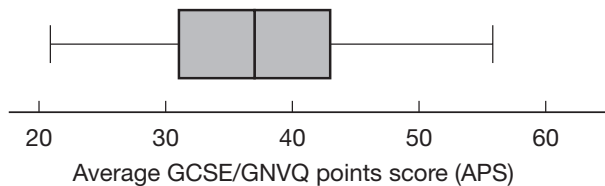


Check-Up

41

Box-and-whisker diagrams



The box-and-whisker diagram above provides information about the spread of average points scores (APS) in GCSE/GNVQ examinations for about 3200 maintained mainstream secondary schools in England in one particular year.

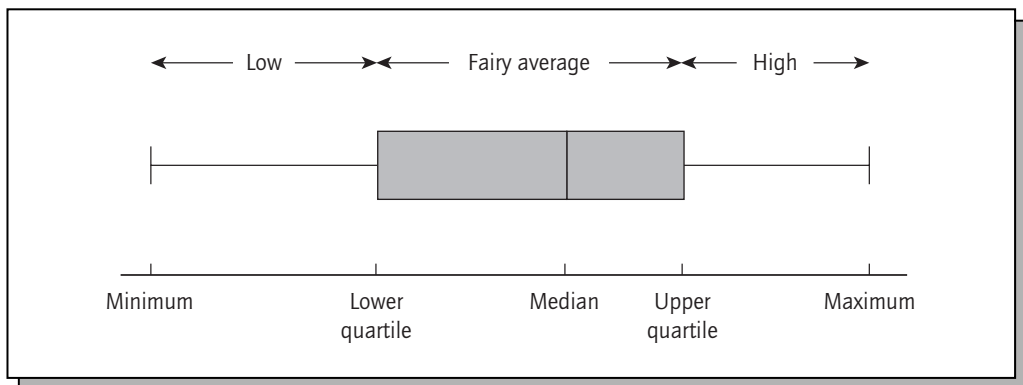
- What was the median value of the APS for these schools?
- What were the highest APS and the lowest APS achieved by any school?
- What was the range of APSs for the 1600 or so schools in the 'fairly average' bracket?
- How well did a school with an APS of 28 do compared to all schools nationally?
- About how many schools had an APS of 43 or more?

Answers to check-up 41

- a) about 37. b) about 56 and 21. c) from about 31 to 43.
 d) in the bottom 25%, i.e. they were a 'low-scoring' school.
 e) about 800 schools at least.

Discussion and explanation of check-up 41

A box-and-whisker diagram (also called a box-and-whisker plot) is simply a way of putting the information from a five-number summary into pictorial form. The 'box' extends from the lower quartile to the upper quartile. A line is usually drawn in the box to indicate the position of the median. The two 'whiskers' then extend out from the sides of the box from the LQ to the minimum value and from the UQ to the maximum value. At a glance you can then see the spread of values lying in the three groups that we have referred to in previous check-ups as low, fairly average and high. The left-hand whisker represents the bottom 25%, the box represents the middle 50% and the right-hand whisker represents the top 25%.



I have chosen to draw these diagrams with the scale going horizontally. You will also see them drawn vertically, but the principles are just the same.

Looking at the diagram in the check-up on the previous page, I can read off that, in terms of APS, the bottom 25% of schools (i.e. roughly the bottom 800 schools) had scores ranging from about 21 to 31. The middle 50% of schools (i.e. roughly the middle 1600 schools) had scores ranging from about 31 to 43. The top 25% (i.e. roughly the top 800 schools) had scores ranging from about 43 to 56. We cannot be too precise in reading off this information. The point

of the diagram is to give a quick overview, at a glance, of the distribution of scores. I can't tell from the diagram, for example, whether the scores of the top 800 schools were spread across the whole range from 43 to 56, or whether 799 of them scored 43 and just one scored 56!

A particular school can then look at the data represented in this form and identify their position in relation to the results nationally. A school with an APS of 28 is in the left-hand whisker, clearly in the bottom 25%. A school with an APS of 32 is in the box, the middle 50%, but well below the median. A school with an APS of 38 is pretty much in the middle, definitely 'fairly average'. A school with an APS of 48 is clearly in the high-scoring, top 25%.

These box-and-whisker diagrams are often used for making comparisons between sets of data, by drawing two or more such diagrams side by side, as in Further Practice question 41.1.

Summary of key ideas

- ◆ A box-and-whisker diagram is a picture of the way in which the values in a set of numerical data are distributed, using the minimum, the LQ, the median, the UQ and the maximum values as reference points.
- ◆ The box extends from the LQ to the UQ and represents the middle 50% of the set.
- ◆ One whisker extends from the LQ to the minimum value and represents the lowest 25% of the set.
- ◆ The other whisker extends from the UQ to the maximum value and represents the highest 25% of the set.

Further practice

- 41.1 The diagram shows the distributions of marks out of 100 in literacy and numeracy tests given to about 200 Year 7 pupils on entry to a secondary school. Are the following statements basically valid or invalid inferences from this data?
- A At least one pupil scored 100 marks for numeracy, but no pupil scored 100 marks for literacy.
 - B The median marks for literacy and numeracy were about the same.
 - C Nine pupils scored higher in numeracy than in literacy.
 - D A boy scoring 40 in both tests did equally well in literacy and numeracy compared to other pupils.
 - E A girl scoring 70 in both tests did better in numeracy than in literacy relative to other pupils.
 - F In general, these pupils found the numeracy test harder than the literacy test.
 - G The lowest 100 literacy scores ranged from 20 to 40.
 - H There was a much greater spread of scores in the top 25% for numeracy than for literacy.

