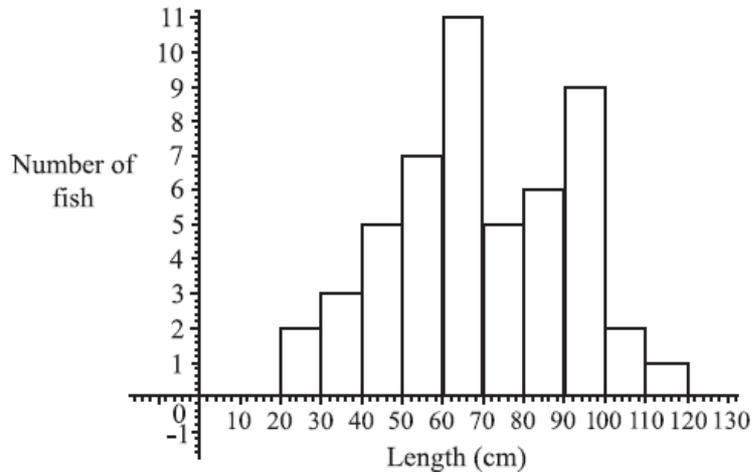


Topic 4 Part 4 [94 marks]

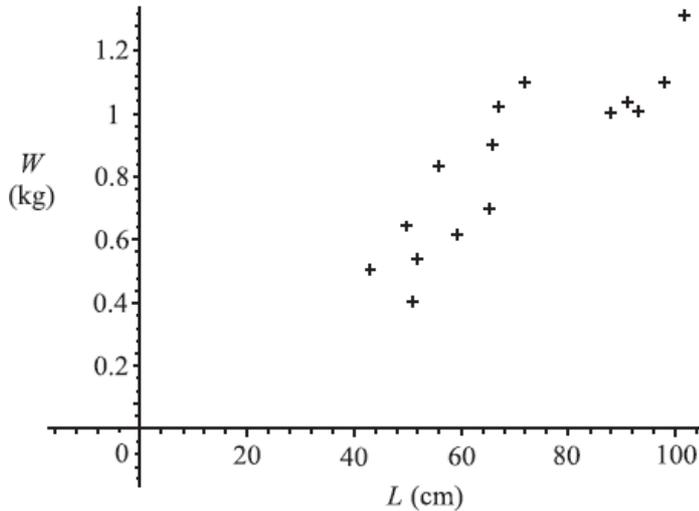
The figure below shows the lengths in centimetres of fish found in the net of a small trawler.



- 1a. Find the total number of fish in the net. [2 marks]
- 1b. Find (i) the modal length interval, [5 marks]
(ii) the interval containing the median length,
(iii) an estimate of the mean length.
- 1c. (i) Write down an estimate for the standard deviation of the lengths. [3 marks]
(ii) How many fish (if any) have length **greater than** three standard deviations **above** the mean?
- 1d. The fishing company must pay a fine if more than 10% of the catch have lengths less than 40cm. [2 marks]
Do a calculation to decide whether the company is fined.

1e. A sample of 15 of the fish was weighed. The weight, W was plotted against length, L as shown below.

[2 marks]



Exactly **two** of the following statements about the plot could be correct. Identify the two correct statements.

Note: You do **not** need to enter data in a GDC **or** to calculate r exactly.

- (i) The value of r , the correlation coefficient, is approximately 0.871.
- (ii) There is an exact linear relation between W and L .
- (iii) The line of regression of W on L has equation $W = 0.012L + 0.008$.
- (iv) There is negative correlation between the length and weight.
- (v) The value of r , the correlation coefficient, is approximately 0.998.
- (vi) The line of regression of W on L has equation $W = 63.5L + 16.5$.

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)	$0 \leq t < 15$	$15 \leq t < 30$	$30 \leq t < 45$	$45 \leq t < 60$	$60 \leq t < 75$	$75 \leq t < 90$
Number of people	21	32	35	41	27	11

2a. State the modal group.

[1 mark]

2b. 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]

2c. 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**.

2d. Draw a table with 2 rows and 4 columns of data so that Manuel can perform a chi-squared test.

[3 marks]

2e. State Manuel's null hypothesis and alternative hypothesis.

[1 mark]

2f. 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]

2g. Using your graphic display calculator, or otherwise, find the chi-squared statistic for Manuel's data. [3 marks]

2h. (i) State the number of degrees of freedom available for this calculation. [3 marks]

(ii) State his conclusion.

The local park is used for walking dogs. The sizes of the dogs are observed at different times of the day. The table below shows the numbers of dogs present, classified by size, at three different times last Sunday.

	Small	Medium	Large
Morning	9	18	21
Afternoon	11	6	13
Evening	7	8	9

3a. Write a suitable null hypothesis for a χ^2 test on this data. [1 mark]

3b. Write down the value of χ^2 for this data. [2 marks]

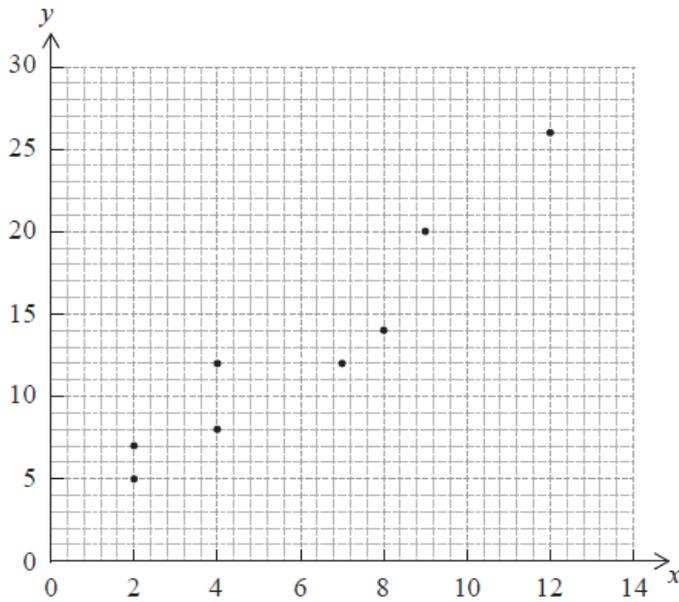
3c. The number of degrees of freedom is 4. Show how this value is calculated. [1 mark]

3d. The critical value, at the 5% level of significance, is 9.488. [2 marks]

What conclusion can be drawn from this test? Give a reason for your answer.

Consider the following set of data which is plotted on the scatter diagram below.

x	2	4	7	12	4	8	9	2
y	5	8	12	26	12	14	20	7



- 4a. Write down the coordinates of the mean point (\bar{x}, \bar{y}) . [2 marks]
- 4b. Write down the value of r , the Pearson's product-moment correlation coefficient for this set of data. [2 marks]
- 4c. Draw the regression line for y on x on the set of axes above. [2 marks]

Members of a certain club are required to register for one of three sports, badminton, volleyball or table tennis. The number of club members of each gender choosing each sport in a particular year is shown in the table below.

A

χ^2 (Chi-squared) test at the

5% significance level is used to determine whether the choice of sport is independent of gender.

	Badminton	Volleyball	Table tennis
Male	40	20	10
Female	20	15	15

- 5a. Find the expected number of female volleyball players under this hypothesis. [2 marks]
- 5b. Write down the p -value for the test. [2 marks]
- 5c. State, with a reason, the conclusion of the test. [2 marks]

The Brahma chicken produces eggs with weights in grams that are normally distributed about a mean of 55 g with a standard deviation of 7 g. The eggs are classified as small, medium, large or extra large according to their weight, as shown in the table below.

Size	Weight (g)
Small	Weight < 53
Medium	$53 \leq \text{Weight} < 63$
Large	$63 \leq \text{Weight} < 73$
Extra Large	Weight ≥ 73

6a. Sketch a diagram of the distribution of the weight of Brahma chicken eggs. On your diagram, show clearly the boundaries for the classification of the eggs. [3 marks]

6b. An egg is chosen at random. Find the probability that the egg is [4 marks]
 (i) medium;
 (ii) extra large.

