



H21-10min Box Plots from Cumulative Frequency and Outliers: 10-minute Pack

EDEXCEL 1ST0 · Calculator allowed · about 10 minutes

Total Marks

Name: _____ Date: ____ / ____ / ____

Answer ALL questions. Show all your working.

- 1** A PE teacher at a large secondary school wants to find out how many hours per week Year 10 students spend on organised sport, and whether this differs between boys and girls. There are 180 boys and 220 girls in Year 10. The teacher plans to select a stratified sample of 40 students.
- (a) Work out how many boys and how many girls should be included in the sample. (2)
- _____
- (b) State one advantage of using a stratified sample here rather than a simple random sample. (1)
- (c) The teacher collects the data using a self-completed questionnaire, then groups the responses into a frequency table before drawing a cumulative frequency graph. Explain why grouping this continuous data is a necessary part of processing it before a cumulative frequency graph can be drawn. (2)
- (d) Suggest one way the teacher could check whether the sample of 40 students is representative of the whole of Year 10. (1)

(Total for Question 1 is 6 marks)

2 A quality-control inspector weighs a sample of 50 bags of crisps from a factory. For this sample, the lower quartile of the masses is 24 g and the upper quartile is 31 g.

(a) Calculate the interquartile range of the masses. (1)

(b) Use the $1.5 \times \text{IQR}$ rule to calculate the lower and upper boundaries for outliers. (2)

(c) Five bags were set aside for closer inspection, with masses 12 g, 22 g, 28 g, 33 g and 44 g. Determine, showing your reasoning, which of these bags (if any) contain an outlier mass. (2)

(Total for Question 2 is 5 marks)

Mark scheme · H21-10min Box Plots from Cumulative Frequency and Outliers: 10-minute Pack

Question 1

- (a) **M1** 40/400 ($=0.1$) used as the sampling fraction, oe
- (a) **A1** 18 boys and 22 girls cao
- (a) **Answer: 18 boys, 22 girls**
- (b) **B1** valid advantage, e.g. it guarantees both boys and girls are represented in the sample in proportion to their numbers in the year group, making a comparison between them fairer/ more reliable oe
- (b) **Answer: It ensures boys and girls are represented in proportion to the year group, making comparisons between them fairer.**
- (c) **B1** recognises the number of hours is continuous/could take almost any value, so a list of 40 individual results would be hard to summarise/organise oe
- (c) **B1** recognises a cumulative frequency graph is built from running totals of grouped class frequencies, not raw individual values, so the data must be grouped into class intervals first oe
- (c) **Answer: Grouping organises the continuous raw data into class intervals, which is what a cumulative frequency graph is built from (running totals of grouped frequencies).**
- (d) **B1** valid check, e.g. compare a known characteristic of the sample (such as the gender split or average age) with the same characteristic for the whole of Year 10 to see if they are similar oe
- (d) **Answer: Compare a known characteristic of the sample (e.g. the gender split) with the whole year group to check they match.**

Question 2

- (a) **B1** 7 cao
- (a) **Answer: 7 g**
- (b) **M1** $1.5 \times 7 = 10.5$, oe
- (b) **A1** ft: lower boundary = 13.5 and upper boundary = 41.5
- (b) **Answer: Lower boundary = 13.5 g, upper boundary = 41.5 g**
- (c) **M1** ft: compares each of the 5 masses with both boundaries
- (c) **A1** ft: 12 g (below 13.5, a low outlier) and 44 g (above 41.5, a high outlier); 22 g, 28 g and 33 g are not outliers
- (c) **Answer: 12 g and 44 g are outliers; 22 g, 28 g and 33 g are not.**