

R 기초

R 기본연산자



R 기본 연산자

◇ 변수설정

◇ **Assign operators, arithmetic operators, logical operators, relational operators**

📋 변수설정

```
> 1+1  
2  
> X = 1+1  
> X  
2
```

🔍 R 객체가 저장된 위치의 이름

- ✓ 변수명은 **영문자, 한글 또는 .**로 시작하고 이후 숫자, _ 사용가능
- ✓ 일부 예약어는 사용할 수 없음
 - if, for, while, TRUE, FALSE ...
- ✓ 대문자와 소문자 구분
 - If, For, While, True, False는 변수명으로 사용가능

Assign Operators

◇ `=, <-, ->`

✓ `=` : 오른쪽의 연산 값을 왼쪽의 변수에 할당(대입)

✓ `<-` : 오른쪽의 연산 값을 왼쪽의 변수에 할당(대입)

✓ `->` : 왼쪽의 연산 값을 오른쪽의 변수에 할당(대입)

```
> X1 = 1  
> X2 <- 2  
> 3 -> X3
```

◇ `c()`: 여러 개 자료를 묶을 때 사용하는 함수

```
> X = c(1, 2, 3)
```

Arithmetic Operators

✧ $+, -, *, /, ^, **, \% \%$ (나머지), $\%/\%$ (몫)

$2+3$; $2-3$; $2*3$; $2/3$
 $2**3$; $13\%\%5$; $13\%/\%5$

✓ 연산순위: $(^, **)$ > $(\%\%, \%/\%)$ > $(*, /)$ > $(+, -)$

✓ () 안의 내용을 우선 연산

✓ 같은 우선 순위인 경우 왼쪽부터 먼저 실시

$1+6/2$; $(1+6)/2$
 $2*3**2$; $(2*3)**2$
 $12/3*4$; $12/(3*4)$



Logical Operators

◇ 논리합: “|”로 표시

	결과
True True	True
True False	True
False True	True
False False	False

◇ 논리곱: “&”로 표시

&	결과
True & True	True
True & False	False
False & True	False
False & False	False

✓ TRUE는 T, FALSE는 F로 표시할 수 있음

✓ TRUE는 1, FALSE는 0으로 처리

0이면 FALSE, 0이 아닌 숫자(Inf 포함)는 TRUE로 처리

Logical Operators

◇ `!, &, &&, |, ||, xor(x, y)`

✓ 결과는 TRUE, FALSE

◇ `&`와 `&&`, `|`와 `||`의 차이점

`c(TRUE, FALSE) | TRUE`
`c(TRUE, FALSE) & TRUE`

`(TRUE, TRUE)`
`(TRUE, FALSE)`

`c(TRUE, FALSE) || TRUE`
`c(TRUE, FALSE) && TRUE`

TRUE
TRUE

Logical Operators

◇ 논리연산 순서는 산술연산과 동일

TRUE FALSE TRUE
TRUE TRUE & FALSE
(TRUE FALSE) & TRUE
(TRUE TRUE) & FALSE

TRUE
TRUE
TRUE
FALSE

- ✓ 논리합과 논리곱이 혼재되어 있는 논리연산의 순서는 논리곱을 먼저 수행
- ✓ 연산순위가 같은 경우 왼쪽부터 수행
- ✓ () 안에 논리연산식 우선 수행

Relational Operators

◇ $>$, $<$, $>=$, $<=$, $==$, $!=$

✓ 비교는 일종의 의문문으로 이해

✓ 관계결과는 TRUE, FALSE

R활용 기본분석과정

분석자료 불러오기-scan



통계분석과정

주요 단계	세부 내용
연구주제설정	연구대상 선정
자료수집	조사, 실험, 관찰, ...
자료 전처리	자료 불러오기
	자료 수정, 보완
	변수, 관측값 선택
	자료 병합
자료분석	기술 통계(시각적, 수치적 해석)
	통계적 추론
결론 도출 및 활용	위의 과정에 따라 일반화 가능여부 결정

📋 R 분석에서의 기본 자료 구조

🔗 데이터 프레임(data frame)

변 수

관
측
개
체

번호	성	연령	신장	체중	비만도	혈액형	총치
001	남	25	181	75	정상	B	5
002	남	23	175	75	과체중	A	1
003	여	19	161	48	정상	A	4
004	여	23	178	67	정상	A	2
005	여	21	165	55	정상	A	3
006	여	20	165	47	저체중	O	2
007	남	23	170	78	비만	A	4
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

📋 R 분석에서의 기본 자료 구조

📌 변수(variable)

- ✓ 일변량 자료(univariate data)

- ✓ 다변량 자료(multivariate data)

- 변수들 간 관련성(예측, 분류)
- 변수값에 따른 비교

📌 관측개체(observation, 관측)

