

# Introducción al análisis de datos con R. Ejemplo: mtcars

Probabilidad y Estadística (c)

14 de Abril, 2020

## ¿De que trata conjunto de datos “mtcars”?

mtcars es un conjunto de datos que viene con R (hay muchos más).

- Para una descripción de los datos escribir en la consola:

```
help(mtcars)
```

- Para ver los datos completos escribir en la consola

```
mtcars
```

## Explorando los datos

- Para ver las primeras filas

```
head(mtcars)
```

```
##           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
```

- Si, en cambio, queremos ver las últimas, escribimos

```
tail(mtcars)
```

## Para ver los nombres de las columnas

```
colnames(mtcars)
```

```
## [1] "mpg" "cyl" "disp" "hp" "drat" "wt" "qsec" "vs" "am" "gear"
## [11] "carb"
```

## Para ver los nombres de las filas

```
rownames(mtcars)
```

```
## [1] "Mazda RX4"           "Mazda RX4 Wag"       "Datsun 710"
## [4] "Hornet 4 Drive"      "Hornet Sportabout"   "Valiant"
## [7] "Duster 360"          "Merc 240D"           "Merc 230"
## [10] "Merc 280"            "Merc 280C"           "Merc 450SE"
## [13] "Merc 450SL"          "Merc 450SLC"         "Cadillac Fleetwood"
## [16] "Lincoln Continental" "Chrysler Imperial"   "Fiat 128"
## [19] "Honda Civic"         "Toyota Corolla"      "Toyota Corona"
## [22] "Dodge Challenger"    "AMC Javelin"         "Camaro Z28"
## [25] "Pontiac Firebird"    "Fiat X1-9"           "Porsche 914-2"
## [28] "Lotus Europa"        "Ford Pantera L"      "Ferrari Dino"
## [31] "Maserati Bora"       "Volvo 142E"
```

## Cómo extraer algunas filas

- Extraemos las filas 3 y 5:

```
mtcars[c(3,5),]
```

```
##           mpg cyl disp  hp drat   wt  qsec vs am gear carb
## Datsun 710  22.8   4  108  93 3.85  2.32 18.61  1  1   4    1
## Hornet Sportabout 18.7   8  360 175 3.15  3.44 17.02  0  0   3    2
```

## Cómo extraer algunas columnas

- Usando su número

```
mtcars[,4]
```

```
## [1] 110 110  93 110 175 105 245  62  95 123 123 180 180 180 205 215 230  66  52
## [20]  65  97 150 150 245 175  66  91 113 264 175 335 109
```

- Usando su nombre

```
mtcars$hp
```

```
## [1] 110 110  93 110 175 105 245  62  95 123 123 180 180 180 205 215 230  66  52
## [20]  65  97 150 150 245 175  66  91 113 264 175 335 109
```

## ¿Cuántos autos tienen 4 velocidades?

```
mtcars$gear
```

```
## [1] 4 4 4 3 3 3 3 4 4 4 4 3 3 3 3 3 4 4 4 3 3 3 3 4 5 5 5 5 4
```

```
mtcars$gear == 4
```

```
## [1] TRUE TRUE TRUE FALSE FALSE FALSE FALSE TRUE TRUE TRUE TRUE FALSE
## [13] FALSE FALSE FALSE FALSE FALSE TRUE TRUE TRUE FALSE FALSE FALSE
```

```
## [25] FALSE TRUE FALSE FALSE FALSE FALSE FALSE TRUE
```

```
sum(mtcars$gear == 4)
```

```
## [1] 12
```

## ¿Qué autos tienen 4 velocidades?

```
which( mtcars$gear == 4)
```

```
## [1] 1 2 3 8 9 10 11 18 19 20 26 32
```

```
rownames(mtcars)[c(1,2,3,8,9,10,11,18,19,20,26,32)]
```

```
## [1] "Mazda RX4"      "Mazda RX4 Wag"  "Datsun 710"     "Merc 240D"
## [5] "Merc 230"      "Merc 280"       "Merc 280C"      "Fiat 128"
## [9] "Honda Civic"   "Toyota Corolla" "Fiat X1-9"      "Volvo 142E"
```

## ¿Qué autos tienen 4 velocidades? Otra forma.

```
rownames(mtcars)[which( mtcars$gear == 4)]
```

```
## [1] "Mazda RX4"      "Mazda RX4 Wag"  "Datsun 710"     "Merc 240D"
## [5] "Merc 230"      "Merc 280"       "Merc 280C"      "Fiat 128"
## [9] "Honda Civic"   "Toyota Corolla" "Fiat X1-9"      "Volvo 142E"
```

o lo hacemos en dos pasos

```
cuales <- which( mtcars$gear == 4)
rownames(mtcars)[cuales]
```

```
## [1] "Mazda RX4"      "Mazda RX4 Wag"  "Datsun 710"     "Merc 240D"
## [5] "Merc 230"      "Merc 280"       "Merc 280C"      "Fiat 128"
## [9] "Honda Civic"   "Toyota Corolla" "Fiat X1-9"      "Volvo 142E"
```

## ¿Qué autos tienen 4 velocidades y transmisión manual?

```
cuales <- which( mtcars$gear == 4 & mtcars$am == 1)
rownames(mtcars)[cuales]
```

```
## [1] "Mazda RX4"      "Mazda RX4 Wag"  "Datsun 710"     "Fiat 128"
## [5] "Honda Civic"   "Toyota Corolla" "Fiat X1-9"      "Volvo 142E"
```

Hallar la cantidad media de millas por galón de los autos de este conjunto de datos.

```
mean(mtcars$mpg)
```

```
## [1] 20.09062
```

Hallar la cantidad media de millas por galón de los autos con 2 y 3 carburadores.

```
mean(mtcars$mpg[mtcars$carb == 2])
```

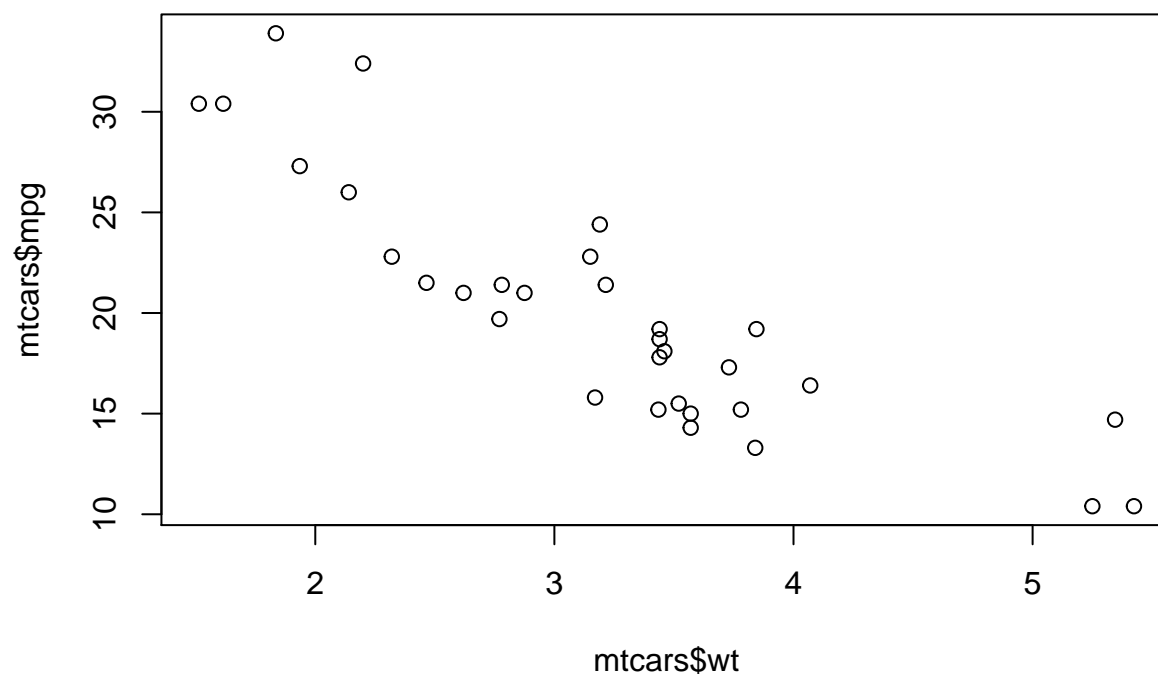
```
## [1] 22.4
```

```
mean(mtcars$mpg[mtcars$carb == 3])
```

```
## [1] 16.3
```

Explorar la relación entre millas por galón y peso

```
plot(mtcars$wt, mtcars$mpg)
```



¿Qué podemos decir de la relación entre millas por galón y peso?

La relación es aproximadamente lineal. Podríamos arriesgarnos a decir que por cada tonelada de peso, las millas por galon disminuyen en 10 (muuuuy a ojo).