The Line of Best Fit

The line of best fit (least squares regression line) makes the total squared error smallest. Technology finds m and b. Read them in the story, predict, and ask whether the prediction makes sense.

1 THE VOCABULARY
Words we use

Line of best fit — the line closest to all the points: the least squares regression line.

Regression — finding the best-fitting line or curve.

Predicted value — what the model says y should be.

Interpolate — predict inside the data: fairly safe.

Extrapolate — predict outside the data: risky.

- Watch out for these.
- Saying the rat gains exactly 40 grams. Say about.
- Trusting a prediction far outside the data.
- Reading the y-intercept as the slope.
- Forgetting the units on the slope: grams per week.

2 THE METHOD
Use a linear model

1. FIT the tool or a calculator gives m, b, and r.
2. SLOPE the change in y for each 1 of x, in the story's units.
3. Y-INTERCEPT the value when x = 0.
4. PREDICT forward: put x in. Backward: set equal and solve.
5. JUDGE inside the data or outside? Does it make sense?

WHY SQUARED DISTANCES
Misses above the line are positive and misses below are negative. Added plainly they cancel. Squared, every miss counts.

3 THE LAB RAT
W = 100 + 40t

slope 40	gains about 40 grams a week
y-intercept 100	about 100 grams at birth
t = 10: W = 500	forward
100 + 40t = 760, so t = 16.5	backward, the Unit 1 way
t = 260: W = 10,500	extrapolation: not believable

USED CARS
y = −0.14x + 23.2: the price drops about $140 per thousand miles. At 70 thousand miles, about $13,400.

SAY IT OUT LOUD
A model is a family member that fits best. Read its numbers in the story, and ask if the prediction makes sense.