Topic 2 Part 3 *[189 marks]*

The grades obtained by a group of 13 students are listed below.

5 3 6 5 7 3 2 6 4 6 6 6 4

1a. Write down the modal grade.   *[1 mark]*

1b. Find the mean grade.   *[2 marks]*

1c. Write down the standard deviation.   *[1 mark]*

1d. Find the interquartile range.   *[2 marks]*

The length, in cm, of six baseball bats was measured. The lengths are given below.

104.5, 105.1, 104.8, 105.2, 104.9, 104.9

2a. Calculate the **exact value** of the mean length.   *[2 marks]*

2b. Write your answer to part (a) in the form $a \times 10^k$ where $1 \leq a < 10$ and $k \in \mathbb{Z}$.   *[2 marks]*

2c. Marian calculates the mean length and finds it to be 105 cm.   *[2 marks]*

Calculate the percentage error made by Marian.

The cumulative frequency curve shows the percentage marks, given correct to the nearest integer, gained by 500 students in an examination.
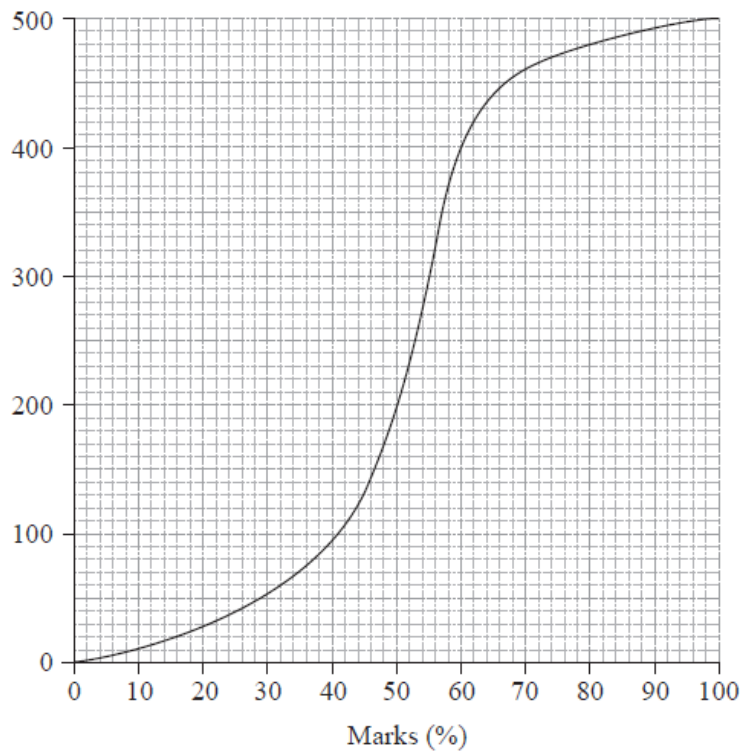
The passing grades were determined as given below.

85 to 100 %, grade A

66 to 84 %, grade B

57 to 65 %, grade C

50 to 56 %, grade D

Those scoring less than 50 % failed the examination.

3a. Find the number of students who failed the examination. [2 marks]

3b. Find the number of students who were awarded grade C or better. [2 marks]

3c. The top 20 % of the students are eligible for further study. [2 marks]

Find the lowest mark required to be eligible for further study.

Toronto's annual snowfall, x, in cm, has been recorded for the past 176 years. The results are shown in the table.

Snowfall (cm)	$2 \leq x < 6$	$6 \leq x < 10$	$10 \leq x < 14$	$14 \leq x < 18$	$18 \leq x < 22$	$22 \leq x < 26$	$26 \leq x < 30$
Frequency	30	26	29	32	18	27	14

4a. Write down the modal class. [1 mark]

4b. Write down the mid interval value for the class $6 \leq x < 10$. [1 mark]

4c. Calculate an estimate of the mean annual snowfall. [2 marks]

4d. Find the number of years for which the annual snowfall was at least 18 cm. [2 marks]

The table below shows the scores for 12 golfers for their first two rounds in a local golf tournament.

