Lab2 Instruction

Jun Cai

September 18, 2015

Note

- All R codes are present in boxes with grey background. You can run them in your R Console. Lines leading by `##` are outputs of R codes.
- All R codes are run correctly on my own Mac OS X. When you try them on your own computer, please customize your own working directory.
- The functions in **bold** are recommended to use.
- The Lab Instruction in different formats including .Rmd, .md, .html and .pdf are available on my GitHub. For your convenience, the data used and results produced in the Lab are also provided in the Github.

Lab2 contains basics for file and directory manipulation, and R data input and output. Rather than a complete collection of functions, I will introduce the frequenty-used functions from my own R experience.

Part I File and Directory Manipulation

R has a variety of functions for file and directory manipulation. The following are a few examples:

`setwd()` and `getwd()`: used to change or determine the current working directory. It's a good habit to set working directory before your data analysis as all results during your data analysis will be stored in the working directory.

`list.files()` and `list.dirs()`: returns a character vector of names of files or directories under the given directory.

`file.info()`: gives file size, creation time, directory vs. ordinary file status, and so on for each file whose name is in the argument, a character vector.

`file.create()` and `dir.create()`: creates files or directories with the given names if they do not already exist.

`file.exists()` and `dir.exists()`: returns a logical vector indicating whether the given file exists for each name in the first argument, a character vector.

`file.copy()` and `file.rename()`: moves files `from` source path `to` destination path.

`file.remove()` and **`unlink()`**: deletes the files or directories specified by the first argument, a character vector.

```
# set current working directory to DAE
setwd("/Users/tonytsai/Documents/R/DAE")
getwd()
```

```
## [1] "/Users/tonytsai/Documents/R/DAE"
```

```r
# list all files including directories under current working directory DAE
list.files()
```

```
## [1] "Lab1note.pdf"             "Lab2"
## [3] "Lecture1Introduction.pdf" "LICENSE"
## [5] "README.md"                "reference"
## [7] "script"                   "生态数据分析课程大纲.pdf"
```

```r
# extract file information for those files
file.info(list.files())
```

```
##                               size isdir mode               mtime
## Lab1note.pdf                553464 FALSE  777 2015-09-20 12:59:30
## Lab2                           306  TRUE  755 2015-09-20 20:06:58
## Lecture1Introduction.pdf   1134795 FALSE  640 2015-09-20 12:59:16
## LICENSE                       1077 FALSE  644 2015-09-18 09:40:04
## README.md                      244 FALSE  644 2015-09-18 09:58:11
## reference                      170  TRUE  777 2015-09-16 11:22:01
## script                         136  TRUE  755 2015-09-20 20:06:58
## 生态数据分析课程大纲.pdf    253722 FALSE  777 2015-09-12 18:07:12
##                                           ctime               atime uid gid
## Lab1note.pdf              2015-09-20 14:24:21 2015-09-20 14:24:19 501  20
## Lab2                      2015-09-20 20:06:58 2015-09-20 20:06:58 501  20
## Lecture1Introduction.pdf  2015-09-20 12:59:29 2015-09-20 12:59:15 501  20
## LICENSE                   2015-09-18 09:42:06 2015-09-20 12:49:52 501  20
## README.md                 2015-09-18 09:58:11 2015-09-18 09:58:04 501  20
## reference                 2015-09-16 13:27:58 2015-09-20 20:06:58 501  20
## script                    2015-09-20 20:06:58 2015-09-20 20:06:58 501  20
## 生态数据分析课程大纲.pdf  2015-09-16 13:27:58 2015-09-16 13:27:58 501  20
##                               uname grname
## Lab1note.pdf              tonytsai  staff
## Lab2                      tonytsai  staff
## Lecture1Introduction.pdf  tonytsai  staff
## LICENSE                   tonytsai  staff
## README.md                 tonytsai  staff
## reference                 tonytsai  staff
## script                    tonytsai  staff
## 生态数据分析课程大纲.pdf  tonytsai  staff
```

```r
# list only directories under DAE
list.dirs()
```

```
## [1] "."
## [2] "./Lab2"
## [3] "./Lab2/data"
## [4] "./Lab2/data/CMDSSS"
## [5] "./Lab2/data/CMDSSS/SURF_CLI_CHN_MUL_DAY_V3.0"
## [6] "./reference"
## [7] "./script"
```

```r
# find all R scripts under DAE and give their full path names (or absolute paths)
list.files(recursive = TRUE, pattern = ".R$", full.names = TRUE)
```

```
## [1] "./script/20150916.R"
```

```r
# create a recursive directory under DAE/Lab2, which stores the TXT data
# that will be read in Part III.
if(!dir.exists("Lab2/data/CMDSSS/SURF_CLI_CHN_MUL_DAY_V3.0"))
  dir.create("Lab2/data/CMDSSS/SURF_CLI_CHN_MUL_DAY_V3.0", recursive = TRUE)

# create a temporary directory under script
if(!dir.exists("script/tmp")) dir.create("script/tmp")
# create a temporary R script under tmp to say Hello World, Hello R!
file.create("script/tmp/tmp.R")
```

```
## [1] TRUE
```

```r
cat("print('Hello World, Hello R!')", file = "script/tmp/tmp.R")
# excute the R script
source("script/tmp/tmp.R")
```

```
## [1] "Hello World, Hello R!"
```

```r
# copy tmp.R to helloworld.R
file.copy("script/tmp/tmp.R", "script/helloworld.R")
```

```
## [1] TRUE
```

```r
list.files("script", recursive = TRUE)
```

```
## [1] "20150916.R"   "helloworld.R" "tmp/tmp.R"
```

```r
# rename helloworld.R to hello.R
file.rename("script/helloworld.R", "script/hello.R")
```

```
## [1] TRUE
```

```r
list.files("script", recursive = TRUE)
```

```
## [1] "20150916.R" "hello.R"    "tmp/tmp.R"
```

```r
# delete all R scripts under script directory except for 20150916.R
# attempt to delete inexistent hellworld.R
file.remove(c("script/hello.R", "script/helloworld.R"))
```

```
## Warning in file.remove(c("script/hello.R", "script/helloworld.R")): cannot
## remove file 'script/helloworld.R', reason 'No such file or directory'
```

```
## [1]  TRUE FALSE
```

```
list.files("script", recursive = TRUE)
```

```
## [1] "20150916.R" "tmp/tmp.R"
```

```
# delete the temporary directory that is not empty.
unlink("script/tmp", recursive = TRUE)
list.files("script")
```

```
## [1] "20150916.R"
```

