

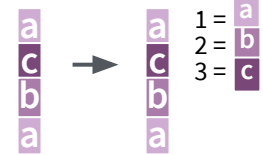
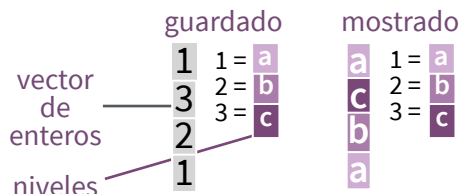
Factores con forcats :: HOJA DE REFERENCIA



El paquete **forcats** provee herramientas para trabajar con factores, la estructura de datos de R para datos categóricos.

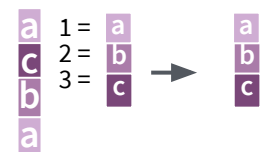
Factores

R representa datos categóricos con factores. Un **factor** es un vector de enteros con un atributo de **niveles** que guarda mapeos entre enteros y valores categóricos. Cuando ves un factor, R no muestra enteros, sino los valores asociados con ellos.



Crear un factor con factor()
factor(x = character(), levels, labels = levels, exclude = NA, ordered = is.ordered(x), nmax = NA) Convierte un vector a un factor. También **as_factor**.
f <- factor(c("a", "b", "c", "a"),

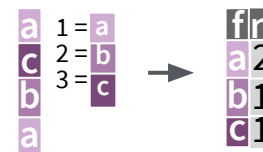
levels = c("a", "b", "c"))



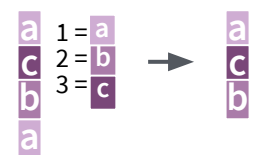
Devolver sus niveles con levels()
levels(x) Devuelve/fija los niveles de un factor. **levels(f); levels(f) <- c("x", "y", "z")**

Usa **unclass()** para ver su estructura

Examina Factores

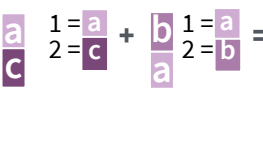


fct_count(f, sort = FALSE)
Cuenta el número de valores de cada nivel. **fct_count(f)**

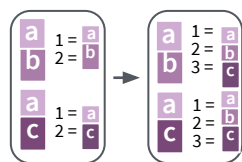


fct_unique(f) Devuelve los valores únicos, removiendo duplicados. **fct_unique(f)**

Combina Factores

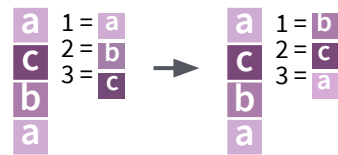


fct_c(...) Combina factores con niveles distintos.
f1 <- factor(c("a", "c"))
f2 <- factor(c("b", "a"))
fct_c(f1, f2)



fct_unify(fs, levels = lvs_union(fs)) Estandariza niveles en una lista de factores.
fct_unify(list(f2, f1))

Cambia el orden de los niveles



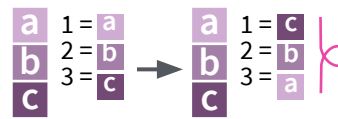
fct_relevel(f, ..., after = 0L)
Reordena niveles manualmente.
fct_relevel(f, c("b", "c", "a"))



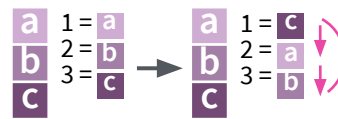
fct_infreq(f, ordered = NA)
Reordena niveles por la frecuencia con la que aparecen en los datos (mayor frecuencia primero).
f3 <- factor(c("c", "c", "a"))
fct_infreq(f3)



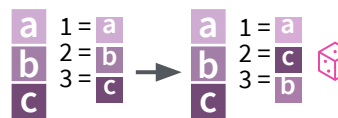
fct_inorder(f, ordered = NA)
Reordena niveles por el orden en el cual aparecen en los datos.
fct_inorder(f2)



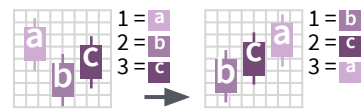
fct_rev(f) Ordena los niveles en reversa.
f4 <- factor(c("a", "b", "c"))
fct_rev(f4)



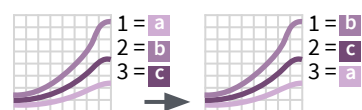
fct_shift(f) Desplaza niveles a izquierda o derecha en forma envolvente.
fct_shift(f4)



fct_shuffle(f, n = 1L) Permuta el orden de los niveles al azar.
fct_shuffle(f4)

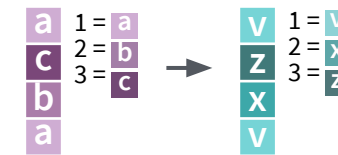


fct_reorder(f, x, .fun = median, ..., .desc = FALSE) Reordena los niveles según su asociación con otra variable.
boxplot(data = iris, Sepal.Width ~ fct_reorder(Species, Sepal.Width))

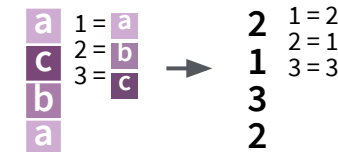


fct_reorder2(f, x, y, .fun = last2, ..., .desc = TRUE) Reordena niveles por su valor final cuando se grafican con otras dos variables.
ggplot(data = iris, aes(Sepal.Width, Sepal.Length, color = fct_reorder2(Species, Sepal.Width, Sepal.Length))) + geom_smooth()

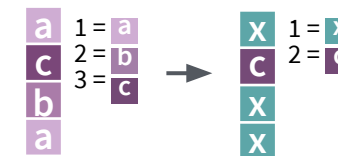
Cambia el valor de los niveles



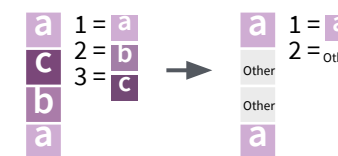
fct_recode(f, ...) Cambia manualmente los niveles. También **fct_relabel** que tiene sintaxis de purrr::map para aplicar una función o expresión a cada nivel.
fct_recode(f, v = "a", x = "b", z = "c")
fct_relabel(f, ~ paste0("x", .x))



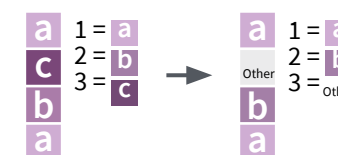
fct_anon(f, prefix = "")
Anonimiza niveles con números enteros al azar. **fct_anon(f)**



fct_collapse(f, ...) Colapsa niveles en grupos definidos manualmente.
fct_collapse(f, x = c("a", "b"))



fct_lump(f, n, prop, w = NULL, other_level = "Other", ties.method = c("min", "average", "first", "last", "random", "max")) Combina niveles menos o más comunes en un único nivel. También **fct_lump_min**.
fct_lump(f, n = 1)



fct_other(f, keep, drop, other_level = "Other") Reemplaza niveles con "other." **fct_other(f, keep = c("a", "b"))**

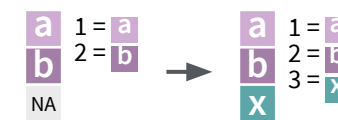
Agrega o descarta niveles



fct_drop(f, only) Descarga niveles no usados.
f5 <- factor(c("a", "b"), c("a", "b", "x"))
f6 <- fct_drop(f5)



fct_expand(f, ...) Agrega niveles a un factor. **fct_expand(f6, "x")**



fct_explicit_na(f, na_level = "(Missing)") Asigna un nivel a NAs para incluirlos, por ejemplo en gráficos.
fct_explicit_na(factor(c("a", "b", NA)))