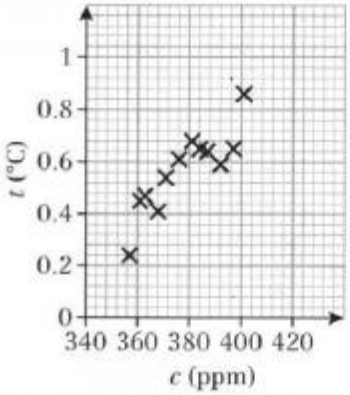
# Answers - Correlation and Regression

**1** The data shows that the number of serious road accidents in a week strongly correlates with the number of fast food restaurants. However, it does not show whether the relationship is causal. Both variables could correlate with a third variable, e.g. the number of roads coming into a town.

**2 a**




**b** Strong positive correlation.

**c** As mean $CO_2$ concentration in the atmosphere increases, mean global temperatures also increase.

**3 a** Strong positive correlation.

**b** If the number of items increases by 1, the time taken increases by approximately 2.64 minutes.

**4** (1) 3500 is outside the range of the data (extrapolation).
(2) The regression equation should only be used to predict a value of GNP ($y$) given energy consumption ($x$).

**5 a** Mean + 2SD = $15.2 + 2 \times 11.4 = 38$; $50 > 38$

**b** The outlier should be omitted as it is very unlikely that the average temperature was 50 °C.

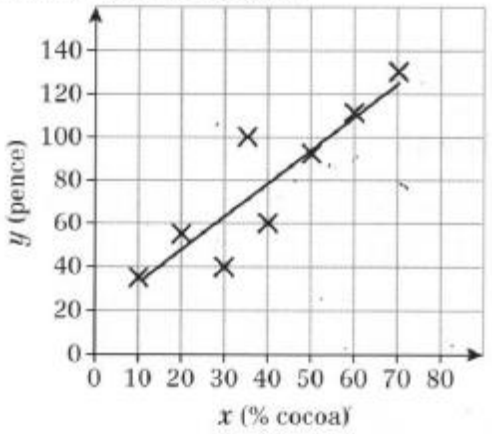
**c** If the temperature increases by approximately 1 °C, the number of pairs of gloves sold each month decreases by 0.18.

**6 a** 44 is the length in centimetres of the spring with no mass attached. If a mass of 1 g is attached, the spring would increase in length by approximately 0.2 cm.

**b** **i** Outside the range of the data (extrapolation)

**ii** The regression equation should only be used to predict a value of $s$ given $m$

**7 a & b**




**c** Brand D is overpriced, since it is a long way above the line.

**d** The regression equation should be used to predict a value for $y$ given $x$ so the student's method is valid.