

How to read in data files

Usually data are sent stored in other files and you need to read them in.
Here are some ways of reading in data.

How to read in data files

Case 1. Suppose your data are located at a website, such as the `acid.txt` data for homework 2. The easiest way to read it in is by telling R the web address directly like this:

```
>d1<-read.table("http://www.math.unm.edu/~james/acid.txt",header=T)
> d1 <- read.table("http://www.math.unm.edu/~james/acid.txt")
> head(d1)
  conc exper
1 0.123 Acid1
2 0.109 Acid1
3 0.110 Acid1
4 0.109 Acid1
```

Here `d1` is an arbitrary name given to a variable which stores the data as a data frame. The option `header=T` means that it is true that the first line consists of variable names instead of data. The default for the `read.table()` function is to assume that the first line is data. You can extract a column using `d1$conc` or `d1$exper`.

How to read in data files

Case 1, continued. You should be aware of what happens if you forget to specify `header=T`.

```
> head(d1)
      V1    V2
1 conc exper
2 0.123 Acid1
3 0.109 Acid1
4  0.11 Acid1
5 0.109 Acid1
6 0.112 Acid1
```

R thinks that `conc` and `exper` are character data instead of variable names. When variable names aren't specified, R automatically creates variables called `V!`, `V2` etc for the columns. You can refer to `d1$V1` for the first column. Because the first observation is a string, you will have an extra observation and your data will be treated as non-numerical.

How to read in data files

Case 2. You download the data file to your computer.

If you download the data file to your computer, then you can also read it in using the `read.table()` statement for plain text files (such as files ending in `.txt`).

The important thing here is that if you specify the name of the file, R assumes that the file is in the current working directory (i.e., folder) for your R session, or that you specify the path for the file.

```
> getwd()
[1] "/Users/superjames/Documents/TeachingUNM/STAT427"
> dir()
[1] "*"                "#email.txt#"
[3] "accommodation"   "acid.txt"
```

Here I see that the `acid.txt` file is in my current directory, so I can read it in. If your file isn't there, you should put it in your current directory. It's often helpful to have a directory dedicated to the class or to a specific project.

How to read in data files

Instead of moving your data, you can also change your current working directory. This can either be done through the function `setwd()` or through the menus for R. In R studio it is under

session – > Set Working Directory

in the menus.

In RStudio, you can also use the menus to open a data set, under File – > Open File.

A final approach is to use `read.table()` and specify the path for the file, such as

```
d1 <- read.table("/Users/superjames/Downloads/acid.txt",header=T)
```

How to read in data files

If you have a comma separated file with extension `.csv`, you can use

```
read.csv()
```

This works basically the same way as `read.table()` but assumes a comma separated `.csv` file.