

## 1강. R데이터처리(1)

◆ 담당교수 : 김성수 교수

### ■ 주요용어

| 용어       | 해설   |
|----------|--|
| R        | 인터넷에서 자유롭게 다운받아 사용할 수 있는 통계소프트웨어이면서도 기능이 매우 뛰어난 통계시스템임. R은 프로그램의 기능을 가지고 있어 자체로서도 유용한 기능들을 구현할 수 있을 뿐만 아니라 자료처리 및 그래픽스 분야에 탁월한 기능을 가지고 있다.     |
| R Studio | R Studio는 사용자가 좀 더 쉽게 R을 사용할 수 있도록 개발된 R 통합환경 시스템이다. R Studio는 <a href="http://www.rstudio.com">http://www.rstudio.com</a> 에서 다운받아 설치하면 된다.    |
| 작업 영역    | R에서 데이터파일을 읽어들이기 위해 설정된 영역을 말한다. R의 작업 영역을 보기 위해서는 "getwd()"를 이용한다. 작업영역을 지정하려면 setwd()를 이용한다.   |
| 텍스트파일 읽기 | 텍스트 파일을 데이터 프레임(data frame)으로 읽기 위해서는 read.table() 명령을 사용한다. 변수값들이 “,” 로 구분된 경우에는 read.csv()를 사용한다. 데이터 객체를 파일로 저장할 때는 write.table() 명령을 이용한다. |

### ■ 연습문제

1. R에서 데이터를 가져오고 저장할 작업영역인 폴더를 지정하고자 한다. 작업영역을 대화형으로 지정하기 위한 명령은?

정답 및 해설 : setwd(choose.dir())

2. R에서 데이터의 작업영역을 알고자 한다. R 명령은?

정답 및 해설 : `getwd()`

3. 다음과 같은 텍스트 파일을 알고자 한다. 적합한 명령 (a)는 ?



| id | sex | job | religion | edu | amount | salary |
|----|-----|-----|----------|-----|--------|--------|
| 1  | m   | 1   | 3        | 7.0 | 110    |        |
| 2  | m   | 2   | 1        | 4   | 12.0   | 135    |
| 3  | f   | 2   | 3        | 5   | 8.5    | 127    |
| 4  | f   | 3   | 3        | 5   | 5.0    | 150    |
| 5  | m   | 1   | 3        | 3   | 4.5    | 113    |
| 6  | m   | 2   | 1        | 2   | 3.5    | 95     |
| 7  | m   | 3   | 2        | 4   | 4.0    | 102    |
| 8  | f   | 3   | 2        | 4   | 4.0    | 122    |
| 9  | f   | 2   | 3        | 4   | 4.5    | 110    |
| 10 | m   | 1   | 3        | 5   | 17.0   | 200    |
| 11 | f   | 1   | 1        | 3   | 22.0   | NA     |
| 12 | m   | 2   | 1        | 2   | 5.5    | 105    |
| 13 | m   | 3   | 2        | 1   | 4.5    | 130    |
| 14 | m   | 3   | 2        | 5   | 7.0    | 150    |
| 15 | m   | NA  | 3        | 4   | 6.0    | 110    |
| 16 | f   | 1   | 3        | NA  | 7.0    | 88     |
| 17 | m   | 1   | 1        | 4   | 6.0    | 138    |
| 18 | f   | 2   | 1        | 5   | 5.0    | 110    |
| 19 | m   | 2   | 3        | 3   | 7.0    | 85     |
| 20 | m   | 3   | 3        | 4   | 9.5    | 110    |
| 21 | m   | 3   | 1        | 4   | 10.0   | 95     |
| 22 | m   | 3   | 2        | 3   | 12.0   | 88     |

`insurance.data = ( a )"c:/Rfolder/data/insurance.txt", header=T)`

정답 및 해설 : `read.table`

#### ■ 참고자료

- Oscar T.-R. (2013), Introduction to R Studio,
- Pruim, R. (2011), Computational Statistics Using R and R Studio,
- W.N.Venables, D.M.Smith and the R Core Team(2014), An Introduction to R. Notes on R: A Programming Environment and Data Analysis and Graphics.
- R : <https://www.r-project.org/>
- R Studio : <https://www.rstudio.com/>

## ■ 코스웨어

- R의 기본 코스웨어 학습방법 안내

다음은 R에 대한 기본적인 학습을 할 수 있는 방법을 안내 해 드립니다. 아직 R에 대하여 익숙하지 못한 학습자는 “R의 입문 코스웨어”를 통하여 R에 대하여 익숙해지기 바랍니다.

### 1) [학교 홈페이지에 접속] - [우측하단 “학과” 선택]



### 2) [“정보통계학과” 선택]



### 3) [상단메뉴 학습정보 내의 “학습자료실” 선택]



### 4) [좌측메뉴의 “코스웨어” 선택]





7) 각 장의 좌측메뉴의 학습목차를 참고하여 학습



“R의 입문 코스웨어” 과정을 통하여 R의 기본개념 및 구조를 이해하고 실제 작업에 필요한 자료처리 및 분석기법을 응용할 수 있는 능력을 학습하기 바랍니다.