

GRADE 8

STATISTICS & DATA ANALYSIS

Mean • Median • Mode • Range • Distributions

6 Practice Worksheets • 6 Answer Keys

Worksheet 1 — Choosing the Right Statistic

1. A school wants to describe the typical number of books students read, but a few students read unusually large numbers. Which measure would you consider and why?
2. For the data 12, 13, 14, 15, 16, 17, 48, calculate the mean and median. Which gives a better description of the centre? Explain.
3. A shoe shop wants to know the most common shoe size sold. Which statistic should it use?
4. The data are 8, 9, 9, 10, 11, 11, 11, 12, 13. Find the mode, median and mean.
5. Give one situation where mode is more useful than mean.
6. Two data sets have the same mean. Does that mean they have the same median and range? Explain.
7. Challenge: Create a data set where the mean is greater than the median by at least 5.

Worksheet 2 — Outliers & Their Effect

1. For 14, 15, 16, 17, 18 find the mean and range.
2. Add the value 60 to the data. Find the new mean and range.
3. Explain exactly what changed and why.
4. A player scores 18, 20, 21, 22 and 24 points. In the next game she scores 70. Which measure of centre is affected most by the new score?
5. A student says, 'An outlier always changes the median dramatically.' Is this true? Explain.
6. Replace one value in 10, 12, 14, 15, 16 so that the mean increases but the median stays 14.
7. Challenge: Give an example where adding an extreme value changes the mean but leaves the median unchanged.

Worksheet 3 — Five-Number Summaries & Distribution

1. For the ordered data 4, 7, 9, 12, 15, 18, 21, identify the minimum, median and maximum.
2. For 3, 5, 8, 9, 12, 14, 17, 20, 24, find the median and identify the lower and upper halves.
3. Find the lower quartile and upper quartile for the data in Question 2 using the median-of-halves method.
4. Find the interquartile range (IQR) for Question 2.
5. Explain what the IQR tells us that the range does not.
6. Two data sets have the same range, but one has a much larger IQR. What could this tell you about their distributions?
7. Challenge: Construct an ordered data set of nine values with median 20, $Q1 = 12$ and $Q3 = 30$.

Worksheet 4 — Comparing Distributions

1. Class A test scores: 62, 65, 68, 70, 72, 74, 76. Class B: 40, 55, 65, 70, 75, 85, 100. Find the mean, median and range of each.
2. Which class has greater spread? Use more than one statistic to justify your explanation.
3. Which class appears more consistent? Explain what 'consistent' means in this context.
4. Two athletes have the same mean score over six matches. Athlete A has range 4 and Athlete B has range 18. What can you conclude?
5. Explain why reporting only the mean can sometimes hide important information.
6. Challenge: Create two seven-value data sets with the same median but different ranges.

Worksheet 5 — Missing Data & Reverse Reasoning

1. The mean of five numbers is 24. Four values are 17, 21, 25 and 29. Find the missing value.
2. The mean of six numbers is 18. Five values total 73. Find the sixth value.
3. The median of five ordered values is 18. The values are 7, 12, 18, x , 31. Give two possible values of x .
4. A data set has minimum 8, maximum 41 and range 33. If the maximum changes to 50, what is the new range?
5. The mean of four values is 15. One value is increased by 8 while the others stay the same. What is the new mean?
6. A set of five numbers has mean 20. One number is removed and the mean becomes 17. What can you say about the removed number? Explain.
7. Challenge: Find three different five-number data sets with mean 20.

Worksheet 6 — Statistics Detective Challenge

1. A report says: 'The average household income is \$62,000, so a typical family earns about \$62,000.' Explain why this conclusion may be misleading.
2. Data set A has mean 30, median 28 and range 12. Data set B has mean 30, median 34 and range 40. Describe two important differences between the distributions.
3. A school reports only the mean test score. Write three questions you would ask before deciding whether the result represents the whole class well.
4. A data set has five values. Its median is 16, its range is 20 and its mode is 12. Construct one possible data set.
5. A company changes one salary from \$40,000 to \$140,000. Explain which statistics are likely to change and which may remain unchanged.
6. A student claims: 'A larger range always means the data are less reliable.' Explain why this is not necessarily a valid conclusion.
7. Challenge: Design a mini statistical investigation. State the question, population, sample, variable, statistic(s) you would calculate, and how you would communicate the result.

