

Data Frames en R. Tarea 01

Alberto Simón

22/04/2025

Nota: el enlace suministrado no existe. He utilizado el fichero medals.csv de la carpeta data del curso.

```
df = read.csv("../data/medals.csv")
str(df)
```

```
## 'data.frame':    2311 obs. of  9 variables:
## $ X              : int  1 2 3 4 5 6 7 8 9 10 ...
## $ Year            : int  1924 1924 1924 1924 1924 1924 1924 1924 1924 1924 ...
## $ City            : chr  "Chamonix" "Chamonix" "Chamonix" "Chamonix" ...
## $ Sport           : chr  "Skating" "Skating" "Skating" "Bobsleigh" ...
## $ Discipline      : chr  "Figure skating" "Figure skating" "Figure skating" "Bobsleigh" ...
## $ NOC              : chr  "AUT" "AUT" "AUT" "BEL" ...
## $ Event            : chr  "individual" "individual" "pairs" "four-man" ...
## $ Event.gender     : chr  "M" "W" "X" "M" ...
## $ Medal            : chr  "Silver" "Gold" "Gold" "Bronze" ...
```

Pregunta 1

¿Cuántos deportistas aparecen en el data frame? - No hay campo con el nombre de los deportistas, así que no se puede saber cuántos deportistas distintos hay. Sólo eventos.

```
print(nrow(df))
```

```
## [1] 2311
```

Pregunta 2

¿Cuántos deportistas han ganado medallas de oro? ¿Cuántos han ganado medallas de plata? ¿Y cuántos de bronce?

```
table(df$Medal)
```

```
##
## Bronze    Gold Silver
##    764     774     773
```

Pregunta 3

¿En cuántos lugares se han hecho olimpiadas de invierno?

```
length(unique(df$City))
```

```
## [1] 17
```

Pregunta 4

¿Cuántos hombres y cuántas mujeres hay?

```
table(df$Event.gender)
```

```
##  
##      M      W      X  
## 1386   802   123
```

Pregunta 5

¿En qué año participaron más deportistas?

```
year_table = table(df$Year)  
year_table
```

```
##  
## 1924 1928 1932 1936 1948 1952 1956 1960 1964 1968 1972 1976 1980 1984 1988 1992  
##   49   41   42   51   68   67   72   81  103  106  105  111  115  117  138  171  
## 1994 1998 2002 2006  
##  183  205  234  252
```

```
year_table[which.max(year_table)]
```

```
## 2006  
## 252
```

Pregunta 6

El campo NOC indica el país del ganador de la medalla. ¿Qué país puede presumir de haber ganado más medallas de oro en los juegos de invierno entre 1960 y 1996?

```
subset_df = subset(df, Year >= 1960 & Year <= 1996 & Medal == "Gold")  
tabla_oro = table(subset_df$NOC)  
tabla_oro
```

```
##  
## AUT CAN ESP EUA EUN FIN FRA FRG GBR GDR GER ITA JPN KAZ KOR LIE NED NOR POL RUS  
##  25  12   1   7   9  20  12  11   4  39  19  22   3   1   6   2  14  40   1  11  
## SUI SWE TCH UKR URS USA UZB  
##  19  24   2   1  71  34   1
```

```
top <- sort(tabla_oro, decreasing = TRUE)[1]  
top
```

```
## URS  
## 71
```