Spearman's *[41 marks]*

The Malvern Aquatic Center hosted a 3 metre spring board diving event. The judges, Stan and Minsun awarded 8 competitors a score out of 10. The raw data is collated in the following table.

Competitors	A	B	C	D	E	F	G	H
Stan's score (x)	4.1	3	4.3	6	7.1	6	7.5	6
Minsun's score (y)	4.7	4.6	4.8	7.2	7.8	9	9.5	7.2

1a. Write down the value of the Pearson's product–moment correlation coefficient, r. *[2 marks]*

1b. Using the value of r, interpret the relationship between Stan's score and Minsun's score. *[2 marks]*

1c. Write down the equation of the regression line y on x. *[2 marks]*

1d. Use your regression equation from part (b) to estimate Minsun's score when Stan awards a perfect 10. *[2 marks]*

1e. State whether this estimate is reliable. Justify your answer. *[2 marks]*

The Commissioner for the event would like to find the Spearman's rank correlation coefficient.

1f. **Copy** and complete the information in the following table. *[2 marks]*

Competitors	A	B	C	D	E	F	G	H
Stan's Rank		8					1	4
Minsun's Rank		8					1	4.5

1g. Find the value of the Spearman's rank correlation coefficient, r_s. *[2 marks]*

1h. Comment on the result obtained for r_s. *[2 marks]*

1i. The Commissioner believes Minsun's score for competitor G is too high and so decreases the score from 9.5 to 9.1. *[1 mark]*

Explain why the value of the Spearman's rank correlation coefficient r_s does not change.

Kayla wants to measure the extent to which two judges in a gymnastics competition are in agreement. Each judge has ranked the seven competitors, as shown in the table, where 1 is the highest ranking and 7 is the lowest.

Competitor	*A*	*B*	*C*	*D*	*E*	*F*	*G*
Judge 1	1	2	3	3	5	6	6
Judge 2	2	3	1	4	5	5	7

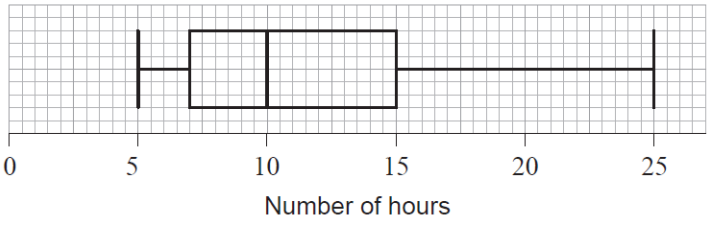
2a. Calculate Spearman's rank correlation coefficient for this data. *[5 marks]*

2b. State what conclusion Kayla can make from the answer in part (a). *[1 mark]*

As part of his mathematics exploration about classic books, Jason investigated the time taken by students in his school to read the book *The Old Man and the Sea*. He collected his data by stopping and asking students in the school corridor, until he reached his target of 10 students from **each** of the literature classes in his school.

3a. State which of the two sampling methods, systematic or quota, Jason has used. *[1 mark]*

Jason constructed the following box and whisker diagram to show the number of hours students in the sample took to read this book.



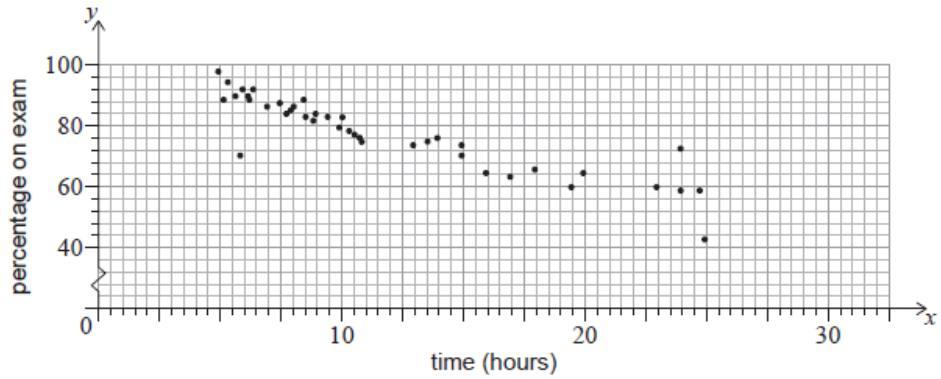
3b. Write down the median time to read the book. *[1 mark]*

3c. Calculate the interquartile range. *[2 marks]*

Mackenzie, a member of the sample, took 25 hours to read the novel. Jason believes Mackenzie's time is not an outlier.

3d. Determine whether Jason is correct. Support your reasoning. *[4 marks]*

For each student interviewed, Jason recorded the time taken to read *The Old Man and the Sea* (x), measured in hours, and paired this with their percentage score on the final exam (y). These data are represented on the scatter diagram.



3e. Describe the correlation. *[1 mark]*

Jason correctly calculates the equation of the regression line y on x for these students to be

$$y = -1.54x + 98.8.$$

He uses the equation to estimate the percentage score on the final exam for a student who read the book in 1.5 hours.

3f. Find the percentage score calculated by Jason. *[2 marks]*

3g. State whether it is valid to use the regression line y on x for Jason's estimate. Give a reason for your answer. *[2 marks]*

Jason found a website that rated the 'top 50' classic books. He randomly chose eight of these classic books and recorded the number of pages. For example, Book H is rated $44\mathrm{th}$ and has 281 pages. These data are shown in the table.

Book	A	B	C	D	E	F	G	H
Number of pages (n)	4215	863	585	1225	366	209	624	281
Top 50 rating (r)	1	2	5	7	13	22	40	44

Jason intends to analyse the data using Spearman's rank correlation coefficient, r_s .

3h. Copy and complete the information in the following table. *[2 marks]*

Book	A	B	C	D	E	F	G	H
Rank – Number of pages	1							
Rank – Top 50 Rating	1							

3i. Calculate the value of r_s. *[2 marks]*

3j. Interpret your result. *[1 mark]*