Residuals, Checking the Fit

Residual = actual minus predicted. Use the site's regression tool or a graphing calculator for the fits. Random scatter in the residual plot means a line fits; a pattern means it does not.

Name _____ Date _____

Words we use

say the math word, then the plain English

WORD	PLAIN ENGLISH	WHAT YOU DO
Residual	Actual minus predicted: how far off the model is for one point.	Subtract in that order. Positive is above the line.
Predicted value	What the model says y should be.	Put x into the model.
Residual plot	A graph of the residuals against x, with a flat line at 0.	Use it to check whether a line fits.
Random scatter	No pattern: dots above and below 0 in no order.	A line is a good model.
Pattern	A curve in the residuals, like a U or a hill.	A line is the wrong family; try a curve.
Above, below the line	Positive residual: the model guessed too low. Negative: too high.	Read the sign.

Part A Tuition

Use the model R(t) = 326t + 6440. Fill in the predicted and residual columns.

T	ACTUAL	PREDICTED	RESIDUAL	ABOVE OR BELOW
0	6,500			
2	7,187			
4	7,874			
5	7,985			
8	8,928			
10	9,665			

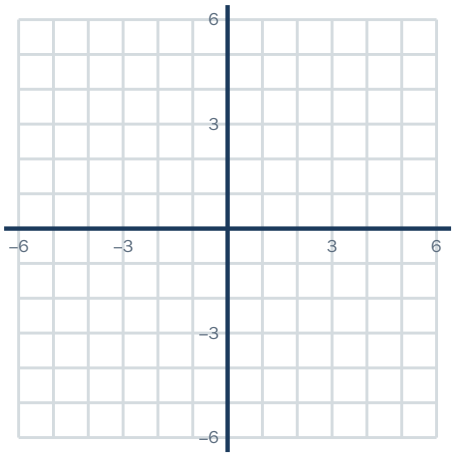
Part B Small models

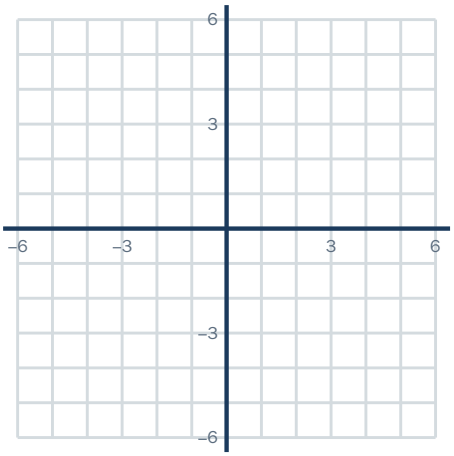
1. Model y = 2x + 3, data point (4, 13). Predicted value? Residual?

2. Model y = 5x + 10, data (1, 17), (2, 19), (3, 26). Find the three residuals.

Part C Read the residual plots

Plot each list of residuals on a grid (x = 1 to 8 across, residual up and down). Then decide.





RESIDUALS, LEFT TO RIGHT	PATTERN OR RANDOM SCATTER	IS A LINE APPROPRIATE?
+3, –2, +1, –4, +2, –1, +3, –2		
+6, +2, –3, –5, –4, –1, +3, +6		

Part D The gnats

Gnats after 0, 1, 2, 3, 4 hours: 12, 20, 35, 60, 80. Fit a line with the regression tool or a calculator.

1. What line do you get? What is r?

2. Write the five residuals and their signs.

3. Is a line the right family? Explain, using the residuals.

Part E Exit ticket

1. A model predicts 50; the actual value is 44. Residual? Above or below?

2. What three things do you check together before trusting a line?

Answer key

Part A. 6,440, +60, above. 7,092, +95, above. 7,744, +130, above. 8,070, –85, below. 9,048, –120, below. 9,700, –35, below.

Part B. 1. Predicted 11; residual +2. 2. +2, –1, +1.

Part C. First: random scatter; a line is appropriate. Second: a U-shaped pattern; a line is not appropriate.

Part D. 1. $y = 17.6x + 6.2$; r is about 0.98. 2. +5.8, –3.8, –6.4, +1.0, +3.4: plus, minus, minus, plus, plus. 3. No: the residuals make a U, so the data curves, even though r is big.

Part E. 1. 44 – 50 = –6; below the line. 2. The scatter plot, r, and the residual plot.