**1.** Marc took a random sample of 16 students from a school and for each student recorded

- the number of letters, $x$, in their last name
- the number of letters, $y$, in their first name

His results are shown in the scatter diagram on the next page.

(a) Describe the correlation between $x$ and $y$.

**(1)**

Marc suggests that parents with long last names tend to give their children shorter first names.

(b) Using the scatter diagram comment on Marc's suggestion, giving a reason for your answer.

**(1)**

The results from Marc's random sample of 16 observations are given in the table below.

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $x$ | 3 | 6 | 8 | 7 | 5 | 3 | 11 | 3 | 4 | 5 | 4 | 9 | 7 | 10 | 6 | 6 |
| $y$ | 7 | 7 | 4 | 4 | 6 | 8 | 5 | 5 | 8 | 4 | 7 | 4 | 5 | 5 | 6 | 3 |

(c) Use your calculator to find the product moment correlation coefficient between $x$ and $y$ for these data.

**(1)**

(d) Test whether or not there is evidence of a negative correlation between the number of letters in the last name and the number of letters in the first name.

You should

- state your hypotheses clearly
- use a 5% level of significance

**(3)**

a) weak negative (1) (just "negative" is also accepted)

b) mark's suggestion is compatible · The graph shows negative correlation. (1)

(c) $r = -0.54458266\ldots$

$= -0.545$ (3 s.f.) (1)

(d) $H_0 : \rho = 0$, $H_1 : \rho < 0$ (1)

$r = -0.54458\ldots$

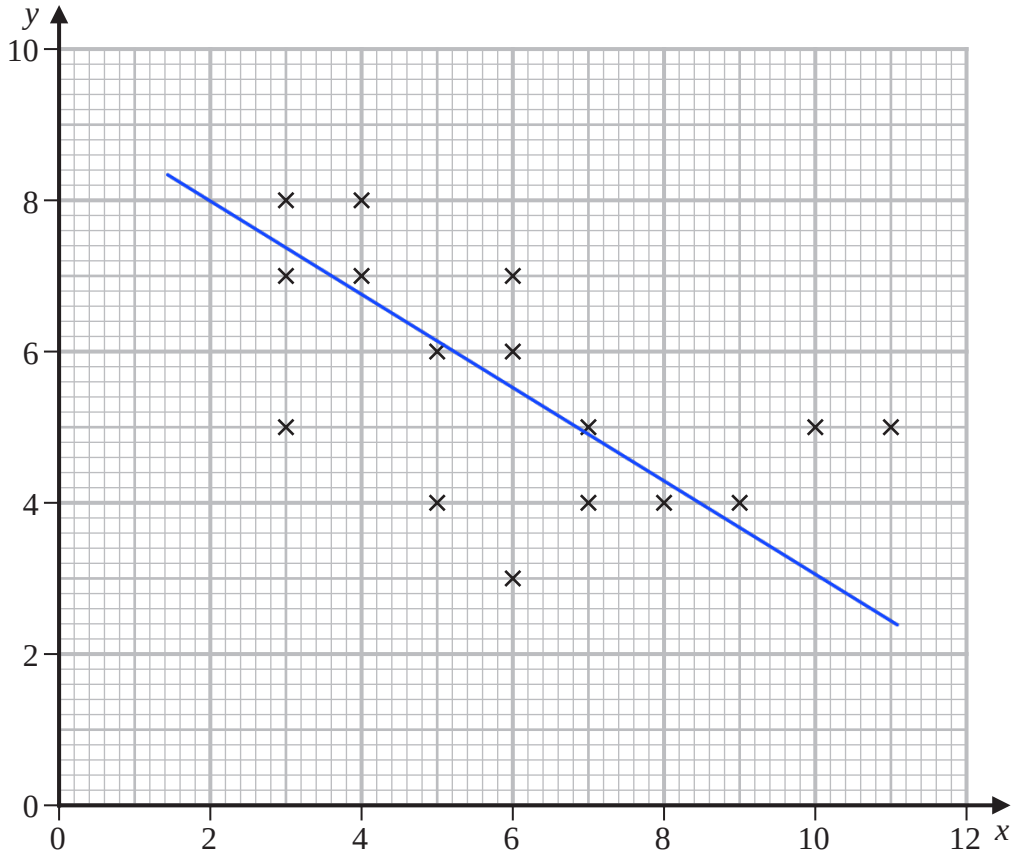
5%, 1-tail → $r_c = -0.4259$ (1)

critical value

$-0.5446 < -0.4259$

(significant)

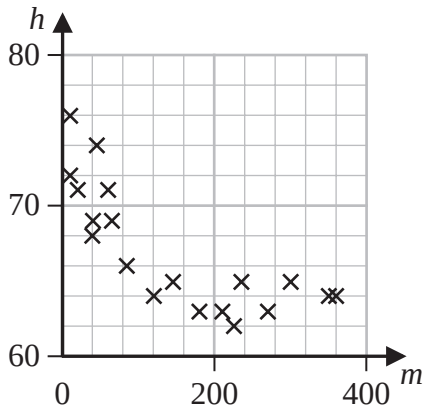
∴ reject $H_0$. There is evidence of negative correlation between number of letters in student's last name and first name. (1)


y
10
8
6
4
2
0
0
2
4
6
8
10
12
x

**2.** Anna is investigating the relationship between exercise and resting heart rate. She takes a random sample of 19 people in her year at school and records for each person

- their resting heart rate, $h$ beats per minute
- the number of minutes, $m$, spent exercising each week

Her results are shown on the scatter diagram.




(a) Interpret the nature of the relationship between $h$ and $m$

**(1)**

Anna codes the data using the formulae

$$x = \log_{10} m$$
$$y = \log_{10} h$$

The product moment correlation coefficient between $x$ and $y$ is $-0.897$

(b) Test whether or not there is significant evidence of a negative correlation between $x$ and $y$
You should

- state your hypotheses clearly
- use a 5% level of significance
- state the critical value used

**(3)**

The equation of the line of best fit of $y$ on $x$ is

$$y = -0.05x + 1.92$$

(c) Use the equation of the line of best fit of $y$ on $x$ to find a model for $h$ on $m$ in the form

$$h = am^k$$

where $a$ and $k$ are constants to be found.

**(5)**

a) As the number of minutes of exercise increases, the resting heart rate decreases (1)

b) $H_0: \rho = 0$ $H_1: \rho < 0$ (1)

sample size = 19
significance value = 0.05

∴ critical value is -0.3887 (1)

$-0.3887 > -0.897$

So there is sufficient evidence to suggest there is a negative correlation between h and m. (1)

c) $y = -0.05x + 1.92$

$\log_{10} h = -0.05 \log_{10} m + 1.92$ (1)

$h = 10^{-0.05 \log_{10} m + 1.92}$ (1)

$h = 10^{\log_{10} m^{-0.05}} \times 10^{1.92}$ (1)

$h = m^{-0.05} \times 10^{1.92}$ (1)

$h = 83.17\, m^{-0.05}$ (1)