Round 1 (x)	71	79	66	73	69	76	68	75	82	67	69	74
Round 2 (y)	73	81	68	75	70	79	69	77	83	68	72	76

5a. (i) Write down the mean score in Round 1.　　　　　　　　　　　　　　　　　　[5 marks]

(ii) Write down the standard deviation in Round 1.

(iii) Find the number of these golfers that had a score of more than one standard deviation above the mean in Round 1.

5b. Write down the correlation coefficient, r.　　　　　　　　　　　　　　　　　[2 marks]

5c. Write down the equation of the regression line of y on x.　　　　　　　　　[2 marks]

5d. Another golfer scored 70 in Round 1.　　　　　　　　　　　　　　　　　　　[2 marks]

Calculate an estimate of his score in Round 2.

5e. Another golfer scored 89 in Round 1.　　　　　　　　　　　　　　　　　　　[2 marks]

Determine whether you can use the equation of the regression line to estimate his score in Round 2. Give a reason for your answer.
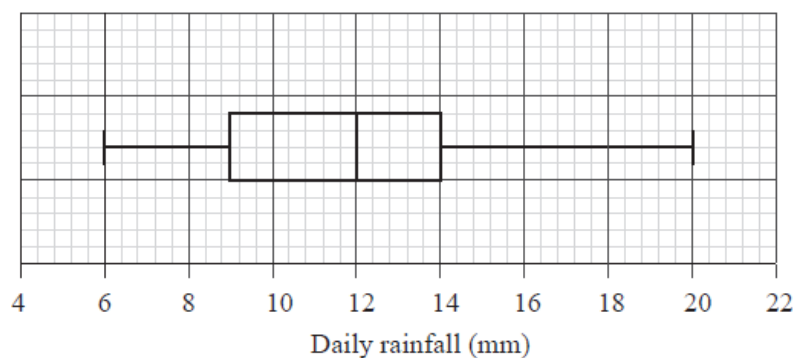
The following **six** integers are arranged from smallest to largest

$1 , x , 3 , y , 14 , z$

The mode is 1 , the median is 5 and the mean is 7.

6a. Find x ;　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　[1 mark]

6b. Find y ;　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　[2 marks]

6c. Find z .　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　[3 marks]

The daily rainfall for the town of St. Anna is collected over a 20-day period of time. The collected data are represented in the box and whisker plot below.

7a. Write down　　　　　　　　　　　　　　　　　　　　　　　　　　　　　[2 marks]

(i) the lowest daily rainfall;

(ii) the highest daily rainfall.

7b. State what the value of 12 mm represents on the given diagram.   *[1 mark]*

7c. Find the interquartile range.   *[2 marks]*

7d. Write down the percentage of the data which is less than the upper quartile.   *[1 mark]*

The resting pulse rates of a group of 10 students who exercise regularly are given below.

$$65, 62, 75, 63, 69, 58, 65, 67, 55, 60$$

8a. Find the median resting pulse rate of the students.   *[2 marks]*

8b. Find the mean resting pulse rate of the students.   *[2 marks]*

8c. A new student joins the class and the mean resting pulse rate of the group of 11 students becomes 65.   *[2 marks]*

Find the resting pulse rate of the student who joined the group.

200 people were asked the amount of time T (minutes) they had spent in the supermarket. The results are represented in the table below.

Time (T)	$0 < T \leq 10$	$10 < T \leq 20$	$20 < T \leq 30$	$30 < T \leq 40$	$40 < T \leq 50$
Number of people	23	57	93	21	6

9a. State if the data is discrete or continuous.   *[1 mark]*

9b. State the modal group.   *[1 mark]*

9c. Write down the midpoint of the interval $10 < T \leq 20$.   *[1 mark]*

9d. Use your graphic display calculator to find an estimate for   *[3 marks]*

(i) the mean;

(ii) the standard deviation.

