

Estadística aplicada al deporte. Varios ejemplos con datos de baloncesto

Guillermo Vinué
Matemático. Doctor en Estadística
Analista de datos

Universidad de Verano
Universidad de Valencia
Gandía

13 de Julio 2023

universitat d'estiu UVgàndia

VNIVERSITAT
D VALÈNCIA

Resumen

1 Motivación

2 Recolección de datos

3 Estadística descriptiva

4 Estadística predictiva

5 Conclusiones

6 Referencias

Aclaración

- **Estadística:** Disciplina matemática dedicada al estudio y análisis de datos de cualquier tipo.
- **Estadísticas:** Datos numéricos que representan la actividad de los equipos y jugadores.

La Estadística nos permite estudiar las estadísticas...

Motivación



- Afición personal por el deporte, y, en particular, por el baloncesto.
- Interés en unir mi afición con mi profesión.
- El baloncesto genera una gran cantidad de datos (*estadísticas*), tanto a nivel individual como colectivo.
- En la **NBA** (la liga profesional masculina en EEUU), la recolección de datos y su análisis ha sido siempre práctica habitual.



- La NBA cuenta con una excelente página web en la que se puede consultar una gran cantidad de información.

www.nba.com/stats

The screenshot shows the NBA Advanced Stats website interface. At the top, there is a navigation bar with the NBA logo and links for Games, Schedule, Watch, News, Stats, Standings, Teams, Players, Fantasy, and NBA Bet. On the right side of the navigation bar, there are links for League Pass, Store, Tickets, and Sign In. Below the navigation bar, there is a secondary menu with links for Home, Players, Teams, Leaders, Stats 101, Cume Stats, Lineups Tool, Media Central Game Stats, Draft, Quick Links, and Contact Us. The main content area is divided into several sections:

- NBA Advanced Stats**: A header section with the text "Stats Home" and a sub-menu on the right containing links for Franchise History, Statistical Minimums, Glossary, Transactions, Fantasy News, Articles, Weekly Stats Archive, FAQ, Video & Tracking Status, and Video Rulebook.
- PLAYERS**: A section titled "YESTERDAY'S LEADERS" for the date 03/22/2023. It features three columns of data:

POINTS	REBOUND	ASSISTS
1. Lauri Markkanen UTA 40	1. Giannis Antetokounmpo MIL 14	1. Luka Doncic DAL 17
2. Jaren Jackson Jr. MEM 37	2. Andre Drummond CHI 12	2. Stephen Curry GSW 13
3. Jimmy Butler MIA 35	3. Rudy Gobert MIN 12	3. Damian Lillard POR 12
4. Devin Booker PHX 33	4. Nikola Jokic DEN 12	4. Austin Reaves LAL 11
5. Jalen Green HOU 32	2. Kevon Looney GSW 12	4. Fred VanVleet TOR 11
- MORE STATS**: A section titled "TOTAL POINTS" listing the top scorers:

Jayson Tatum	2044
Joel Embiid	1958
Shai Gilgeous-Alexander	1915

 Below this is a section titled "TOTAL REBOUNDS" listing the top rebounders:

Domantas Sabonis	875
Nikola Vucevic	810
Nikola Jokic	774

- En España, el tratamiento de los datos a nivel oficial sigue siendo muy rudimentario:
 - La liga profesional masculina es la liga **ACB (Asociación de Clubes de Baloncesto)** (por motivos de patrocinio la ACB se conoce como Liga Endesa) y es la que ofrece al público las estadísticas.
 - Estadísticas básicas, que, aunque son informativas, **pueden ser ampliadas.**
 - Representación con tablas, sin posibilidad de interacción. **Ausencia de visualizaciones.**

