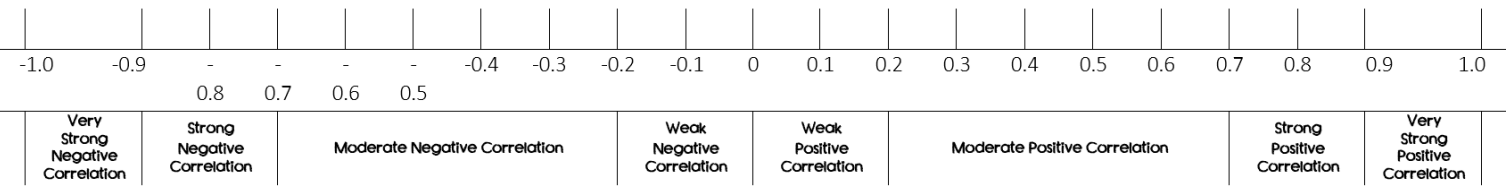# Key Vocab

Bivariate - Data which has two variables

Line of Best Fit - A straight line that passes through the centre of the plotted coordinates

Correlation - The way in which the data is related on a scatter graph

PMCC - Product Moment Correlation Coefficient

SRCC - Spearman's Rank Correlation Coefficient

| -1.0 | -0.9 | -0.8 | -0.7 | -0.6 | -0.5 | -0.4 | -0.3 | -0.2 | -0.1 | 0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1.0 |

| Very Strong Negative Correlation | Strong Negative Correlation | Moderate Negative Correlation | Weak Negative Correlation | Weak Positive Correlation | Moderate Positive Correlation | Strong Positive Correlation | Very Strong Positive Correlation |

Key Knowledge

PMCC

The PMCC is a value between 1 and + 1 (where 0 indicates no correlation)

- ➢ A negative value implies a negative correlation
- ➢ A positive value implies a positive correlation
- ➢ The closer the value is to zero, the weaker the correlation is

SRCC

$$1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

- Where d is the differences between the ranks
- Where n is the total number of pairs of data

Product Moment Correlation Coefficient	Spearman's Rank Correlation Coefficient
Both measure the strength of the correlation of 2 variables	
Both calculate a value from - 1 to + 1	
PMCC calculates the relationship between the 2 sets of data values	SRCC calculates the relationship between the rankings of 2 sets of data
PMCC can only be calculated for numerical data	SRCC can be calculated for both numerical and non-numerical data
You will not be expected to calculate PMCC, only interprets its value	You MUST rank the data before you calculate the SRCC
	You need to be able to calculate the SRCC using the formula given in the front of the exam book

It is common for SRCC to be stronger than PMCC as it is more likely for ranks to be equal than actual values

Correlation Coefficients

Example – I DO

A food critic was asked to compare six mince pies (labelled A to F) and to rank them in order of quality.

Jacques wants to see if the price of a mince pie depends on its quality.

The tables show information about these six mince pies.

Quality rank	Mince pie
1 (highest quality)	B
2	A
3	C
4	F
5	D
6 (lowest quality)	E

Price rank	Mince pie
1 (highest price)	C
2	B
3	A
4	F
5	E
6 (lowest price)	D

Jacques calculates Spearman's rank correlation coefficient for the quality ranks and the price ranks.

(a) Explain whether or not this is a sensible statistic for Jacques to calculate.

----

(2)

The value of Spearman's rank correlation coefficient calculated by Jacques is 0.77

(b) Based on this value, write down a conclusion that Jacques could reach. You must justify your answer.

Example – WE DO

In a television talent contest, 9 acts were given the following ranks by the judges and by a public telephone vote.

Act	Judges' rank	Public vote rank		
A	9	9		
B	3	2		
C	1	3		
D	4	5		
E	5	8		
F	8	6		
G	6	4		
H	2	1		
I	7	7		

(a) Use calculations to determine how much agreement there is between the judges and the public.

YOU DO a

There are 11 727 students at a university.
Their nationality is classified as UK, EU or International.

The table shows information about the nationality of these students.

Nationality	UK	EU	International	Total
Number of students	9393	979	1355	11 727

(Source: www.ox.ac.uk)

The manager of a book shop wants to carry out a survey into the books read by the students at this university.

Three students, Amy, Beth and Carlos, bought the same eight books from the book shop.

The manager asked Amy, Beth and Carlos to rank these eight books in increasing order of how much they each enjoyed them.

The manager calculated the Spearman's rank correlation coefficient for the ranks given by Amy and Beth.

She got a result of 1.2

(a) Explain how you know that this result is not correct.

----

(1)

The manager also calculated the Spearman's rank correlation coefficient for the ranks given by Beth and Carlos.

She got a result of 0.74

(b) (i) What type of correlation is shown by this result?

(ii) Interpret this result.

----

(2)

YOU DO b

Raina has been watching the judging of a cake baking competition.

Two judges ranked the 10 bakers for their sponge cakes.

Raina calculated the Spearman's rank correlation coefficient for the ranks given by the judges.

She got a value of 0.8

(a) (i) What type of correlation is shown by the value 0.8?
   Put a cross in one of the boxes below.

   Negative correlation ☐   **No correlation** ☐   **Positive correlation** ☐

(ii) Interpret Raina's value.

----

(2)

The same two judges will also be judging a flower-arranging competition.

(b) Is it possible to say anything about the ranks they are likely to give for flower arranging based on the value of Spearman's rank correlation coefficient that Raina calculated?

   Give a reason for your answer.

----

(1)

YOU DO c

In a town music competition, 6 groups competed against each other.

The table shows the marks awarded to each group by the invited independent judge.

The table also shows what the Mayor thought the rank order of the groups should be.
(Best group is given rank 1)

Group	Marks from judge	Mayor's rank			
Artex Monkeys	37	2			
Brevity	39	1			
Carfax	36	5			
Deft Ducks	29	4			
Extinct	27	6			
Flaming Friars	34	3			

Using suitable calculations, investigate how much agreement there is between the judge and the Mayor.

You may use the blank columns in the table for your working.

YOU DO d

The World Happiness Report 2017 is based on results obtained by surveying a sample of 1000 people in each of 155 countries.

In the report each country is given a happiness score based on the answers the sampled people give to a series of questions.

The table shows the happiness scores for the top 8 countries and the Gross Domestic Product (GDP) per capita of these countries.

For the sample data, the highest happiness score (Happiness rank 1) represents the country with the happiest people and the highest GDP represents the wealthiest country.

Country	GDP per capita (US Dollars)	Happiness score		Happiness rank	
Norway	70 392	7.537		1	
Denmark	53 744	7.522		2	
Iceland	59 629	7.504		3	
Switzerland	79 242	7.494		4	
Finland	43 169	7.469		5	
Netherlands	45 283	7.377		6	
Canada	42 210	7.316		7	
New Zealand	38 345	7.314		8	

(Source: www.worldhappiness.report)

(a) Calculate Spearman's rank correlation coefficient for these data.
   Give your answer correct to 2 decimal places.

----

(4)

(b) Describe and interpret the correlation found in part (a).

----

(2)

George wants to compare the happiness scores for these top 8 countries with average life expectancies for those countries.

He works out Spearman's rank correlation coefficient for these data and gets a result of 1.3

(c) Explain how you know that this result is not correct.

----

(1)

George also compared the happiness scores for these countries with two other variables.

George calculated Spearman's rank correlation coefficient in each case.

His values are –0.8 and –0.5

(d) Compare these two correlation coefficients.
   Give **two** comparisons.

----

(2)