9e. The results are represented in the cumulative frequency table below, with upper class boundaries of 10, 20, 30, 40, 50.   *[2 marks]*

Upper class boundaries	10	20	30	40	50
Cumulative frequency	23	80	173	q	r

Write down the value of

(i) q;

(ii) r.

9f. The results are represented in the cumulative frequency table below, with upper class boundaries of 10, 20, 30, 40, 50.   *[4 marks]*

Upper class boundaries	10	20	30	40	50
Cumulative frequency	23	80	173	q	r

On graph paper, draw a cumulative frequency graph, using a scale of 2 cm to represent 10 minutes (T) on the horizontal axis and 1 cm to represent 10 people on the vertical axis.

9g. **Use your graph** from part (f) to estimate   *[6 marks]*

   (i) the median;

   (ii) the 90th percentile of the results;

   (iii) the number of people who shopped at the supermarket for more than 15 minutes.

Francesca is a chef in a restaurant. She cooks eight chickens and records their masses and cooking times. The mass m of each chicken, in kg, and its cooking time t, in minutes, are shown in the following table.

Mass m (kg)	Cooking time t (minutes)
1.5	62
1.6	75
1.8	82
1.9	83
2.0	86
2.1	87
2.1	91
2.3	98

10a. Draw a scatter diagram to show the relationship between the mass of a chicken and its cooking time. Use 2 cm to represent   *[4 marks]*
0.5 kg on the horizontal axis and 1 cm to represent 10 minutes on the vertical axis.

10b. Write down for this set of data   *[2 marks]*

   (i) the mean mass,
   $\bar{m}$;

   (ii) the mean cooking time,
   $\bar{t}$.

10c. Label the point   *[1 mark]*
   $M(\bar{m}, \bar{t})$ on the scatter diagram.

10d. Draw the line of best fit on the scatter diagram.   *[2 marks]*

10e. Using your line of best fit, estimate the cooking time, in minutes, for a 1.7 kg chicken.   *[2 marks]*

10f. Write down the Pearson's product–moment correlation coefficient, r .   *[2 marks]*

10g. Using your value for r , comment on the correlation.   *[2 marks]*

10h. The cooking time of an additional 2.0 kg chicken is recorded. If the mass and cooking time of this chicken is included in the   *[2 marks]*
data, the correlation is weak.

   (i) Explain how the cooking time of this additional chicken might differ from that of the other eight chickens.

   (ii) Explain how a new line of best fit might differ from that drawn in part (d).

