Digital SAT Math

Data Analysis — Practice Set II

Two-Variable Data and Probability: Medium to Hard

6 Questions | Full Solutions | Mistake-Pattern Analysis

Question 1 [MEDIUM]

A line of best fit is given by y = 2.5x + 10, where x is hours studied and y is a test score. What is the predicted score for 4 hours of studying?

(Student-produced response)

Algebraic: y = 2.5(4) + 10 = 10 + 10 = 20.

Desmos: Graph y = 2.5x + 10 and trace to x = 4 to read the y-value.

Faster path: Algebraic, direct substitution is faster than tracing a graph for one point.

Mistake: Concept gap: student multiplies 2.5 by 4 and forgets to add the 10, reporting 10 instead of 20.

Question 2 [MEDIUM]

Using the line of best fit y = 1.8x + 32, what is the predicted value of y when x = 10?

(Student-produced response)

Algebraic: y = 1.8(10) + 32 = 18 + 32 = 50.

Desmos: Graph the line and trace to x = 10 to read the y-value.

Faster path: Algebraic, direct substitution is faster than tracing a graph for one point.

Mistake: Concept gap: student multiplies 1.8 by 10 correctly but then multiplies the result by 32 instead of adding it, applying the wrong operation.

Question 3 [HARD]

A line of best fit is y = 0.5x + 3. A data point has an actual y-value of 10 when x = 8. What is the residual for this point (actual minus predicted)?

(Student-produced response)

Algebraic: Predicted value: y = 0.5(8) + 3 = 4 + 3 = 7. Residual = actual − predicted = 10 − 7 = 3.

Desmos: Graph the line and plot the point (8, 10); the vertical distance between the point and the line is the residual.

Faster path: Algebraic, computing the predicted value and subtracting is faster than reading a vertical distance precisely off a graph.

Mistake: Concept gap: student computes predicted minus actual instead of actual minus predicted, getting the residual's sign backward.

Question 4 [MEDIUM]

A survey of 200 students is summarized in a two-way table: of the 100 ninth-graders, 40 like math and 60 do not; of the 100 tenth-graders, 70 like math and 30 do not. What is the probability that a randomly selected student from this survey is in ninth grade?

(Student-produced response)

Algebraic: There are 100 ninth-graders out of 200 total students, so the probability is 100/200 = 0.5, or 50%.

Desmos: Not useful, this is a direct ratio read from a table.

Faster path: Algebraic, always, for a single-cell-to-total table read.

Mistake: Concept gap: student divides the total by the ninth-grade count instead of the other way around, inverting the probability.

Question 5 [MEDIUM]

Using the same survey (100 ninth-graders: 40 like math, 60 do not; 100 tenth-graders: 70 like math, 30 do not), what is the probability that a randomly selected student likes math?

(Student-produced response)

Algebraic: The total who like math is 40 + 70 = 110, out of 200 total students: 110/200 = 0.55, or 55%.

Desmos: Not useful, this is a direct table read requiring only one addition.

Faster path: Algebraic, always.

Mistake: Concept gap: student uses only one grade's count (for example, just the 70 tenth-graders) instead of combining both grades before dividing by the total.

Question 6 [HARD]

Using the same survey, given that a student is in ninth grade, what is the probability that the student likes math?

(Student-produced response)

Algebraic: This is a conditional probability restricted to the ninth-grade row: 40 out of 100 ninth-graders like math, so the probability is 40/100 = 0.40, or 40%.

Desmos: Not useful, this is a direct table read within a single row.

Faster path: Algebraic, always. Reading within the given row directly avoids any need for the full joint-probability formula.

Mistake: Concept gap: student divides by the full survey total (200) instead of the ninth-grade row total (100), since the condition "given that a student is in ninth grade" restricts the denominator.