

Clase 1.3

Acomodo de datos

Marcos Rosetti y Luis Pacheco-Cobos

Estadística y Manejo de Datos con R (EMDR) — Virtual

Acomodo de datos: **forcats**



Acomodo de datos: **forcats**

- Una herramienta para resolver problemas relacionados con factores, como cambiar el orden de los niveles o los niveles mismos.

```
library(forcats)
```

```
data(starwars, package = "dplyr")  
?starwars  
head(starwars)
```

Acomodo de datos: forcats

- `fct_infreq()` reordena un factor por la frecuencia de sus valores.

```
starwars %>%  
  mutate(eye_color = fct_infreq(eye_color))
```

```
## # A tibble: 87 × 14  
##   name      height mass hair_color skin_color eye_color birth_year sex  gender  
##   <chr>    <int> <dbl> <chr>    <chr>    <fct>    <dbl> <chr> <chr>  
## 1 Luke S...   172    77 blond    fair      blue      19    male  masculi  
## 2 C-3PO      167    75 <NA>     gold      yellow    112   none  masculi  
## 3 R2-D2      96     32 <NA>     white, bl... red        33    none  masculi  
## 4 Darth ...  202   136 none     white     yellow    41.9  male  masculi  
## 5 Leia O...  150    49 brown    light     brown     19    fema.. feminin  
## 6 Owen L...  178   120 brown, grey light     blue      52    male  masculi  
## 7 Beru W...  165    75 brown    light     blue      47    fema.. feminin  
## 8 R5-D4      97     32 <NA>     white, red red        NA    none  masculi  
## 9 Biggs ...  183    84 black    light     brown     24    male  masculi  
## 10 Obi-Wa...  182    77 auburn, wh... fair      blue-gray  57    male  masculi  
## # ... with 77 more rows, and 5 more variables: homeworld <chr>, species <chr>,  
## #   films <list>, vehicles <list>, starships <list>
```

Acomodo de datos: forcats

- `fct_lump()` colapsa los valores menos frecuentes en un factor de “otros”.

```
starwars %>%  
  filter(!is.na(species)) %>%  
  mutate(species = fct_lump(species, n = 3)) %>%  
  count(species)
```

```
## # A tibble: 4 × 2  
##   species      n  
##   <fct>    <int>  
## 1 Droid      6  
## 2 Gungan     3  
## 3 Human     35  
## 4 Other     39
```

Acomodo de datos: forcats

- `fct_relevel()` reordena los niveles de los factores.

```
swars <- starwars %>% filter(!is.na(gender)) %>% mutate(gender=factor(gender))  
levels(swars$gender)
```

```
## [1] "feminine" "masculine"
```

```
swars$gender <- fct_relevel(swars$gender, "female", after = 2)
```

```
## Warning: Unknown levels in `f`: female
```

```
levels(swars$gender)
```

```
## [1] "feminine" "masculine"
```

```
swars$gender <- fct_relevel(swars$gender, sort)  
levels(swars$gender)
```

```
## [1] "feminine" "masculine"
```

Licencia CC BY



Estadística y Manejo de Datos con R (EMDR) por Marcos F. Rosetti S. y Luis Pacheco-Cobos se distribuye bajo una [Licencia Creative Commons Atribución 4.0 Internacional](https://creativecommons.org/licenses/by/4.0/).