

Homework without the hassle

Sukindi is an Australian learning platform that supports Teachers, Parents and Students with an interactive homework app and downloadable resources.



Scan to start your
free 7-day trial

www.sukindi.com.au

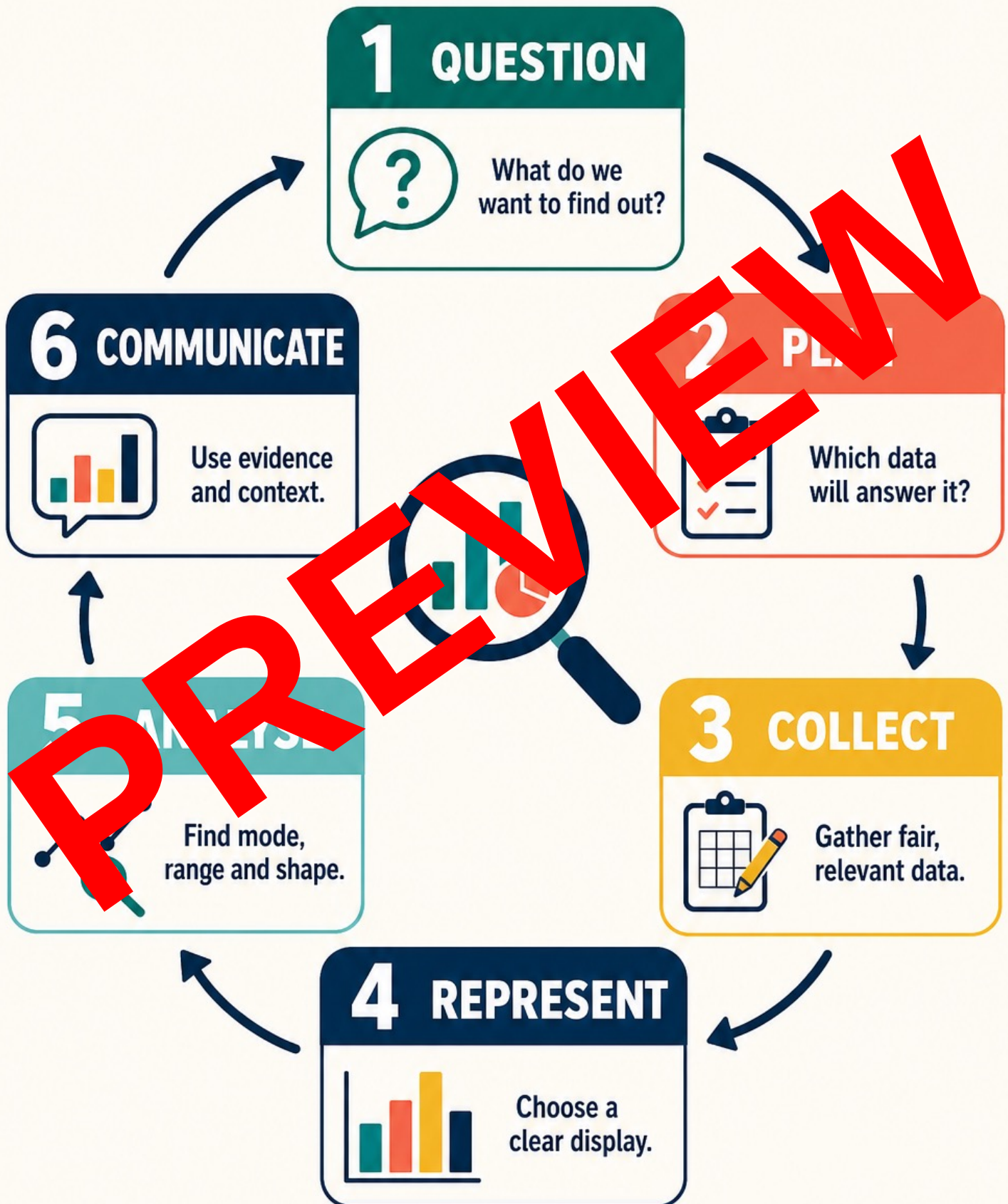


Investigating and Interpreting Data - Year 6

Table of Contents	
Presentation: Investigating and Interpreting Data	
Display: Statistical Investigation Cycle	
Display Set: Know Your Data + Trust the Graph?	
Activity: Data Detective Trail	
Activity: Distribution Match-Up	
Worksheet: Question to Data	
Worksheet: Mode, Range and Shape	
Worksheet: Comparing Distributions	
Worksheet: Media Claims Under the Microscope	
Worksheet: Statistical Investigation Planner	
Quiz: Investigating and Interpreting Data	
Assessment: Community Recreation Survey	
Rubric: Investigating and Interpreting Data Rubric	



STATISTICAL INVESTIGATION CYCLE



KNOW YOUR DATA

The type of variable guides the display.

NOMINAL

Names or groups with no order



Example: favourite fruit

ORDINAL

Categories with an order



Example: satisfaction rating

DISCRETE

Counted whole numbers



Example: goals scored

CONTINUOUS

Measured values



Example: travel time



ASK: What does each value represent?

TRUST THE GRAPH?

Before you believe a statistical claim, check:



METHOD

- ✓ Who was asked?
- ✓ How many?
- ✓ Was the question fair?



DISPLAY

- ✓ Is the scale even?
- ✓ Are labels clear?
- ✓ Is anything missing?



CONCLUSION

- ✓ Does the claim match the data?
- ✓ Are limits stated?
- ✓ Is another explanation possible?



**Strong conclusions
use evidence
and context.**

DATA DETECTIVE TRAIL



HOW TO PLAY

1. Roll and move.
2. Draw the matching card.
3. Solve it and explain your evidence.



1

CARD 1

PET VOTES



Cat 8 | Dog 12 | Fish 5

Which category is the mode?

2

CARD 2

READING MINUTES



10, 15, 20, 30, 35

Find the mode and range.

3

CARD 3

RAINFALL (mm)



18, 20, 21, 19, 42

Which value may be an outlier?

