AQA - Correlation and Regression – GCSE Statistics – 2021

1. June/2022/Paper_8382/1H/No.1

Two variables have no correlation.

Which of these could be the value of the Spearman's Rank Correlation Coefficient between the two variables?

Circle your answer.

[1 mark]

1 $\quad$ -1 $\quad$ 0.05 $\quad$ 0.5

2. June/2022/Paper_8382/1H/No.12

Large car parks often have signs indicating the number of spaces available.

Councils can collect these data to judge whether they have enough car parks.

Negan, a council worker, records data hourly from 8 am to 5 pm about the spaces available in the main city centre car park.

The car park has 800 spaces.

Number of hours after 8 am	0	1	2	3	4
Spaces available	632	176	34	8	0

Number of hours after 8 am	5	6	7	8	9
Spaces available	0	0	25	106	447

(a) Estimate the **percentage** of this car park's spaces available at 9.30 am.

You **must** show your working.

[2 marks]

Answer _____ %

(b) The Product Moment Correlation Coefficient (PMCC) is calculated for the 5 data pairs in the top table and separately for the 5 data pairs in the bottom table.

By looking at patterns in the data, use **one** of the values listed below to complete each statement.

[2 marks]

1.25	0.831	0	–0.016	–0.845

The PMCC for the top table is _____

The PMCC for the bottom table is _____

(c) Interpret your answers to **part (b)** in context.

[2 marks]

The PMCC for the top table shows _____

_____

_____

The PMCC for the bottom table shows _____

_____

_____

(d) Lucille looks at the data and says,

"This car park is full for most of the day which shows there are not enough car parks."

Criticise both parts of her statement.

[2 marks]

"This car park is full for most of the day…" _____

_____

_____

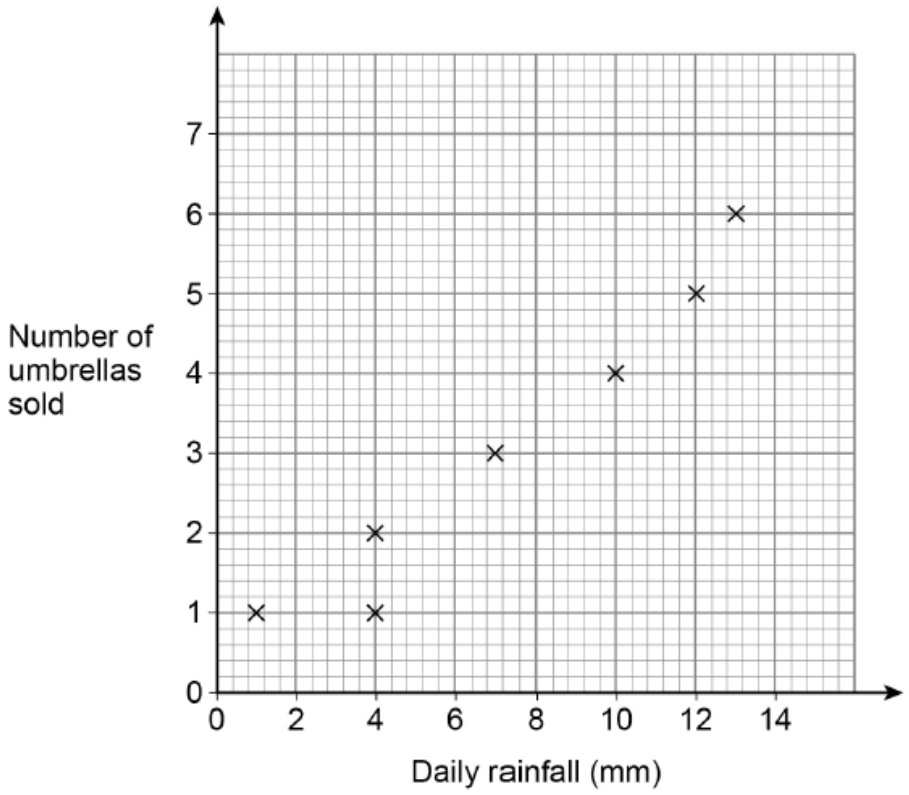
"…which shows there are not enough car parks." _____

_____

_____

3. June/2022/Paper_8382/2F/No.12

Caro has plotted this scatter diagram to show the numbers of umbrellas she sold against the amount of rain that fell each day.



(a) Caro can see from the scatter diagram that the greater the rainfall, the greater the number of umbrellas sold.

Tess says that if Caro sells more umbrellas, there will be greater rainfall.

Give a reason why Tess is **not** correct.

[1 mark]

(b) The weather forecast for tomorrow suggests that there will be 20 mm of rain throughout the day.

Caro says she can use the scatter diagram to predict how many umbrellas she would sell if the rainfall was 20 mm.

Is she correct?

Tick (✓) a box.

[1 mark]

Yes ☐ No ☐

Give a reason for your answer.