ESTADÍSTICAS

VALENCIA BASKET

TEMPORADA 2022-2023

LIGA ENDESA

TODAS LAS FASES

MEDIA

Liga Endesa - TEMPORADA 2022-2023 - TODAS LAS FASES

	PAR			MIN			PT			TIROS DE 3			TIROS DE 2			TIROS LIBRES			REBOTES			ASI			BALONES			TAPONES			FALTAS			VAL
	Jug	Jug	5i	Con	Int	%	Con	Int	%	Con	Int	%	Def	Ofe	Tot	Efe	Rec	Per	Fav	Con	Mat	Com	Rec	+/-										
Jasiel Rivero	18	21:21	0,2	14,4	0,6	1,1	55,0%	4,7	6,7	70,0%	3,2	5,3	60,4%	2,5	2,0	4,5	0,7	0,7	1,0	0,8	0,1	0,3	2,1	4,4	4,0	17,8								
Bojan Dubljevic	22	21:34	0,9	9,9	1,4	3,7	37,0%	1,9	3,5	53,2%	2,0	2,5	83,3%	4,1	2,0	6,0	1,5	0,8	1,6	0,3	0,5	0,0	1,7	2,9	1,0	13,2								
Chris Jones	25	22:04	0,8	12,2	1,2	3,0	40,5%	3,4	6,3	54,1%	1,8	2,1	84,6%	1,6	0,2	1,8	4,8	1,0	2,2	0,0	0,3	0,0	1,5	2,2	3,0	13,0								
Xabi López-Arostegui	25	20:16	0,3	8,6	1,0	2,9	36,1%	2,0	3,7	52,7%	1,6	1,8	88,6%	2,8	0,7	3,5	1,4	1,5	1,2	0,1	0,2	0,0	1,9	2,5	0,0	10,5								
James Webb III	18	19:42	0,7	9,1	1,7	4,4	37,5%	1,7	2,6	63,8%	0,8	1,0	77,8%	4,4	1,1	5,5	0,4	0,8	1,1	0,4	0,2	0,4	2,3	1,3	1,0	10,1								
Jared Harper	13	15:11	0,1	8,6	1,4	3,5	40,0%	0,8	2,4	32,3%	2,9	3,3	88,4%	0,5	0,1	0,5	2,4	0,5	1,6	0,1	0,2	0,0	2,0	2,9	-2,0	7,2								
Sam Van Rossom	6	16:39	0,3	7,8	1,5	3,7	40,9%	0,8	1,7	50,0%	1,7	1,8	90,9%	1,0	0,0	1,0	1,8	0,3	1,0	0,0	0,3	0,0	1,2	1,5	-4,0	6,8								
Klemen Prepelic	21	16:16	0,2	6,2	0,9	3,9	22,0%	0,8	1,7	48,6%	2,0	2,6	76,4%	1,0	0,2	1,2	2,9	0,3	1,6	0,0	0,2	0,0	1,3	2,8	0,0	5,8								
Jaime Pradilla	23	15:02	0,3	5,1	0,4	1,0	43,5%	1,6	3,3	48,1%	0,6	0,8	73,7%	1,2	1,1	2,3	1,0	0,7	0,7	0,1	0,4	0,2	1,1	1,0	0,0	5,6								
Kyle Alexander	21	14:29	0,2	5,2	0,0	0,4	11,1%	2,1	3,1	66,7%	0,9	1,4	63,3%	1,7	1,0	2,7	0,4	0,3	0,9	0,6	0,2	0,9	2,1	1,3	-2,0	5,3								
Shannon Evans	10	16:50	0,0	5,9	0,7	1,9	36,8%	1,3	2,7	48,1%	1,2	1,8	66,7%	1,4	0,0	1,4	2,4	0,8	1,9	0,1	0,0	0,1	1,3	1,5	-6,0	5,7								
Jonah Radebaugh	26	19:02	0,7	6,6	1,0	3,0	32,5%	1,3	2,7	50,7%	1,0	1,1	89,7%	1,3	0,3	1,6	1,2	0,7	0,8	0,1	0,2	0,0	2,0	1,0	0,0	4,8								
Víctor Claver	23	18:26	0,3	4,0	0,8	2,0	39,1%	0,7	1,6	40,5%	0,3	0,3	87,5%	2,2	0,5	2,7	0,8	0,5	0,8	0,3	0,1	0,1	1,6	0,9	-1,0	4,6								
Martin Hermannsson	5	17:30	0,2	2,6	0,4	2,0	20,0%	0,4	2,0	20,0%	0,6	0,6	100,0%	1,4	0,0	1,4	3,6	0,6	1,0	0,0	0,0	0,0	0,8	1,0	0,0	4,2								
Josep Puerto	25	16:10	0,7	5,2	0,8	2,4	34,4%	1,1	1,6	65,9%	0,5	0,7	76,5%	1,4	0,5	2,0	0,4	0,4	0,6	0,0	0,2	0,1	1,8	0,8	0,0	4,0								
Guillem Ferrando	8	12:54	0,3	3,9	1,1	2,5	45,0%	0,1	0,1	100,0%	0,3	0,4	66,7%	0,4	0,4	0,8	1,9	0,5	1,4	0,0	0,0	0,0	2,6	0,4	3,0	1,9								
Millán Jiménez	0	0:00	0,0	0,0	0,0	0,0	0,0%	0,0	0,0	0,0%	0,0	0,0	0,0%	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0								
Sergio De Larrea	0	0:00	0,0	0,0	0,0	0,0	0,0%	0,0	0,0	0,0%	0,0	0,0	0,0%	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0								
Gonzalo Bressan	1	1:54	0,0	0,0	0,0	0,0	0,0%	0,0	1,0	0,0%	0,0	0,0	0,0%	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	0,0	-7,0	-2,0									
Lucas Mari	3	1:00	0,0	0,0	0,0	0,7	0,0%	0,0	0,0	0,0%	0,0	0,0	0,0%	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,3	0,0	0,0	-1,0									
Totales	26			83,5	10,2	28,6	35,7%	19,0	34,6	55,1%	14,8	19,2	77,2%	23,5	9,5	32,9	17,5	10,5	13,4	2,2	2,4	1,7	19,5	20,5		93,5								