4

CARD 4

JOURNEY TIMES (min)



Route A: 18, 20, 22, 24

Route B: 12, 20, 28, 36

Which route has less variation?

5

CARD 5

LIBRARY VISITS



Mon 44 | Tue 47 | Wed 50

Thu 46 | Fri 80

How does 80 affect the range?

6

CARD 6

SPORT CHOICE



Football 15 | Netball 11

Basketball 9 | Other 5

Write one supported conclusion.

CARD 7

PLANT HEIGHTS (cm)

12, 13, 14, 14, 15



Where do the values cluster?

CARD 8

FOURTEEN SPEND (\$)

Class A: 3, 4, 4, 5, 5

Class B: 2, 3, 4, 6, 7



Compare the ranges.

CARD 9

DEVICE HOURS

1, 2, 2, 3, 8



Check the value 8. What could explain it?

CARD 10

CHESS CLUB SURVEY

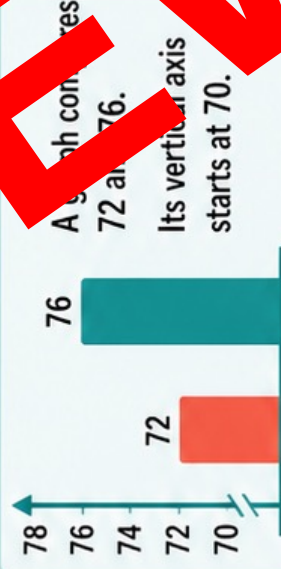


9 of 10 chess club members prefer chess.

Can this represent all Year 6 students? Why?

CARD 11

GRAPH SCALE



How might the scale exaggerate the difference?

CARD 12

TEMPERATURES (°C)



Melbourne:

21, 23, 24, 22, 25



Darwin:

31, 32, 31, 33, 32

Compare typical values and ranges.

DATA CARD A



BOOKS BORROWED

4, 5, 5, 5, 6

DATA CARD B



TRAINING LAPS

3, 4, 5, 6, 17

DATA CARD C



QUIZ SCORES

Class A: 10, 12, 14, 16
Class B: 11, 12, 13, 14

DATA CARD D



HOUSE COLOUR VOTES

Red 7 | Blue 12 | Green 5

DATA CARD E



SATISFACTION

Poor 2 | Fair 4
Good 9 | Excellent 5

DATA CARD F



GAME SCORES

Team X: 6, 6, 7, 7, 8
Team Y: 4, 6, 7, 8, 10

CLAIM CARD 1

Most values cluster around 5.



CLAIM CARD 2

One unusual value greatly increases the range.



CLAIM CARD 3

Class B is less variable than Class A.



CLAIM CARD 4

Blue is the modal category.



CLAIM CARD 5

Good is the most common ordered response.



CLAIM CARD 6

Both teams centre on 7, but Team X is more consistent.



DISTRIBUTION MATCH-UP



HOW TO PLAY

1. Pair one Data Card with one Claim Card.
2. Record the match and cite numerical evidence.

PAIR	DATA CARD	CLAIM CARD	NUMERICAL EVIDENCE
1			
2			
3			
4			
5			
6			



CHALLENGE

Could another claim also fit? Explain.

QUESTION TO DATA

Name: _____

Date: _____

A. IDENTIFY THE VARIABLE

Circle: nominal, ordinal, discrete or continuous.

- | | | | | |
|--|---------|---------|----------|------------|
| 1. Favourite lunch choice | nominal | ordinal | discrete | continuous |
| 2. Number of books read | nominal | ordinal | discrete | continuous |
| 3. Satisfaction: poor, fair, good or excellent | nominal | ordinal | discrete | continuous |
| 4. Travel time in minutes | nominal | ordinal | discrete | continuous |

B. REFINE THE QUESTION

Write one specific statistical question for each topic.

5. School travel: _____

6. Reading habits: _____

C. CHOOSE A METHOD

7. How long do Year 6 students read each night?

- A Ask only the library club.
B Randomly sample students across Year 6.
C Ask one friend.

Best choice: _____ Why? _____

8. Which lunch option is most popular in Year 6?

Variable: _____ Population: _____

Fair collection method: _____

QUESTION TO DATA

ANSWERS

Name: _____

Date: _____

A. IDENTIFY THE VARIABLE

Circle: nominal, ordinal, discrete or continuous.

- | | | | | |
|--|----------------|----------------|-----------------|-------------------|
| 1. Favourite lunch choice | <u>nominal</u> | ordinal | discrete | continuous |
| 2. Number of books read | nominal | ordinal | <u>discrete</u> | continuous |
| 3. Satisfaction: poor, fair, good or excellent | nominal | <u>ordinal</u> | discrete | continuous |
| 4. Travel time in minutes | nominal | ordinal | discrete | <u>continuous</u> |

B. REFINE THE QUESTION

Write one specific statistical question for each topic.

5. School travel: How many minutes do Year 6 students spend travelling to school?
6. Reading habits: How many minutes do Year 6 students read on a typical school night?

C. CHOOSE A METHOD

7. How long do Year 6 students read each night?

- A Ask only the library club.
B Randomly sample students across Year 6.
C Ask one friend.

Best choice: B Why? A random sample across Year 6 is more representative and reduces bias.

8. Which lunch option is most popular in Year 6?

Variable: lunch choice (nominal) Population: all Year 6 students

Fair collection method: Survey all Year 6 students, or a random sample across Year 6.

MODE, RANGE AND SHAPE

PAGE 1 • CALCULATE

Name: _____

Date: _____

Show your calculations beside each question.

DATA SET A • DAILY STEPS (thousands)

6, 8, 8, 9, 10, 11, 13

1. Mode: _____
2. Minimum: _____ Maximum: _____
3. Range: _____
4. Where do most values cluster?

DATA SET B • WATER BOTTLE REFILLS

1, 2, 2, 2, 3, 3, 8

5. Mode: _____
6. Range: _____
7. Which value may be an outlier? _____
8. How does that value affect the range?

COMPARE

9. Compare the two ranges. What do you notice?

MODE, RANGE AND SHAPE

PAGE 1 • CALCULATE

Name: **ANSWERS** _____

Date: _____

Show your calculations beside each question.

