

Primer ejercicio P2P

El fichero *Halterofilia.csv*, que encontrarás en el repositorio del curso (en el url <https://miriadax.net/documents/28098821/74010125/Halterofilia.csv/f2fb1ddc-6de0-4942-870e-fbe5e9d2eca8>), contiene los resultados del campeonato del mundo de halterofilia de 2010. Cada fila corresponde a un halterófilo participante en el campeonato. Consta de 6 columnas:

- *Peso*: El peso corporal del halterófilo.
- *Arrancada*: El peso máximo que levantó en la modalidad de Arrancada.
- *Dos.Tiempos*: El peso máximo que levantó en la modalidad de Dos Tiempos.
- *Total*: La suma de los dos pesos anteriores.
- *Categoría*: La categoría en la que compitió. Son categorías por peso. Por ejemplo “menos 62” quiere decir que el halterófilo pesaba menos de 62 Kg.
- *Sexo*: “M” para masculino y “F” para femenino.

Para esta tarea debes:

a) Guardar el fichero *Halterofilia.csv* en un data frame llamado **Resultados**. La variable *Categoría* ha de ser un factor ordenado en el orden creciente de los pesos:

“menos 48”, “menos 53”, “menos 56”, “menos 58”, “menos 62”, “menos 63”, “menos 69”,
“menos 75”, “mas 75”, “menos 77”, “menos 85”, “menos 94”, “menos 105”, “mas 105”

```
>Resultados=read.table("https://miriadax.net/documents/28098821/74010125/Halterofilia.csv/f2fb1ddc-6de0-4942-870e-fbe5e9d2eca8", header=TRUE, sep =";")
```

```
>Resultados$Categoría=sort(Resultados$Categoría, deacresing=FALSE)
```

```
> str(Resultados)
```

```
'data.frame': 462 obs. of 6 variables:
 $ Peso      : num  55.6 55.6 55.9 55.7 55.9 ...
 $ Arrancada  : int   132 127 130 123 120 127 118 112 121 115 ...
 $ Dos.Tiempos: int   160 161 150 150 149 140 146 152 142 146 ...
 $ Total      : int   292 288 280 273 269 267 264 264 263 261 ...
 $ Categoría  : Factor w/ 14 levels "mas 105","mas 75",...: 1 1 1 1 1 1 1 1 1 1 ...
 $ Sexo       : Factor w/ 2 levels "F","M": 2 2 2 2 2 2 2 2 2 2 ...
```

Peso Arrancada Dos.Tiempos Total Categoría Sexo

1	55.61	132	160	292	menos 48	M
2	55.64	127	161	288	menos 48	M
3	55.87	130	150	280	menos 48	M
4	55.73	123	150	273	menos 48	M
5	55.93	120	149	269	menos 48	M

6	55.87	127	140	267	menos 48	M
7	55.86	118	146	264	menos 48	M
8	55.88	112	152	264	menos 48	M
9	55.19	121	142	263	menos 48	M
10	55.93	115	146	261	menos 48	M

[...]

452	94.31	106	136	242	mas 105	F
453	110.08	101	138	239	mas 105	F
454	100.55	98	138	236	mas 105	F
455	80.95	106	128	234	mas 105	F
456	98.79	102	125	227	mas 105	F
457	102.36	98	124	222	mas 105	F
458	89.60	100	120	220	mas 105	F
459	97.51	95	123	218	mas 105	F
460	104.36	72	105	177	mas 105	F
461	83.24	77	98	175	mas 105	F
462	101.29	78	97	175	mas 105	F

b) Crear una tabla en la que, para cada combinación de categoría y sexo, se dé el peso medio en la modalidad de Arrancada y el peso medio en la modalidad de Dos Tiempos de los participantes de esa categoría y ese sexo. Esta tabla tiene que tener cuatro columnas (“Categoría”, “Sexo”, “Peso medio Arrancada” y “Peso medio Dos Tiempos”) y una fila por combinación de categoría y sexo. Te recomendamos usar la función **aggregate** para calcular los pesos medios.

```
> Resultados2=aggregate(cbind(Resultados$Arrancada, Resultados$Dos.Tiempos)~Cate-
goria+Sexo, data=Resultados,FUN=mean)
> names(Resultados2)=c("Categoría","Sexo","Peso medio Arrancada","Peso medio Dos
Tiempos")
```

	Categoría	Sexo	Peso medio Arrancada	Peso medio Dos Tiempos
1	mas 75	F	82.55556	106.22222

2	menos 77	F	75.91667	95.83333
3	menos 85	F	82.37143	103.94286
4	menos 94	F	93.24390	116.85366
5	menos 105	F	99.89474	122.60526
6	mas 105	F	101.60000	130.36667
7	menos 48	M	113.39286	139.00000
8	menos 53	M	123.92308	151.69231
9	menos 56	M	131.16129	160.54839
10	menos 58	M	141.89655	174.20690
11	menos 62	M	146.03125	178.40625
12	menos 63	M	154.92308	187.76923
13	menos 69	M	163.17241	198.55172
14	menos 75	M	169.68966	208.44828
15	mas 75	M	157.71429	194.92857

c) Extraer de **Resultados** dos dataframes: uno llamado **ResM**, con los datos de los hombres, y otro llamado **ResF**, con los datos de las mujeres. Ambos dataframes deben contener sólo dos columnas: "Peso" y "Total".

```
>ResF=Resultados[Resultados$Sexo=="F",c("Peso","Total")]
```

```
>ResM=Resultados[Resultados$Sexo=="M",c("Peso","Total")]
```

