6.10 Statistical Enquiry and Misleading Graphs: 1-hour Pack

Calculators not allowed · about 60 minutes

Name: _____ Date: ____ / ____ / _____

Answer ALL questions. Show all your working.

1 A bar chart shows the pupils at a school: June = 47, July = 52. The vertical axis starts at 40. Explain why this could be misleading, and compare the visible height ratio with the real ratio of the two values, each as July ÷ June. Give both as decimals to 1 dp.

(3)

2 Mia wants to find out what pupils think about the canteen. She asks the first 50 pupils in the canteen queue. Give one reason why this sample may be biased.

(2)

3 A bar chart shows the sales of a phone: 2022 = 57, 2023 = 69. The vertical axis starts at 50 instead of 0, so the 2023 bar is drawn 19/7 times as tall as the 2022 bar.
Explain why the chart is misleading and work out how many times larger 2023 really is than 2022, to 2 decimal places.

(3)

© Revision Library | revisionlibrary.co.uk

4 A list has 600 names. A systematic sample of 50 names is wanted, starting from a random name. How many names are skipped between each selected name, and what is the gap between selected positions?

(2)

---

5 Hamza wants to find the mean number of hours of sleep of pupils at a school of 790 pupils. He asks the 9 members of the school football team. Give a reason why this might not give a reliable answer.

(2)

---

6 A bar chart shows the profits of a shop: Year 1 = 77, Year 2 = 83. The vertical axis starts at 70. Explain why this could be misleading, and compare the visible height ratio with the real ratio of the two values, each as Year 2 ÷ Year 1. Give both as decimals to 1 dp.

(3)

---

7 A questionnaire for pupils includes this question: "Do you agree that homework is a waste of time?"
Give one reason why this is a poor question.

(2)

© Revision Library | revisionlibrary.co.uk

8 Here are summary statistics for goals.
Team A: mean 26, range 15.
Team B: mean 14, range 10.
Compare the two sets of data.

(2)

9 Here are summary statistics for marks.
Class 8X: mean 11, range 11.
Class 8Y: mean 13, range 19.
Compare the two sets of data.

(2)

10 A bar chart shows the sales of a phone: 2022 = 47, 2023 = 56. The vertical axis starts at 40. Explain why this could be misleading, and compare the visible height ratio with the real ratio of the two values, each as 2023 ÷ 2022. Give both as decimals to 1 dp.

(3)

11 Hamza wants to find the mean number of hours of sleep of pupils at a school of 1145 pupils. He asks the 8 members of the school football team. Give a reason why this might not give a reliable answer.

(2)

© Revision Library | revisionlibrary.co.uk

12 A questionnaire for pupils includes this question: "Do you exercise often?"
Give one reason why this is a poor question.

(2)

13 A line of best fit for age of a car in years (x) and price in hundreds of pounds (y) has equation $y = 3x + 7$. Write down what the gradient 3 means in this context.

(2)

14 Here are summary statistics for minutes.
Ali: mean 30, range 7.
Bea: mean 15, range 11.
Compare the two sets of data.

(2)

15 A bar chart shows the pupils at a school: June = 45, July = 54. The vertical axis starts at 40. Explain why this could be misleading, and compare the visible height ratio with the real ratio of the two values, each as July ÷ June. Give both as decimals to 1 dp.

(3)

16 A list has 600 names. A systematic sample of 20 names is wanted, starting from a random name. How many names are skipped between each selected name, and what is the gap between selected positions?

(2)

© Revision Library | revisionlibrary.co.uk

17 The scatter graph for hours of revision (x) and test mark (y) uses data with x values from 6 to 23. The line of best fit is y = 4x + 23. Katie uses the line to predict y when x = 33. Is this prediction likely to be reliable? Give a reason.

(2)

18 Here are summary statistics for marks.
Class 8X: mean 24, range 13.
Class 8Y: mean 17, range 9.
Compare the two sets of data.

(2)

19 The scatter graph for hours of revision (x) and test mark (y) uses data with x values from 4 to 23. The line of best fit is y = 3x + 5. Katie uses the line to predict y when x = 33. Is this prediction likely to be reliable? Give a reason.

(2)

20 Hamza wants to find the mean number of hours of sleep of pupils at a school of 1910 pupils. He asks the 10 members of the school football team. Give a reason why this might not give a reliable answer.

(2)

21 A line of best fit for temperature (x) and ice creams sold (y) has equation y = 2x + 22. Use the line to estimate y when x = 13.