DATA SET A • DAILY STEPS (thousands)

6, 8, 8, 9, 10, 11, 13

1. Mode: 8

2. Minimum: 6 Maximum: 13

3. Range: 7

4. Where do most values cluster?

Most values cluster from 8 to 11.

$$13 - 6 = 7$$

DATA SET B • WATER BOTTLE REFILLS

1, 2, 2, 2, 3, 3, 8

5. Mode: 2

6. Range: 7

7. Which value may be an outlier? 8

8. How does that value affect the range?

It increases the range from 2 (without 8) to 7.

$$8 - 1 = 7$$

COMPARE

9. Compare the two ranges. What do you notice?

Both data sets have a range of 7, but Data Set B has an unusual value.

Name: _____

Date: _____

DATA SET C · TRAVEL TIMES (minutes)

9, 10, 10, 11, 12, 13, 30

1. Where do the values cluster?

2. Describe the gap in the data.

3. Which value may be an outlier? _____

4. Write a cautious conclusion about these travel times.

DATA SET D · QUIZ SCORES

Class A: 14, 15, 16, 17, 18, 19, 20

Class B: 10, 14, 16, 17, 18, 22, 25

The middle value of both sets is 17.

5. Range of Class A: _____ Range of Class B: _____

6. Which class has more spread? _____

7. How can equal middle values hide a difference?

8. Write one comparison using the middle value and range.

Name: _____ **ANSWERS**

Date: _____

DATA SET C · TRAVEL TIMES (minutes)

9, 10, 10, 11, 12, 13, 30

1. Where do the values cluster?

9 to 13 minutes.

2. Describe the gap in the data.

There are no values from 14 to 29.

3. Which value may be an outlier? **30**

4. Write a cautious conclusion about these travel times.

Most travel times are 9–13 minutes; 30 minutes is unusual and should be checked.

DATA SET D · QUIZ SCORES

Class A: 11, 15, 16, 17, 18, 19, 20

Class B: 10, 14, 16, 17, 18, 22, 25

The middle value of both sets is 17.

5. Range of Class A: **6** Range of Class B: **15**

6. Which class has more spread? **Class B**

7. How can equal middle values hide a difference?

The same middle value can occur with very different spreads.

8. Write one comparison using the middle value and range.

Both have middle value 17, but Class B has greater range (15 vs 6), so its scores vary more.

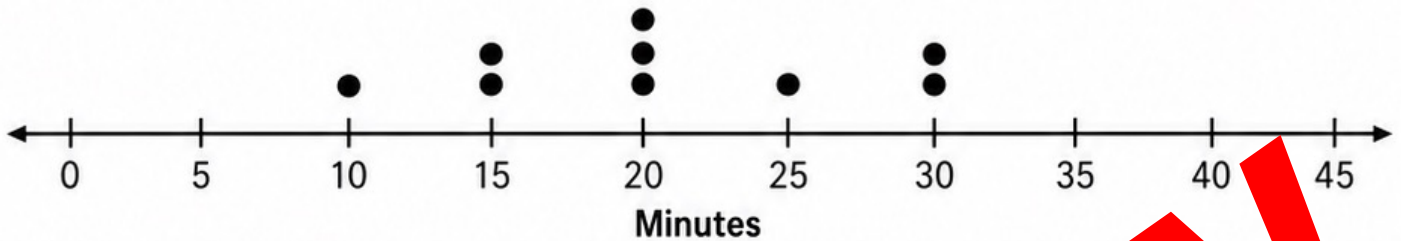
COMPARING DISTRIBUTIONS

Name: _____

Date: _____

Minutes read on one school night

Class W: 10, 15, 15, 20, 20, 20, 25, 30, 30



Class Z: 5, 10, 15, 20, 25, 30, 35, 40, 45



The middle value of each set is 20.

1. Mode of Class W: _____ Mode of Class Z: _____

2. Range of Class W: _____ Range of Class Z: _____

3. Which class is more consistent? Use evidence.

4. The middle values are equal. How are the distributions different?

5. Write one supported comparison conclusion.

COMPARING DISTRIBUTIONS

Name: **ANSWERS**

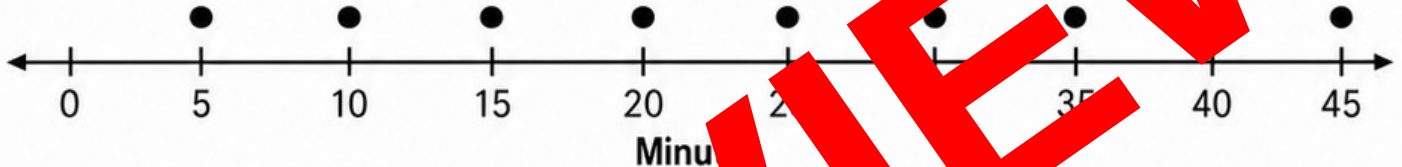
Date: _____

Minutes read on one school night

Class W: 10, 15, 15, 20, 20, 20, 25, 30, 30



Class Z: 5, 10, 15, 20, 25, 30, 35, 40, 45



The middle value of each set is 20.

1. Mode of Class W: 20 Mode of Class Z: no mode

2. Range of Class W: 20 Range of Class Z: 40

3. Which class is more consistent? Use evidence.

Class W; its range is 20 rather than 40, and its values cluster from 15 to 30.

4. The middle values are equal. How are the distributions different?

Both middle values are 20, but Class W clusters near 20 while Class Z spreads from 5 to 45.

5. Write one supported comparison conclusion.

Both classes centre on 20, but Class W is more consistent because its range is smaller.

