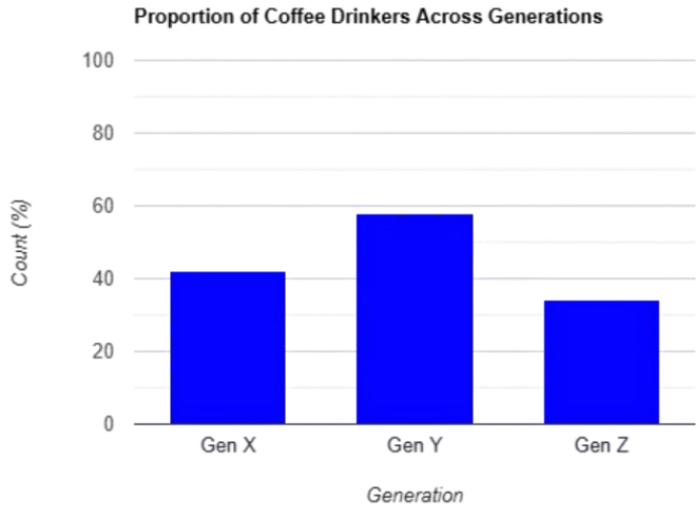
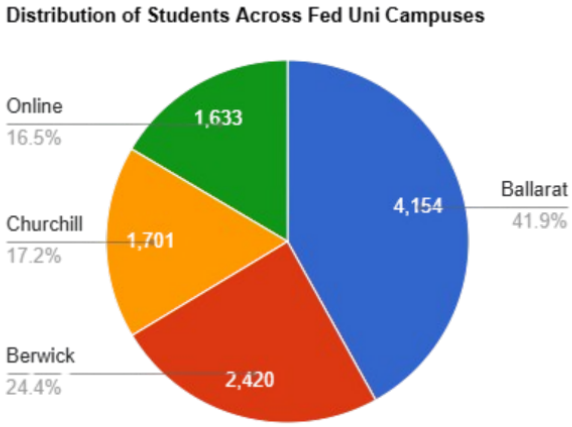
Interpreting Graphs & Charts

Understanding Graph Types

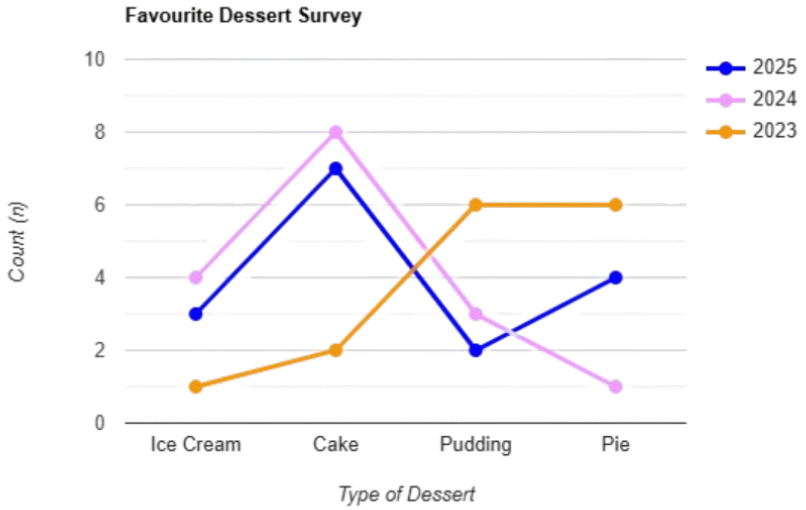
The common types of graphs you'll encounter in your studies are typically the following:



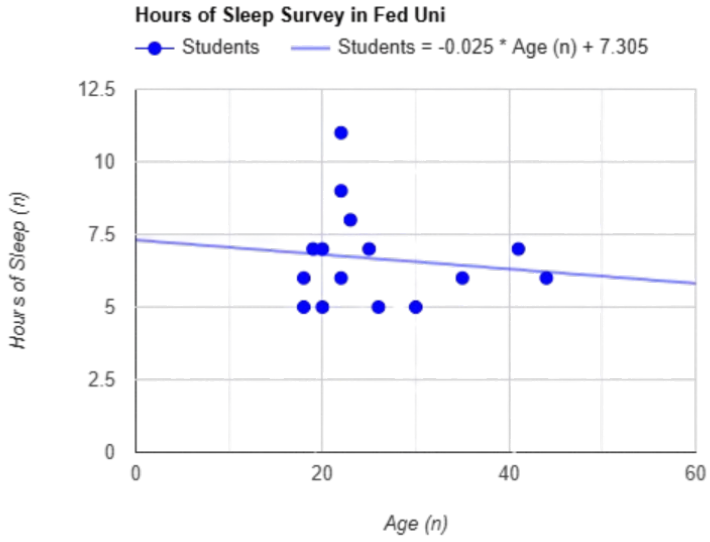
- **Bar Charts:** Useful in comparing quantities across multiple categories.
    - *What to look for:* Differences in the height of each bar which will highlight comparisons.



