BIVARIATE STATISTICS

A BIVARIATE VARIABLES

A.1 IDENTIFYING INDEPENDENT AND DEPENDENT VARIABLES

Ex 1: A botanist is exploring the effect of daily sunlight on plant growth. She collects data over a week from several plants, measuring the average hours of sunlight each plant receives daily and the resulting height increase (in centimeters). The data represents a pair of numerical variables analyzed together. For this study, identify the roles of the variables:

- "Sunlight Hours per Day" is the:
  ☐ independent variable (x)
  ☐ dependent variable (y)
- "Height Increase (cm)" is the:
  ☐ independent variable (x)
  ☐ dependent variable (y)

Ex 2: A fitness coach is examining how weekly exercise time impacts heart rate. He gathers data over a month from several clients, measuring the average hours of exercise per week and the resting heart rate (in beats per minute, bpm). The data represents a pair of numerical variables analyzed together. For this study, identify the roles of the variables:

- "Resting Heart Rate (bpm)" is the:
  ☐ independent variable (x)
  ☐ dependent variable (y)
- "Exercise Hours per Week" is the:
  ☐ dependent variable (y)
  ☐ independent variable (x)

Ex 3: A baker is studying the impact of oven temperature on bread rise. She collects data from multiple batches, recording the oven temperature (in degrees Celsius) and the height of the bread rise (in centimeters). The data represents a pair of numerical variables analyzed together. For this study, identify the roles of the variables:

- "Oven Temperature (°C)" is the:
  ☐ dependent variable (y)
  ☐ independent variable (x)
- "Bread Rise Height (cm)" is the:
  ☐ dependent variable (y)
  ☐ independent variable (x)

Ex 4: A student is investigating how study time influences quiz performance. He collects data over a semester, tracking the hours spent studying per week and the quiz scores (out of 20). The data represents a pair of numerical variables analyzed together. For this study, identify the roles of the variables:

- "Quiz Score (out of 20)" is the:
  ☐ dependent variable (y)
  ☐ independent variable (x)
- "Study Hours per Week" is the:
  ☐ independent variable (x)
  ☐ dependent variable (y)

B SCATTER PLOTS

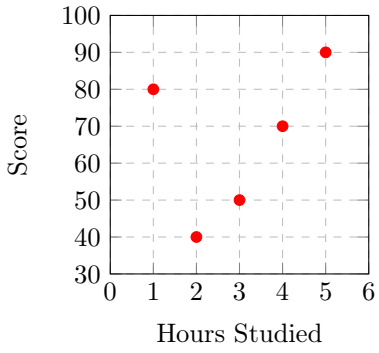
B.1 FINDING SCATTER PLOT

MCQ 5: The table below shows hours studied (x) and exam scores (y, out of 100) for five students:

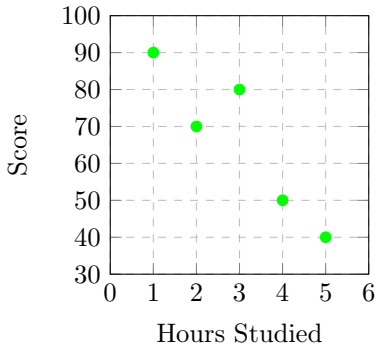
Student	O	P	Q	R	S
x: Hours Studied	1	2	4	3	5
y: Score Obtained	40	50	80	70	90

Below are three scatter plots. Each plots hours studied (x) against scores (y).

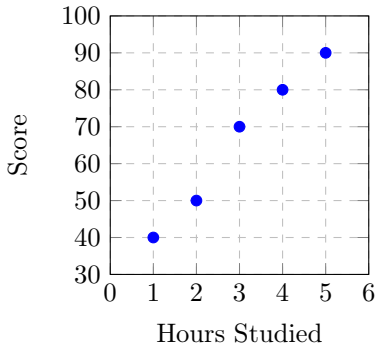
Graph E



Graph F



Graph G



Which scatter plot correctly matches the table? Check the correct one:

☐ Graph E

☐ Graph F

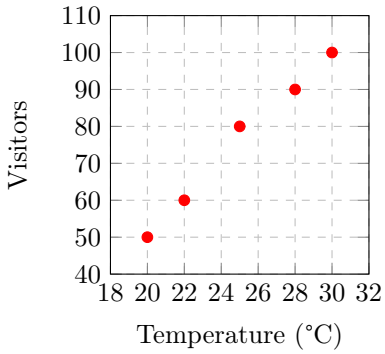
☐ Graph G

MCQ 6: The table below shows the daily temperature (x, in °C) and the number of beach visitors (y) at a local beach over five days:

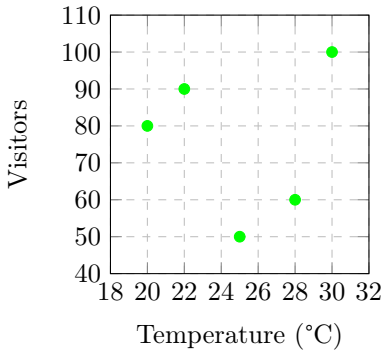
Day	A	B	C	D	E
x: Temperature (°C)	20	25	22	28	30
y: Beach Visitors	50	80	60	90	100

Below are three scatter plots. Each plots temperature (x) against the number of beach visitors (y).

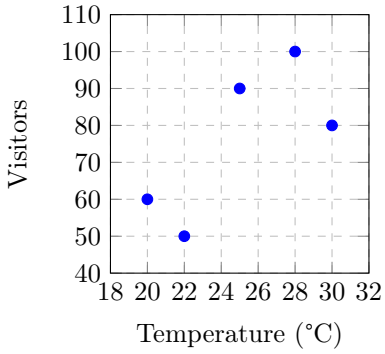
Graph A



Graph B



Graph C



Which scatter plot correctly matches the table? Check the correct one:

☐ Graph A

☐ Graph B

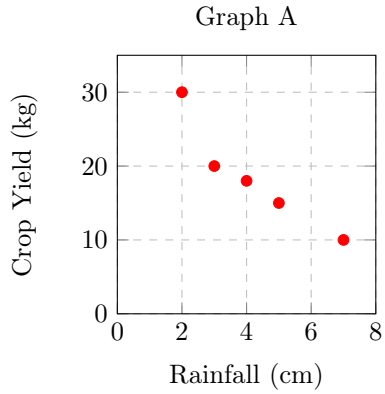
☐ Graph C

MCQ 7: The table below shows monthly rainfall (x, in cm) and crop yield (y, in kg) for five fields:

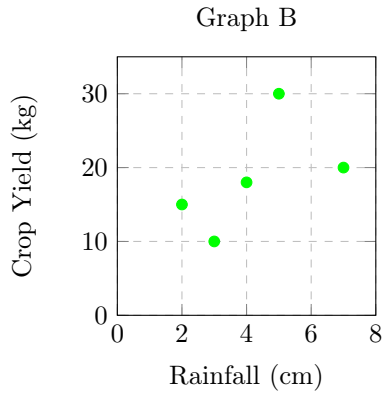
Field	A	B	C	D	E
x: Rainfall (cm)	5	3	7	2	4
y: Crop Yield (kg)	20	15	30	10	18

Below are three scatter plots. Each plots rainfall (x) against crop yield (y).

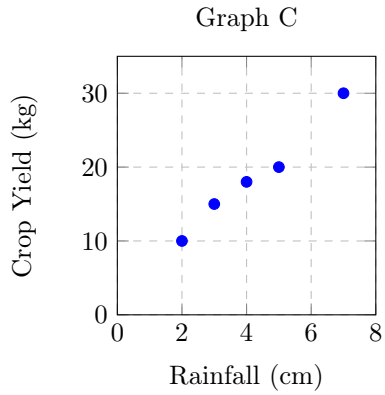
Graph A



Graph B



Graph C



Which scatter plot correctly matches the table? Check the correct one:

☐ Graph A

☐ Graph B

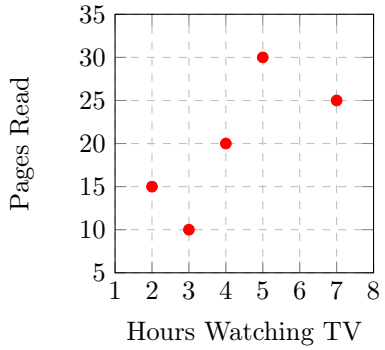
☐ Graph C

MCQ 8: The table below shows the number of hours spent watching TV (x) and the number of pages read in a book (y) by five students over a week:

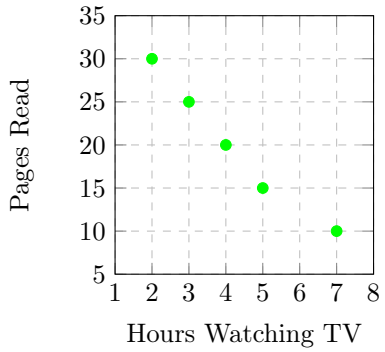
Student	V	W	X	Y	Z
x: Hours Watching TV	2	5	3	7	4
y: Pages Read	30	15	25	10	20

Below are three scatter plots. Each plots hours watching TV (x) against pages read (y).

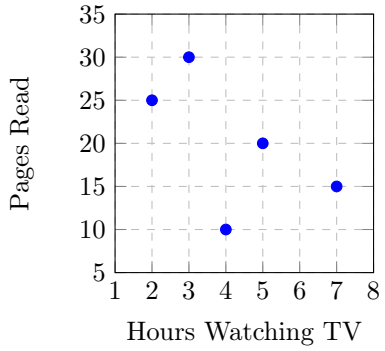
Graph P



Graph Q



Graph R



Which scatter plot correctly matches the table? Check the correct one:

☐ Graph P

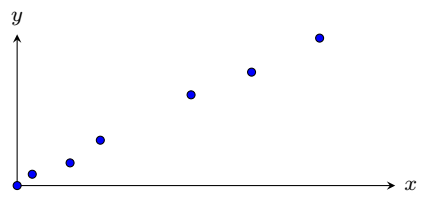
☐ Graph Q

☐ Graph R

C CORRELATION

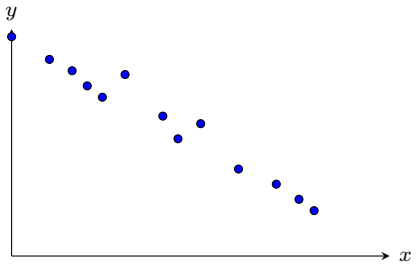
C.1 DETERMINING CORRELATION DIRECTION

Ex 9:



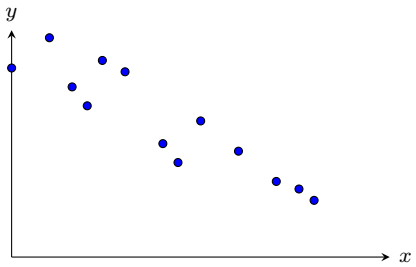
For the scatter plot shown above, the correlation is ☐ positive ☐ negative .