COMPARING DISTRIBUTIONS

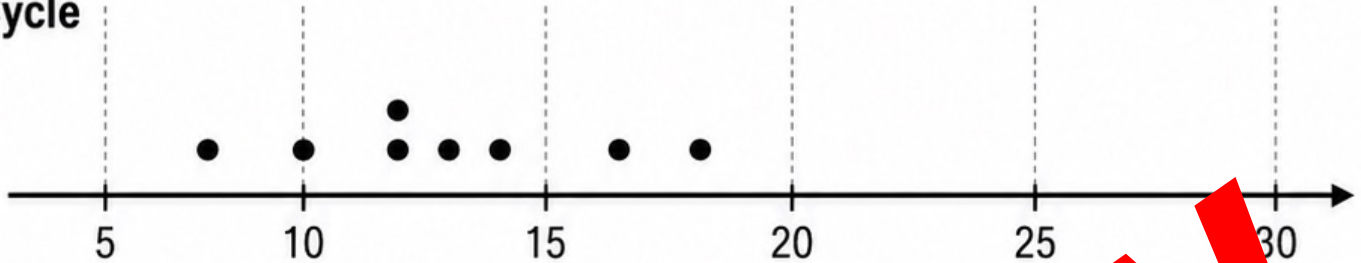
PAGE 2 • TRAVEL TO SCHOOL

Name: _____

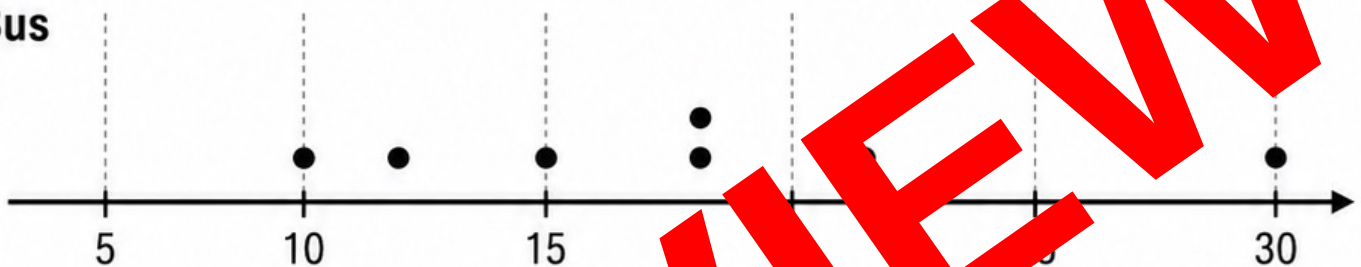
Date: _____

Travel time in minutes

Cycle



Bus



Cycle: 8, 10, 12, 12, 13, 14, 16, 18

Bus: 10, 12, 15, 18, 18, 24, 30

1. Mode of Cycle: _____ Mode of Bus: _____
2. Range of Cycle: _____ Range of Bus: _____
3. Which method has the higher typical time? Use evidence.

4. Which method is more variable? Use evidence.

5. A principal says, "Travel time is about the same for both methods."
Is this claim supported? Explain and state one limitation.

COMPARING DISTRIBUTIONS

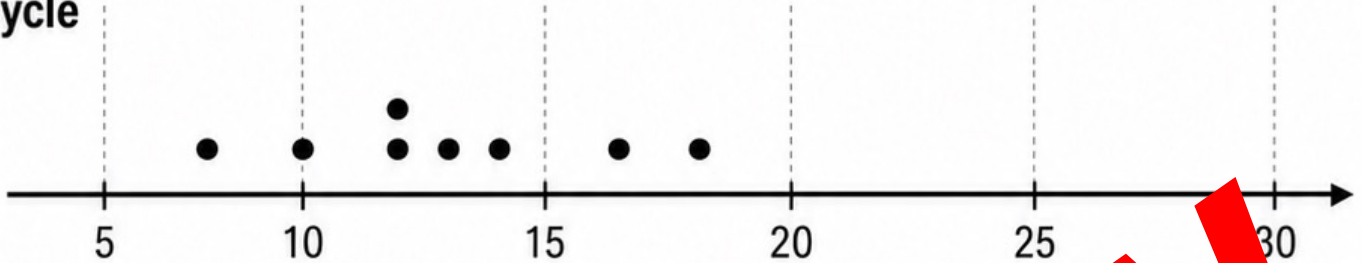
PAGE 2 • TRAVEL TO SCHOOL

Name: **ANSWERS**

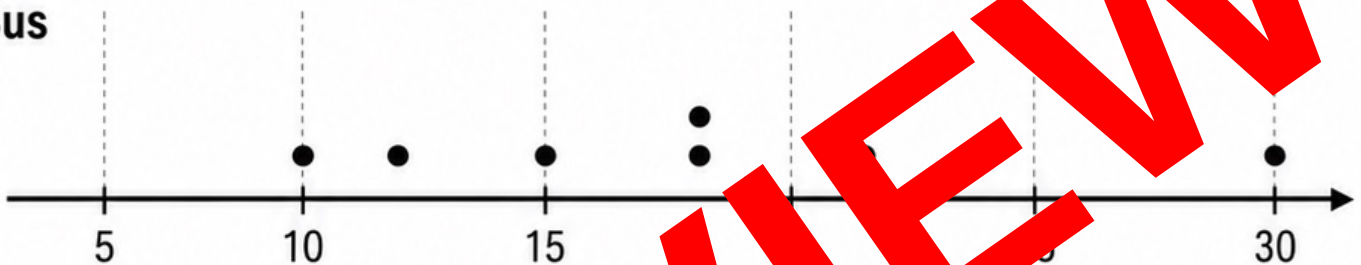
Date: _____

Travel time in minutes

Cycle



Bus



Cycle: 8, 10, 12, 12, 13, 14, 16, 18

Bus: 10, 12, 15, 18, 18, 24, 30

1. Mode of Cycle: 12 Mode of Bus: 18
2. Range of Cycle: 10 Range of Bus: 20
3. Which method has the higher typical time? Use evidence.

Bus; its mode is 18 rather than 12, and its values are generally higher.

4. Which method is more variable? Use evidence.

Bus; its range is 20 compared with 10 for Cycle.

5. A principal says, "Travel time is about the same for both methods."
Is this claim supported? Explain and state one limitation.

