Topic 2 Part 4 *[47 marks]*

A random sample of 167 people who own mobile phones was used to collect data on the amount of time they spent per day using their phones. The results are displayed in the table below.

Time spent per day (t minutes)	$0 \le t < 15$	$15 \le t < 30$	$30 \le t < 45$	$45 \le t < 60$	$60 \le t < 75$	$75 \le t < 90$
Number of people	21	32	35	41	27	11

1a. State the modal group. *[1 mark]*

1b. Use your graphic display calculator to calculate approximate values of the mean and standard deviation of the time spent per day on these mobile phones. *[3 marks]*

1c. On graph paper, draw a fully labelled histogram to represent the data. *[4 marks]*

Manuel conducts a survey on a random sample of 751 people to see which television programme type they watch most from the following: Drama, Comedy, Film, News. The results are as follows.

	Drama	Comedy	Film	News
Males under 25	22	65	90	35
Males 25 and over	36	54	67	17
Females under 25	22	59	82	15
Females 25 and over	64	39	38	46

Manuel decides to ignore the ages and to test at the 5 % level of significance whether the most watched programme type is independent of **gender.**

1d. Draw a table with 2 rows and 4 columns of data so that Manuel can perform a chi-squared test. *[3 marks]*

1e. State Manuel's null hypothesis and alternative hypothesis. *[1 mark]*

1f. Find the expected frequency for the number of females who had 'Comedy' as their most-watched programme type. Give your answer to the nearest whole number. *[2 marks]*

1g. Using your graphic display calculator, or otherwise, find the chi-squared statistic for Manuel's data. *[3 marks]*

1h. (i) State the number of degrees of freedom available for this calculation. *[3 marks]*

(ii) State his conclusion.

The lengths (l) in centimetres of 100 copper pipes at a local building supplier were measured. The results are listed in the table below.

Length l (cm)	Frequency
17.5	12
32.5	26
47.5	32
62.5	21
77.5	9

2a. Write down the mode. *[1 mark]*

2b. Using your graphic display calculator, write down the value of *[4 marks]*
(i) the mean;
(ii) the standard deviation;
(iii) the median.

2c. Find the interquartile range. *[2 marks]*

2d. Draw a box and whisker diagram for this data, on graph paper, using a scale of 1 cm to represent 5 cm. *[4 marks]*

2e. Sam estimated the value of the mean of the measured lengths to be 43 cm. *[2 marks]*

Find the percentage error of Sam's estimated mean.

The number of bottles of water sold at a railway station on each day is given in the following table.

Day	0	1	2	3	4	5	6	7	8	9	10	11	12
Temperature ($T°$)	21	20.7	20	19	18	17.3	17	17.3	18	19	20	20.7	21
Number of bottles sold (n)	150	141	126	125	98	101	93	99	116	121	119	134	141

3a. Write down *[2 marks]*
(i) the mean temperature;
(ii) the standard deviation of the temperatures.

3b. Write down the correlation coefficient, r, for the variables n and T. *[1 mark]*

3c. Comment on your value for r. *[2 marks]*

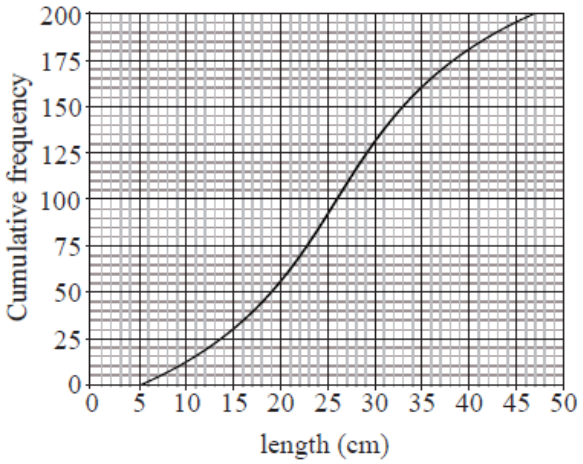
3d. The equation of the line of regression for n on T is $n = dT - 100$. *[2 marks]*

(i) Write down the value of d.

(ii) Estimate how many bottles of water will be sold when the temperature is $19.6°$.

3e. On a day when the temperature was $36°$ Peter calculates that 314 bottles would be sold. Give one reason why his answer might be unreliable. *[1 mark]*

A random sample of 200 females measured the length of their hair in cm. The results are displayed in the cumulative frequency curve below.



4a. Write down the median length of hair in the sample. *[1 mark]*

4b. Find the interquartile range for the length of hair in the sample. *[2 marks]*

4c. Given that the shortest length was 6 cm and the longest 47 cm, draw and label a box and whisker plot for the data on the grid provided below. *[3 marks]*