Regression, correlation and hypothesis testing Mixed exercise 1

1 a

Base area, x (cm^2)	1.1	1.3	1.9	2.2	2.5	3.7
Time, t (hours)	0.7	0.9	1.5	1.8	2.2	3.8
log x	0.0414	0.114	0.279	0.342	0.398	0.568
log t	−0.155	−0.0458	0.176	0.255	0.342	0.580

Calculating the PMCC for log x and log t: $r = 0.9998$.

b r is close to 1, so a graph of log t against log x shows a straight line, suggesting that the relationship is in the form $t = ax^n$.

c $\log t = -0.215 + 1.38 \log x$

$\Rightarrow t = 10^{-0.215+1.38\log x} = 10^{-0.215} \times 10^{1.38\log x}$

$\Rightarrow t = 10^{-0.215} \times 10^{\log x^{1.38}} = 10^{-0.215} \times x^{1.38}$

Therefore $a = 10^{-0.215} \approx 0.617$ (3 s.f.) and $n = 1.38$

2 a

Temperature, t (°C)	38	51	72	83	89	94
Dry residue, d (grams)	4.3	11.7	58.6	136.7	217.0	318.8
y = log d	0.633	1.07	1.77	2.14	2.34	2.50

$y = -0.635 + 0.0334x$

$\Rightarrow \log d = -0.635 + 0.0334t$

$\Rightarrow d = 10^{(-0.635+0.0334)^t} = 10^{0.635} \times 10^{0.0334t}$

$\Rightarrow d = 10^{-0.635} \times \left(10^{0.0334}\right)^t$

Therefore $a = 10^{-0.635} = 0.232$ (3 s.f.) and $b = 10^{0.0334} = 1.08$ (3 s.f.)

b 151 °C is outside the range of the data (extrapolation).

3 As a person's age increases their score on a memory test decreases.

4 a Each cow should be given 7 units. The yield levels off at this point. This can be seen even more clearly by drawing a scatter plot.

b $r = 0.952$ (3 s.f.)

c It would be less than 0.952. The yield of the last three cows is no greater than that of the seventh cow.

5 a $r = -0.972$ (to 3 s.f.)

b There is strong negative correlation. As c increases, f decreases.

6 a $r = 0.340$ (3 s.f.)

b $H_0: \rho = 0$
$H_1: \rho \neq 0$

Sample size = 10

Significance level in each tail = 0.025

From the table, critical values of r for a 2.5% significance level with a sample size of 10 are $r = \pm 0.6319$

So the critical region is $r < -0.6319$ and $r > 0.6319$

$0.340 < 0.6319$ so do not reject H_0.

There is not sufficient evidence, at the 5% level of significance, of correlation between age and salary. This means that an older person in this profession does not necessarily earn more than a younger person.

7 a $r = 0.937$ (3 s.f.)

b H_0: $\rho = 0$, H_1: $\rho \neq 0$, critical value = ± 0.6319. Reject H_0. There is evidence that there is a correlation between the age of a machine and its maintenance costs.

8

$\left.\begin{matrix} H_0: \rho = 0 \\ H_1: \rho < 0 \end{matrix}\right\}$ 1-tail $\alpha = 0.05$

Test statistic = –0.975

$n = 9$, critical value = –0.5822

Lower tail test, t.s. < c.v. since $-0.975 < -0.5822$ reject H_0.

Conclude there is evidence of negative correlation. There is evidence that the greater the height above sea level, the lower the temperature at 7.00 a.m. is likely to be.

9

$\left.\begin{matrix} H_0: \rho = 0 \\ H_1: \rho > 0 \end{matrix}\right\}$ 1-tail $\alpha = 0.05$

Test statistic = $r = 0.972$

$n = 9$, critical value = 0.5822

Upper tail test, t.s. > c.v. since $0.972 > 0.5822$ so reject H_0.

Conclude there is evidence of a positive association between age and weight. This means the older a baby is, the heavier it is likely to be.

10 a $r = 0.940$ (to 3 s.f.)

b H_0: $\rho = 0$, H_1: $\rho > 0$, critical value 0.7293. Reject H_0. There is evidence that sunshine hours and ice cream sales are positively correlated.

11 $r = 0.843$ (3 s.f.), H_0: $\rho = 0$, H_1: $\rho > 0$, critical value 0.8054. Reject H_0. There is evidence that mean windspeed and daily maximum gust are positively correlated.

12 $r = -0.793$ (3 s.f.), H_0: $\rho = 0$, H_1: $\rho < 0$, critical value -0.7545. Reject H_0. There is evidence that temperature and pressure are negatively correlated.