



GuideMyGrades

— YOUR GOALS. OUR EXPERTISE. —

EDEXCEL INTERNATIONAL A LEVEL MATHS: QUESTIONS BY TOPIC

Correlation (S3.5)

International A Level Mathematics · S3: Statistics 3 (WST03)

13 questions · 131 marks · Mark schemes included

Past papers October 2020 - June 2025 · 1 sub-topic

 guidemygrades.com

 [@Guidemygrades](https://www.instagram.com/Guidemygrades)

 [@Guidemygrades](https://twitter.com/Guidemygrades)

 [@Guidemygrades](https://www.youtube.com/Guidemygrades)

 info@guidemygrades.com

 Piccadilly Business Centre, Unit C, Aldow Enterprise Park
Blackett Street, Manchester, M12 6AE

HOW TO USE THIS BOOKLET

Questions are grouped by sub-topic and run from the lowest to the highest mark tariff, then oldest paper first.

Click any row below to jump straight to that sub-topic or question.

On every question page, click ANSWER in the top-right corner to jump to its mark scheme; < QUESTION brings you back.

A gold bar in the left margin marks exactly where the question sits on the original page.

Diagrams, graphs, grids and tables are printed exactly as they appear in the original papers.

Each question is filed under its main topic; longer questions may also test skills from other topics.

QUESTION	SOURCE	MARKS	PAGE
S3.5a	SPEARMAN'S RANK & PRODUCT MOMENT CORRELATION TESTS	13 Qs · 131 marks	2
1. Q2	IAL S3 (WST03) June 2025	7	3
2. Q3	IAL S3 (WST03) (unused paper) January 2022	8	5
3. Q1	IAL S3 (WST03) January 2025	8	7
4. Q2	IAL S3 (WST03) January 2021	9	11
5. Q1	IAL S3 (WST03) June 2021	9	13
6. Q1	IAL S3 (WST03) June 2023	9	15
7. Q1	IAL S3 (WST03) June 2022	10	19
8. Q2	IAL S3 (WST03) June 2024	10	23
9. Q3	IAL S3 (WST03) October 2020	11	27
10. Q3	IAL S3 (WST03) January 2022	12	31
11. Q2	IAL S3 (WST03) January 2023	12	35
12. Q3	IAL S3 (WST03) January 2024	12	39
13. Q3	IAL S3 (WST03) October 2021	14	43



QUESTIONS

13 questions · 131 marks · 1 sub-topic

SUB-TOPIC S3.5a

SPEARMAN'S RANK & PRODUCT MOMENT CORRELATION TESTS

Sub-topic S3.5a · 13 questions · 131 marks · lowest tariff first



2. At a local talent show 2 judges were asked to rank 10 performances, in order of preference. The results are shown in the following table.

Performance	A	B	C	D	E	F	G	H	I	J
Judge 1	1	3	5	7	9	10	2	6	8	4
Judge 2	2	1	3	7	9	4	5	8	10	6

- (a) Calculate Spearman's rank correlation coefficient for these data. Show your working clearly.

(3)

- (b) Stating your hypotheses clearly, test at the 1% level of significance, whether or not there is evidence of a positive correlation between the ranks of Judge 1 and the ranks of Judge 2

(4)



Question 2 continued

(Total for Question 2 is 7 marks)



Question 3 continued

Leave
blank

Lined area for writing the answer to Question 3.

(Total 8 marks)

Q3

- The table also shows some of the average match day attendance ranks.

Team	Final position	Average match day attendance	Average match day attendance rank
Arsenal	5	59 776	2
Brighton and Hove Albion	9	30 988	
Chelsea	3	36 424	
Leicester City	8	32 440	
Liverpool	2	53 027	
Manchester City	1	53 739	
Manchester United	6	72 992	1
Tottenham Hotspur	4	56 523	
West Ham United	7	57 915	
Wolverhampton Wanderers	10	30 725	

(Source from: www.statista.com)

By completing the ranks for average match day attendance,

- (a) calculate the value of the Spearman's rank correlation coefficient for these data.

Show your working clearly.

(4)

- (b) Stating your hypotheses clearly, test at the 1% level of significance, whether or not

team in the English Premier League and its average match day attendance.

(4)





Question 1 continued

(Total for Question 1 is 8 marks)



- | Student | A | B | C | D | E | F | G | H | I | J | K |
|-----------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| No. of objects | 8 | 11 | 9 | 15 | 17 | 6 | 10 | 14 | 12 | 13 | 5 |
| % in maths test | 30 | 62 | 57 | 80 | 75 | 43 | 65 | 51 | 48 | 55 | 32 |

- (5)

- (3)

(1)

Question 2 continued

Leave
blank

Q2

(Total 9 marks)



Question 1 continued

Lined area for writing the answer to Question 1.

(Total for Question 1 is 9 marks)



Question 1 continued



Question 1 continued

(Total for Question 1 is 9 marks)



1. The table below shows the number of televised tournaments won and the total number of tournaments won by the top 10 ranked darts players in 2020

Player's rank	Televised tournaments won	Total tournaments won
1	55	135
2	7	33
3	5	17
4	2	14
5	4	9
6	2	5
7	9	36
8	0	15
9	3	3
10	0	13

Michael did not want to calculate Spearman's rank correlation coefficient between player's rank and the rank in televised tournaments won because there would be tied ranks.

