

##### 분석자료불러오기: scan

# 직접입력: c()

## \* 벡터생성

x <- c(1,2,3,4)      c(1, 2, 3, 4, "a") : 모두 문자가 됨

y <- c("A","B","C","D") "나' 상관없이 둘 다 문자로 처리

# 직접입력: scan()

## \* 옵션이 없는 경우 수치형 벡터 생성

data1 <- scan()

1

2

3 4

5

## \* 문자형 자료

data2 <- scan()

A B C      ERROR: scan은 문자를 읽어들이지 못함

? scan      도움말      몇번째 줄부터지      몇번째 줄까지      끝까지 읽어들이는 뜻

## \* 주요 옵션: scan(file = "", what = double(0), n = -1, sep = "",  
## skip = 0, nlines = 0, fileEncoding = "", ... )

## \* 문자 자료인 경우 what = "" 또는 what = "character"

data2 <- scan(what="") 문자 입력 지정

A B C

## \* 다변량 자료 입력

##### \* 예제: 첫번째 숫자, 두번째 문자

data3 <- scan(what=list(0,"")) 숫자, 문자

1 A

2 B

3 4 5 T

6 F

# 텍스트 파일 자료 불러오기: scan()

## \* 작업 디렉토리 변경

getwd()      # 현재 작업공간

위치 중요

dir(getwd())      # 현재 작업공간에 있는 파일목록

setwd("E:/통계프로그래밍")      # 작업공간 변경: 경로는 '₩'가 아닌 '/'로 표시

x1 <- scan("data1.txt")

str(x1)

## \* 부분자료 읽기      숫자로 되어있고, 총 몇 개의 값이 있는지와 최초 5개의 값을 보여줌

x2 <- scan("data1.txt", n=10)      앞부터 10개의 데이터만 읽어오기

x3 <- scan("data1.txt", nlines = 10)      앞부터 10줄을 읽어오기

## 문자형 자료 읽기

gender <- scan("gender.txt", "")      # 파일명이 있는 경우

## \* 다변량자료 파일 읽기      가장 첫 줄은 변수명이므로  
2번째 줄에서부터 시작해야함

x4 <- scan("data2.txt", what=list("", "", 0))

-> 1번째 줄은 skip

x4 <- scan("data2.txt", what=list("", "", 0), skip=1)      # Read 474 records

str(x4)

x5 <- scan("data3.txt", what=list("", "", 0), skip=1)      # 저장방식이 UTF-8

x5 <- scan("data3.txt", what=list("", "", 0), skip=1, fileEncoding = "UTF-8")

한글이 있는 경우 저장방식에 맞게 불러와야함