- ✓ 관측개체 간 관련성 유무에 따라 분석방법이 다름

📄 분석자료 불러오기 I

🔑 입력방식

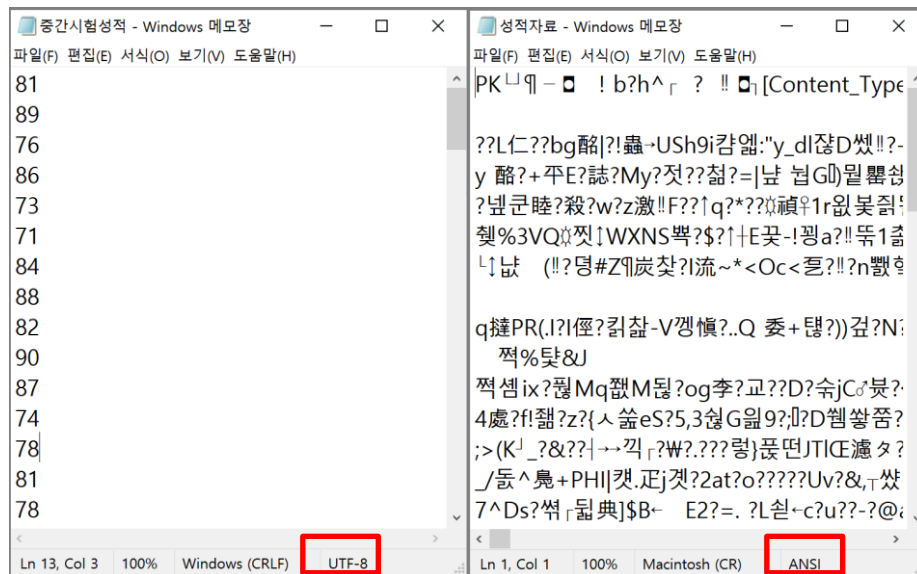
✓ 직접입력

✓ 파일 자료 불러오기

○ 텍스트 파일 : 저장방식(ANSI, UTF-8, ...)

○ 특수형태파일(Excel, ...)

⇒ package 필요



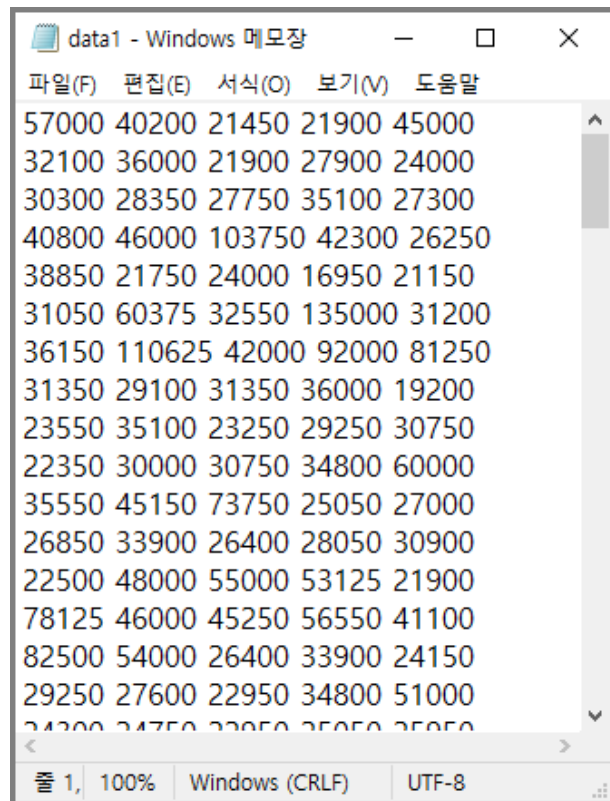
📋 분석자료 불러오기 I

🔑 텍스트 파일인 경우

✓ Table 형태가 아닌 경우

✓ Table 형태인 경우

- 탭(tab), 빈칸(space)로 분리
- 구분 기호로 분리: 콤마(,) 등
- 너비가 일정



57000	40200	21450	21900	45000
32100	36000	21900	27900	24000
30300	28350	27750	35100	27300
40800	46000	103750	42300	26250
38850	21750	24000	16950	21150
31050	60375	32550	135000	31200
36150	110625	42000	92000	81250
31350	29100	31350	36000	19200
23550	35100	23250	29250	30750
22350	30000	30750	34800	60000
35550	45150	73750	25050	27000
26850	33900	26400	28050	30900
22500	48000	55000	53125	21900
78125	46000	45250	56550	41100
82500	54000	26400	33900	24150
29250	27600	22950	34800	51000
24200	24750	22050	25050	25050

📋 자료 불러오기: 정리

🔍 입력방식

☑️ 직접입력 : **c()** 함수, **scan()** 함수

```
> x <- c(1,2,3,4)
> y <- c("A","B",'C','D')
> x
[1] 1 2 3 4
> y
[1] "A" "B" "C" "D"
> data1 <- scan()
1: 1
2: 2
3: 3 4
5: 5
6:
Read 5 items
> data1
[1] 1 2 3 4 5
```


📋 자료 불러오기: 정리

💎 입력방식

✓ scan()

- 주요 옵션:

- scan(file = "", what = double(0), n = -1, sep="", skip = 0, nlines = 0, fileEncoding = "", ...)

- 문자인 경우 what = "character"

- 다변량 자료(숫자,문자)인 경우 what=list(0,"")

- ⇒ list 형태로 만듦

```
> x1 <- scan("data1.txt")
Read 474 items
> str(x1)
num [1:474] 57000 40200 21450 21900 45000 ...
```

✓ 한글이 들어있는 경우 저장방식에 맞게 불러오기