- (a) Explain how Michael could have dealt with these tied ranks.

(1)

Given that the largest number of total tournaments won is ranked number 1

- (b) calculate the value of Spearman's rank correlation coefficient between player's rank and the rank in total tournaments won.

(4)

- (c) Stating your hypotheses and critical value clearly, test at the 5% level of significance, whether or not there is evidence of a positive correlation between player's rank and the rank in total tournaments won for these darts players.

(4)

Michael does not believe that there is a positive correlation between player's rank and the rank in total number of tournaments won.

- (d) Find the largest level of significance, that is given in the tables provided, which could be used to support Michael's claim.
You must state your critical value.

(1)







Question 1 continued

(Total for Question 1 is 10 marks)







Question 2 continued

(Total for Question 2 is 10 marks)



Question 3 continued

Leave
blank

Question 3 continued

Leave
blank

Question 3 continued

Leave
blank

(Total 11 marks)

Q3



3. A medical research team carried out an investigation into the metabolic rate, MR, of men aged between 30 years and 60 years.

A random sample of 10 men was taken from this age group.

The table below shows for each man his MR and his body mass index, BMI. The table also shows the rank for the level of daily physical activity, DPA, which was assessed by the medical research team.

Rank 1 was assigned to the man with the highest level of daily physical activity.

Man	A	B	C	D	E	F	G	H	I	J
MR (x)	6.24	5.94	6.83	6.53	6.31	7.44	7.32	8.70	7.88	7.78
BMI (y)	19.6	19.2	23.6	21.4	20.2	20.8	22.9	25.5	23.3	25.1
DPA rank	10	7	9	8	6	3	1	4	5	2

[You may use $S_{xy} = 15.1608$ $S_{xx} = 6.90181$ $S_{yy} = 45.304$]

- (a) Calculate the value of the product moment correlation coefficient between MR and BMI for these 10 men. (2)
- (b) Use your value of the product moment correlation coefficient to test, at the 5% significance level, whether or not there is evidence of a positive correlation between MR and BMI.
State your hypotheses clearly. (3)
- (c) State an assumption that must be made to carry out the test in part (b). (1)
- (d) Calculate the value of Spearman's rank correlation coefficient between MR and DPA for these 10 men. (4)
- (e) Use a two-tailed test and a 5% level of significance to assess whether or not there is evidence of a correlation between MR and DPA. (2)

Question 3 continued





(Total for Question 3 is 12 marks)



- 2 The table shows the season's best times, x seconds, for the 8 athletes who took part in the 200 m final in the 2021 Tokyo Olympics. It also shows their finishing position in the race.

Athlete	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>H</i>
Season's best time	19.89	19.83	19.74	19.84	19.91	19.99	20.13	20.10
Finishing position	1	2	3	4	5	6	7	8

Given that the fastest season's best time is ranked number 1

- (a) calculate the value of the Spearman's rank correlation coefficient for these data. (4)
- (b) Stating your hypotheses clearly, test, at the 1% level of significance, whether or not there is evidence of a positive correlation between the rank of the season's best time and the finishing position for these athletes. (4)

Chris suggests that it would be better to use the actual finishing time, y seconds, of these athletes rather than their finishing position.

Given that

$$S_{xx} = 0.1286875 \quad S_{yy} = 0.55275 \quad S_{xy} = 0.225175$$

- (c) calculate the product moment correlation coefficient between the season's best time and the finishing time for these athletes.
Give your answer correct to 3 decimal places. (2)
- (d) Use your value of the product moment correlation coefficient to test, at the 1% level of significance, whether or not there is evidence of a positive correlation between the season's best time and the finishing time for these athletes. (2)

[illegible]

This image shows a single sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.



This image shows a single sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.



(Total for Question 2 is 12 marks)



3. The table shows the annual tea consumption, t (kg/person), and population, p (millions), for a random sample of 7 European countries.

Country	A	B	C	D	E	F	G
Annual tea consumption, t (kg/person)	0.27	0.15	0.42	0.06	1.94	0.78	0.44
Population, p (millions)	5.4	5.8	9	10.2	67.9	17.1	8.7

(You may use $S_{tt} = 2.486$ $S_{pp} = 3026.234$ $S_{pt} = 83.634$)

Angela suggests using the product moment correlation coefficient to calculate the correlation between annual tea consumption and population.

- (a) Use Angela's suggestion to test, at the 5% level of significance, whether or not there is evidence of any correlation between annual tea consumption and population. State your hypotheses clearly and the critical value used.

(5)

Johan suggests using Spearman's rank correlation coefficient to calculate the correlation between the rank of annual tea consumption and the rank of population.

- (b) Calculate Spearman's rank correlation coefficient between the rank of annual tea consumption and the rank of population.

(4)

- (c) Use Johan's suggestion to test, at the 5% level of significance, whether or not there is evidence of a positive correlation between annual tea consumption and population. State your hypotheses clearly and the critical value used.

(3)

Question 3 continued





Question 3 continued

(Total for Question 3 is 12 marks)



Leave
blank

3. A cafe owner wishes to know whether the price of strawberry jam is related to the taste of the jam. He finds a website that lists the price per 100 grams and a mark for the taste, out of 100, awarded by a judge, for 9 different strawberry jams A, B, C, D, E, F, G, H and I . He then ranks the marks for taste and the prices.

