Standard form. Very small and very large quantities are tedious (and error prone) to write out using full decimal notation. Standard form: e.g. 6.67×10^{-11} is an integer between 1 and 9 followed by $N-1$ digits, where N is the number of significant figures of the quantity. The power of 10 (the 'exponent') gives you an immediate sense of scale.

Precision vs accuracy. A *precise* measurement is performed to a high number of significant figures. This means the random error in the measurement (i.e. the standard deviation) is very small compared to the mean value. In calculations, one should quote a answer to the worst precision (i.e. lowest number of significant figures) of the input values. *Accuracy* relates to the degree of *systematic error.* A time of 12.345s may be very precise, but could easily be 2.000s out from a true value of 10.345s if there is some form of accidental offset in the timing system.

Errors. All measurable quantities will be subject to *uncertainty*. If quantities $x, y...$ are within a known range, we can use upper and lower bounds to determine the range of combined quantities. e.g. $x_- \le x \le x_+$, $y_- \le y \le y_+$
Therefore: $x_-^{\,2} y_- \le x^2 y \le x_+^{\,2} y_+$ and $x_-^{\,2}/y_+ < x^2/y < x_+^{\,2}/y_-$. Note the mixing of bounds in the last example.
If errors are *normally distributed*, the **'Law of Errors'** can be useful (although may result in an artificially tighter uncertainty than upper and lower bounds). Let $f(x,y,z..)$ be a function of measureable quantities e.g. $x=\bar{x}\pm\sigma_x$.

$$f=\bar{f}\pm\sigma_f \text{ where } \bar{f}=f(\bar{x},\bar{y},\bar{z}...) \quad : \qquad \sigma_f^2=\left(\frac{\partial f}{\partial x}\sigma_x\right)^2+\left(\frac{\partial f}{\partial y}\sigma_y\right)^2+\left(\frac{\partial f}{\partial z}\sigma_z\right)^2+...$$

If $f(x,y....)=kx^a y^b... \Rightarrow \left(\frac{\sigma_f}{f}\right)^2=\left(\frac{a\sigma_x}{x}\right)^2+\left(\frac{b\sigma_y}{y}\right)^2+...$ You *add* the (power weighted) squares of fractional errors.

If a quantity x is subject to random error and N independent measurements $\{x_i\}$ are made, the *unbiased estimate* of the mean value of x is: $\bar{x}=\frac{1}{N}\sum_{i=1}^{N}x_i$. Since the mean value is used in the calculation of the *standard deviation*, the unbiased estimate of the standard deviation in x (i.e. the 'error' in x) is: $\sigma_x=\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(x_i-\bar{x}\right)^2}$. We can quote: $x=\bar{x}\pm\sigma_x$

Units and dimensional analysis. *Equations* relate quantities in the physical sciences. For an equation to be valid it must be dimensionally consistent as well as numerically. In other words if the left side of the equation is a mass (in kg), then the combination of quantities on the right hand side of the equation must also yield kg. This idea can be used to guess the form of various equations. e.g. the period of a pendulum T may depend upon pendulum length l , the strength of gravity g and pendulum bob mass M . Therefore assert $T=kl^a g^b M^c$. Noting dimensions: $[T]=s$
$\therefore \mathrm{s}=[k]\times\mathrm{m}^a\times\mathrm{m}^b\times\mathrm{s}^{-2b}\times\mathrm{kg}^c$. If constant k is assumed to be dimensionless, by considering the powers of SI (International System or Metric) units s, m, kg : $T=k\sqrt{l/g}$. The constant k is actually 2π . (See the *Simple Harmonic Motion* sheet).

For many equations it is useful to write them in dimension scaled, i.e. '*dimensionless'* form. This can greatly aid comprehension, and 'ready-reckoning' quick-calculation usage. For example *Kepler's Third law* of orbital motion about a single massive star of mass M can be written as: $(T/\mathrm{Yr})^2=\left(M_\odot/M\right)\left(a/\mathrm{AU}\right)^3$ where Yr is an earth year ($365\times24\times3600\mathrm{s}$), $M_\odot$ is the mass of the Sun and $\mathrm{AU}=1.496\times10^{11}\mathrm{m}$ ('Astronomical Unit') is the average radius of the Earth's orbit about the Sun. T is the orbital period and a is the semi-major axis of the orbit ('radius' if a circular orbit).

