

Data Management Review

Answer key Final answers only. Try first, then check.

CORE SKILLS

- K1** a) (i) categorical; (ii) discrete; (iii) continuous
b) Walk 35%, Bus 30%, Car 25%, Bike 10%
c) Walk 126°, Bus 108°, Car 90°, Bike 36°
- K2** a) 2.12 books
b) 2 books
c) 1 book
- K3** a) Min 12, $Q_1 = 15$, median 20, $Q_3 = 24$, max 45
b) $IQR = 9$ minutes
c) 1.5 and 37.5; 45 is an outlier (nothing is below 1.5)
- K4** a) Negative, strong
b) $y = -0.5x + 10$
c) 7 hours
- K5** a) 91.5 cm; interpolation
b) 229 cm; extrapolation
c) $t = 5$ days; extrapolation

REAL WORLD

- A1** a) 73.2%
b) 82.5%
c) 120%: impossible, since no mark is above 100%. The Period 1 mean can reach at most about 71.3%.
- A2** a) Slope -1.8 : the price drops by about \$1800 for each year of age. Intercept 24: a new car (age 0) costs about \$24,000.
b) \$15,900; interpolation
c) $P = -3$, a price of $-\$3000$: impossible. 15 years is far outside the data (1 to 8 years), so the model should not be used there.
- A3** a) Route 7: median 24 min, $IQR = 6$ min. Route 12: median 26 min, $IQR = 4$ min.
b) About 5 trips (one quarter of the data)
c) Route 12: all 20 days. Route 7: at least 15 and at most 19 days; the box plot cannot give the exact number.

THINK IT THROUGH

- T1** a) Pool A: 420; Pool B: 480; about 14.3% greater
b) The vertical axis starts at 400, not 0. The bars show only the 20 and 80 visitors above 400, and $80 = 4 \times 20$.
c) Start the vertical axis at 0; the heights are then in the ratio $480 : 420 = 8 : 7 \approx 1.14$.
- T2** a) e.g. (0, 2), (10, 3), (20, 4), (30, 5)
b) Stronger correlation: A. Steeper line: B.
c) Strength describes how closely the points lie to a line, not how steep the line is. (a) and (b) are counterexamples: a shallow line can fit perfectly, and a steep one can fit poorly.

BOSS LEVEL

- Boss** a) 96
b) 36 (the list 1, 2, 10, 11, 36)
c) 17 (the list 8, 9, 10, 16, 17)

