

Tarea8

Imanol

17/2/2021

Carga en R o en Python el CSV de la web “Pincha aquí” y responde a las preguntas de la tarea.

Preguntas de esta tarea

1. ¿Cuántos deportistas aparecen en el data frame?
2. ¿Cuántos han ganado medallas de oro, cuantos de plata y cuantos de bronce?
3. ¿En cuantos lugares se han hecho olimpiadas de invierno?
4. ¿Cuántos hombres y cuantas mujeres hay?
5. ¿En qué año participaron más deportistas?
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?

```
df = read.csv("http://winterolympicsmedals.com/medals.csv", header = TRUE)
```

```
str(df)
```

```
## 'data.frame':    2311 obs. of  8 variables:
## $ Year          : int  1924 1924 1924 1924 1924 1924 1924 1924 1924 1924 ...
## $ City          : Factor w/ 17 levels "Albertville",...: 3 3 3 3 3 3 3 3 3 3 ...
## $ Sport         : Factor w/ 7 levels "Biathlon","Bobsleigh",...: 6 6 6 2 4 1 6 6 6 6 ...
## $ Discipline    : Factor w/ 15 levels "Alpine Skiing",...: 6 6 6 3 8 2 6 15 15 15 ...
## $ NOC           : Factor w/ 45 levels "AUS","AUT","BEL",...: 2 2 2 3 6 15 15 15 15 15 ...
## $ Event         : Factor w/ 67 levels "10000m","1000m",...: 48 48 56 41 47 54 56 1 1 7 ...
## $ Event.gender  : Factor w/ 3 levels "M","W","X": 1 2 3 1 1 1 3 1 1 1 ...
## $ Medal         : Factor w/ 3 levels "Bronze","Gold",...: 3 2 2 1 2 3 3 2 3 2 ...
```

```
names(df)
```

```
## [1] "Year"          "City"          "Sport"         "Discipline"    "NOC"
## [6] "Event"         "Event.gender"  "Medal"
```

```
nDeportistas = nrow(df)
nDeportistas
```

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

```
sum(df$Medal == "Gold")
```

```
## [1] 774
```

```
table(df$Medal == "Gold")
```

```
##  
## FALSE TRUE  
## 1537 774
```

```
sum(df$Medal == "Silver")
```

```
## [1] 773
```

```
table(df$Medal == "Silver")
```

```
##  
## FALSE TRUE  
## 1538 773
```

```
sum(df$Medal == "Bronze")
```

```
## [1] 764
```

```
table(df$Medal == "Bronze")
```

```
##  
## FALSE TRUE  
## 1547 764
```

```
levels(df$City)
```

```
## [1] "Albertville"      "Calgary"           "Chamonix"  
## [4] "Cortina d'Ampezzo" "Garmisch-Partenkirchen" "Grenoble"  
## [7] "Innsbruck"        "Lake Placid"       "Lillehammer"  
## [10] "Nagano"           "Oslo"              "Salt Lake City"  
## [13] "Sapporo"          "Sarajevo"          "Squaw Valley"  
## [16] "St. Moritz"       "Turin"
```

```
length(levels(df$City))
```

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

```
sum(df$Event.gender == "M")
```

```
## [1] 1386
```

```
table(df$Event.gender == "M")
```

```
##  
## FALSE TRUE  
## 925 1386
```

```
sum(df$Event.gender == "W")
```

```
## [1] 802
```

```
table(df$Event.gender == "W")
```

```
##  
## FALSE TRUE  
## 1509 802
```

```
table(df$Year)
```

```
##  
## 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
```

```
which.max(table(df$Year))
```

```
## 2006  
## 20
```

```
tabla = table(df$Year)  
names(tabla[tabla == max(tabla)])
```

```
## [1] "2006"
```

```
df2 = df[df$Year >= 1960 & df$Year <= 1996,]  
head(sort(table(df2$NOC[df2$Medal == "Gold"]), dec = TRUE), n=1L)
```

```
## URS  
## 61
```

```
tabla = table(df$NOC)  
names(tabla[tabla == max(tabla)])
```

```
## [1] "NOR"
```