

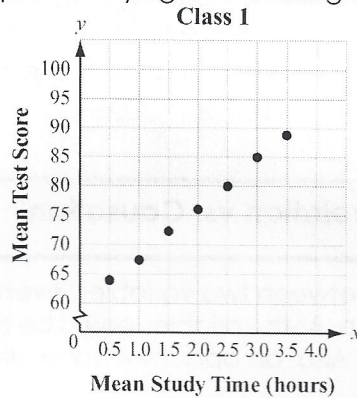
Name: Answer Key Date: \_\_\_\_\_

## Correlation

**MCC9-12.S.ID.6** Represent data on two quantitative variables on a scatter plot, and describe how the variables are related.  
**MCC9-12.S.ID.9** Distinguish between correlation and causation.

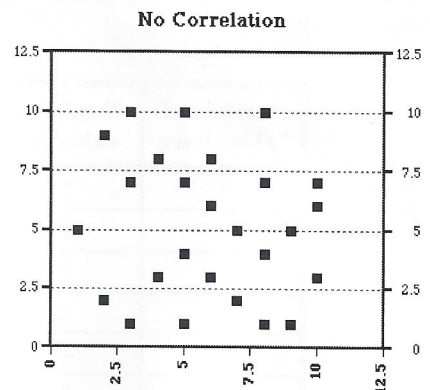
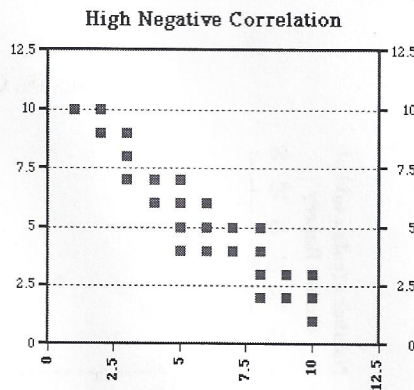
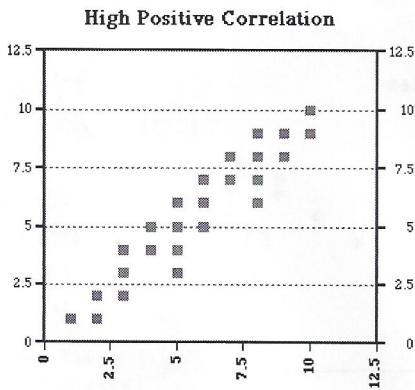
A **scatter plot** is often used to present bivariate **quantitative** data. Each variable is represented on an axis and the axes are labeled accordingly.

A scatter plot displays data as points on a grid using the associated numbers as coordinates or ordered pairs  $(x, y)$ . The way the points are arranged by themselves in a scatter plot may or may not suggest a relationship between the two variables. For instance, by reading the graph below, do you think there is a relationship between the hours spent studying and exam grades?



If  $y$  tends to increase as  $x$  increases, then the data have **positive** correlation.

If  $y$  tends to decrease as  $x$  increases, then the data have **negative** correlation.

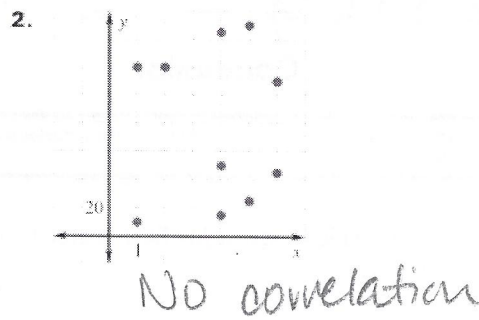
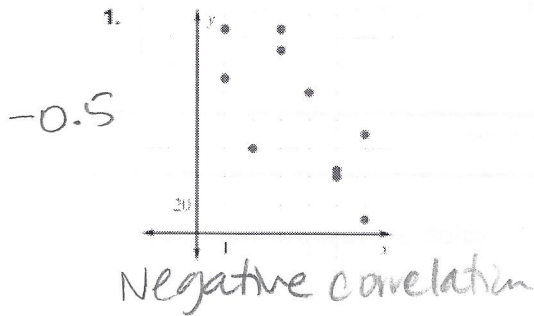