120 Mathematics students in a school sat an examination. Their scores (given as a percentage) were summarized on a cumulative frequency diagram. This diagram is given below.
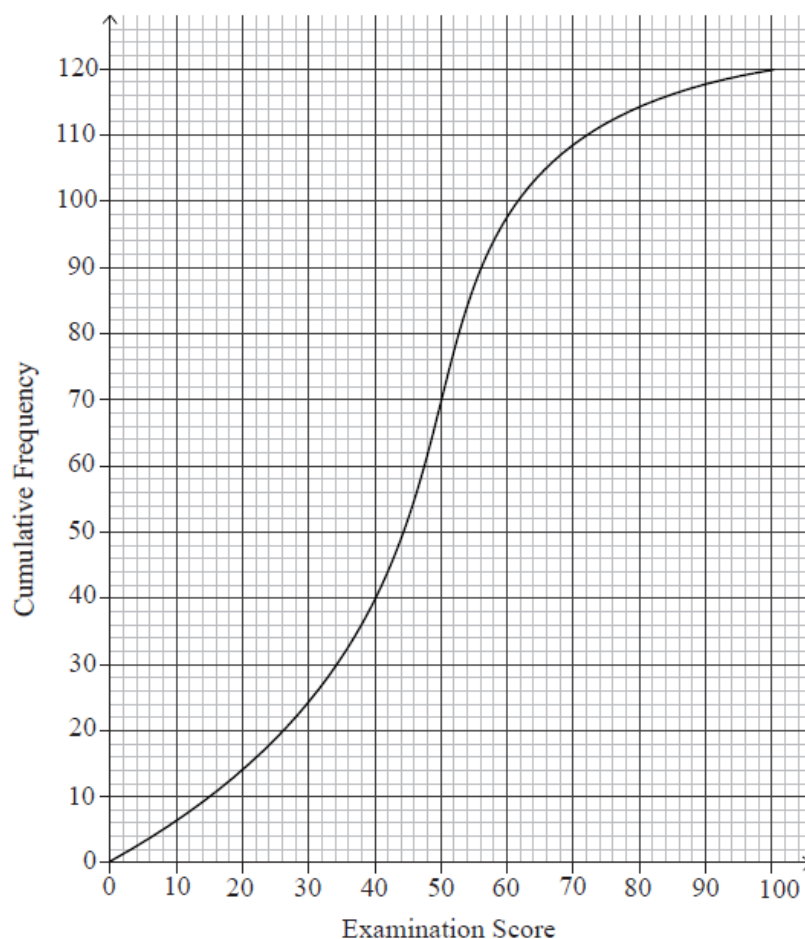
11a. Complete the grouped frequency table for the students.   [3 marks]

Examination Score x (%)	$0 \leq x \leq 20$	$20 < x \leq 40$	$40 < x \leq 60$	$60 < x \leq 80$	$80 < x \leq 100$
Frequency	14	26			

11b. Write down the mid-interval value of the $40 < x \leqslant 60$ interval.   [1 mark]

11c. Calculate an estimate of the mean examination score of the students.   [2 marks]

