

Graphs R Us – Data Visualisation with R

Project

For each of the provided data sets, produce a publication-quality figure in R and save those in PNG or PDF format. Submit the figures in their PNG or PDF formats along with a single R script containing the code that produces these figures via *Turnitin* by **Friday December 14, 6 pm**. The R code also needs to show all data acquisition, graphical data exploration and data aggregation steps.

Multi-panel scatterplot

The first data set is called 'chilling_sensitivity.csv' and contains data from an experiment on the chilling sensitivity of four grass species. Plants were grown at a control (nonchilling) and a cool temperature (chilling) and at maturity their photosynthetic rate ($\mu\text{mol m}^{-2} \text{s}^{-1}$) was measured as a function of the ambient CO_2 concentration (ppm) set in the chamber of the measurement device.

Use graphical data exploration tools to explore this data set.

Create separate panels for each species and colour-code the two treatment levels (nonchilled vs. chilled) in each panel (your choice of panel arrangement in terms of number of rows or columns, respectively). Include the units given above in the axis titles and label the panels with italicised species names.

Produce two versions of this multi-panel figure using the ***traditional graphics system*** (R's built-in graphics) AND the ***ggplot2*** package.

Do not forget to include a legend when plotting with the ***traditional graphics system*** (*ggplot* adds one automatically).

Variables in the data frame

plant	factor with 3 levels (1, 2, 3) indicating the replicate
spec	factor with 4 levels indicating the grass species
treat	factor with 2 levels (nonchilled, chilled)
conc	numeric variable giving the CO_2 concentration set in the instrument chamber
photo	photosynthetic rate in $\mu\text{mol m}^{-2} \text{s}^{-1}$

Multi-panel barplot with insets

Use only the **traditional graphics system** to create this figure.

The second data set is called 'drought_elevation.csv' and stems from a global change study testing the impact of drought (control vs. rain exclusion) on the height growth (cm) of seedlings of three native tree species (*Agathis australis*, *Podocarpus totara* and *Beilschmiedia tawa*) across three elevations (low, mid, high).

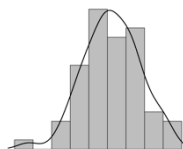
Use graphical data exploration tools to explore this data set.

Create separate panels for each elevation. Within each panel display the three species and for each species plot bars for control and drought treatment side by side.

The bars should represent means \pm standard errors indicated by error bars. Add italicised species names below the corresponding groups of bars at a 45° angle.

Do not forget to include a legend.

Add a small inset to each of the panels showing a *density histogram* (a normal histogram overplotted with a density line) for the seedling heights pooled across the drought treatment levels and species (i.e. form a subset for each elevation and use the height values from each subset in the histograms). Remove all axes and annotation in the histograms, so that your insets look like this:



Variables in the data frame

spec	factor with 3 levels indicating the tree species
elevation	factor with 3 levels (low, mid, high)
treat	factor with 2 levels indicating the control and the drought treatment (ctrl, drought)
height	seedling height in cm

Multi-panel map

Using the **ggplot2** package, create a two-panel figure showing a map of Australia with the state boundaries and abbreviations of the states printed in the center of each state in the left-hand panel. In the right-hand panel, plot a digital elevation model of Australia including a north arrow and colour bar indicating the elevation gradient.