The ranks are shown in the table below.

Rank	1	2	3	4	5	6	7	8	9
Price	A	B	E	C	D	F	G	H	I
Taste	A	B	F	E	H	G	I	C	D

- (a) Calculate Spearman's rank correlation coefficient for these data. (5)
- (b) Test, at the 5% level of significance, whether or not there is a relationship between the price and the taste of these strawberry jams. State your hypotheses clearly. (3)

A friend suggests that it would be better to use the price per 100 grams, c , and the mark for the taste, m , for each strawberry jam rather than rank them.

Given that

$$S_{cc} = 2.0455 \quad S_{mm} = 243.5556 \quad S_{cm} = 16.4943$$

- (c) calculate the product moment correlation coefficient between the price and the mark for taste of these strawberry jams, giving your answer correct to 3 decimal places. (2)
- (d) Use your value of the product moment correlation coefficient to test, at the 5% level of significance, whether or not there is evidence of a positive correlation between the price and the mark for taste of these 9 strawberry jams. State your hypotheses clearly. (3)
- (e) State which of the tests in parts (b) and (d) is more appropriate for the cafe owner to use. Give a reason for your answer. (1)

Question 3 continued

Leave
blank



Question 3 continued

Leave
blank



Question 3 continued

Leave
blank

(Total 14 marks)

Q3



MARK SCHEMES

official mark schemes, in the same order as the questions



Question Number	Scheme		Marks
2 (a)	$\sum d^2 = 1+4+4+0+0+36+9+4+4+4 [= 66]$		M1
	$r_s = 1 - \frac{6 \times '66'}{10(10^2 - 1)} = 0.6$		M1 A1
			(3)
(b)	$H_0 : \rho = 0 \quad H_1 : \rho > 0$		B1
	Critical Value $r_s = 0.7455$ or CR: $r_s \dots 0.7455$		B1
	e.g. Do not reject H_0 or e.g. not significant or e.g. does not lie in the critical region		M1
	e.g. there is insufficient evidence of a positive correlation between the ranks of Judge 1 and Judge 2		A1ft
			(4)
	Notes		Total 7
(a)	M1	For finding the difference between ranks of judge 1 and judge 2 and evaluating $\sum d^2$. May be implied by 66. (Allow one error). May be implied by 0.6 or 0.4	
	M1	For using $1 - \frac{6 \sum d^2}{10(99)}$ with their $\sum d^2$ (this value cannot just be an attempt at adding the differences). May be implied by 0.6	
	A1	0.6 oe (a correct answer scores 3/3)	
(b)	B1	For both hypotheses correct. Must be in terms of ρ or ρ_s (condone if it appears as p) Must be attached to H_0 and H_1 (do not accept h for H) Do not allow hypotheses in words on their own.	
	B1	For CV of 0.7455 Allow $\pm$ and do not be concerned by use of a strict or inclusive inequality if a critical region is stated	
	M1	A correct statement independent of hypotheses ft part (a) $ r_s < 1$ and their CV where $ CV < 1$ with compatible signs. No context needed but do not allow contradicting non contextual statements. Condone e.g. “accept H_0 ”, “reject H_1 ” lies in acceptance region. A correct contextual conclusion may imply this mark.	
	A1ft	Dependent on previous method mark and $ r_s < CV < 1$. For a correct conclusion suggesting that there is insufficient evidence of a positive correlation between the ranks of the judges. Must mention <u>positive correlation</u> , <u>ranks</u> and the <u>judges</u> oe Condone e.g. “there isn’t positive correlation between the ranks of the judges”. Follow through their r_s with 0.7455 Note that M0A1ft is not possible. Note that a correct contextual conclusion on its own scores M1A1ft	
	Note	Two-tailed test Applying a two-tailed test scores a maximum of B0B0M1A0 allowing access to the M1 mark only	

Question Number	Scheme											Marks
3 (a)	Driver	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>H</i>	<i>I</i>	<i>J</i>	M1
	Rank FQL	1	5	3	2	6	4	8	9	10	7	
	FP	1	2	3	4	5	6	7	8	9	10	
	$\sum d^2 = 0 + 9 + 0 + 4 + 1 + 4 + 1 + 1 + 1 + 9 = 30$											M1
	$r_s = 1 - \frac{6(30)}{10(99)}$											dM1
	$= 0.8181818\dots$ awrt 0.818											A1
												(4)
(b)	$H_0: \rho = 0, H_1: \rho > 0$											B1
	Critical Value $r_s = 0.7455$ or CR: $r_s \dots 0.7455$											B1
	Reject H_0 or significant or lies in the critical region											M1
	There is sufficient evidence of a positive correlation between fastest qualifying lap time and finishing position for these Formula One racing drivers											A1
												(4)
	Notes											Total 8
3 (a)	M1	attempt to rank fastest qualifying lap (at least four correct).										
	M1	finding the difference between each of the ranks and evaluating $\sum d^2$										
	dM1	dependent on 1 st M1. Using $1 - \frac{6 \sum d^2}{10(99)}$ with their $\sum d^2$										
	A1	$\frac{9}{11}$ or awrt 0.818										
(b)	B1	both hypotheses correct. Must be in terms of ρ . Must be attached to H_0 and H_1										
	B1	critical value of 0.7455										
	M1	A correct statement comparing their CV with their r_s - no context needed but do not allow contradicting non contextual comments.										
	A1	correct conclusion which is rejecting H_0 , which must mention lap time and finishing position .										

