1 A football coach measured the heights and weights of 12 players, The data is shown below.

Height (cm)	188	194	178	175	185	175	188	193	180	190	181	169
Weight (kg)	70	100	83	69	77	58	90	86	71	94	68	61

(a) Draw a scatter graph for this information. **(2)**

(b) Give an interpretation of the correlation between the height and weight of the footballers. **(1)**

The equation of the regression line is $w = 1.37h - 173$

(c) Give an interpretation of the gradient of this regression line. **(1)**

The product moment correlation coefficient is calculated to be 0.81.

(d) Stating your hypotheses clearly test, at the 5% significance level, whether there is a positive **(4)** correlation between height and weight

(e) Determine whether you would reach the same conclusion at the 1% significance level. **(2)**
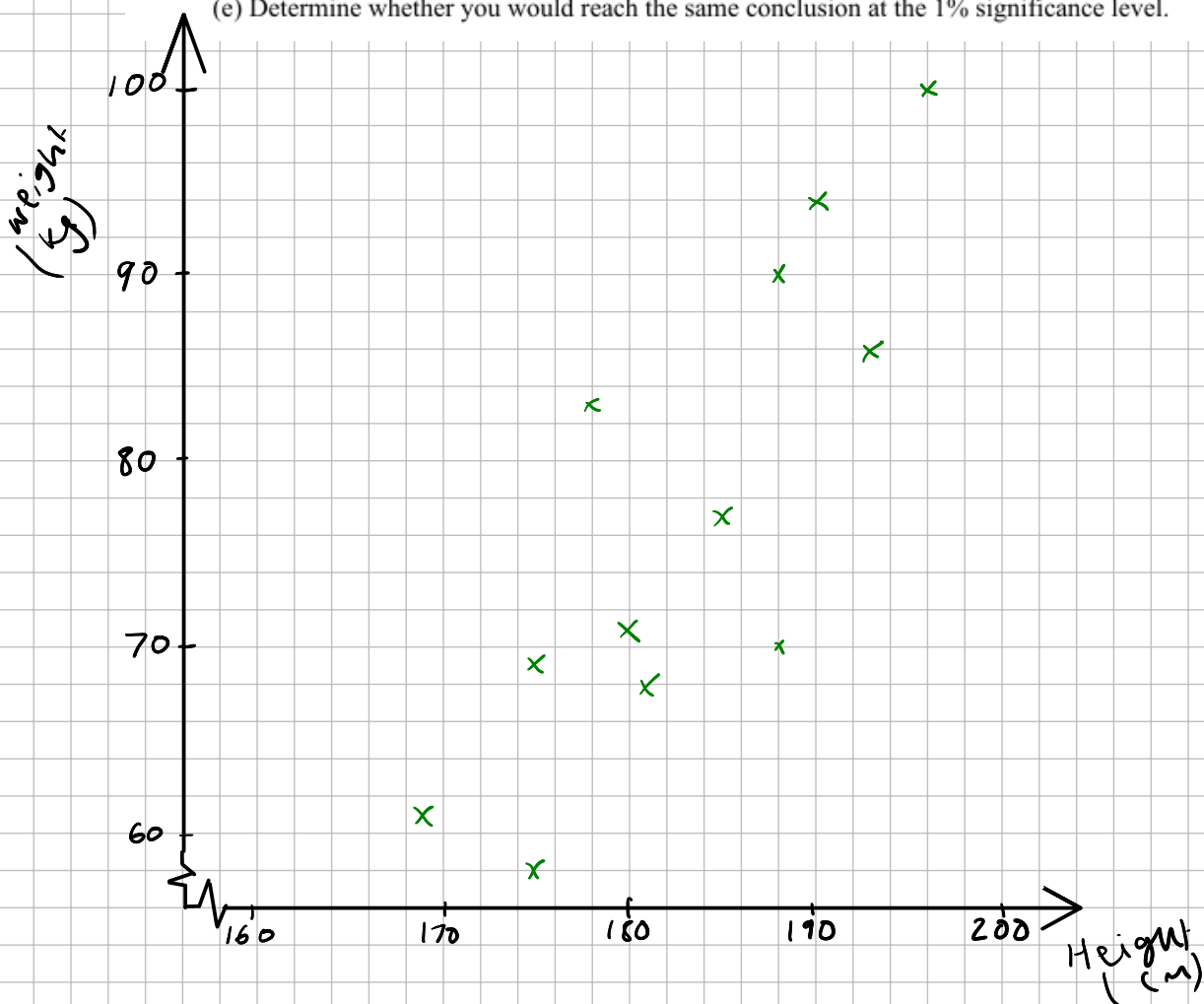
b/ As height of the footballers increases the weight increases.

