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)**
- "Height Increase (cm)" is the: **dependent variable (y)**

Answer: In this study, the botanist is investigating how sunlight influences plant growth, using two numerical variables:

- "Sunlight Hours per Day" is the independent variable (x) because it is the factor controlled or varied to observe its impact on growth. It acts as the input or predictor.
- "Height Increase (cm)" is the dependent variable (y) because it is the outcome or response that changes based on the amount of sunlight received. It depends on the independent variable.

Thus, the correct choices are:

- "Sunlight Hours per Day": independent variable (x).
- "Height Increase (cm)": 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: **dependent variable (y)**
- "Exercise Hours per Week" is the: **independent variable (x)**

Answer: In this study, the fitness coach is investigating how exercise influences heart rate, using two numerical variables:

- "Exercise Hours per Week" is the independent variable (x) because it is the factor varied to observe its effect on heart rate. It serves as the input or predictor.
- "Resting Heart Rate (bpm)" is the dependent variable (y) because it is the outcome that changes based on the amount of exercise. It depends on the independent variable.

Thus, the correct choices are:

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

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: **independent variable (x)**
- "Bread Rise Height (cm)" is the: **dependent variable (y)**

Answer: In this study, the baker is examining how oven temperature affects bread rise, using two numerical variables:

- "Oven Temperature (°C)" is the independent variable (x) because it is the factor controlled to observe its effect on the bread's rise. It acts as the input or predictor.
- "Bread Rise Height (cm)" is the dependent variable (y) because it is the outcome that changes based on the temperature. It depends on the independent variable.

Thus, the correct choices are:

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

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)**
- "Study Hours per Week" is the: **independent variable (x)**

Answer: In this study, the student is investigating how study time affects quiz performance, using two numerical variables:

- "Study Hours per Week" is the independent variable (x) because it is the factor varied to observe its impact on quiz scores. It serves as the input or predictor.
- "Quiz Score (out of 20)" is the dependent variable (y) because it is the outcome that changes based on the study time. It depends on the independent variable.

Thus, the correct choices are:

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

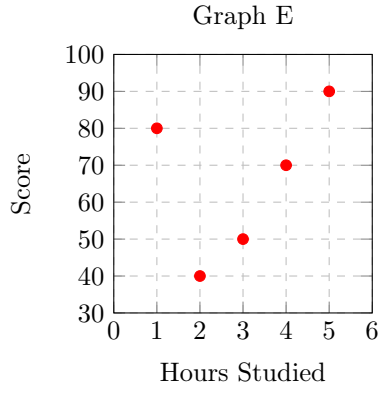
B SCATTER PLOTS

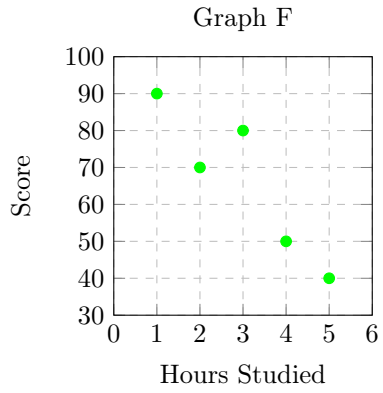
B.1 FINDING SCATTER PLOT

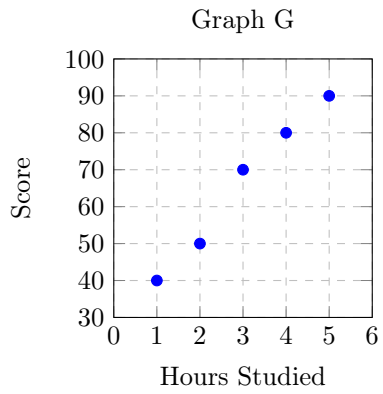
MCQ 5: The table below shows hours studied (x) and exam scores (y, out of 100) for five students:

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).







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

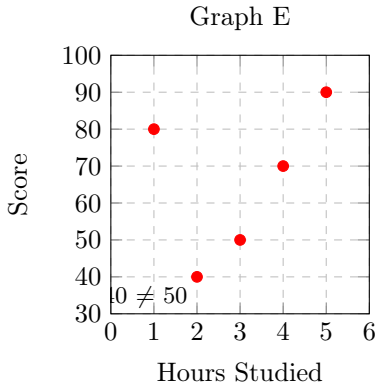
☐ Graph E

☐ Graph F

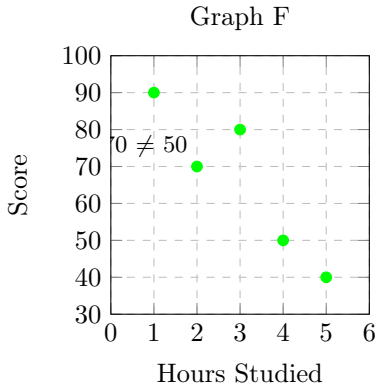
☒ Graph G

Answer: The correct answer is Graph G. Here's why:

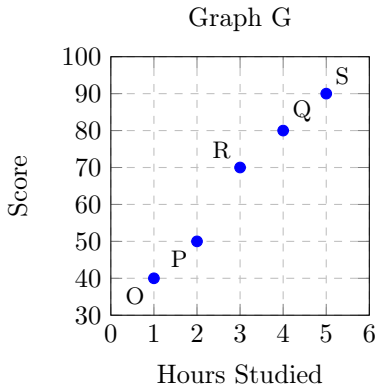
- **Graph E (Incorrect)**: Points don't match the table. For example, P (2 hours) scores 40, not 50; O (1 hour) scores 80, not 40.



- **Graph F (Incorrect)**: Points are mismatched. For example, P (2 hours) scores 70, not 50; S (5 hours) scores 40, not 90.



- **Graph G (Correct)**: All points match the table: O (1, 40), P (2, 50), Q (4, 80), R (3, 70), S (5, 90), .

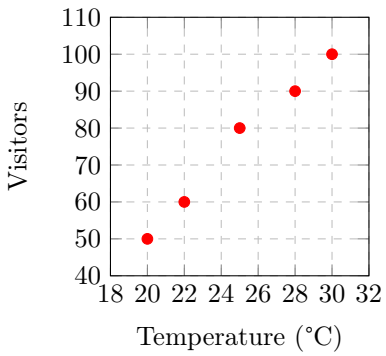


MCQ 6: The table below shows the daily temperature (x, in °C) and the number of beach visitors (y) at a local beach over five days:

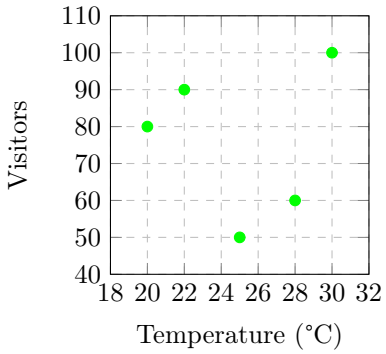
Day	A	B	C	D	E
x: Temperature (°C)	20	25	22	28	30
y: Beach Visitors	50	80	60	90	100

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

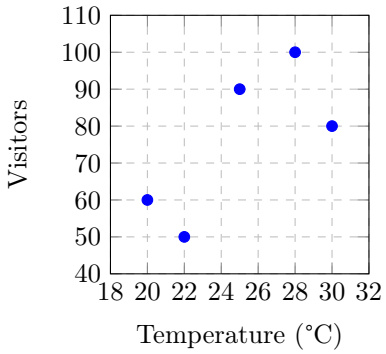
Graph A



Graph B



Graph C



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

☒ Graph A

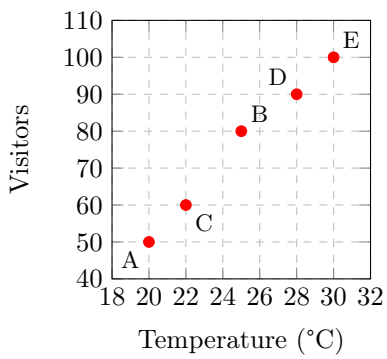
☐ Graph B

☐ Graph C

Answer: The correct answer is Graph A. Here's why:

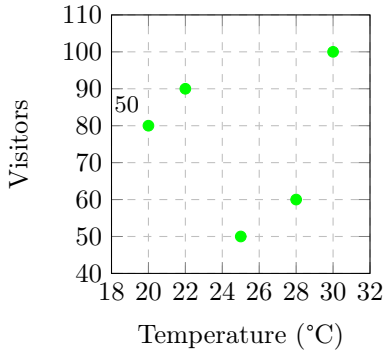
- **Graph A (Correct)**: All points match the table: A (20, 50), B (25, 80), C (22, 60), D (28, 90), E (30, 100).

Graph A



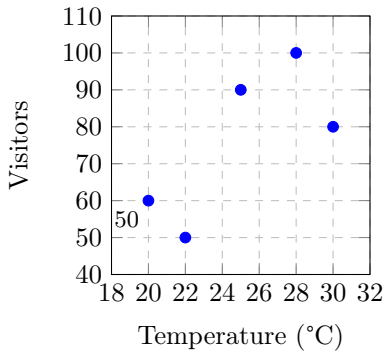
- **Graph B (Incorrect)**: Points don't match. For example, A (20°C) has 80 visitors, not 50; C (22°C) has 90, not 60.

Graph B



- **Graph C (Incorrect)**: Points are mismatched. For example, A (20°C) has 60 visitors, not 50; E (30°C) has 80, not 100.

Graph C

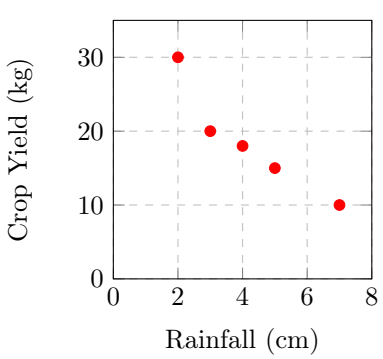


MCQ 7: The table below shows monthly rainfall (x, in cm) and crop yield (y, in kg) for five fields:

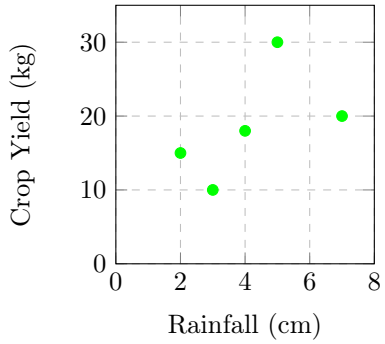
Field	A	B	C	D	E
x: Rainfall (cm)	5	3	7	2	4
y: Crop Yield (kg)	20	15	30	10	18

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

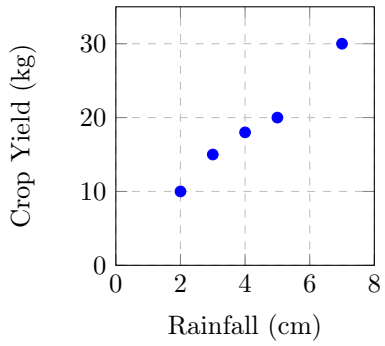
Graph A



Graph B



Graph C



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

☐ Graph A

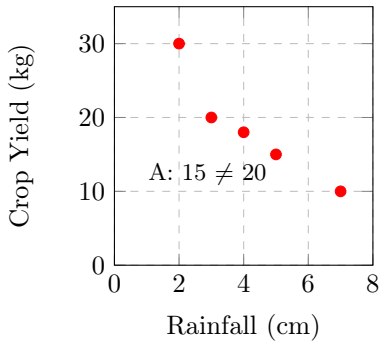
☐ Graph B

☒ Graph C

Answer: The correct answer is Graph C. Here's why:

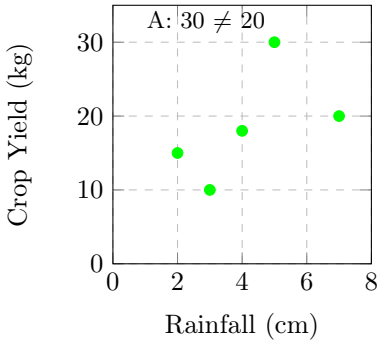
- **Graph A (Incorrect)**: Points don't match. For example, A (5 cm) yields 15 kg, not 20 kg; D (2 cm) yields 30 kg, not 10 kg.

Graph A



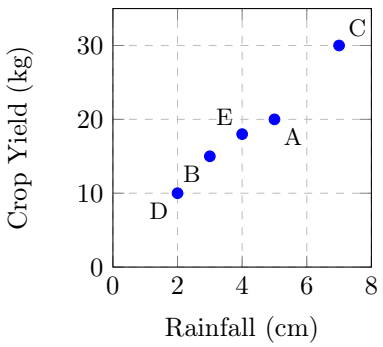
- **Graph B (Incorrect)**: Points are mismatched. For example, A (5 cm) yields 30 kg, not 20 kg; B (3 cm) yields 10 kg, not 15 kg.

Graph B



- **Graph C (Correct)**: All points match: A (5, 20), B (3, 15), C (7, 30), D (2, 10), E (4, 18).

Graph C

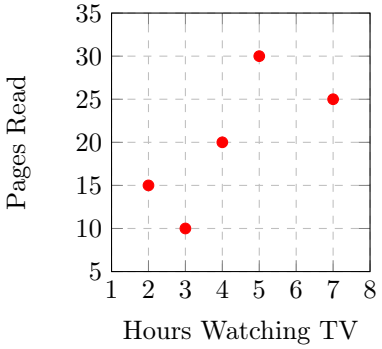


MCQ 8: The table below shows the number of hours spent watching TV (x) and the number of pages read in a book (y) by five students over a week:

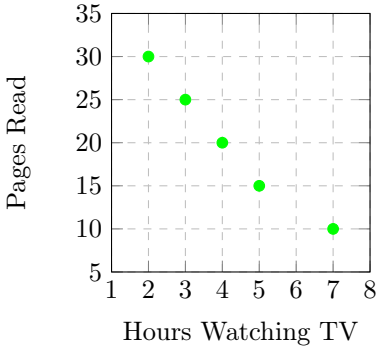
Student	V	W	X	Y	Z
x: Hours Watching TV	2	5	3	7	4
y: Pages Read	30	15	25	10	20

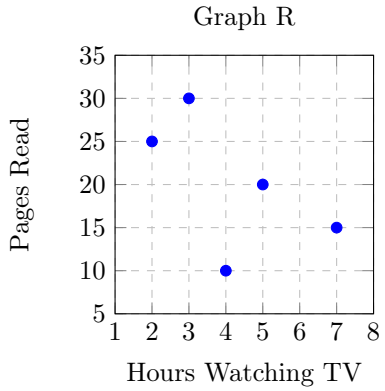
Below are three scatter plots. Each plots hours watching TV (x) against pages read (y).

Graph P



Graph Q





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

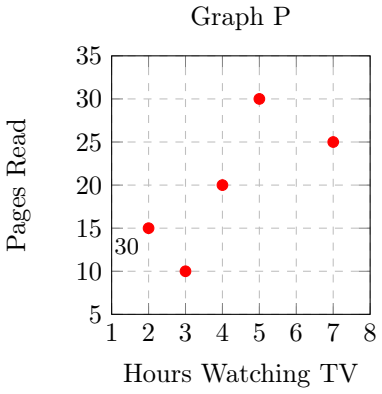
☐ Graph P

☒ Graph Q

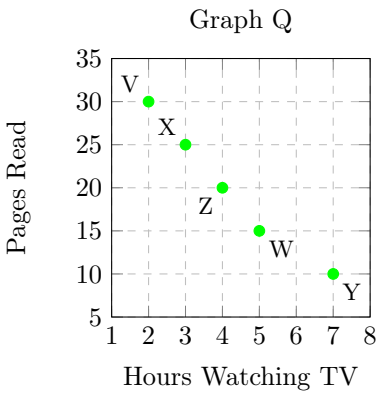
☐ Graph R

Answer: The correct answer is Graph Q. Here's why:

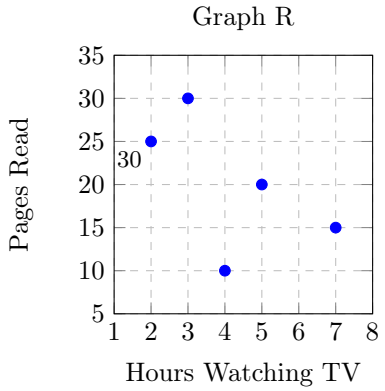
- **Graph P (Incorrect)**: Points don't match the table. For example, V (2 hours) has 15 pages, not 30; W (5 hours) has 30 pages, not 15.



