Name: _____ Date: _____

UNIT 5

2.5.8 - Introduction to Scatter Graphs

Vocabulary • Plotting Paired Data

A. KEY VOCABULARY (FILL IN THE BLANKS)

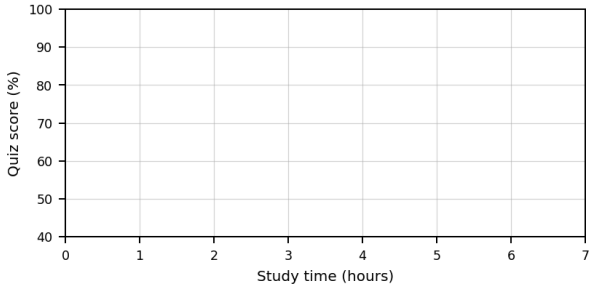
Use a short mathematical definition or description.

1. A **scatter graph** displays _____ .
2. **Positive correlation**: as one variable increases, _____ .
3. **Negative correlation**: as one variable increases, _____ .
4. **No correlation** means _____ .
5. An **outlier** is _____ .

B. PLOT PAIRED DATA

Each point represents one student. Plot all eight points accurately.

Study time (hours)	1	1.5	2	2.5	3	4	5	6
Quiz score (%)	52	58	61	65	69	78	84	91



(a) Describe the direction and strength: _____ (b) What does one point represent? _____

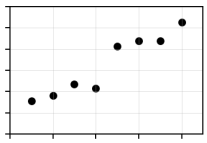
Name: _____ Date: _____

UNIT 5

2.5.8 - Introduction to Scatter Graphs

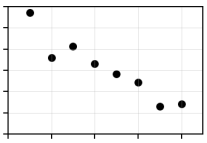
Correlation Patterns • Reading Scatter Graphs

C. QUICK CHECK: NAME THE PATTERN



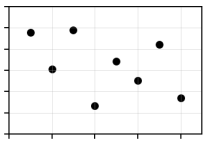
A

_____



B

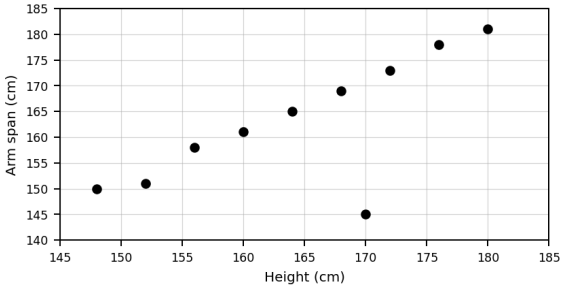
_____



C

_____

D. READ AND DESCRIBE A SCATTER GRAPH



1. Describe the correlation using direction and strength.

_____

2. Identify the approximate outlier coordinates. _____

3. Explain what the outlier means in context.

_____

Name: _____ Date: _____

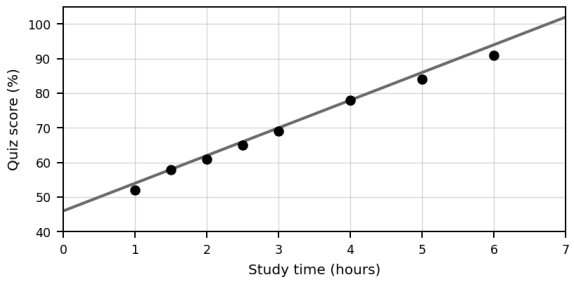
UNIT 5

2.5.8 - Introduction to Scatter Graphs

Line of Best Fit • Correlation and Causation

E. A LINE OF BEST FIT SUMMARISES THE TREND

The model shown is $\hat{y} = 8x + 46$, where x is study time and y is quiz score.



