

## OBJECTIUS

Aquests exercicis pretenen familiaritzar a l'alumne amb l'ús de ggplot per crear scatter plots, gràfics de barres, histogrames i boxplots.

## REQUERIMENTS

Per poder fer els exercicis es recomana primer veure els apartats 3.2, 3.2.1, 3.2.2 d'aquest mòdul, i haver testejat exemples de 3.2.3. S'ha d'haver completat també la M2T01

## DADES

Continuem amb el conjunt de dades *mtcars* de R que vaig utilitzar al vídeo. Aquest dataframe conté informació de 32 cotxes, recollides en 11 variables diferents.

Com el conjunt de dades és petit, poseu `view(mtcars)` per veure el *dataframe* en forma de taula. La funció `view()` ens permet obtenir informació sobre el conjunt de dades de manera complementària a les dues maneres que vam veure l'altre dia: l'ajuda `?mtcars` o la funció mostrant l'estructura, `str(mtcars)`. Us apareixerà:

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
Mazda RX4	21.0	6	160.0	110	3.90	2.620	16.46	0	1	4	4
Mazda RX4 Wag	21.0	6	160.0	110	3.90	2.875	17.02	0	1	4	4
Datsun 710	22.8	4	108.0	93	3.85	2.320	18.61	1	1	4	1
Hornet 4 Drive	21.4	6	258.0	110	3.08	3.215	19.44	1	0	3	1
Hornet Sportabout	18.7	8	360.0	175	3.15	3.440	17.02	0	0	3	2
Valiant	18.1	6	225.0	105	2.76	3.460	20.22	1	0	3	1
Duster 360	14.3	8	360.0	245	3.21	3.570	15.84	0	0	3	4
Merc 240D	24.4	4	146.7	62	3.69	3.190	20.00	1	0	4	2
Merc 230	22.8	4	140.8	95	3.92	3.150	22.90	1	0	4	2
Merc 280	19.2	6	167.6	123	3.92	3.440	18.30	1	0	4	4
Merc 280C	17.8	6	167.6	123	3.92	3.440	18.90	1	0	4	4
Merc 450SE	16.4	8	275.8	180	3.07	4.070	17.40	0	0	3	3
Merc 450SL	17.3	8	275.8	180	3.07	3.730	17.60	0	0	3	3
Merc 450SLC	15.2	8	275.8	180	3.07	3.780	18.00	0	0	3	3

Showing 1 to 15 of 32 entries, 11 total columns

NOTA: Si obriu R de nou, primer de tot recordeu que heu de tornar a carregar la llibreria tidyverse: `library(tidyverse)`.

## EXERCICIS

1.- D'una manera semblant a la que se us va explicar al vídeo, feu un scatter plot que compari el pes d'un cotxe (*wt*), amb el desplaçament (*disp*) i la cilindrada (*cyl*). Si alguna d'aquestes variables numèriques es pot categoritzar amb factor, feu-ho i diferencieu els seus valors/nivells amb l'estètica corresponent a la forma (*shape*).

Poseu etiquetes als eixos i a la llegenda fent ús de la capa *scale* (com se us va ensenyar al vídeo també).

Podeu extreure alguna conclusió entre la relació de les tres variables veient el gràfic?

2.- Utilitzant una mateixa gràfica, volem saber quants cotxes hi ha de tres grups/nivells diferents segons el nombre de cilindres (4 cilindres, 6 cilindres i 8 cilindres). Feu un

gràfic de barres on l'alçada de les barres sigui proporcional al nombre de cotxes de cada grup/nivell. Quina informació us dona la gràfica?

a) Poseu la llegenda dels eixos utilitzant `xlab()` i `ylab()`. Com no heu utilitzat encara `xlab()` i `ylab()`, feu ús de l'ajuda de R per saber com s'utilitzen.

b) Pinteu la part interior de les tres barres que us surten segons el nombre de cilindres (4, 6 o 8). Per fer-ho utilitzeu `fill` dins d'`aes()`, és a dir, la vostra estètica serà de la forma: `aes(***,fill=factor(cyl))`, on en `***` hi posareu la variable que voleu mostrar en l'eix x del vostre gràfic. Què passa si utilitzeu `color` enlloc de `fill`?

Ens dona alguna informació extra pintar les barres? Per què?

c) Feu les barres més estretes especificant l'amplada (`width`) de les barres. Nota: `width` actua com argument de `geom_bar()`. Posem doncs `geom_bar(width=0.5)`. Proveu altres valors de `width`. Què us passa si poseu 1.5?

**3.-** Feu un gràfic on pugueu veure la proporció de vehicles amb transmissió manual versus automàtica segons el nombre de cilindres. Què observeu?

Utilitzeu l'argument `dodge` de manera que cada tipus de transmissió aparegui en l'eix x, una al costat de l'altra d'acord amb el nombre de cilindres. Per fer això últim especificarem l'argument posició en `geom_bar(position="dodge")`. Observeu més coses?

**4.-** Mostreu la distribució de cotxes amb una certa potencia bruta ('hp'). Quina informació en podeu extreure?

**5.-** Dibuixeu un boxplot on l'eix de les x correspongui a la variable 'cyl' (cilindres) i l'eix y a la variable 'mpg' (milles de galó). A més:

a) Poseu la llegenda dels eixos utilitzant `xlab()` i `ylab()`

b) Gireu els eixos de coordenades, fent que els boxplots quedin en horitzontal utilitzant la comanda `coord_flip()`

c) Pinteu els boxplots de tres maneres diferents:

c1) Escollint un únic color amb `fill`, `color` i `alpha` per als tres grups:

`geom_boxplot(color="red", fill="blue", alpha=0.2)`

Fixeu-vos en el resultat de canviar els tres paràmetres `color`, `fill`, `alpha`. Què fa cadascun?

c2) Donant color només el contorn del boxplot per a cada grup (cotxes amb els mateixos cilindres)

c3) Donant color a l'interior de cada caixa per a cada grup (cotxes amb els mateixos cilindres)

Els canvis fets en c2) i c3) ens aporten alguna informació nova? Doneu un exemple, en el que afegir color us podria servir per tal d'ampliar la informació, per exemple, donant la relació amb una nova variable

NOTA: Per c2) i c3) al diferenciar els grups posant color, recordeu que afegim la informació a l'estètica, és a dir, dins d'`aes()`.