- España es una potencia en baloncesto a nivel mundial, a la altura del baloncesto NBA (vigentes campeones de Europa y del Mundo).



- Pero en el ámbito de las estadísticas, la NBA todavía nos lleva ventaja. Si ellos utilizan sus datos de manera avanzada...

¿Por qué no hacerlo nosotros también?



- Posibilidad de innovación a partir de la información generada en cada partido:
 - **Nueva representación de las estadísticas tradicionales.**
 - **Nuevos análisis a partir de los datos jugada a jugada.**
- **Importancia de visualizar datos:** Identificar tendencias y patrones en los datos con más facilidad.
- Herramienta de software: R, www.r-project.org



Breve comentario sobre R



Es que aquí usamos SAS...

**Es que pedimos saber
Tableau...**

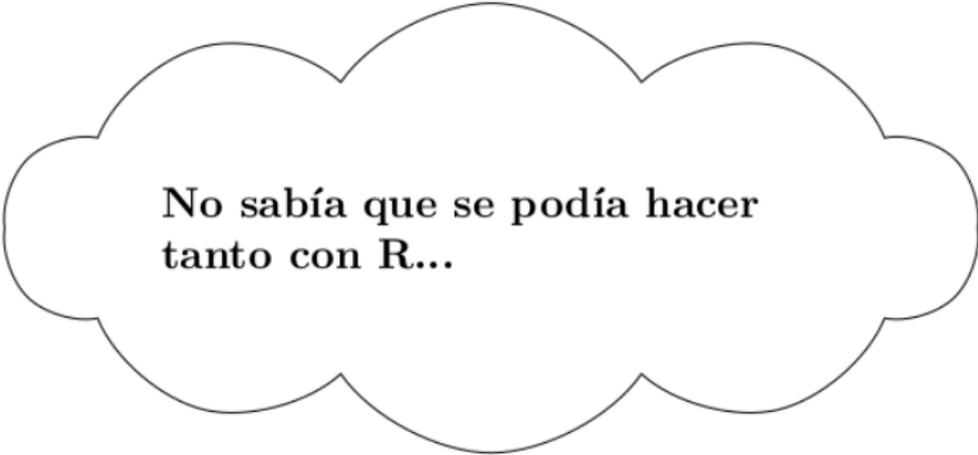


- R es un lenguaje de programación desarrollado **específicamente** para realizar **análisis estadísticos de datos**.
- Se basa en un sistema de paquetes, que son colecciones de funciones desarrolladas por y para la comunidad de usuarios.
- Se almacenan en su mayoría en el repositorio CRAN (The Comprehensive R Archive Network). Unos 20.000 paquetes¹.
- Lo utilizan los mejores estadísticos del mundo (Rob Tibshirani, Julia Silge, Rob Hyndman, etc).
- Es gratis.

