Year 9 · Statistics: Reading the Story in Data

Name: _____ Date: _____

This sheet helps you get to know a data set by finding its five-point summary and interquartile range, and by spotting what kind of statistical question is being asked. Take your time ordering the numbers, that's where most marks are won or lost!

Warm-up

1. Order this data set from smallest to largest, then write down the minimum and maximum: 12, 7, 15, 9, 20, 4.

2. Find the median of: 3, 8, 5, 10, 6.

3. A survey records each student's eye colour and their height in centimetres. Which variable is categorical and which is numerical?

4. Mia counts the birds in her own backyard in Avondale. Wiremu uses bird numbers published on the StatsNZ website. Who is collecting primary data and who is using secondary data?

Practice

5. For the data set 4, 6, 7, 9, 11, 12, 15, 18, 20, find the five-point summary and the interquartile range.

6. These are the heights (cm) of 10 students: 152, 155, 158, 160, 162, 165, 168, 170, 172, 180. Find the five-point summary and the IQR.

7. Label each question as a summary, comparison, relationship, or time-series investigation. (a) What is the typical number of pets owned by students at Avondale Intermediate? (b) Do Year 9 boys spend more time on screens than Year 9 girls? (c) Is there a connection between hours studied and test score? (d) How has the price of a flat white in Avondale changed each year from 2020 to 2025?

8. Aroha plans an observational study at her school. For each student she records their favourite sport and how many hours per week they exercise. (a) Name one categorical and one numerical variable. (b) Is this primary or secondary data? (c) Write a good summary statistical question she could investigate.

9. The Avondale netball team scored these goals across 11 games: 18, 22, 15, 30, 25, 20, 28, 19, 24, 21, 26. Find the five-point summary and the IQR of the goals scored.

10. Eight students recorded how many minutes it took them to get to school: 5, 12, 8, 20, 15, 10, 25, 18. Find the five-point summary and the IQR, then state what the IQR tells you.

Challenge

11. Screen-time (hours per week) was recorded for two groups. Year 9: 2, 3, 3, 4, 5, 6, 7, 8. Year 10: 1, 3, 5, 6, 7, 8, 9, 11. Find the five-point summary and IQR for each group, then compare the two groups.

12. Waiting times (minutes) at a medical centre were: 5, 7, 8, 9, 10, 11, 12, 45. Find the five-point summary, the range, and the IQR. Explain which measure of spread better describes most patients' experience, and why.

Answer key (for parents and tutors)

1. Order this data set from smallest to largest, then write down the minimum and maximum: 12, 7, 15, 9, 20, 4.
   Ordered: 4, 7, 9, 12, 15, 20. Minimum = 4, Maximum = 20.

2. Find the median of: 3, 8, 5, 10, 6.
   Order first: 3, 5, 6, 8, 10. There are 5 values, so the median is the middle (3rd) value. Median = 6.

3. A survey records each student's eye colour and their height in centimetres. Which variable is categorical and which is numerical?
   Eye colour is categorical (it sorts people into named groups). Height in centimetres is numerical (it is a measured number).

4. Mia counts the birds in her own backyard in Avondale. Wiremu uses bird numbers published on the StatsNZ website. Who is collecting primary data and who is using secondary data?
   Mia is collecting primary data (she gathers it first-hand). Wiremu is using secondary data (it was collected by someone else).

5. For the data set 4, 6, 7, 9, 11, 12, 15, 18, 20, find the five-point summary and the interquartile range.
   The data is already ordered and has 9 values. Median (Q2) is the middle (5th) value = 11. Lower half (below the median): 4, 6, 7, 9 → Q1 = (6+7)÷2 = 6.5. Upper half (above the median): 12, 15, 18, 20 → Q3 = (15+18)÷2 = 16.5. Min = 4, Max = 20. Five-point summary: 4, 6.5, 11, 16.5, 20. IQR = Q3 − Q1 = 16.5 − 6.5 = 10.

6. These are the heights (cm) of 10 students: 152, 155, 158, 160, 162, 165, 168, 170, 172, 180. Find the five-point summary and the IQR.
   Already ordered, 10 values. Median (Q2) is between the 5th and 6th values = (162+165)÷2 = 163.5. Lower half (first 5): 152, 155, 158, 160, 162 → Q1 = middle value = 158. Upper half (last 5): 165, 168, 170, 172, 180 → Q3 = middle value = 170. Min = 152, Max = 180. Five-point summary: 152, 158, 163.5, 170, 180. IQR = 170 − 158 = 12 cm.

