



Programación y estadística con R

| Equipo 16

Ana Elizabeth Guzmán Jiménez

Carlos Paz

Fred Jordi Miramontes Arias

Luis Arturo Rosas León

Ludim Sánchez

Norma Arrazola Herrera

Contenido

PostWork 1: Introducción a R y Software	3
Ejercicio 1: Importa los datos de soccer de la temporada 2019/2020 de la primera división de la liga española a R, los datos los puedes encontrar en el siguiente enlace: https://www.football-data.co.uk/spainm.php	3
Ejercicio 2: Del data frame que resulta de importar los datos a R, extrae las columnas que contienen los números de goles anotados por los equipos que jugaron en casa (FTHG) y los goles anotados por los equipos que jugaron como visitante (FTAG)	3
Ejercicio 3: Consulta cómo funciona la función table en R al ejecutar en la consola ?table	4
Ejercicio 4: Posteriormente elabora tablas de frecuencias relativas para estimar las siguientes probabilidades:	5
Parte 4.1: La probabilidad (marginal) de que el equipo que juega en casa anote x goles ($x = 0, 1, 2, \dots$)	5
Parte 4.2: La probabilidad (marginal) de que el equipo que juega como visitante anote y goles ($y = 0, 1, 2, \dots$)	8
Parte 4.3: La probabilidad (conjunta) de que el equipo que juega en casa anote x goles y el equipo que juega como visitante anote y goles ($x = 0, 1, 2, \dots, y = 0, 1, 2, \dots$)	8

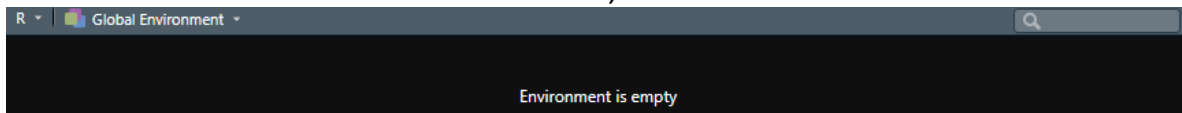
PostWork 1: Introducción a R y Software

Ejercicio 1: Importa los datos de soccer de la temporada 2019/2020 de la primera división de la liga española a R, los datos los puedes encontrar en el siguiente enlace: <https://www.football-data.co.uk/spainm.php>

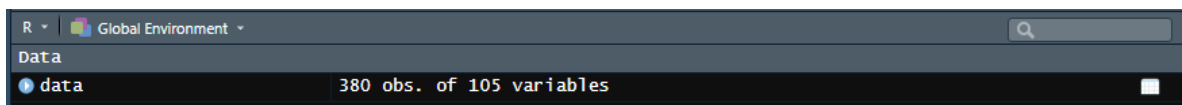
```
#<variable> <- read.csv("enlace o ubicación donde se encuentra el  
archivo csv")
```

```
#La variable almacenara los datos que contenga el archivo el cual se  
#obtendrá mediante la lectura del método read.csv que recibe como  
#parámetro tanto como el enlace del csv como la dirección local donde  
#se encuentre el archivo csv.
```

```
data <- read.csv("https://www.football-  
data.co.uk/mmz4281/1920/SP1.csv")
```



```
data <- read.csv("https://www.football-data.co.uk/mmz4281/1920/SP1.csv")  
> data <- read.csv("https://www.football-data.co.uk/mmz4281/1920/SP1.csv")  
> |
```



Ejercicio 2: Del data frame que resulta de importar los datos a R, extrae las columnas que contienen los números de goles anotados por los equipos que jugaron en casa (FTHG) y los goles anotados por los equipos que jugaron como visitante (FTAG)

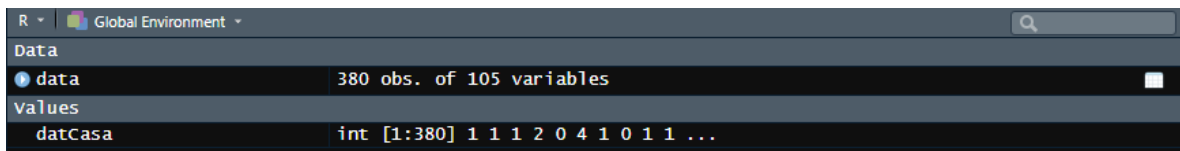
```
#<variable> <- <DataFrame>$<columna>
```

#La variable almacenara los datos los cuales se extraen de la
#<columna> del <DataFrame> que se obtienen mediante el símbolo \$.

```
datCasa <- data$FTHG
```

```
datCasa<-data$FTHG
```

```
> datCasa<-data$FTHG  
> |
```



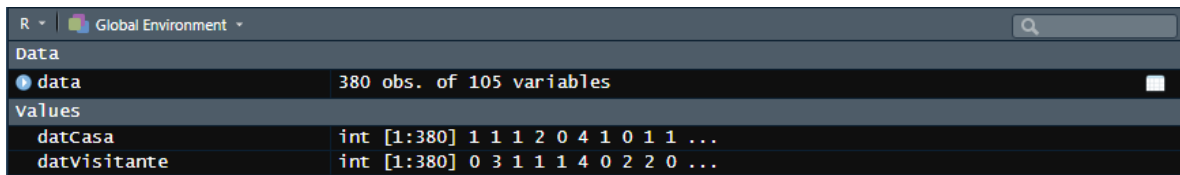
R Global Environment

Data	
data	380 obs. of 105 variables
Values	
datCasa	int [1:380] 1 1 1 2 0 4 1 0 1 1 ...

```
datVisitante <- data$FTAG
```

```
datVisitante<-data$FTAG
```

```
> datVisitante<-data$FTAG  
> |
```



R Global Environment

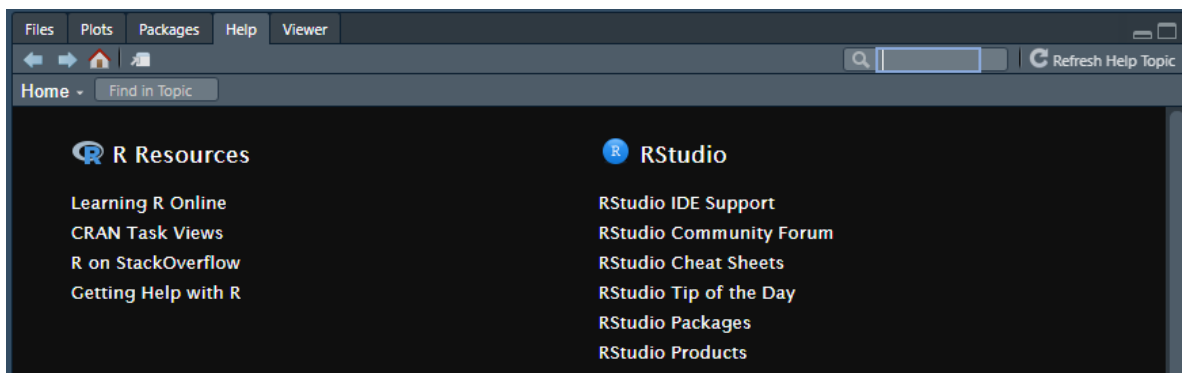
Data	
data	380 obs. of 105 variables
Values	
datCasa	int [1:380] 1 1 1 2 0 4 1 0 1 1 ...
datVisitante	int [1:380] 0 3 1 1 1 4 0 2 2 0 ...

Ejercicio 3: Consulta cómo funciona la función table en R al ejecutar en la consola ?table

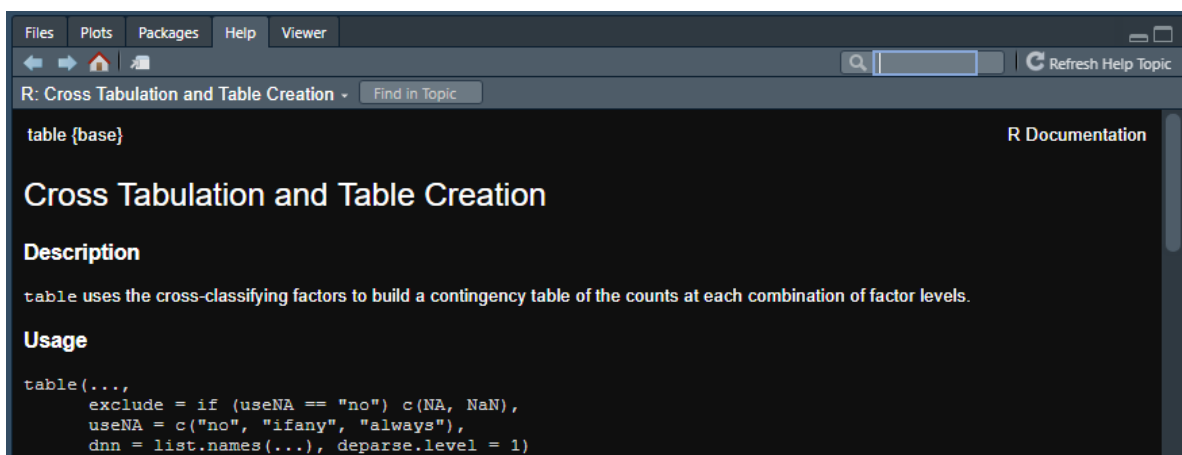
#?<método>

#Esta es una de las formas de poder acceder a la ayuda que se ofrece
#para saber el funcionamiento del <método> tanto como lo que realiza
#como los parámetros que se requieran para utilizarlo.

?table



```
> ?table
> |
```



Ejercicio 4: Posteriormente elabora tablas de frecuencias relativas para estimar las siguientes probabilidades:

Parte 4.1: La probabilidad (marginal) de que el equipo que juega en casa anote x goles ($x = 0, 1, 2, \dots$)

```
#(<variable> <- table(<vector>))
```

```
#El método table va a convertir los datos del vector en una matriz la se
#termina almacenando en la variable para su posterior manipulación
#/y/o consulta, al estar encerrados entre paréntesis al final de
#almacenar los datos en la variable se terminarán consultando.
```

```
(tablaCasa<-table(datCasa))
```

```
(tablaCasa<-table(datCasa))
```

```
> (tablaCasa<-table(datCasa))
datCasa
  0    1    2    3    4    5    6
88 132  99  38  14    8    1
> |
```

Environment	History	Connections	Tutorial
Import Dataset 112 MiB			
R Global Environment			
Data			
data 380 obs. of 105 variables			
Values			
datCasa	int [1:380]	1 1 1 2 0 4 1 0 1 1 ...	
datVisitante	int [1:380]	0 3 1 1 1 4 0 2 2 0 ...	
tablaCasa	'table' int [1:7(1d)]	88 132 99 38 14 8 1	

```
(totalGoles<-table(datVisitante,datCasa))
```

```
(totalGoles<-table(datvisitante,datCasa))
```

```
> (totalGoles<-table(datvisitante,datCasa))
      datCasa
datvisitante  0  1  2  3  4  5  6
  0  33 43 39 14  4  2  1
  1  28 49 35 14  5  3  0
  2  15 32 20  7  4  3  0
  3   8  5  3  2  0  0  0
  4   2  3  2  1  1  0  0
  5   2  0  0  0  0  0  0
> |
```

Environment	History	Connections	Tutorial
Import Dataset 112 MiB			
R Global Environment			
Data			
data 380 obs. of 105 variables			
Values			
datCasa	int [1:380]	1 1 1 2 0 4 1 0 1 1 ...	
datvisitante	int [1:380]	0 3 1 1 1 4 0 2 2 0 ...	
tablaCasa	'table' int [1:7(1d)]	88 132 99 38 14 8 1	
totalGoles	'table' int [1:6, 1:7]	33 28 15 8 2 2 43 49 32 5 ...	

```
#(<variable> <- sum(<table>))
```

#El método sum va retornar la suma de todos los elementos
#contenidos en la <table>.

```
(totalFrecAbsoluta <- sum(totalGoles))
```

```
(totalFrecAbsoluta <- sum(totalGoles))
```

```
> (totalFrecAbsoluta <- sum(totalGoles))  
[1] 380  
>
```

R Global Environment	
Data	
data	380 obs. of 105 variables
values	
datCasa	int [1:380] 1 1 1 2 0 4 1 0 1 1 ...
datVisitante	int [1:380] 0 3 1 1 1 4 0 2 2 0 ...
tablaCasa	'table' int [1:7(1d)] 88 132 99 38 14 8 1
totalFrecAbsoluta	380L
totalGoles	'table' int [1:6, 1:7] 33 28 15 8 2 2 43 49 32 5 ...

```
#(<variable> <- round(<table>/<int>,<int>))
```

#El método round recibe 2 parámetros el primero es una tabla que
#contiene la cantidad de goles por columna el cual se divide en el
#número total de goles y el segundo parámetro es para limitar el
#número de decimales a imprimir por lo que round retornara una tabla
#con los resultados por columna limitados al numero de decimales
#colocados en el segundo parámetro.

```
(FrecRelCasa <- round (tablaCasa/totalFrecAbsoluta,4))
```

```
(FrecRelCasa<-round (tablaCasa/totalFrecAbsoluta,4))
```

```
> (FrecRelCasa<-round (tablaCasa/totalFrecAbsoluta,4))  
datCasa  
      0      1      2      3      4      5      6  
0.2316 0.3474 0.2605 0.1000 0.0368 0.0211 0.0026  
>
```

R Global Environment	
Data	
data	380 obs. of 105 variables
values	
datCasa	int [1:380] 1 1 1 2 0 4 1 0 1 1 ...
datVisitante	int [1:380] 0 3 1 1 1 4 0 2 2 0 ...
FrecRelCasa	'table' num [1:7(1d)] 0.2316 0.3474 0.2605 0.1 0.0368 ...
tablaCasa	'table' int [1:7(1d)] 88 132 99 38 14 8 1
totalFrecAbsoluta	380L
totalGoles	'table' int [1:6, 1:7] 33 28 15 8 2 2 43 49 32 5 ...

Parte 4.2: La probabilidad (marginal) de que el equipo que juega como visitante anote y goles ($y = 0, 1, 2, \dots$)

```
(tablaCasa <- table(datVisitante))
```

```
(tablaCasa<-table(datvisitante))
datvisitante
  0   1   2   3   4   5
136 134  81  18   9   2
>
```

R - Global Environment	
Data	
data	380 obs. of 105 variables
Values	
datCasa	int [1:380] 1 1 1 2 0 4 1 0 1 1 ...
datVisitante	int [1:380] 0 3 1 1 1 4 0 2 2 0 ...
FrecRelCasa	'table' num [1:7(1d)] 0.2316 0.3474 0.2605 0.1 0.0368 ...
tablaCasa	'table' int [1:6(1d)] 136 134 81 18 9 2
totalFrecAbsoluta	380L
totalGoles	'table' int [1:6, 1:7] 33 28 15 8 2 2 43 49 32 5 ...

```
(FrecRelVisitante <- round(tablaCasa/totalFrecAbsoluta,4))
```

```
(FrecRelVisitante<-round (tablaCasa/totalFrecAbsoluta,4))
> (FrecRelVisitante<-round (tablaCasa/totalFrecAbsoluta,4))
datvisitante
  0   1   2   3   4   5
0.3579 0.3526 0.2132 0.0474 0.0237 0.0053
>
```

R - Global Environment	
Data	
data	380 obs. of 105 variables
Values	
datCasa	int [1:380] 1 1 1 2 0 4 1 0 1 1 ...
datVisitante	int [1:380] 0 3 1 1 1 4 0 2 2 0 ...
FrecRelCasa	'table' num [1:7(1d)] 0.2316 0.3474 0.2605 0.1 0.0368 ...
FrecRelVisitante	'table' num [1:6(1d)] 0.3579 0.3526 0.2132 0.0474 0.0237 ...
tablaCasa	'table' int [1:6(1d)] 136 134 81 18 9 2
totalFrecAbsoluta	380L
totalGoles	'table' int [1:6, 1:7] 33 28 15 8 2 2 43 49 32 5 ...

Parte 4.3: La probabilidad (conjunta) de que el equipo que juega en casa anote x goles y el equipo que juega como visitante anote y goles ($x = 0, 1, 2, \dots, y = 0, 1, 2, \dots$)


```
(FrecRelCon <- round(totalGoles/totalFrecAbsoluta,4))
```

```
(FrecRelCon<-round(totalGoles/totalFrecAbsoluta,4))
```

```
> (FrecRelCon<-round(totalGoles/totalFrecAbsoluta,4))
      datCasa
datvisitante  0    1    2    3    4    5    6
      0 0.0868 0.1132 0.1026 0.0368 0.0105 0.0053 0.0026
      1 0.0737 0.1289 0.0921 0.0368 0.0132 0.0079 0.0000
      2 0.0395 0.0842 0.0526 0.0184 0.0105 0.0079 0.0000
      3 0.0211 0.0132 0.0079 0.0053 0.0000 0.0000 0.0000
      4 0.0053 0.0079 0.0053 0.0026 0.0026 0.0000 0.0000
      5 0.0053 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
> |
```

R Global Environment	
Data	
data	380 obs. of 105 variables
Values	
datCasa	int [1:380] 1 1 1 2 0 4 1 0 1 1 ...
datVisitante	int [1:380] 0 3 1 1 1 4 0 2 2 0 ...
FrecRelCasa	'table' num [1:7(1d)] 0.2316 0.3474 0.2605 0.1 0.0368 ...
FrecRelCon	'table' num [1:6, 1:7] 0.0868 0.0737 0.0395 0.0211 0.0053 ...
FrecRelVisitante	'table' num [1:6(1d)] 0.3579 0.3526 0.2132 0.0474 0.0237 ...
tablacasa	'table' int [1:6(1d)] 136 134 81 18 9 2
totalFrecAbsoluta	380L
totalGoles	'table' int [1:6, 1:7] 33 28 15 8 2 2 43 49 32 5 ...