¹<https://cran.r-project.org/web/packages/>

Ejemplo de paquetes de R

- **tidyverse**: Procesamiento y minería de datos.
- **tidymodels**: Modelos y sistemas predictivos con *machine learning*.
- **shiny, flexdashboard**: Cuadros de mando para la visualización de datos.
- **roxygen2**: Desarrollo de paquetes personales.
- **RMarkdown**: Escritura de documentos (interactivos y con \LaTeX).



**No sabía que se podía hacer
tanto con R...**

Recolección de datos



Tipos de datos en la ACB

- **Estadísticas tradicionales:** Resumen de la actividad de los jugadores en distintas facetas del juego (puntos, rebotes, etc) **una vez acabado el partido.**
- **Datos jugada a jugada:** Transcripción de cada evento (jugador, acción y momento temporal (minuto y segundo)) que sucede **durante el partido.**

1 **Estadísticas tradicionales.**

2 Datos jugada a jugada.

Estadísticas tradicionales

- Se muestran como tablas estáticas (*box score data*, en inglés).
- No se ofrece la posibilidad de descarga (por ejemplo, como un archivo Excel).
- Inviabile copiar y pegar.

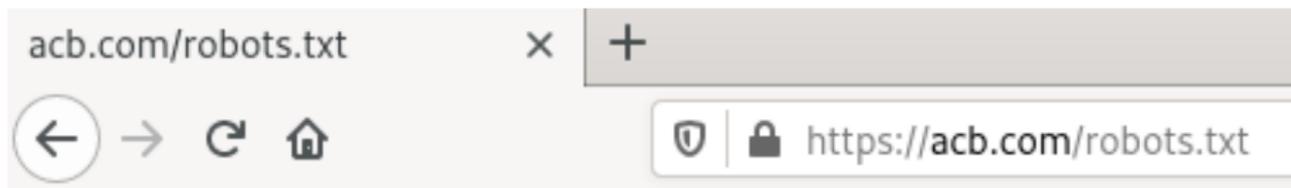
<http://www.acb.com/partido/estadisticas/id/103170>

Valencia Basket 75																						
D	Nombre	Min	P	REB						TAP						FP						
				T2	T2 %	T3	T3 %	T1	T1 %	T	D+0	A	BR	BP	C	F	C	M	F	C	+/-	V
1	V. Claver	21:55	14	0/2	0%	4/4	100%	2/2	100%	7	5+2	0	0	0	0	0	0	0	1	2	-1	20
*2	J. Puerto	15:19	7	2/4	50%	1/4	25%	0/0	0%	1	1+0	1	1	1	2	0	1	0	1	0	1	2
4	J. Pradilla	07:41	4	2/3	67%	0/1	0%	0/1	0%	0	0+0	0	0	1	0	0	0	0	1	1	-2	0
*5	J. Webb	14:50	0	0/1	0%	0/3	0%	0/0	0%	3	2+1	1	1	3	0	0	1	0	5	0	-4	-8
6	López-Arostegui	19:41	5	0/2	0%	1/5	20%	2/2	100%	4	4+0	1	1	1	0	0	1	0	2	2	-10	3
7	C. Jones	18:44	14	3/5	60%	2/5	40%	2/4	50%	2	2+0	2	0	2	0	0	0	0	1	2	-9	10
*12	J. Radebaugh	22:18	4	1/1	100%	0/3	0%	2/2	100%	4	2+2	0	1	0	0	0	0	0	2	2	-5	6
13	S. Evans	14:34	4	2/2	100%	0/1	0%	0/0	0%	1	1+0	2	1	0	0	0	0	0	1	0	-10	6
*14	B. Dubljevic	26:31	15	3/3	100%	2/6	33%	3/4	75%	3	1+2	2	3	1	0	0	0	0	3	5	-7	19
*24	M. Hermannsson	19:03	2	1/4	25%	0/2	0%	0/0	0%	2	2+0	6	0	1	0	0	0	0	2	2	-1	4
41	J. Rivero	19:24	6	2/4	50%	0/1	0%	2/5	40%	4	2+2	1	1	2	0	1	0	0	1	4	-7	8
55	L. Marí																					
	Equipo									2	0+2		3	1					0	0		4
	Total	200:00	75	16/31	52%	10/35	29%	13/20	65%	33	22+11	16	12	13	2	1	3	0	20	20	-55	74
E	Álex Mumbrú																					
5f	J. Webb																					

- **Solución:** Recurrir al así llamado *web scraping amigable*.
- El *web scraping* es una técnica de análisis de datos que permite extraer información de páginas web mediante programas de software.
- Pero... ¿a qué se refiere *amigable*?