Scaling equations often results in various important *dimensionless groups*. These pure numbers characterize a purely mathematical relationship, and can be very useful in building numerical models - particularly when the equations themselves cannot be solved easily. e.g. Optics the *refractive index* n is the speed of light in a vacuum / speed of light in a medium, in fluid dynamics the *Reynold's number* (Re) is the fluid density $\times$ speed $\times$ characteristic length / dynamic viscosity. Once the Reynold's number exceeds a certain range of values, any fluid will become turbulent rather than laminar ('smoothly flowing'). In other words, dimensionless scaling can tease out universal behaviour of physical laws, as encoded by their mathematical relationships.

Linear regression

Perhaps the most important analytical tool in the physical sciences is the ability to quantify the validity of a model relating a set of measurable parameters. The idea is as follows:

- Rearrange the model in such a way that it becomes a linear equation of the form $y = mx + c$ or $y = mx$ if *direct proportion* is assumed. e.g. 'potential difference is proportional to current in a fixed resistor.'
- Plot experimental (x, y) data on a graph and determine the line of best fit through the data.
- Determine gradient m and vertical intercept c from the line of best fit.
- Determine the standard deviations ('errors') of both gradient m and intercept c, and a quantitative measure of how good the fit is (this is called the product moment correlation coefficient r).

The concept is to find m, c *minimize* a sum of squared deviations $S = \sum_{i=1}^{N}(y_i - mx_i - c)^2$ or $S = \sum_{i=1}^{N}(y_i - mx_i)^2$.

The recipe for linear regression (which can readily be inputted into a spreadsheet or computer program) is:

$$\bar{x} = \tfrac{1}{N}\sum_{i=1}^{N} x_i, \quad \bar{y} = \tfrac{1}{N}\sum_{i=1}^{N} y_i, \quad \overline{x^2} = \tfrac{1}{N}\sum_{i=1}^{N} x_i^2, \quad \overline{y^2} = \tfrac{1}{N}\sum_{i=1}^{N} y_i^2, \quad \overline{xy} = \tfrac{1}{N}\sum_{i=1}^{N} x_i y_i$$

$$V[x] = \overline{x^2} - \bar{x}^2, \qquad V[y] = \overline{y^2} - \bar{y}, \qquad \mathrm{cov}[x,y] = \overline{xy} - \bar{x}\bar{y}$$

For a $y = mx + c$ line of best fit:

$$m = \frac{\mathrm{cov}[x,y]}{V[x]}, \quad c = \bar{y} - m\bar{x}, \quad s = \sqrt{\frac{1}{N-2}\sum_{i=1}^{N}(y_i - mx_i - c)^2}, \quad \Delta m = \frac{s}{\sqrt{N}}\frac{1}{\sqrt{V[x]}}, \quad \Delta c = \frac{s}{\sqrt{N}}\sqrt{1 + \frac{\bar{x}^2}{V[x]}}$$

For a $y = mx$ line of best fit: $$m = \frac{\overline{xy}}{\overline{x^2}}, \quad s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(y_i - mx_i)^2}, \quad \Delta m = \frac{s}{\sqrt{N}}\frac{1}{\sqrt{V[x]}}$$

The *product moment correlation coefficient* is: $r = \dfrac{\mathrm{cov}[x,y]}{\sqrt{V[x]V[y]}}$. This is +1 for a perfect *positive* gradient line of best fit, and -1 for a perfect negative fit. It is zero for no discernible correlation.

Question 1

(i) In Greek mythology, King Sisyphus was doomed to push a large boulder endlessly up a hill. Unfortunately he would never reach the top, and the boulder would always roll back down again. Sisyphus (who was *outrageously* punished for outsmarting the somewhat prideful God Zeus) decides to time the rolls to make him feel better. His measurements are: 7.52s, 7.86s, 7.15s, 7.33, 7.44s. Determine the time t in the form $t = \bar{t} \pm \sigma_t$.

(ii) Sam trains for the Senior cross-country by running the route several times. He always gets a little lost and also forgets to bring his watch. He reckons he runs distance x in the range $9.8\mathrm{km} \le x \le 10.3\mathrm{km}$ in time t which is in the range $39\,\mathrm{min} \le t \le 43\,\mathrm{min}$. Determine the range of possible speeds v in (a) m/s and (b) km/h.