6c. There is a probability of 0.3 that a randomly chosen egg weighs more than w grams. [2 marks]
 Find w .

6d. The probability that a Brahma chicken produces a large size egg is 0.121. Frank's Brahma chickens produce 2000 eggs each month. [2 marks]
 Calculate an estimate of the number of large size eggs produced by Frank's chickens each month.

6e. The selling price, in US dollars (USD), of each size is shown in the table below. [3 marks]

Size	Selling price (USD)
Small	0.30
Medium	0.50
Large	0.65
Extra Large	0.80

The probability that a Brahma chicken produces a small size egg is 0.388.

Estimate the monthly income, in USD, earned by selling the 2000 eggs. Give your answer correct to two decimal places.

Tania wishes to see whether there is any correlation between a person's age and the number of objects on a tray which could be remembered after looking at them for a certain time.

She obtains the following table of results.

Age (x years)	15	21	36	40	44	55
Number of objects remembered (y)	17	20	15	16	17	12

7a. Use your graphic display calculator to find the equation of the regression line of y on x . [2 marks]

- 7b. Use your equation to estimate the number of objects remembered by a person aged 28 years. [1 mark]
- 7c. Use your graphic display calculator to find the correlation coefficient r . [1 mark]
- 7d. Comment on your value for r . [2 marks]

A survey of 400 people is carried out by a market research organization in two different cities, Buenos Aires and Montevideo. The people are asked which brand of cereal they prefer out of Chocos, Zucos or Fruti. The table below summarizes their responses.

	Chocos	Zucos	Fruti	Total
Buenos Aires	43	85	62	190
Montevideo	57	35	118	210
Total	100	120	180	400

- 8a. One person is chosen at random from those surveyed. Find the probability that this person [4 marks]
- (i) does not prefer Zucos;
- (ii) prefers Chocos, given that they live in Montevideo.
- 8b. Two people are chosen at random from those surveyed. Find the probability that they both prefer Fruti. [3 marks]
- 8c. The market research organization tests the survey data to determine whether the brand of cereal preferred is associated with a city. A chi-squared test at the 5% level of significance is performed. [1 mark]
- State the null hypothesis.
- 8d. The market research organization tests the survey data to determine whether the brand of cereal preferred is associated with a city. A chi-squared test at the 5% level of significance is performed. [1 mark]
- State the number of degrees of freedom.
- 8e. The market research organization tests the survey data to determine whether the brand of cereal preferred is associated with a city. A chi-squared test at the 5% level of significance is performed. [2 marks]
- Show that the expected frequency for the number of people who live in Montevideo and prefer Zucos is 63.
- 8f. The market research organization tests the survey data to determine whether the brand of cereal preferred is associated with a city. A chi-squared test at the 5% level of significance is performed. [2 marks]
- Write down the chi-squared statistic for this data.
- 8g. The market research organization tests the survey data to determine whether the brand of cereal preferred is associated with a city. A chi-squared test at the 5% level of significance is performed. [2 marks]
- State whether the market research organization would accept the null hypothesis. Clearly justify your answer.

The following table shows the cost in AUD of seven paperback books chosen at random, together with the number of pages in each book.

Book	1	2	3	4	5	6	7
Number of pages (x)	50	120	200	330	400	450	630
Cost (y AUD)	6.00	5.40	7.20	4.60	7.60	5.80	5.20

- 8h. Plot these pairs of values on a scatter diagram. Use a scale of 1 cm to represent 50 pages on the horizontal axis and 1 cm to represent 1 AUD on the vertical axis. *[3 marks]*
- 8i. Write down the linear correlation coefficient, r , for the data. *[2 marks]*
- 8j. Stephen wishes to buy a paperback book which has 350 pages in it. He plans to draw a line of best fit to determine the price. State whether or not this is an appropriate method in this case and justify your answer. *[2 marks]*