- Antes de obtener datos desde la web, se debe averiguar si existe permiso para ello.
- En el archivo **robots.txt**, el propietario de la web especifica quién puede extraer datos y quién no.



```
User-agent: Twitterbot  
Disallow:
```

- Una vez sabemos que tenemos permiso para obtener los datos, debemos averiguar si los datos están disponibles en formato *html*.
- Procedimiento: Botón derecho del ratón → Ver código fuente de la página:

```

<td class="dorsal">1</td>
<td class="nombre jugador ellipsis"><a href="/jugador/ver/20201901-V.-Claver">V. Claver</a></td>
<td class="">21:55</td>
<td class="grisclaro">14</td>
<td class="">>0/2</td>
<td class="">>0%</td>
<td class="grisclaro">4/4</td>
<td class="grisclaro">>100%</td>
<td class="">>2/2</td>
<td class="">>100%</td>
<td class="grisclaro">7</td>
<td class="grisclaro">>5+2</td>
<td class="">>0</td>
<td class="grisclaro">>0</td>
<td class="">>0</td>
<td class="grisclaro">>0</td>
<td class="">>0</td>
<td class="grisclaro">>0</td>
<td class="">>1</td>
<td class="grisclaro">>2</td>
<td class="">>-1</td>
<td class="grisclaro">>20</td>
</tr>
<tr>
<td class="dorsal">*2</td>
<td class="nombre jugador ellipsis"><a href="/jugador/ver/20210392-J.-Puerto">J. Puerto</a></td>
<td class="">>15:19</td>
<td class="grisclaro">7</td>
<td class="">>2/4</td>
<td class="">>50%</td>
<td class="grisclaro">1/4</td>
<td class="grisclaro">>25%</td>
<td class="">>0/0</td>
<td class="">>0%</td>

```

- Ahora que sabemos que el formato *html* está disponible, el paquete de R **rvest**² nos ofrece las funciones necesarias.

```
library(rvest)
url_html <- read_html("http://www.acb.com/partido/estadisticas/id/103170")
url_html %>%
  html_table(fill = TRUE) %>%
  `[[`[3]
```

²<https://CRAN.R-project.org/package=rvest>

data_acb x

Filter

										REB	REB					TAP	TAP		FP	FP	+/-	V	
1	D	Nombre	Min	P	T2	T2 %	T3	T3 %	T1	T1 %	T	D+0	A	BR	BP	C	F	C	M	F	C	+/-	V
2	1	V. Claver	21:55	14	0/2	0%	4/4	100%	2/2	100%	7	5+2	0	0	0	0	0	0	0	1	2	-1	20
3	*2	J. Puerto	15:19	7	2/4	50%	1/4	25%	0/0	0%	1	1+0	1	1	1	2	0	1	0	1	0	1	2
4	4	J. Pradilla	07:41	4	2/3	67%	0/1	0%	0/1	0%	0	0+0	0	0	1	0	0	0	0	1	1	-2	0
5	*5	J. Webb	14:50	0	0/1	0%	0/3	0%	0/0	0%	3	2+1	1	1	3	0	0	1	0	5	0	-4	-8
6	6	López-Arostegui	19:41	5	0/2	0%	1/5	20%	2/2	100%	4	4+0	1	1	1	0	0	1	0	2	2	-10	3
7	7	C. Jones	18:44	14	3/5	60%	2/5	40%	2/4	50%	2	2+0	2	0	2	0	0	0	1	2	-9	10	
8	*12	J. Radebaugh	22:18	4	1/1	100%	0/3	0%	2/2	100%	4	2+2	0	1	0	0	0	0	0	2	2	-5	6
9	13	S. Evans	14:34	4	2/2	100%	0/1	0%	0/0	0%	1	1+0	2	1	0	0	0	0	1	0	-10	6	
10	*14	B. Dubljevic	26:31	15	3/3	100%	2/6	33%	3/4	75%	3	1+2	2	3	1	0	0	0	0	3	5	-7	19
11	*24	M. Hermansson	19:03	2	1/4	25%	0/2	0%	0/0	0%	2	2+0	6	0	1	0	0	0	0	2	2	-1	4
12	41	J. Rivero	19:24	6	2/4	50%	0/1	0%	2/5	40%	4	2+2	1	1	2	0	1	0	0	1	4	-7	8
13	55	L. Marí																					
14		Equipo									2	0+2		3	1					0	0		4
15		Total	200:00	75	16/31	52%	10/35	29%	13/20	65%	33	22+11	16	12	13	2	1	3	0	20	20	-55	74

