An exercise in data collection and display

David Hawcroft

Published online: 13 Dec 2010.

Submit your article to this journal ⬀

Article views: 37

View related articles ⬀

Interactive learning

An exercise in data collection and display

David Hawcroft

Dundalk Institute of Technology, Dundalk, Ireland

It is common to find students believing that they have a high skill level in computer use yet this does not normally include proficiency in the use of computer packages for data presentation. This exercise develops team work in data collection, requires the production of a simple spreadsheet, and the exporting of the data from that into one or more visual displays. Depending on the biological material used it can also form the basis of a discussion on sub-populations within a given biological area.
Key words: Student; Data; Displays; Excel

Introduction

Students commonly have difficulties with manual production of good quality data presentations (eg.. result graphs) yet in due course will also be expected to produce these using PC software packages. The work below is an exercise in the collection, production and display of scientific data, using easily obtained biological material, an Excel spreadsheet and graphical displays derived from that. Since in the author's experience many students have difficulty with the latter two areas, full details are given here but could be omitted from the instructions ultimately given to students.

The primary data for this exercise are obtained from measurements of parts of the leaf of the horse chestnut (*Aesculus hippocastanum*). Horse chestnut has the advantage of having leaves that are produced early in spring, well within the academic year, yet also extend into the autumn long enough to be useful at the beginning of an academic year. In addition they are robust enough to withstand storage by freezing and, indeed, leaves shed into the leaf litter can also be used. An important benefit of the species is that the leaflets fall into at least two size populations, giving an interesting result on a scatter plot, and allowing for a useful discussion on population distributions. In localities where horse chestnut is not available, other species could be used equally well.

Instructions

Data Generation

You will either be provided with about 50 (or 30 for younger students) horse chestnut leaves, or be asked to collect them yourself. The first task is to measure the length and width of both the main (central and largest) leaflet, and <u>one</u> of the basal (the smallest) leaflets. It is most convenient to do this as a pair of students, one measuring and the other recording the data.

A possible way of doing the measurement systematically

The 'measurer' should find a flat surface and place on it a rule showing a millimetre scale. To reduce the chance of mis-recording data, work in a particular sequence; for example Main Length, Main Width, Basal Length and then Basal Width. Place a leaflet on the ruler and determine the dimension in mm from the very tip (or if this is damaged, some reproducible point) of the leaflet to the point where the edge of the leaflet lamina joins the petiole (i.e. ignoring the petiole length). The width of the leaflet at its maximum point should also be measured.

The 'recorder' should draw up a table of columns (Main Length, Main Width, Basal Length and then Basal Width) and record the recited measurements in millimetres (remember 9.6cm = 96mm). However, it may be possible to enter these directly into an Excel spreadsheet, which would save time (and eliminate a definite and common source of error – called a transcription error – which could arise if they were transferred later).

So if you did not do it while taking the measurements, your first job is to transfer the data to an Excel spreadsheet with four columns (Main Length, Main Width, Basal Length and Basal Width).

Creating a simple Excel spreadsheet

Assuming you can see a toolbar on your PC screen, double click on the green X button

This should open a screen of boxes with
'Microsoft Excel – Book 1' in the top left corner

Click once in box A1 and
type 'MAIN LENGTH' (it should appear in the A1 box, and in the bar labelled X√ fx)
Use the right cursor arrow key to move to box B1
type MAIN WIDTH
Use the right cursor arrow to move to box C1
type BASAL LENGTH
Use the right cursor arrow to move to box D1
type BASAL WIDTH
You should save the spreadsheet at this point by going to 'File', 'Save', creating a file name, selecting where you want

it saving, and then clicking once on 'Save'.

Using the mouse to move the screen cursor arrow, click into box A2 and type in the first result

Use the down cursor arrow key to move to A3 and type in the next result.

Continue until all results are entered into the four columns. Note that it is useful to save the developing work periodically (perhaps every 10 data sets) and this is done simply by pressing 'Ctrl' and 'S' at the same time.

This finishes off a simple spreadsheet. There are ways of getting mean values and other calculations from these columns and these can be explored with a simple written guide or a tutor.

Data display
To set up the display
There are numerous ways in which the data can be displayed, for example as a scatter plot with both sets on the same axes. To do this:

Click once and drag across and down to highlight the first pair of data columns (e.g. Main Length, Main Width)

Click once on the chart wizard icon (in the middle of my top tool bar; it looks like a New York skyline)

Click twice on 'XY (Scatter)' this should then show a scatter plot

Click once on 'Series'

Click 'Add' in the Series box (this should show 'Series 2')

Click once in the 'X values' box; (and also delete any contents)

Click once on the blue bar at the top of the display box, hold and drag to one side for better viewing of the data columns

Click once and drag down to highlight (*) the next set of data to go on the x axis of the graph (e.g. Basal Length)

Click in the 'Y values' box (and delete any contents)

Click once and drag down to highlight (*) the next set of data to go on the y axis of the graph (eg Basal Width)

The plot should now show the two sets of data.

Click once on the blue bar to drag the chart display back to the centre if you wish

Click 'Next'

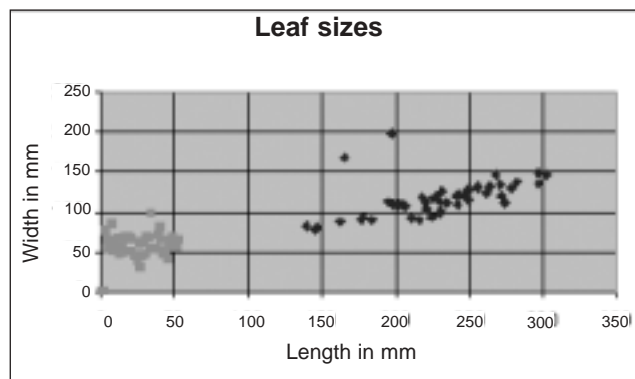(*) if the chart disappears while you do this it will reappear straight after the next operation.

Figure 1. A specimen scatter plot

To improve the titles
Select 'chart title' – delete any contents and change this to 'Leaf Sizes'

In 'Value (X) axis' add 'Length in mm'

In 'Value (Y) axis' add 'Width in mm'

Click 'Gridlines' tab and add gridlines to X and Y axes, or change as desired

Go to 'Legend' tab and delete at 'Show Legend'

Click 'Next' and 'Finish'

Alternatives
Other plot types can be found by using the chart wizard icon. You could also use the spreadsheet to manipulate the data. For example the scatter plot should show that the data from the two leaflet types fall into two distinct regions, but do the mean ratios differ between the main and basal leaflets? To determine a ratio click once in the next cell (ie E1), go to the fx bar and type =(A1/B1). After determining the first ratio value, left click and pull down to the bottom of the column to create a blue bar. Then ctrl + V will paste in all the ratrio values for that column. Give it a tile of MAIN RATIOS. Repeat in column F for the BASAL RATIOS. To determine the column mean values, go to the end of the column, click once in a vacant cell, type into the box at the top the following exactly as shown

=Average(top column cell code:bottom column cell code)

eg
=Average(B4:B54)

David Hawcroft is a lecturer in biochemistry in the Department of Applied Sciences, Dundalk Institute of Technology, Dundalk, Co Louth, Ireland. Tel: 00 353 42 931738. Fax: 00 353 42 9370566. Email: davidhawcroft@yahoo.co.uk