(a) Estimate the score for 4.5 hours of study. _____

(b) Estimate the study time for a score of 80%. _____

(c) Is the 4.5-hour estimate interpolation or extrapolation? _____

(d) A student studies 8 hours. Would using the line be interpolation or extrapolation? _____

(e) Why is the 8-hour prediction less reliable? _____

F. CORRELATION IS NOT CAUSATION

Ice-cream sales and sunburn cases both increase during some months.

1. **Suggest a third variable that could affect both.**

   _____

2. **Complete the statement: correlation shows an _____ , but it does not by itself prove _____ .**

Name: _____ Date: _____

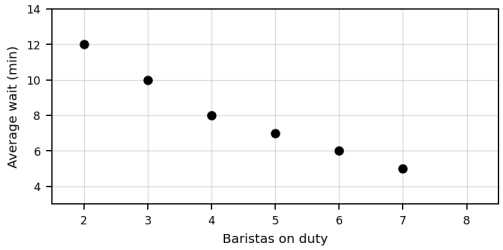
UNIT 5

2.5.8 - Introduction to Scatter Graphs

Real-World Modelling • Choosing Appropriate Variables

G. REAL-WORLD CONNECTION: CAFÉ WAITING TIMES

A café records the number of baristas on duty and the average customer wait.



(a) Describe the correlation. _____

(b) Explain the relationship in context. _____

(c) Estimate the wait with 8 baristas. _____ Is this interpolation or extrapolation?

_____

(d) Give one reason the estimate might be unreliable.

_____

H. IS A SCATTER GRAPH APPROPRIATE?

A scatter graph is most meaningful when both variables are quantitative measurements.

Pair of variables	Suitable?	Reason
Height and arm span		
Outdoor temperature and heating cost		
Shoe size and favourite colour		
Student ID number and mathematics score		

Name: _____ Date: _____

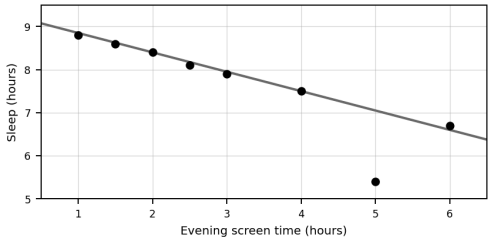
UNIT 5

2.5.8 - Introduction to Scatter Graphs

Extended Problems • Trend-Line Reasoning

I. EXTENDED PROBLEM: SCREEN TIME AND SLEEP

A line of best fit is modelled by $\hat{y} = 9.3 - 0.45x$.



(a) Predict sleep when x = 5 hours. _____

(b) The observed point at x = 5 is 5.4 hours. Find observed minus predicted. _____

(c) Why might this point be described as unusual? _____

(d) Predict at x = 9, then comment on reliability. _____

_____

K. CHALLENGE: INTERPRETING A TREND-LINE EQUATION

A coach models weekly training distance x (km) and marathon race time y (minutes) with $\hat{y} = 305 - 1.8x$. The observed training distances were 30 to 70 km.

1. **State the direction of correlation suggested by the model.** _____
2. **Estimate the race time for x = 55 km.** _____ **minutes.**
3. **Explain the meaning of the number -1.8 in context.**

   _____
4. **Estimate the race time for x = 90 km.** _____ **minutes. Is this interpolation or extrapolation?** _____
5. **Give two reasons not to treat the 90 km estimate as a guaranteed result.**

