Vertex Academy

NCEA Level 3 Maths workbook

12 worksheets · answers at the back

Contents

1. Complex arithmetic, conjugates and equating parts
2. The derivative toolkit: chain, product and quotient rules
3. The integration toolkit: surds, e, trig and the chain rule
4. Margin of error: the three rules of thumb
5. Tables, trees and Venn diagrams from written information
6. Choosing a distribution, and naming it every time
7. Trig functions and radians
8. Trig identities and proofs
9. Features of a time series
10. Scatter plots and correlation
11. Sampling variability and bootstrapping
12. Solving three equations

NCEA Level 3 · AS 91577: Complex Numbers & Their Algebra

Name: _____ Date: _____

Hi! This standard is all about working confidently with complex numbers in both rectangular (a + bi) and polar (r cis θ) form, using De Moivre's theorem, and interpreting loci. Work through the sheet in order, the Key skills warm you up, then push into Merit and Excellence where you'll link ideas and justify your reasoning.

Revision summary

- Rectangular form: $z = a + bi$ where $i^2 = -1$. Conjugate $\bar{z} = a - bi$, and $z \cdot \bar{z} = a^2 + b^2 = |z|^2$.
- Polar form: $z = r\,\text{cis}\,\theta = r(\cos\theta + i\sin\theta)$, where $r = |z| = \sqrt{(a^2 + b^2)}$ and $\theta = \arg(z)$ (choose the quadrant carefully; $-\pi < \theta \leq \pi$).
- Multiply in polar: multiply moduli, add arguments. Divide: divide moduli, subtract arguments. To divide in rectangular form, multiply top and bottom by the conjugate of the denominator.
- De Moivre's theorem: $(r\,\text{cis}\,\theta)^n = r^n\,\text{cis}(n\theta)$ for all rational n.
- nth roots of a complex number: take the nth root of the modulus and use angles $(\theta + 2\pi k)/n$ for $k = 0, 1, \ldots, n-1 \rightarrow n$ roots equally spaced around a circle.
- Loci: $|z - a|$ = distance from the point a; $|z - a| = |z - b|$ is the perpendicular bisector of a and b; $|z - a| = r$ is a circle centre a radius r.
- Surds/indices: $\sqrt{(-n)} = i\sqrt{n}$; rationalise denominators by multiplying by the conjugate; a^(m/n) = ($\sqrt[n]{a}$)^m.

Key skills

1. Simplify $\sqrt{(-48)}$.

2. Write $z = 2 + 2i$ in polar form.

3. Expand and simplify $(3 + 2i)(1 - 4i)$.

4. Use De Moivre's theorem to evaluate $(2\,\text{cis}(\pi/6))^3$, giving your answer in rectangular form.

Achieved

5. Express $(4 + 3i)/(2 - i)$ in the form $a + bi$.

6. Write $z = -\sqrt{3} + i$ in polar form with $-\pi < \theta \leq \pi$.

7. Solve $z^2 = -9$ for z.

8. Rationalise the denominator and simplify: $2/(\sqrt{7} - \sqrt{3})$.

Merit

9. Find all three cube roots of $8i$, giving your answers in rectangular form.

10. Given $z = 1 + i$, use De Moivre's theorem to find z^6 in rectangular form.

11. Describe and identify the locus of points z satisfying $|z - 2| = |z + 2i|$.

Excellence

12. Prove, using mathematical induction, that $(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$ for all positive integers n.

13. Show that the sum of all n distinct nth roots of unity is 0 for $n \geq 2$.

14. A complex number z satisfies $z + 1/z = 2\cos\theta$. Prove that $z^n + 1/z^n = 2\cos n\theta$ for every positive integer n.

NCEA Level 3 · AS 91578: Differentiation Toolkit

Name: _____ Date: _____

Hi! This standard is all about differentiating trickier functions (using product, quotient and chain rules, plus trig, e^x and ln) and then using those derivatives to solve real problems, rates, optimisation, tangents, curve shapes and motion. Work through the sheet top to bottom: warm up on the key skills, then build toward the Merit and Excellence reasoning.

Revision summary

- Rules: product $(uv)' = u'v + uv'$; quotient $(u/v)' = (u'v - uv')/v^2$; chain $dy/dx = dy/du \cdot du/dx$.
- Standard derivatives: $d/dx(\sin x)=\cos x$, $d/dx(\cos x)=-\sin x$, $d/dx(\tan x)=\sec^2 x$, $d/dx(e^x)=e^x$, $d/dx(\ln x)=1/x$.
- Chain versions: $d/dx(e^{(f)})=f' \cdot e^{(f)}$; $d/dx(\ln f)=f'/f$; $d/dx(\sin f)=f' \cdot \cos f$.
- Stationary points: solve $f'(x)=0$. Classify with sign of $f'(x)$ or with $f''(x)$: $f''>0$ minimum, $f''<0$ maximum, test if $f''=0$.
- Optimisation: write the quantity in one variable using a constraint, differentiate, set derivative = 0, then justify it is a max/min.
- Related rates: differentiate with respect to t; chain the rates together, e.g. $dV/dt = dV/dr \cdot dr/dt$.
- Kinematics: velocity $v = ds/dt$, acceleration $a = dv/dt = d^2s/dt^2$. Object at rest when $v = 0$.
- Tangent at $x = a$: gradient $m = f'(a)$, line $y - f(a) = m(x - a)$.

Key skills

1. Differentiate $y = (3x^2 + 1)^5$.

2. Differentiate $y = x^2 e^x$.

3. Differentiate $y = (2x + 1)/(x - 3)$.

4. Differentiate $y = \ln(\sin x)$.

Achieved

5. Find the gradient of the curve $y = 3\sin(2x)$ at $x = \pi/6$.

6. Find the coordinates of the stationary points of $y = x^3 - 6x^2 + 9x + 2$.

7. A trolley moves so its position is $s = t^3 - 4t^2 + 2t$ metres after t seconds. Find its velocity at $t = 3$ s.

8. Find the equation of the tangent to $y = e^{(2x)}$ at the point where $x = 0$.

Merit

9. A farmer near Pukekohe uses 120 m of fencing to make a rectangular paddock against a straight river. No fence is needed along the river. Find the dimensions that give the maximum area, and state that maximum area.

10. A closed cylindrical can must hold 500 cm³. Find the radius that minimises the surface area (give it to 2 d.p.).

11. A spherical weather balloon is inflated so its radius increases at 2 cm/s. Find the rate at which its volume is increasing when the radius is 5 cm.

Excellence

12. For the curve $y = x^3 - 3kx + 2$, determine all values of the constant k for which the curve has two distinct stationary points. Justify your answer.

13. Using calculus, prove that among all rectangles with a fixed perimeter P, the square gives the greatest area.

14. A particle moves along a line with displacement $x(t) = e^{(-t)} \sin t$ metres for $t \geq 0$. Show that the particle is instantaneously at rest whenever $\tan t = 1$, and explain why the exponential factor never causes it to stop.

NCEA Level 3 · AS 91579: Integration Methods

Name: _____ Date: _____

Hi! This standard is all about anti-derivatives, finding areas, using substitution, and solving differential equations. Work through each section in order, the questions get harder as you go, so use the early ones to warm up your fluency before tackling the Merit and Excellence problems.

Revision summary

- Power rule: $\int x^n\,dx = x^{n+1}/(n+1) + c$, for $n \neq -1$. Always add + c for indefinite integrals.
- Standard forms: $\int e^{kx}\,dx = (1/k)e^{kx} + c$; $\int (1/x)\,dx = \ln|x| + c$; $\int \sin(kx)\,dx = -(1/k)\cos(kx) + c$; $\int \cos(kx)\,dx = (1/k)\sin(kx) + c$.
- Definite integral (area): $\int$ from a to b of f(x) dx = F(b) − F(a). Area below the x-axis counts as negative, so split the integral at the roots for total area.
- Substitution: let u = inner function, then du = (du/dx) dx. Replace everything in x with u (including limits for definite integrals) before integrating.
- Differential equations: dy/dx = f(x) is solved by integrating, y = ∫f(x) dx + c. Use a given point (initial condition) to find c.
- Rate problems: if dV/dt is given, then V = ∫(dV/dt) dt; the constant represents the starting amount.
- Chain-rule check: to integrate f(g(x))·g'(x), the derivative of the inner function must appear (up to a constant) as a factor.
- Area between curves: ∫ from a to b of (top curve − bottom curve) dx, where a and b are the intersection points.

Key skills

1. Find $\int (6x^2 - 4x + 5)\,dx$.

2. Find $\int (e^{3x} + 2\cos x)\,dx$.

3. Evaluate $\int$ from 1 to 3 of $(2x)\,dx$.

4. Find $\int (4/x)\,dx$ for $x > 0$.

Achieved

5. The gradient function of a curve is $dy/dx = 3x^2 - 8x$. The curve passes through the point (2, 1). Find the equation of the curve.

6. Find the area between the curve $y = x^2 + 1$ and the x-axis from $x = 0$ to $x = 2$.

7. Use the substitution $u = x^2 + 3$ to find $\int 2x(x^2 + 3)^4\,dx$.

8. Water flows into a tank at a rate of $dV/dt = 6t + 2$ litres per minute, where t is in minutes. If the tank starts empty, find the volume of water after 5 minutes.

Merit

9. Find the area enclosed between the curves $y = x^2$ and $y = 6x - x^2$.

10. Find $\int x\sqrt{(2x^2 + 1)}\,dx$ using a suitable substitution, showing your choice.

11. The acceleration of a drone rising vertically is $a = 12 - 6t$ m s^{-2} at time t seconds. At $t = 0$ its velocity is 5 m s^{-1} and its height is 0 m. Find the height when the drone's velocity first returns to 5 m s^{-1}.

Excellence

12. A curve has gradient function $dy/dx = kx + 1$, where k is a constant. The curve passes through (0, 3) and (2, 9). Find the value of k and the equation of the curve, justifying your working.

13. Show that $\int$ from 0 to a of $(2x)\,dx = \int$ from 0 to a of $(3x^2)\,dx$ has exactly one positive solution for a, and find it.

14. A population grows so that $dP/dt = 0.4P$, where P is the number of insects and t is in days. Given that $P = 200$ when $t = 0$, show that $P = 200e^{(0.4t)}$, and find how long it takes the population to reach 1000. (Hint: rewrite as $(1/P)\,dP = 0.4\,dt$ and integrate both sides.)

NCEA Level 3 · AS 91584: Evaluating Statistical Reports

Name: _____ Date: _____

Reports turn a sample into a claim about a population. This sheet practises the 1/√n rules of thumb, then asks you to judge how far a claim can be trusted. All contexts are invented. Write every interval in full and interpret it in context.

Revision summary

- One proportion: margin of error ≈ 1/√n.
- Difference within one group or poll: margin of error ≈ 2 × 1/√n.
- Difference between two independent groups: margin of error ≈ 1.5 × the average of the two margins.
- The rules suit random samples with proportions from about 30% to 70%.
- A margin of error covers sampling variation only. It does not cover bias, non-response or question wording.
- Compare groups with one interval for the difference, not by checking whether two intervals overlap.
- An interval all above or all below zero supports a claim of a difference. Zero inside it means no claim either way.
- Interpret first: name the population, the variable and the parameter. Make any claim separately.
- A claim about cause needs a randomised experiment.

Key skills

1. A random poll surveys 1600 voters. Find the margin of error.

2. A random poll of 400 people finds 46% support a policy. Form the interval.

3. A cycling club emails its members to ask whether the town wants a new cycleway. Name the main problem with this sample.

4. Name two things a margin of error does not cover.

Achieved

5. A random sample of 900 residents finds 54% support a new library. Form the interval.

6. Interpret the interval from question 5.

7. In the same poll, 54% prefer site A and 38% prefer site B. Find an interval for the difference.

8. A headline says, "Coffee drinkers live longer", based on a survey of adults' habits. Why can it not show that coffee causes longer life?

Merit

9. Town A: n = 625, 52% support a plan. Town B: n = 400, 44% support it. Find an interval for the difference and make a claim.

10. A student says the Town A and Town B intervals overlap, so the towns cannot differ. Explain the error.

11. In a study, 200 students are randomly allocated to use a revision app or to revise as usual. Why does random allocation matter, and why is the second group needed?

Excellence

12. A random sample of 1000 finds 4% of people own an electric bike. Can the $1/\sqrt{n}$ rule be used? Explain.

13. A website poll gets 10 000 votes, with 62% in favour, and reports a margin of error of 1%. Critique this.

14. A random poll of 2500 finds Party A on 48% and Party B on 44%. The report says Party A leads. Evaluate the claim.

NCEA Level 3 · AS 91585: Probability Concepts in Action

Name: _____ Date: _____

Hey! This standard is all about digging into probability, conditional probability, independence, risk, and reading two-way tables and trees. Work through each section in order, check your working against the answers, and pay attention to the 'reasoning' style questions in Merit and Excellence, because that's where the top grades live.

Revision summary

- Conditional probability: $P(A|B) = P(A \text{ and } B) \div P(B)$. It answers 'given B has happened, what's the chance of A?'
- Independence: A and B are independent if $P(A|B) = P(A)$, or equivalently $P(A \text{ and } B) = P(A) \times P(B)$. If these don't hold, the events are dependent.
- Two-way tables: read counts off the table, use row/column totals for conditional probabilities, and the grand total for unconditional ones.
- Tree diagrams: multiply probabilities ALONG branches; add across separate branches that give the same outcome. Second-layer branches are conditional probabilities.
- Relative risk = (risk for exposed group) ÷ (risk for baseline/unexposed group). Risk = proportion with the outcome. RR > 1 means increased risk, RR = 1 means no difference, RR < 1 means reduced risk.
- Absolute risk is the actual proportion (e.g. 0.02); relative risk is a ratio comparing two groups, always state which group is the baseline.
- Probability distribution: probabilities sum to 1. Expected value $E(X) = \Sigma(x \times P(X = x))$.
- Always define events clearly and check whether a probability is conditional (restricted to a subgroup) or unconditional (out of the whole population).

Key skills

1. A survey of 200 Avondale residents recorded whether they own a bike and whether they cycle to work. 60 own a bike; of those, 45 cycle to work. Find the probability that a randomly chosen bike owner cycles to work.

2. For two events, $P(A) = 0.4$, $P(B) = 0.5$, and $P(A \text{ and } B) = 0.2$. Are A and B independent?

3. In a trial, 8% of people taking a new supplement caught a cold, compared with 20% of people not taking it. Calculate the relative risk of catching a cold for supplement-takers compared with non-takers.

4. A tree diagram shows $P(\text{rain}) = 0.3$, and $P(\text{bus is late} \mid \text{rain}) = 0.6$, $P(\text{bus is late} \mid \text{no rain}) = 0.1$. Find the probability the bus is late.

Achieved

5. At Vertex Academy, 150 students were asked about their transport. The two-way table: Drive, Year 12: 20, Year 13: 40; Take bus, Year 12: 50, Year 13: 40. Find the probability that a randomly selected student is a Year 13 student who drives.

6. A café's coffee machine fails on 5% of mornings. If it works, 90% of customers are served on time; if it fails, only 30% are served on time. Draw the tree in your head and find the probability a customer is served on time.

7. In a health study, 15 of 300 people who exercise regularly developed high blood pressure, compared with 48 of 400 people who don't. Calculate the relative risk for non-exercisers compared with exercisers.

8. The number of texts X a student receives in a lunch break has this distribution: P(X=0)=0.2, P(X=1)=0.3, P(X=2)=0.4, P(X=3)=k. Find k, then find the expected number of texts.

Merit

9. A factory in Penrose has two machines producing USB drives. Machine A makes 70% of drives with a 3% defect rate; Machine B makes 30% with a 6% defect rate. A drive is found defective. What is the probability it came from Machine B? Show your reasoning.

10. A study reports the relative risk of a skin condition for people with high sun exposure is 3.0 compared with low exposure. The risk for the low-exposure group is 0.02. A news headline says 'high sun exposure triples your chance of the condition, huge danger!'. Using the numbers, discuss whether the headline is misleading.

11. Of 500 Auckland commuters surveyed, 180 use a bike-share app. Of app users, 63 had a minor accident last year; of non-users, 64 had one. Determine whether having an accident is independent of using the app, showing the comparison.

Excellence

12. A medical test for a rare condition affecting 1 in 1000 people has sensitivity 99% (correctly positive if you have it) and specificity 95% (correctly negative if you don't). A person tests positive. Find the probability they actually have the condition, and explain why the answer seems surprisingly low.

