

# 05 데이터 처리 기본

## 학습 내용

\*

### 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")`

## 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 [12]:

```
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 [13]:

```
new_val
```

1 2 3 4 5 6

## 값의 삭제

- 데이터[-위치] : 해당되는 위치의 값을 삭제한다.

In [14]:

```
new_val[-2]      # 2번째 값 삭제
```

1 3 4 5 6

In [15]:

```
new_val
```

1 2 3 4 5 6

In [17]:

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

1 2 3 4 5

## 값의 변경

In [22]:

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

In [23]:

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

1 2 30 5 6

In [24]:

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

2 30 6

In [25]:

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

1 200 200 5 200

## (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
<b>abind</b>	abind	C:/Anaconda3/Lib/R/library		1.4-5	NA	R (>= 1.5.0)	methods, utils	NA
<b>assertthat</b>	assertthat	C:/Anaconda3/Lib/R/library		0.2.0	NA	NA	tools	NA
<b>backports</b>	backports	C:/Anaconda3/Lib/R/library		1.1.2	NA	R (>= 3.0.0)	utils	NA
<b>base</b>	base	C:/Anaconda3/Lib/R/library		3.5.1	base	NA	NA	NA
<b>base64enc</b>	base64enc	C:/Anaconda3/Lib/R/library		0.1-3	NA	R (>= 2.9.0)	NA	NA
<b>BH</b>	BH	C:/Anaconda3/Lib/R/library		1.66.0-1	NA	NA	NA	NA
<b>bindr</b>	bindr	C:/Anaconda3/Lib/R/library		0.1.1	NA	NA	NA	NA
<b>bindrcpp</b>	bindrcpp	C:/Anaconda3/Lib/R/library		0.2.2	NA	NA	bindr (>= 0.1.1), Rcpp (>= 0.12.16)	NA
						R (>= 3.0.0),		

In [34]:

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

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

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