- **Pie Charts:** More common in conveying proportions or percentages of something (e.g. funding of government projects, demographics of a population)
    - *What to look for:* The proportion of each 'slice' within the pie chart.



- **Line Graphs:** Helpful in showing how a variable changes across specific conditions.
    - *What to look for:* Consistency or changes in the data.



- **Scatter Plots:** Shows the relationship between two variables using multiple data points. Used to explore possible patterns in datasets.
    - *What to look for:* 'Clustering' within the dataset or outliers that may need to be omitted from the data.

Analysing the Contents of a Graph or Chart

Here are some considerations to help thoroughly analyse a graph or chart:

- **Title & Axes:** Found above or below the data, it summarises the purpose of the graph/chart (i.e. what the data represents). Most graphs have two axes:
    - the vertical or "y-axis" which looks at the dependent variable (the one being changed).
    - the horizontal or "x-axis" which usually provides data on the independent variable (the one not changing).

It's important to look at what the label of each axis is as it provides information on what is being measured (time, dose, concentration, etc.). It's also helpful to look at the units being used in each axis (mins, mg/L, degrees Celsius) as this helps give more context to the data being presented.

- **Legend:** Indicates any symbols and/or colours that represents different datasets and categories. Legends are a great reference point to stop assumptions being made about the data which can lead to drawing wrong conclusions.

  TIP: Always cross-reference the legend with the data and axis to ensure proper interpretation.

- **Scale:** Usually denoted by smaller lines within each axes. This is important to consider if you're looking at a linear graph (the spacing of the scale is consistent) but especially if you're looking at **logarithmic** (which normally compresses larger values and expands smaller ones) graphs which spans multiple orders of magnitude.

