

Tema 3 - Estructuras de datos con R

Curso Estadística Descriptiva

2022-10-05

Factores

Un Factor es similar al vector, con una estructura interna más rica que permite clasificar las observaciones.

- **levels:** atributo del factor. Cada elemento del factor es igual a un nivel. Los niveles clasifican las entradas del factor. Se ordenan por orden alfabético
- Para definir un factor, definimos un vector y lo transformamos con la función `factor()` o `as.factor()`

La función `factor()`

- `factor(vector, levels=...)`: define un factor a partir del vector y dispone de algunos parámetros que permiten modificar el factor que se crea:
 - **levels:** permite especificar los niveles e incluso añadir niveles que no aparecen en el vector
 - **labels:** permite cambiar los nombres de los niveles
- `levels(factor)`: para obtener los niveles del factor

Cuando tenemos un vector y lo queremos utilizar para clasificar los datos, es conveniente utilizar los factores.

```
nombres = c("Pedro", "Juan", "Carlos", "Diego", "Sergio", "Pedro", "Diego")
nombres.factor = factor(nombres)
nombres.factor
```

```
## [1] Pedro Juan Carlos Diego Sergio Pedro Diego
## Levels: Carlos Diego Juan Pedro Sergio
```

En contraste con `as.factor()`, con `factor` podemos definir niveles aún cuando no aparecen en el vector.

```
g = c("H", "M", "M", "H", "M")
g.factor <- factor(g, levels = c("H", "M", "B"))

g.flabel <- factor(g, levels = c("H", "M", "B"), labels = c("Hombre", "Mujer", "Hermafrodita"))
```

Factores Ordenados

Factor ordenado. Es un factor donde los niveles siguen un orden

- `ordered(vector, levels=...)`: función que define un factor ordenado y tiene los mismos parámetros que `factor`

```
fac = factor(c(1,1,1,2,2,3,2,4,1,3,3,4,2,3,4,4),
            levels = c(1,2,3,4), labels = c("Sus","Apr","Not","Exc"))
fac
```

```
[1] Sus Sus Sus Apr Apr Not Apr Exc Sus Not Not Exc Apr Not Exc Exc
Levels: Sus Apr Not Exc
```

```
facOrd = ordered(c(1,1,1,2,2,3,2,4,1,3,3,4,2,3,4,4),
                 levels = c(1,2,3,4), labels = c("Sus","Apr","Not","Exc"))
facOrd
```

```
[1] Sus Sus Sus Apr Apr Not Apr Exc Sus Not Not Exc Apr Not Exc Exc
Levels: Sus < Apr < Not < Exc
```

Figure 1: factores ordenados y no ordenados

Listas

List

List. Lista formada por objetos no necesariamente del mismo tipo, cada uno con un nombre interno

- `list()`. Crea una lista
 - Para obtener una componente concreta usamos la instrucción `list$componente`
 - También podemos indicar el objeto por su posición usando dobles corchetes: `list[[i]]`. Lo que obtendremos es una list formada por esa única componente, no el objeto que forma la componente

```
x = c(1,2,5,8, 9, 10, -12, -21)
```

```
L = list(name = "Temperaturas", data = x, mean = mean(x), sums = cumsum(x))
```

```
L
```

```
## $name
## [1] "Temperaturas"
##
## $data
## [1]  1  2  5  8  9 10 -12 -21
##
## $mean
## [1] 0.25
##
## $sums
## [1]  1  3  8 16 25 35 23  2
```

```
#Para acceder a las variables o componentes de la lista.
```

```
L$name
```

```
## [1] "Temperaturas"
```

```
L$mean
```

```
## [1] 0.25
```

Obtener información de una lista

- `str(list)`: para conocer la estructura interna de una list
- `names(list)`: para saber los nombres de la list

```
#names  
names(L)
```

```
## [1] "name" "data" "mean" "sums"
```