13. Events A and B satisfy P(A) = a and P(B) = b, and are independent. Show that A and the complement of B (B') are also independent, giving a general algebraic argument.

14. A game charges $2 to play. You roll a fair die: rolling a 6 wins $10, rolling a 4 or 5 wins $3, otherwise nothing. Determine whether the game is worth playing in the long run, then find the winning payout the $10 prize would need to become for the game to be exactly fair (expected profit zero), keeping all else the same.

NCEA Level 3 · AS 91586: Working with Probability Distributions

Name: Date:

Hi! This standard is all about choosing the right probability model (normal, binomial, Poisson, uniform or triangular) and using it to answer real questions. Work through each section in order, check your working against the answers, and pay attention to when you have to justify your choice of model, that's where the top grades live.

Revision summary

- Normal: standardise with $z = (x - \mu)/\sigma$, then use tables/calculator for $P(Z < z)$. Inverse: given a probability, find z first, then $x = \mu + z\sigma$.
- Binomial $X \sim B(n, p)$: fixed n independent trials, constant p. $P(X = k) = C(n,k)\, p^k (1-p)^{(n-k)}$; mean = np, variance = np(1−p). Use when counting successes in a set number of trials.
- Poisson $X \sim Po(\lambda)$: events at a constant average rate over a fixed interval. $P(X = k) = e^{(-\lambda)} \lambda^k / k!$; mean = variance = λ. Scale λ in proportion to the interval length.
- Uniform (continuous on [a,b]): constant density $1/(b-a)$; $P(c < X < d) = (d-c)/(b-a)$; mean = (a+b)/2.
- Triangular(a, b, c) with min a, max b, mode c: peak height $2/(b-a)$. $P(X \le c) = (c-a)/(b-a)$. Areas are triangles, so use ½ × base × height.
- Model choice: binomial = fixed trials with success/failure; Poisson = counts over time/space with a rate (use as approx to binomial when n large, p small, λ = np); normal = continuous, symmetric, bell-shaped data.
- Useful z-values: $P(Z<z)=0.90 \rightarrow z=1.2816$; $0.95 \rightarrow 1.6449$; $0.975 \rightarrow 1.96$; 0.02 lower tail → z=−2.054.
- Normal approximation to binomial ($np \ge 5$ and $n(1-p) \ge 5$): $\mu = np$, $\sigma = \sqrt{(np(1-p))}$, and apply a continuity correction of ±0.5.

Key skills

1. X is normally distributed with mean 50 and standard deviation 8. Find $P(X < 60)$.

2. $X \sim$ Binomial($n = 10$, $p = 0.3$). Find $P(X = 4)$.

3. The number of texts Aroha gets per hour follows a Poisson distribution with mean 3. Find $P(X = 2)$.

4. A bus arrival delay is uniformly distributed on the interval [0, 20] minutes. Find $P(5 < X < 12)$.

Achieved

5. The weight of a kererū is normally distributed with mean 650 g and standard deviation 40 g. Find the probability a randomly chosen kererū weighs more than 700 g.

6. A 12-question multiple-choice quiz has four options per question. Manu guesses every answer, so the chance each is correct is 0.25. Find the probability he gets exactly 3 correct.

7. A café receives online orders at an average rate of 5 per hour, following a Poisson distribution. Find the probability that fewer than 3 orders arrive in a given hour.

8. Delivery times for a courier follow a triangular distribution with minimum 20 min, mode 30 min and maximum 50 min. Find the probability a delivery takes less than 30 minutes.

Merit

9. A machine fills juice bottles so the volume is normally distributed with standard deviation 5 mL. The manager wants only 2% of bottles to fall below the labelled 500 mL. What mean fill volume should be set?

10. A factory line produces components that are defective 1.5% of the time, independently. A batch of 200 is checked. Explain why a Poisson model is suitable and find the probability that at most 2 components in the batch are defective.

11. Bags of chips have weights that are normally distributed with mean 150 g and standard deviation 4 g. In a shipment of 500 bags, how many would you expect to weigh between 145 g and 155 g?

Excellence

12. A test score X is normally distributed. It is known that 10% of students score below 40 and 5% score above 75. Find the mean and standard deviation of the distribution.

13. Minor earthquakes felt in Auckland occur at an average rate of 0.8 per year, modelled as Poisson. Find the probability that at least 5 are felt over a 5–year period, and state one assumption the model relies on.

14. A fair coin is tossed 400 times. Using an appropriate distribution, estimate the probability of getting at least 210 heads, and justify why your chosen method is valid.

Trig functions and radians

NCEA Level 3 · AS 91575 · Worksheet

Name _____ Date _____

Achieved

1. Convert to radians: (a) 60° (b) 210°

2. Convert to degrees: (a) $3\pi/2$ (b) $2\pi/3$

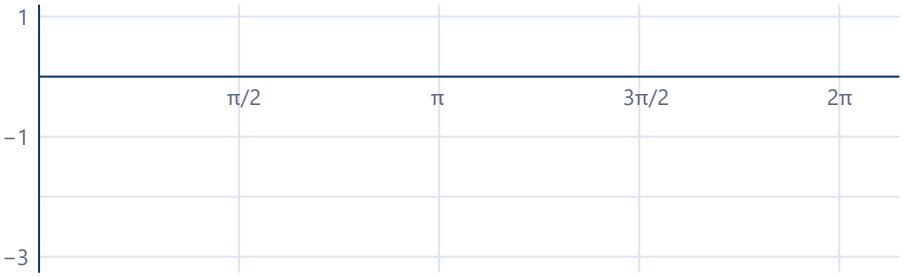
3. Find the exact value: (a) $\sin(\pi/6)$ (b) $\cos(3\pi/4)$ (c) $\tan(5\pi/4)$

4. State the amplitude, period, midline and range of $y = 2\sin(3x) + 5$.

5. State the amplitude, period, midline and range of $y = -4\cos(x/2)$.

Merit

6. Sketch $y = 2\cos(2x) - 1$ for $0 \le x \le 2\pi$. Mark its maximum and minimum points.



7. A sine graph has a maximum of 6, a minimum of -2 and a period of 6π. It starts on its midline, rising. Write its equation.

8. Solve $2 \sin x = -\sqrt{3}$ for $0 \le x \le 2\pi$, giving exact answers.

9. Solve $\tan x = \sqrt{3}$ for $0 \le x \le 2\pi$, giving exact answers.

Excellence

10. $y = a \cos(bx) + c$ has a maximum of 8 at $x = 0$, and its next minimum is 2 at $x = \pi/4$. Find a, b and c.

11. Solve $2 \sin(2x) + 1 = 0$ for $0 \le x \le \pi$, giving exact answers.

12. The depth of water in a harbour is $h = 2 \sin(\pi t/6) + 3$ metres, t hours after noon. (a) State the shallowest and deepest depths. (b) For how long in the first 12 hours is the water deeper than 4 m?

Answers

1. (a) $\pi/3$ (b) $7\pi/6$
2. (a) $270°$ (b) $120°$
3. (a) $1/2$ (b) $-1/\sqrt{2}$ (c) 1
4. Amplitude 2, period $2\pi/3$, midline $y = 5$, range $3 \le y \le 7$
5. Amplitude 4, period 4π, midline $y = 0$, range $-4 \le y \le 4$
6. Maxima of 1 at $x = 0$, π and 2π. Minima of -3 at $x = \pi/2$ and $3\pi/2$.
7. Midline 2, amplitude 4, $b = 2\pi \div 6\pi = 1/3$, so $y = 4\sin(x/3) + 2$
8. $\sin x = -\sqrt{3}/2$. Reference angle $\pi/3$, third and fourth quadrants: $x = 4\pi/3$ or $5\pi/3$
9. Reference angle $\pi/3$, first and third quadrants: $x = \pi/3$ or $4\pi/3$
10. Midline $c = 5$ and amplitude $a = 3$. Half a period is $\pi/4$, so the period is $\pi/2$ and $b = 4$: $y = 3\cos(4x) + 5$
11. $\sin(2x) = -1/2$ with $0 \le 2x \le 2\pi$, so $2x = 7\pi/6$ or $11\pi/6$, and $x = 7\pi/12$ or $11\pi/12$
12. (a) Shallowest 1 m, deepest 5 m. (b) $\sin(\pi t/6) = 1/2$ gives $t = 1$ and $t = 5$, so the water is deeper than 4 m for 4 hours.

NCEA Level 3 · AS 91575: Trig Methods & Periodic Models

Name: _____ Date: _____

Kia ora! This standard is all about solving trig equations (in radians), using compound and double-angle formulae, proving identities, and modelling things that repeat like tides and Ferris wheels. Work top to bottom, nail the Key skills first, then push into the graded questions.

Revision summary

- Radians: π rad = 180°. To convert degrees→radians ×π/180; radians→degrees ×180/π. Common: π/6=30°, π/4=45°, π/3=60°, π/2=90°.
- Compound angles: $\sin(A\pm B)=\sin A\cos B \pm \cos A\sin B$; $\cos(A\pm B)=\cos A\cos B \mp \sin A\sin B$.
- Double angles: $\sin 2\theta=2\sin\theta\cos\theta$; $\cos 2\theta=\cos^2\theta-\sin^2\theta=1-2\sin^2\theta=2\cos^2\theta-1$.
- General solutions: $\sin x=k \rightarrow x=\alpha+2n\pi$ or $(\pi-\alpha)+2n\pi$; $\cos x=k \rightarrow x=\pm\alpha+2n\pi$; $\tan x=k \rightarrow x=\alpha+n\pi$ (n is any integer).
- Pythagorean identity: $\sin^2\theta+\cos^2\theta=1$, so $\tan\theta=\sin\theta/\cos\theta$, $\cot\theta=\cos\theta/\sin\theta$.
- Model form $y=A\sin(B(t-C))+D$: amplitude=$|A|$, period=$2\pi/B$, D=vertical shift (midline), C=horizontal shift. Max=D+A, min=D−A.
- Wave combination: $a\sin x + b\cos x = R\sin(x+\varphi)$, where $R=\sqrt{(a^2+b^2)}$ and $\tan\varphi=b/a$.
- Solving strategy: use identities to get ONE trig ratio, factorise if quadratic-in-trig, solve, then list all solutions in the required interval.

Key skills

1. Convert 135° to radians (exact), and convert $5\pi/6$ radians to degrees.

2. Write down the exact value of $\cos(2\pi/3)$.

3. Use a compound-angle formula to find the exact value of $\sin(75°)$.

4. Simplify $1 - 2\sin^2\theta$ and simplify $2\sin\theta\cos\theta$.

Achieved

5. Solve $2\sin x - 1 = 0$ for $0 \le x \le 2\pi$.

6. Write the general solution (in radians) of $\cos\theta = -1/2$.

7. A and B are acute angles with $\sin A = 3/5$ and $\cos B = 5/13$. Find the exact value of $\sin(A + B)$.

8. A Ferris wheel at Ōrākei has radius 25 m and its centre is 30 m above the ground. It completes one revolution every 20 seconds, starting from the lowest point. Its height is $h(t) = 30 - 25\cos(\pi t/10)$ metres. State the amplitude and period, and find the height after 5 seconds.

Merit

9. Solve $2\cos^2 x + \cos x - 1 = 0$ for $0 \le x \le 2\pi$.

10. Solve $\sin 2x = \cos x$ for $0 \le x \le 2\pi$.

11. The depth of water at a Manukau Harbour jetty is modelled by $d = 8 + 3\sin(\pi t/6)$ metres, where t is hours after midnight. A boat needs a depth of at least 9.5 m to enter. During the first 12 hours ($0 \le t \le 12$), for what time interval is the depth at least 9.5 m?

Excellence

12. Prove the identity $(1 + \cos 2\theta)/\sin 2\theta = \cot\theta$, and state any values of θ for which it is undefined.

13. Express $3\sin x + 4\cos x$ in the form $R\sin(x + \varphi)$ with $R > 0$ and $0 < \varphi < \pi/2$, then find the general solution of $3\sin x + 4\cos x = 2.5$.

14. Prove that $\sin 3\theta = 3\sin\theta - 4\sin^3\theta$.

NCEA Level 3 · AS 91580: Time Series Investigations

Name: _____ Date: _____

Hi! This standard is all about spotting patterns in data collected over time, trend, seasonal ups and downs, and using them to forecast. Work through the sheet in order, and always link your maths back to the real-world context of the data.

Revision summary

- Time series data = measurements recorded at regular time intervals (e.g. quarterly sales). Look for: trend (long-term direction), seasonality (regular repeating pattern), and irregular/random variation.
- Additive model: Data = Trend + Seasonal + Irregular ($y = T + S + I$). Use when seasonal swings stay roughly constant in size over time.
- Centred Moving Average (CMA) smooths out seasonality to reveal the trend. For quarterly data (period 4) use a 4-point moving average then centre it (average two consecutive 4-point means).
- Seasonal effect for a season = average of (actual value − CMA) for that season. Seasonal effects should sum to ≈ 0 across the cycle; if not, adjust each by subtracting the mean.
- Forecasting: extend the trend line (from the CMA or a fitted trend), then add the relevant seasonal effect for that season: Forecast = Trend estimate + Seasonal effect.
- Interpret in context: describe trend (increasing/decreasing, by how much), seasonal pattern (which seasons are high/low and why), and comment on the reliability of forecasts.
- Statistical enquiry cycle (PPDAC): Problem, Plan, Data, Analysis, Conclusion, state a purpose, explore the data, and write conclusions in context.

Key skills

1. A café in Devonport records quarterly hot-chocolate sales (in hundreds of cups): Q1 = 8, Q2 = 5, Q3 = 4, Q4 = 9. Calculate the 4-point moving average for these four quarters.

2. State what the 'seasonal effect' of a particular quarter tells you in an additive time series model.

3. The additive model gives a trend estimate of 42 units for the next winter quarter, and winter has a seasonal effect of −6.5. Forecast the value for that quarter.

4. Explain why we 'centre' a moving average when working with quarterly data.

Achieved

5. A garden centre in Henderson records quarterly plant sales ($000): 2022 Q1 22, Q2 30, Q3 18, Q4 14; 2023 Q1 26, Q2 34, Q3 22, Q4 18. Describe the trend and the seasonal pattern you can see.

6. Using the garden-centre data above, calculate the first two 4-point moving averages and the first centred moving average.

7. A power company's data shows the CMA (trend) for electricity use rising by about 5 units per quarter, reaching 210 units at Q4 2023. The Q1 seasonal effect is +18. Forecast Q1 2024.

8. The seasonal effects for a quarterly series are found to be Q1 = +6, Q2 = +11, Q3 = −4, Q4 = −11. Check whether they need adjusting, and adjust if required.

Merit

9. A Ponsonby ice-cream shop's quarterly revenue shows a clear seasonal pattern with Q2 (summer months Jan–Mar) highest. The CMA trend is increasing steadily. Explain how you would produce a forecast for next summer (Q2), and why the additive model is suitable here.

10. A tourism operator sees quarterly visitor numbers with a strong seasonal pattern, but the summer peaks are getting larger and larger each year as the business grows. Explain why forecasts from a simple additive model with constant seasonal effects could become unreliable, and what this means for the operator.

11. For a quarterly dataset, a student calculated seasonal effects but forgot to centre their 4-point moving average. Explain what problem this causes and how it would show up when they compare their trend to the actual data.

Excellence

12. Two analysts model the same quarterly rainfall series for Whangārei. One uses an additive model, the other a multiplicative model (Data = Trend × Seasonal × Irregular). The trend is roughly flat over the period. Argue what you would expect about the two sets of forecasts, and under what condition they would give nearly identical results.

13. A student claims: 'If I add up all the CMA (centred moving average) values and compare with the sum of the original data over the same complete years, they should be about equal.' Evaluate whether this claim is justified.

14. A local council forecasts quarterly water demand using an additive model built on 3 years of data. The forecasts extend 8 quarters (2 years) ahead. Discuss, with justification, the factors that limit how much confidence should be placed in the 8th forecast compared with the 1st.

NCEA Level 3 · AS 91581: Investigating Bivariate Data

Name: _____ Date: _____

This internal is all about exploring the relationship between two measurement variables using the statistical enquiry cycle. Work through the fluency questions first, then use the graded sections to practise describing, modelling, predicting and evaluating like a statistician.