- **Graph Q (Correct)**: All points match the table: V (2, 30), W (5, 15), X (3, 25), Y (7, 10), Z (4, 20). This also shows a weak negative correlation, as more TV hours generally correspond to fewer pages read, though not perfectly linear.



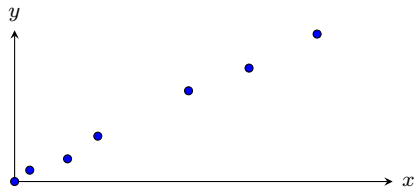
- **Graph R (Incorrect)**: Points are mismatched. For example, V (2 hours) has 25 pages, not 30; Y (7 hours) has 15 pages, not 10.



C CORRELATION

C.1 DETERMINING CORRELATION DIRECTION

Ex 9:

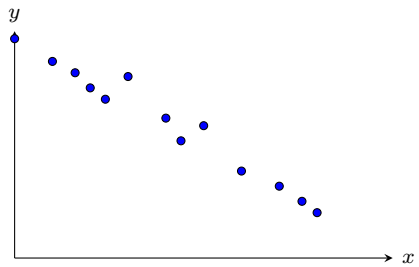


For the scatter plot shown above, the correlation is **positive**.

Answer:

- The data points follow an **upward** trend.
- As the variable x increases $\rightarrow$, the variable y also tends to **increase** $\uparrow$.
- Therefore, the correlation is **positive**.

Ex 10:

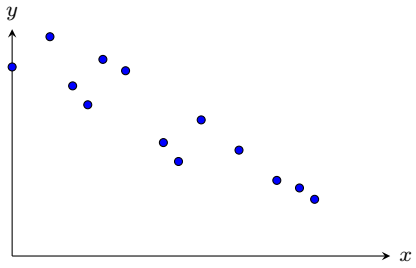


For the scatter plot above, the correlation is **negative**.

Answer:

- The data points exhibit a clear **downward** trend.
- As the variable x increases $\rightarrow$, the variable y tends to **decrease** $\downarrow$.
- Thus, the correlation is **negative**.

Ex 11:

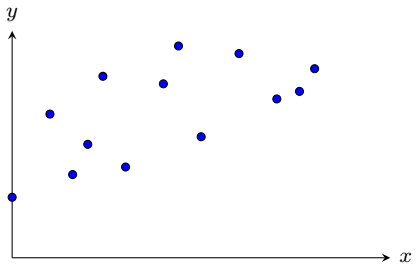


For the scatter plot above, the correlation is **negative**.

Answer:

- The data points are scattered but suggest a subtle **downward** trend.
- As the variable x increases $\rightarrow$, the variable y generally tends to **decrease** $\downarrow$.
- Therefore, the correlation is **negative**, though it appears weaker than a strong negative correlation.

Ex 12:



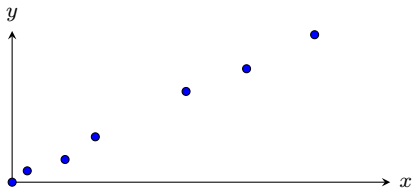
For the scatter plot above, the correlation is **positive**.

Answer:

- The data points are dispersed but indicate a slight **upward** trend.
- As the variable x increases $\rightarrow$, the variable y tends to **increase** $\uparrow$.
- Thus, the correlation is **positive**.

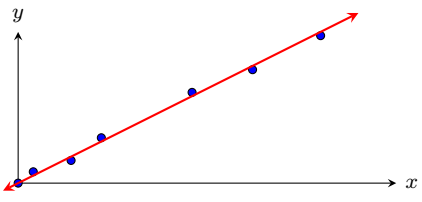
C.2 DETERMINING CORRELATION STRENGTH

Ex 13:



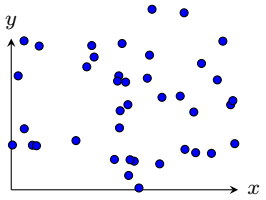
For the scatter plot above, the correlation strength is **high**.

Answer:

- 

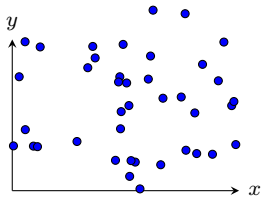
- The points align closely with a straight line.
- Therefore, the correlation strength is **high**.

Ex 14:



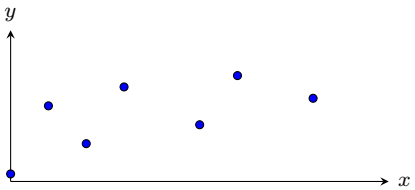
For the scatter plot above, the correlation strength is **no**.

Answer:

- 

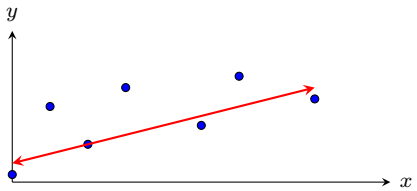
- The points are widely scattered with no discernible trend.
- Thus, there is **no** correlation.

Ex 15:



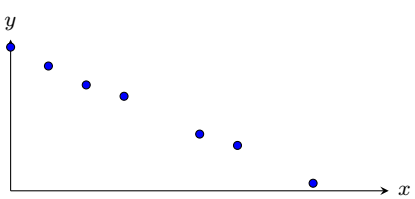
For the scatter plot above, the correlation strength is **low**.

Answer:

- 

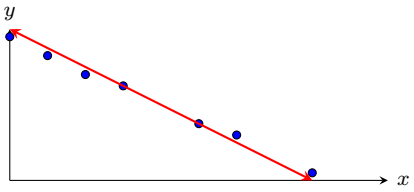
- The points are scattered around the line, showing a weak upward trend.
- Therefore, the correlation strength is **low**.

Ex 16:



For the scatter plot above, the correlation strength is **high**.

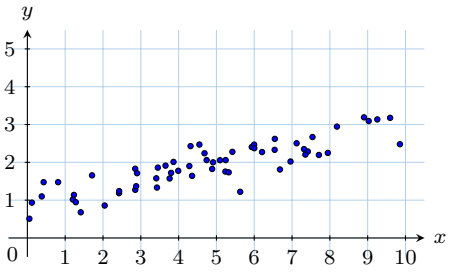
Answer:



- The points closely follow a straight line.
- Thus, the correlation strength is **high**.

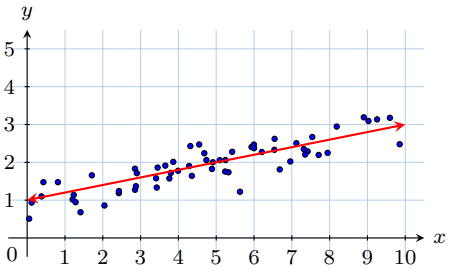
C.3 SPOTTING LINEAR CORRELATION

Ex 17:



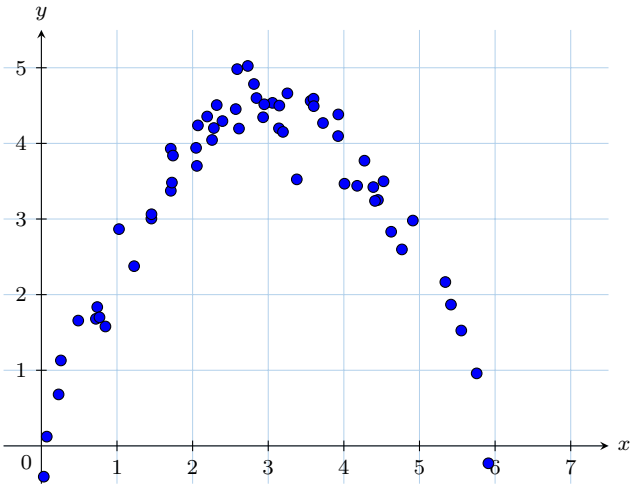
For the scatter plot above, the correlation is **linear**.

