Year 9 · Statistics: Picture the Data

Name: _____ Date: _____

Hi! This sheet is all about turning data into pictures that tell a story. You'll build different graphs, pick sensible scales, and even draw a line of best fit to make a prediction. Grab a ruler and a pencil!

Warm-up

1. Name the best type of graph to show how a class's favourite lunch (pie, sushi, sandwich, wrap) is spread across the group.

2. A dot plot shows the number of siblings for 10 students: 0, 1, 1, 1, 2, 2, 3, 0, 1, 4. How many students have exactly 1 sibling?

3. For a scatter plot investigating whether hours of study predicts test score, which variable goes on the x-axis?

4. A frequency table records: Walk 8, Bus 6, Car 5, Bike 1. What is the total number of students surveyed?

Practice

5. The heights (cm) of 12 Year 9 students are: 150, 155, 155, 160, 162, 162, 162, 165, 168, 170, 172, 180. What is a sensible scale for the vertical axis of a histogram of these heights, and why?

6. Complete this frequency table for the number of pets owned by 15 whānau: 0,1,2,1,3,0,2,1,0,4,1,2,0,1,2. Fill in the frequency for each value from 0 to 4.

7. Word problem: Aroha records the temperature at Avondale each day at noon over one week: Mon 14°C, Tue 16°C, Wed 15°C, Thu 18°C, Fri 20°C, Sat 19°C, Sun 17°C. Which graph best shows how temperature changed over the week, and what does the graph show?

8. Word problem: A dairy on Rosebank Road sells hot pies. The owner tracks daily pies sold against the daily high temperature: (12°C, 40 pies), (14°C, 36), (16°C, 30), (18°C, 26), (20°C, 20), (22°C, 15). Describe the relationship you would see on a scatter plot and which variable should be on each axis.

9. Using the dairy data above, draw an eyeballed line of best fit and use it to predict how many pies would be sold on a 17°C day.

10. A survey asked 40 students two questions: Do you play a sport? (Yes/No) and Do you play a musical instrument? (Yes/No). Results: sport & instrument 9, sport only 15, instrument only 6, neither 10. Set up a two-way table and find how many students play a sport in total.

Challenge

11. Ten Year 9 students recorded their weekly screen time (hours) and their hours of sleep per night: (5,9), (10,8.5), (15,8), (20,7.5), (25,7), (30,6.5), (12,8.2), (22,7.3), (8,8.7), (28,6.8). (a) Which variable is explanatory? (b) Describe the relationship. (c) Draw a line of best fit and predict the sleep for a student with 18 hours of screen time.

12. A journalist shows a bar graph where the vertical axis starts at 40 (not 0) to compare test averages of two classes: Class A = 62, Class B = 68. Explain why this scale can mislead readers, and describe a better visualisation choice for comparing the full spread of each class's scores.

Answer key (for parents and tutors)

1. Name the best type of graph to show how a class's favourite lunch (pie, sushi, sandwich, wrap) is spread across the group.
   A bar graph. The variable 'favourite lunch' is categorical, so a bar graph shows the frequency of each category with bars of equal width and gaps between them. (A dot plot could also be used for a small group.)

2. A dot plot shows the number of siblings for 10 students: 0, 1, 1, 1, 2, 2, 3, 0, 1, 4. How many students have exactly 1 sibling?
   Count the 1s in the list: positions 2, 3, 4 and 9 → that is four 1s. So 4 students have exactly 1 sibling.

3. For a scatter plot investigating whether hours of study predicts test score, which variable goes on the x-axis?
   Hours of study goes on the x-axis. It is the explanatory (independent, 'predictive') variable. Test score is the response variable and goes on the y-axis.

4. A frequency table records: Walk 8, Bus 6, Car 5, Bike 1. What is the total number of students surveyed?
   Add the frequencies: 8 + 6 + 5 + 1 = 20 students.

5. The heights (cm) of 12 Year 9 students are: 150, 155, 155, 160, 162, 162, 162, 165, 168, 170, 172, 180. What is a sensible scale for the vertical axis of a histogram of these heights, and why?
   The values range from 150 cm to 180 cm, a range of 30 cm. A sensible scale is to start the axis at 150 (or a little below) and use class intervals of 5 cm: 150–154, 155–159, 160–164, 165–169, 170–174, 175–179. Marking the axis every 5 cm keeps the graph readable without too many empty gaps. Frequencies would be 1, 2, 4, 2, 2, 1 (the 180 falls in a 180–184 bin, frequency 1).