Revision summary

- The PPDAC cycle: Problem (pose a relationship question), Plan, Data, Analysis (scatter plot, trend, model), Conclusion (answer in context with limitations).
- Scatter plot description covers TAFOS: Trend (linear/non-linear), Association (positive/negative), strength of correlation, Form/features, Outliers, Scatter (spread about the trend).
- Correlation coefficient r ranges from −1 to +1. Near ±1 = strong linear relationship; near 0 = weak/none. Correlation is NOT causation.
- Linear model: $y = a + bx$ (or predicted variable = intercept + slope × explanatory variable). Slope b = change in y per 1-unit increase in x; intercept a = predicted y when x = 0.
- Interpolation (predicting inside the data range) is more reliable than extrapolation (predicting outside the range), which is risky.
- Always define which variable is explanatory (x) and which is the response (y); the model predicts the response.
- Evaluate predictions by checking r, the scatter about the line, sample size, whether it is interpolation or extrapolation, and real-world context.

Key skills

1. A study of 40 Avondale cafés recorded the daily temperature (°C) and the number of iced coffees sold. The correlation coefficient is r = 0.82. Describe the strength and direction of the linear relationship.

2. A linear model for petrol used (litres) against distance driven (km) is: litres = 2.5 + 0.09 × distance. Interpret the slope in context.

3. Using the model litres = 2.5 + 0.09 × distance, predict the petrol used for a 150 km trip.

4. A scatter plot of house size (m²) versus sale price ($) shows points curving upward and getting more spread out for larger houses. Name two features you would describe from this plot.

Achieved

5. Mere collects data on the age (years) and resale value ($) of 25 used utes in West Auckland. She fits the model: value = 42000 − 3100 × age. Describe the relationship and predict the value of a 6-year-old ute.

6. A scatter plot shows the height (cm) and arm span (cm) of 30 students at a Henderson high school. The points lie close to a straight line rising from left to right, with one point far below the others. Describe the trend, direction, strength and any unusual features.

7. For a data set of weekly rainfall (mm) and weekly grass growth (cm) at a Waikato farm, r = 0.15. What does this tell you about the linear relationship?

8. A model for a car's stopping distance (m) against speed (km/h) is: distance = −4 + 0.85 × speed, fitted from data for speeds between 30 and 100 km/h. Explain why using this model to predict the stopping distance at 5 km/h could be unreliable.

Merit

9. A café owner in Avondale investigates the relationship between the number of staff on shift (2 to 8) and the average customer wait time (minutes). Her model is: wait = 14.2 − 1.6 × staff, and r = −0.88. Using the model and the correlation, discuss what the data suggests, and comment on how much you would trust a prediction of wait time for 5 staff versus for 12 staff.

10. Two variables, daily ice-cream sales and daily drowning incidents at NZ beaches, show a strong positive correlation (r = 0.9). A student concludes that buying ice-cream causes drownings. Explain why this reasoning is flawed and suggest a better explanation for the relationship.

11. A researcher compares two possible linear models for the same data on tree age (years) and trunk diameter (cm): Model A has r = 0.72; Model B (after removing one influential outlier) has r = 0.91. Explain what the change in r suggests and discuss whether removing the point is justified.

Excellence

12. A regional council models nitrogen level (mg/L) in a stream against the percentage of surrounding land used for dairy farming, based on 50 catchments. The model is nitrogen = 0.8 + 0.12 × dairy%, with r = 0.68. A colleague argues that because r is only moderate, the model is 'useless' for the council's decisions. Evaluate this claim, discussing the value AND limitations of the model for guiding policy.

13. For a linear model $y = a + bx$ fitted to data, explain algebraically and conceptually why the fitted line always passes through the point $(\bar{x}, \bar{y})$, the mean of the explanatory variable and the mean of the response. Then explain what this tells you about interpreting the intercept a.

14. Two students investigate the same relationship between weekly study hours and exam mark (%). One uses hours as the explanatory variable and gets slope 3.2; the other swaps the axes, using mark as the explanatory variable, and gets a different slope. They are confused that $1 \div 3.2 \neq$ their slope. Explain why swapping which variable is explanatory generally produces a different line that is NOT simply the algebraic inverse.

NCEA Level 3 · AS 91582: Making the Call with Confidence Intervals

Name: _____ Date: _____

Hi! This standard is about using a bootstrap confidence interval to make a formal inference comparing two groups back in their populations. Work through each section in order, and always finish by writing your 'call' in context.

Revision summary

- Formal inference = making a claim about a POPULATION difference using a SAMPLE.
- Bootstrap: resample each sample WITH replacement (same size), record the difference in statistics (e.g. median A − median B), repeat ~1000 times to build a bootstrap distribution.
- Confidence interval (CI): the middle 90% of the bootstrap distribution, from the 5th percentile to the 95th percentile of the resampled differences.
- Making the call: if the whole CI is above 0, infer group A's population parameter is larger; if entirely below 0, it is smaller; if the CI contains 0, you cannot make a call (the difference could plausibly be zero).
- Interpretation: 'It is a fairly safe bet that the median [variable] for [population A] is between X and Y greater/less than for [population B].'
- Sampling variability: different samples give different statistics; larger samples give narrower CIs and more reliable calls.
- Always compare like with like, use the sample size, spread, shape and any outliers to support or temper your inference.
- Good justification links the CI position relative to 0, the sample sizes, and the practical context.

Key skills

1. A bootstrap confidence interval for the difference in median weekly grocery spend (Avondale households − Blockhouse Bay households) is (−$5, $42). Does this interval let you make a call about which population has the higher median spend? Briefly say why.

2. Explain what 'resampling with replacement' means when bootstrapping from a sample of 30 commute times.

3. A bootstrap distribution of 1000 differences in mean rainfall (mm) is sorted. The 50th value is 2.1 and the 950th value is 18.4. Write down the 90% bootstrap confidence interval.

4. A CI for the difference in median screen time (Year 13 students − Year 9 students) is (0.4 h, 2.7 h). Write the formal call in context.

Achieved

5. A researcher samples 40 kūmara from a Dargaville farm and 40 from a Kaipara farm, recording weight (g). A bootstrap gives a 90% confidence interval for the difference in median weight (Dargaville − Kaipara) of (12 g, 58 g). Make a formal inference in context.

6. From samples of daily bus wait times (minutes) in Onehunga and Ōtāhuhu, a bootstrap CI for the difference in medians (Onehunga − Ōtāhuhu) is (−9.2, −1.5). What can you conclude?

7. Explain why we make an inference about the POPULATION and not just report which SAMPLE median is higher.

8. A CI for the difference in mean weekly earnings (part-time workers aged 18–20 − aged 21–24) in a survey is (−$85, −$10). Give the call and state one reason a wider sample might change your confidence.

Merit

9. Two samples of marathon finish times (minutes) are taken: Auckland runners (n = 60, sample median 248, IQR 42) and Wellington runners (n = 55, sample median 261, IQR 39). A bootstrap 90% CI for the difference in medians (Auckland − Wellington) is (−24, −2). Discuss what this tells you, linking the sample summaries to the inference.

10. A student takes two samples from the same population by mistake and bootstraps a CI for the difference in medians, getting (−3.1, 3.4). They conclude 'there is no difference in the populations.' Explain why this wording is not quite right and give a better statement.

11. Explain, with reference to how the bootstrap distribution is built, why increasing each sample size from 20 to 80 tends to give a narrower confidence interval.

Excellence

12. A council claims 'water use per household is the same in Titirangi and Glen Eden.' Samples give a 90% bootstrap CI for the difference in median daily water use (Titirangi − Glen Eden) of (18 L, 61 L). A councillor argues this 'proves' the council is wrong. Critique both the council's claim and the councillor's argument, and state what can properly be concluded.

13. Two analysts study the same two populations. Analyst A uses samples of size 25 and gets CI (−2, 30) for the difference in medians; Analyst B uses samples of size 200 and gets CI (6, 22). Explain how both results can be consistent, and discuss which supports a stronger inference and why.

14. Explain why a bootstrap confidence interval built by resampling can be trusted to reflect population uncertainty, even though we only ever resample from a single original sample and never from the population itself.

NCEA Level 3 · AS 91587: Systems of Simultaneous Equations

Name: _____ Date: _____

This standard is all about solving and modelling problems with three unknowns using Gaussian elimination, and knowing when a system has one solution, infinitely many, or none. Work through the sections in order, the Key skills warm up your method, then push into the Merit and Excellence reasoning that earns the top grades.

Revision summary

- A system of 3 equations in 3 unknowns (x, y, z) usually represents three planes; the solution is where they all meet.
- Gaussian elimination: use row operations (swap rows, multiply a row by a non-zero number, add a multiple of one row to another) to reach upper-triangular form, then back-substitute.
- Unique solution: elimination gives a leading variable in every row (three planes meet at one point).
- Infinite solutions: a row becomes 0 = 0 (a redundant equation); write the answer parametrically, e.g. let z = t and express x and y in terms of t.
- No solution: a row becomes 0 = k where k ≠ 0 (a contradiction, the planes have no common point).
- To model: define each unknown clearly, turn each condition into one linear equation, then solve and interpret the answer in context.
- For a system with a parameter, isolate the parameter in the final row: (parameter)·variable = value decides unique / infinite / none.
- Always check a solution by substituting back into every original equation.

Key skills

1. Solve this triangular system by back-substitution: $x + 2y - z = 3$; $y + 2z = 5$; $z = 2$.

2. Use elimination to solve: $2x + y - z = 3$; $x - y + 2z = 1$; $x + y + z = 4$.

3. A system reduces so that its final row becomes: (a) $0 = 0$, (b) $0 = 7$, (c) a leading variable in each of the three rows. State the number of solutions in each case.

4. Write the solution set in parametric form: $x + y + z = 5$; $x + 2y + 3z = 8$; $2x + 3y + 4z = 13$.

Achieved

5. Solve the system: $x + y + z = 100$; $2x + y + 3z = 180$; $x + 2y + z = 150$.

6. At a Avondale café, three orders are placed. Order 1: 2 pies, 3 coffees and 1 muffin cost $26. Order 2: 1 pie, 2 coffees and 2 muffins cost $20. Order 3: 3 pies, 1 coffee and 1 muffin cost $20. Find the price of each item.

7. Determine whether the system has a unique solution, infinitely many, or none: $x + y + z = 4$; $2x + 2y + 2z = 8$; $x - y + z = 2$. If infinite, give the solution parametrically.

8. $10 000 is split between three accounts earning 3%, 4% and 5% per year. The total interest for the year is $370, and the amount in the first account equals the total of the other two. How much is in each account?

Merit

9. A concert sells adult, student and child tickets at $25, $15 and $10. A total of 500 tickets are sold for total revenue of $9000, and the number of students is twice the number of children. Find how many of each ticket were sold.

10. A parabola $y = ax^2 + bx + c$ passes through the points (1, 6), (2, 11) and (3, 18). Find its equation, connecting the points to a system of equations.

11. Reduce and interpret: $x + 2y - z = 3$; $2x - y + 3z = 4$; $x - 3y + 4z = 1$. State the solution type and give the general solution.

Excellence

12. For the system $x + y + z = 6$; $x + 2y + 3z = 10$; $x + 2y + kz = m$, find the value(s) of k and m for which the system has (i) a unique solution, (ii) infinitely many solutions, (iii) no solution. Justify your answer.

13. Investigate the system $ax + y + z = 1$; $x + ay + z = 1$; $x + y + az = 1$. Determine all values of a giving a unique solution, infinitely many, or none, and justify.

14. Balancing the combustion reaction $a\cdot C_3H_8 + b\cdot O_2 \rightarrow c\cdot CO_2 + d\cdot H_2O$ leads to a system for a, b, c, d. Set up the equations, show it has infinitely many solutions, and explain how a valid balanced equation is chosen.

Answer key (for parents and tutors)

1. Simplify $\sqrt{(-48)}$.
   $\sqrt{(-48)} = \sqrt{48} \times \sqrt{(-1)} = \sqrt{(16 \times 3)} \times i = 4\sqrt{3}\,i$.
   Achieved fluency: rewrites the square root of a negative using i.

2. Write $z = 2 + 2i$ in polar form.
   $r = \sqrt{(2^2 + 2^2)} = \sqrt{8} = 2\sqrt{2}$. The point is in the first quadrant, so $\arg = \arctan(2/2) = \pi/4$. Therefore $z = 2\sqrt{2}\,\mathrm{cis}(\pi/4)$.
   Achieved fluency: converts rectangular to polar with correct modulus and argument.

3. Expand and simplify $(3 + 2i)(1 - 4i)$.
   $(3 + 2i)(1 - 4i) = 3 - 12i + 2i - 8i^2 = 3 - 10i + 8 = 11 - 10i$.
   Achieved fluency: multiplies complex numbers using $i^2 = -1$.

4. Use De Moivre's theorem to evaluate $(2\,\mathrm{cis}(\pi/6))^3$, giving your answer in rectangular form.
   $(2\,\mathrm{cis}(\pi/6))^3 = 2^3\,\mathrm{cis}(3 \times \pi/6) = 8\,\mathrm{cis}(\pi/2) = 8(\cos \pi/2 + i \sin \pi/2) = 8(0 + i) = 8i$.
   Achieved fluency: applies De Moivre's theorem to a power.

5. Express $(4 + 3i)/(2 - i)$ in the form $a + bi$.
   Multiply numerator and denominator by the conjugate $(2 + i)$: numerator $= (4 + 3i)(2 + i) = 8 + 4i + 6i + 3i^2 = 8 + 10i - 3 = 5 + 10i$. Denominator $= (2 - i)(2 + i) = 4 + 1 = 5$. So the result $= (5 + 10i)/5 = 1 + 2i$.
   Achieved: selects the conjugate method to divide complex numbers.

6. Write $z = -\sqrt{3} + i$ in polar form with $-\pi < \theta \leq \pi$.
   $r = \sqrt{((-\sqrt{3})^2 + 1^2)} = \sqrt{(3 + 1)} = 2$. The point is in the second quadrant. Reference angle $= \arctan(1/\sqrt{3}) = \pi/6$, so $\arg = \pi - \pi/6 = 5\pi/6$. Therefore $z = 2\,\mathrm{cis}(5\pi/6)$.
   Achieved: converts to polar and places the argument in the correct quadrant.

7. Solve $z^2 = -9$ for z.
   $z^2 = -9$ means $z = \pm\sqrt{(-9)} = \pm\sqrt{9} \times i = \pm 3i$. So $z = 3i$ or $z = -3i$.
   Achieved: solves a simple quadratic giving complex roots.

8. Rationalise the denominator and simplify: $2/(\sqrt{7} - \sqrt{3})$.
   Multiply top and bottom by the conjugate $(\sqrt{7} + \sqrt{3})$: $2(\sqrt{7} + \sqrt{3}) / ((\sqrt{7} - \sqrt{3})(\sqrt{7} + \sqrt{3})) = 2(\sqrt{7} + \sqrt{3})/(7 - 3) = 2(\sqrt{7} + \sqrt{3})/4 = (\sqrt{7} + \sqrt{3})/2$.
   Achieved: applies surd manipulation by rationalising a denominator.

9. Find all three cube roots of $8i$, giving your answers in rectangular form.
   Write $8i$ in polar form: $8i = 8\,\mathrm{cis}(\pi/2)$. Cube roots have modulus $\sqrt[3]{8} = 2$ and arguments $(\pi/2 + 2\pi k)/3$ for $k = 0, 1, 2$. $k = 0$: $\arg = \pi/6 \to 2\,\mathrm{cis}(\pi/6) = 2(\sqrt{3}/2 + i/2) = \sqrt{3} + i$. $k = 1$: $\arg = \pi/6 + 2\pi/3 = 5\pi/6 \to 2\,\mathrm{cis}(5\pi/6) = 2(-\sqrt{3}/2 + i/2) = -\sqrt{3} + i$. $k = 2$: $\arg = \pi/6 + 4\pi/3 = 3\pi/2 \to 2\,\mathrm{cis}(3\pi/2) = 2(0 - i) = -2i$. Roots: $\sqrt{3} + i$, $-\sqrt{3} + i$, $-2i$.
   Merit: links polar form, De Moivre and root spacing across multiple steps.

10. Given $z = 1 + i$, use De Moivre's theorem to find z^6 in rectangular form.
    Convert: $r = \sqrt{(1^2 + 1^2)} = \sqrt{2}$, $\arg = \pi/4$, so $z = \sqrt{2}\,\mathrm{cis}(\pi/4)$. Then $z^6 = (\sqrt{2})^6\,\mathrm{cis}(6 \times \pi/4) = 8\,\mathrm{cis}(3\pi/2) = 8(\cos 3\pi/2 + i \sin 3\pi/2) = 8(0 - i) = -8i$.
    Merit: connects conversion to polar with De Moivre to evaluate a high power efficiently.