(iii) A beta particle (a very fast moving electron) has an energy of between 5.0 and 9.8 MeV. An *electron-volt* $1\mathrm{eV} = 1.602 \times 10^{-19}\,\mathrm{J}$. If the beta particle has mass $m = 9.109 \times 10^{-31}\,\mathrm{kg}$, determine the range of speeds of the electron (to appropriate precision) as a fraction of the speed of light using both the Classical formula for kinetic energy $E = \tfrac{1}{2}mv^2$ and the Relativistic formula $E = \left(1 - v^2/c^2\right)^{-\frac{1}{2}} mc^2 - mc^2$. Speed of light $c = 2.998 \times 10^8\,\mathrm{ms}^{-1}$.

(iv) The gravitational force between two masses m , M separated by distance r is $F = GMm/r^2$. Given 'force is mass $\times$ acceleration' , determine the SI units (kg,m,s) of constant G .

(v) The speed of sound in a gas of density ρ and pressure p is given by the formula $c = k\rho^a p^b$ where k is a dimensionless constant. Use dimensional analysis to find powers a, b .

(vi) The power P of a wind turbine is given by $P = kr^2v^3$ where r is the blade radius and v is the wind speed. In the county of Windyshire, blades vary by $\pm 10\%$ in radius, and wind speed vary by $\pm 30\%$ from their respective mean values. What is the expected % error in power? [Perform using upper and lower bounds, and then assume normally distributed errors. Compare the results.

(vii) Homer is measuring the resistance R of light bulbs in the Springfield Nuclear Power Station given by the equation $P = V^2/R$ where P is the electrical power transformed and V is the potential difference across them. His measurements are: 60.2W, 59.4W, 60.5W, 59.8W for power and 110.1V, 110.4V, 109.8V, 109.9V for potential difference. Showing all calculations, determine the mean and uncertainty in resistance (in Ω).

(viii) A set of experimental data for quantities x and y is given below:

$x = \{-0.20, 0.63, 2.86, 3.96, 2.71, 5.50, 6.34, 9.03, 11.08, 12.22\}$

$y = \{-1.70, 0.50, -12.65, -15.44, -6.18, -6.82, -20.86, -23.24, -32.35, -29.07\}$

Construct a spreadsheet and input these numbers as columns. Work out x^2, y^2, xy columns and hence determine $\bar{x}, \bar{y}, \overline{x^2}, \overline{y^2}, \overline{xy}, V[x], V[y], \text{cov}[x,y]$. Hence determine m, c for a $y = mx + c$ line of best fit. Determine the product moment correlation coefficient r and also the errors Δm and Δc in the gradient and intercept.

Plot the (x, y) as crosses and overlay the (mean) line of best fit, and also the 'envelope' of lines based upon the values calculated for Δm and Δc.

Question 2 A student is investigating *Kepler's Third Law* $(T/\text{Yr})^2 \approx (a/\text{AU})^3$ for the planets + Pluto in the solar system. The data is as follows: (Note it is deliberately imprecise! We known a and T to many more significant figures).

Planet	**Distance from sun *a* in AU**	**Orbital period *T* /years**
Mercury	0.4	0.2
Venus	0.7	0.6
Earth	1.0	1.0
Mars	1.5	1.9
Jupiter	5.2	11.9
Saturn	9.8	29.6
Uranus	19.3	84.7
Neptune	30.2	166
Pluto	39.5	248

(a) Type the data into a spreadsheet and plot (a, T) as crosses (+) on a graph. *Don't join them with lines*, as we predict an underlying model. Make a separate column of a values between 0 and 40 with a step size of 0.1 or smaller. Determine the corresponding T values. Plot this as a curved line (no markers) on the same graph as the data. Ideally this model curve should *underlay* the data points.

(b) Determine an appropriate linearization of Kepler III in the form $y = mx$. Work out x^2, y^2, xy columns and hence determine $\bar{x}, \bar{y}, \overline{x^2}, \overline{y^2}, \overline{xy}, V[x], V[y], \text{cov}[x,y]$. Hence determine m for a $y = mx$ line of best fit. Determine the product moment correlation coefficient r and also the gradient error Δm. Overlay the (mean) line of best fit, and also the 'envelope' of lines of gradients $m \pm \Delta m$.

(c) *Bode's Law* is an approximate 'numerological' relationship between the planet number n in the Solar System and the orbital radius from the Sun a: $10 \times (a/\text{AU}) \approx 4 + 3 \times 2^n$. Estimate (or if you can use logarithms, calculate) n for each planet. Is there anything wrong with the pattern? Note the *asteroid belt* between Mars and Jupiter is at (approximately) $a \approx 2.1\text{AU}$, although is much more dispersed than a fixed radius.