**No. Bus times are generally higher and more variable. Limitation:
only eight students were recorded in each group.**

MEDIA CLAIMS UNDER THE MICROSCOPE

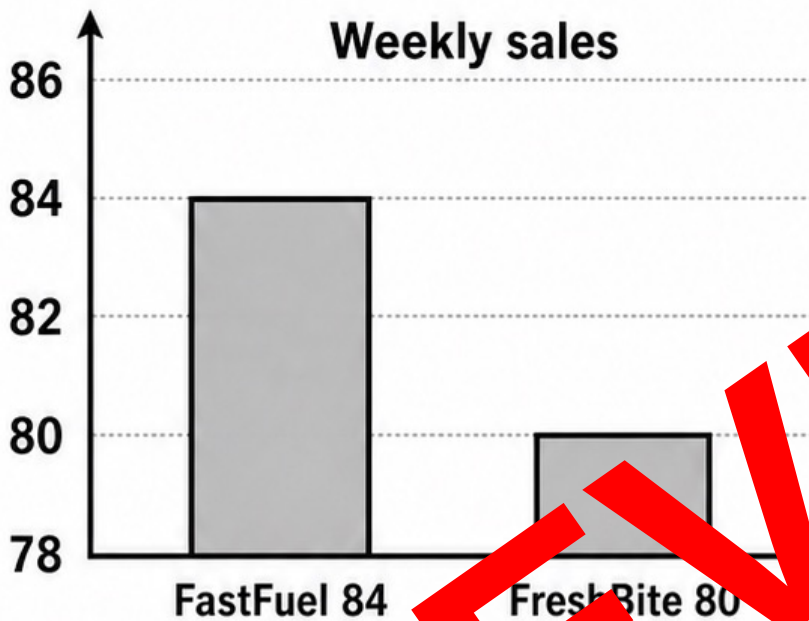
PAGE 1 • DISPLAY CHECK

Name: _____

Date: _____

FASTFUEL ADVERTISEMENT

Students choose FastFuel!



*FastFuel
is almost
twice as
popular!*

1. What is the actual difference in weekly sales? _____

2. Does the vertical axis start at zero? _____

3. How does the axis make the difference look larger?

4. Is the advertisement's claim supported? Use evidence.

5. Suggest one fairer way to display the data.

MEDIA CLAIMS UNDER THE MICROSCOPE

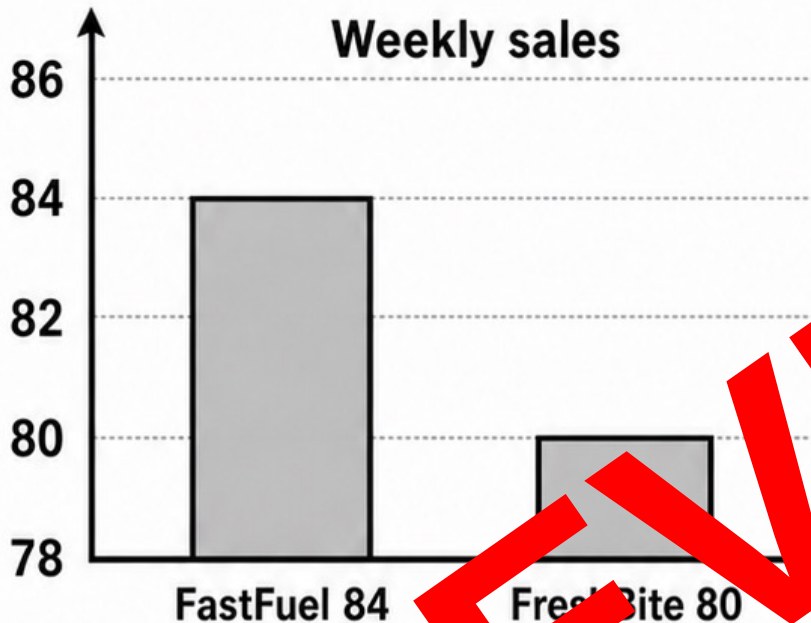
PAGE 1 • DISPLAY CHECK

Name: **ANSWERS**

Date: _____

FASTFUEL ADVERTISEMENT

Students choose FastFuel!



**FastFuel
is almost
twice as
popular!**

1. What is the actual difference in weekly sales? **4**

2. Does the vertical axis start at zero? **No**

3. How does the axis make the difference look larger?

It starts at 78, so a difference of 4 fills half the chart height.

4. Is the advertisement's claim supported? Use evidence.

No. 84 is only 4 more than 80, or 5% more—not almost twice.

5. Suggest one fairer way to display the data.

Use a vertical axis starting at 0 with equal intervals.

MEDIA CLAIMS UNDER THE MICROSCOPE



PAGE 2 • METHOD AND CONCLUSION

Name: _____

Date: _____



STREAMLINE NEWS



9 in 10 students want later school starts!

An online poll was posted on a night-owl forum.

90 responses: 81 yes, 9 no.

1. Which population does the headline seem to describe?

2. Who actually formed the sample?

3. Identify two possible sources of bias.

4. What percentage selected yes? _____ Can this result represent all students? _____

5. Rewrite the conclusion so it matches the evidence.

6. Plan a stronger follow-up investigation.

Population: _____

Sample: _____

Neutral question and method: _____

MEDIA CLAIMS UNDER THE MICROSCOPE

PAGE 2 • METHOD AND CONCLUSION

Name: **ANSWERS**

Date: _____



STREAMLINE NEWS



9 in 10 students want later school starts!

An online poll was posted on a night-owl forum.
90 responses: 81 yes, 9 no.

1. Which population does the headline seem to describe?
All students.
2. Who actually formed the sample?
The 90 people who chose to respond on the night-owl forum.
3. Identify two possible sources of bias.
Voluntary-response bias; the forum audience is not representative.
4. What percentage selected yes? **90%** Can this result represent all students? **No**
5. Rewrite the conclusion so it matches the evidence.
Among 90 forum respondents, 90% selected a later start; this cannot represent all students.
6. Plan a stronger follow-up investigation.
Population: **All students at the school.**
Sample: **A random or stratified sample across year levels.**
Neutral question and method: **Ask "Which school start time do you prefer?" in an anonymous school-wide survey.**

STATISTICAL INVESTIGATION PLANNER

PAGE 1 · QUESTION AND PLAN

Name: _____

Date: _____

A. FOCUS

1. Topic or problem: _____

2. Draft statistical question: _____

Check: Does it expect variation? Is it specific? Can data be collected safely?