11. Describe and identify the locus of points z satisfying $|z - 2| = |z + 2i|$.
    Let $z = x + iy$. Then $|z - 2|^2 = (x - 2)^2 + y^2$ and $|z + 2i|^2 = x^2 + (y + 2)^2$. Setting them equal: $(x - 2)^2 + y^2 = x^2 + (y + 2)^2$. Expand: $x^2 - 4x + 4 + y^2 = x^2 + y^2 + 4y + 4$. Simplify: $-4x = 4y$, so $y = -x$. The locus is the straight line $y = -x$, which is the perpendicular bisector of the segment joining the points 2 (i.e. $2 + 0i$) and $-2i$ (i.e. $0 - 2i$).
    Merit: connects the algebraic condition to a geometric locus and identifies it.

12. Prove, using mathematical induction, that $(\cos \theta + i \sin \theta)^n = \cos n\theta + i \sin n\theta$ for all positive integers n.
    Base case $n = 1$: $(\cos \theta + i \sin \theta)^1 = \cos \theta + i \sin \theta = \cos(1 \cdot \theta) + i \sin(1 \cdot \theta)$. True. Inductive step: assume it holds for $n = k$, i.e. $(\cos \theta + i \sin \theta)^k = \cos k\theta + i \sin k\theta$. Then $(\cos \theta + i \sin \theta)^{(k+1)} = (\cos \theta + i \sin \theta)^k (\cos \theta + i \sin \theta) = (\cos k\theta + i \sin k\theta)(\cos \theta + i \sin \theta)$. Expanding: $= \cos k\theta \cos \theta + i \cos k\theta \sin \theta + i \sin k\theta \cos \theta + i^2 \sin k\theta \sin \theta = (\cos k\theta \cos \theta - \sin k\theta \sin \theta) + i(\sin k\theta \cos \theta + \cos k\theta \sin \theta)$. Using the compound-angle formulae: $= \cos(k\theta + \theta) + i \sin(k\theta + \theta) = \cos((k+1)\theta)$

+ i sin((k+1)θ). So it holds for n = k + 1. By induction it holds for all positive integers n.
Excellence: constructs a rigorous inductive proof using compound-angle identities.

13. Show that the sum of all n distinct nth roots of unity is 0 for n ≥ 2.
The nth roots of unity are ω^k for k = 0, 1, …, n−1, where ω = cis(2π/n). These are the roots of $z^n - 1 = 0$. Their sum is a geometric series: S = 1 + ω + ω^2 + … + ω^(n−1) = (ω^n − 1)/(ω − 1). Since ω = cis(2π/n), ω^n = cis(2π) = 1, so the numerator is 0. Because n ≥ 2, ω ≠ 1, so the denominator is non-zero. Therefore S = 0/(ω − 1) = 0. (Equivalently, by Vieta's formulas the sum of roots of $z^n - 1$ equals the negative of the coefficient of z^(n−1), which is 0.)
Excellence: generalises and justifies a result about roots of unity using summation or Vieta's.

14. A complex number z satisfies z + 1/z = 2cos θ. Prove that $z^n + 1/z^n$ = 2cos nθ for every positive integer n.
From z + 1/z = 2cos θ, multiply by z: z^2 − 2cos θ · z + 1 = 0. Solving: z = (2cos θ ± √(4$\cos^2\theta$ − 4))/2 = cos θ ± √($\cos^2\theta$ − 1) = cos θ ± √(−$\sin^2\theta$) = cos θ ± i sin θ. Take z = cos θ + i sin θ = cis θ (the other root gives the same result by symmetry). By De Moivre's theorem, z^n = cis(nθ) = cos nθ + i sin nθ. Also $1/z^n$ = z^(−n) = cis(−nθ) = cos nθ − i sin nθ. Adding: $z^n + 1/z^n$ = (cos nθ + i sin nθ) + (cos nθ − i sin nθ) = 2cos nθ. Hence proven for all positive integers n.
Excellence: solves an unfamiliar problem by deducing polar form and generalising with De Moivre.

Answer key (for parents and tutors)

1. Differentiate $y = (3x^2 + 1)^5$.
   Use the chain rule with $u = 3x^2 + 1$, so $y = u^5$. $dy/du = 5u^4$ and $du/dx = 6x$. $dy/dx = 5(3x^2 + 1)^4 \times 6x = 30x(3x^2 + 1)^4$.
   Key skill: applies the chain rule to a power of a polynomial.

2. Differentiate $y = x^2e^x$.
   Product rule with $u = x^2$, $v = e^x$. $u' = 2x$, $v' = e^x$. $dy/dx = u'v + uv' = 2x \cdot e^x + x^2 \cdot e^x = e^x(x^2 + 2x) = x\,e^x(x + 2)$.
   Key skill: applies the product rule combining a polynomial and e^x.

3. Differentiate $y = (2x + 1)/(x - 3)$.
   Quotient rule with $u = 2x + 1$, $v = x - 3$. $u' = 2$, $v' = 1$. $dy/dx = (u'v - uv')/v^2 = [2(x - 3) - (2x + 1)(1)]/(x - 3)^2 = (2x - 6 - 2x - 1)/(x - 3)^2 = -7/(x - 3)^2$.
   Key skill: applies the quotient rule and simplifies.

4. Differentiate $y = \ln(\sin x)$.
   Chain rule for a log: $d/dx(\ln f) = f'/f$ with $f = \sin x$, $f' = \cos x$. $dy/dx = \cos x / \sin x = \cot x$.
   Key skill: differentiates a logarithm of a trig function using the chain rule.

5. Find the gradient of the curve $y = 3\sin(2x)$ at $x = \pi/6$.
   $dy/dx = 3 \times \cos(2x) \times 2 = 6\cos(2x)$. At $x = \pi/6$: $6\cos(2 \times \pi/6) = 6\cos(\pi/3) = 6 \times 0.5 = 3$. The gradient is 3.
   Achieved: selects the chain rule for a trig function and evaluates at a point.

6. Find the coordinates of the stationary points of $y = x^3 - 6x^2 + 9x + 2$.
   $dy/dx = 3x^2 - 12x + 9 = 3(x^2 - 4x + 3) = 3(x - 1)(x - 3)$. Set = 0: $x = 1$ or $x = 3$. $y(1) = 1 - 6 + 9 + 2 = 6$; $y(3) = 27 - 54 + 27 + 2 = 2$. Stationary points are (1, 6) and (3, 2).
   Achieved: finds and evaluates stationary points from $f'(x) = 0$.

7. A trolley moves so its position is $s = t^3 - 4t^2 + 2t$ metres after t seconds. Find its velocity at $t = 3$ s.
   Velocity $v = ds/dt = 3t^2 - 8t + 2$. At $t = 3$: $v = 3(9) - 8(3) + 2 = 27 - 24 + 2 = 5$. The velocity is 5 m/s.
   Achieved: differentiates a position function to obtain velocity at a time.

8. Find the equation of the tangent to $y = e^{(2x)}$ at the point where $x = 0$.
   $dy/dx = 2e^{(2x)}$. At $x = 0$ the gradient is $2e^0 = 2$. The point is $(0, e^0) = (0, 1)$. Tangent: $y - 1 = 2(x - 0)$, so $y = 2x + 1$.
   Achieved: finds a tangent line using the derivative of an exponential.

9. A farmer near Pukekohe uses 120 m of fencing to make a rectangular paddock against a straight river. No fence is needed along the river. Find the dimensions that give the maximum area, and state that maximum area.
   Let the two sides perpendicular to the river be x and the side parallel be y. Fencing used: $2x + y = 120$, so $y = 120 - 2x$. Area $A = xy = x(120 - 2x) = 120x - 2x^2$. $dA/dx = 120 - 4x = 0 \rightarrow x = 30$, so $y = 120 - 60 = 60$. $A'' = -4 < 0$, confirming a maximum. Dimensions 30 m × 60 m; maximum area = $30 \times 60 = 1800$ m².
   Merit: links a constraint to the area function and justifies the maximum.

10. A closed cylindrical can must hold 500 cm³. Find the radius that minimises the surface area (give it to 2 d.p.).
    Volume: $\pi r^2 h = 500$, so $h = 500/(\pi r^2)$. Surface area $S = 2\pi r^2 + 2\pi rh = 2\pi r^2 + 2\pi r(500/(\pi r^2)) = 2\pi r^2 + 1000/r$. $dS/dr = 4\pi r - 1000/r^2$. Set = 0: $4\pi r^3 = 1000 \rightarrow r^3 = 250/\pi \approx 79.577 \rightarrow r \approx 4.30$ cm. Since dS/dr changes from negative to positive (or $S'' = 4\pi + 2000/r^3 > 0$), this is a minimum. Minimising radius ≈ 4.30 cm.
    Merit: connects volume constraint to surface area and solves the optimisation.

11. A spherical weather balloon is inflated so its radius increases at 2 cm/s. Find the rate at which its volume is increasing when the radius is 5 cm.
    $V = (4/3)\pi r^3$, so $dV/dr = 4\pi r^2$. By the chain rule $dV/dt = dV/dr \times dr/dt = 4\pi r^2 \times 2$. At $r = 5$: $dV/dt = 4\pi(25)(2) = 200\pi \approx 628.3$ cm³/s.
    Merit: chains related rates together using the chain rule.

12. For the curve $y = x^3 - 3kx + 2$, determine all values of the constant k for which the curve has two distinct stationary points. Justify your answer.
    $dy/dx = 3x^2 - 3k$. Stationary points occur where $3x^2 - 3k = 0$, i.e. $x^2 = k$. This equation has two distinct real solutions $x = \pm\sqrt{k}$ only when $k > 0$ (if $k = 0$ there is a single repeated solution $x = 0$, a stationary point of inflection; if

$k < 0$ there are no real solutions). Therefore the curve has two distinct stationary points exactly when $k > 0$.
Excellence: generalises a condition on a parameter and justifies each case.

13. Using calculus, prove that among all rectangles with a fixed perimeter P, the square gives the greatest area.
Let one side be x, so the adjacent side is $P/2 - x$ (from $2x + 2(\text{other}) = P$). Area $A(x) = x(P/2 - x) = (P/2)x - x^2$, valid for $0 < x < P/2$. $A'(x) = P/2 - 2x$. Setting $A'(x) = 0$ gives $x = P/4$, and then the other side $= P/2 - P/4 = P/4$, so the rectangle is a square. $A''(x) = -2 < 0$, so this stationary point is a maximum. Hence the square maximises the area for a given perimeter.
Excellence: proves a general optimisation result with full justification.

14. A particle moves along a line with displacement $x(t) = e^{(-t)} \sin t$ metres for $t \geq 0$. Show that the particle is instantaneously at rest whenever $\tan t = 1$, and explain why the exponential factor never causes it to stop.
Velocity: $x'(t) = d/dt[e^{(-t)} \sin t]$. Product rule: $x'(t) = (-e^{(-t)})\sin t + e^{(-t)}\cos t = e^{(-t)}(\cos t - \sin t)$. The particle is at rest when $x'(t) = 0$. Since $e^{(-t)} > 0$ for all real t, it is never zero, so it can never make the velocity zero on its own. Thus $x'(t) = 0$ requires $\cos t - \sin t = 0$, i.e. $\cos t = \sin t$, i.e. $\tan t = 1$. This occurs at $t = \pi/4, 5\pi/4, 9\pi/4, \ldots$ ($t = \pi/4 + n\pi$). Because $e^{(-t)}$ stays strictly positive, the only way the velocity vanishes is through the trig factor, confirming the condition $\tan t = 1$.
Excellence: differentiates an unfamiliar product and reasons about why a factor cannot be zero.

Answer key (for parents and tutors)

1. Find $\int(6x^2 - 4x + 5)\,dx$.
   Integrate term by term: $6x^2 \to 6x^3/3 = 2x^3$; $-4x \to -4x^2/2 = -2x^2$; $5 \to 5x$. So $\int(6x^2 - 4x + 5)\,dx = 2x^3 - 2x^2 + 5x + c$.
   Achieved: applies the power rule to a polynomial.

2. Find $\int(e^{3x} + 2\cos x)\,dx$.
   $\int e^{3x}\,dx = (1/3)e^{3x}$; $\int 2\cos x\,dx = 2\sin x$. So the answer is $(1/3)e^{3x} + 2\sin x + c$.
   Achieved: recalls exponential and trig anti-derivatives.

3. Evaluate $\int$ from 1 to 3 of $(2x)\,dx$.
   Anti-derivative of $2x$ is x^2. Evaluate: $(3^2) - (1^2) = 9 - 1 = 8$.
   Achieved: applies the fundamental theorem to a definite integral.

4. Find $\int(4/x)\,dx$ for $x > 0$.
   $\int(4/x)\,dx = 4\int(1/x)\,dx = 4\ln|x| + c = 4\ln x + c$ (since $x > 0$).
   Achieved: recognises the 1/x standard form giving a logarithm.

5. The gradient function of a curve is $dy/dx = 3x^2 - 8x$. The curve passes through the point $(2, 1)$. Find the equation of the curve.
   Integrate: $y = \int(3x^2 - 8x)\,dx = x^3 - 4x^2 + c$. Substitute $(2, 1)$: $1 = 2^3 - 4(2^2) + c = 8 - 16 + c = -8 + c$, so $c = 9$. The curve is $y = x^3 - 4x^2 + 9$.
   Achieved: integrates a gradient function and uses a point to find c.

6. Find the area between the curve $y = x^2 + 1$ and the x-axis from $x = 0$ to $x = 2$.
   Area = $\int$ from 0 to 2 of $(x^2 + 1)\,dx = [x^3/3 + x]$ from 0 to 2 $= (8/3 + 2) - (0) = 8/3 + 6/3 = 14/3 \approx 4.67$ square units.
   Achieved: sets up and evaluates a definite integral for area.

7. Use the substitution $u = x^2 + 3$ to find $\int 2x(x^2 + 3)^4\,dx$.
   Let $u = x^2 + 3$, so $du = 2x\,dx$. The integral becomes $\int u^4\,du = u^5/5 + c$. Substitute back: $(x^2 + 3)^5/5 + c$.
   Achieved: applies a given substitution correctly.

8. Water flows into a tank at a rate of $dV/dt = 6t + 2$ litres per minute, where t is in minutes. If the tank starts empty, find the volume of water after 5 minutes.
   $V = \int(6t + 2)\,dt = 3t^2 + 2t + c$. Tank starts empty: at $t = 0$, $V = 0$, so $c = 0$. Thus $V = 3t^2 + 2t$. At $t = 5$: $V = 3(25) + 2(5) = 75 + 10 = 85$ litres.
   Achieved: integrates a rate and applies an initial condition.

9. Find the area enclosed between the curves $y = x^2$ and $y = 6x - x^2$.
   Find intersections: $x^2 = 6x - x^2 \to 2x^2 - 6x = 0 \to 2x(x - 3) = 0$, so $x = 0$ and $x = 3$. Between these, $6x - x^2$ is the top curve. Area = $\int$ from 0 to 3 of $((6x - x^2) - x^2)\,dx = \int$ from 0 to 3 of $(6x - 2x^2)\,dx = [3x^2 - 2x^3/3]$ from 0 to 3 $= (3(9) - 2(27)/3) - 0 = (27 - 18) = 9$ square units.
   Merit: links intersection points and the difference of functions to find enclosed area.

10. Find $\int x\sqrt{(2x^2 + 1)}\,dx$ using a suitable substitution, showing your choice.
    Let $u = 2x^2 + 1$, so $du = 4x\,dx$, meaning $x\,dx = du/4$. Then $\int x\sqrt{(2x^2 + 1)}\,dx = \int\sqrt{u} \cdot (1/4)\,du = (1/4)\int u\text{^}(1/2)\,du = (1/4)\cdot(u\text{^}(3/2)/(3/2)) + c = (1/4)\cdot(2/3)u\text{^}(3/2) + c = (1/6)(2x^2 + 1)\text{^}(3/2) + c$.
    Merit: chooses an appropriate substitution and works through multiple steps.

