AP Stats Unit 3 Date _____

Section 3.1 – Scatterplots and Correlation:
Each member of a small statistics class ran a 40-yard sprint and then did a long jump (with a running start). The table below shows the sprint time (in seconds) and the long-jump distance (in inches)

Sprint time (s)	5.41	5.05	7.01	7.17	6.73	5.68	5.78	6.31	6.44	6.50	6.80	7.25
Long-jump distance (in)	171	184	90	65	78	130	173	143	92	139	120	110

Problem: Make a scatterplot of the relationship between sprint time and long-jump distance.

Hint – Which variable makes more sense as the explanatory variable vs the response variable?

Use our work from yesterday to describe your scatterplot and any relationships in the data.

Direction:

Form:

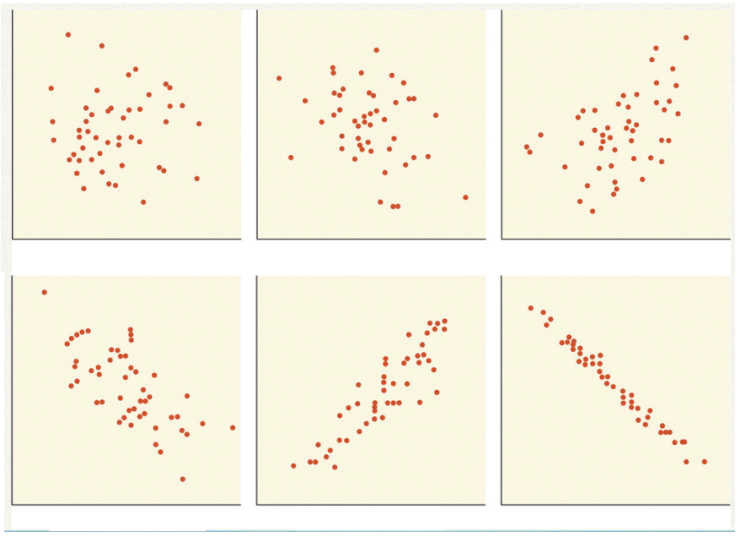
Strength:

Outliers:

Measuring Linear Association
Correlation (r):

Key properties of r:

Estimate the Correlation value for these graphs



How Correlation is calculated:

Suppose that we have data on variables x and y for n individuals. The values for the first individual are x_1 and y_1, the values for the second individual are x_2 and y_2, and so on. The means and standard deviations of the two variables are *x-bar* and s_x for the x-values and *y-bar* and s_y for the y-values.

The correlation r between x and y is:

$$r = \frac{1}{n-1}\left[\left(\frac{x_1-\bar{x}}{s_x}\right)\left(\frac{y_1-\bar{y}}{s_y}\right)+\left(\frac{x_2-\bar{x}}{s_x}\right)\left(\frac{y_2-\bar{y}}{s_y}\right)+\ldots+\left(\frac{x_n-\bar{x}}{s_x}\right)\left(\frac{y_n-\bar{y}}{s_y}\right)\right]$$

$$r = \frac{1}{n-1}\sum\left(\frac{x_i-\bar{x}}{s_x}\right)\left(\frac{y_i-\bar{y}}{s_y}\right)$$

What do you notice about this formula?

Comments on Correlation