   _____

   _____

ANSWER KEY

Name: _____ Date: _____

UNIT 5

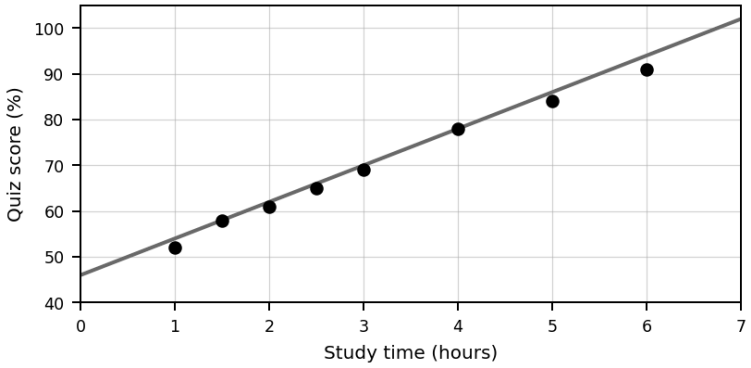
ANSWER KEY – 2.5.8 Scatter Graphs

Solutions: Vocabulary • Plotting Paired Data

A. KEY VOCABULARY

1. **A graph of paired values for two variables, shown as points.**
2. **The other variable tends to increase.**
3. **The other variable tends to decrease.**
4. **There is no clear increasing or decreasing pattern.**
5. **A point that does not fit the general pattern/trend.**

B. PLOT PAIRED DATA



(a) Strong positive correlation. (b) One student's paired study time and quiz score.

Name: _____ Date: _____

UNIT 5

ANSWER KEY – 2.5.8 Scatter Graphs

Solutions: Correlation Patterns • Reading Scatter Graphs

C. QUICK CHECK

A: positive correlation. B: negative correlation. C: no correlation.

D. READ AND DESCRIBE

1. **Strong positive correlation overall.**
2. **Approximately (170, 145).**
3. **The person is about 170 cm tall but has an arm span much lower than the trend would suggest.**

Name: _____ Date: _____

UNIT 5

ANSWER KEY - 2.5.8 Scatter Graphs

Solutions: Line of Best Fit • Correlation and Causation

E. LINE OF BEST FIT

(a) $y = 8(4.5)+46 = 82\%$, so about 82%.
(b) $80 = 8x+46$ gives $x = 4.25$ hours.
(c) Interpolation: 4.5 is inside the observed 1 to 6 hour range.
(d) Extrapolation: 8 is outside the observed range.
(e) The relationship may change outside the data range, so extending the line is less well supported.

F. CORRELATION IS NOT CAUSATION

1. **Hot weather / temperature / sunny weather.**
2. **Association; causation.**

Name: _____ Date: _____

UNIT 5

ANSWER KEY - 2.5.8 Scatter Graphs

Solutions: Real-World Modelling • Appropriate Variables

G. CAFÉ WAITING TIMES

(a) Strong negative correlation.
(b) As more baristas are on duty, average waiting time tends to decrease.
(c) A sensible visual estimate is around 4-5 minutes; extrapolation because 8 is beyond the observed 2-7 range.
(d) The pattern may not continue, the café may become crowded, or other factors such as order size may affect waiting time.

H. IS A SCATTER GRAPH APPROPRIATE?

Pair	Answer
Height and arm span	**Yes - both are quantitative measurements.**
Temperature and heating cost	**Yes - both are quantitative measurements.**
Shoe size and favourite colour	**No - favourite colour is categorical.**
Student ID and maths score	**No as a meaningful relationship - an ID is a label, not a measured quantity.**

Name: _____ Date: _____

UNIT 5

ANSWER KEY - 2.5.8 Scatter Graphs

Solutions: Extended Problems • Trend-Line Reasoning

I. SCREEN TIME AND SLEEP

(a) Predicted sleep = 9.3 - 0.45(5) = 7.05 hours.
(b) Observed - predicted = 5.4 - 7.05 = -1.65 hours.
(c) It lies well below the line and the general cluster, so it may be an outlier/unusual case.
(d) At x=9, predicted sleep = 5.25 hours. This is extrapolation beyond the observed 1-6 hour range, so it is less reliable.

K. TREND-LINE EQUATION

1. **Negative correlation.**
2. **305 - 1.8(55) = 206 minutes.**
3. **Within the model, each additional 1 km of weekly training is associated with about 1.8 fewer minutes of race time.**
4. **305 - 1.8(90) = 143 minutes; extrapolation.**
5. **90 km is outside the observed 30-70 km range, and correlation/model fit does not guarantee an individual's outcome; other variables also affect race time.**