



Super Computación y Cálculo Científico

Estructuras de Datos en R

Gilberto Díaz

gilberto.diaz@uis.edu.co

Universidad Industrial de Santander
Bucaramanga - Colombia



R puede cargar instrucciones desde un archivo texto

```
source("camino/script.R")
```

También, las puede ejecutar desde la línea de comandos

```
R -f script.R
```



R puede manejar los valores VERDADERO (TRUE) y FALSO (FALSE)

Estos sirven para indicar que algún otro valor asociado puede ser verdadero o falso según convenga.

consentimiento `c(FALSE, TRUE, FALSE, TRUE)`



R maneja vectores de cadenas de caracteres

Estos sirven para manejar nombres por ejemplo

nombre `c("Andrea", "Yanet", "Manuel", "Juan")`

Un vector en R sólo puede manejar un sólo tipo de dato

edad `c(16, 34, 43, 50)`

Si se combinan número y caracteres, todo será tratado como caracteres



Es común que se necesite categorizar datos según un aspecto determinado. R permite hacer esto con la función **factor**

```
sexo c("Femenino", "Femenino", "Masculino",  
"Masculino")
```

```
factor(sexo)
```

```
[1] Femenino Femenino Masculino Masculino
```

```
Levels: Femenino Masculino
```

R convierte las cadenas de caracteres en categorías de los datos. El tipo de dato es **level**



Para conocer el tipo de dato almacenado en un vector se utiliza la función class

```
class(consentimiento)  
[1] "logical"
```

```
class(edad)  
[1] "numeric"
```

```
class(nombre)  
[1] "character"
```



R maneja los vectores como la estructura de almacenamiento de datos principal.

Sin embargo, frecuentemente se utiliza los **Data Frames**, los cuales son un conjunto de observaciones de un conjunto de variables. Es decir, los **resultados de un experimento**



Ejemplo:

Se requiere analizar información sobre un conjunto de pacientes. Podemos crear una estructura para organizar la información

```
pacientes = data.frame(nombre, consentimiento,  
sexo, edad)
```



Se puede acceder directamente a las variables con el siguiente operador

```
pacientes$edad
```

Se puede asignar nombres a las variables del data frame

```
names(pacientes) = c("Nombre", "Consentimiento",  
"Sexo", "Edad")
```



Para eliminar una variable utilice

```
rm(edad)
```

Para eliminar varias variables

```
rm(list = c("edad","nombre"))
```

Para agregar una variable

```
pacientes$ID = paste("No", 1:4)
```