

Práctica de R

Probabilidad y Estadística

Ingeniería Industrial - Facultad de Ciencias Agrarias

Tipos de datos y Estructura de datos

Tipo de datos

En R, todo son objetos

Estos objetos se crean a partir de objetos más simples.

Los objetos más simples se llaman atómicos:

- `character`(cadena de caracteres)
- `numeric` (números reales)
- `integer`(números enteros)
- `complex` (números complejos)
- `logical` (lógicos o booleanos, toman valores de `TRUE` o `FALSE`)

Estructura de datos

Vectores

Los vectores son arreglos ordenados en los cuales se puede almacenar información de tipo numérico (variable cuantitativa), alfanumérico (variable cualitativa) o lógico (`TRUE` o `FALSE`). La función de R para crear un vector es `c()` y que significa concatenar; dentro de los paréntesis de esta función se ubica la información a almacenar. Una vez construido el vector se acostumbra a etiquetarlo con un nombre corto y representativo de la información que almacena, la asignación se hace por medio del operador `<-` entre el nombre y el vector.

- Vector de enteros

```
c(3,2,5)
```

```
## [1] 3 2 5
```

```
numeros<-c(3,2,5)
numeros
```

```
## [1] 3 2 5
```

- Vector de decimales

```
c(3.2,2.4,5.8)
```

```
## [1] 3.2 2.4 5.8
```

- Vector de caracteres

```
c("Ana","Sofía","Pedro","Juan")
```

```
## [1] "Ana" "Sofía" "Pedro" "Juan"
```

- Vector de booleanos

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

```
## [1] TRUE FALSE TRUE FALSE
```

Matrices

Las matrices son arreglos rectangulares de filas y columnas con información numérica, alfanumérica o lógica. Para construir una matriz se usa la función **matrix()**. Por ejemplo, para crear una matriz de 4 filas y 5 columnas (de dimensión 4×5) con los primeros 20 números positivos se escribe el código siguiente en la consola.

```
mimatriz <- matrix(data=1:20, nrow=4, ncol=5)
mimatriz
```

```
##      [,1] [,2] [,3] [,4] [,5]
## [1,]    1    5    9   13   17
## [2,]    2    6   10   14   18
## [3,]    3    7   11   15   19
## [4,]    4    8   12   16   20
```

Dataframes o Marco de datos

Data.frame es uno de los objetos más utilizados en R, porque permite agrupar vectores con información de diferente tipo (numérica, alfanumérica o lógica) en un mismo objeto, la única restricción es que los vectores deben tener la misma longitud. Para crear un marco de datos se usa la función **data.frame()**

- Paso 1: creamos los vectores

```
#Creamos en primer lugar tres vectores
alumnos<-c("Ana","Sofía","Luis", "Pedro","Juan")
edad<-c(15, 19, 13, 18, 20)
deporte<-c(TRUE, TRUE, FALSE, FALSE, TRUE)
comic.fav<-c('Superman', 'Superman', 'Batman', 'Spiderman', 'Batman')
```

- Paso 2: creamos un data.frame a partir de los vectores

```
#Usamos la función data.frame
mimarco <- data.frame(alumnos, edad, deporte, comic.fav)
mimarco
```

```
##  alumnos edad deporte comic.fav
## 1     Ana   15    TRUE  Superman
## 2   Sofía   19    TRUE  Superman
## 3     Luis   13   FALSE   Batman
## 4   Pedro   18   FALSE Spiderman
## 5     Juan   20    TRUE   Batman
```

Bibliografía

- Hernández, Freddy; Usuga, Olga (2019) Manual de R. Disponible en <https://fhernanb.github.io/Manual-de-R/>
- Santana, Julio Sergio; Farfán, Efraín Mateos (2014) El arte de programar en R. Un lenguaje para la estadística. https://cran.r-project.org/doc/contrib/Santana_El_arte_de_programar_en_R.pdf