Question 3 The charge to mass ratio e/m_e of an electron can be determined by measuring the radius r of a circular beam of ionized hydrogen gas in a Fine Beam Tube. Other inputs are the accelerating potential V and magnetic field strength B. The relationship is: . $e/m_e = 2V/B^2r^2$. $V = (170 \pm 5)\text{V}$, $B = (8.0 \pm 0.2) \times 10^{-4}\text{T}$ and $r = (5.50 \pm 0.30)\text{cm}$. (a) Determine e/m_e in the form $k_- < e/m_e < k_+$ where $k_\pm$ are the upper and lower bounds.

(b) Assuming all the errors are normally distributed, determine $e/m_e = \bar{k} \pm \sigma_k$.

(c) $e = 1.602 \times 10^{-19}\text{C}$ and $m_e = 9.109 \times 10^{-31}\text{kg}$. Is this result within the error range of the experiment?

Question 4 An air-filled syringe is slowly compressed via a screw system, such that the heat transfer between the air and the surroundings maintains thermal equilibrium (i.e. the same temperature). In this situation one might expect the relationship between air pressure and volume to follow Boyle's Law i.e. $pV = \text{constant}$. The syringe is cylindrical and has a marked scale, so volume can be measured. The syringe also has a pressure gauge. The following data is recorded:

Pressure $/10^5\,\text{Pa}$	**Volume** $/10^{-4}\,\text{m}^3$
3.25	0.98
2.78	1.18
2.28	1.37
2.00	1.57
1.78	1.77
1.57	1.96
1.45	2.16
1.33	2.36
1.23	2.55
1.13	2.75
1.07	2.95
0.99	3.14
0.93	3.34
0.88	3.53
0.85	3.73
0.80	3.93
0.76	4.12
0.72	4.32
0.69	4.52
0.66	4.71
0.60	4.91

The *Ideal Gas Equation* is $pV = nRT$ where n is the number of moles of gas, T is the absolute temperature in Kelvin (K) and the *Molar Gas Constant* $R = 8.314\text{Jmol}^{-1}\text{K}^{-1}$. The fixed gas temperature when the syringe was compressed was 294K.

$\log_{10} p = \log_{10}(nRT) - \log_{10} V$, so one predicts a graph of $y = \log_{10} p$ vs $x = \log_{10} V$ will yield a straight line with gradient -1 and y intercept $\log_{10}(nRT)$

(a) Input the recorded data into a spreadsheet and determine extra columns for $x = \log_{10} V$ and $y = \log_{10} p$.

(b) Perform a *linear regression* and determine $m \pm \Delta m$ and $c \pm \Delta c$ for the gradient and intercept of a line of best fit $y = mx + c$. (Do this from scratch and don't cheat by using the built-in Excel 'trendline' function! However, you could use this *afterwards* to check your calculations are correct)

(c) Underlay the line of best fit to the (x, y) data and calculate the product-moment correlation coefficient r. Does Boyle's Law correlate with the data?

(d) Determine from the y intercept the number of moles of gas in the syringe n.

(e) Hence determine a model (p,V) curve and plot this as a smooth underlay to the data points (as crosses).

Question 5 Protactinium-234 decays to Uranium 234 via beta-minus emission. The activity A of counts per second (Becquerel or Bq) decays according to the equation $A = A_0 \times 2^{-t/T}$ where t is the time in seconds and T is the *half-life* of Pa-234. Times and activities for Pa-234 decay are recorded in the table below.

Use the data to determine the half-life of Pa-234, and the uncertainty in this measurement. Do the same for the initial activity of the source A_0.

Time /s	**Activity /Bq**
0	29.9
13	28.2
26	24.9
39	18.6
52	21.3
65	17.3
78	14.3
91	15.5
104	12.3
117	7.8
130	10.6
143	8.4
156	8.9
169	7.3
182	6.2
195	4.8
208	4.4
221	3.0
234	4.8

Construct a spreadsheet for this process. Structure this and plot suitable graphs as per the norms hopefully established in questions 1 (viii), 2 and 4.

The fractional activities are due to an (integer!) count in 10s being recorded, and then adjusted for background radiation. The background count was 182 in 100s.

Note the 'official' value for the half-life of Pa-234 is 70.2s. How does this compare to the half life (in the form $T = \bar{T} \pm \sigma_T$) determined from the data?