Ex 10:



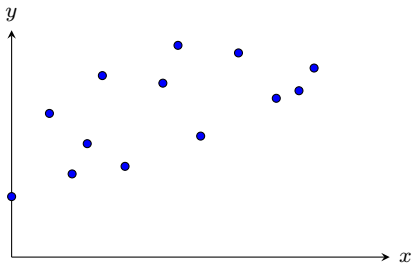
For the scatter plot above, the correlation is ☐ positive ☐ negative .

Ex 11:



For the scatter plot above, the correlation is ☐ positive ☐ negative .

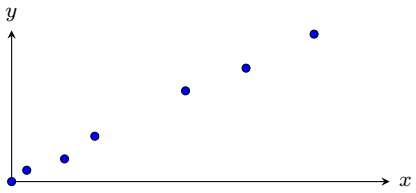
Ex 12:



For the scatter plot above, the correlation is ☐ positive ☐ negative .

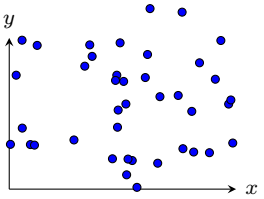
C.2 DETERMINING CORRELATION STRENGTH

Ex 13:



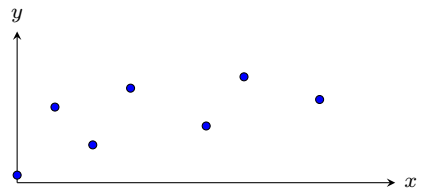
For the scatter plot above, the correlation strength is ☐ high ☐ low ☐ no .

Ex 14:



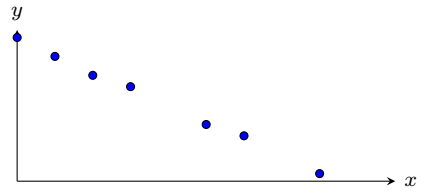
For the scatter plot above, the correlation strength is ☐ high ☐ low ☐ no .

Ex 15:



For the scatter plot above, the correlation strength is ☐ high ☐ low ☐ no .

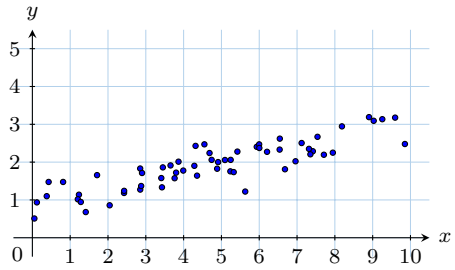
Ex 16:



For the scatter plot above, the correlation strength is ☐ high ☐ low ☐ no .

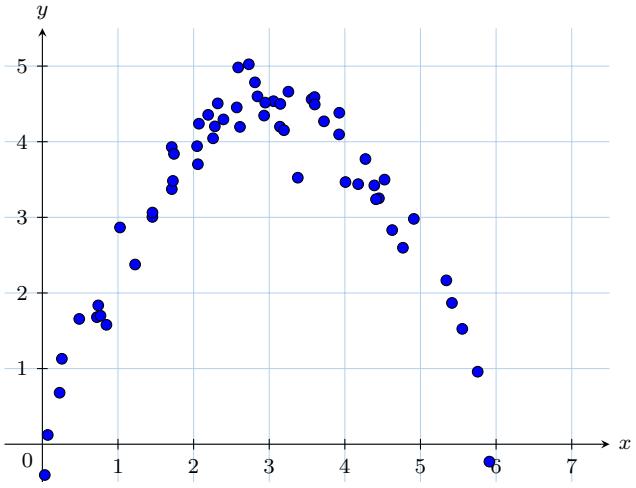
C.3 SPOTTING LINEAR CORRELATION

Ex 17:



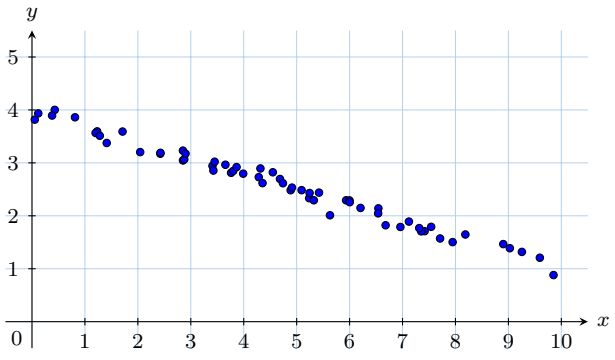
For the scatter plot above, the correlation is ☐ linear ☐ non-linear .

Ex 18:



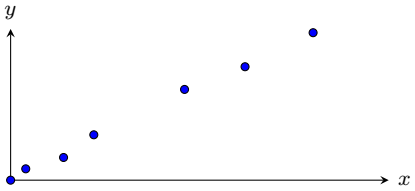
For the scatter plot above, the correlation is ☐ linear ☐ non-linear .

Ex 19:



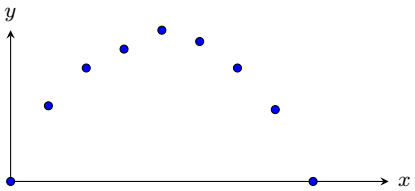
For the scatter plot above, the correlation is ☐ linear ☐ non-linear .

Ex 20:



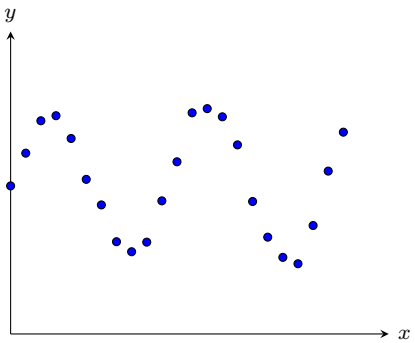
For the scatter plot above, the correlation is ☐ linear ☐ non-linear .

Ex 21:



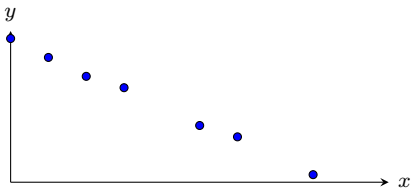
For the scatter plot above, the correlation is ☐ linear ☐ non-linear .

Ex 22:



For this scatter plot, the correlation is ☐ linear ☐ non-linear .

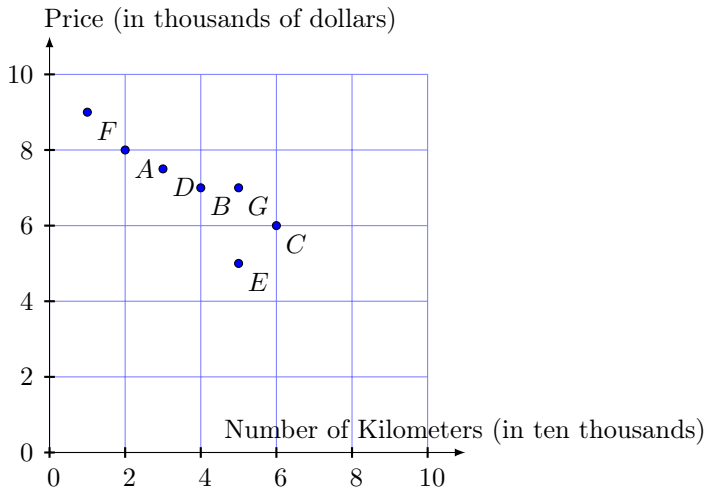
Ex 23:



For the scatter plot above, the correlation is ☐ linear ☐ non-linear .

C.4 OUTLIER DETECTION IN REAL-WORLD SCENARIOS

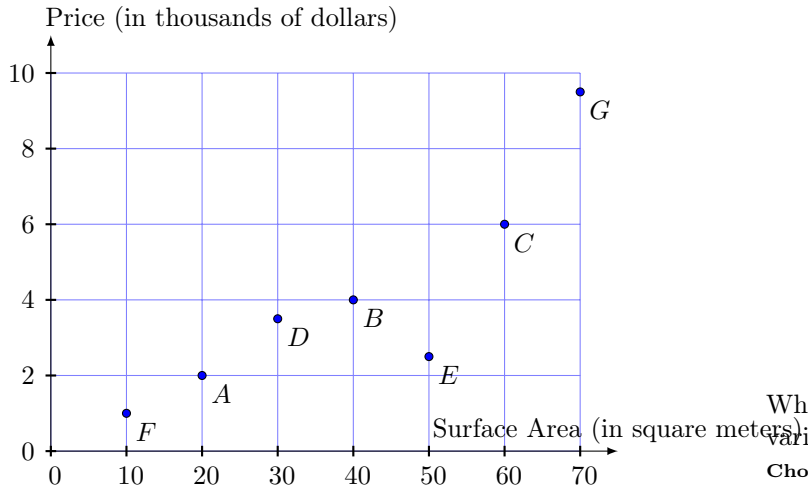
Ex 24: You're shopping for a used car and have collected data on available options, focusing on their prices and mileage. The scatter plot below displays the price of cars versus the number of kilometers driven (in ten thousands):



Which car offers a good deal, with a price significantly lower than expected for its mileage?

Select one:
- ☐ A
- ☐ B
- ☐ E
- ☐ G
- ☐ D

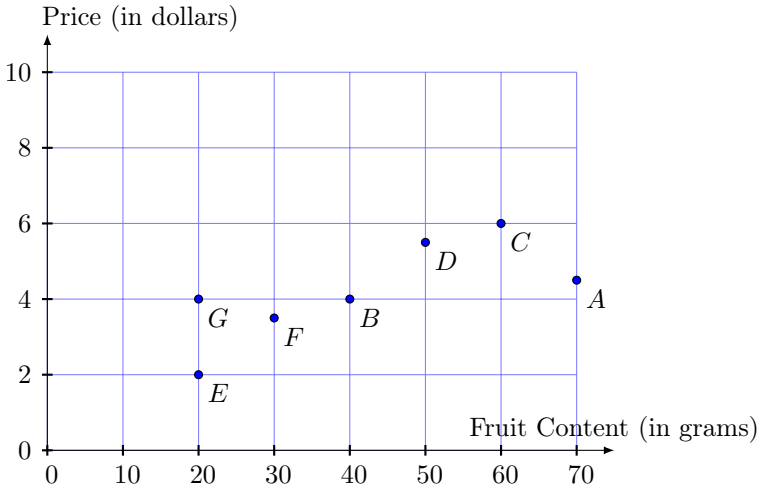
Ex 25: You're looking to buy a house and have gathered data on available properties, noting their prices and surface areas. The scatter plot below shows house prices versus surface area (in square meters):



Which house is a good deal, with a price much lower than expected for its surface area?

Select one:
- ☐ A
- ☐ B
- ☐ E
- ☐ G
- ☐ D

Ex 26: You're launching a homemade jam business and researching competitors. You've collected data on various jam brands, focusing on their prices and fruit content. The scatter plot below shows jam prices versus fruit content (in grams):

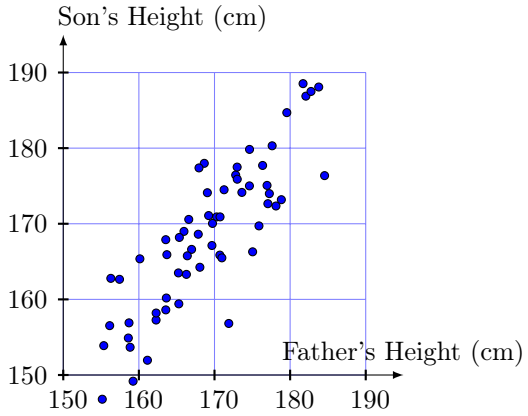


Which jam jar is a good deal, with a price much lower than expected for its fruit content?

Select one:
- ☐ A
- ☐ B
- ☐ E
- ☐ G
- ☐ D

C.5 SPOTTING TRENDS IN DATA

MCQ 27: Raymundo wants to explore whether sons of taller fathers tend to be taller. The scatter plot below shows the relationship between fathers' heights and their sons' heights (in centimeters) based on Raymundo's sample:

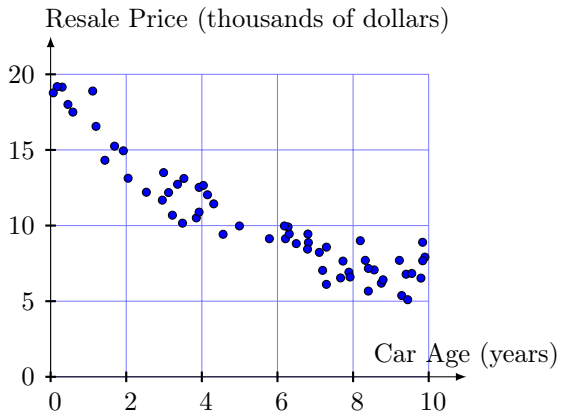


Which statement best describes the association between these variables?

Choose one answer:

- ☐ Taller fathers tend to have taller sons.
- ☐ Taller fathers tend to have shorter sons.
- ☐ There is no clear relationship between fathers' and sons' heights.

MCQ 28: Marcelo is studying how the age of used cars affects their resale prices. The scatter plot below illustrates the relationship between car age and resale price:



Which statement best describes the association between these variables?

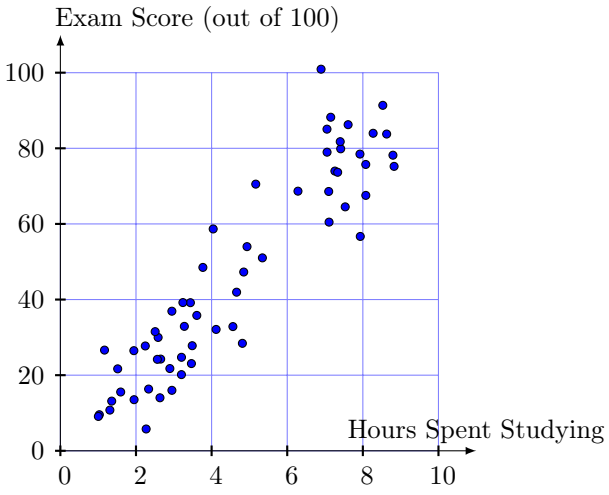
Choose one answer:

☐ Older cars tend to have higher resale prices.

☐ Older cars tend to have lower resale prices.

☐ There is no clear relationship between car age and resale price.

MCQ 29: Lorena is investigating whether spending more time studying improves exam scores. The scatter plot below shows the relationship between hours spent studying and exam scores:



Which statement best describes the association between these variables?

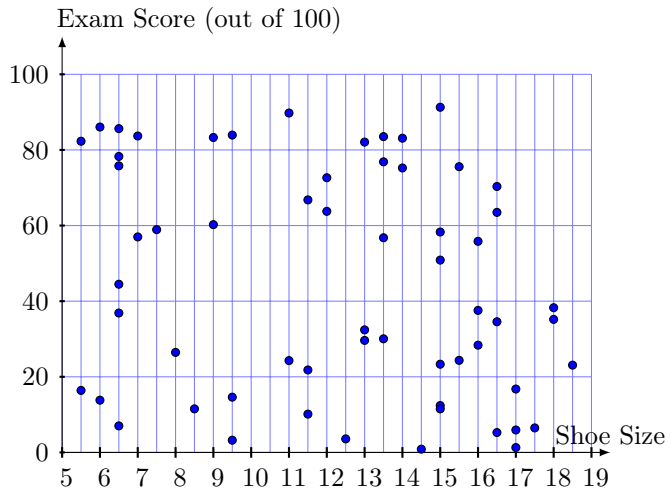
Choose one answer:

☐ Students who study more hours tend to score higher on the exam.

☐ Students who study more hours tend to score lower on the exam.

☐ There is no clear relationship between study hours and exam scores.

MCQ 30: Nina is exploring whether a student's shoe size relates to their exam score. The scatter plot below shows the relationship between students' shoe sizes and their exam scores:



Which statement best describes the association between these variables?

Choose one answer:

☐ Students with larger shoe sizes tend to score higher on the exam.

☐ Students with larger shoe sizes tend to score lower on the exam.

☐ There is no clear relationship between shoe size and exam score.

D CAUSALITY

D.1 CLASSIFYING: CORRELATION OR CAUSATION?

Ex 31: Determine whether each relationship is best described as "just a correlation" (an association without direct cause) or "a causation" (one causes the other).

Select the correct label for each statement:

- Eating healthy and feeling strong: ☐ just a correlation ☐ a causation .
- Feeling strong and living longer: ☐ just a correlation ☐ a causation .
- Eating healthy and living longer: ☐ just a correlation ☐ a causation .

Ex 32: Determine whether each relationship is best described as "just a correlation" (an association without direct cause) or "a causation" (one causes the other).

Select the correct label for each statement:

- Sunny weather and eating ice cream: ☐ just a correlation ☐ a causation .
- Eating ice cream and getting sunburned: ☐ just a correlation ☐ a causation .
- Sunny weather and getting sunburned: ☐ just a correlation ☐ a causation .

Ex 33: Determine whether each relationship is best described as "just a correlation" (an association without direct cause) or "a causation" (one causes the other).

Select the correct label for each statement:

- Working hard and getting good marks:
  ☐ just a correlation
  ☐ a causation .
- Getting good marks and wearing glasses:
  ☐ just a correlation
  ☐ a causation .
- Working hard and wearing glasses:
  ☐ just a correlation
  ☐ a causation .

Ex 34: Determine whether each relationship is best described as "just a correlation" (an association without direct cause) or "a causation" (one causes the other).

Select the correct label for each statement:

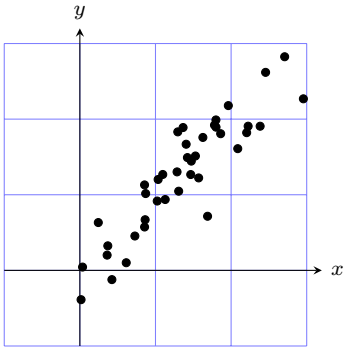
- Exercising regularly and performing well in sports:
  ☐ just a correlation
  ☐ a causation .
- Exercising regularly and having more energy:
  ☐ just a correlation
  ☐ a causation .
- Having more energy and performing well in sports:
  ☐ just a correlation
  ☐ a causation .

E MEASURING CORRELATION

E.1 MEASURING CORRELATION

Ex 35: Examine the scatter plot below, which shows the relationship between two variables, x and y.

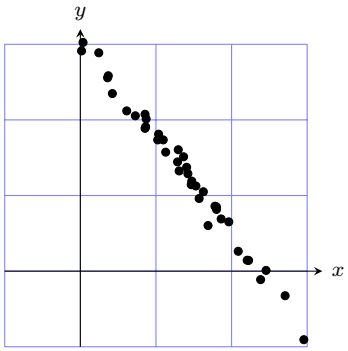
Choose the correct correlation coefficient (r) from the options provided:



The correlation coefficient is $r =$
☐ 1
☐ 0.9
☐ 0.7
☐ 0
☐ −0.7
☐ −0.9
☐ −1 .

Ex 36: Examine the scatter plot below, which shows the relationship between two variables, x and y.

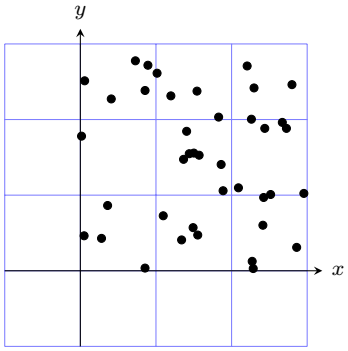
Choose the correct correlation coefficient (r) from the options provided:



The correlation coefficient is $r =$
☐ 1
☐ 0.9
☐ 0.7
☐ 0
☐ −0.7
☐ −0.9
☐ −1 .

Ex 37: Examine the scatter plot below, which shows the relationship between two variables, x and y.

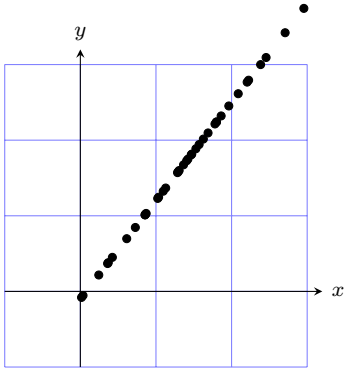
Choose the correct correlation coefficient (r) from the options provided:



The correlation coefficient is $r =$
☐ 1
☐ 0.9
☐ 0.7
☐ 0
☐ −0.7
☐ −0.9
☐ −1 .

Ex 38: Examine the scatter plot below, which shows the relationship between two variables, x and y.

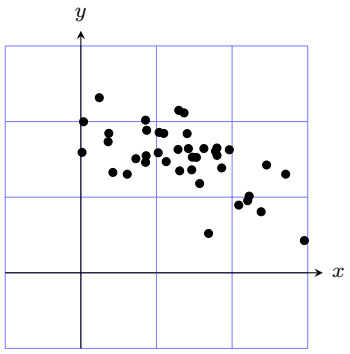
Choose the correct correlation coefficient (r) from the options provided:



☐ 1
☐ 0.9
☐ 0.7
The correlation coefficient is $r =$ ☐ 0 .
☐ −0.7
☐ −0.9
☐ −1

Ex 39: Examine the scatter plot below, which shows the relationship between two variables, x and y.

Choose the correct correlation coefficient (r) from the options provided:

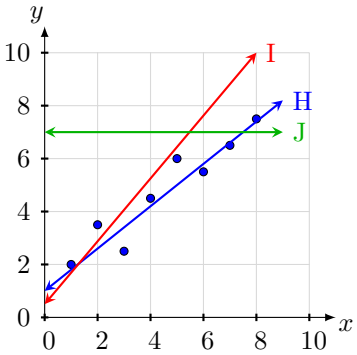


☐ 1
☐ 0.9
☐ 0.7
The correlation coefficient is $r =$ ☐ 0 .
☐ −0.7
☐ −0.9
☐ −1

F LINE FITTING

F.1 FINDING THE LINE OF BEST FIT VISUALLY

MCQ 40: Which line best fits the data shown in the scatter plot below?



Choose one answer:

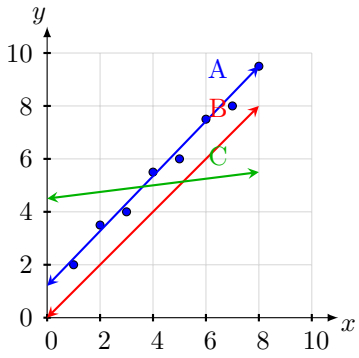
☐ H

☐ I

☐ J

☐ None of the lines fit the data.

MCQ 41: Which line best fits the data shown in the scatter plot below?



Choose one answer:

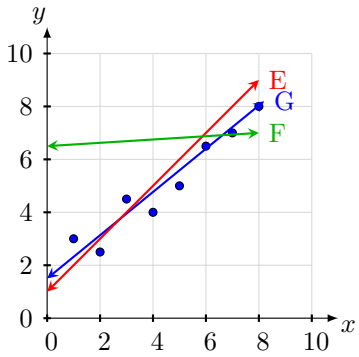
☐ A

☐ B

☐ C

☐ None of the lines fit the data.

MCQ 42: Which line best fits the data shown in the scatter plot below?



Choose one answer:

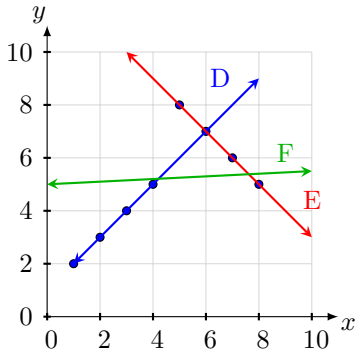
☐ E

☐ F

☐ G

☐ None of the lines fit the data.

MCQ 43: Which line best fits the data shown in the scatter plot below?



Choose one answer:

☐ D

☐ E

☐ F

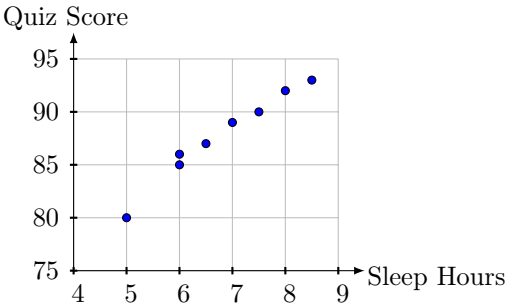
☐ None of the lines fit the data.

F.2 EVALUATING THE FIT OF A LINEAR MODEL

MCQ 44: Lisa wanted to explore if the number of hours she sleeps affects her quiz performance. She recorded her sleep hours and quiz scores over 8 days:

Sleep Hours	6	7.5	5	8	6.5	7	8.5	6
Quiz Score	85	90	80	92	87	89	93	86

She plotted the data as follows:



Is it reasonable to fit a line of best fit to this data?

Choose one answer:

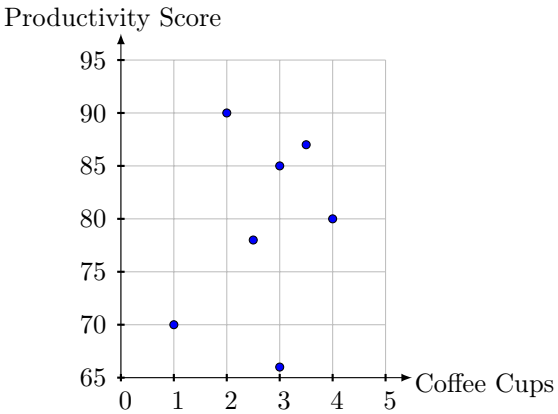
☐ Yes

☐ No

MCQ 45: John wanted to check if coffee consumption affects his productivity. He recorded coffee cups and productivity scores over 7 days:

Coffee Cups	1	3	2	4	2.5	3	3.5
Productivity Score	70	85	90	80	78	66	87

He plotted the data as follows:



Is it reasonable to fit a line of best fit to this data?

Choose one answer:

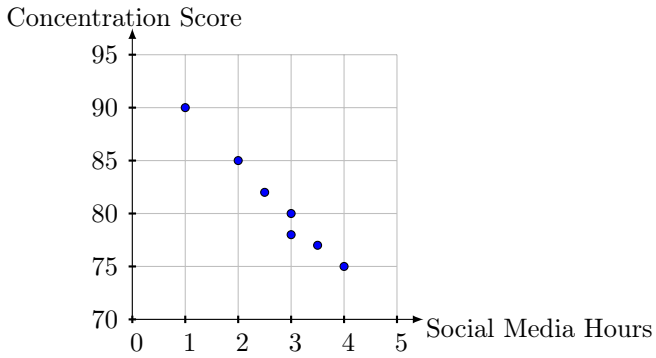
☐ Yes

☐ No

MCQ 46: Sophia wanted to see if social media use impacts her concentration. She recorded social media hours and concentration scores (y, out of 100) over 7 days:

Social Media Hours	1	3	2	4	2.5	3	3.5
Concentration Score	90	80	85	75	82	78	77

She plotted the data as follows:



Is it reasonable to fit a line of best fit to this data?

Choose one answer:

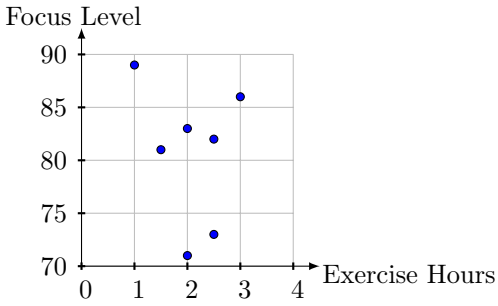
☐ Yes

☐ No

MCQ 47: Anna wanted to investigate if exercise impacts her mental focus. She recorded exercise hours and focus levels (y, out of 100) over 7 days:

Exercise Hours	1	2.5	2	3	1.5	2	2.5
Focus Level	89	82	71	86	81	83	73

She plotted the data as follows:



Is it reasonable to fit a line of best fit to this data?