3. Refined statistical question: _____

B. PEOPLE AND DATA

4. Population: _____

5. Sample and why it is fair: _____

6. Variables

VARIABLE	TYPE	UNIT OR CATEGORIES

C. COLLECTION

7. Collection method: _____

8. How will you reduce bias and keep collection consistent?

D. ETHICS CHECK

Participation is voluntary. Keep responses private. Ask neutral questions.

STATISTICAL INVESTIGATION PLANNER

PAGE 1 • QUESTION AND PLAN

Name: ANSWERS

Date: _____

A. FOCUS

1. Topic or problem: Year 6 travel time to school.
2. Draft statistical question: How long do Year 6 students take to travel to school?

Check: Does it expect variation? Is it specific? Can data be collected safely?
3. Refined statistical question: How do travel times differ between active travel and motor travel for Year 6 students?

B. PEOPLE AND DATA

4. Population: All Year 6 students at the school.
5. Sample and why it is fair: A stratified random sample from every Year 6 class; it includes varied travel methods.
6. Variables:

VARIABLE	TYPE	UNIT OR CATEGORIES
Travel time	continuous	minutes
Travel method	nominal	active or motor

C. COLLECTION

7. Collection method: Anonymous survey using the same question about a typical school day.
8. How will you reduce bias and keep collection consistent?
Sample across classes, use neutral wording and use one agreed definition of a typical day.

D. ETHICS CHECK

Participation is voluntary. Keep responses private. Ask neutral questions.

STATISTICAL INVESTIGATION PLANNER

PAGE 2 • ANALYSE AND COMMUNICATE

Name: _____

Date: _____

Use your collected data.

A. REPRESENT

1. Chosen display: _____

2. Why does it suit the variables and comparison?

B. ANALYSE

3. Record useful evidence.

Mode or typical value: _____

Range or variation: _____

Shape, clusters, gaps or outliers: _____

C. INTERPRET

4. Claim: _____

5. Evidence: _____

6. Limitation or uncertainty: _____

D. COMMUNICATE

7. Audience and format: _____

8. Final conclusion: _____

FINAL CHECK

My conclusion answers the question, uses evidence and stays within the data.

STATISTICAL INVESTIGATION PLANNER

PAGE 2 • ANALYSE AND COMMUNICATE

Name: **ANSWERS**

Date: _____

Use your collected data.

A. REPRESENT

- Chosen display: **Comparative dot plots.**
- Why does it suit the variables and comparison?
They show two numerical distributions on the same scale.

B. ANALYSE

- Record useful evidence.
Mode or typical value: **Active 12 min; motor 20 min.**
Range or variation: **Active 15 min; motor 18 min.**
Shape, clusters, gaps or outliers: **Active clusters lower;
motor is more spread with the high value.**

C. INTERPRET

- Claim: **Active travel times usually report shorter, less variable times.**
- Evidence: **Active: mode 12, range 15. Motor: mode 20, range 28.**
- Limitation or uncertainty: **Small school sample;
weather and traffic may change times.**

D. COMMUNICATE

- Audience and format: **Year 6 families; a one-page report
with comparative dot plots.**
- Final conclusion: **In this sample, active travel times were generally
shorter and less variable, but results may differ on other days.**

FINAL CHECK

My conclusion answers the question, uses evidence and stays within the data.

COMMUNITY RECREATION SURVEY

ASSESSMENT • PAGE 1 • DATA AND SUMMARIES

Name: _____

Date: _____

A local council recorded weekly active minutes for two after-school programs.

Riverside: 45, 50, 55, 60, 60, 60, 65, 70, 75, 80

Hillview: 30, 45, 50, 55, 60, 60, 65, 70, 75, 110

ACTIVE MINUTES	RIVERSIDE FREQUENCY	HILLVIEW FREQUENCY
30	0	1
45	1	1
50	1	1
55	1	1
60	3	2
65	1	1
70	1	1
75		1
80	1	0
110	0	1

1. What is the numerical variable? _____

What type of variable is it? _____

Mode of Riverside: _____ Mode of Hillview: _____

3. Range of Riverside: _____ Range of Hillview: _____

4. Which value may be an outlier? _____

5. Describe one cluster or shape feature in each distribution.

Riverside: _____

Hillview: _____

COMMUNITY RECREATION SURVEY

ASSESSMENT · PAGE 1 · DATA AND SUMMARIES

Name: ANSWERS

Date: _____

A local council recorded weekly active minutes for two after-school programs.

Riverside: 45, 50, 55, 60, 60, 60, 65, 70, 75, 80

Hillview: 30, 45, 50, 55, 60, 60, 65, 70, 75, 110

ACTIVE MINUTES	RIVERSIDE FREQUENCY	HILLVIEW FREQUENCY
30	0	1
45	1	1
50	1	1
55	1	1
60	3	2
65	1	1
70	1	1
75		1
80	1	0
110		1

1. What is the numerical variable? weekly active minutes

What type of variable is it? continuous

Mode of Riverside: 60 Mode of Hillview: 60

3. Range of Riverside: 35 Range of Hillview: 80

4. Which value may be an outlier? 110

5. Describe one cluster or shape feature in each distribution.

Riverside: Values cluster from 45 to 80, with three values at 60.

Hillview: Most values cluster from 45 to 75, with a large gap to 110.

