter the following lists into your calculator:

x	4	5	6	7	8
y	2	5	4	6	9

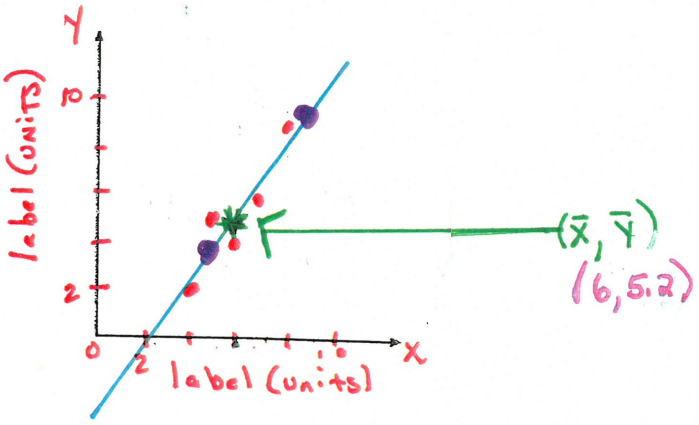
1) Sketch a scatterplot.

$y =$ $\hat{y} = -3.8 + 1.5x$

TABLE

x	5	9
y	3.7	9.7

NOTICE $(\bar{x}, \bar{y}) \rightarrow (6, 5.2)$ is on LSRL



a) Explain correlation coefficient $r = .916$

The association based on "r" is positive and strong. The graph indicates a linear relationship.

b) Explain slope $b = 1.5$

(in context)

The predicted y increases about 1.5 UNITS FOR AN increase of 1 UNIT IN X. (remember Context + hedging words)

c) Explain y-intercept $a = -3.8$

When x IS ZERO, typically the starting point or resting point. (sometimes the yint is meaningless)

d) Explain YHAT $(\hat{y})$

$\hat{Y}$ is the predicted value. It must where a hat!!

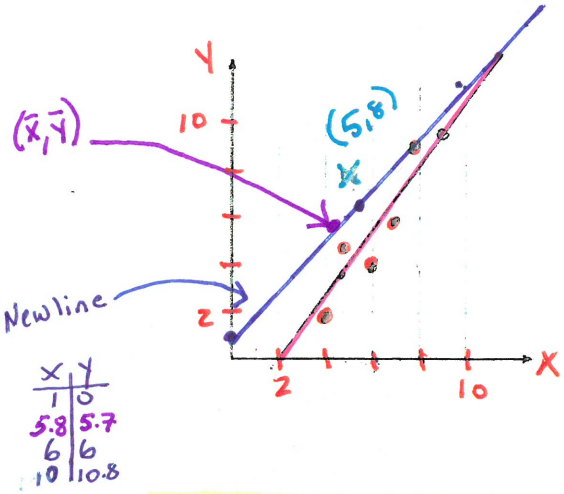
e) Find the least squares regression equation. Write equation in context.

$\hat{y} = -3.8 + 1.5x$ where x = _____

y = _____

OUTLIER

2) **Add the data point (5,8)** to the lists. Sketch a scatterplot. Identify the point (5,8) with an X on your plot. Find the LSRL and the correlation coefficient. Notice what effect this point has on the LSRL and r.



x	y
1	0
5.8	5.7
6	6
10	10.8

ORIG LSRL

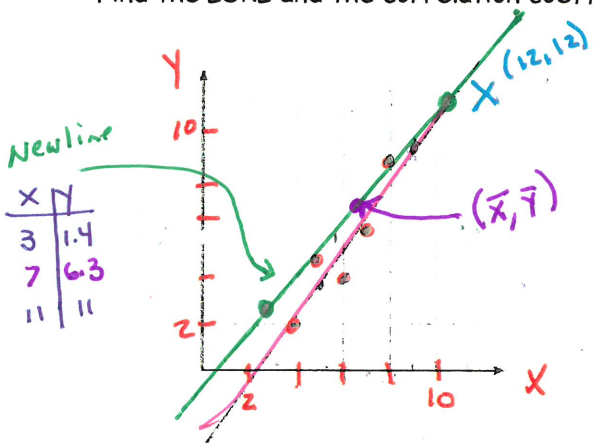
x	y
5	3.7
9	9.7

$\hat{y} = -1.2 + 1.2x$ $\quad r = .667$

① The pt (5,8) decreased the strength of the association

② The LSRL shifts
- the y-intercept moved up and the line became less steep

3) **Change the point (5,8) to (12,12).** Sketch a scatterplot. Identify the point (12,12) with an X on your plot. Find the LSRL and the correlation coefficient. Notice what effect this point has on the LSRL and r.



x	y
3	1.4
7	6.3
11	11

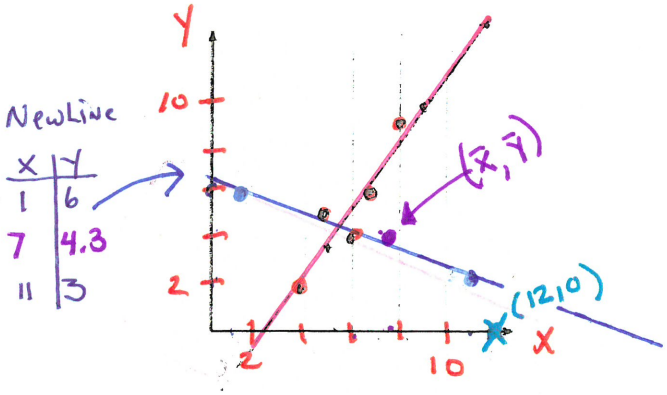
$\hat{y} = -2.2 + 1.2x$ $\quad r = .9585$

① The pt (12,12) increased the strength of the association because it added a point that was close to the original LSRL

② The LSRL's are quite similar.

INFLUENCIAL POINT

4) **Change the point (12,12) to (12,0).** Sketch a scatterplot. Identify the point (12,0) with an X on your plot. Find the LSRL and the correlation coefficient. Notice what effect this point has on the LSRL and r.



x	y
1	6
7	4.3
11	3

$\hat{y} = 6.3 - 0.3x$ $\quad r = -.248$

① The pt (12,0) had a huge impact on the association.
- r has switched from positive to negative
- r is now very weak.

② The new LSRL has a negative slope and dramatically incre

TIP: Reread section 3.1 to understand concepts covered in this activity !

y-int.