Answer:



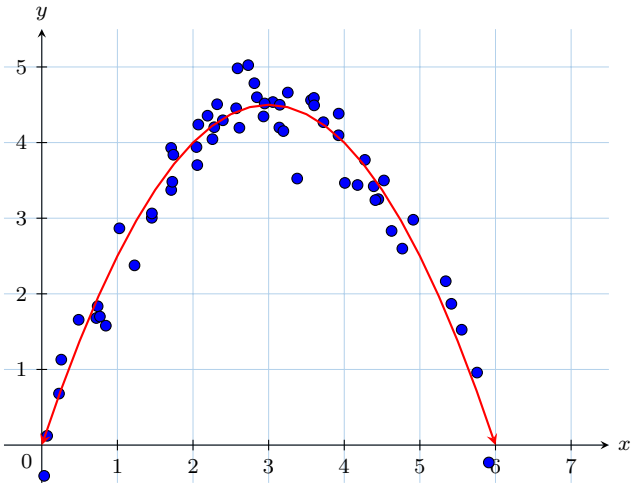
- The points closely align with a straight line.
- Thus, the correlation is **linear**.

Ex 18:



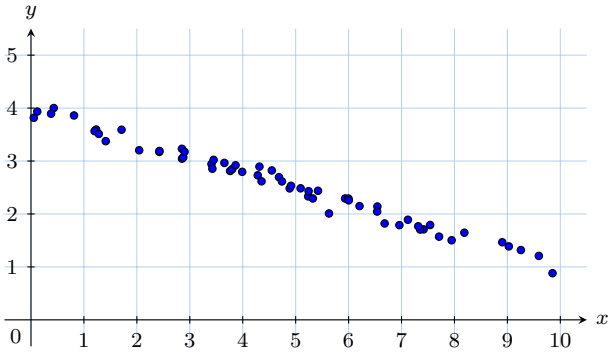
For the scatter plot above, the correlation is **non-linear**.

Answer:



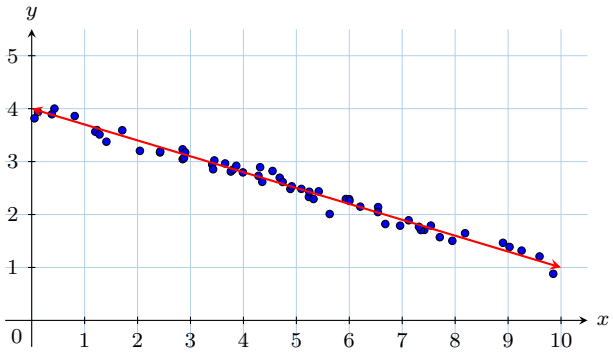
- The points follow a parabolic curve rather than a straight line.
- Therefore, the correlation is **non-linear**.

Ex 19:



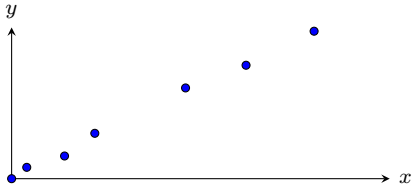
For the scatter plot above, the correlation is **linear**.

Answer:



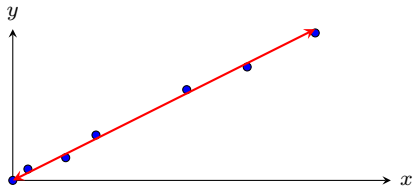
- The points closely follow a straight line.
- Thus, the correlation is **linear**.

Ex 20:



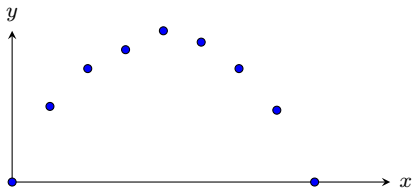
For the scatter plot above, the correlation is **linear**.

Answer:



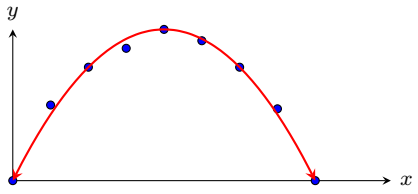
- The points align closely with a straight line.
- Therefore, the correlation is **linear**.

Ex 21:



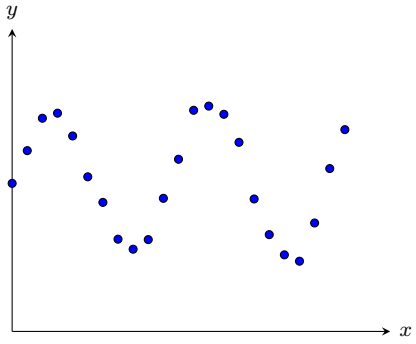
For the scatter plot above, the correlation is **non-linear**.

Answer:



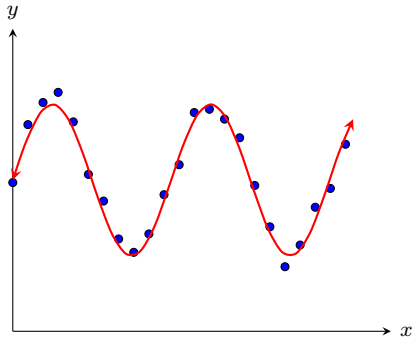
- The points form a parabolic pattern rather than a straight line.
- Thus, the correlation is **non-linear**.

Ex 22:



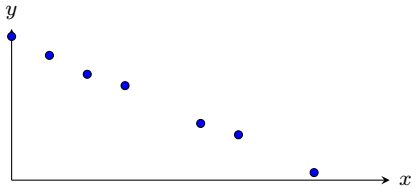
For this scatter plot, the correlation is **non-linear**.

Answer:



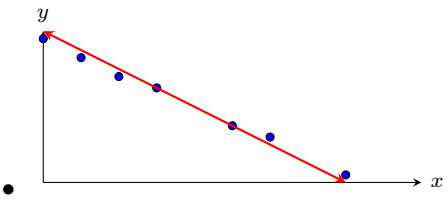
- The points follow a sinusoidal pattern with slight random variation, not a straight line.
- Therefore, the correlation is **non-linear**.

Ex 23:



For the scatter plot above, the correlation is **linear**.

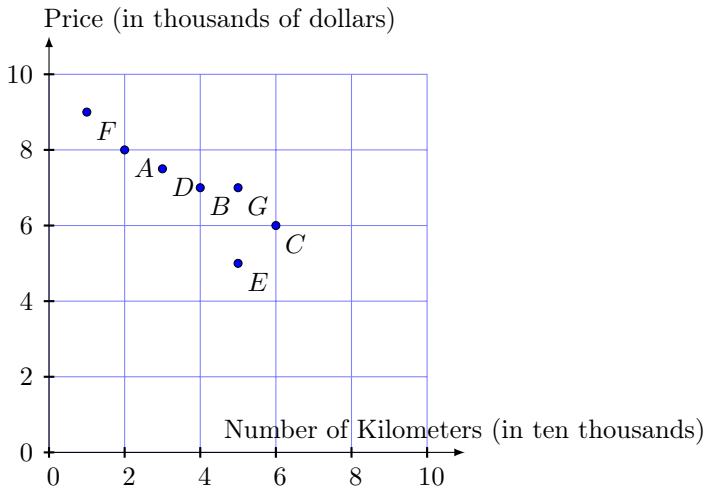
Answer:



- The points closely align with a straight line.
- Thus, the correlation is **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):

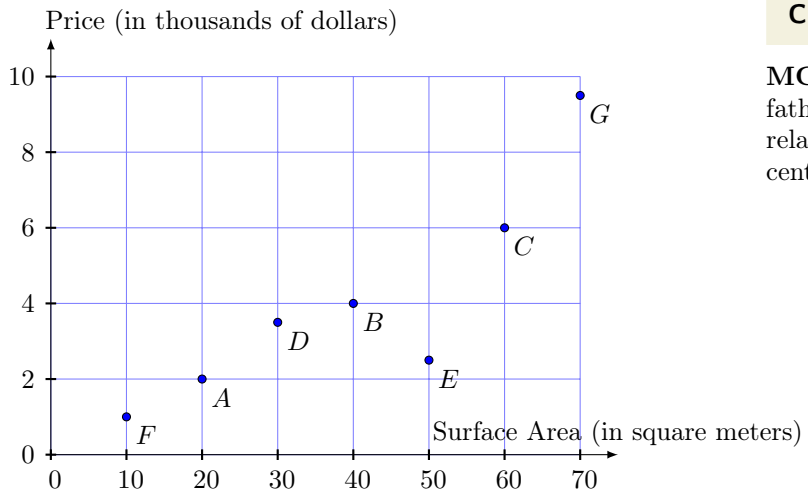


Which car offers a good deal, with a price significantly lower than expected for its mileage?
Select one: *E*.

