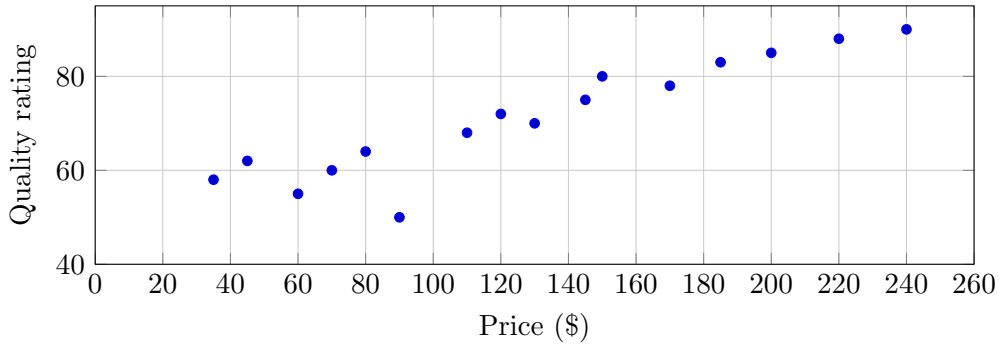
Unit 2: Extra Practice

Mr. Merrick · October 21, 2025

Problem 1. Describing a relationship

A study recorded the price (in dollars) and expert quality rating (0–100) for $n = 16$ Bluetooth speakers. A scatterplot (below) shows the data.



(a) Describe the direction, form, and strength of the association, and identify any notable features.

(b) Based solely on the scatter plot, would a least-squares regression of rating on price be reasonable? Justify using the graph.

(c) In context, explain what a point near $(\$90, 50)$ suggests to a shopper.

Problem 2. Interpreting slope and intercept

The least-squares line for predicting rating y from price x (in dollars) for a different set of speakers is

$$\widehat{y} = 41.3 + 0.21x, \qquad s = 5.6, \quad r^2 = 68\%.$$

These data came from speakers priced between about \$30 and \$250.

(a) Interpret the slope and the intercept in context.

(b) Estimate the rating for a \$150 speaker and interpret $s = 5.6$.

(c) Is a \$500 prediction advisable? Explain.

Problem 3. From output to equation

A computer regresses weekly study hours y on weekly work hours x for $n = 12$ students (working between 5 and 25 hours per week) and reports:

Predictor	Coef	SE Coef	t	p
Constant	18.2	2.9	6.28	< 0.001
Work	-0.41	0.15	-2.73	0.021
$S = 3.7$	$R^2 = 42\%$		$R^2_{\text{adj}} = 36\%$	

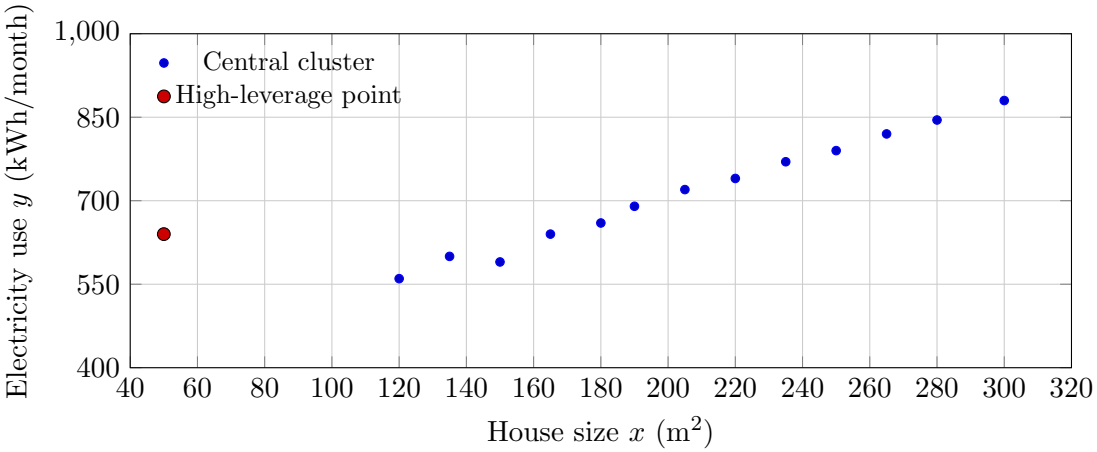
(a) Write the least-squares equation and interpret the slope.

(b) If a student works 20 hours, what is the predicted study time? Comment on practical reasonableness.

(c) Compute and interpret the residual for a student who worked $x = 10$ hours and studied $y = 12$ hours. Then sketch or describe the residual plot pattern you would expect.

Problem 4. Influential point vs. outlier

The scatterplot shows monthly electricity use (y, in kWh) versus house size (x, in m^2). Most houses fall between 100–300 m^2.



(a) Explain why the leftmost point is likely *influential* for the regression line.

(b) If that point were removed, what would you expect to happen to the slope and to R^2? Explain your reasoning using the figure.

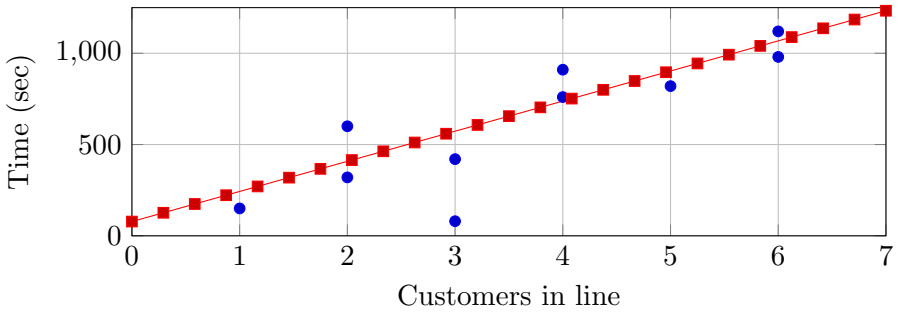
Problem 5. AP-style free response

A manager samples $n = 10$ checkout lines. Let x be the number of customers ahead of a shopper and y the total checkout time (sec). The regression output is:

Predictor	Coef	SE Coef	t	p
Constant	78	96	0.81	0.44
Customers in line	165	28	5.89	< 0.001
$S = 190$	$R^2 = 78\%$		$R^2_{\text{adj}} = 75\%$	

(a) Write the least-squares equation. Interpret the slope in context.

(b) Circle on the sketch the most likely outlier and explain why:



(c) Interpret $R^2 = 78\%$.

Problem 6. Using residual to recover an observed value

For wolves, the fitted line for weight (kg) on length (m) is $\widehat{y} = -16.46 + 35.02x$. A wolf of length 1.40 m has residual -9.67 kg.

(a) What is this wolf's actual weight?

(b) Interpret the residual.

Problem 7. Multiple parts, mixed skills

Biologists measured mass y (g) and length x (mm) for 11 frogs and obtained the regression line

$$\widehat{y} = -546 + 6.086x, \qquad r^2 \approx 0.819.$$

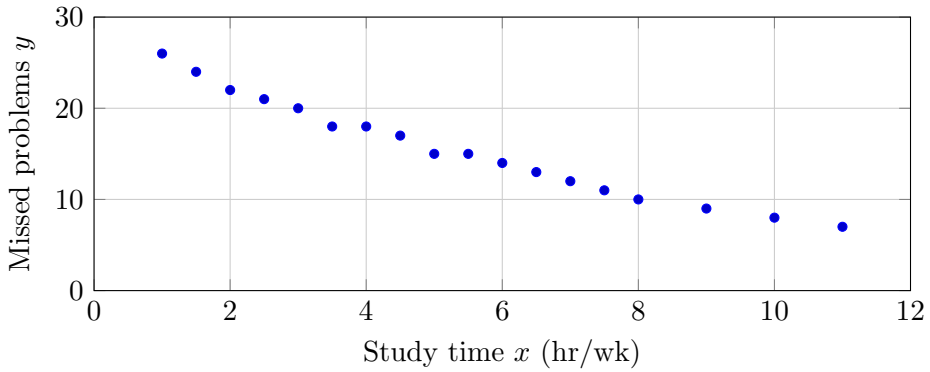
(a) Interpret the slope in context.

(b) Interpret r^2 in context.

(c) On a residual plot, which frog would have the larger magnitude residual: one with $(x = 130, y = 220)$ or one with $(x = 170, y = 530)$? Show work.

Problem 8. Correlation: sign, magnitude, and meaning

The plot shows a relationship between study time (x, hours/week) and number of missed homework problems (y) for $n = 18$ students.



(a) Based on the plot, state the *direction*, *form*, and *strength* of the association and give a rough estimate of the *sign* of r.

(b) A computer output (not shown) reports $R^2 = 0.92$ and a *negative* slope. Compute r and interpret R^2 in context.

Problem 9. Changing units: what changes, what doesn't

For $n = 25$ headphones, the least-squares line for predicting quality rating y (0–100 points) from price x (US dollars) is

$$\hat{y} = 12.0 + 0.45x, \qquad R^2 = 64\%.$$

Answer the following about unit changes.