c/ For every 1cm the height increases the weight increases by 1.37kg.

d/ $H_0 : \rho = 0$ $H_1 : \rho > 0$

$n = 12$ From table   critical value = 0.4973

$$0.81 > 0.4973$$

There is evidence to suggest $~~_{there~is}$ a positive correlation between height and weight.

e/ critical value = 0.6581

$$0.81 > 0.6581$$ Yes, The conclusion would be the same

---

2   The temperature and the rainfall on 15 days is recorded.

(a) Suggest a null and alternative hypothesis for a two tailed test to investigate whether there is a correlation between temperature and rainfall. (2)

The product moment correlation coefficient is calculated to be r = 0.37.

(b) Test your hypotheses at the 10% significance level. (2)

2a/ $H_0 : \rho = 0$ $H_1 : \rho \neq 0$ $n = 15$

b/   critical values -0.4409 and 0.4409

$0.37 < 0.4409$ There is not enough evidence to suggest that there is a correlation between temperature and hours of sunshine.

---

3   The temperature and the number of hours of sunshine on 12 days is recorded.
The product moment correlation coefficient is calculated to be r = 0.636.

Stating your hypotheses clearly test, at the 1% significance level, whether there is a positive correlation between temperature and hours of sunshine

$H_0 : \rho = 0$ $H_1 : \rho > 0$ $n = 12$

Critical value = 0.6581 $0.636 < 0.6581$

There is not enough evidence to say that there is a positive correlation between temperature and hours of sunshine.

4 Amy wants to find out if there is a correlation between daily maximum relative humidity and daily mean pressure.

(a) Suggest a suitable null and alternative hypothesis for a two-tail test.

Amy takes a sample of 14 days and finds a product moment correlation coefficient of -0.55.

(b) Carry out the hypothesis test at the 5% significance level.

a) $H_0: \rho = 0$ $H_1: \rho \neq 0$

b) $n = 14$ critical values -0.5324 and 0.5324

$-0.55 < -0.5324$

There is evidence to suggest there is a correlation between daily maximum relative humidity and daily mean pressure.

5   James runs a fish and chip shop. He records the average daily temperature, $t°C$, and the daily sales, $£S$, for 10 days in the summer.
The product moment correlation for these data is 0.67

(a) Stating your hypotheses clearly and using a 5% significance level, test whether there is a positive correlation between daily sales and daily temperature.

James suggests that a linear regression model could be used to model the data.

(b) State, giving a reason, whether or not the correlation coefficient is consistent with James's suggestion.

(c) State, giving a reason, which variable would be the explanatory variable.

James calculated the linear regression equation as $S = 5148 + 141t$

(d) Give an interpretation of the gradient of this regression equation.

a/ $H_0 : \rho = 0$ $H_1 : \rho > 0$ $n = 10$

Critical value = 0.5494

0.67 > 0.5494   ∴ There is evidence to suggest a positive correlation between daily sales and daily temporature.

b/   Yes. There is a positive correlation so a line of best fit can be drawn.

c/   The temperature. A change in the temperature can cause the sales to increase (not the other way around).

d/ For every 1°C increase in temperature sales increase by £141.

6 Chris is investigating the relationship between the length of time an employee has worked at his company, x years, and the amount of time it takes them to complete a task, y hours.

Chris records the time it takes 12 employees to complete the task and finds the product moment correlation coefficient between the length of time the employee has worked at the company and the time taken to complete the task to be -0.47

(a) Stating your hypotheses clearly and using a 5% level of significance, test whether or not the correlation between the length of time an employee has worked the company and the amount of time it takes them to complete the task is negative.

Chris believes that a non-linear model would be a better fit to the data. He codes the data using the coding $m = \log_{10}x$ and $n = \log_{10}y$ and obtains the model $n = 0.25 - 0.16m$

The product moment correlation coefficient between m and n is found to be -0.83

(b) Explain how this value supports Chris's belief.

