

Chapter 1

파일 입출력

1.1 데이터 파일 입출력

데이터 입력과 출력은 R을 이용한 분석에서의 첫번째 단계와 마지막 단계라고도 할 수 있습니다. 데이터가 입력된 형식은 매우 다양하지만, 일반적으로 R을 이용하여 데이터의 입출력을 하는데 있어 안전한 방법은 .csv이라는 파일확장자를 가진 파일을 이용하는 것입니다. 따라서, 가급적이면 다른종류의 파일확장자를 .csv로 먼저 변경한 뒤에 사용하는 것이 좋습니다.

```
mydata <- read.table(file="./filename.csv", header=TRUE, sep=",")
```

여기에서 filename.csv 은 파일명입니다.

입력과 관련된 문제해결법 read.table() 함수를 이용하여 데이터를 불러오는데 있어서 많이 발생하는 오류는 “데이터 파일을 작업디렉토리로부터 찾을 수 없다” 또는 “데이터 파일이 존재한 파일경로가 올바르지 않다” 라는 것입니다.

파일을 입력받을 때 R은 일반적으로 첫번째 인자에 주어진 파일명과 현재 작업디렉토리의 파일경로를 함께 묶어 절대경로를 생성한 뒤, 이 절대경로를 이용하여 파일명을 찾습니다. 이러한 원리때문에 운영체제가 영어가 아닌 컴퓨터의 경우, 이 절대경로를 올바르게 생성하지 못할 경우가 있습니다. 또한, 파일경로명에 띄어쓰기가 있는 경우 및 특수문자가 포함된 경우에 이러한 문제가 발생할 경우가 있습니다. 따라서, 사용자는 간혹 문법에서 틀린 점도 없고, 불러오고자 하는 데이터 파일도 올바른 파일경로에 위치하고 있음에도 불구하고, 데이터를 찾을 수 없다는 에러 메시지를 보게 되는 경우가 있습니다. 이러한 경우에 보다 안전한 방법으로 read.table() 사용하고자 한다면 아래와 같이 file.choose() 함수 또는 file.path() 함수를 이용하시길 바랍니다.

```
mydata <- read.table(file.choose(), header=TRUE, sep=",")
```

file.choose()는 탐색기를 띄워 사용자가 원하고자 하는 파일을 찾을 수 있도록 도와줍니다.

```
mydata <- read.table(file.path(), header=TRUE, sep=",")
```

```
mydata1 <- read.table(file=url(site_address), header=TRUE, sep=",")
```

file.path()는 절대경로를 보다 안전하게 R이 이해할 수 있도록 도와줍니다.

```
age <- scan()  
32 33 39 28 20 20
```

1.1.1 출력

저장하기

`.RData`

`.CSV`

`.HTML`

```
write(t(mydata), file="./where/should/be/saved", ncolumns)
```

데이터셋 또는 변수에 주석첨가하기