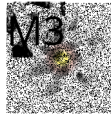


# Taller Evaluable 1, FIFA 2022-23



## Taller evaluable datos FIFA 2023

Cada pregunta vale 1 punto menos las dos últimas que valen 1.5 puntos. Se puntúa la presentación, la claridad y que los dibujos están completos. Este taller está pensado para resolver con R-base pero podéis utilizar tidyverse u otros paquetes de R.

En la web de kaggle FIFA23 OFFICIAL DATASET. Contiene todos los data sets de de datos básicos de FIFA17 to FIFA23 del videojuego.

Las siguientes preguntas son relativas al data set `players_23.csv`, que se adjunta con la práctica.

Hay que contestar con código R explicar adecuadamente cada salida. Subid a la activad el Rmd y el html.

Rellenad estos datos:

### PONED NOMBRE DEL GRUPO

- Florit Ensenyat, Jordi
- Ferrer Fernández, Marc
- Girón Rodríguez, Pau
- Fornés Reynés, Josep Gabriel

## Carga de datos

Tenéis que carga los datos con el código que se ve a continuación. Visualizar y explorara el data set y averiguar de qué tipo son cada una de las variables y en qué tipo de fichero están guardadas. El código carga los datos en un data frame llamado `datos` con la función `read.csv`. Debéis entender clases y tipos de datos de cada de cada columna de datos. El parámetro `encoding` es necesario para cargar debidamente los acentos y caracteres especiales. Lo que obtenemos es un data frame de 18539 observaciones (filas) y 90 variables (columnas).

Cargaremos todas las variables de texto como factor con el parámetro `stringsAsFactors = TRUE`

```
datos = read.csv("players_fifa23.csv",  
  encoding="UTF-8",stringsAsFactors = TRUE)# cambia tu path  
#str(datos)  
#names(datos)
```

Las variables de la 1 (`NationalTeam`) a la 31(`NationalPosition`) son variables de perfil del jugador: su nombre, su equipo su sueldo su número de camiseta... El resto de variables de la 34 (`pace`) a la 90 (`rb`) son variables numéricas enteras con valores de 0 a 100 que parametrizan cómo es el jugador el el juego FIFA player 2023.

## Pregunta 1

Las selecciones europeas que han ganado un mundial son

```
eur=c("England","France","Germany","Italy","Portugal","Spain")
eur
```

```
## [1] "England" "France" "Germany" "Italy" "Portugal" "Spain"
```

Generar un data frame con el nombre `fifa23_eur` con los jugadores de estas selecciones.