11. The acceleration of a drone rising vertically is $a = 12 - 6t$ m s^{-2} at time t seconds. At $t = 0$ its velocity is 5 m s^{-1} and its height is 0 m. Find the height when the drone's velocity first returns to 5 m s^{-1}.
    Velocity: $v = \int(12 - 6t)\,dt = 12t - 3t^2 + c_1$. At $t = 0$, $v = 5$, so $c_1 = 5$, giving $v = 12t - 3t^2 + 5$. Height: $h = \int v\,dt = 6t^2 - t^3 + 5t + c_2$. At $t = 0$, $h = 0$, so $c_2 = 0$, giving $h = 6t^2 - t^3 + 5t$. Velocity returns to 5 when $12t - 3t^2 + 5 = 5 \to 12t - 3t^2 = 0 \to 3t(4 - t) = 0$, so $t = 4$ (excluding $t = 0$). Height at $t = 4$: $h = 6(16) - 64 + 20 = 96 - 64 + 20 = 52$ m.
    Merit: integrates twice, links acceleration→velocity→displacement with conditions.

12. A curve has gradient function $dy/dx = kx + 1$, where k is a constant. The curve passes through $(0, 3)$ and $(2, 9)$. Find the value of k and the equation of the curve, justifying your working.
    Integrate: $y = \int(kx + 1)\,dx = kx^2/2 + x + c$. Through $(0, 3)$: $3 = 0 + 0 + c$, so $c = 3$. Through $(2, 9)$: $9 = k(4)/2 + 2 + 3 = 2k$

+ 5, so 2k = 4, k = 2. The equation is $y = x^2 + x + 3$. Check: $dy/dx = 2x + 1 = kx + 1$ with k = 2. ✓ Both points satisfy: (0,3) and (2, 4+2+3=9). ✓

Excellence: solves for an unknown constant within an integrated model and justifies with both conditions.

13. Show that ∫ from 0 to a of (2x) dx = ∫ from 0 to a of ($3x^2$) dx has exactly one positive solution for a, and find it.
Left side: $[x^2]$ from 0 to $a = a^2$. Right side: $[x^3]$ from 0 to $a = a^3$. Set equal: $a^2 = a^3 \rightarrow a^3 - a^2 = 0 \rightarrow a^2(a - 1) = 0$, so a = 0 or a = 1. The only positive solution is a = 1. Justification: for $a > 0$, $a^2(a - 1) = 0$ requires a = 1, since $a^2 > 0$; hence exactly one positive value. Both integrals then equal 1.

Excellence: forms and justifies an equation from two definite integrals and argues uniqueness.

14. A population grows so that dP/dt = 0.4P, where P is the number of insects and t is in days. Given that P = 200 when t = 0, show that P = 200e^(0.4t), and find how long it takes the population to reach 1000. (Hint: rewrite as (1/P) dP = 0.4 dt and integrate both sides.)
Separate: (1/P) dP = 0.4 dt. Integrate both sides: ∫(1/P) dP = ∫0.4 dt → ln P = 0.4t + c (P > 0). So P = e^(0.4t + c) = Ae^(0.4t) where A = e^c. At t = 0, P = 200: $200 = A \cdot e^0 = A$, so A = 200. Therefore P = 200e^(0.4t). To reach 1000: 1000 = 200e^(0.4t) → 5 = e^(0.4t) → ln 5 = 0.4t → t = ln 5 / 0.4 ≈ 1.609/0.4 ≈ 4.02 days.

Excellence: solves a separable differential equation from first principles and interprets the result.

Answer key (for parents and tutors)

1. A random poll surveys 1600 voters. Find the margin of error.
   $1/\sqrt{1600} = 1/40 = 0.025$, so about 2.5%.
2. A random poll of 400 people finds 46% support a policy. Form the interval.
   $1/\sqrt{400} = 1/20 = 5\%$. $46 - 5 = 41$ and $46 + 5 = 51$, so 41% to 51%.
3. A cycling club emails its members to ask whether the town wants a new cycleway. Name the main problem with this sample.
   Selection bias: club members are not representative of the town, and are more likely to want a cycleway.
4. Name two things a margin of error does not cover.
   Any two of: bias in how the sample was chosen, non-response, poorly worded questions, people not answering truthfully.
5. A random sample of 900 residents finds 54% support a new library. Form the interval.
   $1/\sqrt{900} = 1/30 = 3.3\%$. $54 - 3.3 = 50.7$ and $54 + 3.3 = 57.3$, so 50.7% to 57.3%.
6. Interpret the interval from question 5.
   I am fairly sure that the proportion of all residents who support a new library is between 50.7% and 57.3%.
7. In the same poll, 54% prefer site A and 38% prefer site B. Find an interval for the difference.
   Same poll: margin $\approx 2/\sqrt{900} = 2/30 = 6.7\%$. Difference $54 - 38 = 16$ points. $16 - 6.7 = 9.3$ and $16 + 6.7 = 22.7$, so 9.3 to 22.7 points.
8. A headline says, "Coffee drinkers live longer", based on a survey of adults' habits. Why can it not show that coffee causes longer life?
   It is observational. Coffee drinkers may differ in other ways, such as income or exercise, and these confounding variables could explain the link. A causal claim needs a randomised experiment.
9. Town A: $n = 625$, 52% support a plan. Town B: $n = 400$, 44% support it. Find an interval for the difference and make a claim.
   Margins: $1/\sqrt{625} = 4\%$ and $1/\sqrt{400} = 5\%$. Average 4.5%, so margin $\approx 1.5 \times 4.5 = 6.75\%$. Difference 8 points: 1.25 to 14.75 points. The whole interval is above zero, so support for the plan is higher in Town A.
10. A student says the Town A and Town B intervals overlap, so the towns cannot differ. Explain the error.
    Overlap is not the right test. The comparison needs one interval for the difference. That interval, 1.25 to 14.75 points, does not include zero, so a difference can be claimed.
11. In a study, 200 students are randomly allocated to use a revision app or to revise as usual. Why does random allocation matter, and why is the second group needed?
    Random allocation balances other variables across the groups, so a difference can be linked to the app. The second group is a comparison: without it, improvement could be due to practice or time rather than the app.
12. A random sample of 1000 finds 4% of people own an electric bike. Can the $1/\sqrt{n}$ rule be used? Explain.
    Not reliably. 4% is outside the range of about 30% to 70%, and the rule would overstate the margin: $1/\sqrt{1000}$ is about 3.2%, which is large compared with 4%.
13. A website poll gets 10 000 votes, with 62% in favour, and reports a margin of error of 1%. Critique this.
    Voters chose to take part, so the sample is self-selected and not random. The 1% margin assumes a random sample and covers sampling variation only. The large sample does not remove bias, so 62% cannot be trusted as an estimate for the population.
14. A random poll of 2500 finds Party A on 48% and Party B on 44%. The report says Party A leads. Evaluate the claim.
    Same poll: margin $\approx 2/\sqrt{2500} = 2/50 = 4\%$. Difference 4 points: 0 to 8 points. Zero is at the edge of the interval, so the data do not show that Party A leads. The report's claim goes further than the evidence.

Answer key (for parents and tutors)

1. A survey of 200 Avondale residents recorded whether they own a bike and whether they cycle to work. 60 own a bike; of those, 45 cycle to work. Find the probability that a randomly chosen bike owner cycles to work.
   This is conditional, restricted to bike owners. P(cycles | owns bike) = 45 ÷ 60 = 0.75.
   Achieved-level fluency: applies the conditional probability formula from counts.

2. For two events, P(A) = 0.4, P(B) = 0.5, and P(A and B) = 0.2. Are A and B independent?
   Check P(A) × P(B) = 0.4 × 0.5 = 0.2. This equals P(A and B) = 0.2, so A and B are independent.
   Key skill: tests independence using the multiplication condition.

3. In a trial, 8% of people taking a new supplement caught a cold, compared with 20% of people not taking it. Calculate the relative risk of catching a cold for supplement-takers compared with non-takers.
   Relative risk = risk (exposed) ÷ risk (baseline) = 0.08 ÷ 0.20 = 0.4. Supplement-takers had 0.4 times the risk (a 60% reduction).
   Key skill: computes relative risk with a stated baseline group.

4. A tree diagram shows P(rain) = 0.3, and P(bus is late | rain) = 0.6, P(bus is late | no rain) = 0.1. Find the probability the bus is late.
   Multiply along branches and add: P(late) = P(rain)×P(late|rain) + P(no rain)×P(late|no rain) = 0.3×0.6 + 0.7×0.1 = 0.18 + 0.07 = 0.25.
   Key skill: combines tree branches using the law of total probability.

5. At Vertex Academy, 150 students were asked about their transport. The two-way table: Drive, Year 12: 20, Year 13: 40; Take bus, Year 12: 50, Year 13: 40. Find the probability that a randomly selected student is a Year 13 student who drives.
   Total = 20 + 40 + 50 + 40 = 150. Year 13 drivers = 40. P(Year 13 and drives) = 40 ÷ 150 = 4/15 ≈ 0.267.
   Achieved: selects and applies a method to read a joint probability from a two-way table.

6. A café's coffee machine fails on 5% of mornings. If it works, 90% of customers are served on time; if it fails, only 30% are served on time. Draw the tree in your head and find the probability a customer is served on time.
   P(on time) = P(works)×0.90 + P(fails)×0.30 = 0.95×0.90 + 0.05×0.30 = 0.855 + 0.015 = 0.87.
   Achieved: applies tree-diagram method with the law of total probability.

7. In a health study, 15 of 300 people who exercise regularly developed high blood pressure, compared with 48 of 400 people who don't. Calculate the relative risk for non-exercisers compared with exercisers.
   Risk (non-exercisers) = 48 ÷ 400 = 0.12. Risk (exercisers) = 15 ÷ 300 = 0.05. Relative risk = 0.12 ÷ 0.05 = 2.4. Non-exercisers were 2.4 times as likely to develop high blood pressure.
   Achieved: calculates relative risk with a clearly identified baseline.

8. The number of texts X a student receives in a lunch break has this distribution: P(X=0)=0.2, P(X=1)=0.3, P(X=2)=0.4, P(X=3)=k. Find k, then find the expected number of texts.
   Probabilities sum to 1: 0.2+0.3+0.4+k = 1, so k = 0.1. E(X) = 0×0.2 + 1×0.3 + 2×0.4 + 3×0.1 = 0 + 0.3 + 0.8 + 0.3 = 1.4 texts.
   Achieved: completes a probability distribution and computes expected value.

9. A factory in Penrose has two machines producing USB drives. Machine A makes 70% of drives with a 3% defect rate; Machine B makes 30% with a 6% defect rate. A drive is found defective. What is the probability it came from Machine B? Show your reasoning.
   P(defect) = 0.70×0.03 + 0.30×0.06 = 0.021 + 0.018 = 0.039. P(B | defect) = P(B and defect) ÷ P(defect) = 0.018 ÷ 0.039 ≈ 0.462. So about 46.2% of defective drives come from Machine B, even though B makes less of the total output, its higher defect rate lifts its share of defects.
   Merit: links tree probabilities to a reverse conditional (Bayes-style) and interprets the result.

10. A study reports the relative risk of a skin condition for people with high sun exposure is 3.0 compared with low exposure. The risk for the low-exposure group is 0.02. A news headline says 'high sun exposure triples your chance of the condition, huge danger!'. Using the numbers, discuss whether the headline is misleading.
    Risk (high exposure) = relative risk × baseline risk = 3.0 × 0.02 = 0.06. So the absolute risk rises from 2% to 6%, a 4 percentage-point increase. While 'triples' (relative risk 3) is technically correct, the actual chance is still only 6%,

so the headline exaggerates by quoting the relative risk without the low absolute risk. A more honest statement links both: the risk tripled but remains fairly small (6 in 100).
Merit: connects relative and absolute risk to critique a claim.

11. Of 500 Auckland commuters surveyed, 180 use a bike-share app. Of app users, 63 had a minor accident last year; of non-users, 64 had one. Determine whether having an accident is independent of using the app, showing the comparison.
P(accident) = (63+64) ÷ 500 = 127 ÷ 500 = 0.254. P(accident | app user) = 63 ÷ 180 = 0.35. Since P(accident | app user) = 0.35 ≠ P(accident) = 0.254, the events are NOT independent, app users have a higher accident rate, so accident and app use are dependent.
Merit: links a two-way table to a formal independence check and interprets it.

12. A medical test for a rare condition affecting 1 in 1000 people has sensitivity 99% (correctly positive if you have it) and specificity 95% (correctly negative if you don't). A person tests positive. Find the probability they actually have the condition, and explain why the answer seems surprisingly low.
Let D = has condition. P(D) = 0.001, P(not D) = 0.999. P(positive) = P(D)×0.99 + P(not D)×0.05 = 0.001×0.99 + 0.999×0.05 = 0.00099 + 0.04995 = 0.05094. P(D | positive) = 0.00099 ÷ 0.05094 ≈ 0.0194, about 1.9%. It is low because the condition is very rare: even a small 5% false-positive rate applied to the huge healthy majority (999 people) produces far more false positives than the true positives from the tiny affected group. The base rate dominates.
Excellence: justifies a counter-intuitive conditional result by reasoning about base rates.

13. Events A and B satisfy P(A) = a and P(B) = b, and are independent. Show that A and the complement of B (B') are also independent, giving a general algebraic argument.
Independence gives P(A and B) = ab. Now P(A and B') = P(A) − P(A and B) = a − ab = a(1 − b). Also P(A)×P(B') = a×(1 − b) = a(1 − b). Since P(A and B') = P(A)×P(B'), A and B' are independent for any a, b. This generalises: if A is independent of B, it is automatically independent of B'.
Excellence: constructs a general algebraic proof of a probability property.

14. A game charges $2 to play. You roll a fair die: rolling a 6 wins $10, rolling a 4 or 5 wins $3, otherwise nothing. Determine whether the game is worth playing in the long run, then find the winning payout the $10 prize would need to become for the game to be exactly fair (expected profit zero), keeping all else the same.
Let W = winnings before the $2 fee. P(W=10)=1/6, P(W=3)=2/6, P(W=0)=3/6. E(W) = 10×1/6 + 3×2/6 + 0 = 10/6 + 6/6 = 16/6 ≈ $2.67. Expected profit = E(W) − 2 = 2.67 − 2 = $0.67 per play, so it is worth playing (positive in the long run). For a fair game, expected profit = 0, so E(W) must equal $2. Let the '6' prize be x: x×1/6 + 3×2/6 = 2, so x/6 + 1 = 2, x/6 = 1, x = $6. Lowering the top prize from $10 to $6 makes the game exactly fair.
Excellence: applies expected value to evaluate a game and solves an inverse problem to generalise fairness.

Answer key (for parents and tutors)

1. X is normally distributed with mean 50 and standard deviation 8. Find P(X < 60).
   Standardise: $z = (60 - 50)/8 = 1.25$. $P(Z < 1.25) = 0.8944$. So $P(X < 60) \approx 0.894$.
   Achieved: standardises and reads a normal probability.

2. X ~ Binomial(n = 10, p = 0.3). Find P(X = 4).
   $P(X = 4) = C(10,4) \times 0.3^4 \times 0.7^6 = 210 \times 0.0081 \times 0.117649 = 0.2001$ (4 dp).
   Achieved: applies the binomial probability formula.

3. The number of texts Aroha gets per hour follows a Poisson distribution with mean 3. Find P(X = 2).
   $P(X = 2) = e^{-3} \times 3^2 / 2! = 0.049787 \times 9 / 2 = 0.2240$.
   Achieved: applies the Poisson probability formula.

4. A bus arrival delay is uniformly distributed on the interval [0, 20] minutes. Find P(5 < X < 12).
   For a continuous uniform distribution, $P(5 < X < 12) = (12 - 5)/(20 - 0) = 7/20 = 0.35$.
   Achieved: uses the uniform distribution to find a probability.

5. The weight of a kererū is normally distributed with mean 650 g and standard deviation 40 g. Find the probability a randomly chosen kererū weighs more than 700 g.
   $z = (700 - 650)/40 = 1.25$. $P(Z > 1.25) = 1 - 0.8944 = 0.1056$. So $P(X > 700) \approx 0.106$.
   Achieved: selects and applies the normal model to an upper tail.

6. A 12-question multiple-choice quiz has four options per question. Manu guesses every answer, so the chance each is correct is 0.25. Find the probability he gets exactly 3 correct.
   $X \sim B(12, 0.25)$. $P(X = 3) = C(12,3) \times 0.25^3 \times 0.75^9 = 220 \times 0.015625 \times 0.075085 = 0.258$.
   Achieved: identifies binomial conditions and computes a probability.