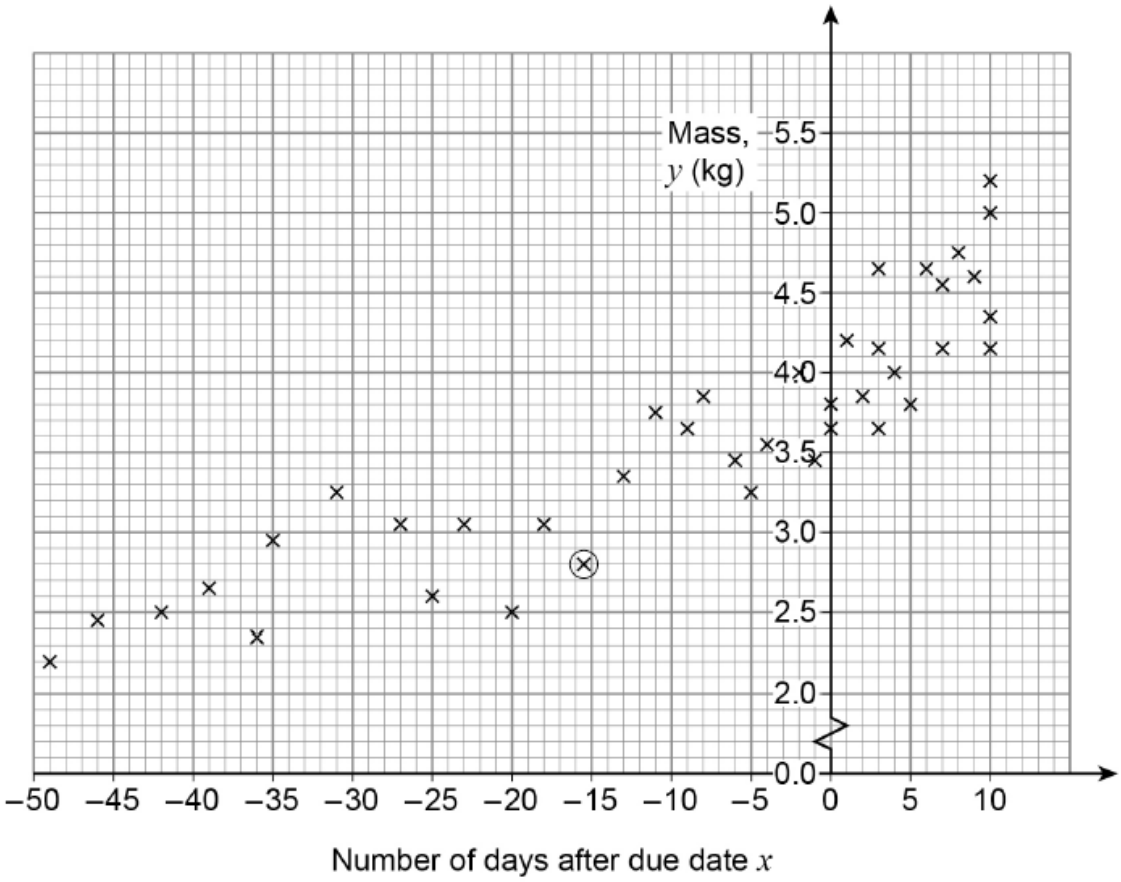
4. June/2022/Paper_8382/2H/No.11

The due date of a baby is the date on which it is expected to be born.

The scatter diagram shows the mass of 40 new-born babies **born** on March 20th 2021 plotted against the number of days the babies were born after their due date in whole days.

For example,

Value of number of days after due date	Interpretation
–10	Baby born 10 days **before** due date
5	Baby born 5 days **after** due date



(a) How many of **these** babies had a due date of March 10th 2021?

Circle your answer.

1 2 3 4

[1 mark]

(b) The circled point is an incorrect plot.

Give a reason how you know this.

[1 mark]

(c) The error is corrected.

The equation of the line of best fit for the data is $y = 4.01 + 0.04x$

(c) (i) Interpret the value 4.01 in the context of this scatter graph.

[1 mark]

(c) (ii) Interpret the value 0.04 in the context of this scatter graph.

[1 mark]

(c) (iii) Draw the line of best fit on the graph.

[2 marks]

(d) It is later found that there were two more babies born on March 20th.

- Sam was born 15 days before his due date.
- Nim was born 15 days after her due date.

Discuss the valid use of the scatter graph to estimate the birth mass of each of these babies.

Where an estimate is possible, explain your method and give the value.

[3 marks]

Sam _____

_____

_____

Nim _____

_____

_____

5. June/2022/Paper_8382/2H/No.13

A teacher marks some mock GCSE Statistics exam papers.

Each student sits two papers.

The teacher wants to see how the marks scored on the two papers are related.

She decides to calculate the value of Spearman's rank correlation coefficient (SRCC) using the formula

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

5 students sit the mock papers and the value of r_s = 0.8

(a) Interpret the value of SRCC in this context.

[1 mark]

(b) A sixth student now sits the papers.

Both papers are ranked 6th.

(b) (i) How will including the additional pair of papers change the value of SRCC?

[1 mark]

(b) (ii) Calculate the new value of SRCC for all 6 students.

[4 marks]

Answer _____