## Solución 1

```
fifa23_eur = datos[datos$NationalTeam %in% eur,]
fifa23_eur$NationalTeam = droplevels(fifa23_eur$NationalTeam)
unique(fifa23_eur$Name)
```

```
## [1] K. Benzema      K. Mbappé      M. Neuer      Cristiano Ronaldo
## [5] H. Kane         J. Kimmich     N. Kanté     Rúben Dias
## [9] G. Donnarumma   Bernardo Silva João Cancelo   M. ter Stegen
## [13] A. Rüdiger      Rodri         H. Lloris     T. Müller
## [17] De Gea         M. Verratti    M. Maignan    L. Goretzka
## [21] C. Immobile     R. Sterling    K. Trapp      Bruno Fernandes
## [25] A. Laporte      K. Coman       N. Barella     C. Nkunku
## [29] N. Süle        S. Gnabry      İ. Gündoğan    K. Walker
## [33] Jordi Alba      Sergio Busquets P. Pogba      Jorginho
## [37] Gerard Moreno   Pedri         P. Foden      T. Hernández
## [41] Diogo Jota      M. Reus       L. Hernández  Sergio Ramos
## [45] D. Berardi      J. Grealish    Carvajal      R. Varane
## [49] W. Ben Yedder   L. Insigne     K. Trippier    L. Bonucci
## [53] J. Bellingham   F. Chiesa      Rafael Leão    J. Koundé
## [57] L. Sané         R. James       A. Bastoni     K. Havertz
## [61] João Félix      D. Rice       M. Mount      J. Sancho
## [65] Oyarzabal       L. Pellegrini  Marcos Llorente Rúben Neves
## [69] P. Kimpembe     Unai Simón     Carlos Soler   Pau Torres
## [73] J. Henderson    A. Griezmann   J. Stones      Koke
## [77] G. Di Lorenzo   M. Locatelli   R. Gosens      Gonçalo Guedes
## [81] J. Brandt       T. Werner      Gayà           Otávio
## [85] R. Guerreiro    T. Abraham     N. Schlotterbeck B. Saka
## [89] Dani Olmo       A. Tchouaméni  Ferran Torres  A. Ramsdale
## [93] B. Chilwell     André Silva    J. Pickford     Rui Patrício
## [97] Azpilicueta     L. Digne       L. Spinazzola   A. Lopes
## [101] F. Acerbi       D. Raum        A. Areola      J. Musiala
## [105] Pablo Sarabia   H. Maguire     Palhinha       N. Pope
## [109] João Moutinho   Pepe          J. Hofmann     K. Phillips
## [113] M. Guendouzi    Nuno Mendes    W. Saliba      J. Clauss
## [117] Danilo Pereira  Morata        José Fonte     L. Klostermann
## [121] B. Cristante    A. Belotti     A. Rabiot      C. Gallagher
## [125] Matheus Nunes   G. Raspadori   B. Pavard      Diogo Costa
## [129] C. Coady        A. Florenzi     M. Kean        A. Meret
## [133] L. Nmecha       S. Sirigu      Eric García     Robert Sánchez
## [137] Emerson         T. Kehrer
## 17530 Levels: A. Abaz A. Abdallah A. Abdenmour A. Abdi A. Abdiyanovic ... Zubimendi
```

## Pregunta 2

Calcular la media y la desviación típica del valor de cada selección nacional cada equipo del data frame `fifa23_eur`.

Calcular la media y la desviación típica EN MILES de euros del valor de cada jugador `ValueEUR` de cada selección nacional del frame `fifa23_eur` por posición en el campo delantera media y defensa.

## Solucion 2

## Pregunta 3

Discretiza la variable `ValueEUR` de `fifa23_eur` en los 4 niveles siguientes: “Cuartil\_1”, “Cuartil\_2”, “Cuartil\_3” y “Cuartil\_4”, según los cortes por la función `quantile` para 0.25, 0.5 y 0.75. La variable resultante `Value_Level` tiene que ser un factor ordenado en orden creciente de valor.

## Pregunta 4

¿Qué selección tiene a más jugadores en los intervalos de Valor máximo calculado en el ejercicio anterior?

Estudia la función `droplevels` para quitar los niveles de las selecciones que no aparecen.

## Pregunta 5

¿Respecto al tiro cuántos zurdos, diestros y ambidiestros (3) (buscad qué variable es e interpretar su valor de 1 a 5 hay entre todos los jugadores de `fifa23_eur`? Construir una variable llamada `foot` que tenga por niveles “left”, “right”, “ambidextrous” ¿Qué selección tiene mayor cantidad de zurdos (decidid que es zurdo diestro y ambidiestro)?

## Pregunta 6

Calcular la tabla de contingencia (frecuencias absolutas) por posición `NationalPosition` contra `foot`. Contingencia con las variables `foot`. Calcular la tabla global de proporciones de `NationalPosition` y `foot`. Calcular la tabla de proporciones marginales de `foot` por (condicionada a) `NationalPosition`.

## Pregunta 7

Calcular diagramas de barras adosados para la primera tabla del ejercicio anterior y un diagrama de mosaico de la segunda tabla. Poner una leyenda y nombre del gráfico y comentar los resultados con un pequeño párrafo.

## Pregunta 8

Comparar la distribución del score `Overall` con un `boxplot` para las 6 selecciones. Decorar adecuadamente el resultado. Comentar los resultados.

