6.10 Statistical Enquiry and Misleading Graphs: 2-hour Pack

Calculators not allowed · about 120 minutes

Name: _____ Date: ____ / ____ / _____

Answer ALL questions. Show all your working.

1 Hamza wants to find the mean number of hours of sleep of pupils at a school of 390 pupils. He asks the 19 members of the school football team. Give a reason why this might not give a reliable answer.

(2)

2 Hamza wants to find the mean number of hours of sleep of pupils at a school of 845 pupils. He asks the 28 members of the school football team. Give a reason why this might not give a reliable answer.

(2)

3 A bar chart shows the sales of a phone: 2022 = 76, 2023 = 83. The vertical axis starts at 70 instead of 0, so the 2023 bar is drawn 13/6 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 Hamza wants to find the mean number of hours of sleep of pupils at a school of 795 pupils. He asks the 24 members of the school football team. Give a reason why this might not give a reliable answer.

(2)

---

5 Hamza wants to find the mean number of hours of sleep of pupils at a school of 870 pupils. He asks the 25 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 = 56, Year 2 = 62. The vertical axis starts at 50. 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 Here are summary statistics for minutes.
Ali: mean 25, range 14.
Bea: mean 23, range 17.
Compare the two sets of data.

(2)

© Revision Library | revisionlibrary.co.uk

8 A bar chart shows the sales of a phone: 2022 = 39, 2023 = 46. The vertical axis starts at 30 instead of 0, so the 2023 bar is drawn 16/9 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)

---

9 Here are summary statistics for marks.
Class 8X: mean 17, range 18.
Class 8Y: mean 14, range 19.
Compare the two sets of data.

(2)

---

10 A line of best fit for age of a car in years (x) and price in hundreds of pounds (y) has equation y = 6x + 14. Write down what the gradient 6 means in this context.

(2)

---

11 Here are summary statistics for marks.
Class 8X: mean 17, range 11.
Class 8Y: mean 14, range 6.
Compare the two sets of data.

(2)

© Revision Library | revisionlibrary.co.uk

12 The scatter graph for age of a car in years (x) and price in hundreds of pounds (y) uses data with x values from 3 to 20. The line of best fit is y = 2x + 30. Katie uses the line to predict y when x = 30. Is this prediction likely to be reliable? Give a reason.

(2)

---

13 Here are summary statistics for marks.
Class 8X: mean 24, range 5.
Class 8Y: mean 13, range 7.
Compare the two sets of data.

(2)

---

14 Here are summary statistics for minutes.
Ali: mean 18, range 20.
Bea: mean 15, range 9.
Compare the two sets of data.

(2)

---

15 Hamza wants to find the mean number of hours of sleep of pupils at a school of 1355 pupils. He asks the 35 members of the school football team. Give a reason why this might not give a reliable answer.

(2)

© Revision Library | revisionlibrary.co.uk

16 A bar chart shows the profits of a shop: Year 1 = 37, Year 2 = 44. The vertical axis starts at 30 instead of 0, so the Year 2 bar is drawn 2 times as tall as the Year 1 bar.
Explain why the chart is misleading and work out how many times larger Year 2 really is than Year 1, to 2 decimal places.

(3)

---

17 Here are summary statistics for goals.
Team A: mean 30, range 5.
Team B: mean 12, range 19.
Compare the two sets of data.

(2)

---

18 Here are summary statistics for marks.
Class 8X: mean 20, range 9.
Class 8Y: mean 19, range 10.
Compare the two sets of data.

(2)

---

19 A line of best fit for hours of revision (x) and test mark (y) has equation $y = 5x + 14$. Write down what the gradient 5 means in this context.

(2)

© Revision Library | revisionlibrary.co.uk

20 A questionnaire for pupils includes this question: "How old are you? Tick one box: 0-10, 10-20, 20-30"
Give one reason why this is a poor question.

(2)

21 Hamza wants to find the mean number of hours of sleep of pupils at a school of 490 pupils. He asks the 35 members of the school football team. Give a reason why this might not give a reliable answer.

(2)

22 A questionnaire for pupils includes this question: "How many hours of sport do you do? Tick one: 1 to 2, 3 to 4"
Give one reason why this is a poor question.