Answer Key 1 — Choosing the Right Statistic

1. Median may be useful because unusually large values have less effect on it.
2. Mean = 19.3 (approximately), median = 15; the median better represents the centre because 48 is an outlier.
3. Mode.
4. Mode = 11, median = 11, mean = 10.9 (approximately).
5. Example: most common shoe size.
6. No. Different data sets can share a mean while having different medians and ranges.
7. Example: 10, 10, 10, 10, 35 gives mean 15 and median 10.

Answer Key 2 — Outliers & Their Effect

1. Mean = 16; range = 4.
2. New mean = 23.33 (approximately); range = 46.
3. The mean rises substantially and the range becomes much larger because 60 is far from the other values.
4. Mean is affected more strongly than median.
5. No. An outlier may have little or no effect on the median, especially in a large ordered data set.
6. Example: replace 10 with 30; data become 12, 14, 15, 16, 30. Mean increases while median remains 15, so this does not meet the target. A valid example is 8, 10, 14, 14, 16: mean 12.4, median 14. For the original data, replace 16 with 30: 10, 12, 14, 15, 30, median 14.
7. Example: 2, 3, 4, 5, 100; adding 100 changes the mean strongly while the median remains 4.

Answer Key 3 — Five-Number Summaries & Distribution

1. Minimum 4, median 12, maximum 21.
2. Median = 12; lower half = 3,5,8,9; upper half = 14,17,20,24.
3. $Q1 = 6.5$; $Q3 = 18.5$.
4. $IQR = 12$.
5. IQR measures the spread of the middle 50%, while range uses only the minimum and maximum.
6. The data may have similar extremes but one distribution may be more spread out in its middle values.
7. Example: 5, 12, 12, 16, 20, 24, 30, 30, 35.

Answer Key 4 — Comparing Distributions

1. Class A mean ≈ 69.6 , median 70, range 14. Class B mean ≈ 70 , median 70, range 60.
2. Class B has much greater spread; its range is 60 compared with 14.
3. Class A is more consistent because its scores are clustered more closely.
4. Their average performance is similar, but Athlete A is much more consistent.
5. The mean does not show spread, outliers or how values are distributed.
6. Example: 10, 12, 15, 18, 20, 22, 24 and 2, 12, 15, 18, 20, 22, 50 have the same median 18 but different ranges.

Answer Key 5 — Missing Data & Reverse Reasoning

1. 28.
2. 35.
3. Any x from 18 through 31 maintains the stated ordering; for example 20 or 25.
4. 42.
5. Original total = 60; new total = 68; new mean = 17.
6. The removed value was below 17, because removing a value below the original mean raises the mean. Since the new mean is lower, this statement needs careful comparison: the original total is 100; remaining four total 85, so removed value = 15.
7. Answers vary; any three five-number sets with total 100 are valid.

Answer Key 6 — Statistics Detective Challenge

1. Income data can be skewed by very high earners; median and spread would provide additional context.
2. Both have mean 30, but A has median 28 and small range, while B has median 34 and much larger range; their distributions differ substantially.
3. Examples: What is the median? What is the range/IQR? Are there outliers? How many students were included?
4. Example: 12, 12, 16, 20, 32. Median 16, range 20, mode 12.
5. The mean will rise; the median and mode may remain unchanged depending on the rest of the data.
6. Range measures spread, not reliability or data quality. A large range may be a genuine feature of the population.
7. Answers vary, but a complete investigation should identify the population, sample, variable, suitable statistics and a clear method of presenting findings.