

Exercises R for Beginners

Day 2 – 13/03/2018

1. – Learning R by comparing MS Excel with R: Translate an MS Excel application to R.

Load the dataset provided as `treesExample.xls`. Take a look at the data. Find minimum, maximum, number of trees and variance. Check if there is any missing value and find out how many missing values there are for total tree height (m) and dbh (diameter at breast height, cm).

What is the mean tree height and dbh in the given stand data? What are the values of variance, standard deviation, minimum and maximum values and the coefficient of variation?

2. – Examining vectors.

You have measured 20 trees in a circular plot of 15 meter radius. The measurements include the trees' stem diameters at breast height (dbh, in cm):

42.5, 34.4, 45.7, 43.9, 36.0, 28.1, 28.8, 38.3, 59.2, 32.7, 35.3, 39.4,
45.4, 40.5, 31.4, 35.1, 33.7, 26.8, 28.3, 28.6.

The corresponding total heights (in m) are:

26.0, 25.0, 29.3, 27.6, 24.5, 22.8, 24.1, 27.1, 28.8, 23.3, 22.9, 25.5,
26.7, 22.4, 24.5, 29.8, 20.5, 17.5, 19.8, 17.9

Read these data into two vectors (and give them recognisable names). Then explore the relationship between both variables. Compute the stand basal area per hectare.

Calculate mean standard deviation and coefficient of variation of the individual-tree basal areas.

Aggregate the three vectors (dbh, height and basal area) in a data frame. Add tree numbers and another variable with species information. The first ten trees have species code 1 = *Pinus spec* and the next ten trees have species code 2 = *Picea abies* (L.) H. Karst.

3. – Subsetting data

Use the same data set from Exercise 2.

Select the first eight observations. Select by variable values: Take a subset of trees where dbh is larger than 30cm. In addition, subset also those trees that are smaller than 30cm dbh. What is the mean height of those trees where dbh is larger than 30cm?

Select by species: Create two separate data frames for each species.

4. – Visualising data

You have measured 36 trees in a circular plot of 20 m of radius. There are two species. The first 18 trees correspond to *Pinus sylvestris* L. and the remaining two to *Picea abies* (L.) H. Karst. You have measured their dbhs:

26.0, 19.3, 27.1, 21.0, 23.0, 21.4, 23.2, 13.6, 8.4, 25.4, 28.5, 28.3,
22.0, 15.1, 24.9, 26.8, 28.3, 28.6, 42.5, 34.4, 45.7, 47.1, 51.9, 36.8,
32.8, 43.0, 43.9, 36.0, 28.1, 28.8, 38.3, 59.2, 32.7, 35.3, 39.4, 45.4

And these are the corresponding total heights:

18.8, 16.5, 20.8, 16.3, 17.6, 17.1, 19.1, 13.3, 9.3, 20.0, 20.3, 17.8,
20.5, 15.8, 18.1, 17.5, 19.8, 17.9, 26.0, 25.0, 29.3, 27.3, 28.9, 25.6,
23.9, 26.8, 27.6, 24.5, 22.8, 24.1, 27.1, 28.8, 23.3, 22.9, 25.5, 26.7

Produce a scatter plot of dbh versus height for each species.

Produce another scatter plot where the data points of both species appear in one single plot with black solid circles for pine and red ones for Norway spruce. Also add a legend.

Final points: Check your scripts.

Tidy up your R script by writing a few comments starting with #. Add your name and the current date. Explain in one comment sentence what the script is intended to achieve.

Now make sure your script is entirely reproducible: First clear the workspace using `rm(list=ls())` or click “clear” from the environment tab). Then run the entire script (by clicking “source” in the script window, top right).