(2)

© Revision Library | revisionlibrary.co.uk

Mark scheme · 6.10 Statistical Enquiry and Misleading Graphs: 1-hour Pack

Question 1

axis not starting at zero	B1
visible ratio 12/7 and real ratio 52/47	M1
visible 1.7, real 1.1	A1

Answer: **Axis starts at 40; visible ratio 1.7, real ratio 1.1**

Question 2

identifies the sampling method as non-random	M1
only pupils who use the canteen are asked (not representative)	A1

Answer: **Only pupils who use the canteen are asked, so the sample is not representative**

Question 3

the axis does not start at zero, so the difference looks bigger	B1
69 ÷ 57	M1
1.21 times cao	A1

Answer: **The axis starts at 50, exaggerating the difference; the real ratio is 1.21**

Question 4

600 ÷ 50	M1
every 12th name (gap of 12)	A1

Answer: **every 12th name**

Question 5

identifies that the sample is not random or not representative	M1
only footballers are asked, who may sleep differently, and the sample is small relative to the school	A1

Answer: **The sample is not representative of all pupils**

Question 6

axis not starting at zero	B1
visible ratio 13/7 and real ratio 83/77	M1
visible 1.9, real 1.1	A1

Answer: **Axis starts at 70; visible ratio 1.9, real ratio 1.1**

Question 7

identifies a flaw	M1
it is a leading question: the wording suggests agreement that homework is a waste of time (accept equivalent explanations of bias or lack of neutral wording)	A1

Answer: **It is a leading question**

Question 8

comparison of the averages: Team A has the higher mean	B1
comparison of the spread: Team B is more consistent (smaller range)	B1

Answer: **Team A has the higher mean; Team B is more consistent as it has the smaller range**

Question 9

comparison of the averages: Class 8Y has the higher mean	B1
comparison of the spread: Class 8X is more consistent (smaller range)	B1

Answer: **Class 8Y has the higher mean; Class 8X is more consistent as it has the smaller range**

Question 10

axis not starting at zero	B1
visible ratio 16/7 and real ratio 56/47	M1
visible 2.3, real 1.2	A1

Answer: **Axis starts at 40; visible ratio 2.3, real ratio 1.2**

Question 11

identifies that the sample is not random or not representative	M1
only footballers are asked, who may sleep differently, and the sample is small relative to the school	A1

Answer: **The sample is not representative of all pupils**

Question 12

identifies a flaw	M1
any valid flaw explained, e.g. 'often' has no defined frequency or time period and different pupils may interpret it differently; accept no response options or an equivalent explanation	A1

Answer: **'often' is vague, with no time period**

Question 13

refers to the change in y for each 1 unit increase in x	M1
y increases by 3 for each extra unit of x	A1

Answer: **price in hundreds of pounds increases by 3 for each extra unit of age of a car in years**

Question 14

comparison of the averages: Ali has the higher mean	B1
comparison of the spread: Ali is more consistent (smaller range)	B1

Answer: **Ali has the higher mean; Ali is more consistent as it has the smaller range**

Question 15

axis not starting at zero	B1
visible ratio 14/5 and real ratio 54/45	M1
visible 2.8, real 1.2	A1

Answer: **Axis starts at 40; visible ratio 2.8, real ratio 1.2**

Question 16

600 ÷ 20	M1
every 30th name (gap of 30)	A1

Answer: **every 30th name**

© Revision Library | revisionlibrary.co.uk

Question 17

notes that 33 is outside the range of the data	**M1**
No, because it is extrapolation	**A1**

Answer: **No, it is outside the range of the data (extrapolation)**

Question 18

comparison of the averages: Class 8X has the higher mean	**B1**
comparison of the spread: Class 8Y is more consistent (smaller range)	**B1**

Answer: **Class 8X has the higher mean; Class 8Y is more consistent as it has the smaller range**

Question 19

notes that 33 is outside the range of the data	**M1**
No, because it is extrapolation	**A1**

Answer: **No, it is outside the range of the data (extrapolation)**

Question 20

identifies that the sample is not random or not representative	**M1**
only footballers are asked, who may sleep differently, and the sample is small relative to the school	**A1**

Answer: **The sample is not representative of all pupils**

Question 21

substitutes x = 13	**M1**
48 cao	**A1**

Answer: **48**

© Revision Library | revisionlibrary.co.uk