7. A café receives online orders at an average rate of 5 per hour, following a Poisson distribution. Find the probability that fewer than 3 orders arrive in a given hour.
   $\lambda = 5$. $P(X < 3) = P(0) + P(1) + P(2) = e^{-5}(1 + 5 + 12.5) = 0.0067379 \times 18.5 = 0.1247$.
   Achieved: applies the Poisson model to a cumulative 'fewer than' event.

8. Delivery times for a courier follow a triangular distribution with minimum 20 min, mode 30 min and maximum 50 min. Find the probability a delivery takes less than 30 minutes.
   For Triangular(a=20, c=30, b=50), $P(X \le \text{mode}) = (c - a)/(b - a) = (30 - 20)/(50 - 20) = 10/30 = 0.333$.
   Achieved: applies the triangular distribution to find an area up to the mode.

9. A machine fills juice bottles so the volume is normally distributed with standard deviation 5 mL. The manager wants only 2% of bottles to fall below the labelled 500 mL. What mean fill volume should be set?
   We need $P(X < 500) = 0.02$, so the z-value for the 0.02 lower tail is $z = -2.054$. Then $500 = \mu + (-2.054)(5)$, so $\mu = 500 + 10.27 = 510.3$ mL. The machine should be set to about 510.3 mL.
   Merit: works backwards from a tail probability using inverse normal to find a parameter.

10. A factory line produces components that are defective 1.5% of the time, independently. A batch of 200 is checked. Explain why a Poisson model is suitable and find the probability that at most 2 components in the batch are defective.
    With $n = 200$ large and $p = 0.015$ small, the binomial can be approximated by a Poisson with $\lambda = np = 200 \times 0.015 = 3$. $P(X \le 2) = e^{-3}(1 + 3 + 3^2/2) = e^{-3}(1 + 3 + 4.5) = 0.049787 \times 8.5 = 0.423$.
    Merit: links binomial conditions to a Poisson approximation and applies it.

11. Bags of chips have weights that are normally distributed with mean 150 g and standard deviation 4 g. In a shipment of 500 bags, how many would you expect to weigh between 145 g and 155 g?
    z-values: $(145 - 150)/4 = -1.25$ and $(155 - 150)/4 = 1.25$. $P(-1.25 < Z < 1.25) = 0.8944 - 0.1056 = 0.7888$. Expected number = $0.7888 \times 500 = 394.4$, so about 394 bags.
    Merit: links a normal probability to an expected frequency over many trials.

12. A test score X is normally distributed. It is known that 10% of students score below 40 and 5% score above 75. Find the mean and standard deviation of the distribution.
    For the lower 10%, $z = -1.2816$, giving $40 = \mu - 1.2816\sigma$. For the upper 5%, $z = 1.6449$, giving $75 = \mu + 1.6449\sigma$.

Subtracting the equations: 75 − 40 = (1.6449 + 1.2816)σ, so 35 = 2.9265σ, giving σ = 11.96. Then μ = 40 + 1.2816 × 11.96 = 55.3. So μ ≈ 55.3 and σ ≈ 12.0.

Excellence: sets up and solves simultaneous equations from two inverse-normal conditions.

13. Minor earthquakes felt in Auckland occur at an average rate of 0.8 per year, modelled as Poisson. Find the probability that at least 5 are felt over a 5-year period, and state one assumption the model relies on.
Over 5 years, λ = 0.8 × 5 = 4. P(X ≥ 5) = 1 − P(X ≤ 4). P(0)=e^(−4)=0.018316, P(1)=0.073263, P(2)=0.146525, P(3)=0.195367, P(4)=0.195367; sum = 0.628838. So P(X ≥ 5) = 1 − 0.6288 = 0.371. Assumption: earthquakes occur independently at a constant average rate, questionable, since aftershocks cluster, which could make the Poisson model less valid.

Excellence: scales the rate, evaluates a cumulative tail, and critically evaluates a model assumption.

14. A fair coin is tossed 400 times. Using an appropriate distribution, estimate the probability of getting at least 210 heads, and justify why your chosen method is valid.
Exactly, X ~ B(400, 0.5). Since np = 200 and n(1−p) = 200 are both well above 5, a normal approximation is valid: μ = np = 200, σ = √(np(1−p)) = √100 = 10. Apply a continuity correction: P(X ≥ 210) ≈ P(Y ≥ 209.5) where Y ~ N(200, 10^2). z = (209.5 − 200)/10 = 0.95. P(Z > 0.95) = 1 − 0.8289 = 0.1711. So P ≈ 0.171. The approximation is justified because n is large and p is not extreme, so the binomial is close to symmetric and bell-shaped.

Excellence: justifies a normal approximation to the binomial and applies a continuity correction.

Answer key (for parents and tutors)

1. Convert 135° to radians (exact), and convert $5\pi/6$ radians to degrees.
   $135° = 135 \times \pi/180 = 3\pi/4$ radians. $5\pi/6 = 5\pi/6 \times 180/\pi = 5 \times 30 = 150°$.
   Achieved: fluent conversion between degrees and radians.

2. Write down the exact value of $\cos(2\pi/3)$.
   $2\pi/3 = 120°$, which is in the second quadrant. $\cos(120°) = -\cos(60°) = -1/2$.
   Achieved: recalls exact values and quadrant sign.

3. Use a compound-angle formula to find the exact value of $\sin(75°)$.
   $\sin75° = \sin(45° + 30°) = \sin45°\cos30° + \cos45°\sin30° = (\sqrt{2}/2)(\sqrt{3}/2) + (\sqrt{2}/2)(1/2) = (\sqrt{6} + \sqrt{2})/4$.
   Achieved: applies the compound-angle expansion.

4. Simplify $1 - 2\sin^2\theta$ and simplify $2\sin\theta\cos\theta$.
   $1 - 2\sin^2\theta = \cos2\theta$ (double-angle form). $2\sin\theta\cos\theta = \sin2\theta$.
   Achieved: recognises double-angle identities.

5. Solve $2\sin x - 1 = 0$ for $0 \le x \le 2\pi$.
   $2\sin x = 1$, so $\sin x = 1/2$. The reference angle is $\pi/6$. Sine is positive in quadrants 1 and 2, so $x = \pi/6$ and $x = \pi - \pi/6 = 5\pi/6$.
   Achieved: selects and applies a method to solve within a stated interval.

6. Write the general solution (in radians) of $\cos\theta = -1/2$.
   The reference angle is $\pi/3$. Cosine is negative in quadrants 2 and 3, so $\theta = 2\pi/3$ and $\theta = 4\pi/3$ within one revolution. General solution: $\theta = 2\pi/3 + 2n\pi$ or $\theta = 4\pi/3 + 2n\pi$, where n is any integer.
   Achieved: produces a correct general solution using period 2π.

7. A and B are acute angles with $\sin A = 3/5$ and $\cos B = 5/13$. Find the exact value of $\sin(A + B)$.
   Since A is acute, $\cos A = 4/5$. Since B is acute, $\sin B = 12/13$. $\sin(A + B) = \sin A\cos B + \cos A\sin B = (3/5)(5/13) + (4/5)(12/13) = 15/65 + 48/65 = 63/65$.
   Achieved: applies compound-angle formula with found ratios.

8. A Ferris wheel at Ōrākei has radius 25 m and its centre is 30 m above the ground. It completes one revolution every 20 seconds, starting from the lowest point. Its height is $h(t) = 30 - 25\cos(\pi t/10)$ metres. State the amplitude and period, and find the height after 5 seconds.
   Amplitude = 25 m (the coefficient of the cosine). Period $= 2\pi \div (\pi/10) = 20$ seconds. At $t = 5$: $h(5) = 30 - 25\cos(\pi \times 5/10) = 30 - 25\cos(\pi/2) = 30 - 25(0) = 30$ m.
   Achieved: reads model parameters and evaluates a periodic function.

9. Solve $2\cos^2x + \cos x - 1 = 0$ for $0 \le x \le 2\pi$.
   Treat as a quadratic in $\cos x$: $(2\cos x - 1)(\cos x + 1) = 0$. So $\cos x = 1/2$ or $\cos x = -1$. $\cos x = 1/2$ gives $x = \pi/3$ and $x = 5\pi/3$. $\cos x = -1$ gives $x = \pi$. Solutions: $x = \pi/3, \pi, 5\pi/3$.
   Merit: links factorising with solving a trig equation over an interval.

10. Solve $\sin2x = \cos x$ for $0 \le x \le 2\pi$.
    Use the double angle: $2\sin x\cos x = \cos x$. Bring to one side: $2\sin x\cos x - \cos x = 0$, so $\cos x(2\sin x - 1) = 0$. Either $\cos x = 0 \rightarrow x = \pi/2, 3\pi/2$; or $\sin x = 1/2 \rightarrow x = \pi/6, 5\pi/6$. Solutions: $x = \pi/6, \pi/2, 5\pi/6, 3\pi/2$. (Do not divide by $\cos x$ or solutions are lost.)
    Merit: connects double-angle identity to a multi-step solution and avoids losing roots.

11. The depth of water at a Manukau Harbour jetty is modelled by $d = 8 + 3\sin(\pi t/6)$ metres, where t is hours after midnight. A boat needs a depth of at least 9.5 m to enter. During the first 12 hours ($0 \le t \le 12$), for what time interval is the depth at least 9.5 m?
    Require $8 + 3\sin(\pi t/6) \ge 9.5$, so $3\sin(\pi t/6) \ge 1.5$, giving $\sin(\pi t/6) \ge 0.5$. Let $u = \pi t/6$. $\sin u = 0.5$ at $u = \pi/6$ and $5\pi/6$, and $\sin u \ge 0.5$ between these: $\pi/6 \le u \le 5\pi/6$. So $\pi/6 \le \pi t/6 \le 5\pi/6$, giving $1 \le t \le 5$. The depth is at least 9.5 m from 1:00 am to 5:00 am.
    Merit: interprets a model, solves an inequality and translates back to context.

12. Prove the identity $(1 + \cos 2\theta)/\sin 2\theta = \cot\theta$, and state any values of θ for which it is undefined.
    Take the left side and use $\cos 2\theta = 2\cos^2\theta - 1$ and $\sin 2\theta = 2\sin\theta\cos\theta$. LHS $= (1 + 2\cos^2\theta - 1)/(2\sin\theta\cos\theta) = (2\cos^2\theta)/(2\sin\theta\cos\theta) = \cos\theta/\sin\theta = \cot\theta =$ RHS. ■ The identity is undefined where $\sin 2\theta = 0$, i.e. $\theta = n\pi/2$ (any integer n), since that makes the denominator zero (and also where $\sin\theta = 0$ for $\cot\theta$).
    Excellence: constructs a logical proof and states domain restrictions.

13. Express $3\sin x + 4\cos x$ in the form $R\sin(x + \varphi)$ with $R > 0$ and $0 < \varphi < \pi/2$, then find the general solution of $3\sin x + 4\cos x = 2.5$.
    $R\sin(x + \varphi) = R\cos\varphi \sin x + R\sin\varphi \cos x$. Match: $R\cos\varphi = 3$, $R\sin\varphi = 4$. So $R = \sqrt{(3^2 + 4^2)} = 5$ and $\tan\varphi = 4/3$, $\varphi = 0.927$ rad. Thus $3\sin x + 4\cos x = 5\sin(x + 0.927)$. Solve $5\sin(x + 0.927) = 2.5 \rightarrow \sin(x + 0.927) = 0.5$. Reference angle $\pi/6 = 0.524$. So $x + 0.927 = 0.524 + 2n\pi$ or $x + 0.927 = (\pi - 0.524) + 2n\pi = 2.618 + 2n\pi$. Hence $x = -0.403 + 2n\pi$ or $x = 1.691 + 2n\pi$, where n is any integer.
    Excellence: transforms a wave sum and produces a fully justified general solution.

14. Prove that $\sin 3\theta = 3\sin\theta - 4\sin^3\theta$.
    Write $3\theta = 2\theta + \theta$ and expand: $\sin 3\theta = \sin(2\theta + \theta) = \sin 2\theta \cos\theta + \cos 2\theta \sin\theta$. Substitute $\sin 2\theta = 2\sin\theta\cos\theta$ and $\cos 2\theta = 1 - 2\sin^2\theta$: $= (2\sin\theta\cos\theta)\cos\theta + (1 - 2\sin^2\theta)\sin\theta = 2\sin\theta\cos^2\theta + \sin\theta - 2\sin^3\theta$. Now replace $\cos^2\theta = 1 - \sin^2\theta$: $= 2\sin\theta(1 - \sin^2\theta) + \sin\theta - 2\sin^3\theta = 2\sin\theta - 2\sin^3\theta + \sin\theta - 2\sin^3\theta = 3\sin\theta - 4\sin^3\theta$. ■
    Excellence: generalises using compound and double-angle formulae to derive a triple-angle identity.

Answer key (for parents and tutors)

1. A café in Devonport records quarterly hot-chocolate sales (in hundreds of cups): Q1 = 8, Q2 = 5, Q3 = 4, Q4 = 9. Calculate the 4-point moving average for these four quarters.
   4-point moving average = (8 + 5 + 4 + 9) ÷ 4 = 26 ÷ 4 = 6.5 (hundred cups).
   Achieved: applies the moving-average calculation.

2. State what the 'seasonal effect' of a particular quarter tells you in an additive time series model.
   The seasonal effect is the average amount by which that quarter's actual value sits above (+) or below (−) the underlying trend. For example, a seasonal effect of +2.3 means that quarter is typically 2.3 units higher than the trend line for that period.
   Achieved: defines seasonal effect within the additive model.

3. The additive model gives a trend estimate of 42 units for the next winter quarter, and winter has a seasonal effect of −6.5. Forecast the value for that quarter.
   Forecast = Trend + Seasonal effect = 42 + (−6.5) = 35.5 units.
   Achieved: applies Forecast = Trend + Seasonal effect.

4. Explain why we 'centre' a moving average when working with quarterly data.
   A 4-point moving average of quarterly data falls between two quarters (it has no single time point). Centring, averaging two consecutive 4-point means, realigns the smoothed value with an actual quarter, so the trend estimate lines up with the original data points.
   Key skill: understands purpose of the centred moving average.

5. A garden centre in Henderson records quarterly plant sales ($000): 2022 Q1 22, Q2 30, Q3 18, Q4 14; 2023 Q1 26, Q2 34, Q3 22, Q4 18. Describe the trend and the seasonal pattern you can see.
   Trend: sales are generally increasing over time, each 2023 quarter is about $4000 higher than the matching 2022 quarter. Seasonal pattern: Q2 is consistently the highest (30, 34) and Q4 the lowest (14, 18), with Q1 and Q3 in between. This makes sense as spring/summer (Q2) drives higher plant sales.
   Achieved: identifies trend and seasonal features from raw data.

6. Using the garden-centre data above, calculate the first two 4-point moving averages and the first centred moving average.
   First 4-point MA (Q1–Q4 2022) = (22 + 30 + 18 + 14) ÷ 4 = 84 ÷ 4 = 21. Second 4-point MA (Q2 2022 – Q1 2023) = (30 + 18 + 14 + 26) ÷ 4 = 88 ÷ 4 = 22. First CMA (aligned with Q3 2022) = (21 + 22) ÷ 2 = 21.5.
   Achieved: computes 4-point and centred moving averages.

7. A power company's data shows the CMA (trend) for electricity use rising by about 5 units per quarter, reaching 210 units at Q4 2023. The Q1 seasonal effect is +18. Forecast Q1 2024.
   Q1 2024 trend estimate = 210 + 5 = 215 units. Forecast = Trend + Seasonal effect = 215 + 18 = 233 units.
   Achieved: extends trend one step and applies seasonal effect.

8. The seasonal effects for a quarterly series are found to be Q1 = +6, Q2 = +11, Q3 = −4, Q4 = −11. Check whether they need adjusting, and adjust if required.
   Sum = 6 + 11 − 4 − 11 = +2. For an additive model the seasonal effects should sum to 0, so they need adjusting. Mean adjustment = 2 ÷ 4 = 0.5. Subtract 0.5 from each: Q1 = +5.5, Q2 = +10.5, Q3 = −4.5, Q4 = −11.5. These now sum to 0.
   Achieved: applies the seasonal-effect adjustment method.

9. A Ponsonby ice-cream shop's quarterly revenue shows a clear seasonal pattern with Q2 (summer months Jan–Mar) highest. The CMA trend is increasing steadily. Explain how you would produce a forecast for next summer (Q2), and why the additive model is suitable here.
   Method: (1) Fit/extend the CMA trend line to next Q2's time position to get the trend estimate. (2) Take the Q2 seasonal effect (the average of actual − CMA for past Q2s, after adjusting to sum to 0). (3) Forecast = trend estimate + Q2 seasonal effect. The additive model is suitable because the seasonal swings appear to stay roughly constant in absolute size over the years (the summer peak is about the same number of dollars above trend each

year), rather than growing in proportion to the level, which would need a multiplicative model.
Merit: links forecasting steps and justifies model choice from data behaviour.

