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 of the remaining four numbers becomes 17. Find the removed number. Show your working.

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 = 135 ÷ 7 ≈ 19.29; median = 15. The median better represents the centre because 48 is an unusually high value (an outlier).

3. Mode.

4. Mode = 11; median = 11; mean = 94 ÷ 9 ≈ 10.44.

5. Example: identifying the most common shoe size sold in a shop.

6. No. Different data sets can share a mean while having different medians and ranges.

7. Example: 10, 10, 10, 10, 35. Mean = 15 and median = 10, so the mean is 5 greater.

Answer Key 2 — Outliers & Their Effect

1. Mean = (14 + 15 + 16 + 17 + 18) ÷ 5 = 16; range = 18 – 14 = 4.

2. New total = 80 + 60 = 140. New mean = 140 ÷ 6 ≈ 23.33; new range = 60 – 14 = 46.

3. The mean rises from 16 to about 23.33, and the range rises from 4 to 46 because 60 is far above the other values.

4. The mean is affected more strongly than the median because the high score of 70 pulls the mean upwards.

5. No. An outlier may have little or no effect on the median, especially when the data set is large.

6. Replace 16 with 30. The new ordered data are 10, 12, 14, 15, 30. The mean rises from 13.4 to 16.2, while the median remains 14.

7. Example: start with 2, 3, 4, 4, 4, 5. The median is 4. Add 100; the new data are 2, 3, 4, 4, 4, 5, 100. The median remains 4, but the mean increases.

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 = (5 + 8) ÷ 2 = 6.5; Q3 = (17 + 20) ÷ 2 = 18.5.

4. IQR = Q3 – Q1 = 18.5 – 6.5 = 12.

5. IQR measures the spread of the middle 50% of the data, while range uses only the minimum and maximum.

6. The data sets may have similar extremes, but the middle 50% of one set is more spread out.

7. One possible answer: 8, 10, 14, 16, 20, 26, 28, 32, 34. Median = 20; Q1 = (10 + 14) ÷ 2 = 12; Q3 = (28 + 32) ÷ 2 = 30.

Answer Key 4 — Comparing Distributions

1. Class A: total = 487; mean ≈ 69.57; median = 70; range = 14. Class B: total = 490; mean = 70; median = 70; range = 60.

2. Class B has greater spread. Its range is 60, compared with 14 for Class A. Its scores are also more widely distributed.

3. Class A appears more consistent because its scores are clustered more closely together.

4. Both athletes have the same mean, but Athlete A's scores vary less. Athlete A is therefore more consistent.

5. The mean does not show the spread, the presence of outliers or how the values are distributed.

6. Example: 10, 12, 15, 18, 20, 22, 24 and 2, 12, 15, 18, 20, 22, 50. Both have median 18, but their ranges are 14 and 48.

Answer Key 5 — Missing Data & Reverse Reasoning

1. Total of five numbers = 5 × 24 = 120. Known total = 17 + 21 + 25 + 29 = 92. Missing value = 120 – 92 = 28.

2. Total of six numbers = 6 × 18 = 108. Sixth value = 108 – 73 = 35.

3. Any value of x from 18 to 31, inclusive, keeps the values in order and the median at 18. For example, x = 20 or x = 25.

4. New range = 50 – 8 = 42.

5. Original total = 4 × 15 = 60. New total = 60 + 8 = 68. New mean = 68 ÷ 4 = 17.

6. Original total = 5 × 20 = 100. Remaining total = 4 × 17 = 68. Removed number = 100 – 68 = 32.

7. Answers vary. Any three different sets of five numbers with a total of 100 are valid.

Answer Key 6 — Statistics Detective Challenge

1. Income data may be skewed by a small number of very high earners. The median and a measure of spread would give more context about a typical household.

2. Both sets have the same mean of 30. Set A has a lower median (28) and a smaller range (12), so its values are less spread out. Set B has a higher median (34) and a much larger range (40), indicating greater spread.

3. Possible questions: What is the median? What are the range and IQR? Are there outliers? How many students were included?

4. Example: 12, 12, 16, 20, 32. Median = 16; range = 32 − 12 = 20; mode = 12.

5. The mean will increase. The median and mode may remain unchanged, depending on the other salaries and how many employees have each salary.

6. Range measures spread, not reliability or data quality. A large range may be a genuine feature of the population and does not automatically make the data unreliable.

7. Answers vary. A complete investigation should identify a clear question, population, sample and variable, select suitable statistics, and explain how the findings will be communicated.