COMMUNITY RECREATION SURVEY

ASSESSMENT • PAGE 2 • COMPARE AND RECOMMEND

Name: _____

Date: _____

EVIDENCE FROM PAGE 1

Riverside: mode 60, range 35

Hillview: mode 60, range 80

6. Which program has more consistent activity times? Use evidence.

7. How does Hillview's value of 110 affect the comparison?

8. A councillor says, "The programs are equally active because both modes are 60." Evaluate this statement.

9. The council wants more even participation. Which program should it investigate first, and what should it ask?

10. State one limitation of the original investigation.

11. Write a three-sentence recommendation for the council. Use a claim, numerical evidence and a limitation.

COMMUNITY RECREATION SURVEY

ASSESSMENT • PAGE 2 • COMPARE AND RECOMMEND

Name: **ANSWERS**

Date: _____

EVIDENCE FROM PAGE 1

Riverside: mode 60, range 35

Hillview: mode 60, range 80

6. Which program has more consistent activity times? Use evidence.

Riverside; its range is 35 rather than 80, and it has no distant high value.

7. How does Hillview's value of 110 affect the comparison?

It raises Hillview's maximum to 110 and its range to 80. Without 110, the range is 45.

8. A councillor says, "The programs are equally active because both modes are 60." Evaluate this statement.

Not fully. Equal modes show only the most common value; Hillview is much more variable.

9. The council wants more student participation. Which program should it investigate first, and what should it ask?

Investigate Hillview first. Ask what causes very low or high times and whether recording was consistent.

10. State one limitation of the original investigation.

Only ten students per program were recorded, and the sampling method is not stated.

11. Write a three-sentence recommendation for the council. Use a claim, numerical evidence and a limitation.

**Riverside is more consistent: range 35 compared with Hillview's 80.
Check Hillview's unequal participation and the 110-minute value.
Collect a larger representative sample before acting.**

COMMUNITY RECREATION SURVEY

ASSESSMENT • PAGE 3 • CLAIMS AND FOLLOW-UP

Name: _____

Date: _____



RIVERSIDE COMMUNITY PAGE

...

8 in 10 families prefer Riverside!

Poll: 10 followers of the Riverside program page.
8 selected Riverside; 2 selected Hillview.



12. What percentage selected Riverside? _____
13. Which population does the headline seem to describe?

14. Who actually formed the sample?

15. Identify two problems with the method.

16. Is the headline supported? Explain.

17. Rewrite the conclusion so it matches the evidence.

18. Plan a stronger follow-up investigation.
- Statistical question: _____
- Population and sample: _____
- Collection method: _____
- How findings will be communicated: _____

COMMUNITY RECREATION SURVEY

ASSESSMENT • PAGE 3 • CLAIMS AND FOLLOW-UP

Name: ANSWERS

Date: _____



RIVERSIDE COMMUNITY PAGE
8 in 10 families prefer Riverside!
Poll: 10 followers of the Riverside program page.
8 selected Riverside; 2 selected Hillview.



12. What percentage selected Riverside? 80%
13. Which population does the headline seem to describe?
All families in the community.
14. Who actually formed the sample?
Ten followers of the Riverside program page responded.
15. Identify two problems with the method.
Voluntary response bias; the sample favours Riverside; the sample is very small.
16. Is the headline supported? Explain.
No. Eight out of ten people from a biased sample cannot represent all families.
17. Rewrite the conclusion so it matches the evidence.
Among 10 Riverside page followers who responded, 8 selected Riverside.
18. Plan a stronger follow-up investigation.
Statistical question: Which program do community families prefer, and why?
Population and sample: All eligible families; a random sample across neighbourhoods.
Collection method: Use the same neutral anonymous survey for all sampled families.
How findings will be communicated: Report counts, percentages, method and limitations.

INVESTIGATING AND INTERPRETING DATA RUBRIC

Name: _____ Date: _____ Assessment: Community Recreation Survey

CRITERION	BEGINNING	DEVELOPING	ACHIEVING	CONFIDENT	EXCEEDING
QUESTION AND PLAN	Needs help to state a question.	States a question; plan is partly developed.	Poses a clear question and a workable plan.	Refines the question and controls bias.	Designs a rigorous, ethical investigation.
DATA INTERPRETATION	Reads few values accurately.	Describes one feature of a data set.	Finds mode, range and key shape features.	Interprets clusters, gaps and outliers.	Synthesises patterns, variation and context.
COMPARISON AND REASONING	Makes an unsupported comparison.	Uses one partly relevant feature.	Compares distributions with accuracy.	Uses several features to justify conclusions.	Evaluates patterns, exceptions and limitations.
CLAIM CRITIQUE	Accepts claims without checking.	Notices one concern.	Checks one conclusion.	Explains bias or misleading representation.	Evaluates uncertainty and alternative explanations.
COMMUNICATION	Response is incomplete or unclear.	Communicates a simple finding.	Gives a clear contextual conclusion.	Communicates conclusions and limitations.	Communicates precise, responsible recommendations.

Investigating and Interpreting Data



Learning Intention

- Interpret and compare distributions using mode, range and shape.
- Critique the method, display and conclusion behind a statistical claim.
- Plan a fair investigation and communicate the evidence in context.
- Success question: What does the data support—and what does it not support?



The Statistical Investigation Cycle

- QUESTION — pose or refine a question about expected variation.
- PLAN — define population, sample, variables, methods and ethics.
- COLLECT + REPRESENT — gather consistent data and choose a useful display.
- ANALYSE + INTERPRET — compare measures of range, shape, gaps and outliers.
- COMMUNICATE — state a conclusion, evidence and limitations.



What Kind of Data?

- NOMINAL — named groups with no order, such as travel method.
- ORDINAL — ordered categories, such as a satisfaction rating.
- DISCRETE — counted whole numbers, such as books read.
- CONTINUOUS — measured values, such as travel time.
- The variable type guides collection, storage and analysis.



Primary or Secondary?

- Primary data is collected firsthand for the current investigation.
- Secondary data already exists: reports, databases, journals or research.
- Choose a source that fits the purpose, time and available access.
- Always check who collected it, when, how and for which population.



Pose a Useful Question

- Broad topic: school travel.
- Statistical question: 'How do travel times differ between active travel and motor travel for Year 6 students?'
- Expect variation; avoid simple yes/no answer.
- Name the population, variable, comparison and conditions.
- Use neutral words that does not push a preferred response.