Question Number	Scheme		Marks
1 (a)	Rankings 2, 9, 7, 8, 6, 5, 1, 4, 3, 10		B1
	$\sum d^2 = 9 + 0 + 16 + 0 + 16 + 16 + 25 + 0 + 16 + 0 [= 98]$		M1
	$r_s = 1 - \frac{6 \times '98'}{10(10^2 - 1)} = 0.4060...$	awrt 0.406	M1 A1
			(4)
(b)	$H_0 : \rho = 0 \quad H_1 : \rho > 0$		B1
	Critical Value $r_s = 0.7455$ or CR: $r_s \dots 0.7455$		B1
	Not in the critical region/not significant/Do not reject H_0		M1
	There is insufficient evidence of a positive correlation between the final position of a football team in the English Premier League and their average match day attendance .		A1ft
		(4)	
Notes			Total 8
(a)	B1	For all 8 correct missing rankings. If in the table and in the working space and different then award the highest scoring response.	
	M1	For an attempt at $\sum d^2$ (at least 5 correct values seen, with an attempt to add) May be implied by 98	
	M1	For using $1 - \frac{6\sum d^2}{10(99)}$ with their $\sum d^2$ (you will need to check their $\sum d^2$ if no value shown)	
	A1	awrt 0.406 Allow $\frac{67}{165}$ NB awrt 0.406 or $\frac{67}{165}$ scores 4/4	
(b)	B1	For both hypotheses correct. Must be in terms of ρ or ρ_s (Condone p). Must be attached to H_0 and H_1	
	B1	For CV of 0.7455	
	M1	A correct statement ft part (a) and their CV– no context needed but do not allow contradicting non contextual statements. This may be implied by a correct contextual conclusion on its own.	
	A1ft	Correct conclusion in context. Must mention words in bold oe, ft their r in part (a) and their critical value.	

Question Number	Scheme	Marks																																				
2. (a)	<table><tr><td>Student</td><td>A</td><td>B</td><td>C</td><td>D</td><td>E</td><td>F</td><td>G</td><td>H</td><td>I</td><td>J</td><td>K</td></tr><tr><td>Objects rank</td><td>9</td><td>6</td><td>8</td><td>2</td><td>1</td><td>10</td><td>7</td><td>3</td><td>5</td><td>4</td><td>11</td></tr><tr><td>Maths rank</td><td>11</td><td>4</td><td>5</td><td>1</td><td>2</td><td>9</td><td>3</td><td>7</td><td>8</td><td>6</td><td>10</td></tr></table>	Student	A	B	C	D	E	F	G	H	I	J	K	Objects rank	9	6	8	2	1	10	7	3	5	4	11	Maths rank	11	4	5	1	2	9	3	7	8	6	10	M1
	Student	A	B	C	D	E	F	G	H	I	J	K																										
	Objects rank	9	6	8	2	1	10	7	3	5	4	11																										
	Maths rank	11	4	5	1	2	9	3	7	8	6	10																										
		M1																																				
	$\sum d^2 = 4 + 4 + 9 + 1 + 1 + 1 + 16 + 16 + 9 + 4 + 1 = 66$	M1																																				
	$r_s = 1 - \frac{6 \times "66"}{11(11^2 - 1)} \quad ; = \underline{0.7}$	dM1; A1																																				
		(5)																																				
	(b) $H_0 : \rho = 0 \quad H_1 : \rho > 0$ Critical value ($n = 11$ 5% one-tail) is 0.5364 (Significant result so) there is evidence to support the teacher's belief <u>or</u> there is evidence of a positive correlation between short term memory and mathematical ability (o.e.) <u>or</u> evidence that students with strong maths ability also have good short term memory (o.e.)	B1 B1 B1																																				
		(3)																																				
(c) Data shows positive correlation but does not necessarily imply that enhanced short term memory <u>causes</u> increase in mathematical ability.	B1																																					
	(1)																																					
	[9 marks]																																					
Notes																																						
(a)	1 st M1 for attempt to rank one row with at least 5 correct (could be reversed) 2 nd M1 for both rows ranked with at least 5 correct in each row (one or both reversed) 3 rd M1 for an attempt at $\sum d^2$ ft their values and at least 5 correct 4 th dM1 (dep on at least one M1) for use of their $\sum d^2$ in a correct formula A1 for 0.7 or exact equivalent																																					
(b)	1 st B1 for both hypotheses in terms of ρ or ρ_s [If $r_s < 0$ in (a) allow $H_1 : \rho < 0$] 2 nd B1 for critical value of 0.5364 (sign compatible with r_s) [If $r_s < 0$ in (a) need -0.5364] Allow 0.6182 if 1 st B0 for $H_1 : \rho \neq 0$ 3 rd B1 for correct conclusion in context. Penalise contradictory comments e.g. "not significant so supports teacher's belief" [No ft]																																					
(c)	B1 for a comment that states that correlation does <u>not</u> imply <u>causation</u> Need to see "cause" or "causation" clearly mentioned.																																					