ResF

Peso Total

274 47.88 214

275 47.72 205

276 47.83 204

277 47.71 186

278 47.08 185

279 47.94 180

280 47.87 176

281 47.40 175

282 47.20 174

283 47.87 171

284 47.94 170

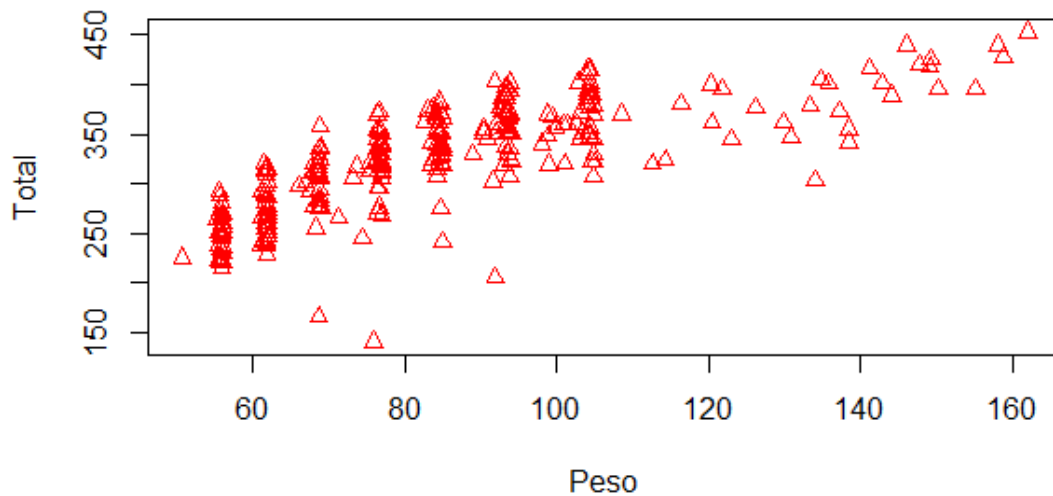
ResM

Peso Total

1	55.61	292
2	55.64	288
3	55.87	280
4	55.73	273
5	55.93	269
6	55.87	267
7	55.86	264
8	55.88	264
9	55.19	263
10	55.93	261

d) Crear un gráfico a partir de **ResM**, con el peso corporal en el eje horizontal y el peso total levantado en el eje vertical. Tienes que usar triángulos vacíos de color rojo. Verás que la distribución de los puntos es bastante peculiar. Encuentra una razón para esta distribución.

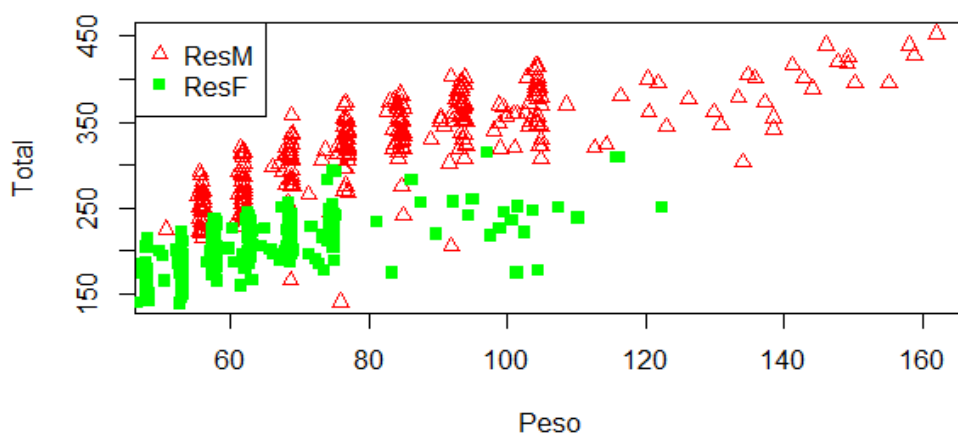
```
> plot(ResM,pch=2,col="red")
```



e) Añadir al gráfico los puntos correspondientes a **ResF**. Tienes que usar cuadrados llenos de color verde.

Añade una leyenda que indique cada tipo de punto qué sexo representa. Asegúrate de que todos los puntos están dentro de la gráfica y ninguno queda sin representar por haber quedado fuera de los límites. Es posible que tengas que redibujar todo usando valores adecuados de "xlim" e "ylim".

```
> points(ResF$Peso,ResF$Total,pch=22,col="green",bg="green")
> legend("topleft",legend=c(expression(ResM),expression(ResF)),col=c("red","green"),pch=c(2,15))
```



f) Calcular la regresión lineal del peso levantado en función del peso corporal tanto para los hombres como para las mujeres. Indica en ambos casos cuánto peso levantan de media por cada Kg. adicional de peso corporal. Indica también los coeficientes de determinación.

Regressió lineal:

```
> lm(ResM$Total~ResM$Peso)
```

Call:

```
lm(formula = ResM$Total ~ ResM$Peso)
```

Coefficients:

(Intercept)	ResM\$Peso
177.480	1.727

```
> lm(ResF$Total~ResF$Peso)
```

Call:

```
lm(formula = ResF$Total ~ ResF$Peso)
```

Coefficients:

(Intercept)	ResF\$Peso
103.05	1.54

Coefficiente de determinación:

```
> summary(lm(ResM$Total~ResM$Peso))$r.squared  
[1] 0.5834273
```

g) Añadir las dos rectas al gráfico anterior con los colores correspondientes.

```
> M=abline(lm(Total~Peso,data = ResM),col="blue",lwd=2)  
> F=abline(lm(Total~Peso,data = ResF),col="orange",lwd=2)  
> legend("bottomright",legend = c(expression(M),expression(F)),col=c("blue","orange"),lwd=2)
```

