Chapter 5 Review

2017100057 / 이영노

October 8, 2022

Chapter 5

Chapter5 Intro: path, working directory/ `readr`,`readxl`/ storing & organizing

Chapter5.1 path, working directory

- example1

```
filename <- "murders.csv" # 이때 murders.csv는 빈공간임. dslabs에서 copy되기 전.
dir <- system.file("exdata", package="dslabs") # data, package들의 디렉토리를 알려줌.
fullpath <- file.path(dir,filename) # combine directories to produce path
file.copy(fullpath,"murders.csv") # from exdata-dslabs, to murders.csv
```

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

this code: Not READ the data into R, but just COPY

Once this file is copied, we can import data

- 결론: 필요한건 (1)디렉토리 (2)파일이름

```
library(tidyverse)
```

```
## Warning: 패키지 'tidyverse'는 R 버전 4.1.3에서 작성되었습니다
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.3.6      v purrr   0.3.4
## v tibble  3.1.7      v dplyr   1.0.10
## v tidyr   1.2.0      v stringr 1.4.0
## v readr   2.1.2      v forcats 0.5.2
```

```
## Warning: 패키지 'ggplot2'는 R 버전 4.1.3에서 작성되었습니다
## Warning: 패키지 'tibble'는 R 버전 4.1.3에서 작성되었습니다
## Warning: 패키지 'tidyr'는 R 버전 4.1.3에서 작성되었습니다
## Warning: 패키지 'readr'는 R 버전 4.1.3에서 작성되었습니다
## Warning: 패키지 'purrr'는 R 버전 4.1.3에서 작성되었습니다
## Warning: 패키지 'dplyr'는 R 버전 4.1.3에서 작성되었습니다
## Warning: 패키지 'stringr'는 R 버전 4.1.3에서 작성되었습니다
```

```
## Warning: 패키지 'forcats'는 R 버전 4.1.3에서 작성되었습니다
```

```
## -- Conflicts --------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
# dat <- read_csv(filename)
```

Chapter5.2 Finding files

- **Directory**: 그릇, 주소지(서울시/성북구/안암동)
- **Path**: 이동경로(서울시-성북구-안암동-21.4)

```
system.file(package="dslabs") # full path
```

```
## [1] "C:/Users/user/Documents/R/win-library/4.1/dslabs"
```

- full path: 원래 패키지, 데이터가 있었던 root directory에서 찾는거

```
C:\Windows\calc.exe
```

- relative path: working directory에서 찾는것

```
calc.exe
```

- example2: Relative path

```
dir <- system.file(package="dslabs")
list.files(path = dir) # dslabs 디렉토리에 존재하는 파일들 리스트(그릇내 음식들)
```

```
##  [1] "data"        "DESCRIPTION" "extdata"     "help"        "html"
##  [6] "INDEX"       "MD5"         "Meta"        "NAMESPACE"   "R"
## [11] "script"
```

- example3: obtaining Full path

```
# getwd() # full path of working directory

# filename <- "murders.csv"
# dir <- system.file("extdata",package="dslabs") # full path
# fullpath <- file.path(dir, filenames)
```

Chapter5.3 Generating path name

- example4

```
dir <- system.file(package = "dslabs")
filename %in% list.files(file.path(dir,"extdata"))
```

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

```
# dslabs-extdata-murders.csv

dir <- system.file("extdata",package="dslabs") # "extdata"is subdirectory
dir
```

```
## [1] "C:/Users/user/Documents/R/win-library/4.1/dslabs/extdata"
```

```
fullpath <- file.path(dir,filename) # combine directories

file.copy(fullpath,"murders.csv") # 요구사항: (1)full path to copy (2)name to paste
```

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

```
# FALSE: directory already has same file.

list.files() %>% head()# see files in working directory
```

```
## [1] "~$22-5^J6^J7월(수정).xlsx"
## [2] "~$22-5^J6^J7월(자동 복구됨).xlsx"
## [3] "~$22-9^10^11월(수정)(자동 복구됨).xlsx"
## [4] "~$22-9^10^11월.xlsx"
## [5] "15-02__1.PDF.pdf"
## [6] "2017100057.jpg"
```

Chapter5.4 Reading files

- example5

```
# (1)
getwd() # Desktop is my current working directory
data=read_table('제조업부가가치.txt')
knitr::kable(as.data.frame(head(data)))

# OR we can write the full path, if working directory is different
# (2)
data2=read_table("C:/Users/user/Desktop/제조업부가가치.txt")
data2 %>% head()

# (3)
url <- ~~~
# dat <- read_csv(url)

# (4) local copy
# download.file(url, "filename")

# (5) scan(sep=',')
path <- system.file("extdata",package="dslabs")
filename <- "murders.csv"
x <- scan(file.path(path,filename),sep=',',what='c')
```