

Grade 8 · Lines of Best Fit and Linear Models · Lesson

Common Core 8.SP.A.2, 8.SP.A.3 · mathovore.fr

A line of best fit is a straight line through a scatter plot that follows the trend of the data. Its equation, $y = mx + b$, is a linear model: the slope m is the change in y per 1-unit rise in x , and b is the predicted y at $x = 0$. Use it to predict.

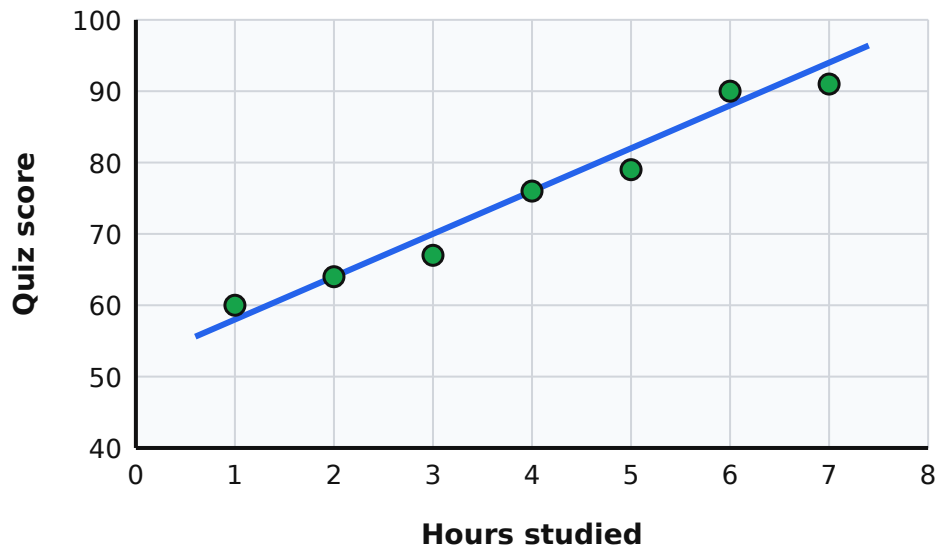
Grade 8 Common Core: 8.SP.A.2 · 8.SP.A.3

Key facts:

- Standards 8.SP.A.2 and 8.SP.A.3: fit a line informally, judge the fit, and interpret the equation of a linear model.
- A good line has points close to it on both sides; a curve or wide scatter means a line is a poor model.
- Slope = change in y per 1 unit of x ; y -intercept = predicted y when $x = 0$.
- Find the equation from two points on the line, not necessarily data points.
- Predictions inside the data range are more reliable than predictions far outside it.

What is a line of best fit?

A **scatter plot** shows pairs of data as points. When the points follow a roughly straight band, you can draw a **line of best fit** through the middle of them. The line is a **linear model**: its equation $y = mx + b$ describes the trend and lets you make predictions.



Hours studied and quiz score for seven students, with the line $y = 6x + 52$.

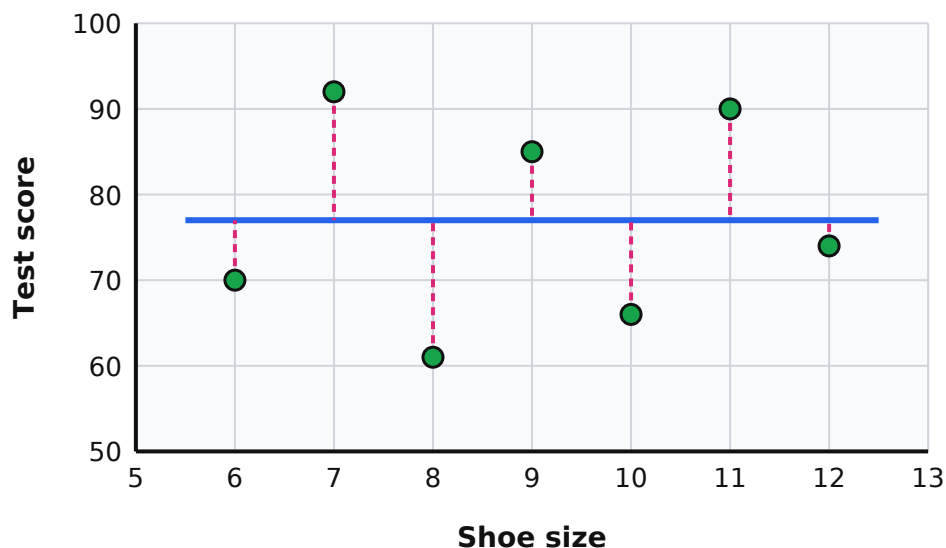
How do you judge how well a line fits?

Look at the vertical gaps between the points and the line. A good line passes close to the points, with some points above it and some below. A poor line leaves points far away, or the points curve in a way a straight line cannot follow.

Vocabulary

Gap (residual): the vertical distance from a data point to the line.

Outlier: a point that sits far from the pattern of the others.

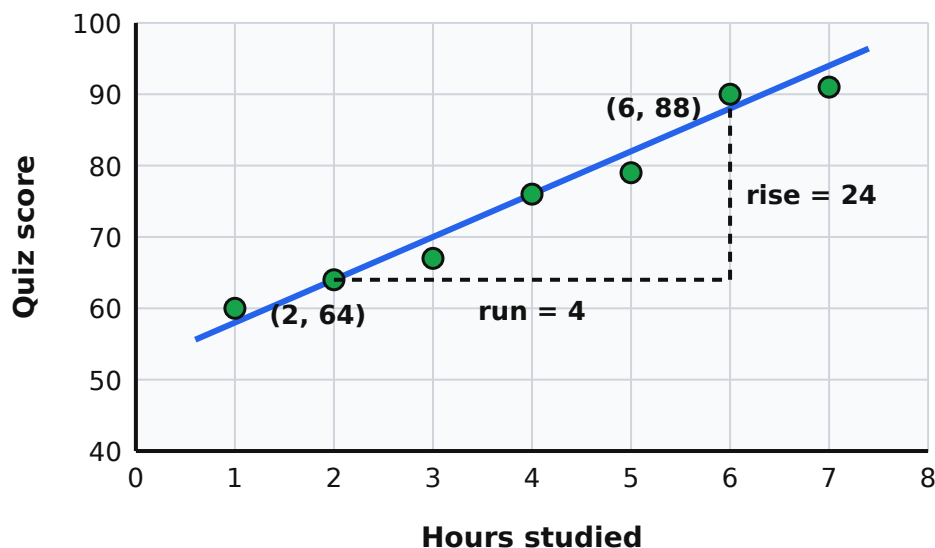


A poor fit: the gaps to the flat line are as large as 16 points, and the points show no trend.

How do you find the equation of a line of best fit?

1. Draw a line close to the points, with about as many points on each side.
2. Pick two points on the line and read their coordinates.
3. Find the slope: $m = \frac{y_2 - y_1}{x_2 - x_1}$.
4. Substitute one point into $y = mx + b$ to find b .

Worked example. The line in the first figure passes through (2,64) and (6,88). Then $m = \frac{88-64}{6-2} = 6$. From $64 = 6(2) + b$ we get $b = 52$, so $y = 6x + 52$.

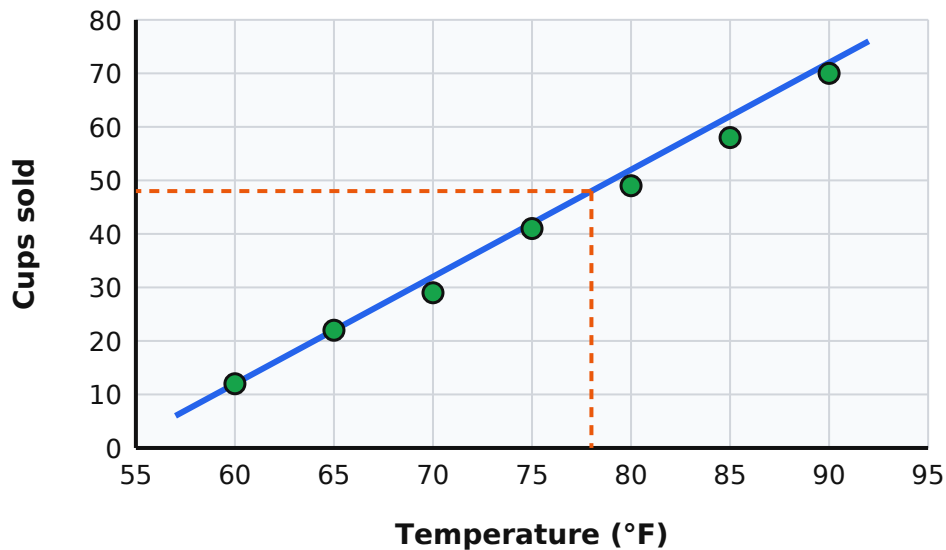


Slope from two points on the line: rise 24 over run 4 gives 6.

What do the slope and y-intercept mean?

The **slope** is the change in y for each 1-unit increase in x . The **y-intercept** is the predicted y when $x = 0$, although that value is not always meaningful. For lemonade sales the model is $y = 2x - 108$, where x is the temperature in $^{\circ}\text{F}$ and y is cups sold. The slope 2 means about 2 more cups for every 1°F rise. The intercept -108 has no real meaning, because a stand would not run at 0°F .

To predict, substitute: at 78°F , $y = 2(78) - 108 = 48$ cups. Predictions inside the range of the data are more reliable than predictions far outside it.



Reading a prediction from the line: at 78°F the model gives 48 cups.

Lines of Best Fit and Linear Models: frequently asked questions

What is a line of best fit?

It is a straight line that follows the overall trend of the points in a scatter plot. It does not have to touch every point. About as many points should lie above it as below it.

How do you tell if a line fits the data well?

Check the vertical gaps between the points and the line. If the gaps are small and the points fall on both sides, the fit is good. If the points curve or are spread widely, the fit is poor.

How do you find the equation of a line of best fit?

Choose two points on the line and compute the slope with rise over run. Then substitute one point into $y = mx + b$ to find the y-intercept.

What does the slope of a linear model mean?

The slope is the predicted change in y for each 1-unit increase in x . In $y = 6x + 52$, each extra hour of study adds about 6 points to the predicted score.

Is the y-intercept always meaningful?

No. It is the predicted value when $x = 0$, which may be outside the data. In the lemonade model $y = 2x - 108$, the intercept -108 is not a real number of cups.