Answer:

- Point G (50,000 km, \$7,000) is an outlier with a price much higher than the trend suggests for its mileage, making it a **bad deal**.
- Point E (50,000 km, \$5,000) is an outlier with a price significantly lower than expected for its mileage, making it a **good deal**.

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):

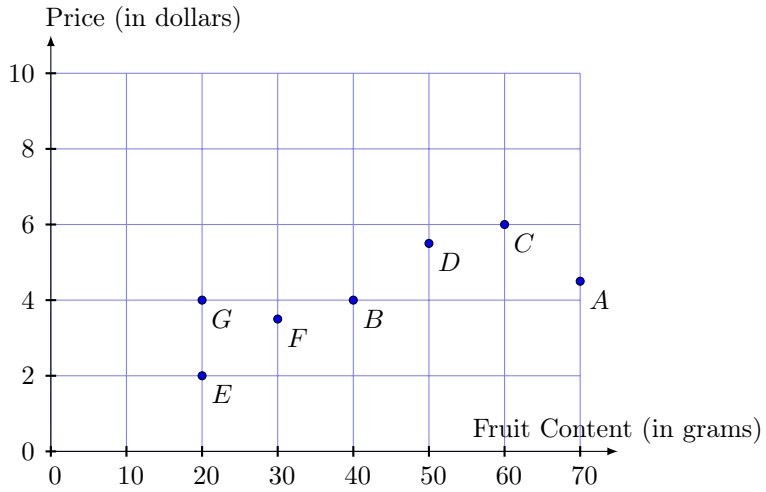


Which house is a good deal, with a price much lower than expected for its surface area?
Select one: $\boxed{E}$.

Answer:

- Point G (70 m², \$9,500) is an outlier with a price far higher than the trend predicts for its surface area, making it a **bad deal**.
- Point E (50 m², \$2,500) is an outlier with a price significantly lower than expected for its surface area, making it a **good deal**.

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):



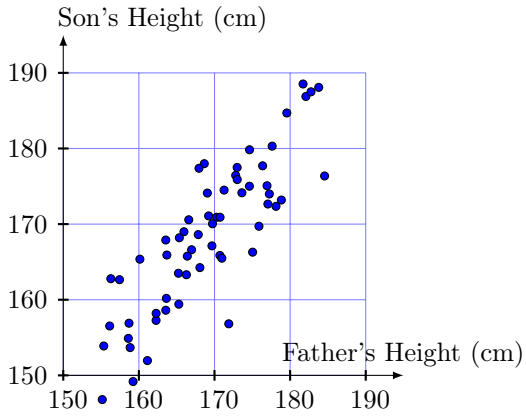
Which jam jar is a good deal, with a price much lower than expected for its fruit content?
Select one: $\boxed{A}$.

Answer:

- Point G (20 g, \$4) is an outlier with a price much higher than the trend suggests for its fruit content, making it a **bad deal**.
- Point A (70 g, \$4.50) is an outlier with a price significantly lower than expected for its fruit content, making it a **good deal**.

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.

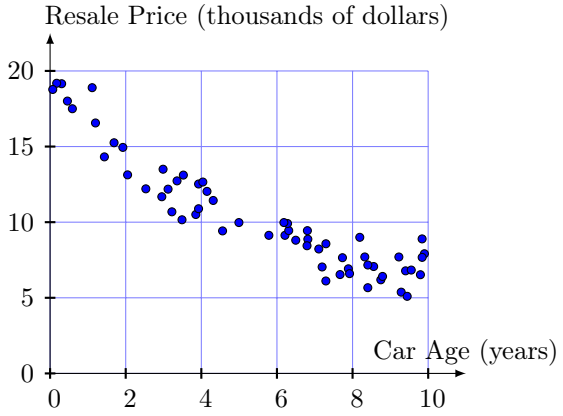
☐ Taller fathers tend to have shorter sons.

☐ There is no clear relationship between fathers' and sons' heights.

Answer:

- The points show an **upward** trend.
- As fathers' height (x) increases →, sons' height (y) tends to **increase** ↑.
- This indicates a positive correlation.
- Thus, taller fathers tend to have taller sons.

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.

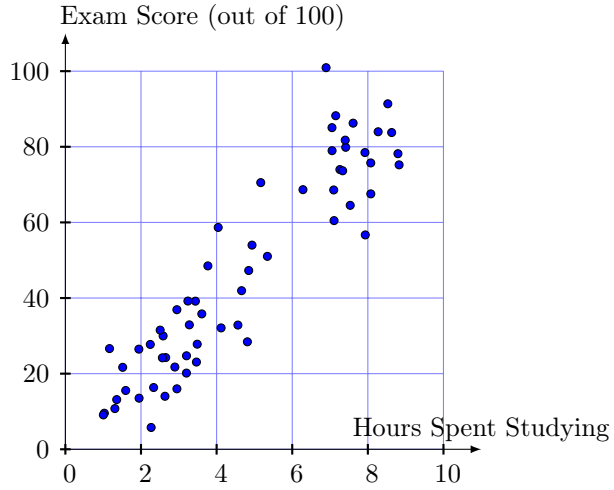
☒ Older cars tend to have lower resale prices.

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

Answer:

- The points exhibit a **downward** trend.
- As car age (x) increases, resale price (y) tends to **decrease** $\downarrow$.
- This indicates a negative correlation.
- Thus, older cars tend to have lower resale prices.

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.

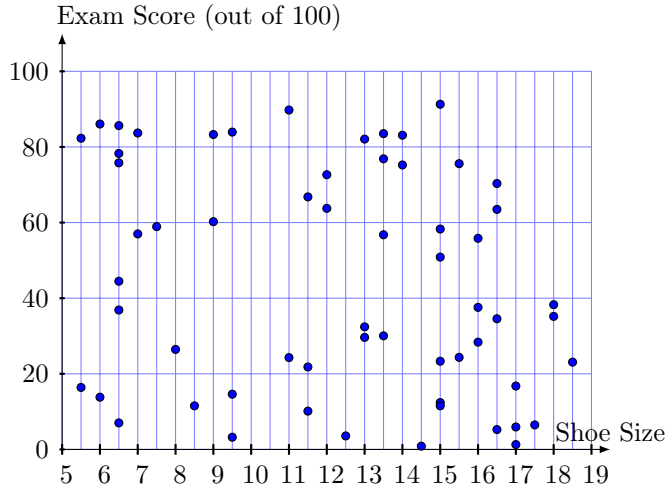
☐ Students who study more hours tend to score lower on the exam.

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

Answer:

- The points display an **upward** trend.
- As study hours (x) increase $\rightarrow$, exam scores (y) tend to **increase** $\uparrow$.
- This shows a positive correlation.
- Thus, students who study more hours tend to score higher on the exam.

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.

Answer:

- The points are widely scattered with no clear trend.
- There is no apparent pattern or correlation between shoe size and exam score.
- Thus, 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: **a causation**.
- Feeling strong and living longer: **just a correlation**.
- Eating healthy and living longer: **a causation**.

Answer:

- **Eating healthy and feeling strong**: This is **a causation**. Eating healthy, with a diet rich in nutrients like vitamins, proteins, and minerals, directly improves physical strength and well-being by supporting muscle growth and energy levels. This is a clear cause-and-effect relationship: a nutritious diet causes a stronger body.

- **Feeling strong and living longer**: This is **just a correlation**. Feeling strong doesn't directly cause a longer life—other factors like genetics, medical care, or lifestyle choices play a bigger role. Similarly, living longer doesn't necessarily cause someone to feel strong. These two might be associated because eating healthy could lead to both, but there's no direct causal link between them.
- **Eating healthy and living longer**: This is **a causation**. Eating healthy reduces the risk of chronic diseases (like heart disease or diabetes) and supports overall bodily function, directly contributing to a longer lifespan. This establishes a cause-and-effect link where a healthy diet causes increased longevity.