The cumulative frequency graph shows the amount of time in minutes, 200 students spend waiting for their train on a particular morning.
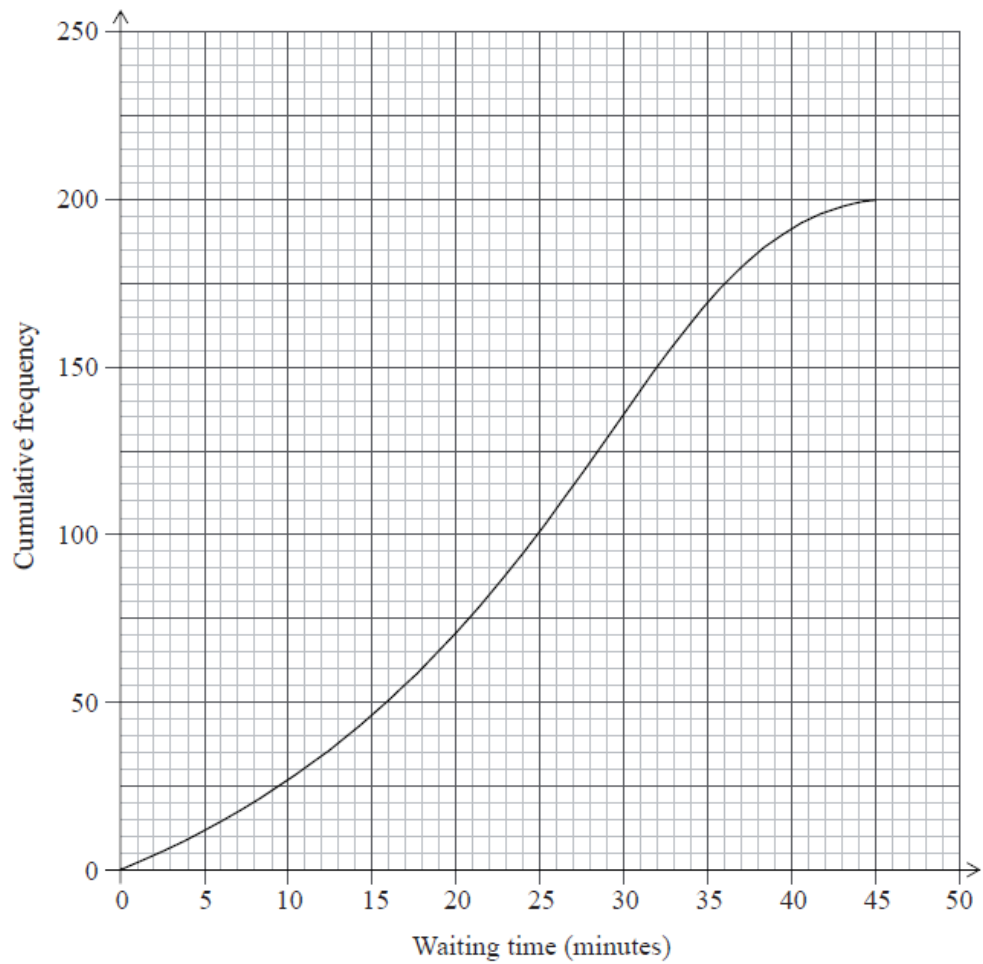
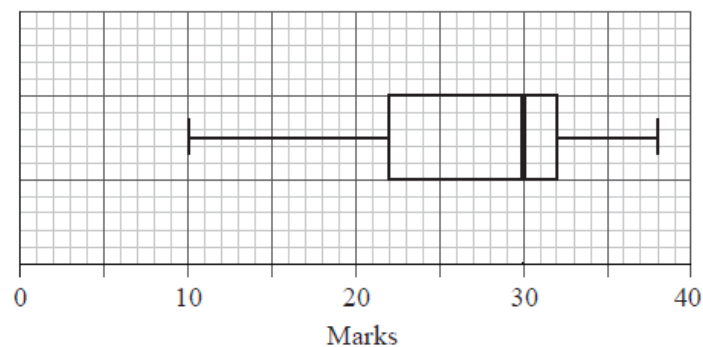12a. Write down the median waiting time.   [1 mark]

12b. Find the interquartile range for the waiting time.   [2 marks]

12c. Draw a box and whisker plot on the grid below to represent this information.   [3 marks]

Waiting time (minutes)
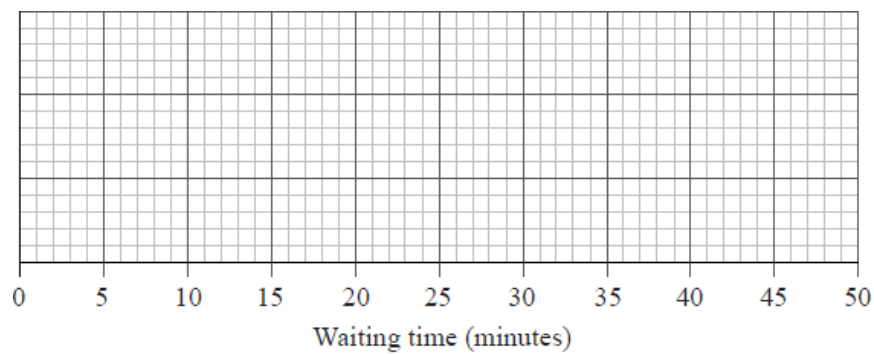
56 students were given a test out of 40 marks. The teacher used the following box and whisker plot to represent the marks of the students.

Marks

13a. Write down the median mark. [1 mark]

13b. Write down the 75th percentile mark. [1 mark]

13c. Write down the range of marks. [2 marks]

13d. Estimate the number of students who achieved a mark greater than 32. [2 marks]

The weights of 90 students in a school were recorded. The information is displayed in the following table.