7. Label each question as a summary, comparison, relationship, or time-series investigation. (a) What is the typical number of pets owned by students at Avondale Intermediate? (b) Do Year 9 boys spend more time on screens than Year 9 girls? (c) Is there a connection between hours studied and test score? (d) How has the price of a flat white in Avondale changed each year from 2020 to 2025?
   (a) Summary, it describes one variable for one group. (b) Comparison, it compares one variable across two groups (boys and girls). (c) Relationship, it looks for a connection between two paired numerical variables. (d) Time-series, it looks at one variable (price) over time.

8. Aroha plans an observational study at her school. For each student she records their favourite sport and how many hours per week they exercise. (a) Name one categorical and one numerical variable. (b) Is this primary or secondary data? (c) Write a good summary statistical question she could investigate.
   (a) Categorical variable = favourite sport; numerical variable = hours of exercise per week. (b) Primary data, Aroha collects it herself. (c) A suitable summary question names the variable, the group, and the intent, for example: "What is the typical number of hours per week that students at Aroha's school exercise?"

9. The Avondale netball team scored these goals across 11 games: 18, 22, 15, 30, 25, 20, 28, 19, 24, 21, 26. Find the five-point summary and the IQR of the goals scored.
   Order the data: 15, 18, 19, 20, 21, 22, 24, 25, 26, 28, 30 (11 values). Median (Q2) is the 6th value = 22. Lower half (first 5): 15, 18, 19, 20, 21 → Q1 = middle value = 19. Upper half (last 5): 24, 25, 26, 28, 30 → Q3 = middle value = 26. Min = 15, Max = 30. Five-point summary: 15, 19, 22, 26, 30. IQR = 26 − 19 = 7 goals.

10. Eight students recorded how many minutes it took them to get to school: 5, 12, 8, 20, 15, 10, 25, 18. Find the five-point summary and the IQR, then state what the IQR tells you.
    Order the data: 5, 8, 10, 12, 15, 18, 20, 25 (8 values). Median (Q2) = (12+15)÷2 = 13.5. Lower half (first 4): 5, 8, 10, 12 → Q1 = (8+10)÷2 = 9. Upper half (last 4): 15, 18, 20, 25 → Q3 = (18+20)÷2 = 19. Min = 5, Max = 25. Five-point summary: 5, 9, 13.5, 19, 25. IQR = 19 − 9 = 10 minutes. The IQR tells you the middle half of the travel times spread across a range of 10 minutes.

11. Screen-time (hours per week) was recorded for two groups. Year 9: 2, 3, 3, 4, 5, 6, 7, 8. Year 10: 1, 3, 5, 6, 7, 8, 9, 11. Find the five-point summary and IQR for each group, then compare the two groups.

Year 9 (8 values, ordered): median = (4+5)÷2 = 4.5. Lower half 2,3,3,4 → Q1 = (3+3)÷2 = 3. Upper half 5,6,7,8 → Q3 = (6+7)÷2 = 6.5. Summary: 2, 3, 4.5, 6.5, 8. IQR = 6.5 − 3 = 3.5. Year 10 (8 values, ordered): median = (6+7)÷2 = 6.5. Lower half 1,3,5,6 → Q1 = (3+5)÷2 = 4. Upper half 7,8,9,11 → Q3 = (8+9)÷2 = 8.5. Summary: 1, 4, 6.5, 8.5, 11. IQR = 8.5 − 4 = 4.5. Comparison: Year 10 students have a higher median screen time (6.5 vs 4.5 hours) and a larger IQR (4.5 vs 3.5 hours), so Year 10 students generally spend more time on screens and their middle half is more spread out.

12. Waiting times (minutes) at a medical centre were: 5, 7, 8, 9, 10, 11, 12, 45. Find the five-point summary, the range, and the IQR. Explain which measure of spread better describes most patients' experience, and why.
Order (already ordered, 8 values): 5, 7, 8, 9, 10, 11, 12, 45. Median = (9+10)÷2 = 9.5. Lower half 5,7,8,9 → Q1 = (7+8)÷2 = 7.5. Upper half 10,11,12,45 → Q3 = (11+12)÷2 = 11.5. Min = 5, Max = 45. Five-point summary: 5, 7.5, 9.5, 11.5, 45. Range = 45 − 5 = 40 minutes. IQR = 11.5 − 7.5 = 4 minutes. The IQR (4 minutes) better describes most patients' experience, because it uses only the middle half of the data and is not stretched by the single unusually long wait of 45 minutes (an outlier). The range of 40 minutes is dragged up by that one extreme value and makes the waits look far more variable than they really are for most people.