6. Complete this frequency table for the number of pets owned by 15 whānau: 0,1,2,1,3,0,2,1,0,4,1,2,0,1,2. Fill in the frequency for each value from 0 to 4.
   Count each value. 0: appears at positions 1,6,9,13 → 4. 1: positions 2,4,8,11,14 → 5. 2: positions 3,7,10,12,15 → 5. 3: position 5 → 1. 4: position 10… check: 4 appears once → 0 pets:4, 1 pet:5, 2 pets:5, 3 pets:1, 4 pets:1. Total = 4+5+5+1+1 = 16? Recount: values are 0,1,2,1,3,0,2,1,0,4,1,2,0,1,2 (15 values). 0s: positions 1,6,9,13 = 4. 1s: positions 2,4,8,11,14 = 5. 2s: positions 3,7,12,15 = 4. 3s: position 5 = 1. 4s: position 10 = 1. Total = 4+5+4+1+1 = 15. So: 0→4, 1→5, 2→4, 3→1, 4→1.

7. Word problem: Aroha records the temperature at Avondale each day at noon over one week: Mon 14°C, Tue 16°C, Wed 15°C, Thu 18°C, Fri 20°C, Sat 19°C, Sun 17°C. Which graph best shows how temperature changed over the week, and what does the graph show?
   A time-series graph is best because the data is ordered over time (days). Plot day on the x-axis and temperature (°C) on the y-axis, then join the points with straight lines. The graph shows temperature generally rising from Monday to Friday (14°C up to 20°C), then dipping slightly over the weekend (19°C, 17°C). The overall trend for the week is an increase in noon temperature to a Friday peak.

8. Word problem: A dairy on Rosebank Road sells hot pies. The owner tracks daily pies sold against the daily high temperature: (12°C, 40 pies), (14°C, 36), (16°C, 30), (18°C, 26), (20°C, 20), (22°C, 15). Describe the relationship you would see on a scatter plot and which variable should be on each axis.
   Temperature is the explanatory variable (x-axis) because it is thought to predict pie sales; pies sold is the response variable (y-axis). On the scatter plot the points fall from upper-left to lower-right, showing a negative relationship: as the temperature rises, the number of pies sold decreases. The points lie close to a straight line, so it is a strong, roughly linear negative relationship.

9. Using the dairy data above, draw an eyeballed line of best fit and use it to predict how many pies would be sold on a 17°C day.
   A straight line drawn through the middle of the points goes from about (12°C, 39) to (22°C, 16). Its slope is about (16 − 39) ÷ (22 − 12) = −23 ÷ 10 ≈ −2.3 pies per °C. At 17°C, moving 5°C up from 12°C: 39 + 5 × (−2.3) = 39 − 11.5 ≈ 27.5. So the line predicts about 27 or 28 pies on a 17°C day. (Any answer between roughly 26 and 29 from a reasonable line is acceptable.)

10. A survey asked 40 students two questions: Do you play a sport? (Yes/No) and Do you play a musical instrument? (Yes/No). Results: sport & instrument 9, sport only 15, instrument only 6, neither 10. Set up a two-way table and find

how many students play a sport in total.
Two-way table (rows = sport, columns = instrument): Instrument Yes | Instrument No | Total Sport Yes 9 | 15 | 24 Sport No 6 | 10 | 16 Total 15 | 25 | 40 Students who play a sport = sport Yes total = 9 + 15 = 24 students. (Check: all cells add to 40 ✓.)

11. Ten Year 9 students recorded their weekly screen time (hours) and their hours of sleep per night: (5,9), (10,8.5), (15,8), (20,7.5), (25,7), (30,6.5), (12,8.2), (22,7.3), (8,8.7), (28,6.8). (a) Which variable is explanatory? (b) Describe the relationship. (c) Draw a line of best fit and predict the sleep for a student with 18 hours of screen time.
(a) Screen time is the explanatory variable (x-axis); sleep is the response variable (y-axis), since more screen time is thought to affect sleep. (b) The points fall steadily from upper-left to lower-right: as screen time increases, sleep decreases, a strong negative, roughly linear relationship. (c) A line of best fit runs from about (5, 9) to (30, 6.5). Slope = (6.5 − 9) ÷ (30 − 5) = −2.5 ÷ 25 = −0.1 hours of sleep per hour of screen time. At 18 hours: starting from (5,9), go 13 hours right: 9 + 13 × (−0.1) = 9 − 1.3 = 7.7 hours of sleep. So about 7.6–7.8 hours is a reasonable prediction.

12. A journalist shows a bar graph where the vertical axis starts at 40 (not 0) to compare test averages of two classes: Class A = 62, Class B = 68. Explain why this scale can mislead readers, and describe a better visualisation choice for comparing the full spread of each class's scores.
Because the axis starts at 40 instead of 0, the bar for Class B (68) looks more than twice as tall as Class A (62), even though the real difference is only 6 marks out of 62 (about 10%). Cutting the axis exaggerates small differences and can mislead viewers into thinking one class did far better. A fairer choice is to start the frequency/score axis at 0 so bar heights are in true proportion. Better still, to compare the full spread of every student's score, use box plots (or dot plots) for each class side by side. Box plots show the median, range, and quartiles, so readers can compare both the middle and the variation of the two classes, not just a single average.