A correlation coefficient, denoted by  $r$ , is a number from  $-1$  to  $1$  that measures how well a line fits a set of data pairs  $(x, y)$ . If  $r$  is near  $1$ , the points lie close to a line with a positive slope. If  $r$  is near  $-1$ , the points lie close to a line with a negative slope. If  $r$  is near  $0$ , the points do not lie close to any line.

Give an example of negative correlation: The temperature outside from August to February.

**Practice Problems:**

For each scatter plot, tell whether the data have a positive correlation, a negative correlation, or no correlation. Then, tell whether the correlation is closest to  $-1$ ,  $-0.5$ ,  $0$ ,  $0.5$ , or  $1$ .



3. Positive, negative, or no correlation?

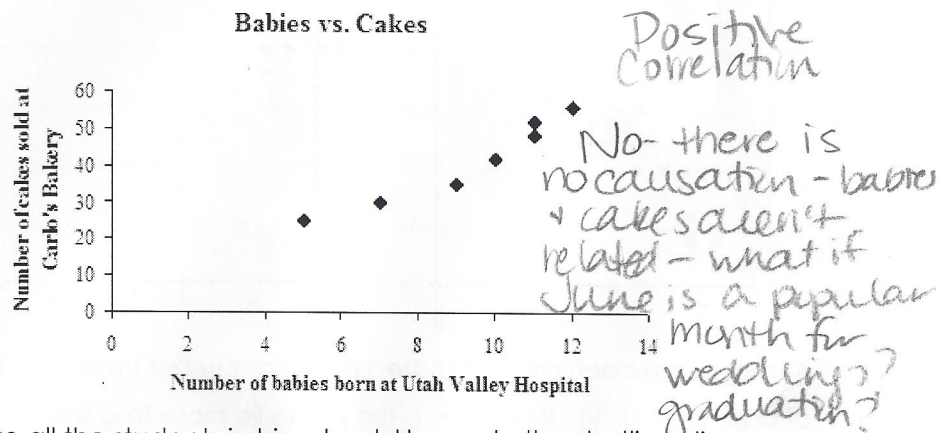
- a. Amount of exercise and percent of body fat Negative
- b. A person's age and the number of medical conditions they have Positive
- c. Temperature and number of ice cream cones sold Positive
- d. The number of students at Hillgrove and the number of dogs in Atlanta No correlation
- e. Age of a tadpole and the length of its tail Negative

### Correlation vs. Causation

When a scatter plot shows a correlation between two variables, even if it's a strong one, there is not necessarily a cause-and-effect relationship. Both variables could be related to some third variable that actually causes the apparent correlation. Also, an apparent correlation simply could be the result of chance.

**Example 1:** During the month of June the number of new babies born at the Utah Valley Hospital was recorded for a week. Over the same time period, the number of cakes sold at Carlo's Bakery in Hoboken, New Jersey was also recorded. What can be said about the correlation? Is there causation? Why or why not?

Number of babies born	Number of cakes sold
5	25
7	30
9	35
10	42
11	48
11	52
12	56



**Example 2:** Mr. Jones gave a math test to all the students in his school. He made the startling discovery that the taller students did better than the short ones. His Causation Statement: *As your height increases, so does your math ability.*

What can be said about the correlation? Is there causation? Why or why not?

*There is a positive correlation but no causation - height has nothing to do w/ math ability.*

**Example 3:** In this present economy families are trying to find ways to save money Families might be thinking about not eating out to spend less money. Causation Statement: *The more you eat out, the more money you spend at restaurants.*

What can be said about the correlation? Is there causation? Why or why not?

*There is a positive correlation + yes there is causation. B/c you eat out, you spend \$. So the more you eat out, the more you spend.*