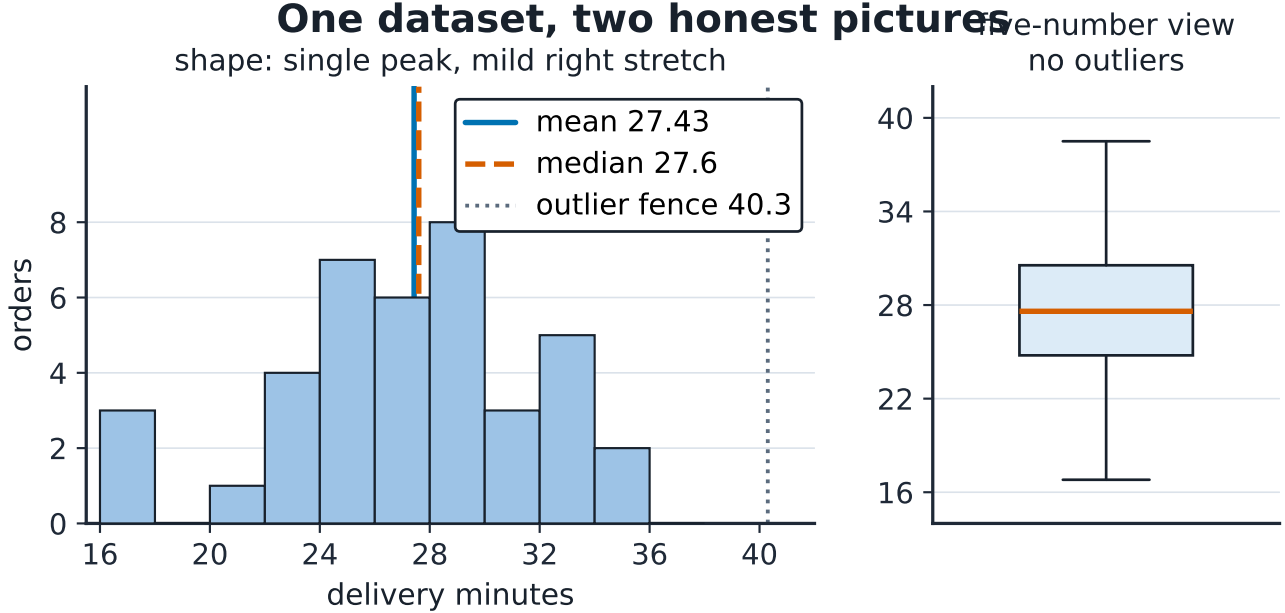
Exercise walk-through

5.7
Decision limitations

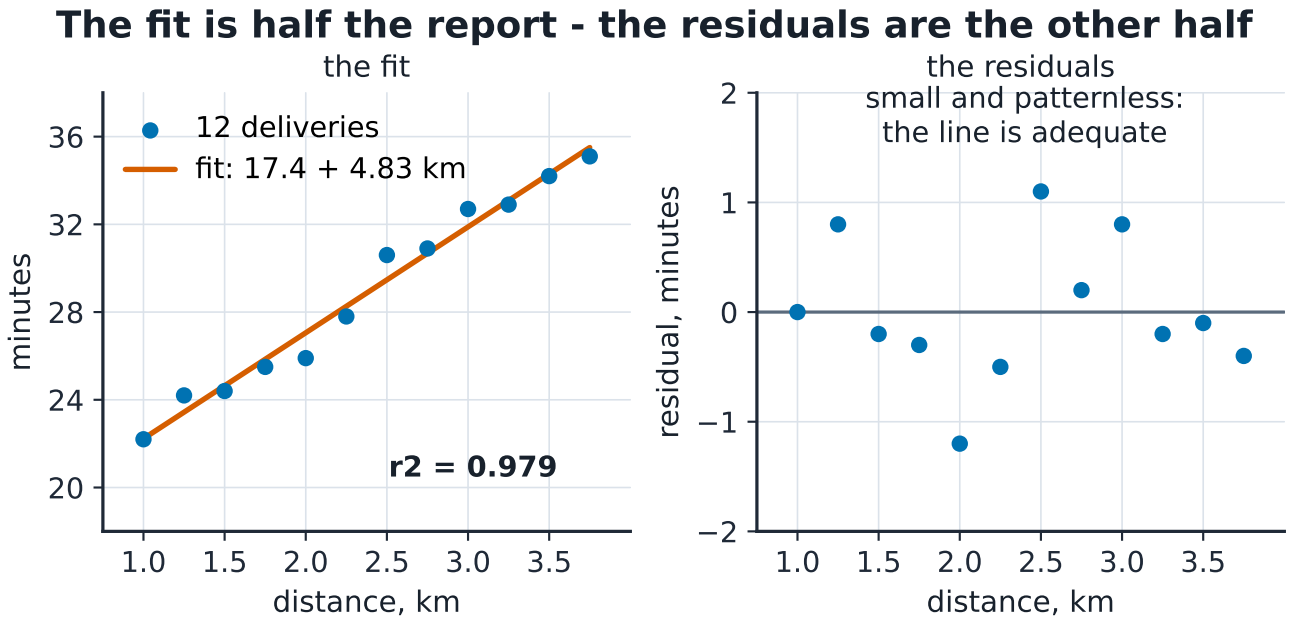
University Business

Descriptive measures and visualisation: the four questions



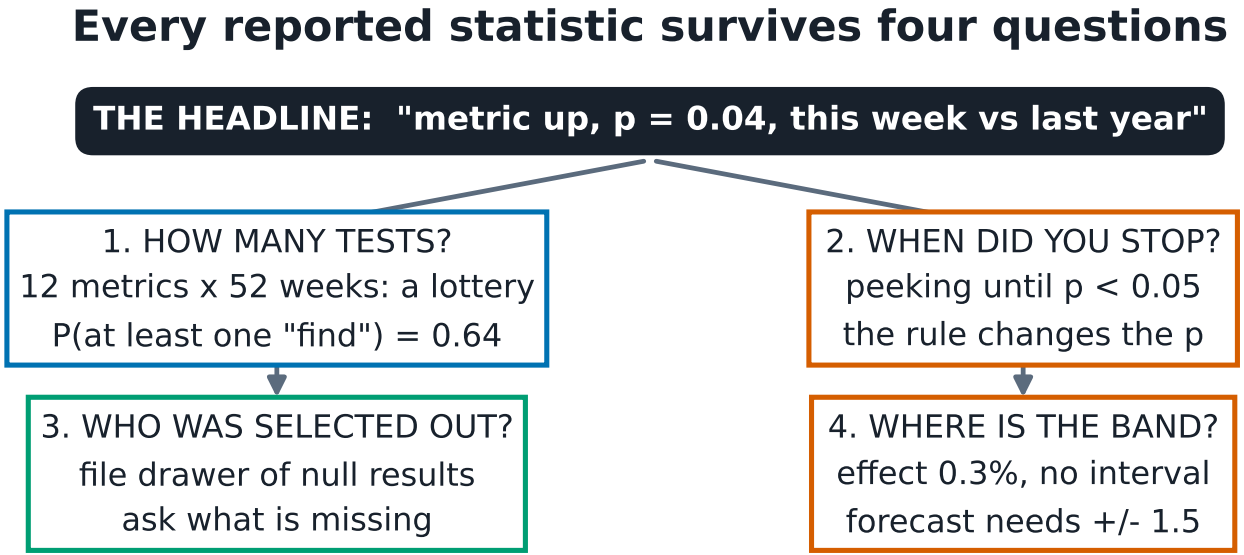
Mean 27.43, **median** 27.6, fence at 40.3: centre, spread and outliers.

Regression: the line and its police



Time = 17.4 + 4.83 km, r2 = 0.979, with patternless residuals.

Limitations: the four gates



verdict: not evidence of change - evidence of a dashboard making headlines

Challenge a headline through four gates: tests, stopping, selection, error.

Exercises 2.1·2.2

2.1 A chain reports “store 7 improved, p = 0.03” – one of its 14 stores, chosen after looking. Compute the chance that at least one of 14 independent tests clears 0.05, and name TWO gates the claim fails. [4]

2.2 An A/B test was checked daily and stopped on day 9, the first morning $p < 0.05$. In two lines, what does the stopping rule do to the nominal error rate, and what rule would have kept it honest? [2]

Exercises 2.1·2.2

2.1 A chain reports "store 7 improved, p = 0.03" – one of its 14 stores, chosen after looking. Compute the chance that at least one of 14 independent tests clears 0.05, and name TWO gates the claim fails. [4]

2.1

- $P(\geq 1) = 1 - 0.95^{14}$.
- $1 - 0.488 = \mathbf{0.51}$ – a coin flip.
- Gate 1: 14 tests (multiple comparisons).
- Gate 3: the store was selected after looking (selection on the result).

