

05 데이터 처리 기본 ¶

학습 내용

- 기본 함수 알아두기

Review

- R과 RStudio 설치
- 패키지 설치 : `install.packages("패키지명")`
- 라이브러리 불러오기 : `library(라이브러리명)`
- R의 자료형 - 벡터 => `a <- c(값1, 값2,...)`
- csv 읽기- 객체 <- `read.csv("파일명.csv")`
- text 읽기- 객체 <- `read.table("파일명.txt", sep=",")`

5-1 기본 함수 알아두기

(가) `is.vector()`

- `is.vector()` : 벡터인가 아닌가? (True: 벡터이다, False: 벡터가 아니다.)

함수	설명
<code>is.vector()</code>	벡터인가?아닌가?
<code>is.factor()</code>	팩터인가?아닌가?
<code>is.double()</code>	double인가?아닌가?
<code>is.data.frame()</code>	데이터프레임인가? 아닌가?
<code>is.integer()</code>	정수인가?아닌가?
<code>is.numeric()</code>	수치형인가?아닌가?
<code>is.NA()</code>	NA인가? 아닌가? NA:결측값
<code>is.null()</code>	null인가 아닌가?
<code>is.nan()</code>	nan인가? 아닌가? nan:수학적으로 정의되지 않은 값
<code>is.infinite()</code>	무한대인가? 아닌가?

In [1]:

```
a <- c(1,2,3,5,6)
is.vector(a)
```

TRUE

(나) 값의 선택, 삭제, 수정, 추가

In [2]:

```
value1 <- c(1,2,3)
value2 <- c(4,5,6)
```

In [3]:

```
new_val <- c(value1, value2)
new_val
```

1 2 3 4 5 6

값의 선택

In [4]:

```
new_val[2]      # 2번째 선택
new_val[3:4]    # 3,4번째 선택
new_val[3:length(new_val)] # 3번째부터 맨 마지막까지
```

2

3 4

3 4 5 6

In [5]:

```
new_val
```

1 2 3 4 5 6

값의 삭제

- 데이터[-위치] : 해당되는 위치의 값을 삭제한다.

In [6]:

```
new_val[-2]      # 2번째 값 삭제
```

1 3 4 5 6

In [7]:

```
new_val
```

1 2 3 4 5 6

In [8]:

```
new_val2 <- new_val[-6] # 6번째 값 삭제  
new_val2
```

1 2 3 4 5

값의 변경

In [9]:

```
new_val <- c(1,2,3,5,6)
```

In [10]:

```
new_val[3] = 30  
new_val
```

1 2 30 5 6

In [11]:

```
### 2의 배수를 200으로 변경  
new_val[new_val %% 2==0]
```

2 30 6

In [12]:

```
### 2의 배수를 200으로 변경  
new_val[new_val %% 2==0] = 200  
new_val
```

1 200 200 5 200

실습과제 05-01

- value를 seq를 이용하여 100이하 짝수의 값을 갖도록 하고, 이 중에서 5의 배수를 1000으로 변경하도록 해보자.

In []:

(다) 변수 확인 및 삭제

In [27]:

```
ls() # 변수 리스트 확인
```

'a' 'new_val' 'new_val2' 'value1' 'value2'

In [28]:

```
ls.str() # 변수 리스트 자세히 보기
```

```
a : num [1:5] 1 2 3 5 6
new_val : num [1:5] 1 200 200 5 200
new_val2 : num [1:5] 1 2 3 4 5
value1 : num [1:3] 1 2 3
value2 : num [1:3] 4 5 6
```

변수 삭제

- rm(변수 또는 객체명)

In [29]:

```
rm(new_val2)
ls.str()
```

```
a : num [1:5] 1 2 3 5 6
new_val : num [1:5] 1 200 200 5 200
value1 : num [1:3] 1 2 3
value2 : num [1:3] 4 5 6
```

