Test answers - Sampling and scatter graphs

Year: Year 11 · Duration: 45 minutes · /20

Answer to Question 1 (4 marks)

a. Fraction $= \frac{50}{1000} = \frac{1}{20}$: **15 men, 25 women, 10 children** (1 each) Add the groups to get 1000 people. The sample is $\frac{50}{1000} = \frac{1}{20}$ of the town, so divide each group by 20. Check: $15 + 25 + 10 = 50$.

b. So that each group is represented in proportion to its size (1) A stratified sample keeps the same proportions of men, women and children as the whole population, so no group is over- or under-represented.

Answer to Question 2 (4 marks)

a. $\frac{30}{N} = \frac{7}{35}$ (1), $N = \frac{30 \times 35}{7}$ (1) $=$ **150** (1) The fraction tagged in the second sample, $\frac{7}{35} = \frac{1}{5}$, estimates the fraction tagged in the whole lake. So 30 is about $\frac{1}{5}$ of the population.

b. The population does not change between the two samples (1) We also assume that tagged fish mix evenly with the others and that the tags do not fall off.

Answer to Question 3 (4 marks)

a. Gradient $= \frac{40 - 12}{7 - 1} = \frac{28}{6} = 4.7$ (1); $y - 12 = 4.7(x - 1)$ (1); $y = 4.7x + 7.3$ (1) The gradient is the change in y divided by the change in x. Substitute the point $(1, 12)$ to find the intercept: $12 - 4.7 = 7.3$.

b. $4.7 \times 4 + 7.3 =$ **26.1** (1) Substitute $x = 4$ into the equation. The value is inside the range of the data, so the estimate is reasonable.

With the exact gradient $\frac{14}{3}$, the estimate is $12 + 3 \times \frac{14}{3} = 26$, so 26.1 is consistent.

Answer to Question 4 (4 marks)

a. Negative correlation (1): as the age increases the value decreases (1)

b. There is correlation, and for cars age really does influence value (wear, technology) (1); but correlation alone does not prove a cause, so other factors such as mileage may also matter (1)

In (a), describe both the type of correlation and its meaning in context to earn both marks. In (b), the key point is that a scatter graph only shows a relationship; it cannot prove that one variable causes the other.

Answer to Question 5 (4 marks)

a. People who work on Saturdays, or who live far away, are missed (1); the people in the centre may be shoppers who are not park users (1)
b. Choose people at random from a list of local residents (1) and contact them at different times (1)

In (a), other valid reasons include asking at only one time and one place, so the sample is not representative. In (b), a larger sample also improves reliability.