10. A tourism operator sees quarterly visitor numbers with a strong seasonal pattern, but the summer peaks are getting larger and larger each year as the business grows. Explain why forecasts from a simple additive model with constant seasonal effects could become unreliable, and what this means for the operator.
An additive model assumes each season sits a fixed amount above/below the trend. If summer peaks are growing (the gap above trend widens each year), a constant seasonal effect will underestimate future summers and the error will grow over time. The seasonal variation is proportional to the level, so a multiplicative model would fit better. For the operator, additive forecasts would understate demand in peak seasons, risking under-staffing or under-stocking, so they should treat far-ahead additive forecasts with caution.
Merit: connects model assumptions to reliability and real consequences.

11. For a quarterly dataset, a student calculated seasonal effects but forgot to centre their 4-point moving average. Explain what problem this causes and how it would show up when they compare their trend to the actual data.
An uncentred 4-point moving average lands between two quarters, so it is time-shifted by half a quarter relative to the actual data points. When the student computes actual − trend to find seasonal effects, they'd be subtracting a trend value that doesn't align with the correct time point. This distorts the seasonal effects (systematically over/under-estimating them) and, when plotted, the smoothed trend line would appear offset from where it should sit against the raw data. Centring fixes the alignment so residuals reflect true seasonal deviations.
Merit: links a procedural step to its effect on results across representations.

12. Two analysts model the same quarterly rainfall series for Whangārei. One uses an additive model, the other a multiplicative model (Data = Trend × Seasonal × Irregular). The trend is roughly flat over the period. Argue what you would expect about the two sets of forecasts, and under what condition they would give nearly identical results.
When the trend is flat (roughly constant level), the additive and multiplicative models will give very similar forecasts. In an additive model, seasonal effect = actual − trend; in a multiplicative model, seasonal index = actual ÷ trend. If the trend level L is constant, then a multiplicative index of, say, 1.2 corresponds to an additive effect of 0.2 × L, a fixed amount, so both describe the same fixed seasonal deviation. The models diverge only when the trend changes over time: then the multiplicative model scales the seasonal swing with the level while the additive keeps it constant. So the near-flat trend here means the two forecasts should be almost identical; the choice matters far more for strongly trending series.
Excellence: generalises the relationship between additive and multiplicative models and justifies the condition for equivalence.

13. A student claims: 'If I add up all the CMA (centred moving average) values and compare with the sum of the original data over the same complete years, they should be about equal.' Evaluate whether this claim is justified.
The claim is broadly justified for complete cycles. A centred moving average is a weighted average of surrounding data that smooths out seasonal highs and lows; over one full cycle the positive seasonal deviations cancel the negative ones (they sum to ≈ 0 in an additive model). So summing the CMA over a whole number of years gives approximately the same total as summing the raw data, because the trend represents the average level and the seasonal parts net out. It is 'about' rather than exactly equal because the CMA cannot be computed for the first and last few data points (end effects), so those points are missing from the CMA sum. If we restrict both sums to the same set of time points where the CMA exists, they will be very close, differing only by the residual irregular variation which also tends to cancel.
Excellence: reasons abstractly about the averaging property and identifies the end-effect limitation.

14. A local council forecasts quarterly water demand using an additive model built on 3 years of data. The forecasts extend 8 quarters (2 years) ahead. Discuss, with justification, the factors that limit how much confidence should be placed in the 8th forecast compared with the 1st.
Several factors reduce confidence as forecasts extend further: (1) Trend extrapolation risk, the trend line is fitted to only 3 years; assuming it continues unchanged 2 years ahead ignores possible turning points (e.g. new water restrictions, population change). Small errors in the trend slope compound over 8 quarters, so the 8th forecast could be far off while the 1st is close. (2) Constant-seasonal assumption, the additive model uses fixed seasonal effects from limited data; if climate or usage patterns shift, later forecasts drift further from reality. (3) Irregular/unpredictable events (droughts, policy changes) accumulate uncertainty over time. (4) Only 3 cycles

gives seasonal effects based on just 3 values each, a small, potentially unrepresentative sample. Overall, the 1st forecast relies mostly on the recent, well-supported trend and is more reliable; the 8th depends on the trend and seasonal pattern holding for two more years, which is far less certain, so its confidence interval would be much wider. The council should update the model as new data arrives rather than commit to the far forecasts.
Excellence: justifies how forecast uncertainty grows and evaluates model limitations in context.

Answer key (for parents and tutors)

1. A study of 40 Avondale cafés recorded the daily temperature (°C) and the number of iced coffees sold. The correlation coefficient is r = 0.82. Describe the strength and direction of the linear relationship.
   r = 0.82 is close to +1, so there is a fairly strong, positive linear relationship: as daily temperature increases, the number of iced coffees sold tends to increase.
   Achieved: interprets a correlation coefficient for strength and direction.

2. A linear model for petrol used (litres) against distance driven (km) is: litres = 2.5 + 0.09 × distance. Interpret the slope in context.
   The slope is 0.09. For every extra 1 km driven, the model predicts the petrol used increases by 0.09 litres (about 9 litres per 100 km).
   Achieved: interprets slope as a rate of change in context.

3. Using the model litres = 2.5 + 0.09 × distance, predict the petrol used for a 150 km trip.
   litres = 2.5 + 0.09 × 150 = 2.5 + 13.5 = 16 litres.
   Key skill: substitutes into a linear model to predict.

4. A scatter plot of house size (m²) versus sale price ($) shows points curving upward and getting more spread out for larger houses. Name two features you would describe from this plot.
   Two features: (1) the trend is non-linear (curved/increasing), and (2) the scatter increases for larger houses (non-constant spread / fanning out). Either of these plus mentioning a positive association would be valid.
   Key skill: identifies features of a scatter plot (form and scatter).

5. Mere collects data on the age (years) and resale value ($) of 25 used utes in West Auckland. She fits the model: value = 42000 − 3100 × age. Describe the relationship and predict the value of a 6-year-old ute.
   The negative slope shows a negative relationship: as age increases, resale value decreases. Prediction: value = 42000 − 3100 × 6 = 42000 − 18600 = $23 400.
   Achieved: applies a linear model to describe a relationship and make a prediction.

6. A scatter plot shows the height (cm) and arm span (cm) of 30 students at a Henderson high school. The points lie close to a straight line rising from left to right, with one point far below the others. Describe the trend, direction, strength and any unusual features.
   Trend: linear. Direction: positive (taller students tend to have larger arm span). Strength: strong, because the points lie close to a straight line. Unusual feature: one outlier lies well below the pattern, possibly a measurement error or an unusual individual.
   Achieved: describes trend, association, strength and outliers from a scatter plot.

7. For a data set of weekly rainfall (mm) and weekly grass growth (cm) at a Waikato farm, r = 0.15. What does this tell you about the linear relationship?
   r = 0.15 is close to 0, so there is a very weak (essentially no) linear relationship between weekly rainfall and grass growth in this data. Knowing the rainfall would not help predict grass growth well using a straight line.
   Achieved: interprets a weak correlation coefficient.

8. A model for a car's stopping distance (m) against speed (km/h) is: distance = −4 + 0.85 × speed, fitted from data for speeds between 30 and 100 km/h. Explain why using this model to predict the stopping distance at 5 km/h could be unreliable.
   5 km/h is well outside the data range of 30–100 km/h, so this is extrapolation. The model was only fitted for that range; the relationship may not hold at very low speeds. In fact it predicts distance = −4 + 0.85 × 5 = 0.25 m, which is a very small (and the intercept gives a negative value at 0 km/h), showing the model breaks down outside the data range.
   Achieved: recognises extrapolation and explains reduced reliability.

9. A café owner in Avondale investigates the relationship between the number of staff on shift (2 to 8) and the average customer wait time (minutes). Her model is: wait = 14.2 − 1.6 × staff, and r = −0.88. Using the model and the correlation, discuss what the data suggests, and comment on how much you would trust a prediction of wait time for 5 staff versus for 12 staff.

The slope −1.6 means each extra staff member is associated with about a 1.6-minute decrease in average wait time. r = −0.88 shows a strong negative linear relationship, so the trend is reliable within the data. Prediction for 5 staff: wait = 14.2 − 1.6 × 5 = 6.2 minutes, this is interpolation (within 2–8), and with strong correlation it is fairly trustworthy. Prediction for 12 staff: wait = 14.2 − 1.6 × 12 = −5 minutes, this is extrapolation beyond the data and gives an impossible negative wait time, so it should NOT be trusted; the linear relationship clearly breaks down outside the observed range.

Merit: links slope, correlation and interpolation/extrapolation to judge prediction reliability.

10. Two variables, daily ice-cream sales and daily drowning incidents at NZ beaches, show a strong positive correlation (r = 0.9). A student concludes that buying ice-cream causes drownings. Explain why this reasoning is flawed and suggest a better explanation for the relationship.
Correlation does not imply causation. A strong r only shows the two variables move together, not that one causes the other. A more sensible explanation is a lurking (confounding) variable: hot summer weather. On hot days more people buy ice-cream AND more people swim, so more drownings occur. The temperature drives both variables, creating the correlation without any direct causal link.
Merit: connects correlation to the causation issue and identifies a confounding variable.

11. A researcher compares two possible linear models for the same data on tree age (years) and trunk diameter (cm): Model A has r = 0.72; Model B (after removing one influential outlier) has r = 0.91. Explain what the change in r suggests and discuss whether removing the point is justified.
The rise from r = 0.72 to 0.91 shows the single removed point was an influential outlier weakening the apparent linear relationship; without it the remaining points follow a straight line much more closely. Removing it strengthens the fit. However, removal is only justified if there is a good reason, e.g. a measurement or recording error. If the point is a genuine observation, it should be kept and investigated, because deleting valid data just to improve r is misleading and hides real variation in the population.
Merit: links the effect of an outlier on r to a reasoned judgement about data handling.

12. A regional council models nitrogen level (mg/L) in a stream against the percentage of surrounding land used for dairy farming, based on 50 catchments. The model is nitrogen = 0.8 + 0.12 × dairy%, with r = 0.68. A colleague argues that because r is only moderate, the model is 'useless' for the council's decisions. Evaluate this claim, discussing the value AND limitations of the model for guiding policy.
The claim is too strong. r = 0.68 indicates a moderate positive linear relationship, dairy land percentage explains a meaningful part of the variation in nitrogen ($r^2 \approx 0.46$, so about 46% of the variability), which is genuinely useful for identifying at-risk catchments and informing policy direction. However, limitations must be acknowledged: (1) 54% of variation is unexplained, so other factors (rainfall, soil type, fertiliser practice, riparian planting) matter and the model alone cannot predict any single catchment precisely; (2) correlation is not proof that dairy farming causes the nitrogen, though there is a plausible causal mechanism here; (3) predictions should stay within the observed range of dairy%. Conclusion: the model is a useful screening tool for policy but should guide, not solely determine, decisions, it must be combined with other evidence and treated as showing a tendency rather than an exact rule.
Excellence: justifies a nuanced evaluation using r^2, explained variation, causation and context.

13. For a linear model y = a + bx fitted to data, explain algebraically and conceptually why the fitted line always passes through the point $(\bar{x}, \bar{y})$, the mean of the explanatory variable and the mean of the response. Then explain what this tells you about interpreting the intercept a.
The least-squares line satisfies $\bar{y} = a + b\cdot\bar{x}$, because the line is fitted to minimise the total squared vertical distance to the points, and the balancing point of that fit is the centroid $(\bar{x}, \bar{y})$, the mean of x and mean of y. Rearranging gives $a = \bar{y} - b\cdot\bar{x}$, showing the intercept is fully determined once the slope and the two means are known. Conceptually, the line 'pivots' about the mean point. For the intercept: a is the predicted y when x = 0. This is only meaningful if x = 0 is within or near the data range; often x = 0 is far outside the data (e.g. a house of 0 m^2), so a is a mathematical anchor for the line rather than a realistic prediction. This is exactly why extrapolation to x = 0 can be misleading.
Excellence: generalises and justifies a structural property of the least-squares line and links it to interpretation.

14. Two students investigate the same relationship between weekly study hours and exam mark (%). One uses hours as the explanatory variable and gets slope 3.2; the other swaps the axes, using mark as the explanatory variable, and

gets a different slope. They are confused that $1 \div 3.2 \neq$ their slope. Explain why swapping which variable is explanatory generally produces a different line that is NOT simply the algebraic inverse.

When fitting a linear model, the line is chosen to minimise the squared distances measured VERTICALLY, in the direction of the response variable. So the model of mark-on-hours minimises errors in mark, while the model of hours-on-mark minimises errors in hours; these are two different optimisation problems using different residuals, so they give two different lines. The slope of one is only the reciprocal of the other in the perfect case $r = \pm 1$ (all points exactly on a line). In general the product of the two regression slopes equals r^2, which is less than 1, so $1 \div 3.2$ will not equal the reversed slope. Conclusion: the choice of explanatory variable matters and must match the investigative question, you fit the model to predict the variable you care about, and you cannot simply invert one model to get the other.

Excellence: abstract reasoning about asymmetry of regression, residual direction and the r^2 relationship.

Answer key (for parents and tutors)

1. A bootstrap confidence interval for the difference in median weekly grocery spend (Avondale households − Blockhouse Bay households) is (−$5, $42). Does this interval let you make a call about which population has the higher median spend? Briefly say why.
   No. The interval (−$5, $42) contains 0, so it is plausible that there is no difference in the population medians. Because zero lies inside the interval, you cannot confidently say which population has the higher median spend.
   Achieved: applies the 'contains 0' rule to decide whether a call can be made.

2. Explain what 'resampling with replacement' means when bootstrapping from a sample of 30 commute times.
   You draw 30 values one at a time from the original sample of 30, but after each draw the value is put back so it can be chosen again. This produces a new sample of size 30 in which some original values may appear more than once and others not at all. This models the variability we would see if we took fresh samples from the population.
   Achieved: states the mechanics of bootstrap resampling.

3. A bootstrap distribution of 1000 differences in mean rainfall (mm) is sorted. The 50th value is 2.1 and the 950th value is 18.4. Write down the 90% bootstrap confidence interval.
   The middle 90% runs from the 5th percentile (the 50th of 1000 values) to the 95th percentile (the 950th value). So the 90% CI is (2.1 mm, 18.4 mm).
   Achieved: reads a confidence interval from percentiles of a bootstrap distribution.

4. A CI for the difference in median screen time (Year 13 students − Year 9 students) is (0.4 h, 2.7 h). Write the formal call in context.
   Because the whole interval is above 0, it is a fairly safe bet that the median daily screen time for the population of Year 13 students is between 0.4 and 2.7 hours greater than for the population of Year 9 students.
   Achieved: writes a contextual call from a CI lying entirely above zero.

5. A researcher samples 40 kūmara from a Dargaville farm and 40 from a Kaipara farm, recording weight (g). A bootstrap gives a 90% confidence interval for the difference in median weight (Dargaville − Kaipara) of (12 g, 58 g). Make a formal inference in context.
   The entire interval (12 g, 58 g) is positive, so 0 is not a plausible value for the difference. Therefore it is a fairly safe bet that the median weight of kūmara from the Dargaville farm population is between 12 g and 58 g greater than the median weight of kūmara from the Kaipara farm population.
   Achieved: applies the CI method to make a straightforward population call.

6. From samples of daily bus wait times (minutes) in Onehunga and Ōtāhuhu, a bootstrap CI for the difference in medians (Onehunga − Ōtāhuhu) is (−9.2, −1.5). What can you conclude?
   The whole interval is negative, so it is a fairly safe bet that the median bus wait time for the Onehunga population is between 1.5 and 9.2 minutes LESS than for the Ōtāhuhu population. In other words, Onehunga tends to have shorter waits.
   Achieved: interprets an entirely negative CI as a directional call.

7. Explain why we make an inference about the POPULATION and not just report which SAMPLE median is higher.
   Sample medians vary from sample to sample (sampling variability), so a difference between two sample medians might just be due to chance. A confidence interval accounts for this variability and lets us say what the difference is likely to be back in the populations, which is what we actually care about, not the specific samples we happened to collect.
   Achieved: distinguishes sample statistic from population parameter and references sampling variability.

