# Linear Model for Scatter Plots

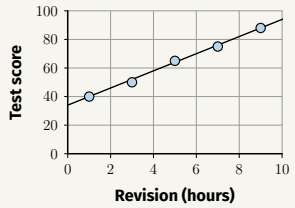
Form the equation of a line of best fit, then use it to predict.

**Name** _____ **Date** _____

**Show your method**

The line of best fit has equation $y = mx + c$: $m$ is the gradient, $c$ is the $y$-intercept. Find the gradient with $m = \dfrac{y_2 - y_1}{x_2 - x_1}$, substitute a point to find $c$, then substitute a known $x$ to predict $y$.

**Worked example** This scatter plot shows revision hours ($x$) against test score ($y$), with a line of best fit through $(1, 40)$ and $(9, 88)$. Find the equation of the line, then predict the score for 6 hours of revision.




Gradient: $\dfrac{88 - 40}{9 - 1} = \dfrac{48}{8} = 6$
$y$-intercept: $40 = 6(1) + c \Rightarrow c = 34$
Equation: $y = 6x + 34$
Predict: $y = 6(6) + 34 = 70$

**1.** A line of best fit has equation $y = 2x + 10$. Predict $y$ when $x = 5$.

**Solution:** $y = 2(5) + 10 = 20$

**2.** A line of best fit has equation $y = 3x + 5$. Predict $y$ when $x = 10$.

**Solution:** $y = 3(10) + 5 = 35$

**3.** A line of best fit has equation $y = 1.5x + 20$. Predict $y$ when $x = 8$.

**Solution:** $y = 1.5(8) + 20 = 32$

**4.** A line of best fit has equation $y = 4x - 2$. Predict $y$ when $x = 6$.

**Solution:** $y = 4(6) - 2 = 22$

**5.** A line of best fit has gradient $2$ and $y$-intercept $15$. Write its equation.

**Solution:** $y = 2x + 15$

**6.** A line has equation $y = 3x + 8$. State its gradient and $y$-intercept.

**Solution:** Gradient $3$, $y$-intercept $8$.

**7.** A line has equation $y = 0.5x + 40$. Predict $y$ when $x = 20$.

**Solution:** $y = 0.5(20) + 40 = 50$

**8.** A line has equation $y = 5x$. Predict $y$ when $x = 12$.

**Solution:** $y = 5(12) = 60$

**9.** A line of best fit passes through $(0, 20)$ and $(10, 70)$. Find its gradient.

**Solution:** $\dfrac{70 - 20}{10 - 0} = 5$

# Linear Model for Scatter Plots (continued)

Apply the method to word contexts and reliability.

**Name** _____ **Date** _____

**10.** Using the line from Question 9 (through $(0, 20)$), write its equation.

**Solution:** $y = 5x + 20$

**11.** Use the equation from Question 10 to predict $y$ at $x = 6$.

**Solution:** $y = 5(6) + 20 = 50$

**12.** Hours studied ($x$) vs test score ($y$) has line $y = 6x + 40$, based on data for $x = 0$ to $10$. Predict the score for 7 hours.

**Solution:** $y = 6(7) + 40 = 82$ (interpolation)

**13.** Using the same line, predict the score for $x = 25$. Explain why this prediction is less reliable.

**Solution:** $y = 6(25) + 40 = 190$. Extrapolation – $x = 25$ is outside the data range, so the trend may not continue.

**14.** A line of best fit passes through $(2, 14)$ and $(6, 26)$. Find its gradient, then its equation given the $y$-intercept is $8$.

**Solution:** $m = \frac{26-14}{6-2} = 3$. $y = 3x + 8$

**15.** Use the equation from Question 14 to predict $y$ when $x = 10$.

**Solution:** $y = 3(10) + 8 = 38$

**16.** A weak correlation means points scatter widely around the line. Explain how this affects confidence in a prediction.

**Solution:** More scatter means less confidence.

**17.** Explain the difference between interpolation and extrapolation.

**Solution:** Interpolation: within the data range. Extrapolation: beyond it, less reliable.

**18.** A line of best fit is $y = -2x + 300$ for $x = 0$ to $50$. Find $x$ when $y = 200$.

**Solution:** $200 = -2x + 300 \Rightarrow x = 50$