

clase 1

Ejercicio clase 15/4/2020

Considerar el conjunto de datos mtcars en R. (Ver help(mtcars))

```
datos<-mtcars  
help(mtcars) #así veo todo el contenido del data frame
```

A)¿Qué variables observa en el data Frame?

```
names(mtcars)
```

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

Usando la función *names* podemos obtener los nombres de las variables, pero como están abreviadas no se entiende bien que es cada una. Puedo usar Help que me da mayor detalle del contenido.

B)¿Cuántas observaciones tiene cada variable?

```
dim(datos)
```

```
## [1] 32 11
```

Observemos que las filas de la tabla de datos corresponden a las observaciones, y las columnas a las variables, por lo tanto mirando la dimensión del data frame **datos** observamos que consiste en 32 observaciones de 11 variables.

C)¿Qué tipo de variables son?

```
str(datos)
```

```
## '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 ...
```

La función *str* nos devuelve la estructura del data frame, en la que como primer elemento luego del nombre de la variable podemos ver cual es el tipo para esa variable. En este caso considero a todas las variables como numéricas.

D)¿Le parece correcto que todas las variables se consideren numéricas? ¿Cual cambiaría?¿cómo lo haría?

La variable **vs** es una variable categórica que vale 0 si el motor tiene forma de V o 1 si es de cilindros alineados. Por lo tanto lo correcto sería considerar a esta variable como un factor. Lo mismo con la variable **am**.

```
datos$vs<-as.factor(datos$vs)
datos$am<-as.factor(datos$am)
```

```
str(datos)
```

```
## '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 : Factor w/ 2 levels "0","1": 1 1 2 2 1 2 1 2 2 2 ...
## $ am : Factor w/ 2 levels "0","1": 2 2 2 1 1 1 1 1 1 1 ...
## $ gear: num 4 4 4 3 3 3 3 4 4 4 ...
## $ carb: num 4 4 1 1 2 1 4 2 2 4 ...
```

E) ¿Qué autos tienen 4 velocidades?

Tenemos que pedir que nos devuelva el nombre de las filas que correspondan a los datos que en la columna **gear** tengan el valor 4

```
rownames(datos[datos$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"
```

F) ¿Qué subconjunto de mtcars es mtcars[mtcars\$disp > 150 & mtcars\$mpg > 20,]?

En esta línea de código pedimos que del data frame mtcars nos muestre todos los datos correspondientes a los autos que tienen un valor de **disp** mayor a 150 y a demás consumen más de 20 millas por galón. Como llamé **datos** a mis datos, cambio la línea de código adaptándola a mi nombre.

```
datos[datos$disp > 150 & datos$mpg > 20,]
```

```
##           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
## Hornet 4 Drive 21.4   6  258 110 3.08 3.215 19.44 1  0   3    1
```

G) ¿Qué autos tienen 4 velocidades y transmisión manual?

Muy similar al ítem (E), pero agregando una condición más.

```
rownames(datos[datos$gear==4 & datos$am==1,])
```

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

H) Hallar la cantidad media de millas por galón de los autos con 2 carburadores.

Usamos la función *mean* para estimar la cantidad media de la variable **mpg**, contemplando que solo queremos el subconjunto correspondiente a los autos con 2 carburadores

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

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

I) Graficar el consumo de combustible de los automóviles (mpg) vs el peso del motor (wt). Analizar lo observado.

```
plot(datos$mpg,datos$wt,pch=20,col="Darkblue", xlab="Millas por galón",  
      ylab="Peso del motor", main="Gráfico para observar tendencia")
```

Gráfico para observar tendencia

