

Clase 1.3

Acomodo de datos

Marcos Rosetti y Luis Pacheco-Cobos

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

Acomodo de datos: **subset** ()

- Una función que hace subconjuntos de las estructuras según los criterios que proveemos.

```
data(Animals, package = "MASS")  
head(Animals)
```

```
##           body brain  
## Mountain beaver  1.35  8.1  
## Cow              465.00 423.0  
## Grey wolf        36.33 119.5  
## Goat             27.66 115.0  
## Guinea pig       1.04  5.5  
## Dipliodocus      11700.00 50.0
```

Acomodo de datos: `subset()`

- Una función que hace subconjuntos de las estructuras según los criterios que proveemos.

```
subset(x = Animals, subset = body / brain > 10)
```

```
##           body brain  
## Dipliodocus 11700  50.0  
## Triceratops  9400  70.0  
## Brachiosaurus 87000 154.5
```

Acomodo de datos: `aggregate()`

- Colapsa los datos usando ciertos criterios (factores) y calculando una función.

```
data(Traffic, package = "MASS")  
head(Traffic)
```

```
##   year day limit  y  
## 1 1961  1    no  9  
## 2 1961  2    no 11  
## 3 1961  3    no  9  
## 4 1961  4    no 20  
## 5 1961  5    no 31  
## 6 1961  6    no 26
```

Acomodo de datos: `aggregate()`

- Colapsa los datos usando ciertos criterios (factores) y calculando una función.

```
aggregate(formula = y ~ year*limit, data = Traffic, FUN = mean)
```

```
##   year limit      y
## 1 1961   no 23.67606
## 2 1962   no 22.25000
## 3 1961  yes 19.66667
## 4 1962  yes 18.58333
```

Paréntesis: el operador tilde ~

- El operador tilde ~ describe una relación particular entre dos variables, por ejemplo: **x** en función de **y**.

```
aggregate (x ~ y)
```

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/).