

Postwork Sesión 2 BEDU Santander Programación en R

José David Vázquez Rojas

Importa los datos de soccer de las temporadas 2017/2018, 2018/2019 y 2019/2020 de la primera division de la liga española a R, los datos los puedes encontrar en el siguiente enlace: <https://www.football-data.co.uk/spainm.php>

```
df1<-read.csv("https://www.football-data.co.uk/mmz4281/1718/SP1.csv")
df2<-read.csv("https://www.football-data.co.uk/mmz4281/1819/SP1.csv")
df3<-read.csv("https://www.football-data.co.uk/mmz4281/1920/SP1.csv")
lista<-list(df1,df2,df3)
```

Obten una mejor idea de las características de los data frames al usar las funciones: str, head, View y summary

```
#lapply(lista,str)
#lapply(lista,head)
#lapply(lista,View)
#lapply(lista,summary)
```

Con la funcion select del paquete dplyr selecciona unicamente las columnas Date, HomeTeam, AwayTeam, FTHG, FTAG y FTR; esto para cada uno de los data frames. (Hint: tambien puedes usar lapply).

```
library(dplyr)

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

nuevo<-lapply(lista, select, Date, HomeTeam, AwayTeam, FTHG, FTAG , FTR)
```

Asegurate de que los elementos de las columnas correspondientes de los nuevos data frames sean del mismo tipo (Hint 1: usa as.Date y mutate para arreglar las fechas). Con ayuda de la funcion rbind forma un unico data frame que contenga las seis columnas mencionadas en el punto 3 (Hint 2: la función do.call podria ser utilizada).

```
nuevo <- lapply(nuevo, mutate)
data <- do.call(rbind, nuevo)
setwd("/Users/DavidRojas/Desktop/Santander/BEDU/Módulos/Programación con R/postworks/sesion 2")
write.csv(data, "dataSesion2.csv")
```