

08. R 기본 Review 및 기본 ¶

학습 내용

- 자료형 알아보기
- 기본 내용 정리

01 R에서 작업을 하면 .RData가 만들어진다.

- 파일명.RData을 정하고 파일을 열어 이어서 작업이 가능하다.

RData 파일을 저장하기

- `save.image("파일이름")`

In [4]:

```
getwd()
```

'C:/Users/WITHJS/Documents/GitHub/RBasic/RLevelUp_Source'

In [5]:

```
save.image('test.RData')
```

내 PC > 문서 > GitHub > RBasic > RLevelUp_Source >

이름

.ipynb_checkpoints

test.RData

Untitled.ipynb

수정한 날짜

2019-08-22 오후 12:00

2019-08-22 오후 12:03

2019-08-22 오후 12:02

유형

파일 폴더

R Workspace

IPYNB 파일

크기

R에서 가장 작은 단위 벡터

In [6]:

```
a <- 1  
is.vector(a)
```

TRUE

In [7]:

```
a <- c(1,2,3,4,"Hello", "World")
is.vector(a)
```

TRUE

In [8]:

```
ls()
ls.str()
```

'a'

a : chr [1:6] "1" "2" "3" "4" "Hello" "World"

In [9]:

```
rm(a)
rm(list=ls())
```

패키지 설치 확인

In [10]:

```
library()
```

R packages available

Packages in library 'C:/Anaconda3/Lib/R/library':

abind	Combine Multidimensional Arrays
assertthat	Easy Pre and Post Assertions
backports	Reimplementations of Functions Introduced Since R-3.0.0
base	The R Base Package
base64enc	Tools for base64 encoding
BH	Boost C++ Header Files
bindr	Parametrized Active Bindings
bindrcpp	An 'Rcpp' Interface to Active Bindings
boot	Bootstrap Functions (Originally by Angelo Canty for S)
broom	Convert Statistical Analysis Objects into Tidy

도움말

- ?명령어
- help('명령어')

데이터 읽어오기

- read.csv("위치,파일명", header=[False or True], fileEncoding=["eur-kr" etc].)
- read.table("위치,파일명", header=T, sep=",")
- read_excel(path="파일위치~파일명", sheet="시트명", col_names=TRUE)
- read.spss("위치~파일명")

데이터 자료 확인

is.[자료형](#)([자료명](#)) (예) is.data.frame(Score)

데이터 자료 변경

as.factor(자료명)

In [11]:

```
a <- c("사과", "바나나", "딸기")
is(a)
```

'character' 'vector' 'data.frameRowLabels' 'SuperClassMethod'

In [12]:

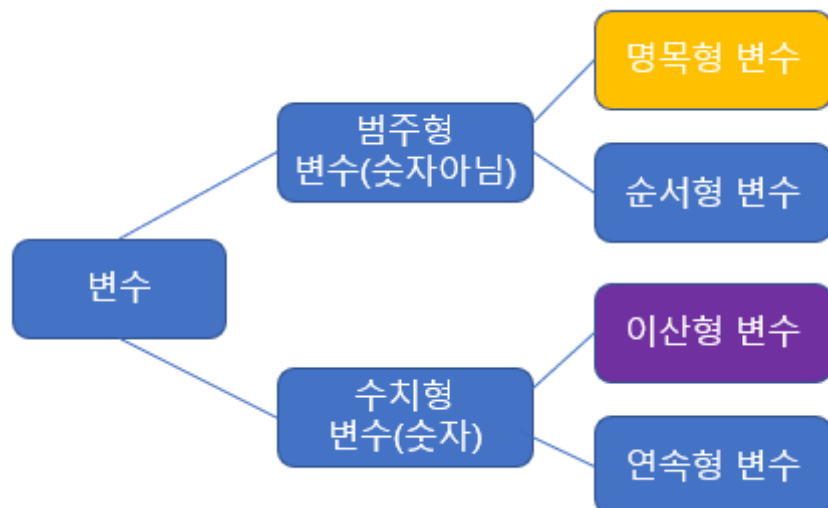
```
af <- as.factor(a)
af
```

사과 바나나 딸기

► Levels:

데이터의 종류

- 변수는 범주형 변수와 수치형 변수로 구분된다.
- 범주형 변수는 명목형 변수와 순서형 변수로 구분된다.
- 수치형 변수는 이산형 변수와 연속형 변수로 구분된다.



- 이산형 변수 : 셀수있는 것이 특징
- 연속형 : 어떤 구분을 사이에 두고 셀수없는 많은 값이 존재
- 명목형 : 숫자가 아니라 문자로 나타내지는 변수값(남, 녀), 혈액형(O, AB, B, A) -> 카테고리를 정해 분류
- 순서형 : 순서형 데이터는 명목형과 비슷하지만 순서가 있다.

R에서의 변수 타입

- character : 문자열 나타내는 타입
- factor : 순서형 변수, 명목형 변수
- integer : 정수값 표현
- numeric : 연속형 변수
- data.frame : 데이터 프레임
- list

In [19]:

```
a1 <- c(1L,2L,3L) # 정수형 벡터  
is(a1)
```

'integer' 'numeric' 'vector' 'data.frameRowLabels'

In [20]:

```
a1 <- c(1,2,3) # 수치형 벡터  
is(a1)
```

'numeric' 'vector'

In [27]:

```
a_str <- c("하나", "둘", "셋") # 문자열 벡터  
is(a1)
```

'factor' 'integer' 'oldClass' 'numeric' 'vector' 'data.frameRowLabels'

In [28]:

```
a1 <- factor(a_str) # 범주형 변수  
is(a1)
```

'factor' 'integer' 'oldClass' 'numeric' 'vector' 'data.frameRowLabels'

In [31]:

```
a1 <- factor(a_str, ordered=T) # 범주형-순서형 변수  
is(a1)
```

'ordered' 'factor' 'oldClass'

In [32]:

```
a1 <- factor(a_str, levels=c("하나", "둘", "셋")) # 범주형-순서형 변수  
is(a1)
```

'factor' 'integer' 'oldClass' 'numeric' 'vector' 'data.frameRowLabels'

- a1 실행 결과

[1] 하나 둘 셋
Levels: 하나 둘 셋

이후에 알아보기

- merge 함수 정리
- list 객체 정리
- rbind() 함수, cbind() 함수
- prop.table()

In []: