

Fitting Lines to Data and Making Predictions

Grade 8 math · Lesson · Standards: 8.SP.A.2, 8.SP.A.3

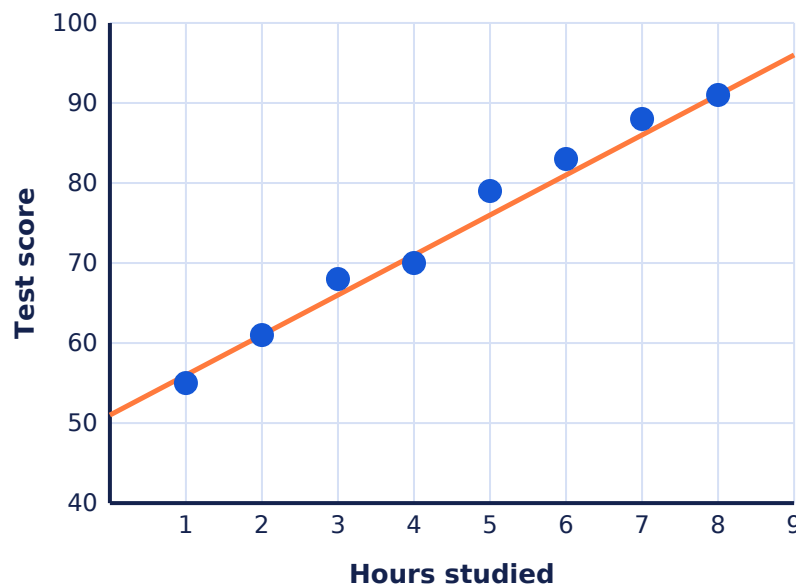
Quick answer: A line of best fit is a straight line drawn through a scatter plot to follow its trend. Its equation $y = mx + b$ gives predictions: the slope m is the change in y for each 1 unit of x , and the intercept b is the predicted y when x is 0.

Standards: 8.SP.A.2, 8.SP.A.3 · **Level:** Grade 8

What is a line of best fit?

Line of best fit

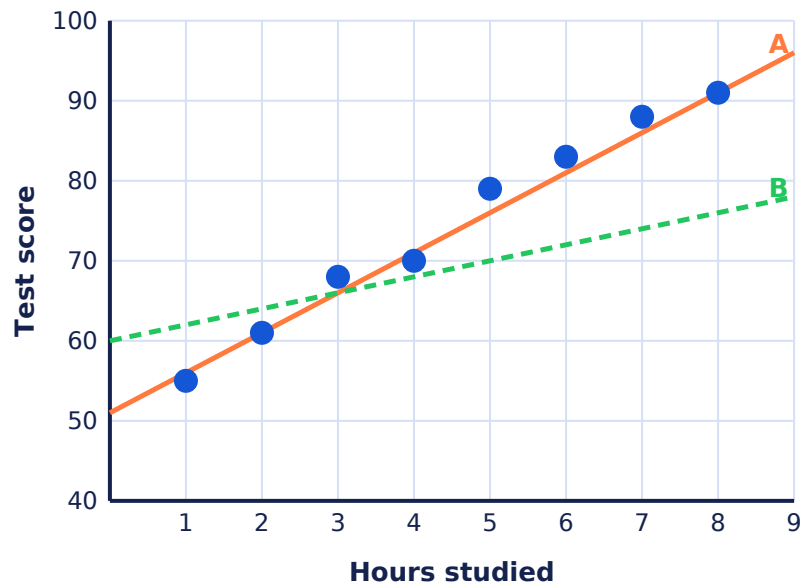
A line of best fit, also called a trend line, is a straight line drawn through a scatter plot so that it follows the overall pattern of the points. About as many points lie above the line as below it, and the line does not need to touch any point.



The line $y = 5x + 51$ follows the trend of the points.

How do you judge whether a line fits well?

A good line stays close to the points all the way across the graph, with points on both sides of it. A poor line runs through the middle on one end but drifts away on the other, so the points lie mostly on one side.



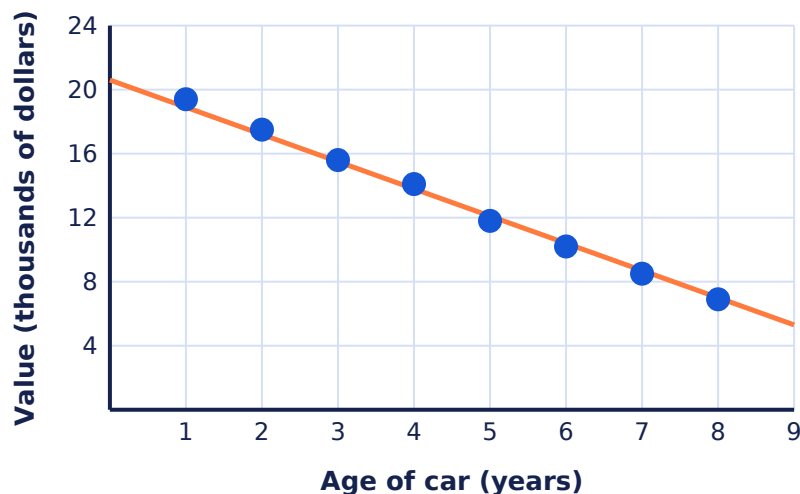
Line A (solid) fits well; line B (dashed) is too flat.

How do you write the equation of the model?

Pick two points on the line, not necessarily data points, and find the slope $m = (y_2 - y_1) \div (x_2 - x_1)$. Read the y-intercept b where the line meets the vertical axis. Then write the model as $y = mx + b$.

Meaning of slope and intercept

- The slope is the change in y for each increase of 1 in x .
- The y-intercept is the predicted value of y when $x = 0$.
- A positive slope means y grows as x grows; a negative slope means it shrinks.



The line $y = -1.7x + 20.6$ has a negative slope.

How do you make predictions?

Method

1. Write the equation of the line.

2. Replace x with the given value and compute y .
3. To find x for a target y , set the equation equal to that value and solve.
4. Check that the answer makes sense in context.

Worked example

For the study model $y = 5x + 51$, the slope 5 says each extra hour adds about 5 points. The intercept 51 is the predicted score with no study. For 6.5 hours, $y = 5 \cdot 6.5 + 51 = 83.5$ points.

Common mistake

Predicting far outside the data can give nonsense. At 20 hours the study model predicts 151 points, which is impossible on a 100-point test. Stay near the range of the data.

Mathbot's tip

Say the slope as a sentence: "for each 1 more hour, the score goes up about 5 points." If the sentence sounds odd, recheck the units.

Key points to remember

- A line of best fit follows the trend with points on both sides; it need not touch any point.
- To write the model, find the slope from two points on the line and read the y -intercept.
- Slope = change in y per 1 unit of x ; intercept = predicted y at $x = 0$.
- Substitute x to predict y , or solve for x to reach a target.
- Predictions far outside the data are unreliable.

Fitting Lines to Data and Making Predictions: frequently asked questions

Does the line of best fit have to pass through the data points?

No. It only has to follow the trend, with about as many points above it as below it. It may pass through several points, one point, or none at all.

What does the slope of a trend line mean?

The slope is the predicted change in y for every increase of 1 in x , in the units of the problem. For $y = 5x + 51$, each extra hour of study predicts about 5 more points.

What does the y -intercept mean?

It is the predicted value of y when x is 0. It does not always make sense in context, so check whether $x = 0$ is a reasonable value before interpreting it.

What is the difference between interpolation and extrapolation?

Interpolation predicts inside the range of the data and is usually reliable. Extrapolation predicts outside that range and can be wrong, because the trend may change.

How can I tell which of two lines fits better?

Compare how far the points sit from each line. The better line runs closer to most points and leaves points on both sides, while the poorer line misses the trend at one end of the graph.