Question Number	Scheme	Notes	Marks																														
1(a)	<table><tr><td></td><td>A</td><td>B</td><td>C</td><td>D</td><td>E</td></tr><tr><td>C</td><td>3</td><td>2</td><td>1</td><td>5</td><td>4</td></tr><tr><td>M</td><td>2</td><td>4</td><td>1</td><td>5</td><td>3</td></tr><tr><td>d</td><td>1</td><td>-2</td><td>0</td><td>0</td><td>1</td></tr><tr><td>d²</td><td>1</td><td>4</td><td>0</td><td>0</td><td>1</td></tr></table>		A	B	C	D	E	C	3	2	1	5	4	M	2	4	1	5	3	d	1	-2	0	0	1	d ²	1	4	0	0	1	Attempt to rank at least 1 row with at least 3 correct. Allow reverse rankings.	M1
			A	B	C	D	E																										
		C	3	2	1	5	4																										
		M	2	4	1	5	3																										
		d	1	-2	0	0	1																										
d ²	1	4	0	0	1																												
Attempt at d^2 for their rankings , can be implied by $\sum d^2 = 6$	M1																																
$\sum d^2 = 6$	Can be implied by correct answer. Must come from correct rankings.	A1																															
$r_s = 1 - \frac{6(6)}{5(24)}$	(dep on 1 st M1) Use of correct formula with their $\sum d^2$	dM1																															
$r_s = 0.7$	0.7 o.e. Must come from correct rankings.	A1																															
			(5)																														
1(b)	$H_0 : \rho = 0, H_1 : \rho > 0$	Both correct in terms of ρ or ρ_s . Must be compatible with their ranking.	B1																														
	cv 0.9 or cr $r_s \geq 0.9$	0.9 sign should match H_1 or their r_s	B1																														
	$r_s = 0.7$ does not lie in cr so do not reject H_0	Correct non-contextual statement e.g. “do not reject H_0 ”, “not in critical region”, “not significant”, “no positive correlation”. $ \text{test value} $ or $ \text{cv} > 1$ award M0	M1																														
	Data does not support plant biologist’s claim .	Correct conclusion in context. Must mention “biologist’s claim” o.e. or moisture and plant coverage . All previous marks in (b) must have been scored.	A1ft																														
SC	For use of two-tailed test: May score B0B1M1A0 for cv = 1(.000) and ‘not significant’ oe																																
			(4)																														
			Total 9																														

Question Number	Scheme		Marks
1 (a)	When the data is ordinal e.g. Judges' ranks		B1
	When a non-linear relationship might be expected		B1
			(2)
(b)	$H_0 : \rho = 0$, $H_1 : \rho \neq 0$		B1
	Critical value $r_s = -0.6485$ or CR: $r_s \leq -0.6485$ (and $r_s \geq 0.6485$)		B1
	Reject H_0 or significant or lies in the critical region		M1
	The Spearman's rank correlation coefficient shows there is sufficient evidence of a correlation [between the length and maximum diameter of the melons]		A1
			(4)
(c)	$H_0 : \rho = 0$, $H_1 : \rho < 0$		B1
	Critical value $r = -0.5494$ or CR: $r \leq -0.5494$		B1
	The product moment correlation coefficient shows there is insufficient evidence of a negative correlation [between the length and maximum diameter of the melons]		B1
			(3)
	Notes		Total 9
(a)	B1	For one correct condition	
	B1	For a second correct condition. Condone not underlying normal	
(b)	B1	For both hypotheses correct. Must be in terms of ρ . Must be attached to H_0 and H_1	
	B1	For critical value of -0.6485 (Allow -0.5636 if a one tailed test is stated for H_1) Condone 0.6485 if compared with 0.673	
	M1	A correct statement – no context needed but do not allow contradicting non contextual comments. ft their CV provided the CV is negative (May be implied by a correct conclusion) Condone a positive CV if a comparison with 0.673 seen	
	A1	For a correct conclusion which is rejecting H_0 Allow negative correlation This mark is independent of the hypotheses	
(c)	B1	For both hypotheses correct. Must be in terms of ρ . Must be attached to H_0 and H_1	
	B1	For critical value of -0.5494 (Allow -0.6319 if a two tailed test is stated for H_1) Condone 0.5494 if compared with 0.525	
	B1	For a correct conclusion which is not rejecting H_0	