(2)

23 A bar chart shows the sales of a phone: 2022 = 66, 2023 = 69. The vertical axis starts at 60 instead of 0, so the 2023 bar is drawn 1.5 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

24 A bar chart shows the profits of a shop: Year 1 = 62, Year 2 = 67. The vertical axis starts at 60 instead of 0, so the Year 2 bar is drawn 3.5 times as tall as the Year 1 bar. Explain why the chart is misleading and work out how many times larger Year 2 really is than Year 1, to 2 decimal places.

(3)

---

25 Here are summary statistics for goals.
Team A: mean 20, range 10.
Team B: mean 17, range 4.
Compare the two sets of data.

(2)

---

26 A list has 600 names. A systematic sample of 60 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)

---

27 A line of best fit for hours of revision (x) and test mark (y) has equation $y = 4x + 16$. Write down what the gradient 4 means in this context.

(2)

© Revision Library | revisionlibrary.co.uk

28 Here are summary statistics for goals.
Team A: mean 30, range 13.
Team B: mean 29, range 6.
Compare the two sets of data.

(2)

29 A line of best fit for temperature (x) and ice creams sold (y) has equation $y = 2x + 8$.
Write down what the gradient 2 means in this context.

(2)

30 Hamza wants to find the mean number of hours of sleep of pupils at a school of 1480 pupils. He asks the 13 members of the school football team. Give a reason why this might not give a reliable answer.

(2)

31 A bar chart shows the sales of a phone: 2022 = 47, 2023 = 51. The vertical axis starts at 40 instead of 0, so the 2023 bar is drawn 11/7 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

32 A questionnaire for pupils includes this question: "How much do you enjoy the fantastic new library?"
Give one reason why this is a poor question.

(2)

---

33 A bar chart shows the sales of a phone: 2022 = 67, 2023 = 79. The vertical axis starts at 60. 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)

---

34 Here are summary statistics for marks.
Class 8X: mean 30, range 17.
Class 8Y: mean 22, range 11.
Compare the two sets of data.

(2)

---

35 Here are summary statistics for minutes.
Ali: mean 22, range 4.
Bea: mean 18, range 14.
Compare the two sets of data.

(2)

© Revision Library | revisionlibrary.co.uk

36 A line of best fit for age of a car in years (x) and price in hundreds of pounds (y) has equation y = 5x + 10. Write down what the gradient 5 means in this context.

(2)

---

37 Mia wants to find out what pupils think about the canteen. She asks the first 20 pupils in the canteen queue. Give one reason why this sample may be biased.

(2)

---

38 A bar chart shows the pupils at a school: June = 33, July = 45. The vertical axis starts at 30. 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)

---

39 Hamza wants to find the mean number of hours of sleep of pupils at a school of 790 pupils. He asks the 10 members of the school football team. Give a reason why this might not give a reliable answer.

(2)

---

40 A line of best fit for hours of revision (x) and test mark (y) has equation y = 5x + 11. Write down what the gradient 5 means in this context.

(2)

© Revision Library | revisionlibrary.co.uk

41 The scatter graph for hours of revision (x) and test mark (y) uses data with x values from 3 to 18. The line of best fit is $y = 3x + 14$. Katie uses the line to predict y when $x = 28$. Is this prediction likely to be reliable? Give a reason.

(2)

42 A line of best fit for age of a car in years (x) and price in hundreds of pounds (y) has equation $y = 4x + 5$. Write down what the gradient 4 means in this context.

(2)

© Revision Library | revisionlibrary.co.uk

Mark scheme · 6.10 Statistical Enquiry and Misleading Graphs: 2-hour Pack

Question 1

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 2

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 3

the axis does not start at zero, so the difference looks bigger	B1
83 ÷ 76	M1
1.09 times cao	A1

Answer: **The axis starts at 70, exaggerating the difference; the real ratio is 1.09**

Question 4

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 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 12/6 and real ratio 62/56	M1
visible 2, real 1.1	A1

Answer: **Axis starts at 50; visible ratio 2, real ratio 1.1**

Question 7

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 8

the axis does not start at zero, so the difference looks bigger	B1
46 ÷ 39	M1
1.18 times cao	A1

Answer: **The axis starts at 30, exaggerating the difference; the real ratio is 1.18**

