The Line of Best Fit

Students learn what the least squares regression line is and why distances are squared, fit a line with technology, interpret the slope and y-intercept in context, predict forward and backward, and judge predictions as interpolation or extrapolation.

TIME	45 minutes (one student with a parent or tutor, or a small group)
WHO	Math 1 (algebra readiness), about age 12. Follows Unit 10, Lesson 1.
MATERIALS	Student worksheet (PDF), the Deep Dive 5 interactive page for this lesson (regression tool), pencil. A graphing calculator can be used instead of the site's tool.
PRIOR KNOWLEDGE	Scatter plots and r (Lesson 1). Slope and y-intercept in context (Unit 3). Solving two-step equations (Unit 1).
STANDARDS	NC.M1.S-ID.6a, NC.M1.S-ID.7, NC.M1.F-LE.5.
KEEP OUT	Computing the regression line by hand. Residual plots (Lesson 3).

Learning goals

1. Explain what the least squares regression line is and why the distances are squared.
2. Fit a line with technology and write it with m and b.
3. Interpret the slope and the y-intercept in the story, using the word about.
4. Predict forward and backward, classify a prediction as interpolation or extrapolation, and judge whether it makes sense.

Vocabulary: say both, every time

The rule for the whole lesson: whenever the adult or the student says a math word, they follow it with the plain-English version. Proper term first, plain English second, until the student does it without being asked.

TERM	LOOKS LIKE	PLAIN ENGLISH
Line of best fit	least squares regression line	Closest to all the points.
Regression	finding the best fit	Technology does it.
Predicted value	what the model says	Say about.
Interpolate, extrapolate	inside, outside the data	Safe, risky.

> THE ONE SENTENCE TO SAY OUT LOUD
>
> A model is a family member that fits the data best. Read its numbers in the story, and always ask if the prediction makes sense.

Lesson timeline (45 minutes)

0:00 to 0:06	Warm-up: draw a line by eye. Use the eyeball tool on the rat data. Compare total squared errors, then reveal the least squares line.
0:06 to 0:14	Mini-lesson: least squares. Many lines are possible; the best makes the total squared miss smallest. Why square: so misses cannot cancel. Technology finds m and b.
0:14 to 0:26	Read the model. W = 100 + 40t: slope, y-intercept, the word about. Forward to week 10. Backward to 760 grams. Interpolation and extrapolation. The 5-year-old rat.
0:26 to 0:34	The regression tool. Used cars: m, b, r, a price from miles, miles from a price. Ask the sense question at 170 thousand miles.
0:34 to 0:41	Practice. The practice problems or Worksheet Parts A to C. Same buttons: Hint, Show next step, Reveal answer, Hide help.
0:41 to 0:45	Exit ticket. Worksheet Part E. Success looks like about 2 cm a week, about 5 cm at the start, 17 cm, and week 10.

Misconceptions to listen for

- **The best line goes through the most points.** It may go through none; it balances the misses.
- **The model is exact.** Say about: the points are not on the line.
- **Any prediction is fine.** Far outside the data can be nonsense.
- **The y-intercept always has a real meaning.** Only if x = 0 makes sense in the story.
- **A strong r means safe extrapolation.** It does not.

Support and stretch

Support

Have the student say the slope sentence with units before any calculation: about 40 grams per week. Mark the data range on a number line and place each prediction on it to decide inside or outside.

Stretch

Ask what happens to the line of best fit when one far-away point is added (try it in the Lesson 1 builder). Ask for a situation where the y-intercept has no real meaning.

Where this goes next

Lesson 3: residuals. The gap between each point and the line is the model's error, and its pattern tells you whether a line was the right family.

> THE HABIT TO BUILD
>
> Fit, read the slope and y-intercept in the story, predict, and ask: inside or outside the data, and does it make sense?