

W02-1: Reading CSV files

This worksheet guides you in reading a CSV data set into one of R most versatile data structures - the data frame. After completing this worksheet you should have gained some experience in reading CSV files and accessing selected rows, columns or cells of data frames.

Things you need for this worksheet

- R — the interpreter can be installed on any operation system. For Linux, you should use the r-cran packages supplied for your Linux distribution. If you use Ubuntu, [this](#) is one of many starting points. If you use windows, you could install R from the official [CRAN](#) web page.
- R Studio — we recommend to use R Studio for (interactive) programming with R. You can download R Studio from the official [web page](#).
- Field survey 2014 subset 01 - a subset of the 2014 field survey data set can be downloaded from [here](#).

The big picture

During September and October 2014, an ecological field survey has been carried out at Chã das Caldeiras on Fogo island. The survey encompassed 161 plots of 5 meters by 5 meters on which data on agricultural and natural vegetation as well as on animal activity and diversity has been collected. In addition, some information has also been collected in a 10 meter by 10 meter area.

During the following worksheets, selected aspects of this data set are analyzed. Later worksheets will build on previous ones and in the end, you will have performed an actual (and hopefully meaningful) ecological analysis for the Fogo natural park.

For simplicity, we will use only a subset of this data set which encompasses all survey plots but for which the individual species records have been aggregated to richness or total activity values.

Learning log assignments

Data analysis always starts with reading the data set. So let's do it.

😬 Please download the field survey data subset and write an R script which reads the content of the data into a data frame. Check if everything is ok by looking at the first few lines of the data once the reading has been completed and get a comprehensive summary of the data set using the `summary()` function.

The R script you have just created will form the basis for the upcoming worksheets, so make sure you save it. For simplicity, please name your script files after the worksheet (i.e. "W02-1.R" in this case).

The following is just a finger exercises and will not directly be used in the upcoming worksheets. So if you want to store it, please make a copy of your script first.

😊 Please perform the following task as a finger exercises on accessing data subsets inside a data frame:

1. Get the values of the first column of the data frame (try both column access methods!)
2. Get the values of the first row of the data frame
3. Get the value of the second column in the first row of the data frame
4. Get the first 10 values of the first column of the data frame
5. Get the first 10 values of the first three columns of the data frame

You might have noticed cells with a value of NA. What does NA mean?

From: <http://bisfogo.environmentalinformatics-marburg.de/> - **BIS-Fogo**

Permanent link: <http://bisfogo.environmentalinformatics-marburg.de/doku.php?id=en:learning:schools:s01:worksheets:ba-ws-02-1>

Last update: **2015/09/22 15:22**