Question Number	Scheme		Marks
1 (a)	You would assign an average rank between the tied ranks		B1
			(1)
(b)	Rank for total tournaments 1 3 4 6 8 9 2 5 10 7		M1
	$\sum d^2 = 0 + 1 + 1 + 4 + 9 + 9 + 25 + 9 + 1 + 9 [= 68]$		M1
	$r_s = 1 - \frac{6 \times '68'}{10(10^2 - 1)}$		dM1
	$= 0.5878...$ awrt 0.588		A1
			(4)
(c)	$H_0: \rho = 0$, $H_1: \rho > 0$		B1
	Critical Value = 0.5636 or CR ... 0.5636		B1
	Reject H_0 or significant or lies in the critical region		dM1
	There is sufficient evidence of a positive correlation between rank and total tournaments won		A1
			(4)
(d)	2.5% and $r_s = 0.6485$ or CR ... 0.6485		B1
			(1)
	Notes		Total 10
(a)	B1	for an appropriate explanation of how to deal with tied ranks. Ignore any comments regarding PMCC Do not allow add 0.5 to both ranks	
(b)	M1	attempt to rank total tournaments (at least four correct) Condone reversed ranks	
	M1	finding the difference between players rank and each of their total tournaments ranks and evaluating $\sum d^2$ May be implied by 68	
	dM1	dependent on 1 st M1. Using $1 - \frac{6 \sum d^2}{10(99)}$ with their $\sum d^2$ (you will need to check their $\sum d^2$ if no value shown)	
	A1	awrt 0.588 Allow $\frac{97}{165}$	
(c)	B1	both hypotheses correct. Must be in terms of ρ . Must be attached to H_0 and H_1 If r_s is negative in part (b) then allow $H_1: \rho < 0$	
	B1	critical value of 0.5636 If r_s is negative in part (b) then allow -0.5636	
	dM1	dependent on 2 nd B1. A correct statement ft their part (b) and their CV– no context needed but do not allow contradicting non contextual comments. This may be implied by a correct contextual conclusion.	
	A1	correct conclusion which is rejecting H_0 , which must mention rank and total tournaments . No hypotheses is A0.	
		NB If they have used $H_1: \rho < 0$ then the maximum they can score is B1B1dM1A0	
(d)	B1	for 2.5% and a correct critical value of 0.6485	

Question Number	Scheme									Marks
2 (a)		Kettle	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	M1
		Price (£)	99.99	14.99	34.97	49.99	19.97	29.99	8.99	
		Aarush Rank	2	5	4	1	7	3	6	
		Actual rank	1	6	3	2	5	4	7	
	$\sum d^2 = 1+1+1+1+4+1+1 [=10]$									M1
	$r_s = 1 - \frac{6 \times "10"}{7 \times 48}$									dM1
	$= 0.8214...$ awrt 0.821									A1cso
										(4)
(b)	$H_0 : \rho = 0 \quad H_1 : \rho > 0$									B1
	Critical Value $r_s = 0.7143$									M1
	Either reject H_0 / result is significant/ r_s does lie in the critical region									M1
	Conclude there is evidence to support that Aarush can rank the kettles in order of price .									A1
										(4)
(c)	Aarush has already ranked them or the order / price is not normally distributed									B1
										(1)
(d)	Kettle <i>A</i> and <i>D</i> would have a tied rank (1.5)									B1
										(1)
	Notes									Total 10
(a)	M1 for an attempt to rank Aarush's order and actual order (at least 4 correct in either row)									
	M1 for finding differences between each pair of their ranks and evaluating $\sum d^2$ imp by A1									
	dM1 dependent on previous 2 M marks being awarded. Using $r_s = 1 - \frac{6 \times \sum d^2}{7 \times 48}$									
	A1 awrt 0.821 or allow $\frac{23}{28}$									
(b)	B1 both hypotheses stated in terms of ρ or ρ_s (condone if appears as p – it cannot be in terms of r) If their $r_s < 0$ condone $H_1 : \rho < 0$									
	M1 allow $ \text{their CV} = 0.7143$ or better									
	M1 for a correct non-contradictory statement consistent with their r_s ($r_s < 1$) and their CV if signs are compatible. Allow eg Accept H_1 or for correct conclusion in context on its own.									
	A1 dependent on all of the previous method marks but independent of hypotheses. For correct conclusion in context which must be rejecting H_0 and has rank , kettles and price and no contradictory statements.									
(c)	B1 for a correct reason e.g. (some of) the data has already been ranked, Aarush has not given the prices (only the ranks) oe. Condone “it is ordinal data”									
(d)	B1 for the idea that both <i>A</i> and <i>D</i> will have same rank (1.5) If the rank is stated it must be 1.5. Must mention both <i>A</i> and <i>D</i> (or imply) and allow explanations such as “equal rank”, “average rank”. Condone “order” for rank. If an incorrect rank or letter pairing is given then B0.									