2.2 An A/B test was checked daily and stopped on day 9, the first morning $p < 0.05$. In two lines, what does the stopping rule do to the nominal error rate, and what rule would have kept it honest? [2]

2.2

- Optional stopping inflates the false-positive rate well above the nominal 5% – each peek is another ticket.
- An honest rule fixes n (or a sequential method) in advance and stops only as declared.

Exercises 2.3·2.4

2.3 “Next month’s revenue: £84,300.” With sd £9,000 across months (n ≈ 30), construct the band the forecast should wear (95%, t ≈ 2), and say in one line how the owner should use it. [3]

2.4 For each report defect – (a) 12 metrics, one flagged; (b) stopped when p dipped; (c) eleven flat stores omitted; (d) £84,300 with no band – name its gate. [4]

Exercises 2.3·2.4

2.3 "Next month's revenue: £84,300." With sd £9,000 across months (n ≈ 30), construct the band the forecast should wear (95%, t ≈ 2), and say in one line how the owner should use it. [3]

2.3

- $SE = 9{,}000/\sqrt{30} = 1{,}643$.
- Band: $84{,}300 \pm 2 \times 1{,}643 = [\mathbf{81{,}000},\ \mathbf{87{,}600}]$.
- Use: plan cash for the band (stress the lower end), and revisit calibration – did past 95% bands hold 95% of the time?

2.4 For each report defect – (a) 12 metrics, one flagged; (b) stopped when p dipped; (c) eleven flat stores omitted; (d) £84,300 with no band – name its gate. [4]

2.4

- (a) How many tests?
- (b) When did you stop?
- (c) Who was selected out?
- (d) Where is the band?