8. A CI for the difference in mean weekly earnings (part-time workers aged 18–20 − aged 21–24) in a survey is (−$85, −$10). Give the call and state one reason a wider sample might change your confidence.
   The interval is entirely negative, so it is a fairly safe bet that the mean weekly earnings for the population of 18–20 year-old part-time workers is between $10 and $85 less than for the 21–24 group. A larger sample would generally produce a narrower interval, giving a more precise (and possibly more confident) estimate of the difference.
   Achieved: makes a call and links sample size to interval width.

9. Two samples of marathon finish times (minutes) are taken: Auckland runners (n = 60, sample median 248, IQR 42) and Wellington runners (n = 55, sample median 261, IQR 39). A bootstrap 90% CI for the difference in medians (Auckland – Wellington) is (−24, −2). Discuss what this tells you, linking the sample summaries to the inference.
The sample median for Auckland (248) is 13 minutes lower than Wellington (261), suggesting Auckland runners finish faster. The bootstrap CI (−24, −2) is entirely below 0, so it is a fairly safe bet that the median finish time for the Auckland population is between 2 and 24 minutes less than for the Wellington population. The similar IQRs (42 vs 39) show comparable spread, so the median difference is a fair comparison. The fairly large samples (60, 55) mean the estimate is reasonably reliable, though the width of the interval (22 minutes) shows there is still meaningful uncertainty about exactly how big the difference is.
Merit: links sample statistics, spread and sample size to interpret the CI.

10. A student takes two samples from the same population by mistake and bootstraps a CI for the difference in medians, getting (−3.1, 3.4). They conclude 'there is no difference in the populations.' Explain why this wording is not quite right and give a better statement.
The CI contains 0, so we cannot make a call about a difference, but 'no difference' is too strong. We have not proved the population medians are equal; we have only found that 0 is a plausible value for the difference, along with a whole range of small positive and negative values. A better statement: 'Because the confidence interval contains 0, it is plausible there is no difference, so we cannot make a call about which population median is larger.'
Merit: distinguishes 'cannot make a call' from 'proved no difference' and reasons about plausibility.

11. Explain, with reference to how the bootstrap distribution is built, why increasing each sample size from 20 to 80 tends to give a narrower confidence interval.
Each bootstrap resample is drawn from the original sample, so its variability reflects the sample size. With larger samples, each resampled median is based on more data and is closer to the true population median, so the resampled differences cluster more tightly. The bootstrap distribution of differences becomes narrower, and since the CI runs from its 5th to 95th percentile, that interval also narrows. This means a larger sample gives a more precise estimate of the population difference and often lets us make a call when a small sample could not.
Merit: connects sample size to the spread of the bootstrap distribution and hence CI width.

12. A council claims 'water use per household is the same in Titirangi and Glen Eden.' Samples give a 90% bootstrap CI for the difference in median daily water use (Titirangi – Glen Eden) of (18 L, 61 L). A councillor argues this 'proves' the council is wrong. Critique both the council's claim and the councillor's argument, and state what can properly be concluded.
The councillor is closer to correct but overstates it. The CI (18, 61) lies entirely above 0, so 0 is not a plausible value for the difference; this gives strong evidence AGAINST the council's claim of equal median use, so the claim is not supported. However, a confidence interval never 'proves' anything with certainty, it is a 90% procedure, meaning about 10% of such intervals would fail to capture the true difference, and different samples would give slightly different intervals. What can properly be concluded: it is a fairly safe bet that the median daily water use for the Titirangi population is between 18 L and 61 L higher than for Glen Eden, so the evidence contradicts the council's equal-use claim, but this is a probabilistic inference, not proof.
Excellence: critically evaluates competing claims and articulates the probabilistic nature of inference.

13. Two analysts study the same two populations. Analyst A uses samples of size 25 and gets CI (−2, 30) for the difference in medians; Analyst B uses samples of size 200 and gets CI (6, 22). Explain how both results can be consistent, and discuss which supports a stronger inference and why.
Both intervals are consistent because they overlap and both are compatible with a true difference somewhere around the low-to-mid teens. Analyst A's small samples give a wide interval (width 32) that includes 0, so A cannot make a call, the sampling variability is too large to rule out a zero difference. Analyst B's much larger samples give a narrower interval (width 16) sitting entirely above 0, so B CAN confidently infer the first population's median is 6 to 22 higher. B's inference is stronger: the larger sample reduces sampling variability, tightens the bootstrap distribution, and produces a more precise estimate that excludes 0. The two are not contradictory, A simply lacked the precision to detect a difference that B's data reveals.
Excellence: generalises the effect of sample size on precision and reconciles apparently different results.

14. Explain why a bootstrap confidence interval built by resampling can be trusted to reflect population uncertainty, even though we only ever resample from a single original sample and never from the population itself.
    The key idea is that a randomly collected sample is our best available picture of the population, its shape, spread and centre approximate the population's. Resampling with replacement from this sample mimics the process of drawing many fresh samples from a population that looks like our sample. Each resample varies because of which values are (re)selected, and that variation stands in for true sampling variability. By repeating this many times, the spread of resampled differences estimates how much the difference-in-statistics would vary across genuine repeated sampling, which is exactly the uncertainty a confidence interval must capture. The method works well when the sample is reasonably large and representative; if the sample is small or biased, the resamples inherit those flaws, so the interval can be misleading. Thus the bootstrap is trustworthy to the extent the original sample faithfully represents the population.
    Excellence: justifies the logical foundation of bootstrapping and states its assumptions and limitations.

Answer key (for parents and tutors)

1. Solve this triangular system by back-substitution: $x + 2y - z = 3$; $y + 2z = 5$; $z = 2$.
   $z = 2$. Substitute into the second equation: $y + 2(2) = 5$, so $y = 1$. Substitute into the first: $x + 2(1) - 2 = 3$, so $x = 3$. Solution: $(x, y, z) = (3, 1, 2)$.
   Achieved: applies back-substitution to a triangular system.

2. Use elimination to solve: $2x + y - z = 3$; $x - y + 2z = 1$; $x + y + z = 4$.
   From equation 3, $x + y + z = 4$. Add equation 2 and equation 3: $(x + x) + (-y + y) + (2z + z) = 1 + 4 \rightarrow 2x + 3z = 5$... instead eliminate y. Equation 1 + equation 2: $3x + z = 4$. Equation 3 + equation 2: $2x + 3z = 5$. From $3x + z = 4$, $z = 4 - 3x$. Substitute: $2x + 3(4 - 3x) = 5 \rightarrow 2x + 12 - 9x = 5 \rightarrow -7x = -7 \rightarrow x = 1$. Then $z = 4 - 3 = 1$, and from equation 3, $y = 4 - 1 - 1 = 2$. Solution: $(1, 2, 1)$.
   Achieved: carries out Gaussian elimination on a 3×3 system.

3. A system reduces so that its final row becomes: (a) $0 = 0$, (b) $0 = 7$, (c) a leading variable in each of the three rows. State the number of solutions in each case.
   (a) $0 = 0$ is a redundant equation, so there are infinitely many solutions. (b) $0 = 7$ is a contradiction, so there is no solution. (c) A leading variable in every row means each unknown is determined, so there is a unique solution.
   Achieved: classifies solution type from row-echelon form.

4. Write the solution set in parametric form: $x + y + z = 5$; $x + 2y + 3z = 8$; $2x + 3y + 4z = 13$.
   Notice equation 3 = equation 1 + equation 2 ($2x + 3y + 4z = 13$), so it is redundant → infinitely many solutions. Equation 2 − equation 1: $y + 2z = 3$, so $y = 3 - 2z$. From equation 1: $x = 5 - y - z = 5 - (3 - 2z) - z = 2 + z$. Let $z = t$. Solution: $(x, y, z) = (2 + t, 3 - 2t, t)$, for any real t.
   Achieved: expresses an infinite solution set with a parameter.

5. Solve the system: $x + y + z = 100$; $2x + y + 3z = 180$; $x + 2y + z = 150$.
   Equation 3 − equation 1: $y = 50$. Equation 2 − 2×equation 1: $(2x - 2x) + (y - 2y) + (3z - 2z) = 180 - 200 \rightarrow -y + z = -20$, so $z = y - 20 = 30$. From equation 1: $x = 100 - 50 - 30 = 20$. Solution: $(x, y, z) = (20, 50, 30)$. Check in equation 2: $40 + 50 + 90 = 180$ ✓.
   Achieved: selects and applies elimination to reach a unique solution.

6. At a Avondale café, three orders are placed. Order 1: 2 pies, 3 coffees and 1 muffin cost $26. Order 2: 1 pie, 2 coffees and 2 muffins cost $20. Order 3: 3 pies, 1 coffee and 1 muffin cost $20. Find the price of each item.
   Let p, c, m be the prices. $2p + 3c + m = 26$; $p + 2c + 2m = 20$; $3p + c + m = 20$. Equation 1 − equation 3: $-p + 2c = 6$. Equation 1 − 2×(equation 2)? Instead: multiply equation 2 by 2: $2p + 4c + 4m = 40$, subtract equation 1: $c + 3m = 14$. Multiply equation 3 by 2: $6p + 2c + 2m = 40$, subtract 3×equation 2 ($3p + 6c + 6m = 60$): $3p - 4c - 4m = -20$. Simpler: from $-p + 2c = 6$, $p = 2c - 6$. Substitute into equation 3: $3(2c - 6) + c + m = 20 \rightarrow 7c + m = 38$. Substitute p into equation 2: $(2c - 6) + 2c + 2m = 20 \rightarrow 4c + 2m = 26 \rightarrow 2c + m = 13$. Subtract: $(7c + m) - (2c + m) = 38 - 13 \rightarrow 5c = 25 \rightarrow c = 5$. Then $m = 13 - 10 = 3$, and $p = 2(5) - 6 = 4$. Pies $4, coffees $5, muffins $3.
   Achieved: models a straightforward context and solves the system.

7. Determine whether the system has a unique solution, infinitely many, or none: $x + y + z = 4$; $2x + 2y + 2z = 8$; $x - y + z = 2$. If infinite, give the solution parametrically.
   Equation 2 is exactly 2×equation 1, so it is redundant. Two independent equations remain in three unknowns → infinitely many solutions. Equation 1 − equation 3: $2y = 2$, so $y = 1$. Then $x + z = 3$. Let $z = t$: $x = 3 - t$. Solution: $(3 - t, 1, t)$ for any real t.
   Achieved: identifies a dependent system and parametrises it.

8. $10 000 is split between three accounts earning 3%, 4% and 5% per year. The total interest for the year is $370, and the amount in the first account equals the total of the other two. How much is in each account?
   Let a, b, c be the amounts. $a + b + c = 10\,000$; $0.03a + 0.04b + 0.05c = 370$ (×100: $3a + 4b + 5c = 37\,000$); $a = b + c \rightarrow a - b - c = 0$. Add equation 1 and the third: $2a = 10\,000$, $a = 5000$, so $b + c = 5000$. Substitute into $3a + 4b + 5c = 37\,000$: $15\,000 + 4b + 5c = 37\,000 \rightarrow 4b + 5c = 22\,000$. With $b = 5000 - c$: $4(5000 - c) + 5c = 22\,000 \rightarrow 20\,000 + c = 22\,000 \rightarrow c = 2000$, $b = 3000$. Accounts: $5000, $3000, $2000.
   Achieved: translates a financial context into equations and solves.

9. A concert sells adult, student and child tickets at $25, $15 and $10. A total of 500 tickets are sold for total revenue of $9000, and the number of students is twice the number of children. Find how many of each ticket were sold.
Let a, s, c be the numbers. a + s + c = 500; 25a + 15s + 10c = 9000; s = 2c. Substitute s = 2c into the first: a + 3c = 500, so a = 500 − 3c. Substitute into the revenue equation: 25(500 − 3c) + 15(2c) + 10c = 9000 → 12 500 − 75c + 30c + 10c = 9000 → 12 500 − 35c = 9000 → 35c = 3500 → c = 100. Then s = 200 and a = 500 − 300 = 200. Adults 200, students 200, children 100. Check revenue: 5000 + 3000 + 1000 = 9000 ✓.
Merit: links a stated relationship into the system and solves a multi-step problem.

10. A parabola $y = ax^2 + bx + c$ passes through the points (1, 6), (2, 11) and (3, 18). Find its equation, connecting the points to a system of equations.
Each point gives one equation: a + b + c = 6; 4a + 2b + c = 11; 9a + 3b + c = 18. Equation 2 − equation 1: 3a + b = 5. Equation 3 − equation 2: 5a + b = 7. Subtract: 2a = 2, so a = 1. Then b = 5 − 3 = 2, and c = 6 − 1 − 2 = 3. The parabola is $y = x^2 + 2x + 3$. Check (3, 18): 9 + 6 + 3 = 18 ✓.
Merit: connects a geometric representation (points on a curve) to a solved system.

11. Reduce and interpret: x + 2y − z = 3; 2x − y + 3z = 4; x − 3y + 4z = 1. State the solution type and give the general solution.
Equation 2 − 2×equation 1: (2x − 2x) + (−y − 4y) + (3z + 2z) = 4 − 6 → −5y + 5z = −2. Equation 3 − equation 1: (x − x) + (−3y − 2y) + (4z + z) = 1 − 3 → −5y + 5z = −2. The two reduced equations are identical, so one equation is redundant → infinitely many solutions (the planes meet in a line). Let z = t: from −5y + 5z = −2, y = t + 2/5. From equation 1: x = 3 − 2y + z = 3 − 2(t + 2/5) + t = 11/5 − t. General solution: (11/5 − t, 2/5 + t, t).
Merit: reduces the system and interprets the redundancy geometrically as a line.

12. For the system x + y + z = 6; x + 2y + 3z = 10; x + 2y + kz = m, find the value(s) of k and m for which the system has (i) a unique solution, (ii) infinitely many solutions, (iii) no solution. Justify your answer.
Equation 2 − equation 1: y + 2z = 4. Equation 3 − equation 1: y + (k − 1)z = m − 6. Subtract these two: (k − 1 − 2)z = (m − 6) − 4 → (k − 3)z = m − 10. (i) If k ≠ 3, we can divide by (k − 3) to find z, then y, then x, a unique solution for any m. (ii) If k = 3, the last row becomes 0 = m − 10. This is consistent only when m = 10, giving 0 = 0, a redundant row → infinitely many solutions. (iii) If k = 3 and m ≠ 10, the last row is 0 = (non-zero), a contradiction → no solution.
Excellence: generalises solution behaviour in terms of two parameters with justification.

13. Investigate the system ax + y + z = 1; x + ay + z = 1; x + y + az = 1. Determine all values of a giving a unique solution, infinitely many, or none, and justify.
Equation 1 − equation 2: (a − 1)x − (a − 1)y = 0 → (a − 1)(x − y) = 0. Equation 2 − equation 3: (a − 1)(y − z) = 0. Case a ≠ 1: then x = y = z. Substitute into any equation: (a + 2)x = 1. If a ≠ −2, x = 1/(a + 2) → unique solution x = y = z = 1/(a + 2). If a = −2, this becomes 0 = 1, a contradiction → no solution. Case a = 1: all three equations become x + y + z = 1 (identical), so two are redundant → infinitely many solutions (a whole plane). Summary: unique for a ≠ 1 and a ≠ −2; no solution for a = −2; infinitely many for a = 1.
Excellence: analyses a symmetric parametric system, proving each solution case.

14. Balancing the combustion reaction $a\cdot C_3H_8 + b\cdot O_2 \rightarrow c\cdot CO_2 + d\cdot H_2O$ leads to a system for a, b, c, d. Set up the equations, show it has infinitely many solutions, and explain how a valid balanced equation is chosen.
Match atoms on each side. Carbon: 3a = c. Hydrogen: 8a = 2d → d = 4a. Oxygen: 2b = 2c + d = 2(3a) + 4a = 10a → b = 5a. So c, d, b are all fixed multiples of a, but a itself is free, this is a homogeneous-type system with one free parameter, hence infinitely many solutions (any scaling works). A valid chemical equation needs the smallest positive whole numbers, so choose the smallest a making all values integers: a = 1 gives b = 5, c = 3, d = 4. The balanced equation is $C_3H_8 + 5O_2 \rightarrow 3CO_2 + 4H_2O$. The infinite solutions correspond to scaling the whole equation (e.g. a = 2 doubles everything).
Excellence: models an unfamiliar context, recognises the infinite-solution structure and justifies the chosen solution.