

# Session 1 - Part 2: Read/write/analyze data

## Reading your data

Before starting with your analysis you have to get your data into R. R allows to work with many different data types, such as .csv, .xlsx, ASCII, .m, but also GIS raster or shape objects. Most of the time your data might be available in Excel spreadsheets. Here I recommend to save the data as .csv for the use in R. In the following two examples I will show how to load data in R. The first example demonstrates how easy reading data is, for the case of a already very clean and well prepared data set is. The second data set holds a very messy (again no offense to your data) data structure. To arrange the data for further analyses then requires some skills and some “creative” playing around with your data. Here one can consider to first prepare the data in Excel before loading it into R. If you gained already some experience in R it is much faster to rearrange your data in R.

## The working directory

For many analyses it is advisable to share all your input data, the R script and the outputs you might produce in one working directory. Especially when starting with R this is the easiest way to work with R. To set a working directory and to maybe query what the current working directory is following commands are helpful:

```
setwd("C:/Users/Christoph/Documents/Projects_R/RCourse")
getwd()
```

```
## [1] "C:/Users/Christoph/Documents/Projects_R/RCourse"
```

The path to your working directory you simply can copy paste from your explorer. As it is a text string it requires quotation marks. Caution here!: When working in Windows you have to replace all back slashes (“\”) in the path with front slashes (“/”). Back slashes have a different purpose in R. With the command *setwd()* you set your path to the directory you are working in. Now when loading or writing a file you do not have to take care about the paths anymore. Simply give the file name and R will look directly under this path for it.

## Reading a .csv file

To read and prepare a .csv file I will demonstrate on the data you provided me with. The first example is very clean data. In such a case there is nothing much to do but loading the data. The second example requires some tricks to bring it in the desired form for further analyses. If you want to follow the steps below, please save the .csv I sent you to a folder on your computer and set this as your working directory.

R offers many ways to load data. For starters the basic *read* family is the recommended way. When typing *read* into the help you will see that there are many different read commands. The most useful for you are the *read.table()* which is a more general read command and the *read.csv()* that is basically a *read.table()* but with default settings optimized for reading spreadsheets (e.g. separator is “,”, the first line is handled as a header: *header = TRUE*). A second command for .csv file might be important for you when you use european pc settings (e.g. separator in csv files is “;” instead of “,”) then the *read.csv2()* is the right choice as it uses european pc settings as default.

## Example 1:

As the data is already well prepared we will use the default *read.csv()* command for loading the data.

```
# Setting the working directory to the folder where the data is located  
setwd("C:/Users/Christoph/Desktop/R Course")  
  
# Read in the csv file with the standard read.csv command  
bio_data <- read.csv("Biomass_compiled.csv")
```