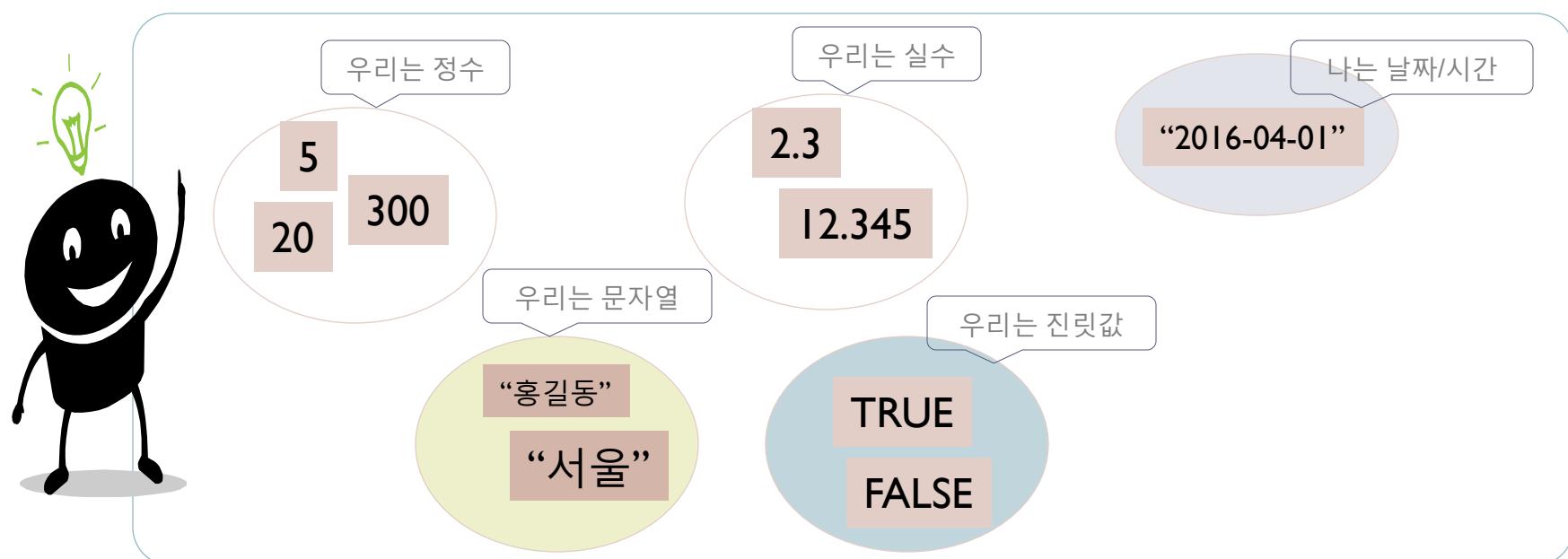
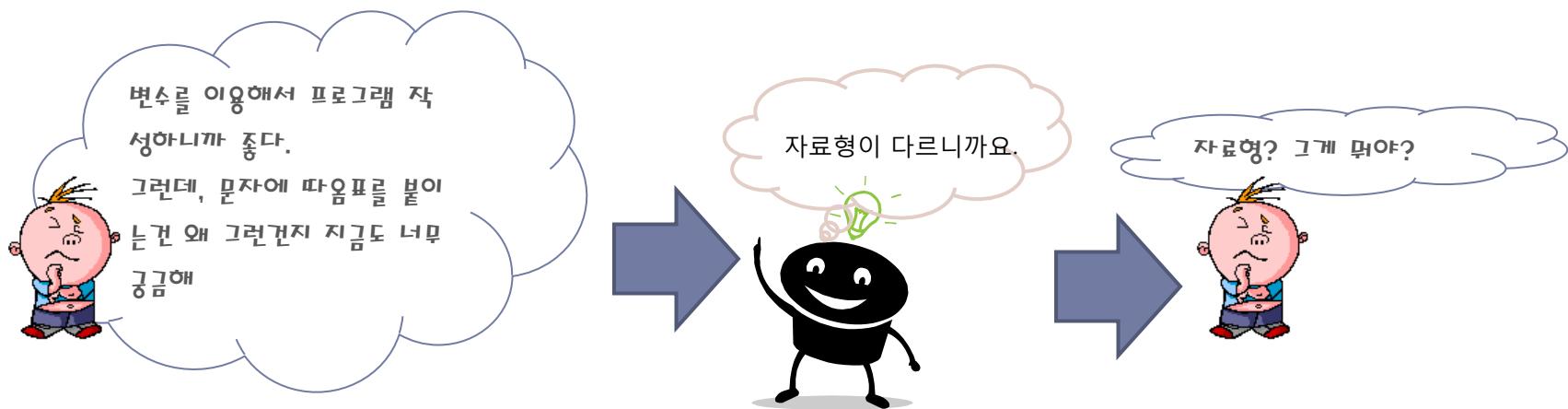


2. 자료 타입

R의 자료형(TYPE)에 대하여 단순 기본타입 및
factor 타입에 대해 학습합니다.



자료형(data type)이란?



```
> x<-5  
> y<-3.14  
> x+y  
[1] 8.14  
> |  
> x>y  
[1] TRUE  
> x=y  
> x==y  
[1] TRUE  
> x<-x-3  
> x>y  
[1] FALSE  
> |
```



두 변수 값을 비교한 결과 당
연히 x 값이 크므로 “참
(true)” 이 됩니다.