Plan Fair Data Collection

- Population = everyone the question is about, sample = the one measured.
- Include a representative spread across all relevant groups.
- Ask the same neutral question and use the same units and definitions.
- Protect privacy, make participation voluntary and record conditions
- A large sample helps but for selection matters too.



Choose a Helpful Display

- Categorical data: frequency table or bar graph.
- Numerical data: dot plot or another display that shows distribution.
- Comparisons need a shared scale, clear labels and units.
- Digital tools can help, but always specify axes and settings.
- The best display makes the question easier to answer.



Mode, Range and Shape

- MODE — the most frequent value or category
- RANGE — maximum minus minimum; a measure of spread.
- SHAPE — clusters, gaps, symmetry, skew and overall distribution.
- No single feature tells the whole story.
- Example: two sets can share mode but have ranges 35 and 80.



Variation and Outliers

- Variation is expected in statistical data.
- An outlier is a value far from the rest—not automatically an error.
- Check data entry, units, conversions and for a plausible explanation.
- Report how the value changes the range or conclusion.
- Never delete an unusual value without evidence.



Compare Two Distributions

- Compare the same variable in the same context and on the same scale.
- Typical value: mode or another meaningful center.
- Spread: range and how tightly values cluster.
- Shape: gaps, clusters and outliers.
- Evidence sentence: 'Both are on 20, but Class W is more consistent because its range is smaller.'



From Evidence to Conclusion

- CLAIM — answer the statistical question.
- EVIDENCE — cite exact values, frequencies, measures, range or shape.
- CONTEXT — name the group, variable and conditions.
- LIMIT — state uncertainty, sample limits or another explanation.
- Prefer 'In this sample...' to 'The data shows...'.



Statistics in the Media

- Statistical arguments appear in headlines, advertising and social posts.
- Ask what the source wants the audience to believe or do.
- Accurate numbers can be treated in a misleading way.
- Separate the raw evidence from the headline claim.



Trust the Method?

- Does the sample represent the claim population?
- How many people were included, and how were they selected?
- Could question wording or voluntary response create bias?
- Were conditions, units of collection procedures consistent?
- Who is missing from the sample?



Trust the Display?

- Does the axis start at a sensible point and use equal intervals?
- Are the title, labels, units and categories clearly displayed?
- Are groups or time periods correctly ordered?
- Do 3D effects, area or pie charts exaggerate differences?
- Would a different histogram display change the impression?



Trust the Conclusion?

- Does the claim match the data—or go further than the evidence?
- Is the conclusion limited to the correct population and context?
- Does it confuse association with cause?
- Are uncertainty and limitations acknowledged?
- Could another explanation also fit?



Guided Claim Critique

- CLAIM: '8 in 10 families prefer Riverside'
- EVIDENCE: 10 followers answered; 8 selected Riverside
- METHOD: a tiny voluntary sample from a Riverside page is biased.
- DISPLAY: 8 in 10 is numerically accurate for respondents.
- REWRITE: 'Among 10 page followers who responded, 8 selected Riverside.'



Your Mini-Investigation

- Pose a specific statistical question that expects variation.
- Define population, fair sample, variables, units and collection method.
- Choose a comparative display and use mode, range and shape.
- Communicate a claim with numerical evidence, context and a limitation.
- Use the plot to keep each decision visible.





Well Done!

Investigating and Interpreting Data Quiz



Question 1 . Variable types

Which variable is continuous?

- A. Favourite sport
- B. Number of siblings
- C. Satisfaction rating
- D. Travel time in minutes



Answer 1

D – Travel time in minutes.

Time is measured and can take any value within a range, so it is continuous.



Question 2 . Primary or secondary

True or false?

Using last year's council report for a new investigation is primary data because you read it yourself.



Answer 2

False.

The council collected the data for another purpose. It is secondary data for your investigation.



Question 3 . Statistical question

Rewrite this as a statistical question that reflects variation.
'Did you walk to school today?'



Answer 3

Example: 'How long does it take Year 6 students to walk to school to complete their journey?'

It names a group and variable and expects different values.



Question 4 · Fair Sample

Which sample best represents all Year 6 students?

- A. The first 20 students at football training
- B. 5 volunteers from one class
- C. Random students from every Year 6 class
- D. Students who follow the school sports page



Answer 4

C – Random students from every Year 6 class.
This spreads selection across the target population and reduces selection bias.



Question 5 · Mode and Range

For the data 6, 8, 8, 9, 11, calculate:

- the mode
- the range



Answer 5

Mode = 8, because 8 occurs most often.
Range = $11 - 6 = 5$.



Question 6 · Compare Distributions

Set A: 17, 19, 20, 20, 21, 23

Set B: 10, 15, 20, 20, 25, 30

Which statement is supported?

A. Set B has the smaller range.

B. Set A is less variable.

C. The sets have different modes.

D. Set A has an outlier.



Answer 6

B — Set A is less variable.
Both modes are 20. Set A has range 10 while Set B has range 20.



Question 7 . Outliers

True or false?

An outlier should always be deleted before analysing a distribution.



Answer 7

False.

Check entry, units and context first. A valid unusual value belongs in the analysis and may affect the range.



Question 8 · Misleading Graph

A bar graph shows 84% and 80%, but its vertical axis starts at 78%.

How could this scale affect a viewer's impression?



Answer 8

The truncated axis visually exaggerates the four percentage-point difference.

Show the full scale or clearly explain the scale before making the comparison.



Question 9 . Media Claim

A page says, '9 in 10 teens are night owls.' After a poll of 900 teenagers, 81 respond yes.

Which conclusion is most reasonable?

- A. All teens are night owls.
- B. Exactly 90% of every teen sleeps in.
- C. Among these 900 teenagers, 81 responded yes.
- D. The page proves late sleep causes tiredness.



Answer 9

C – Among these 90 page followers, 1 respondent replies.
It reports the actual sample and does not generalize beyond the respondents.



Question 10 · Responsible conclusions

Name three features a responsible statistical conclusion should include.



Answer 10

Evidence: exact values, frequencies, model, range, shape.

Context: the correct sample or population and variable.

Limitation: uncertainty, bias, sample size or any other explanation.





Well Done!