Question 9

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

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

Question 10

refers to the change in y for each 1 unit increase in x	M1
y increases by 6 for each extra unit of x	A1

Answer: **price in hundreds of pounds increases by 6 for each extra unit of age of a car in years**

Question 11

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 12

notes that 30 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 13

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

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

Question 14

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

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

© Revision Library | revisionlibrary.co.uk

Question 15

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 16

the axis does not start at zero, so the difference looks bigger **B1**

44 ÷ 37 **M1**

1.19 times cao **A1**

Answer: **The axis starts at 30, exaggerating the difference; the real ratio is 1.19**

Question 17

comparison of the averages: Team A has the higher mean **B1**

comparison of the spread: Team A is more consistent (smaller range) **B1**

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

Question 18

comparison of the averages: Class 8X has the higher mean **B1**

comparison of the spread: Class 8X is more consistent (smaller range) **B1**

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

Question 19

refers to the change in y for each 1 unit increase in x **M1**

y increases by 5 for each extra unit of x **A1**

Answer: **test mark increases by 5 for each extra unit of hours of revision**

Question 20

identifies a flaw **M1**

the response boxes overlap at 10 and 20 (accept that the boxes do not cover all ages, or another valid flaw explained) **A1**

Answer: **The response boxes overlap (and do not cover all ages)**

Question 21

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 22

identifies a flaw **M1**

the response boxes do not cover every possible answer, such as 0 hours, between 2 and 3 hours, or more than 4 hours (accept that no time period is specified, or another valid flaw explained) **A1**

Answer: **The response boxes do not cover every possible answer (such as 0 hours or more than 4)**

Question 23

the axis does not start at zero, so the difference looks bigger **B1**

69 ÷ 66 **M1**

1.05 times cao **A1**

Answer: **The axis starts at 60, exaggerating the difference; the real ratio is 1.05**

Question 24

the axis does not start at zero, so the difference looks bigger **B1**

67 ÷ 62 **M1**

1.08 times cao **A1**

Answer: **The axis starts at 60, exaggerating the difference; the real ratio is 1.08**

Question 25

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 26

600 ÷ 60 **M1**

every 10th name (gap of 10) **A1**

Answer: **every 10th name**

Question 27

refers to the change in y for each 1 unit increase in x **M1**

y increases by 4 for each extra unit of x **A1**

Answer: **test mark increases by 4 for each extra unit of hours of revision**

Question 28

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 29

refers to the change in y for each 1 unit increase in x **M1**

y increases by 2 for each extra unit of x **A1**

Answer: **ice creams sold increases by 2 for each extra unit of temperature**

© Revision Library | revisionlibrary.co.uk

Question 30

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 31

the axis does not start at zero, so the difference looks bigger	B1
51 ÷ 47	M1
1.09 times cao	A1

Answer: **The axis starts at 40, exaggerating the difference; the real ratio is 1.09**

Question 32

identifies a flaw	M1
it is a leading question: 'fantastic' suggests a favourable answer (accept an equivalent explanation or another valid flaw explained)	A1

Answer: **It is a leading question (it suggests an answer)**

Question 33

axis not starting at zero	B1
visible ratio 19/7 and real ratio 79/67	M1
visible 2.7, real 1.2	A1

Answer: **Axis starts at 60; visible ratio 2.7, real ratio 1.2**

Question 34

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 35

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 36

refers to the change in y for each 1 unit increase in x	M1
y increases by 5 for each extra unit of x	A1

Answer: **price in hundreds of pounds increases by 5 for each extra unit of age of a car in years**

Question 37

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 38

axis not starting at zero	B1
visible ratio 15/3 and real ratio 45/33	M1
visible 5, real 1.4	A1

Answer: **Axis starts at 30; visible ratio 5, real ratio 1.4**

Question 39

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 40

refers to the change in y for each 1 unit increase in x	M1
y increases by 5 for each extra unit of x	A1

Answer: **test mark increases by 5 for each extra unit of hours of revision**

Question 41

notes that 28 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 42

refers to the change in y for each 1 unit increase in x	M1
y increases by 4 for each extra unit of x	A1

Answer: **price in hundreds of pounds increases by 4 for each extra unit of age of a car in years**

© Revision Library | revisionlibrary.co.uk