0510Test-statistics *[29 marks]*

The following table shows the sales, y millions of dollars, of a company, x years after it opened.

Time after opening (x years)	2	4	6	8	10
Sales (y millions of dollars)	12	20	30	36	52

The relationship between the variables is modelled by the regression line with equation $y = ax + b$.

1.
(i) Find the value of a and of b. *[4 marks]*

(ii) Write down the value of r.

Markscheme

(i) evidence of set up   *(M1)*

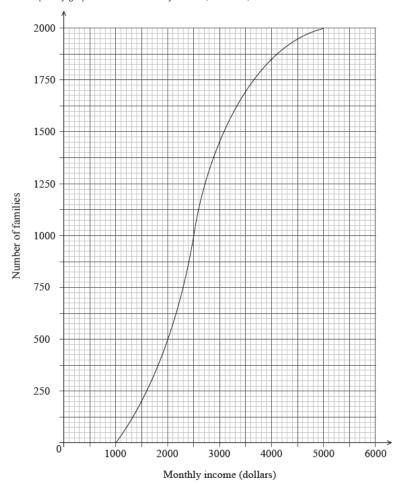*eg* correct value for a, b or r

$a = 4.8$, $b = 1.2$ *A1A1*   *N3*

(ii) $r = 0.988064$

$r = 0.988$ *A1*   *N1*

[4 marks]

The following cumulative frequency graph shows the monthly income, I dollars, of 2000 families.

Find the median monthly income.

2a. [2 marks]

2b. (i)   Write down the number of families who have a monthly income of 2000 dollars or less.   [4 marks]

(ii)   Find the number of families who have a monthly income of more than 4000 dollars.

2c. The 2000 families live in two different types of housing. The following table gives information about the number of families living in   [2 marks]
each type of housing and their monthly income I.

	$1000 < I \leq 2000$	$2000 < I \leq 4000$	$4000 < I \leq 5000$
Apartment	436	765	28
Villa	64	p	122

Find the value of p.

2d. A family is chosen at random.   [2 marks]

(i)   Find the probability that this family lives in an apartment.

(ii)   Find the probability that this family lives in an apartment, given that its monthly income is greater than 4000 dollars.

Markscheme

(i)   correct working   *(A1)*

eg $436 + 765 + 28$

0.6145 (exact)   *A1   N2*

$\frac{1229}{2000}$, 0.615 $[0.614, \ 0.615]$

(ii)   correct working/probability for number of families   *(A1)*

eg $122 + 28$, $\frac{150}{2000}$, 0.075

0.186666

$\frac{28}{150}$ $\left(= \frac{14}{75}\right)$, 0.187 $[0.186, \ 0.187]$ *A1   N2*
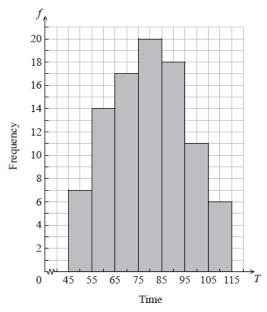
[4 marks]

2e.   Estimate the mean monthly income for families living in a villa.   *[2 marks]*

Markscheme

evidence of using correct mid-interval values $(1500, 3000, 4500)$ *(A1)*

attempt to substitute into $\frac{\sum fx}{\sum f}$ *(M1)*

eg $\frac{1500 \times 64 + 3000 \times p + 4500 \times 122}{64 + 585 + 122}$

3112.84

3110 $[3110, \ 3120]$ (dollars)   *A1   N2*

[3 marks]

Total [15 marks]

The histogram below shows the time T seconds taken by 93 children to solve a puzzle.

The following is the frequency distribution for T.

Time	$45 \leq T < 55$	$55 \leq T < 65$	$65 \leq T < 75$	$75 \leq T < 85$	$85 \leq T < 95$	$95 \leq T < 105$	$105 \leq T < 115$
Frequency	7	14	p	20	18	q	6

3a. (i)   Write down the value of p and of q.   [3 marks]

(ii)   Write down the median class.

Markscheme

(i)
$p = 17$,
$q = 11$ **A1A1**   **N2**

(ii)
$75 \leq T < 85$ **A1**   **N1**

[3 marks]

3b. A child is selected at random. Find the probability that the child takes less than 95 seconds to solve the puzzle.   [2 marks]

Markscheme

evidence of valid approach   **(M1)**

e.g. adding frequencies

$\frac{76}{93} = 0.8172043\ldots$

$P(T < 95) = \frac{76}{93} = 0.817$ **A1**   **N2**

[2 marks]

3c. Consider the class interval
$45 \leq T < 55$.

   (i)   Write down the interval width.

   (ii)   Write down the mid-interval value.

Markscheme

(i) 10   *A1*   *N1*

(ii) 50   *A1*   *N1*

[2 marks]

3d. Hence find an estimate for the

   (i)   mean;

   (ii)   standard deviation.

[4 marks]

Markscheme

(i) evidence of approach using mid-interval values (may be seen in part (ii))   *(M1)*

79.1397849

$\overline{x} = 79.1$ *A2*   *N3*

(ii)
16.4386061

$\sigma = 16.4$ *A1*   *N1*

[4 marks]

3e. John assumes that *T* is normally distributed and uses this to estimate the probability that a child takes less than 95
seconds to solve the puzzle.

[2 marks]

Find John's estimate.

Markscheme

e.g. standardizing,
$z = 0.9648\ldots$

0.8326812

$\mathrm{P}(T < 95) = 0.833$ *A1*   *N2*

[2 marks]