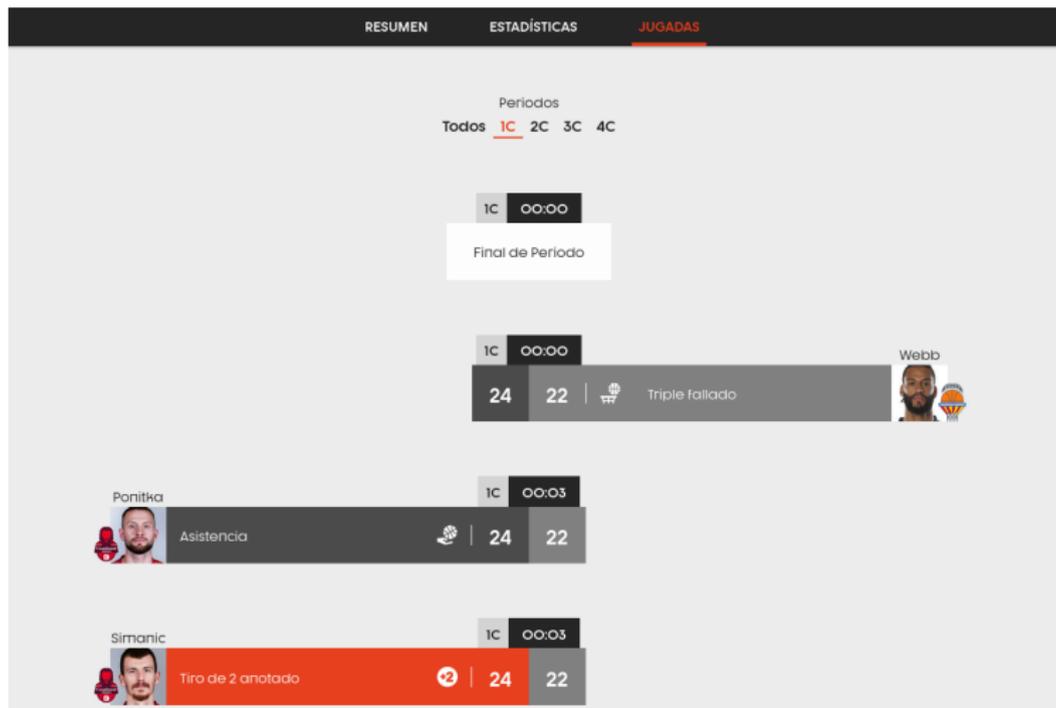
1 Estadísticas tradicionales.

2 **Datos jugada a jugada.**

Datos jugada a jugada

- Se muestran como una sucesión de eventos.
- No se ofrece la posibilidad de descarga (por ejemplo, como un archivo Excel).
- Inviabile copiar y pegar.

