

Statistical Reasoning
Residuals

Name:

A **residual** is the difference between an observed value of the response variable and the value predicted by the regression line.

$$y = 8.07x + 60.44$$

residual = observed y - predicted y

Class Test Scores			
Study Time (in hours)	Actual Test Score	Predicted Test Score	Residual
0.5	63	- 64.48	= -1.48
1	67	- 68.51	= -1.51
1.25	72	- 70.53	= 1.47
2	76	- 76.58	= -0.58
2	80	- 76.58	= 3.42
3	85	- 84.65	= 0.35
3.75	89	- 90.70	= -1.7

- * 1) Using the data above, calculate the regression model. Write the equation below. Round all answers to the nearest hundredth.

$$a = 8.07$$

$$b = 60.44$$

$$y = 8.07x + 60.44$$

- 2) What is the correlation coefficient, r ? What does this tell you?

$$r = 0.979 \quad \text{strong positive correlation}$$

- 3) What is r^2 ? What does this tell you?

$$r^2 = 0.959 \quad 95.9\% \text{ of the data is explained by the regression equation}$$

- 4) Using your regression model from #1, calculate the predicted test score for each student based on their study time.

- 5) Calculate the residuals.

- 6) What do the residuals add up to?

$$-0.03$$

Residuals

- the closer the sum is to 0, the better
- residual plots should be random and have no clear pattern