Question Number	Scheme	Marks
3.		
(a)		
	</	

Question Number	Scheme		Marks
3 (a)	$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} = \frac{15.1608}{\sqrt{6.90181 \times 45.304}}$		M1
	$= 0.8573\dots$ awrt 0.857		A1
			(2)
(b)	$H_0 : \rho = 0$, $H_1 : \rho > 0$		B1
	Critical value 5% = 0.5494		B1
	Significant evidence to suggest that there is a <u>positive correlation</u> between <u>MR</u> and <u>BMI</u>		B1
			(3)
(c)	MR and BMI measurements are normally (or bivariate normal) distributed		B1
			(1)
(d)	Ranks for MR: 9 10 6 7 8 4 5 1 2 3		B1
	$\sum d^2 = 1 + 9 + 9 + 1 + 4 + 1 + 16 + 9 + 9 + 1 [= 60]$		M1
	$r_s = 1 - \frac{6(60)}{10(99)}$		M1
	$= 0.6363$ awrt ($\pm$) 0.636		A1
			(4)
(e)	$[H_0 : \rho = 0$, $H_1 : \rho \neq 0]$		
	Critical value 0.6485		B1
	There is insufficient evidence of a correlation between <u>MR</u> and <u>DPA</u>		B1
			(2)
	Notes		Total 12
(a)	M1	for use of $\frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$	
	A1	awrt 0.857	
(b)	B1	both hypotheses correct. Must be in terms of ρ . Must be attached to H_0 and H_1 Do not allow hypotheses in words on their own.	
	B1	critical value of 0.5494	
	B1	correct conclusion rejecting H_0 which must mention positive correlation, MR and BMI which must be consistent with their CV and their r , with $ \text{their CV} < 1$ and $ \text{their } r < 1$	
(c)	B1	correct assumption referring to MR and BMI needing to be normally distributed	
(d)	B1	attempt to rank MR (at least four correct) (may be implied by correct $\sum d^2$ or correct answer) allow reverse ranks for MR: 2 1 5 4 3 7 6 10 9 8	
	M1	for finding the difference between each of the ranks and evaluating $\sum d^2$ (implied by $\sum d^2 = 60$ or for reverse ranks $\sum d^2 = 270$)	
	M1	using $1 - \frac{6\sum d^2}{10(99)}$ with their $\sum d^2$	
	A1	awrt ($\pm$) 0.636	
(e)	B1	critical value of 0.6485 (or -0.6485 if $r_s < 0$)	
	B1	correct conclusion which is not rejecting H_0 , which must mention MR and DPA which must be consistent with their CV and their r_s , with $ \text{their CV} < 1$ and $ \text{their } r_s < 1$	

Question	Scheme										Marks
2 (a)	Athlete	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>H</i>	M1	
	Rank SBT	4	2	1	3	5	6	8	7		
	FP	1	2	3	4	5	6	7	8		
	$\sum d^2 = 9 + 0 + 4 + 1 + 0 + 0 + 1 + 1 [= 16]$										M1
	$r_s = 1 - \frac{6("16")}{8(63)} = 0.8095....$ awrt 0.81										dM1 A1
											(4)
(b)	$H_0 : \rho = 0 , H_1 : \rho > 0$										B1
	Critical Value $r_s = 0.8333$ or CR: $r_s \geq 0.8333$										B1
	Do not reject H_0 or not significant or does not lie in the critical region or there is no evidence of a positive correlation										M1
	There is no evidence of a positive correlation between season's best time and finishing position for these athletes										A1ft
											(4)
(c)	$r = \frac{0.225175}{\sqrt{0.1286875 \times 0.55275}}$										M1
	$= 0.84428...$ awrt 0.844										A1
											(2)
(d)	Critical Value $r = 0.7887$ or CR: $r \geq 0.7887$										M1
	so there is evidence of a positive correlation between season's best time and finishing time for these athletes										A1 ft
											(2)
Notes											Total 12
(a)	M1	attempt to rank seasonal best time (at least four correct), May be implied by $\sum d^2 = 16$									
	M1	Attempt to find the difference between each of the ranks (at least 3 correct) and evaluating $\sum d^2$ May be implied by awrt 0.81 NB if no ranks for SBT it is M0									
	dM1	dependent on 1 st M1. Using $1 - \frac{6 \sum d^2}{8(63)}$ with their $\sum d^2$									
	A1	$\frac{17}{21}$ or awrt 0.81(0)									
	SC	for reverse rankings May score M1M1dM1A0 order 5 7 8 6 4 3 1 2 $\sum d^2 = 158$									
(b)	B1	both hypotheses correct. Must be in terms of ρ (allow something that looks like rho eg p). Must be attached to H_0 and H_1									
	B1	critical value of 0.8333 Sign should match there H_1 or r_s									
	M1	correct statement comparing their CV with their r_s - no context needed but do not allow contradicting non contextual comments. If no CV or test statistic given or the test value or CV > 1 then it is M0									
	A1ft	correct conclusion in context for their value of r_s from (a) and their stated CV. Conclusion must refer to positive correlation , seasonal best or time and position .									
	SC	For use of two-tailed test: May score B0B1M1A0 CV allow 0.881...)									
(c)	M1	correct method used									
	A1	awrt 0.844									
(d)	M1	Critical value of 0.7887 Allow 0.8343 if hypotheses are two tailed in (b)									

	M1 must be awarded. A correct conclusion for their value of r from (c) Conclusion must refer to A1ft positive correlation, seasonal best or time and finishing time . Do not allow contradicting comments. if the test value or CV > 1 then it is M0
--	---



Question Number	Scheme	Marks																																																
3. (a)	$H_0 : \rho = 0$ $H_1 : \rho \neq 0$ $[r =] \frac{83.634}{\sqrt{2.486 \times 3026.234}} = 0.9642.....$ awrt 0.964 CV = 0.7545 [Reject H_0 /Significant] There is evidence of correlation between annual <u>tea consumption</u> and <u>population</u> .	B1 M1 A1 B1ft A1 (5)																																																
(b)	<table><tr><td>Country</td><td>A</td><td>B</td><td>C</td><td>D</td><td>E</td><td>F</td><td>G</td></tr><tr><td>T Rank</td><td>5</td><td>6</td><td>4</td><td>7</td><td>1</td><td>2</td><td>3</td></tr><tr><td>P Rank</td><td>7</td><td>6</td><td>4</td><td>3</td><td>1</td><td>2</td><td>5</td></tr></table> or <table><tr><td>Country</td><td>A</td><td>B</td><td>C</td><td>D</td><td>E</td><td>F</td><td>G</td></tr><tr><td>T Rank</td><td>3</td><td>2</td><td>4</td><td>1</td><td>7</td><td>6</td><td>5</td></tr><tr><td>P Rank</td><td>1</td><td>2</td><td>4</td><td>5</td><td>7</td><td>6</td><td>3</td></tr></table> $\sum d^2 = 4 + 0 + 0 + 16 + 0 + 4 [= 24]$ $[r_s =] 1 - \frac{6(24)}{7(48)} = 0.571428...$ awrt 0.571	Country	A	B	C	D	E	F	G	T Rank	5	6	4	7	1	2	3	P Rank	7	6	4	3	1	2	5	Country	A	B	C	D	E	F	G	T Rank	3	2	4	1	7	6	5	P Rank	1	2	4	5	7	6	3	M1 M1 dM1A1 (4)
Country	A	B	C	D	E	F	G																																											
T Rank	5	6	4	7	1	2	3																																											
P Rank	7	6	4	3	1	2	5																																											
Country	A	B	C	D	E	F	G																																											
T Rank	3	2	4	1	7	6	5																																											
P Rank	1	2	4	5	7	6	3																																											
(c)	$H_0 : \rho_s = 0$ $H_1 : \rho_s > 0$ CV = 0.7143 [Do not reject H_0 / not significant] There is not enough evidence to suggest a <u>positive</u> correlation between annual <u>tea consumption</u> and <u>population</u> .	B1 M1 A1ft (3) [12]																																																
Notes																																																		
(a)	1 st B1 both hypotheses correct in terms of ρ (must be two-tailed). Condone use of p M1 use of formula for r (May be implied by awrt 0.964) 1 st A1 awrt 0.964 2 nd B1ft 0.7545 (or better) or ft 1-tailed alternative hypothesis (0.6694) 2 nd A1 correct contextual conclusion including tea consumption/ t and population/ p . Must be consistent with their r and their CV. (Ignore any non-contextual conclusion) Allow positive correlation																																																	
(b)	1 st M1 attempt to rank each country for tea and population (at least 4 correct in each) 2 nd M1 for $\sum d^2$ for their ranks (implied by $\sum d^2 = 24$) 2 nd M1 (dep on 1 st M1) use of $1 - \frac{6('24')}{7(48)}$ A1 awrt 0.571 (or $\frac{4}{7}$)																																																	
(c)	B1 both hypotheses correct in terms of ρ or ρ_s . Condone use of p M1 0.7143 (or better) A1ft correct contextual conclusion including positive, tea consumption/ t and population/ p . (Ignore any non-contextual conclusion) ft their part (b)																																																	

Question Number	Scheme										Marks
3(a)	jam	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>H</i>	<i>I</i>	M1
	Price	1	2	4	5	3	6	7	8	9	
	Taste	1	2	8	9	4	3	6	5	7	
	$\sum d^2 = [0+0+16+16+1+9+1+9+4] = 56$										M1A1
	$r_s = 1 - \frac{6(56)}{9(80)} = \frac{8}{15} = 0.5333..$ awrt 0.533										dM1A1
											(5)
(b)	$H_0: \rho = 0, H_1: \rho \neq 0$										B1
	Critical Value = 0.7										B1
	There is no evidence of a relationship between <u>price</u> and <u>taste</u> of strawberry jam										B1ft
											(3)
(c)	$r = \frac{16.4943}{\sqrt{2.0455 \times 243.5556}}$										M1
	$= 0.7389...$ awrt 0.739										A1
											(2)
(d)	$H_0: \rho = 0, H_1: \rho > 0$										B1
	CV = 0.5822										B1
	There is evidence of a <u>positive correlation</u> between <u>price</u> and <u>taste</u> of strawberry jam										B1ft
											(3)
(e)	Spearman's rank as it is unlikely that a joint normal distribution applies.										B1
	or the marks are a judgement or the marks are not a meaningful scale.										
											(1)
	Notes										Total 14
(a)	M1	Attempt to rank each jar for taste and price. At least 4 pairs of ranks correct									
	M1	For an attempt at d^2 row for their ranks (may be implied by $\sum d^2 = 56$)									
	A1	$\sum d^2 = 56$									
	dM1	Dependent on the previous M being awarded. Using $1 - \frac{6\sum d^2}{9(80)}$									
(b)	A1	$\frac{8}{15}$ or awrt 0.533									
	B1	Both hypotheses stated in terms of ρ . Must be two-tail.									
	B1	0.7 for CV. Allow 0.6 if a one tail test is used									
	B1ft	For a correct contextualised comment which has price and taste Follow through their r_s with their 0.7 (provided $ \text{their } r_s < 1$)									
(c)	M1	Correct method used									
	A1	awrt 0.739									
(d)	B1	Both hypotheses stated in terms of ρ . Must be one-tail. If B0 awarded in part (b) then allow any letter instead of ρ that is consistent with part (b)									
	B1	0.5822 Allow 0.6664 if a two-tail test is used.									
	B1ft	Correct conclusion in context which has positive correlation (this may be implied by a correct description of positive correlation), price and taste. Follow through their 0.5822 and 0.739									
	B1	Selecting Spearman's with a suitable reason. Do not allow 'because it is ranked' as a suitable reason									
(e)	B1										