## Pregunta 9

Generar un data frame `fifa23_ame` que contenga exclusivamente a las 6 selecciones de América que van al mundial 2022.

```
ame=c("Argentina","Brazil","Canada","Mexico","Ecuador","United States" )
```

Generar un data frame `fifa23_ame`. Comparar la distribución del score `overall` para TODOS los jugadores de las 6 selecciones de Europa y TODOS los jugadores de las seis selecciones de América. Dibujando un histograma con la función `kernel` en un solo gráfico. Comentar los resultados.

## Solución

```
fifa23_ame=datos[datos$NationalTeam %in% ame,]  
fifa23_ame$NationalTeam = droplevels(fifa23_ame$NationalTeam)  
unique(fifa23_ame$NationalTeam)
```

```
## [1] Argentina      Canada      Brazil      Mexico      United States  
## Levels: Argentina Brazil Canada Mexico United States
```

```
plot(density(fifa23_eur$Overall),cols = "red")
```

```
## Warning in plot.window(...): "cols" is not a graphical parameter
```

```
## Warning in plot.xy(xy, type, ...): "cols" is not a graphical parameter
```

```
## Warning in axis(side = side, at = at, labels = labels, ...): "cols" is not a  
## graphical parameter
```

```
## Warning in axis(side = side, at = at, labels = labels, ...): "cols" is not a  
## graphical parameter
```

```
## Warning in box(...): "cols" is not a graphical parameter
```

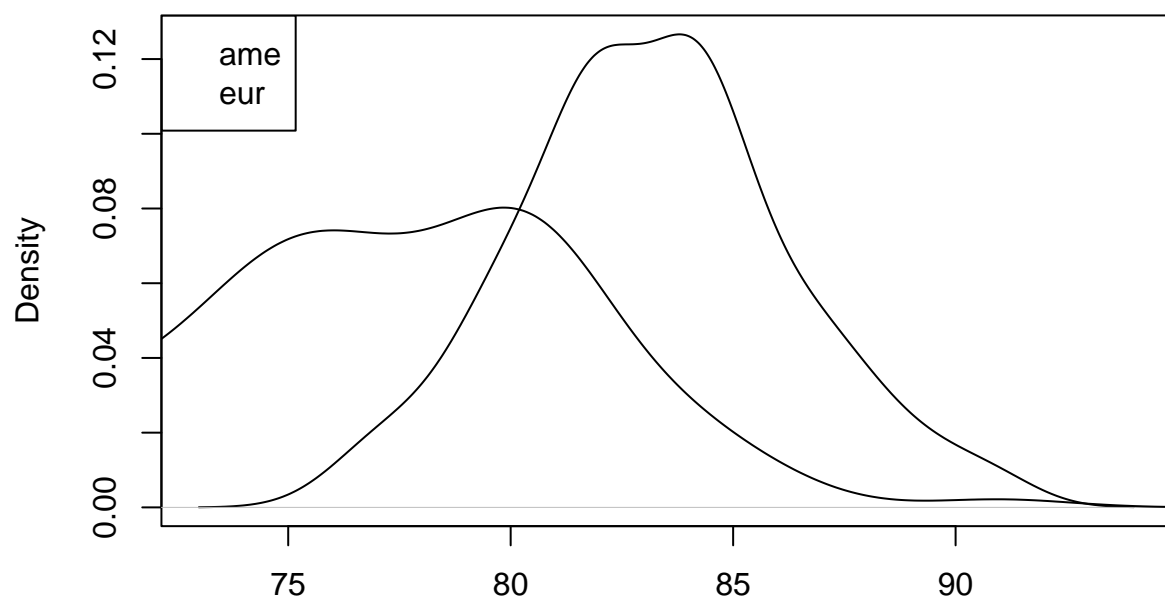
```
## Warning in title(...): "cols" is not a graphical parameter
```

```
lines(density(fifa23_ame$Overall),cols = "blue")
```

```
## Warning in plot.xy(xy.coords(x, y), type = type, ...): "cols" is not a graphical  
## parameter
```

```
legend("topleft",c("ame", "eur"),col=c("blue","red"))
```

**density.default(x = fifa23\_eur\$Overall)**



N = 138 Bandwidth = 1.003