

The **Least-Squares Regression** simulation encourages students to explore data on two quantitative variables, interpret the correlation coefficient, fit a linear function to various data sets, understand how to achieve a best-fit line, and determine whether a linear fit is appropriate.

The screenshot shows the PhET Least-Squares Regression simulation interface. It features a central scatter plot with a best-fit line and a menu of data sets on the right. On the left, there are controls for the best-fit line and correlation coefficient. A blue bowl icon is used to create a custom data set.

OBSERVE the change in the sum of the squared residuals as a point is added or moved.

WATCH r update with each new data point.

DRAG points to create a custom data set.

EXPLORE a data set from the menu.

FIT a line to the data set.

Suggestions for Use

Sample Challenge Prompts

- Create a custom data set with...
 - a correlation coefficient that is *positive*.
 - a correlation coefficient that is *negative*.
 - a correlation coefficient of *zero*.
- Create a custom data set with a linear association. Try to fit a line to it using the My Line controls, explain how you decided on your final line, then show the Best-Fit Line to see how close you were.
- Create a custom data set and show the best-fit line. Choose a point to drag and observe how it influences the best-fit line.
- For each data set in the menu, determine if a linear fit is appropriate, and justify your answer.
- Why do we call this type of regression “least-squares”?

See all activities for Least-Squares Regression [here](#).

For more tips on using PhET sims with your students, see [Tips for Using PhET](#).