Subject-specific examples

Biochemistry

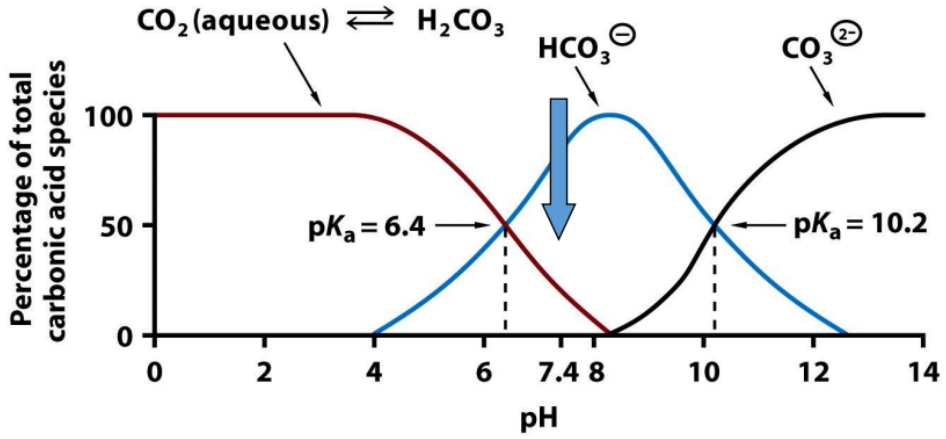
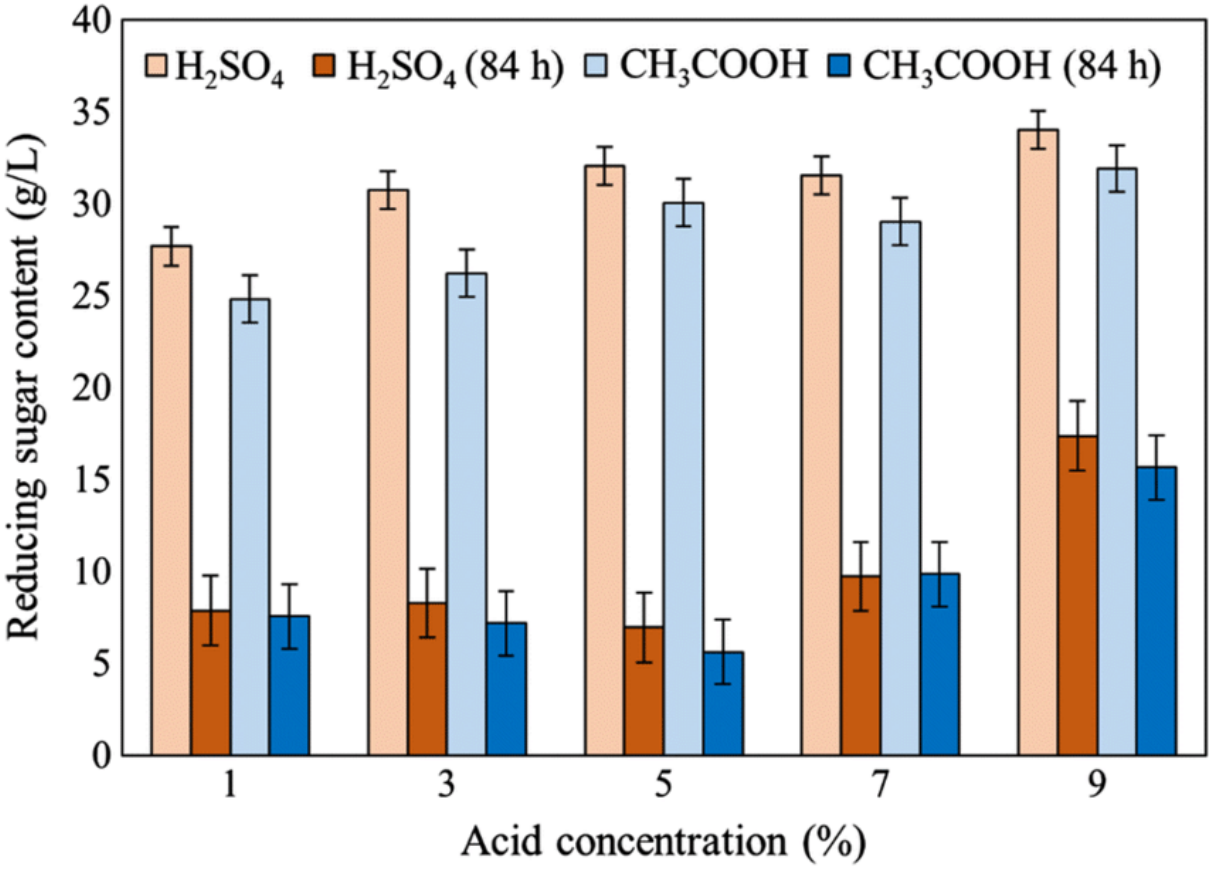


Figure 2-19 Principles of Biochemistry, 4/e
© 2006 Pearson Prentice Hall, Inc.

While lacking a title, the graph above demonstrates how the percentage of total carbonic acid species (y-axis) changes as the pH (x-axis) increases. The H_2CO_3 species (brown line) remains in equilibrium with CO_2 until the pH reaches 4. At this point, equilibrium is disrupted and the formation of HCO_3^- (conjugate base in this example) slowly increases as H_2CO_3 continues to deprotonate (donate H^+). At pH 6.4, notice how the pKa also equals this value (indicating the concentration of the conjugate base and acid are equal at 50% each). This continues until all carbonic acid species exists as HCO_3^- at approx. pH 8.2. As pH continues to increase, HCO_3^- is deprotonated (donating its protons to the now basic solvent) and its concentration decreases whilst the CO_3^{2-} species increases, reaching equal concentrations at pH 10.2, and being the sole carbonic acid species present at approx. pH 12.4.



Caption: Difference between reducing sugar concentration (g/L) of Chlorella after pre-treatment with H2SO4 and CH3COOH and after 84 h fermentation process

Ref: Effects of acids pre-treatment on the microbial fermentation process for bioethanol production from microalgae - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Difference-between-reducing-sugar-concentration-g-L-of-Chlorella-after-pre-treatment_fig4_334810098 [accessed 4 Dec 2024]

In this bar graph, we can see the amount of reduced sugar content, with its units (g/L) along the y-axis whilst 5 different samples with varying acid concentration percentages (1,3,5,7,9) can be found along the x-axis. The title (caption in this example) below the graph provides more context of a microalgae (Chlorella) by explaining the legend (located in the top of the graph) as the chemicals used as a pre/post-treatment to reduce sugar content. In terms of trends, it can be generally observed that there's a positive relationship between acid concentration and reducing sugar content. Furthermore, the highest acid concentration sample (9%) has reduced the most sugar across all the variables introduced. However, we can see that the acid concentration at 5% doesn't follow the general trend in regard to the post-treatment (orange and dark-blue bar graph) as well as the pre-treatment results for the 7% acid concentration of Chlorella. Because we aren't given further information, this is as much of a conclusion as we can draw from this graph.

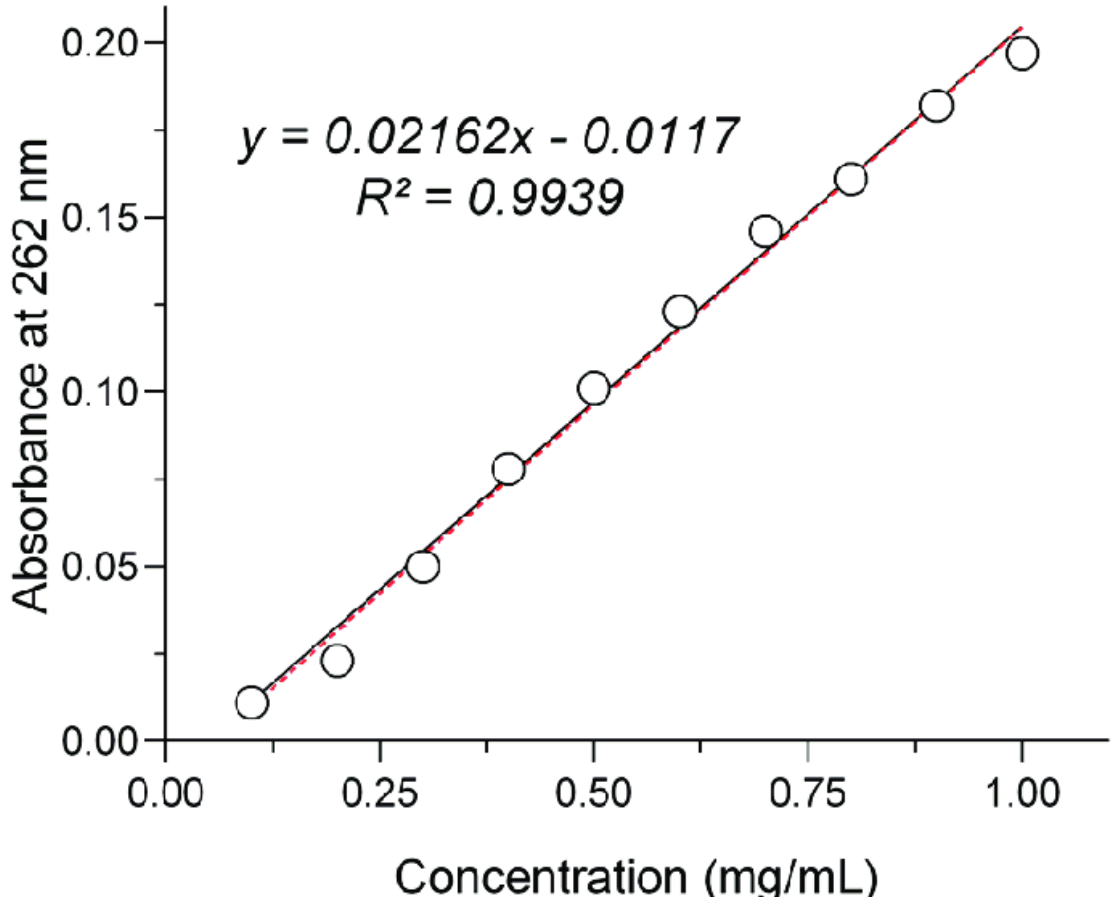


Figure 1. Calibration curve of naproxen reference standard.

Ref: Assessment of Drug Release Kinetics and Quality of Naproxen Generic Tablets in Bangladesh - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Calibration-curve-of-naproxen-reference-standard_fig1_352020271 [accessed 5 Dec 2024]

Linear calibration curves, like the one above, are similar to a line graph except a line of best fit is applied to the data points. The caption/title below this graph provides context as to the compound that is being investigated (naproxen ($C_{14}H_{14}O_3$)). The calibration curve is plotted on a graph in which the concentrations (in this case, mg/mL) of **known standards** are on the x-axis whilst absorbance or intensity at a given wavelength (arbitrary units) is normally found on the y-axis of the graph. From this line of best fit, we can determine the concentrations of unknown substances using the formula of the line (provided the absorbance of our unknown sample falls within the line). The R^2 value is the correlation coefficient and is an indicator of accuracy (the closer R^2 value is to 1).

