Name: Date:

**Student Exploration: Correlation**

**Vocabulary:** correlation, correlation coefficient, least-squares fit line, scatter plot



**Prior Knowledge Questions** (Do these BEFORE using the Gizmo.)

Mr. Bryant asks his students how far they live from school, and their travel time to get to school. He plots a point for each student’s distance in miles (*x*) and travel time in minutes (*y*) to create a **scatter plot** of his data.

1. Suppose Mr. Bryant’s student Elena lives 3 miles from school and it takes her 15 minutes to get to school.

What point will Mr. Bryant plot for Elena? ( , )

1. How do you think increased distance will affect travel time?
2. Sketch points in the grid above to illustrate what you think the data might look like.

**Gizmo Warm-up**

When one variable is related to another, the two variables are said to be **correlated**. In many cases, variables that are correlated have a roughly linear relationship, where the scatter plot approximates a line. You can explore linear correlation with the *Correlation* Gizmo.

The variable *r* is called the **correlation coefficient**. Move the ***r*** slider back and forth and observe the scatter plot.

1. How would you describe the scatter plot when *r* is close to 1?

1. How does the scatter plot look when *r* is near –1?

1. Describe the graph when *r* is near 0.

|  |  |  |
| --- | --- | --- |
| **Activity:** **Correlation and lines of best fit** | Get the Gizmo ready: * Set ***r*** to 1.00. (To quickly set a slider to a specific value, type the value into the text box to the right of the slider, and hit **Enter**.)
 | 101SE2 |

1. In a data set with a strong linear correlation, the points in the scatter plot approximate a line. Turn on **Show least-squares fit line**. The **least-squares fit line** is the “best-fit” line, or the line that most closely “fits” the shape of the data.
2. When *r* = 1, how are the points in the scatter plot related to the least-squares fit line?

1. Slowly decrease ***r***. How does this affect where the points are in relation to the line?

1. With **Show least-squares fit line** still selected, set ***r*** to 0.90. The points should be close to the line, but not right on it. Below **Generate new data set with:** click **Same *r*** several times.
2. Do all the least-squares fit lines for these scatter plots have the same slope?
3. Do all the least-squares fit lines have the same *y*-intercept?
4. What do all the least-squares fit lines have in common?

A positive *r* indicates a *positive correlation*: as *x* increases, *y* also tends to increase.

1. Set ***r*** to –0.90. Click **Same *r*** several times. What do the least-squares fit lines for these scatter plots have in common?

A negative *r* indicates a *negative correlation*: as *x* increases, *y* tends to decrease.

1. Set ***r*** to 0.00. Click **Same *r*** several times.
2. Do all the least-squares fit lines for these scatter plots have the same slope?
3. Do all the least-squares fit lines have the same *y*-intercept?
4. What do all the least-squares fit lines have in common?

When *r* = 0, there is no correlation in the data. This means that the value of *y* does not seem to be at all related to the value of *x*.

**(Activity continued on next page)**

**Activity (continued from previous page)**

1. Turn off **Show least-squares fit line**. Click **New *r***, and sketch the scatter plot to the right.

What is the value of *r*?

Turn on **Fit a line**. Use the slope (***m***) and *y*-intercept (***b***) sliders to estimate the line that fits this data set best. Sketch your line and record its equation below.

Equation of estimated line:

Check your estimate by turning on **Show least-squares fit line**. Record the equation for the actual least-squares fit line.

Least-squares fit line equation: Was your estimate close?

1. Turn off **Show least-squares fit line**. Click **New *r*** several times. For each data set, try to fit the red line to the data, and then check it by turning on **Show least-squares fit line**.

How does the value of *r* relate to how easy it is to estimate the least-squares fit line?

1. Three scatter plots are shown below. Use them to answer the questions below the graphs.



1. For one of the three scatter plots, *r* = –0.83. Which one do you think it is?

Explain.

1. Which graph has a least-squares fit line with the equation *y* = 0.6*x* + 1.75?

Explain.