전체 변수 삭제

In [31]:

```
rm(list=ls())
ls.str()
```

(라) 패키지 설치, 불러오기, 확인

In [32]:

```
# 패키지 설치 확인
library()
```

In [33]:

```
# 패키지 설치 확인(버전)
installed.packages()
```

Package		LibPath	Version	Priority	Depends	Imports	Linking
abind	abind	C:/Anaconda3/Lib/R/library	1.4-5	NA	R (>= 1.5.0)	methods, utils	NA
assertthat	assertthat	C:/Anaconda3/Lib/R/library	0.2.0	NA	NA	tools	I
backports	backports	C:/Anaconda3/Lib/R/library	1.1.2	NA	R (>= 3.0.0)	utils	NA
base	base	C:/Anaconda3/Lib/R/library	3.5.1	base	NA	NA	I
base64enc	base64enc	C:/Anaconda3/Lib/R/library	0.1-3	NA	R (>= 2.9.0)	NA	NA
BH	BH	C:/Anaconda3/Lib/R/library	1.66.0-1	NA	NA	NA	I
bindr	bindr	C:/Anaconda3/Lib/R/library	0.1.1	NA	NA	NA	I
bindrcpp	bindrcpp	C:/Anaconda3/Lib/R/library	0.2.2	NA	NA	bindr (>= 0.1.1), Rcpp (>= 0.12.16)	pl Rc
					R (>= 3.0.0),		

In [34]:

```
# 패키지 설치 확인(버전)
installed.packages()[ , c("Package", "Version")]
```

	Package	Version
abind	abind	1.4-5
assertthat	assertthat	0.2.0
backports	backports	1.1.2
base	base	3.5.1
base64enc	base64enc	0.1-3
BH	BH	1.66.0-1
bindr	bindr	0.1.1
bindrcpp	bindrcpp	0.2.2
boot	boot	1.3-20
broom	broom	0.5.0
callr	callr	2.0.4
caret	caret	6.0-80
cellranger	cellranger	1.1.0
checkpoint	checkpoint	0.4.4
class	class	7.3-14
cli	cli	1.0.0
clipr	clipr	0.4.1
cluster	cluster	2.0.7-1
codetools	codetools	0.2-15
colorspace	colorspace	1.3-2
compiler	compiler	3.5.1
crayon	crayon	1.3.4
curl	curl	3.2
CVST	CVST	0.2-2
data.table	data.table	1.11.4
datasets	datasets	3.5.1
DBI	DBI	1.0.0
dbplyr	dbplyr	1.2.2
ddalpha	ddalpha	1.3.4
DEoptimR	DEoptimR	1.0-8
...
sqlrutils	sqlrutils	1.0.0
SQUAREM	SQUAREM	2017.10-1
stats	stats	3.5.1
stats4	stats4	3.5.1
stringi	stringi	1.1.7

	Package	Version
stringr	stringr	1.3.1
survival	survival	2.42-3
tcltk	tcltk	3.5.1
testthat	testthat	2.0.0
tibble	tibble	1.4.2
tidyr	tidyr	0.8.1
tidyselect	tidyselect	0.2.4
tidyverse	tidyverse	1.2.1
timeDate	timeDate	3043.102
tinytex	tinytex	0.6
tools	tools	3.5.1
translations	translations	3.5.1
TTR	TTR	0.23-3
utf8	utf8	1.1.4
utils	utils	3.5.1
uuid	uuid	0.1-2
viridisLite	viridisLite	0.3.0
whisker	whisker	0.3-2
withr	withr	2.1.2
xfun	xfun	0.3
xml2	xml2	1.2.0
xtable	xtable	1.8-2
xts	xts	0.11-0
yaml	yaml	2.2.0
zoo	zoo	1.8-3

Github로 패키지 설치

- (1) `install.packages("devtools")`
- (2) `library(devtools)`
- (3) `install_github("yihui/knitr")`

로컬에서 설치

- (1) `install.packages("devtools")`
- (2) `library(devtools)`
- (3) `install("c:/knitr-1.9")`

R에서 도움말 보기

- `help("명령어")`

- ?명령어

(마) 파일 읽기

- read.csv : csv파일 (, 구분)
- read.table : txt 파일의 포맷 읽기 (tab 구분)
- read_excel : 엑셀 파일 읽기 (readxl 패키지)
- read.spss : spss 파일 읽기 (foreign 패키지)