(c) Show that the relationship between x and y can be written in the form $y = ax^b$ where a and b are constants to be found.

a/ $H_0 : \rho = 0 \qquad H_1 : \rho < 0 \qquad\qquad n = 12$

critical value $= -0.4973$

$-0.47 > -0.4973 \qquad$ There is not enough evidence to suggest that there is a negative correlation between the length of time an employee has worked for the company and the time taken to complete the task.

b/ The product moment correlation shows a stronger correlation. -0.83 is closer to -1.
$-0.83 < -0.4973 \qquad \rightarrow$ There is evidence to support
$\qquad\qquad\qquad\qquad\qquad\qquad$ Chris's belief.

c/ $\qquad\qquad n = 0.25 - 0.16m$
$\log_{10} y = 0.25 - 0.16 \log_{10} x$
$\log_{10} y = 0.25 - \log_{10} x^{0.16}$

$\log_{10} y + \log_{10} x^{0.16} = 0.25$
$\qquad \log_{10} y x^{0.16} = 0.25$
$\qquad\qquad x^{0.16} y = 10^{0.25}$
$\qquad\qquad\qquad y = 10^{0.25} x^{-0.16}$
$\qquad\qquad\qquad y = 1.78 x^{-0.16}$

7   Two variables x and y are related by the formula $y = ax^b$ where a and b are constants.

Show that this relationship can be written in the form $\log y = \log a + b \log x$

$$\log y = \log a x^b$$

$$\log y = \log a + \log x^b$$

$$\underline{\log y = \log a + b \log x}$$

---

8   The temperature of water ($w°C$) in a kettle t minutes after it was boiled is recorded in the table below

t	1	3	6	9	11	15	20
w	92	76	65	50	46	33	25

The data is coded using the changes of variable $x = t$ and $y = \log_{10} w$

The regression line of x on y is found to be $y = 1.99 - 0.031x$

(a) Given that the data can be modelled by an equation in the form $w = ab^t$ where a and b are constants. Find the values of a and b to 3 significant figures.   **(3)**

(b) Give an interpretation of the constant a in this equation.   **(1)**

(c) Explain why this model is not reliable for calculating the temperature after 1 hour.   **(1)**

a/ $$\log_{10} w = 1.99 - 0.031t$$

$$w = 10^{1.99 - 0.031t}$$

$$w = 10^{1.99} \cdot \left(10^{-0.031}\right)^t$$

$$w = 97.7 \cdot 0.931^t$$

$$\underline{a = 97.7} \quad \underline{b = 0.931}$$

b/   The initial temperature (immediatly after the kettle boiled) was $97.7°C$.

c/   The temperature $\overset{\text{in the model}}{\vee}$ would get closer to zero, instead it should reach room temperature.

9 The variables x and y are recorded and the results are shown in the table below

x	1	2	3	4	5	6	7	8	9	10
y	429	754	871	1119	2478	2653	3050	3279	5470	7439

The data is coded using the changes of variable $r = x$ and $s = \log_{10} y$

(a) The product moment correlation coefficient for the coded data is 0.98
Comment on r for this model and therefore justify the use of a model in the form $y = ab^x$
here a and b are constants

The regression line of r on s is found to be $s = 2.60 + 0.127r$

(b) Find the values of a and b, give your answers to 3 significant figures

a/   0.98 is close to 1
This suggests that there is a non linear
relationship between x and y.

b/ $s = 2.60 + 0.127r$

$\log_{10} y = 2.60 + 0.127x$

$$y = 10^{2.60 + 0.127x}$$
$$= (10^{2.60})(10^{0.127})^x$$
$$= 398 (1.34)^x$$

$a = 398 \qquad b = 1.34$

10   Will sells cars. He thinks there is a positive correlation between the price of a car and its mileage.

He calculates the product moment correlation coefficient between these two variables for a random sample of 15 cars and obtains r = 0.455

Carry out a suitable test to investigate Will's belief at a 5% level of significance.
State clearly
- your hypotheses
- your critical value

$H_0: \rho = 0$ $H_1: \rho > 0$

$n = 15$ critical value = 0.4409

$0.455 > 0.4409$ There is evidence to support Will's claim.

---

11   Davina wants to test to see if there is a correlation between the number of letters in someone's name and their salary.

She calculates the product moment correlation coefficient between these two variables for a random sample of 40 people and obtains r = -0.289

Carry out a suitable test to investigate Davina's belief at a 5% level of significance.
State clearly
- your hypotheses
- your critical value

$H_0: \rho = 0$ $H_1: \rho \neq 0$ (two tailed)

$n = 40$ critical values = 0.3120 and -0.3120

There is not enough evidence to suggest there is a correlation between the number of letters in someone's name and salary.