

Foundation GCSE: Data-Collection Sheets and Questionnaires

Name: _____

Class: _____

Instructions: Answer all questions. Questions are ordered from easiest (1) to hardest (15). Working out should be shown where appropriate.

1. State what tally marks are used for in a data-collection sheet.
2. Write down the tally representation for the number 7.
3. Identify one major flaw in the questionnaire question: "Why is football the best sport?"
4. Design a tally mark row to record a frequency count of 14.
5. Give an example of a closed question with a yes/no response format.

6. The incomplete frequency table below shows pets owned by students. Complete the missing tally and frequency for cats, given that total students equals 30.

Pet	Tally	Frequency
Dog		12
Cat		
Fish		9

7. Explain why a questionnaire response option stating "1–5 hours, 5–10 hours, 10+ hours" is considered ambiguous for someone who studies exactly 5 hours.
8. Identify why the question "How many sweets do you eat, 0, 1, or 2?" is leading or biased.
9. A data sheet tracks animal counts with groups totaling 42 observed specimens divided into 4 equal intervals. Calculate the size ratio per interval given by:

$$\frac{42}{4}$$

10. Suggest a suitable improvement for a survey question that asks respondents for their exact

annual salary without providing income brackets.

11. The frequency table below records modes of transport to school. Calculate the missing frequency value represented by x if the total number of recorded students is 85.

Transport	Frequency
Walk	34
Bus	28
Cycle	x
Other	5

12. State why open-ended questions can be difficult to analyze in large-scale statistical surveys.
13. A survey logs response tallies yielding frequencies of 12, 19, and 15. Calculate the mean frequency per category using:

$$\frac{12 + 19 + 15}{3}$$

14. A questionnaire uses response boxes: '[] 0–10', '[] 10–20', '[] 20–30'. Identify which box a value of exactly 10 belongs to, and explain why overlapping boundaries are poor design.
15. A researcher analyzes questionnaire return rates across three different regions summarized in the table below. Calculate the overall percentage return rate using:

$$\frac{140}{200} \times 100\%$$

Region	Distributed	Completed
North	80	60
South	120	80
Total	200	140

Answers & Worked Solutions

1. To record data quickly and easily as it is gathered.
2. |||| ||.
3. It is a leading question that assumes football is the best sport.
4. |||| |||| ||||.
5. "Do you own a bicycle? Yes / No".
6. Tally: |||| |||| (9); Frequency: 9.
7. It is overlapping (5 falls into both the first and second categories).
8. It restricts typical values and implies a low expectation.
9. 10.5.
10. Use non-overlapping continuous income intervals (e.g., 0 – 10,000, 10,001 – 20,000).
11. $x = 18$ ($85 - 67 = 18$).
12. Text responses are varied, making qualitative data hard to tally and quantify.
13. 15.33

$$\left(\frac{46}{3} \approx 15.33\right)$$
14. It is unclear; categories should be mutually exclusive (e.g., 0 – 9, 10 – 19).
15. 70%

$$\left(\frac{140}{200} = 0.7, 0.7 \times 100 = 70\right)$$