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

Answer:

- **Sunny weather and eating ice cream**: This is **a causation**. Sunny weather causes higher temperatures, which can make people feel hot and thirsty. This heat and thirst directly lead them to want to eat ice cream to cool down or quench their thirst. Unlike a mere association, there's a clear cause-and-effect link here.
- **Eating ice cream and getting sunburned**: This is **just a correlation**. Eating ice cream doesn't cause sunburns—sun exposure does. These two are associated because people often eat ice cream on sunny days when they're also more likely to get sunburned, but there's no direct causal relationship between them.
- **Sunny weather and getting sunburned**: This is **a causation**. Sunny weather directly causes sunburns because UV rays from the sun damage skin when people are exposed without protection. This is a straightforward cause-and-effect relationship.

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

Answer:

- **Working hard and getting good marks**: This is **a causation**. Working hard, such as studying diligently and practicing skills, directly improves understanding and performance, leading to higher marks. There's a clear cause-and-effect relationship: effort causes better academic outcomes.
- **Getting good marks and wearing glasses**: This is **just a correlation**. Getting good marks doesn't cause someone to wear glasses, nor does wearing glasses directly cause good marks. These two might be associated because students who work hard (and get good marks) might also strain their eyes and need glasses, but there's no direct causal link between the two.
- **Working hard and wearing glasses**: This is **a causation**. Working hard, especially through extensive reading or screen time, can strain the eyes over time, leading to vision problems that require glasses. This establishes a cause-and-effect link where the effort directly contributes to the need for glasses.

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

Answer:

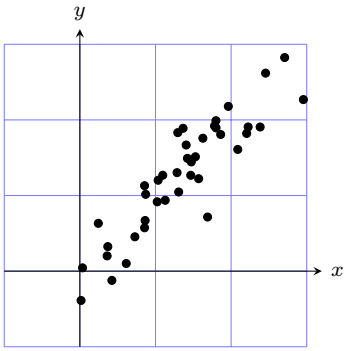
- **Exercising regularly and performing well in sports**: This is **a causation**. Exercising regularly builds strength, endurance, and skills, directly improving athletic performance in sports. This establishes a cause-and-effect link where regular exercise causes better sports outcomes.
- **Exercising regularly and having more energy**: This is **a causation**. Exercising regularly improves physical fitness, boosts metabolism, and enhances blood flow, directly leading to increased energy levels. This is a clear cause-and-effect relationship: consistent exercise causes greater energy.
- **Having more energy and performing well in sports**: This is **just a correlation**. Having more energy doesn't automatically cause better sports performance—skill, training, and technique are also critical. Likewise, performing well in sports doesn't directly cause energy levels to rise. These two might be associated because exercising regularly could lead to both, but there's no direct causal link between them.

E MEASURING CORRELATION

E.1 MEASURING CORRELATION

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

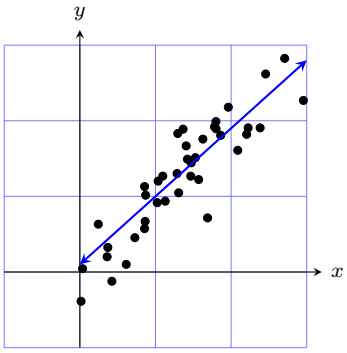
Choose the correct correlation coefficient (r) from the options provided:



The correlation coefficient is $r = \boxed{0.7}$.

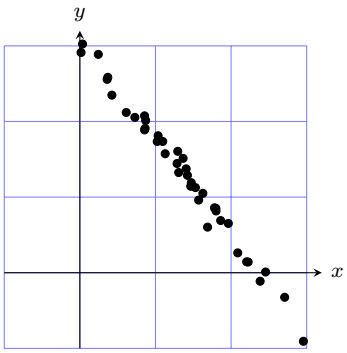
Answer:

- The scatter plot shows a **moderate positive correlation**. As x increases, y tends to increase, but the points are somewhat scattered around a linear trend.
- The correlation coefficient $r = 0.7$ fits this pattern, indicating a clear positive trend that isn't extremely tight.



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

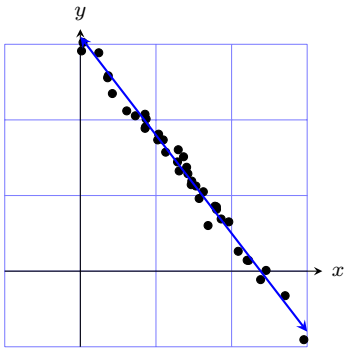
Choose the correct correlation coefficient (r) from the options provided:



The correlation coefficient is $r = \boxed{-0.9}$.

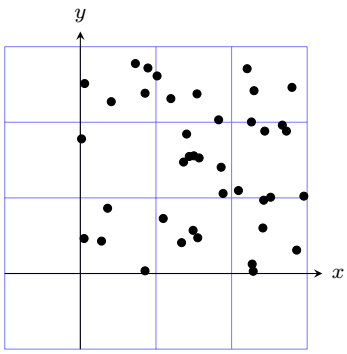
Answer:

- The scatter plot shows a **strong negative correlation**. As x increases, y decreases sharply, with points closely aligned along a downward trend.
- The correlation coefficient $r = -0.9$ reflects this strong negative trend, close to a perfect line but with slight scatter.



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:



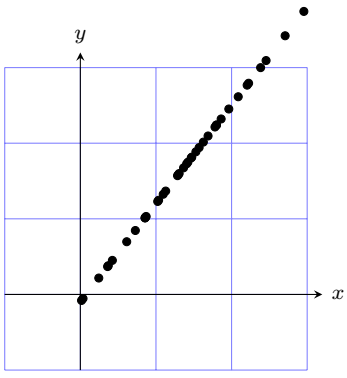
The correlation coefficient is $r = \boxed{0}$.

Answer:

- The scatter plot shows **no correlation**. The points are randomly distributed with no discernible pattern or trend, indicating that changes in x do not predict changes in y.
- The correlation coefficient $r = 0$ is appropriate, as it signifies no consistent association between the variables.

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

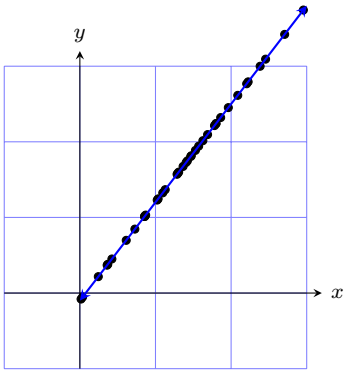
Choose the correct correlation coefficient (r) from the options provided:



The correlation coefficient is $r = \boxed{1}$.

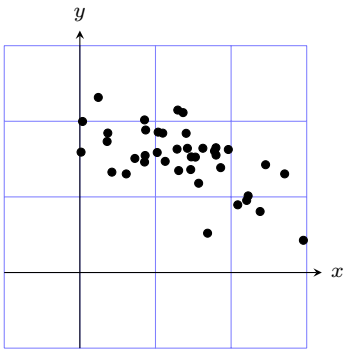
Answer:

- The scatter plot shows a **perfect positive correlation**. As x increases, y increases exactly along a straight line ($y = 1.3x - 0.1$), with no scatter (noise is 0).
- The points form a precise linear pattern, indicating a complete dependence of y on x.
- The correlation coefficient $r = 1$ is correct, as it represents a perfect positive linear relationship.



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

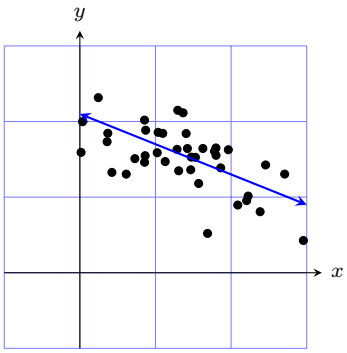
Choose the correct correlation coefficient (r) from the options provided:



The correlation coefficient is $r = \boxed{-0.7}$.

Answer:

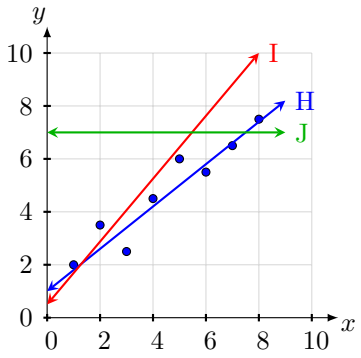
- The scatter plot shows a **moderate negative correlation**. As x increases, y tends to decrease, but the points are moderately scattered around a downward trend.
- The correlation coefficient $r = -0.7$ fits this pattern, indicating a moderate negative trend with some variability.