To see all the file- and directory-related functions, type the following:

```
?files
```

Part II Capturing R Console Output

R provides functions to save the results that appear in your R Console into a text file.

`sink()`: diverts R output to a file connection.

`capture.output()`: sends R output to a character string or file connection.

```
# type following codes in your Console
# divert R output to CO2.txt under data directory
sink(file = "data/CO2.txt")
# load R built-in dataset CO2. type ?CO2 to see the decription about CO2 dataset.
data("CO2")
# divert the first 7 rows of CO2
head(CO2, 7)
# end the divertion
sink()
```

```
# sink() does not work in knitr, because it is already used internally to capture
# results. To make it work, you have to use {} wrap up all aboving code in a
# single expression and print results to Console explicitly.
{
  sink(file = "data/CO2.txt")
  data("CO2")
  print(head(CO2, 7))
  sink()
}
# the content of CO2.txt will be checked in Part III.
```

```
# generate a vector containing a numeric sequence from 1 to 10
1:10
```

```
##  [1]  1  2  3  4  5  6  7  8  9 10
```

```
# send the output to variable x
x <- capture.output(1:10)
x
```

```
## [1] " [1]  1  2  3  4  5  6  7  8  9 10"
```

```
# divert the first 7 rows of CO2 dataset to CO2.txt
capture.output(head(CO2, 7), file = "data/CO2.txt")
```

Part III Imports and Exports

`read.table()` and `write.table()`:

`read.csv()` and `write.csv()`:

`read.xls()` and `write.xls()`:

`read.dbf()` and `write.dbf()`:

`load()` and `save()`:

`data()`:

References

The following are materials on R data import/export that you can access on the Web.

- R Data Import/Export