For example, if we have a naproxen sample with an unknown concentration and find its absorbance to be 0.081 for the data above, we can then rewrite the equation to be:

$0.081 = 0.0216x - 0.0117$

We can then rearrange the equation to make x (concentration) the subject:

x = (0.081+0.0117)/0.0216

and find the concentration of our sample to be:

x = 4.29 mg/mL (3 s.f.)

Note: each data point usually has uncertainty associated with it due to variables like the equipment you used to measure out the samples. Therefore, its important to add error bars in each data point when appropriate. Extrapolating (using values found outside the line of best fit) is not advised as it is likely to be inaccurate but make sure to indicate if you have done so.

Further considerations: Remember to keep in mind the context of your results and refer back to the aims/objectives of the experiment. Are your findings consistent with literature or the advertised amount? If not, why?

Ecology

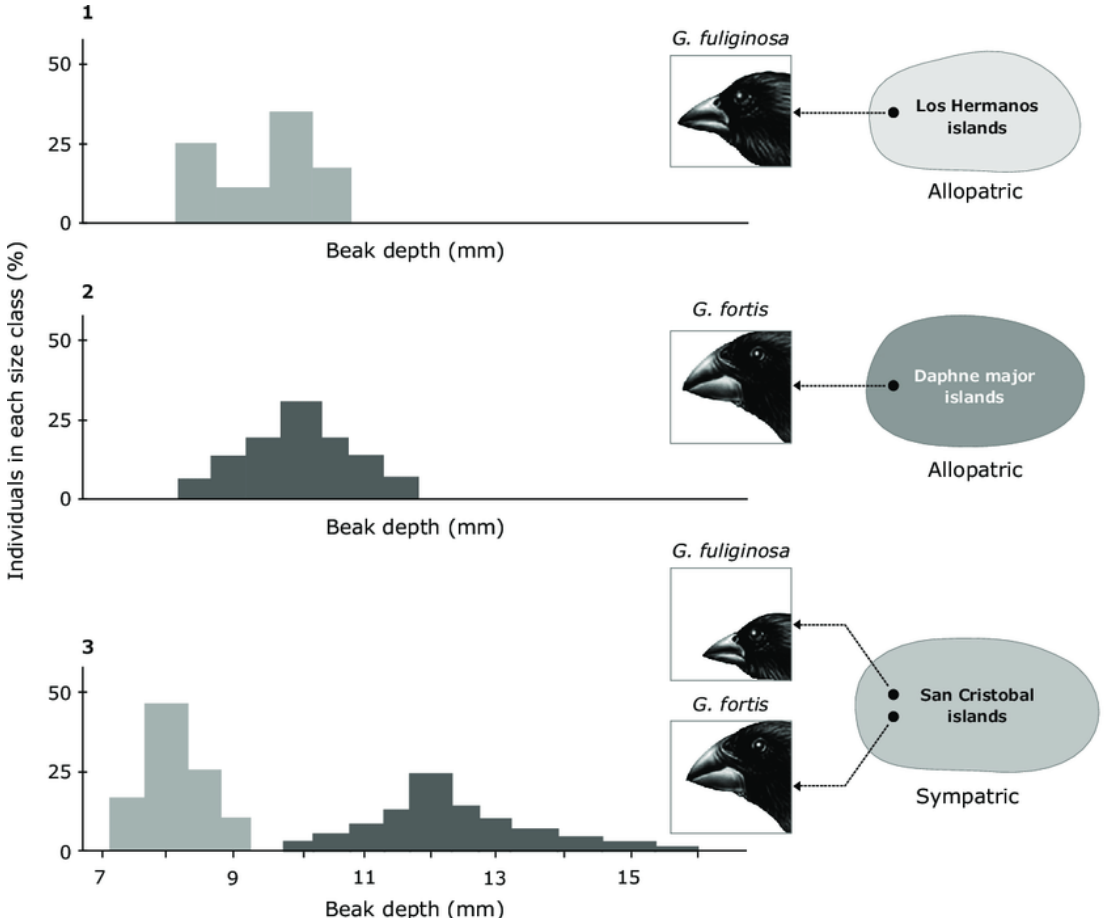


Figure 1. Character displacement in Darwin's finches. These two species of finches (genus Geospiza)

Ref: Fundamentals of Ecology and Environment 3e - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Character-displacement-in-Darwins-finches-These-two-species-of-finches-genus-Geospiza_fig6_355381814 [accessed 10 Dec 2024]

Related helpsheets

- Rearranging Equations