Exercise walk-through

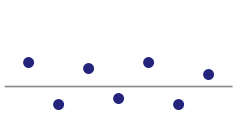
2.7
Residuals

AP Statistics

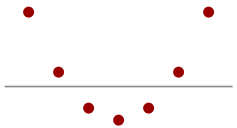
Residuals and fit

A **residual** is the miss: $y - \hat{y}$.

- random **residual plot** $\Rightarrow$ a line fits
- a curve $\Rightarrow$ line is wrong
- s: the standard deviation of the residuals – the typical prediction error, in y's own units
- r^2: the **coefficient of determination**



random: good

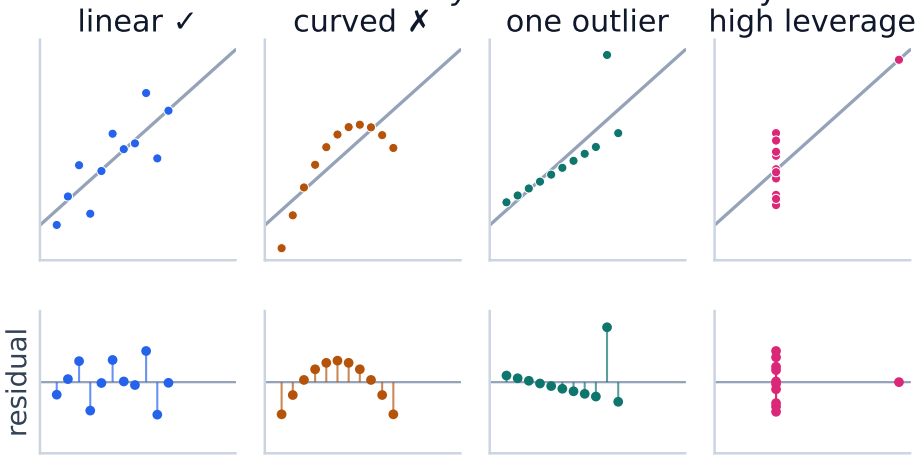


curved: bad

> r^2 is the fraction of y's variation the line explains – report it in context; $b = r\dfrac{s_y}{s_x}$.

A caution about r and the line

every panel has the same $r = 0.82$ and the same line $\hat{y} = 3.0 + 0.5x$ — only the residual plot tells them apart



Anscombe: same stats, different plots — always graph data

All four datasets share $r = 0.82$ and the line $\hat{y} = 3.0 + 0.5x$, yet only the first is genuinely linear – always plot the data before you trust a summary.

Exercises 2.1–2.3

2.1 Write the formula for a residual. [1]

2.2 An observed value is 18 and the model predicts 15. Find the residual. [2]

2.3 State what a residual plot with a clear pattern suggests. [1]

Exercises 2.1–2.3

2.1 Write the formula for a residual. [1]

2.1

- residual $= y - \hat{y}$

2.2 An observed value is 18 and the model predicts 15. Find the residual. [2]

2.2

- $18 - 15 = 3$

2.3 State what a residual plot with a clear pattern suggests. [1]

2.3

- a linear model is not appropriate for the data

Exercises 3.1·3.2

3.1 A residual equals [1]

- **A** $\hat{y} - y$
- **B** $y - \hat{y}$
- **C** $y + \hat{y}$
- **D** $y \times \hat{y}$

3.2 A residual plot showing random scatter suggests the linear model is [1]

- **A** inappropriate
- **B** appropriate
- **C** curved
- **D** wrong

Exercises 3.1·3.2

3.1 A residual equals [1]

- **A** $\hat{y} - y$
- **B** $y - \hat{y}$
- **C** $y + \hat{y}$
- **D** $y \times \hat{y}$

3.1 B

- $y - \hat{y}$.

3.2 A residual plot showing random scatter suggests the linear model is [1]

- **A** inappropriate
- **B** appropriate
- **C** curved
- **D** wrong

3.2 B

- appropriate.

Exercise 3.3 (a)

3.3 A prediction is $\hat{y} = 40$, but the actual value is $y = 46$.

(a) Find the residual. [2]

Exercise 3.3 (a)

3.3 A prediction is $\hat{y} = 40$, but the actual value is $y = 46$.

(a) Find the residual. [2]

(a)

- $46 - 40 = 6$

Exercise 3.3 (b)

3.3 A prediction is $\hat{y} = 40$, but the actual value is $y = 46$.

(b) State whether the point is above or below the line. [1]

Exercise 3.3 (b)

3.3 A prediction is $\hat{y} = 40$, but the actual value is $y = 46$.

(b) State whether the point is above or below the line. [1]

(b)

- above the line (positive residual)