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

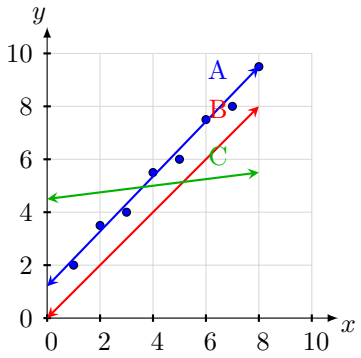
☐ J

☐ None of the lines fit the data.

Answer:

- The line of best fit should pass through the middle of the data points, minimizing the distance to most points and reflecting the general upward trend.
- Line H (blue) closely follows the data's positive slope, balancing points above and below it, making it the best fit.
- Line I (red) is too steep, overshooting most points, and Line J (green) is too flat and high, missing the trend.

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



Choose one answer:

☒ A

☐ B

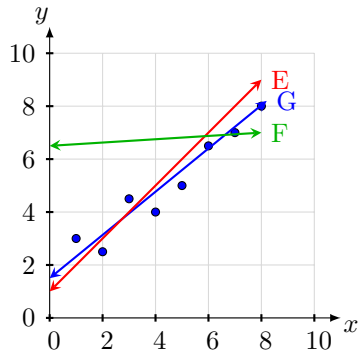
☐ C

☐ None of the lines fit the data.

Answer:

- The line of best fit should align with the data's steep upward trend, passing near the center of the points.
- Line A (blue) matches the data's slope and position, closely fitting the points from bottom-left to top-right.
- Line B (red) is too shallow, missing higher points, and Line C (green) is too flat and misplaced, ignoring the trend.

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



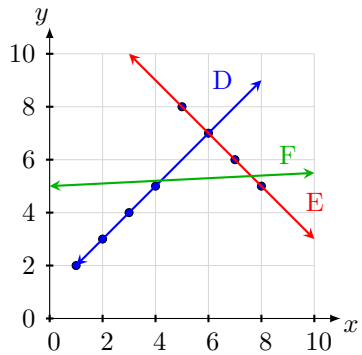
Choose one answer:

- ☐ E
- ☐ F
- ☒ G
- ☐ None of the lines fit the data.

Answer:

- The line of best fit should reflect the data's moderate upward trend, balancing the spread of points.
- Line G (blue) captures the general direction and slope, aligning well with most points across the plot.
- Line E (red) is too steep, overshooting early points, and Line F (green) is nearly flat, missing the upward trend.

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.

Answer:

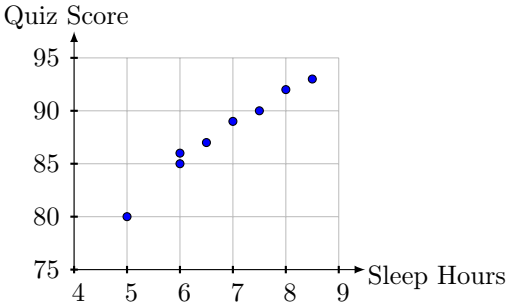
- The data shows two distinct trends: an upward trend (points 1–4) and a downward trend (points 5–8), forming a V-shape.
- No single straight line can fit both trends simultaneously, as the line of best fit assumes a consistent linear pattern.
- None of the lines (D, E, F) adequately represent the data's non-linear behavior.

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:



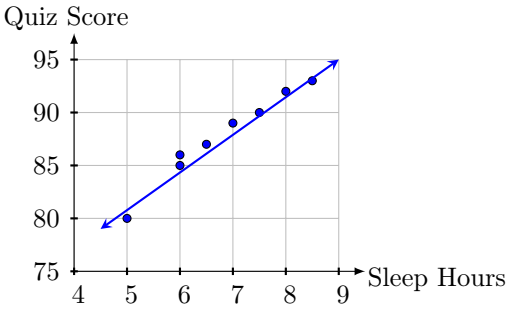
Is it reasonable to fit a line of best fit to this data?

Choose one answer:

- ☒ Yes
- ☐ No

Answer:

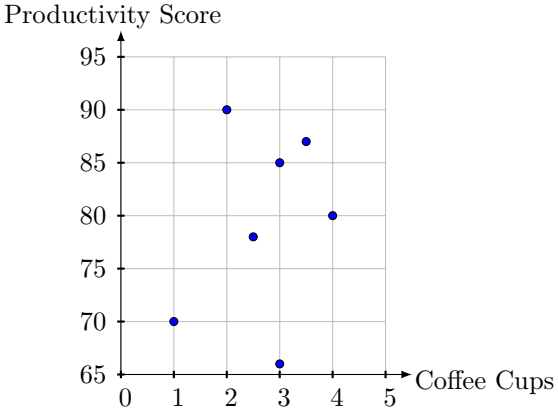
- The scatter plot shows a general upward trend: as sleep hours increase, quiz scores tend to rise, though not perfectly linearly.
- Despite some scatter, the pattern suggests a positive correlation, making it reasonable to fit a line of best fit.
- Example line: A line from (4.5, 79) to (9, 95) approximates the trend, balancing points above and below.



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

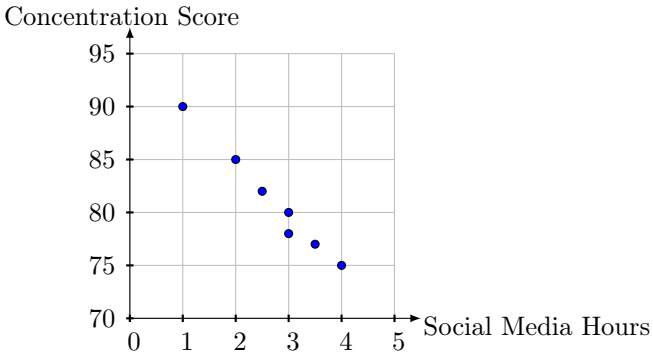
Answer:

- The scatter plot shows no clear trend: productivity scores vary widely without a consistent increase or decrease as coffee cups increase.
- With no discernible linear pattern (e.g., 3 cups yields both 66 and 85), fitting a line of best fit is not reasonable.

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

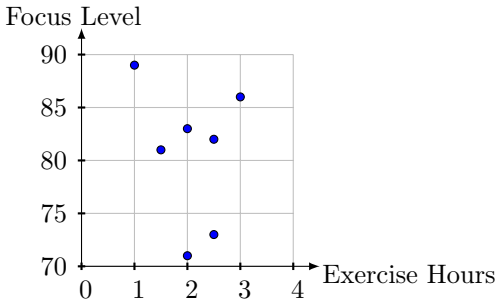
Answer:

- The scatter plot shows a downward trend: as social media hours increase, concentration scores generally decrease.
- The consistent negative pattern suggests a linear relationship, making it reasonable to fit a line of best fit.

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

Answer:

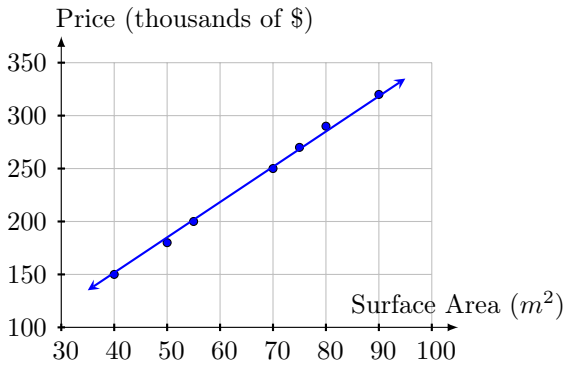
- The scatter plot shows no consistent trend: focus levels fluctuate without a clear increase or decrease as exercise hours rise.
- With no linear pattern (e.g., 2 hours yields 71 and 83), fitting a line of best fit is not reasonable.