Weight (kg)	Number of students
$40 \leq w < 50$	7
$50 \leq w < 60$	28
$60 \leq w < 70$	35
$70 \leq w < 80$	20

14a. Write down the mid interval value for the interval
$50 \leqslant w \leqslant 60$. [1 mark]

14b. Use your graphic display calculator to find an estimate for the mean weight. [2 marks]

14c. Use your graphic display calculator to find an estimate for the standard deviation. [1 mark]

14d. Find the weight that is 3 standard deviations below the mean. [2 marks]

The speed,
s, in
km h^{-1}, of
120 vehicles passing a point on the road was measured. The results are given below.

Speed, s (km h^{-1})	$0 < s \leq 50$	$50 < s \leq 60$	$60 < s \leq 70$	$70 < s \leq 80$	$80 < s \leq 90$	$90 < s \leq 100$
Number of vehicles	30	46	22	12	8	2

15a. Write down the midpoint of the
$60 < s \leqslant 70$ interval.

[1 mark]

15b. Use your graphic display calculator to find an estimate for

[3 marks]

(i) the mean speed of the vehicles;

(ii) the standard deviation of the speeds of the vehicles.

15c. Write down the number of vehicles whose speed is less than or equal to
60 km h^{-1}.

[1 mark]

15d. Consider the cumulative frequency table below.

[2 marks]

Speed, s (km h^{-1})	$s \leq 50$	$s \leq 60$	$s \leq 70$	$s \leq 80$	$s \leq 90$	$s \leq 100$
Number of vehicles	30	a	b	110	c	120

Write down the value of
a, of
b and of
c.

15e. Consider the cumulative frequency table below.

[4 marks]

Speed, s (km h^{-1})	$s \leq 50$	$s \leq 60$	$s \leq 70$	$s \leq 80$	$s \leq 90$	$s \leq 100$
Number of vehicles	30	a	b	110	c	120

Draw a cumulative frequency graph for the information from the table. Use
1 cm to represent
10 km h^{-1} on the horizontal axis and
1 cm to represent
10 vehicles on the vertical axis.

15f. Use your cumulative frequency graph to estimate

[4 marks]

(i) the median speed of the vehicles;

(ii) the number of vehicles that are travelling at a speed less than or equal to
65 km h^{-1}.

15g. All drivers whose vehicle's speed is greater than one standard deviation above the speed limit of
50 km h^{-1} will be fined.

[3 marks]

Use your graph to estimate the number of drivers who will be fined.

The table below shows the frequency distribution of the number of dental fillings for a group of 25 children.

Number of fillings	0	1	2	3	4	5
Frequency	4	3	8	q	4	1

16a. Find the value of q. [2 marks]

16b. Use your graphic display calculator to find [4 marks]
   (i)   the mean number of fillings;
   (ii)   the median number of fillings;
   (iii) the standard deviation of the number of fillings.

The weights in kg, of 80 adult males, were collected and are summarized in the box and whisker plot shown below.
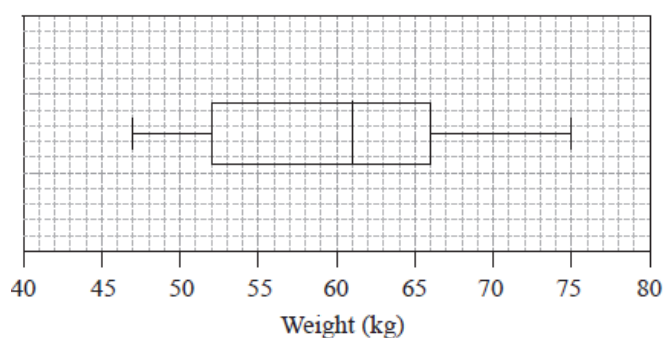
Weight (kg)

17a. Write down the median weight of the males. [1 mark]

17b. Calculate the interquartile range. [2 marks]

17c. Estimate the number of males who weigh between
   61 kg and
   66 kg. [1 mark]

17d. Estimate the mean weight of the lightest
   40 males. [2 marks]

The heat output in thermal units from burning
1 kg of wood changes according to the wood's percentage moisture content. The moisture content and heat output of
10 blocks of the same type of wood each weighing
1 kg were measured. These are shown in the table.