Choose one answer:

☐ Yes

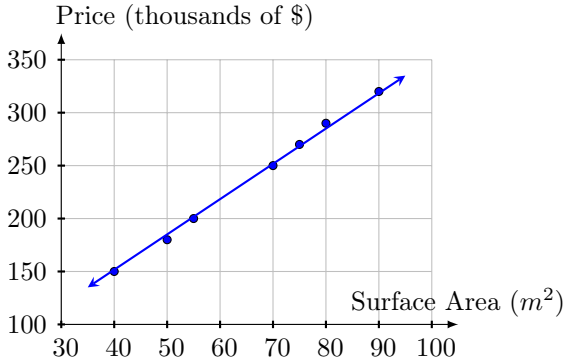
☐ No

F.3 ESTIMATING VALUES GRAPHICALLY

MCQ 48: Sophie, a real estate agent, explored the correlation between house surface area (in m^2) and price (in thousands of dollars). She recorded data from recent sales:

Surface Area (m^2)	40	50	55	70	75	80	90
Price (thousands of $)	150	180	200	250	270	290	320

She plotted the data and drew a line of best fit:



Using the line of best fit, estimate the price of a house with a surface area of 60 m^2.

Choose one answer:

☐ \$165,000

☐ \$200,000

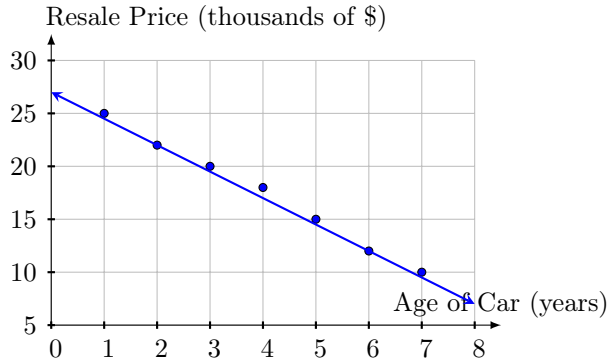
☐ \$225,000

☐ \$240,000

MCQ 49: Caroline, a used car dealer, investigated the correlation between car age (in years) and resale price (in thousands of dollars). She recorded data from recent sales:

Age of Car (years)	1	2	3	4	5	6	7
Resale Price (thousands of \$)	25	22	20	18	15	12	10

She plotted the data and drew a line of best fit:



Using the line of best fit, estimate the resale price of a car that is 4.5 years old.

Choose one answer:

☐ \$16,000

☐ \$19,500

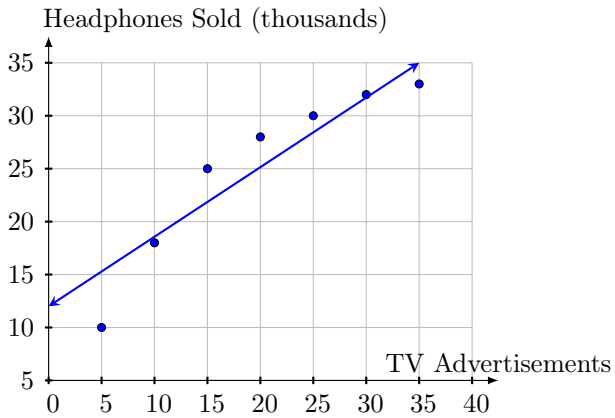
☐ \$21,000

☐ \$13,500

MCQ 50: Alex, a marketing manager, studied the correlation between TV advertisements aired weekly and headphone sales (in thousands). He recorded data over several weeks:

TV Advertisements	5	10	15	20	25	30	35
Headphones Sold (thousands)	10	18	25	28	30	32	33

He plotted the data and drew a line of best fit:



Using the line of best fit, estimate the number of headphones sold if 18 TV advertisements are aired in a week.

Choose one answer:

☐ 22,000

☐ 24,000

☐ 26,000

☐ 28,000

F.4 ESTIMATING VALUES WITH LINEAR EQUATIONS

Ex 51: Dr. Smith, a geneticist, studied the relationship between fathers' and sons' heights (in cm). He collected data from multiple families, where x represents the father's height and y represents the son's height. After analysis, he derived the best-fit line equation: $y = x + 2$. Using this equation, estimate the son's height if the father's height is 177 cm.

☐ cm

Ex 52: Ms. Lopez, a fitness coach, studied the relationship between weekly exercise hours and resting heart rates (in beats per minute, bpm) of her clients. She collected data from several clients, where x represents the hours of exercise per week and y represents the resting heart rate. After analysis, she derived the best-fit line equation: $y = -2x + 80$. Using this equation, estimate the resting heart rate for a client who exercises 6 hours per week.

☐ bpm

Ex 53: Mr. Patel, a nutritionist, investigated the relationship between daily water intake (in liters) and energy levels (on a scale of 0 to 10) among his clients. He collected data from several clients, where x represents the daily water intake and y represents the energy level. After analysis, he derived the best-fit line equation: $y = 1.8x + 2.5$. Using this equation, estimate the energy level for a client who drinks 3.5 liters of water daily.

☐

Ex 54: Ms. Chen, a teacher, explored the relationship between students' study time (in hours per week) and their test scores (out of 100). She collected data from her class, where x represents the study hours per week and y represents the test score. After analysis, she derived the best-fit line equation: $y = 4.2x + 55.6$. Using this equation, estimate the test score for a student who studies 8.5 hours per week.

☐