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

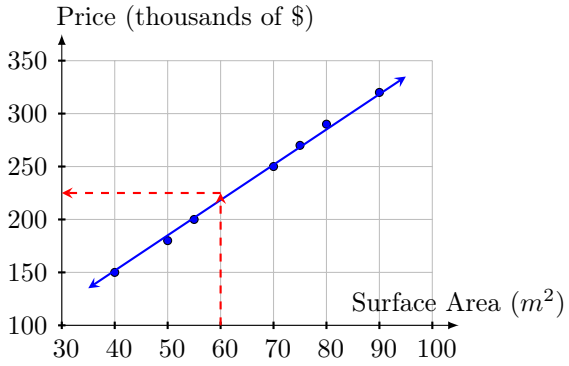
☐ $200,000

☒ $225,000

☐ $240,000

Answer:

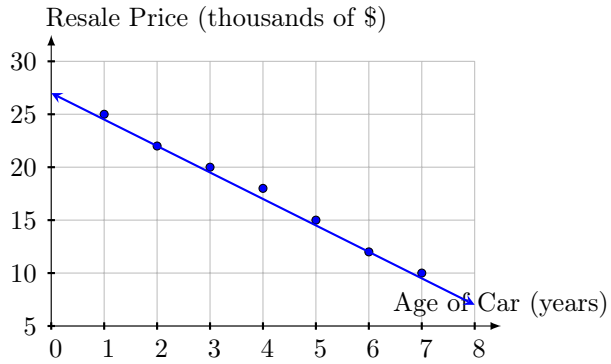
- From the line of best fit, locate 60 m^2 on the x-axis and trace vertically to the line, then horizontally to the y-axis.
- The line passes near (55, 200) and (70, 250); at 60 m^2, it intersects around 225, indicating a price of $225,000.



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

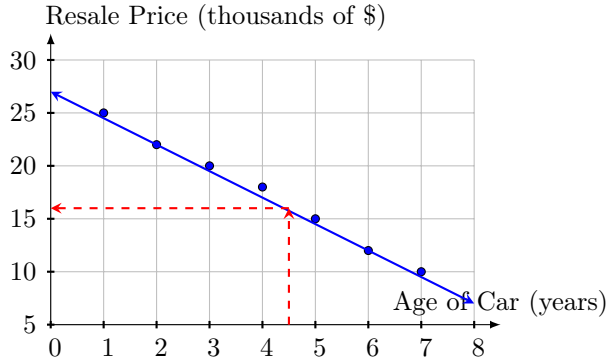
☐ $19,500

☐ $21,000

☐ $13,500

Answer:

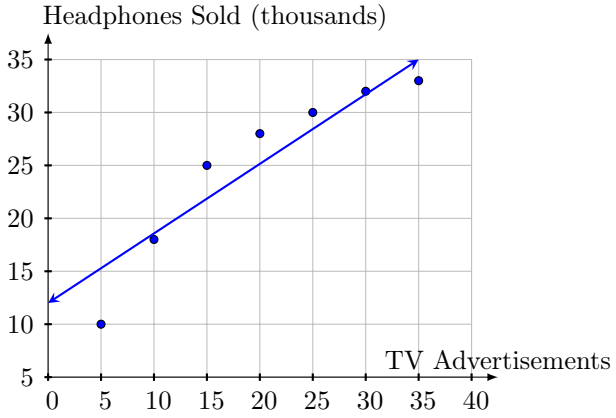
- From the line of best fit, locate 4.5 years on the x-axis and trace to the line, then to the y-axis.
- The line passes near (4, 18) and (5, 15); at 4.5 years, it intersects around 16, suggesting $16,000.



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

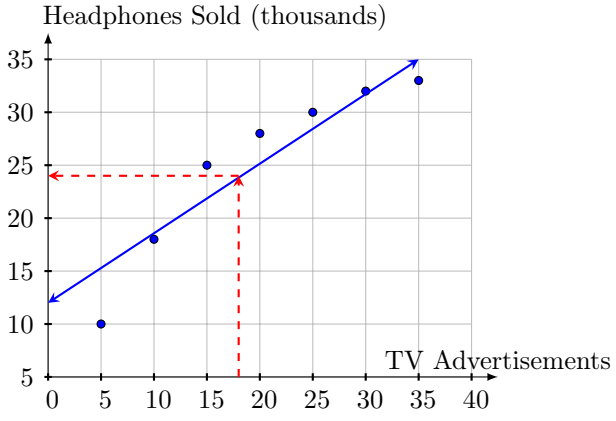
☒ 24,000

☐ 26,000

☐ 28,000

Answer:

- From the line of best fit, locate 18 ads on the x-axis and trace to the line, then to the y-axis.
- The line passes near (15, 25) and (20, 28); at 18 ads, it intersects around 24, suggesting 24,000 headphones sold.



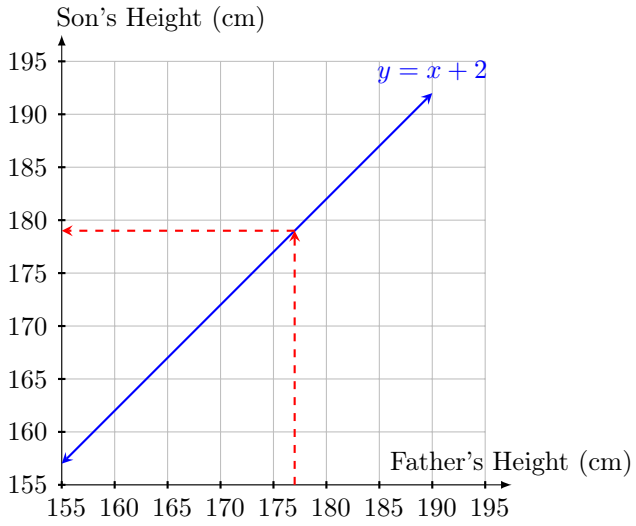
F.4 ESTIMATING VALUES WITH LINEAR EQUATIONS

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.

$\boxed{179}$ cm

Answer:

- Substitute $x = 177$ into the equation $y = x + 2$:
- Calculate: $y = 177 + 2 = 179$ cm.
- Thus, the estimated son's height is 179 cm.
- The graph below illustrates this estimation:

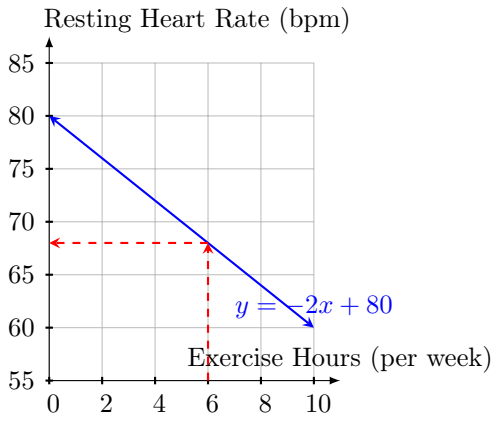


Ex 52: Ms. Lopez, a fitness coach, studied the relationship between weekly exercise hours and resting heart rates (in beats per minute, bpm) of her clients. She collected data from several clients, where x represents the hours of exercise per week and y represents the resting heart rate. After analysis, 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.

$\boxed{68}$ bpm

Answer:

- Substitute $x = 6$ into the equation $y = -2x + 80$:
- Calculate: $y = -2(6) + 80 = -12 + 80 = 68$ bpm.
- Thus, the estimated resting heart rate is 68 bpm.
- The graph below illustrates this estimation:

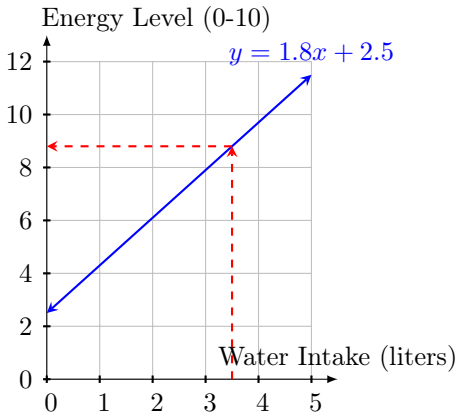


Ex 53: Mr. Patel, a nutritionist, investigated the relationship between daily water intake (in liters) and energy levels (on a scale of 0 to 10) among his clients. He collected data from several clients, where x represents the daily water intake and y represents the energy level. After analysis, he derived 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.

$\boxed{8.8}$

Answer:

- Substitute $x = 3.5$ into the equation $y = 1.8x + 2.5$:
- Calculate: $y = 1.8(3.5) + 2.5 = 6.3 + 2.5 = 8.8$.
- Thus, the estimated energy level is 8.8.
- The graph below illustrates this estimation:



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.

$\boxed{91.3}$

Answer:

- Substitute $x = 8.5$ into the equation $y = 4.2x + 55.6$:
- Calculate: $y = 4.2(8.5) + 55.6 = 35.7 + 55.6 = 91.3$.

- Thus, the estimated test score is 91.3.
- The graph below illustrates this estimation: