

## Introducción

### Tabla de los equipos a participar del campeonato

```
1 library(knitr)
2 library(kableExtra)
3
4 source("source/crear_equipos.R") #Cargar función específica
5 nombre="Ricardo Ramírez" #Emplee el nombre y apellido alumno 1
6 equipos=crear_equipos(nombre) #Función para creación de los equipos
7
8 kable(equipos)%>% # Creación de la tabla
9 kable_styling(latex_options = "striped") # Aplicación de estilos a la tabla
```

EQUIPO	OFF	DEF	SPI
Rayo Almendra	0.54	1.67	52.06
River Cebolla	2.59	0.35	51.34
Arsenal de Ciruela	2.29	2.50	90.99
Independiente de Frambuesa	2.69	1.54	35.10
Sociedad Pera	2.30	2.32	51.10
Alianza de Sandia	0.76	2.08	52.11
Atletico Nispero	2.47	1.08	51.74
Club social Limon	1.39	0.33	15.52
Cruz Higo	1.24	2.16	25.87
Real Coliflor	0.91	1.01	39.98
Agrupacion Damasco	3.14	2.51	86.85
Club Limon	1.03	1.41	47.71
America de Kiwi	0.48	1.94	51.12
Universidad de Rosa Mosqueta	1.06	0.96	39.09
Jaguars de Platano	2.55	0.26	24.72
Puerto Aceituna	3.13	1.38	29.39
San Tuna	3.43	2.22	80.45
Rio Uva	1.38	0.93	33.35
Union Alcachofa	1.27	1.17	77.81
Villa Brocoli	1.83	1.73	55.03

**NOTA:** Según resultado obtenido por la creación de la tabla, nuestro equipo es Rayo Almendra (Primero de la tabla)

- Ataque (OFF)
- Defensa (DEF)
- Poder (SPI)

## Actividad 1

### 1. Realizar análisis cuantitativo y visual de frecuencia absoluta y frecuencia relativa

#### 1.1 Tablas de frecuencia

##### 1.1.1 Distribución de frecuencia de la característica ofensiva (OFF) de los equipos.

```
library(agricolae)
tabla_freq_off=table.freq(hist(equipos$OFF,
                               breaks = "Sturges",plot=FALSE))#Crea Tabla Freq. OFF

kable(tabla_freq_off)%>% # Crea tabla de forma gráfica
kable_styling(latex_options = "striped") # Aplicación de estilos a la tabla
```

Lower	Upper	Main	Frequency	Percentage	CF	CPF
0.0	0.5	0.25	1	5	1	5
0.5	1.0	0.75	3	15	4	20
1.0	1.5	1.25	6	30	10	50
1.5	2.0	1.75	1	5	11	55
2.0	2.5	2.25	3	15	14	70
2.5	3.0	2.75	3	15	17	85
3.0	3.5	3.25	3	15	20	100

De la tabla de frecuencia para la característica ofensiva de los equipos del campeonato, podemos observar que un 50% de los equipos posee entre un 0.0 y hasta 1.5 de puntaje, mientras que el otro 50% posee desde un 1.5 hasta un 3.5 de puntaje, por lo anterior, se podría decir que existe una distribución equitativa, que divide dos hemisferios, los equipos que tienen desde 1.5 hacia abajo y los que tienen desde un 1.5 hacia arriba (50 y 50). También se puede apreciar que el 30% de los equipos solamente tiene entre un puntaje entre 1.0 y 1.5 de ofensiva.

##### 1.1.2 Distribución de frecuencia de la característica defensiva (DEF) de los equipos.

```
library(agricolae)
tabla_freq_def=table.freq(hist(equipos$DEF,
                               breaks = "Sturges",plot=FALSE))#Crea Tabla Freq. DEF

kable(tabla_freq_def)%>% # Crea tabla de forma gráfica
kable_styling(latex_options = "striped") # Aplicación de estilos a la tabla
```

Lower	Upper	Main	Frequency	Percentage	CF	CPF
0.0	0.5	0.25	3	15	3	15
0.5	1.0	0.75	2	10	5	25
1.0	1.5	1.25	5	25	10	50
1.5	2.0	1.75	4	20	14	70
2.0	2.5	2.25	5	25	19	95
2.5	3.0	2.75	1	5	20	100

De la tabla de frecuencia para la característica defensiva (DEF) de los equipos, se puede observar que se divide la distribución de los equipos en partes iguales, es decir, el 50% de los equipos tiene entre un 0.0 hasta un 1.5 de puntaje y el otro 50% tiene desde un 1.5 hasta un 3.0 de puntaje, lo que también indica que existe un rango de 1.5 entre el máximo que alcanza cierto grupo y el mínimo que alcanza el otro grupo. También se puede observar que solamente 1 equipo se encuentra entre el puntaje máximo a obtener (entre 2.5 y 3.0) el que representa al 5% de la población del espacio muestral. Por otra parte, otro dato que se puede observar, es que hay 2 grupos que componen el 50% del total de los equipos, los que poseen entre 1.0 y 1.5 puntos (25%) y los que poseen 2.0 y 2.5 puntos (25%).

### 1.1.3 Distribución de frecuencia de SPI para los equipos

```
library(agricolae)
tabla_freq_spi=table.freq(hist(equipos$SPI,
                               breaks = "Sturges",plot=FALSE)) # Tabla de frecuencia para los SPI
kable(tabla_freq_spi)%>%
kable_styling(latex_options = "striped") # Aplicación de estilos a la tabla
```

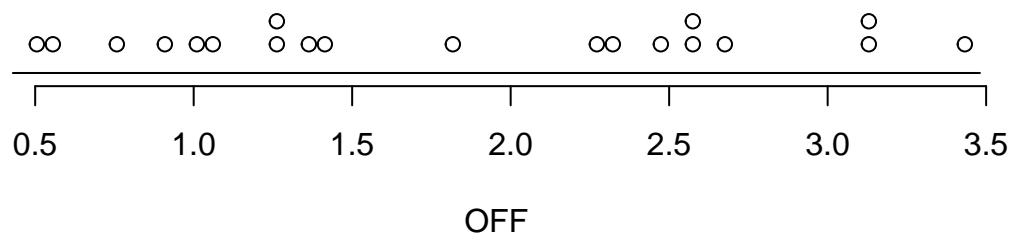
Lower	Upper	Main	Frequency	Percentage	CF	CPF
10	20	15	1	5	1	5
20	30	25	3	15	4	20
30	40	35	4	20	8	40
40	50	45	1	5	9	45
50	60	55	7	35	16	80
60	70	65	0	0	16	80
70	80	75	1	5	17	85
80	90	85	2	10	19	95
90	100	95	1	5	20	100

De la tabla de frecuencia obtenida para el puntaje de poder adquisitivo (SPI) de los equipos, se puede observar que un 35% tiene desde 50 a 60 como máximo, lo que representa a 7 de los 20 equipos del total del espacio muestral y que solamente 1 equipo se encuentra entre el rango de puntaje máximo alcanzado (90 a 100 puntos) y solamente 1 equipo se encuentra entre el rango de puntaje mínimo alcanzado (10 a 20 puntos). También se puede visualizar que un 45% de los equipos alcanza entre 10 a 50 puntos como máximo y un 55% alcanza desde 50 puntos a 100 puntos como máximo, siendo 100 el máximo a alcanzar.

## 1.2 Diagramas de puntos

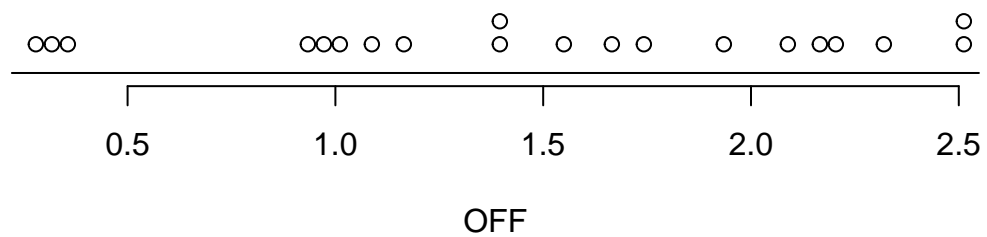
### 1.2.1 Diagrama de puntos para visualizar la distribución de los puntajes ofensivos de los equipos

```
library("BHH2")
dotPlot(equipos$OFF,xlab = "OFF")
```



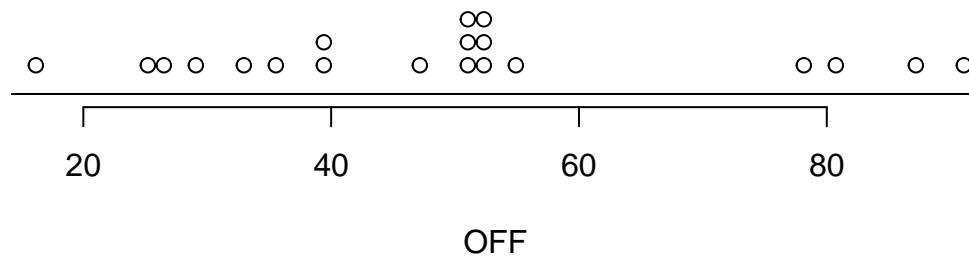
**1.2.2 Diagrama de puntos para visualizar la distribución de los puntajes defensivos de los equipos**

```
library("BHH2")
dotPlot(equipos$DEF,xlab = "OFF")
```



### 1.2.3 Diagrama de puntos para visualizar la distribución de los puntajes de poder adquisitivo de los equipos

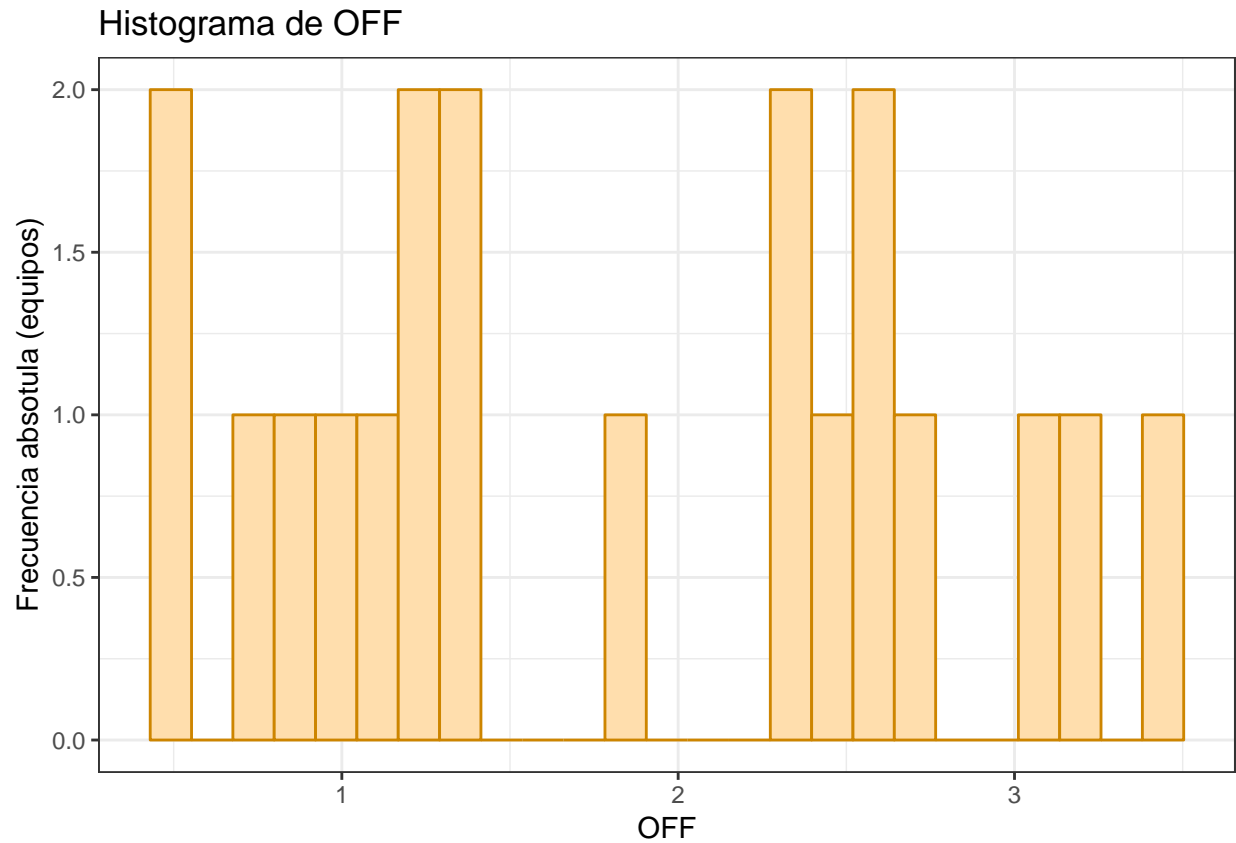
```
library("BHH2")
dotPlot(equipos$SPI,xlab = "OFF")
```



## 1.3 Histogramas

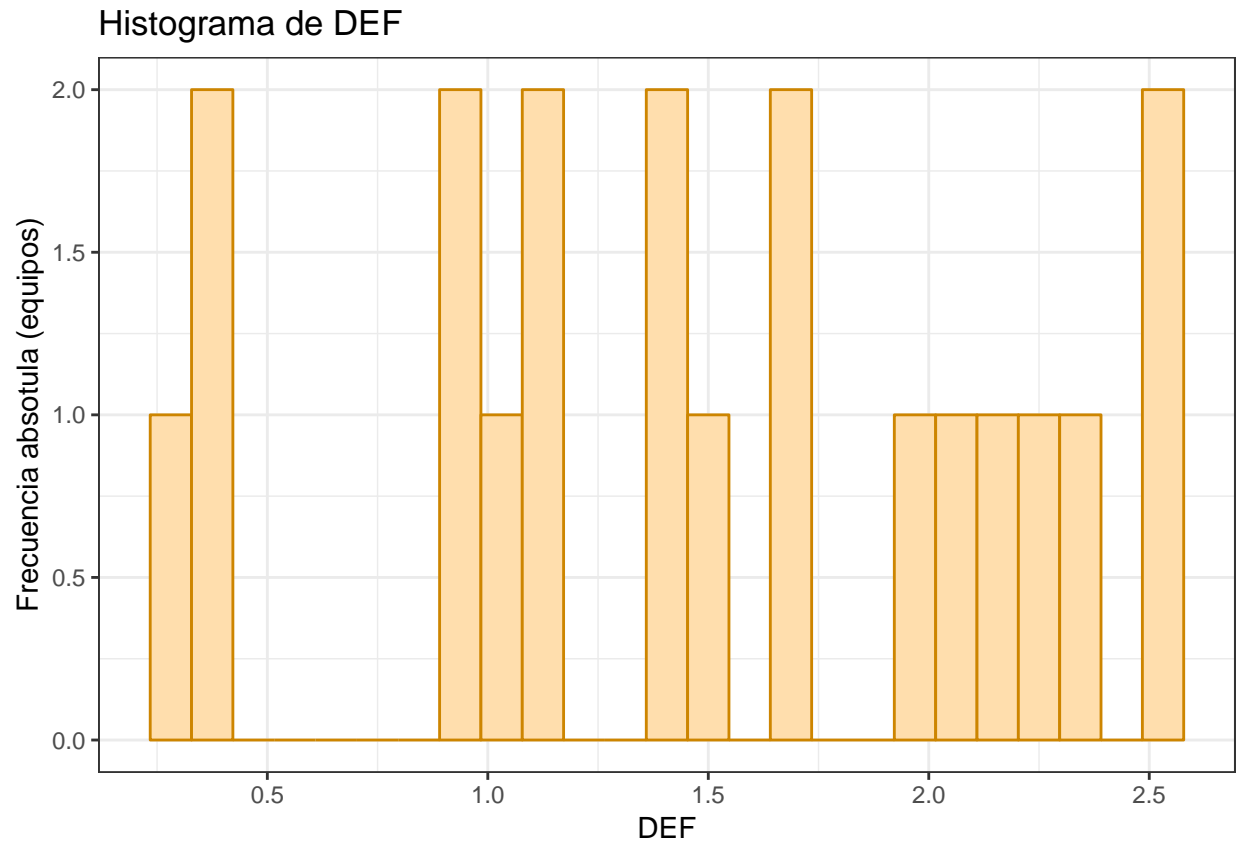
### 1.3.1 Histograma para OFF

```
library("ggplot2")
grafico=ggplot(equipos,aes(equipos$OFF)) # Gráfico y datos base
#Histograma (25 niveles) (colores- http://www.stat.columbia.edu/~tzheng/files/Rcolor.pdf)
grafico = grafico + geom_histogram(bins=25,fill="navajowhite",color="orange3")
grafico = grafico + theme_bw() # Visualización estándar en blanco y negro
grafico = grafico + ylab("Frecuencia absotula (equipos)") + xlab("OFF")
grafico = grafico + ggtitle("Histograma de OFF")
plot(grafico)
```



### 1.3.2 Histograma para DEF

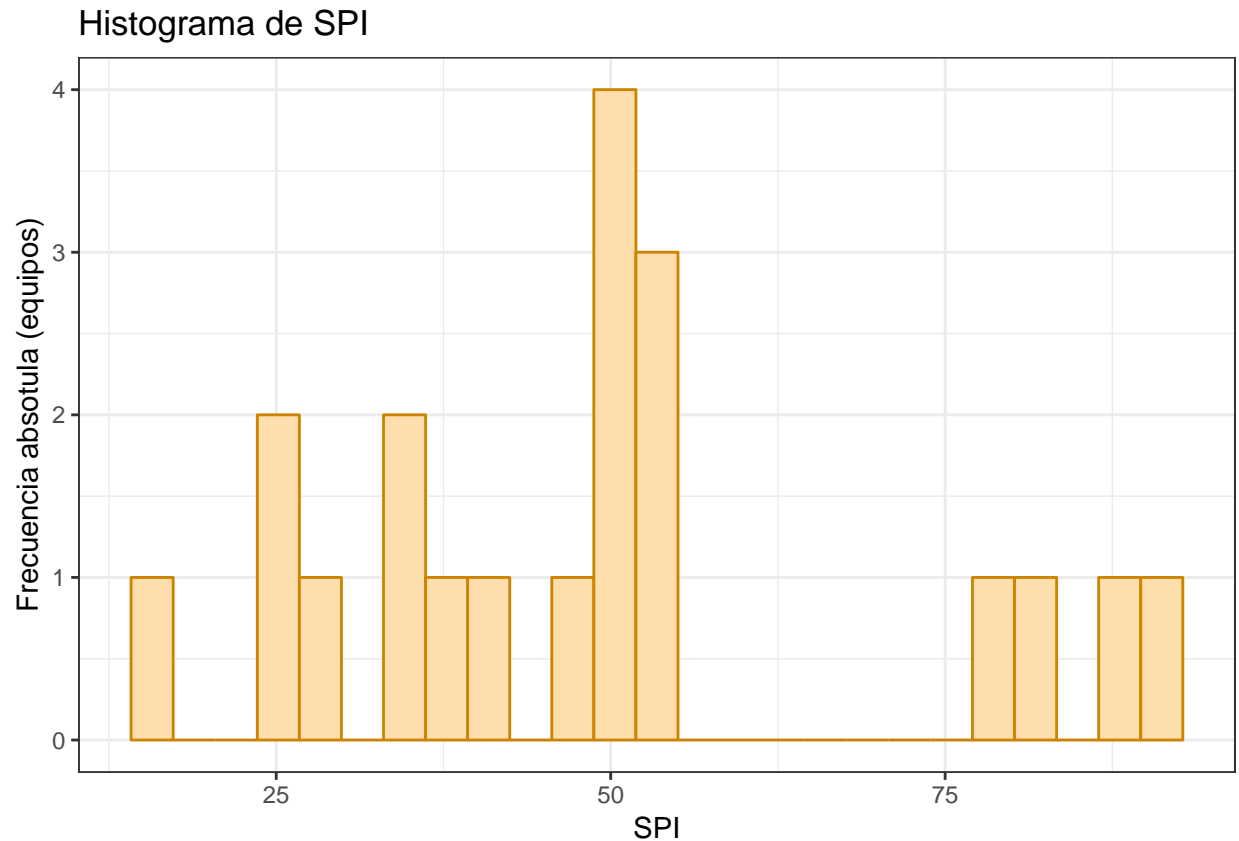
```
library("ggplot2")
grafico=ggplot(equipos,aes(equipos$DEF)) # Gráfico y datos base
#Histograma (25 niveles) (colores- http://www.stat.columbia.edu/~tzheng/files/Rcolor.pdf)
grafico = grafico + geom_histogram(bins=25,fill="navajowhite",color="orange3")
grafico = grafico + theme_bw() # Visualización estándar en blanco y negro
grafico = grafico + ylab("Frecuencia absoluta (equipos)") + xlab("DEF")
grafico = grafico + ggtitle("Histograma de DEF")
plot(grafico)
```



### 1.3.3 Histograma para SPI

```
library("ggplot2")
grafico=ggplot(equipos,aes(equipos$SPI)) # Gráfico y datos base
#Histograma (25 niveles) (colores- http://www.stat.columbia.edu/~tzheng/files/Rcolor.pdf)
grafico = grafico + geom_histogram(bins=25,fill="navajowhite",color="orange3")
grafico = grafico + theme_bw() # Visualización estándar en blanco y negro
grafico = grafico + ylab("Frecuencia absoluta (equipos)") + xlab("SPI")
grafico = grafico + ggtitle("Histograma de SPI")
plot(grafico)
```





## 1.4 Análisis y comparativa entre equipos de mejor y peor rendimiento y nuestro equipo (Rayo Almería)

### 1.4.1 Mínimo y máximo de puntaje ofensivo de los equipos

```
rango=range(equipos$OFF) # Rango mínimo y máximo de OFF de los equipos
rango
```

```
## [1] 0.48 3.43
```

El puntaje que posee el equipo con más bajo rendimiento en ofensiva es de 0.48 puntos, mientras que el equipo que posee el más alto puntaje de ofensiva es 3.43. El equipo Rayo Almería tiene un puntaje ofensivo de 0.54, lo que lo aleja solamente 0.06 puntos del más bajo y una diferencia de 2.89 del equipo con mejor rendimiento ofensivo, lo que posiciona al Rayo Almería entre el 20% con más bajo puntaje de la tabla.

### 1.4.2 Mínimo y máximo de puntaje defensivo de los equipos

```
rango=range(equipos$DEF) # Rango mínimo y máximo de DEF de los equipos
rango
```

```
## [1] 0.26 2.51
```

El puntaje mínimo obtenido para la característica defensiva de los equipos es de 0.26, mientras el máximo es de 2.51 puntos. El equipo Rayo Almendra, para esta categoría tiene un puntaje de 1.67, alejandolo 1.41 puntos del mínimo obtenido y 0.84 del puntaje máximo alcanzado, lo que posiciona a nuestro equipo entre el 50% con mejor rendimiento defensivo del campeonato.

### 1.4.3 Mínimo y máximo de SPI de los equipos

```
rango=range(equipos$SPI) # Rango mínimo y máximo de SPI de los equipos
rango
```

```
## [1] 15.52 90.99
```

El equipo que posee el mínimo SPI tiene 15.52 puntos, mientras que el que posee el máximo puntaje tiene 90.99 puntos. El equipo Rayo Almendra tiene para esta categoría 52.06 puntos, lo que lo aleja 36.54 puntos del peor y 38.93 puntos del mejor. Lo que posiciona a nuestro equipo entre el 55% de mejor puntaje en SPI del campeonato y perteneciendo su puntaje a un 35% del total del espacio muestral.

Generalmente la diferencia para las categorías entre los peores y mejores rendimientos es la siguiente:

- El peor equipo tiene una diferencia de **2.95** puntos del equipo con mejor ofensiva.
- El peor equipo tiene una diferencia de **2.25** puntos del equipo con mejor defensiva.
- El peor equipo tiene una diferencia de **75.47** puntos del equipo con mejor poder adquisitivo.

## 1.5 Análisis de equipos aplicando medidas de centralidad (Media)

### 1.5.1 Media de OFF

```
1 mediana_off = median(equipos$OFF)
2 print(paste(mediana_off))
```

```
## [1] "1.61"
```

El equipo Rayo Almendra tiene un puntaje de 0.54, lo que lo posiciona por debajo de la puntuación media para la característica ofensiva.

### Equipos por sobre la Media de (OFF)

```
1 library("dplyr",warn.conflicts = F)
2 filtro1=filter(equipos, OFF > mediana_off)
3 filtro1=filtro1[order(filtro1$OFF , decreasing = TRUE),]
4 #kable(filtro1)
5 kable(filtro1)%>%
6 kable_styling(latex_options = "striped") # Aplicación de estilos a la tabla
```

	EQUIPO	OFF	DEF	SPI
9	San Tuna	3.43	2.22	80.45
6	Agrupacion Damasco	3.14	2.51	86.85
8	Puerto Aceituna	3.13	1.38	29.39
3	Independiente de Frambuesa	2.69	1.54	35.10
1	River Cebolla	2.59	0.35	51.34
7	Jaguars de Platano	2.55	0.26	24.72
5	Atletico Nispero	2.47	1.08	51.74
4	Sociedad Pera	2.30	2.32	51.10
2	Arsenal de Ciruela	2.29	2.50	90.99
10	Villa Brocoli	1.83	1.73	55.03

```
1 paste(count(filtro1)$n, "equipos por sobre la media OFF de un total de", count(equipos)$n )
```

```
## [1] "10 equipos por sobre la media OFF de un total de 20"
```

### 1.5.2 Media de DEF

```
1 mediana_def = median(equipos$DEF)
2 print(paste(mediana_def))
```

```
## [1] "1.475"
```

El equipo Rayo Almendra para la puntuación de defensiva tiene un puntaje de 1.67, lo que lo posiciona por sobre el puntaje medio (1.475).

### Equipos por sobre la Media de (DEF)

```
1 library("dplyr", warn.conflicts = F)
2 filtro2=filter(equipos, DEF > mediana_def)
3 filtro2=filtro2[order(filtro2$DEF, decreasing = TRUE),]
4 #kable(filtro2)
5 kable(filtro2)%>%
6 kable_styling(latex_options = "striped") # Aplicación de estilos a la tabla
```

	EQUIPO	OFF	DEF	SPI
7	Agrupacion Damasco	3.14	2.51	86.85
2	Arsenal de Ciruela	2.29	2.50	90.99
4	Sociedad Pera	2.30	2.32	51.10
9	San Tuna	3.43	2.22	80.45
6	Cruz Higo	1.24	2.16	25.87
5	Alianza de Sandia	0.76	2.08	52.11
8	America de Kiwi	0.48	1.94	51.12
10	Villa Brocoli	1.83	1.73	55.03
1	Rayo Almendra	0.54	1.67	52.06
3	Independiente de Frambuesa	2.69	1.54	35.10

```
1 paste(count(filtro2)$n, "equipos por sobre la media DEF de un total de", count(equipos)$n )
## [1] "10 equipos por sobre la media DEF de un total de 20"
```

Se puede observar al equipo en la posición nº9 de la tabla de equipos sobre la media en DEF.

## Media de SPI

```
1 mediana_spi = median(equipos$SPI)
2 print(paste(mediana_spi))
## [1] "51.11"
```

El equipo Rayo Almendra tiene un puntaje de 52.06 para SPI, lo que lo posiciona sobre la media la cual es 51.11.

## Equipos por sobre la Media de (SPI)

```
1 library("dplyr", warn.conflicts = F)
2 filtro3=filter(equipos, SPI > mediana_spi)
3 filtro3=filtro3[order(filtro3$SPI, decreasing = TRUE),]
4 #kable(filtro3)
5 kable(filtro3)%>%
6 kable_styling(latex_options = "striped") # Aplicación de estilos a la tabla
```

	EQUIPO	OFF	DEF	SPI
3	Arsenal de Ciruela	2.29	2.50	90.99
6	Agrupacion Damasco	3.14	2.51	86.85
8	San Tuna	3.43	2.22	80.45
9	Union Alcachofa	1.27	1.17	77.81
10	Villa Brocoli	1.83	1.73	55.03
4	Alianza de Sandia	0.76	2.08	52.11
1	Rayo Almendra	0.54	1.67	52.06
5	Atletico Nispero	2.47	1.08	51.74
2	River Cebolla	2.59	0.35	51.34
7	America de Kiwi	0.48	1.94	51.12

```
1 paste(count(filtro3)$n, "equipos por sobre la media SPI de un total de", count(equipos)$n )
## [1] "10 equipos por sobre la media SPI de un total de 20"
```

El equipo Rayo Almendra queda en la posición nº7 de la tabla de los equipos sobre la media de SPI.

## Equipos por sobre la Media de (OFF), Media de (DEF) y Media de (SPI)

```

1 filtro_final=filter(equipos, OFF > mediana_off)
2 filtro_final=filter(filtro_final, DEF > mediana_def)
3 filtro_final=filter(filtro_final, SPI > mediana_spi)
4 #kable(filtro_final)
5 kable(filtro_final)%>%
6 kable_styling(latex_options = "striped") # Aplicación de estilos a la tabla

```

EQUIPO	OFF	DEF	SPI
Arsenal de Ciruela	2.29	2.50	90.99
Agrupacion Damasco	3.14	2.51	86.85
San Tuna	3.43	2.22	80.45
Villa Brocoli	1.83	1.73	55.03

```

1 paste(count(filtro_final)$n, "equipos por sobre todas las medias de un total de", count(equipos)$n )

```

```
## [1] "4 equipos por sobre todas las medias de un total de 20"
```

```

#install.packages('gtools')
library("gtools")
source("source/jugar_partidos.R")
partidos=jugar_partidos(equipos)
head(partidos,40) # Muestra los primeros 20 partidos

```

```

##           EQ1           EQ2 OFF1 DEF1 SPI1 OFF2
## 1 Agrupacion Damasco Alianza de Sandia 3.14 2.51 86.85 0.76
## 2 Agrupacion Damasco America de Kiwi 3.14 2.51 86.85 0.48
## 3 Agrupacion Damasco Arsenal de Ciruela 3.14 2.51 86.85 2.29
## 4 Agrupacion Damasco Atletico Nispero 3.14 2.51 86.85 2.47
## 5 Agrupacion Damasco Club Limon 3.14 2.51 86.85 1.03
## 6 Agrupacion Damasco Club social Limon 3.14 2.51 86.85 1.39
## 7 Agrupacion Damasco Cruz Higo 3.14 2.51 86.85 1.24
## 8 Agrupacion Damasco Independiente de Frambuesa 3.14 2.51 86.85 2.69
## 9 Agrupacion Damasco Jaguares de Platano 3.14 2.51 86.85 2.55
## 10 Agrupacion Damasco Puerto Aceituna 3.14 2.51 86.85 3.13
## 11 Agrupacion Damasco Rayo Almendra 3.14 2.51 86.85 0.54
## 12 Agrupacion Damasco Real Coliflor 3.14 2.51 86.85 0.91
## 13 Agrupacion Damasco Rio Uva 3.14 2.51 86.85 1.38
## 14 Agrupacion Damasco River Cebolla 3.14 2.51 86.85 2.59
## 15 Agrupacion Damasco San Tuna 3.14 2.51 86.85 3.43
## 16 Agrupacion Damasco Sociedad Pera 3.14 2.51 86.85 2.30
## 17 Agrupacion Damasco Union Alcachofa 3.14 2.51 86.85 1.27
## 18 Agrupacion Damasco Universidad de Rosa Mosqueta 3.14 2.51 86.85 1.06
## 19 Agrupacion Damasco Villa Brocoli 3.14 2.51 86.85 1.83
## 20 Alianza de Sandia Agrupacion Damasco 0.76 2.08 52.11 3.14
## 21 Alianza de Sandia America de Kiwi 0.76 2.08 52.11 0.48
## 22 Alianza de Sandia Arsenal de Ciruela 0.76 2.08 52.11 2.29
## 23 Alianza de Sandia Atletico Nispero 0.76 2.08 52.11 2.47
## 24 Alianza de Sandia Club Limon 0.76 2.08 52.11 1.03
## 25 Alianza de Sandia Club social Limon 0.76 2.08 52.11 1.39
## 26 Alianza de Sandia Cruz Higo 0.76 2.08 52.11 1.24

```

## 27	Alianza de Sandia	Independiente de	Frambuesa	0.76	2.08	52.11	2.69
## 28	Alianza de Sandia	Jaguares de	Platano	0.76	2.08	52.11	2.55
## 29	Alianza de Sandia	Puerto	Aceituna	0.76	2.08	52.11	3.13
## 30	Alianza de Sandia	Rayo	Almendra	0.76	2.08	52.11	0.54
## 31	Alianza de Sandia	Real	Coliflor	0.76	2.08	52.11	0.91
## 32	Alianza de Sandia	Rio	Uva	0.76	2.08	52.11	1.38
## 33	Alianza de Sandia	River	Cebolla	0.76	2.08	52.11	2.59
## 34	Alianza de Sandia	San	Tuna	0.76	2.08	52.11	3.43
## 35	Alianza de Sandia	Sociedad	Pera	0.76	2.08	52.11	2.30
## 36	Alianza de Sandia	Union	Alcachofa	0.76	2.08	52.11	1.27
## 37	Alianza de Sandia	Universidad de	Rosa Mosqueta	0.76	2.08	52.11	1.06
## 38	Alianza de Sandia	Villa	Brocoli	0.76	2.08	52.11	1.83
## 39	America de Kiwi	Agrupacion	Damasco	0.48	1.94	51.12	3.14
## 40	America de Kiwi	Alianza de	Sandia	0.48	1.94	51.12	0.76
##	DEF2	SPI2	GEQ1	GEQ2			
## 1	2.08	52.11	3	1			
## 2	1.94	51.12	3	1			
## 3	2.50	90.99	3	2			
## 4	1.08	51.74	3	3			
## 5	1.41	47.71	3	1			
## 6	0.33	15.52	3	2			
## 7	2.16	25.87	4	1			
## 8	1.54	35.10	3	3			
## 9	0.26	24.72	3	3			
## 10	1.38	29.39	3	3			
## 11	1.67	52.06	3	1			
## 12	1.01	39.98	3	1			
## 13	0.93	33.35	3	2			
## 14	0.35	51.34	2	3			
## 15	2.22	80.45	3	3			
## 16	2.32	51.10	3	2			
## 17	1.17	77.81	2	2			
## 18	0.96	39.09	3	1			
## 19	1.73	55.03	3	2			
## 20	2.51	86.85	1	3			
## 21	1.94	51.12	1	1			
## 22	2.50	90.99	1	2			
## 23	1.08	51.74	1	3			
## 24	1.41	47.71	1	1			
## 25	0.33	15.52	1	2			
## 26	2.16	25.87	1	1			
## 27	1.54	35.10	1	3			
## 28	0.26	24.72	1	3			
## 29	1.38	29.39	1	3			
## 30	1.67	52.06	1	1			
## 31	1.01	39.98	1	1			
## 32	0.93	33.35	1	2			
## 33	0.35	51.34	0	3			
## 34	2.22	80.45	1	3			
## 35	2.32	51.10	1	2			
## 36	1.17	77.81	0	2			
## 37	0.96	39.09	1	2			
## 38	1.73	55.03	1	2			
## 39	2.51	86.85	0	3			

```
## 40 2.08 52.11    1    1
```

## 1.6 Actividades Asumiendo probabilidad uniforme

¿Cuántos partidos se jugaron? Formule, explique y resuelva el problema considerando los conceptos de combinación/permutación.

Ya que el orde de los partidos si importa, ya que se juega uno de ida y uno de vuelta, es decir, 2 partidos por par de equipos, se utiliza el concepto de permutación, en donde utilizando la formula

$${}_n P_k = \frac{n!}{(n-k)!}$$

en donde N es la cantidad total de equipos (20) y n es la cantidad de partidos para dos equipos, se reflejará el resultado de el total de permutaciones del campeonato.

```
library(gtools)
n = 2 # Cantidad de partidos a jugar por par de equipos
N = 20 # Cantidad de equipos
permutaciones= nrow(permutations(N,n))
print(permutaciones)
```

```
## [1] 380
```

Analizando los resultados, la cantidad de partidos a jugar fueron 380, considerando que por cada partido existe uno de ida y uno de vuelta.

Cuál es la probabilidad de que su equipo al finalizar el torneo sea el ganador?

por distribución de bernoulli tenemos  $x$  = cantidad de veces que existe un ganador(1); ya que existen 20 equipos, el espacio muestral es de 20. Se considera éxito ganar, entonces la probabilidad es  $1/20$ . Se considera fracaso no ganar entonces  $q=1-P = 1-(1/20) = 19/20$ .

La probabilidad de que el equipo sea ganador es  $x = 1$  entonces  $P(x=1) = (1/20)^1 * (19/20)^{(1-1)} = 0.05$ , lo que nos indica que tenemos una probabilidad de 0.05% de ganar.