> Sys.Date() 시스템 날짜를 표시하기 위한 Sys.Date() 함수

```
[1] "2016-03-30"  
> a1<-"2016-03-30"  
> as.Date(a1)+100  
[1] "2016-07-08"  
> |
```

as.Date() 메소드는 문자열 데이터값을 날짜형으로
변환해줍니다.

NA 데이터 타입(또는 특별한 상수)

- Not Available → 데이터 값이 없다(빠져있다)? → 결측값, 모른다? 의미로 해석해요



여러분이 5명의 키를 입력받았다고 가정해요.
그런데, 그 중 마지막 사람의 키를 모르는 경우 그림처럼 입력해요



```
> a1<-177  
> a2<-180  
> a3<-168  
> a4<-169  
> a5<-NA  
> is.na(a5)  
[1] TRUE
```

is.na() 함수 이용하여
변수에 “NA”값이 있는지
확인합니다.

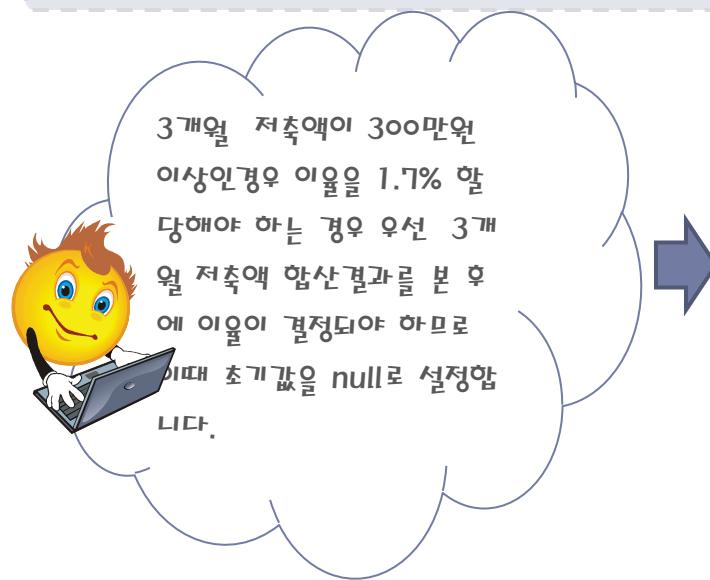


또는 벡터형으로 더 간단히 표현해요

```
> height<-c(177,180,168,169,NA)  
> height  
[1] 177 180 168 169 NA  
> |
```

NULL 데이터 타입

- 데이터 값이 정해지지 않은 상태(undefined)를 의미



```
> rate<-NULL  
> money<-c(1200000,950000,1300000)  
> if(sum(money)>=3000000) {  
+ rate<-0.017  
+ }  
>
```

rate 초기값을 null로 설정

3개월 저축액을 벡터형으로 처리하여 money에 할당

money의 합을 계산하여 300만원 이상이면 rate에 값을 할당


```
> is.null(rate)  
[1] FALSE  
>
```

rate값이 현재 null값인지 확인
null <-TRUE not null <-FALSE

Factor 타입

- 범주형 데이터 – 변수가 취할 수 있는 값이 범주(category)로 주어지는 경우를 의미
- 종류
 - ✓ 명목형 데이터 – 종교 (천주교, 기독교, 불교), 의결(찬성, 반대)
 - ✓ 순서형 데이터 – 크기(대, 중, 소)

Factor 형 관련 함수

1. factor()

x, ←팩터로 표현하고자 하는 값

levels, <-팩터 레벨의 목록을 보여줌

ordered ←TRUE(순서형 데이터), FALSE(명목형 데이터)

)

2. table(x) – level에 따라 몇 개의 값이 존재하는지 count값 출력

3. unclass(x) – 목록의 단순 상수값만을 출력하고, 해당 상수가 의미하는 데이터값을 표시



b명의 종교를 그림처럼 코딩
했습니다. 벡터형으로요..



```
> religion<-c("천주교", "기독교", "기독교", "불교", "불교", "무교")
> religion
[1] "천주교" "기독교" "기독교" "불교" "불교" "무교"
> |
```

입력한 6개 데이터가 모두 표시됩니다.(중복이 되는거죠)



그렇다면, 입력된 데이터의
중복을 제거하여 목록을 만들
수 있어요. 이때 factor타입
이 필요한 거예요.

```
> religion<-factor(c("천주교", "기독교", "기독교", "불교", "불교", "무교"))
> religion
[1] 천주교 기독교 기독교 불교 불교 무교
Levels: 기독교 무교 불교 천주교
> |
```

level인자 생략하면 자동으로 레벨 목록 파악하여 표시

```
> religion<-factor(c("천주교", "천주교", "기독교", "불교", "불교", "기타"), c("천주교", "기독교", "불교", "기타"))
> religion
[1] 천주교 천주교 기독교 불교 불교 기타
Levels: 천주교 기독교 불교 기타
> |
```

factor요소

목록을 표시

순서형 데이터의 factor 생성시 ordered인자에 TRUE값을 설정해줘야 합니다.

```
> width<-factor(c("18평", "24평", "24평", "33평", "33평", "33평"), ordered=TRUE)
> width
[1] 18평 24평 24평 33평 33평 33평
Levels: 18평 < 24평 < 33평
> |
```

목록의 형태가 다르죠?