

# Estructura de datos en R - Factores

Ramon Ceballos

16/1/2021

## FACTORES EN R

Factor: es como un vector, pero con una estructura interna más rica que permite usarlo para clasificar 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, primero hemos de definir un vector y trasformarlo por medio de una de las funciones `factor()` o `as.factor()`.

```
nombres = c("Maria", "Juan", "Jose", "Maria", "Juan")
nombres.factor <- factor(nombres)
nombres.factor
```

```
## [1] Maria Juan  Jose  Maria Juan
## Levels: Jose Juan Maria
```

Los factores sirven para catalogar algún elemento.

### 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

```
gender = c("M", "H", "H", "H", "M", "M", "M", "M")
gender.factor=factor(gender)
gender.factor
```

```
## [1] M H H H M M M M
## Levels: H M
```

```
#as.factor() define también un factor pero no puede incluir niveles extras
gender.factor2 = as.factor(gender)
gender.factor2
```

```
## [1] M H H H M M M M
## Levels: H M
```

```
#Para incluir niveles extras se emplea factor()
gender.factor3=factor(gender, levels =c("M","H","B"))
gender.factor3
```

```
## [1] M H H H M M M M
## Levels: M H B
```

```
#Para aclarar el significado de los niveles se emplea labels
gender.factor4 <- factor(gender, levels =c("M","H","B"), labels = c("Mujer","Hombre","Hermafrodita"))
gender.factor4
```

```
## [1] Mujer  Hombre Hombre Hombre Mujer  Mujer  Mujer  Mujer
## Levels: Mujer Hombre Hermafrodita
```

```
#Consultar los niveles de un factor
levels(gender.factor4)
```

```
## [1] "Mujer"      "Hombre"      "Hermafrodita"
```

```
#Asignar niveles de un factor
levels(gender.factor4) = c("Femenino","Masculino","Hibrido")
gender.factor4
```

```
## [1] Femenino Masculino Masculino Masculino Femenino Femenino Femenino
## [8] Femenino
## Levels: Femenino Masculino Hibrido
```

Hasta ahora hemos visto los factores simples donde el orden de los factores no sería relevante. Este aspecto es diferente para los factores ordenados.

## Factor ordenado

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
```