

Supply, Aggregate y Attach/Detach

Téllez Gerardo Rubén

20/6/2021

Aplicar funciones a data frames

- **sapply(DF, function):** para aplicar una función a todas las columnas de un DF
 - **na.rm = TRUE:** evitar que el valores que devuelva la función contengan NAs
- **aggregate(variables~factores, data = DF, FUN = función):** para aplicar una función a variables de un DF por los niveles de un, o más de un factor.
 - Si queremos aplicar una función a más de una variable, hay que agruparlos con **cbind()**
 - Si queremos separar las variables mediante más de un factor, se agrupan con signos +

```
irisw =iris
```

```
str(irisw)
```

```
## 'data.frame':   150 obs. of  5 variables:
## $ Sepal.Length: num  5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...
## $ Sepal.Width : num  3.5 3 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 ...
## $ Petal.Length: num  1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...
## $ Petal.Width : num  0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...
## $ Species      : Factor w/ 3 levels "setosa","versicolor",...: 1 1 1 1 1 1 1 1 1 1 ...
```

```
sapply(subset(irisw, select=1:4), mean)
```

```
## Sepal.Length Sepal.Width Petal.Length Petal.Width
##      5.843333      3.057333      3.758000      1.199333
```

```
sapply(irisw[, 1:4], sum)
```

```
## Sepal.Length Sepal.Width Petal.Length Petal.Width
##      876.5      458.6      563.7      179.9
```

```
ecm = function(x){sqrt(sum(x^2))}
sapply(subset(irisw, select = 1:4), ecm)
```

```
## Sepal.Length Sepal.Width Petal.Length Petal.Width
##      72.27621      37.82063      50.82037      17.38764
```

```
aggregate(cbind(Sepal.Length, Petal.Length)~Species, data = irisw, FUN = mean)
```

```
##      Species Sepal.Length Petal.Length
## 1    setosa      5.006      1.462
## 2 versicolor      5.936      4.260
## 3 virginica      6.588      5.552
```

```
# Aplica la función a ambas longitudes según las especies
```

Cómo se convierte una columna a factor

```
ncars = mtcars
head(ncars)
```

```
##      mpg  cyl disp  hp drat   wt  qsec vs am gear carb
## Mazda RX4      21.0   6  160 110 3.90 2.620 16.46 0  1    4    4
## Mazda RX4 Wag  21.0   6  160 110 3.90 2.875 17.02 0  1    4    4
## Datsun 710      22.8   4  108  93 3.85 2.320 18.61 1  1    4    1
## Hornet 4 Drive  21.4   6  258 110 3.08 3.215 19.44 1  0    3    1
## Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02 0  0    3    2
## Valiant        18.1   6  225 105 2.76 3.460 20.22 1  0    3    1
```

```
str(ncars)
```

```
## 'data.frame':   32 obs. of  11 variables:
## $ mpg : num  21 21 22.8 21.4 18.7 18.1 14.3 24.4 22.8 19.2 ...
## $ cyl : num  6 6 4 6 8 6 8 4 4 6 ...
## $ disp: num  160 160 108 258 360 ...
## $ hp : num  110 110 93 110 175 105 245 62 95 123 ...
## $ drat: num  3.9 3.9 3.85 3.08 3.15 2.76 3.21 3.69 3.92 3.92 ...
## $ wt : num  2.62 2.88 2.32 3.21 3.44 ...
## $ qsec: num  16.5 17 18.6 19.4 17 ...
## $ vs : num  0 0 1 1 0 1 0 1 1 1 ...
## $ am : num  1 1 1 0 0 0 0 0 0 0 ...
## $ gear: num  4 4 4 3 3 3 3 4 4 4 ...
## $ carb: num  4 4 1 1 2 1 4 2 2 4 ...
```

```
ncars$cyl = as.factor(ncars$cyl)
ncars$gear = as.factor(ncars$gear)
```

```
aggregate(mpg ~ cyl+gear+carb, data=ncars, FUN = mean, na.rm = TRUE)
```

```
##      cyl gear carb    mpg
## 1     4    3    1 21.50
## 2     6    3    1 19.75
## 3     4    4    1 29.10
## 4     8    3    2 17.15
## 5     4    4    2 24.75
```

```
## 6    4    5    2 28.20
## 7    8    3    3 16.30
## 8    8    3    4 12.62
## 9    6    4    4 19.75
## 10   8    5    4 15.80
## 11   6    5    6 19.70
## 12   8    5    8 15.00
```

```
# Según cilindro, marcha y carburador, se tiene el consumo medio del coche.
# Numérica antes de la tilde
# Factor después de la tilde
```

Attach

- **attach(DF)**: para hacer que R entienda sus variables como globales y que las podamos usar por su nombre sin aplicar el DF\$
 - Si ya existe una variable definida con el mismo nombre que en el DF, lanza un mensaje de error
- **detach(DF)**: para eliminar las variables del DF del entornos global

```
attach(mtcars)
cyl
```

```
## [1] 6 6 4 6 8 6 8 4 4 6 6 8 8 8 8 8 4 4 4 4 8 8 8 8 4 4 4 8 6 8 4
```

```
exists("cyl")
```

```
## [1] TRUE
```

```
detach(mtcars)
exists("cyl")
```

```
## [1] FALSE
```