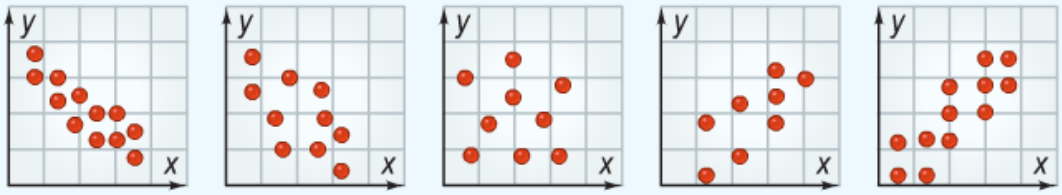
Scatterplots and Linear Regression

A **scatter plot** is a graph that relates two sets of data by plotting the data as ordered pairs. You can use a scatter plot to determine the strength of the relationship, or **correlation**, between data sets. The closer the data points fall along a line with positive slope,

- the stronger the linear relationship and
- the stronger the positive correlation

between the two variables.

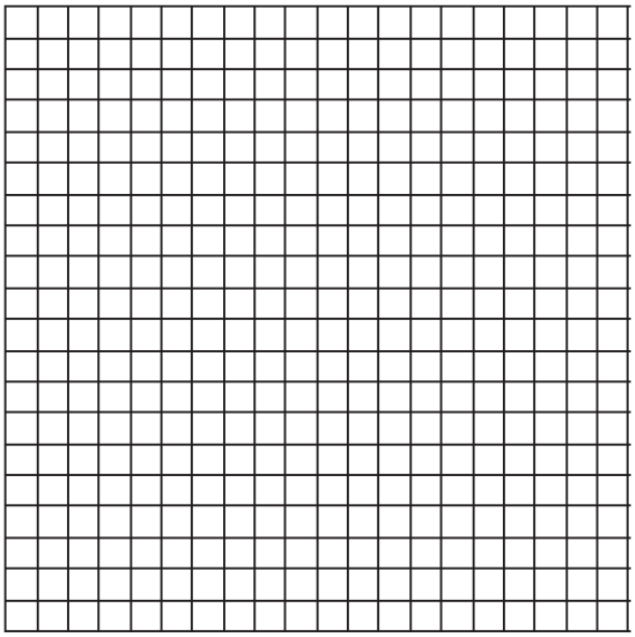


Strong Negative Correlation | Weak Negative Correlation | No Correlation | Weak Positive Correlation | Strong Positive Correlation

1. **Utilities** The table lists average monthly temperatures and electricity costs for a Texas home in 2008. The table displays the values rounded to the nearest whole number. Make a scatter plot. How would you describe the correlation?

Average Temperatures and Electricity Costs

Month	Average Temp. (°F)	Electricity Bill ($)	Month	Average Temp. (°F)	Electricity Bill ($)
January	61	150	July	84	255
February	58	139	August	85	245
March	67	172	September	81	210
April	75	205	October	76	183
May	79	170	November	65	132
June	83	234	December	58	110



A **trend line** is a line that approximates the relationship between the variables, or data sets, of a scatterplot. You can use a trend line to make predictions from the data.

2. Draw a trend line on the scatterplot above. Write an equation of the trend line. Use your equation to predict the electricity cost when the average monthly temperature reaches 100°F.

The trend line that gives the most accurate model of related data is the **line of best fit**. One method for finding a line of best fit is *linear regression*. You can use the **LinReg** function on your graphing calculator to find the line of best fit. The **correlation coefficient**, r, indicates the strength of the correlation. The closer r is to 1 or -1, the more closely the data resembles a line and the more accurate your model is likely to be.

3. Use a graphing calculator to determine the equation for the line of best fit for the data above. How accurate is your line of best fit? Based on your linear model, what is the electricity cost when the average monthly temperature reaches 100°F?