Moisture content % (x)	8	15	22	30	34	45	50	60	74	82
Heat output (y)	80	77	74	69	68	61	61	55	50	45

18a. Draw a scatter diagram to show the above data. Use a scale of
   2 cm to represent
10% on the x-axis and a scale of
   2 cm to represent
10 thermal units on the y-axis. [4 marks]

18b. Write down
   (i)   the mean percentage moisture content,
$\bar{x}$;
   (ii) the mean heat output,
$\bar{y}$.

[2 marks]

18c. Plot the point
$(\bar{x}, \bar{y})$ on your scatter diagram and label this point M .

[2 marks]

18d. Write down the product-moment correlation coefficient,
r .

[2 marks]

18e. The equation of the regression line
y on
x is
$y = -0.470x + 83.7$. Draw the regression line
y on
x on your scatter diagram.

[2 marks]

18f. The equation of the regression line
y on
x is
$y = -0.470x + 83.7$. Estimate the heat output in thermal units of a
1 kg block of wood that has
25% moisture content.

[2 marks]

18g. The equation of the regression line
y on
x is
$y = -0.470x + 83.7$. State, with a reason, whether it is appropriate to use the regression line
y on
x to estimate the heat output in part (f).

[2 marks]

80 matches were played in a football tournament. The following table shows the number of goals scored in all matches.

Number of goals	0	1	2	3	4	5
Number of matches	16	22	19	17	1	5

19a. Find the mean number of goals scored per match.

[2 marks]

19b. Find the median number of goals scored per match.

[2 marks]

19c. A local newspaper claims that the mean number of goals scored per match is two. Calculate the percentage error in the local newspaper's claim.

[2 marks]

Part A

A university required all Science students to study one language for one year. A survey was carried out at the university amongst the 150 Science students. These students all studied one of either French, Spanish or Russian. The results of the survey are shown below.

	French	Spanish	Russian
Female	9	29	12
Male	31	40	29

Ludmila decides to use the χ^2 test at the 5% level of significance to determine whether the choice of language is independent of gender.

20a. State Ludmila's null hypothesis. *[1 mark]*

20b. Write down the number of degrees of freedom. *[1 mark]*

20c. Find the expected frequency for the females studying Spanish. *[2 marks]*

20d. Use your graphic display calculator to find the χ^2 test statistic for this data. *[2 marks]*

20e. State whether Ludmila accepts the null hypothesis. Give a reason for your answer. *[2 marks]*

At the end of the year, only seven of the female Science students sat examinations in Science and French. The marks for these seven students are shown in the following table.

Science (S)	23	51	56	62	12	73	72
French (F)	65	45	45	40	70	36	30

20f. Draw a labelled scatter diagram for this data. Use a scale of 2 cm to represent 10 marks on the x-axis (S) and 10 marks on the y-axis (F). *[4 marks]*

20g. Use your graphic calculator to find *[2 marks]*

(i) $\bar{S}$, the mean of S;

(ii) $\bar{F}$, the mean of F.

20h. Plot the point $M(\bar{S}, \bar{F})$ on your scatter diagram. *[1 mark]*

20i. Use your graphic display calculator to find the equation of the regression line of F on S. [2 marks]

20j. Draw the regression line on your scatter diagram. [2 marks]

20k. Carletta's mark on the Science examination was 44. She did not sit the French examination. [2 marks]

Estimate Carletta's mark for the French examination.

20l. Monique's mark on the Science examination was 85. She did not sit the French examination. Her French teacher wants to use the regression line to estimate Monique's mark. [2 marks]

State whether the mark obtained from the regression line for Monique's French examination is reliable. Justify your answer.