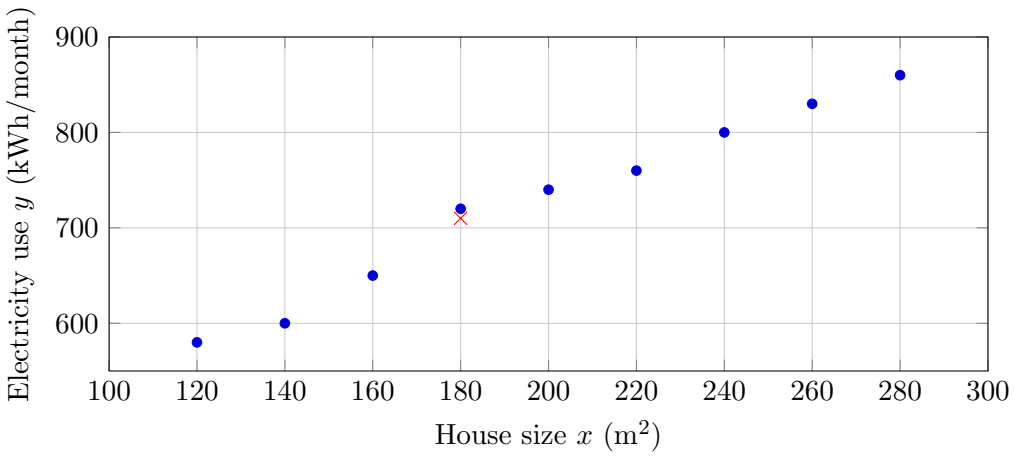
(a) If price is recorded in *cents* ($x_c = 100x$), write the new regression equation $\hat{y}$ in terms of x_c. What happens to r and R^2?

(b) Suppose ratings are converted to a *5-star* scale with $y_\star = y/20$. Write the regression of $y_\star$ on dollars x. What happens to r and R^2?

(c) Briefly explain why r and R^2 are invariant to these linear unit changes.

Problem 10. "Line through the means" & residual properties

For the homes below (electricity vs. size), the sample means are $\bar{x} = 180$ m^2 and $\bar{y} = 710$ kWh. The point $(\bar{x}, \bar{y})$ is marked with a red cross.

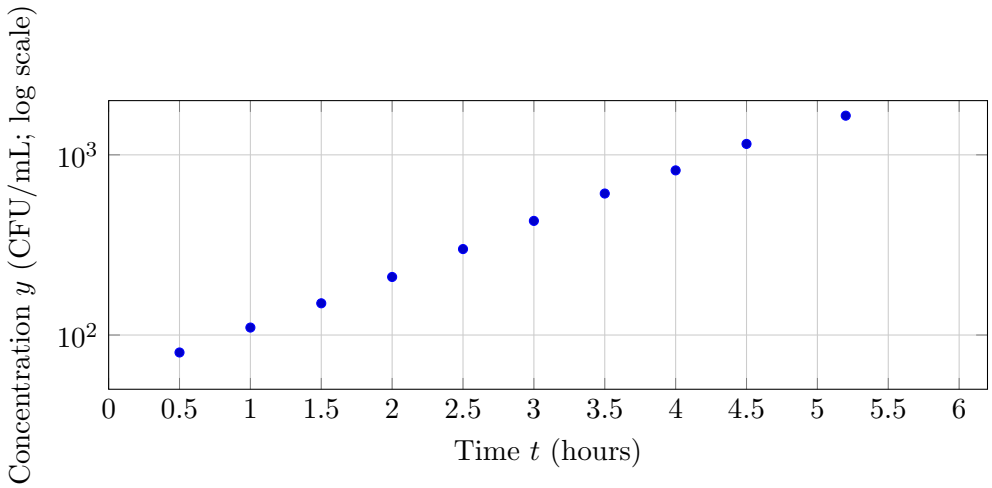


(a) Must the least-squares regression line pass through the cross at $(\bar{x}, \bar{y})$? Explain.

(b) For any fitted least-squares line on this dataset, what is the sum and the mean of the residuals? Briefly justify.

Problem 11. Log re-expression and multiplicative interpretation

The plot shows bacterial concentration y (CFU/mL) versus incubation time t (hours) for $n = 10$ trials.



A linear model was fit to $\log_{10} y$ versus t, giving

$$\log_{10} \hat{y} = 2.10 + 0.18\, t, \qquad R^2 = 94\%.$$

(a) Interpret the slope 0.18 in *multiplicative* terms for y.

(b) Predict the concentration at $t = 3$ hours on the original scale and comment on model fit using R^2.

(c) Briefly explain why the log transformation was appropriate based on the plot.