<http://jv.acb.com/es/103170/jugadas>



- En este caso, el formato *html* no está disponible. En lugar de usar **rvest**, el paquete de R apropiado es **RSelenium**³.
- El procedimiento a seguir es bastante más laborioso, pero funciona.

```
library(RSelenium)
# Abre un servidor:
rD <- rsDriver(browser = "firefox", chromever = NULL)
# Copia el enlace del partido http://jv.acb.com/es/103170/jugadas
# Trabaja cuarto a cuarto. Empieza desde 1C, ve con el ratón hasta al final
# de la página y entonces, ejecuta este código:
remDr <- rD$client # Driver remoto.

# Datos jugada a jugada:
game_elem <- remDr$getPageSource()[[1]]

# Cierra el cliente y el servidor:
remDr$close()
rD$server$stop()
```

³<https://CRAN.R-project.org/package=RSelenium>

data_acb_pbp x

Filter

	period	time_point	player	action	local_score	visitor_score	team
1	1C	00:00	Webb	Triple fallado	24	22	Valencia Basket
2	1C	00:03	Simanic	Tiro de 2 anotado	24	22	Casademont Zaragoza
3	1C	00:03	Ponitka	Asistencia	24	22	Casademont Zaragoza

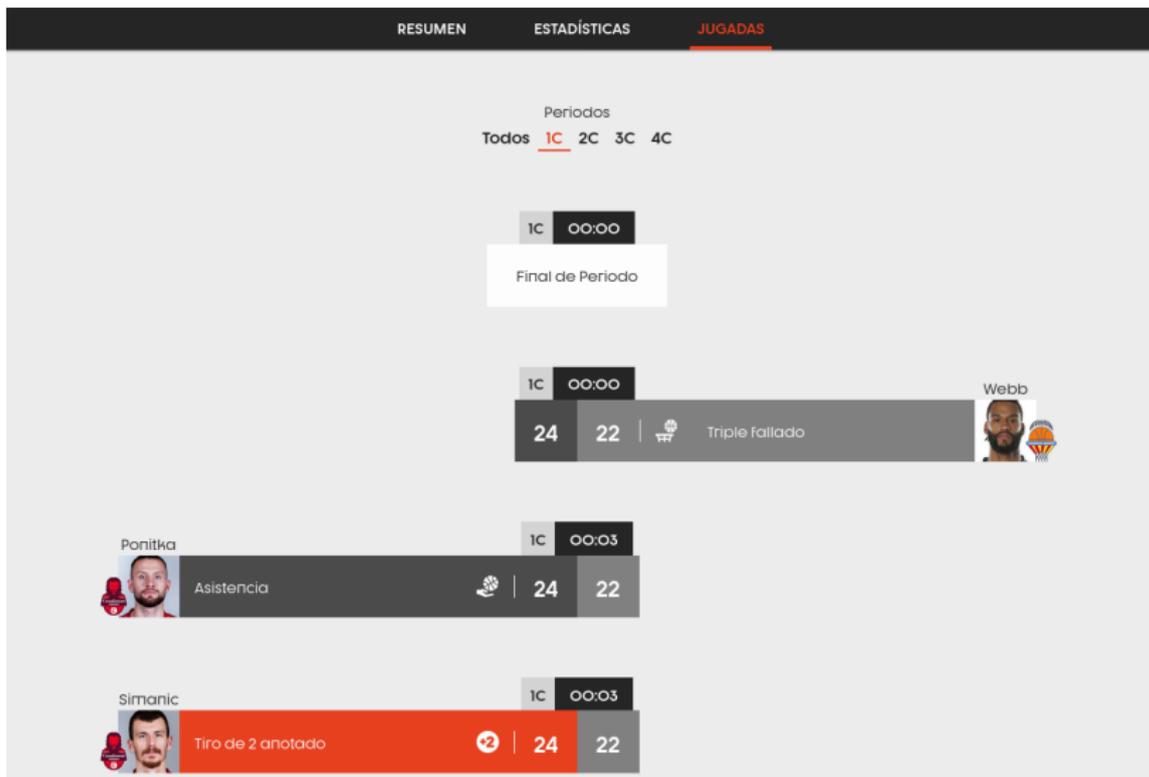
En resumen...



ANTES

Valencia Basket 75																						
D	Nombre	Min	P	REB				TAP				FP				V						
				T2	T2 %	T3	T3 %	T1	T1 %	T	D+0	A	BR	BP	C		F	C	M	F	C	+/-
1	V. Claver	21:55	14	0/2	0%	4/4	100%	2/2	100%	7	5+2	0	0	0	0	0	0	0	1	2	-1	20
*2	J. Puerto	15:19	7	2/4	50%	1/4	25%	0/0	0%	1	1+0	1	1	1	2	0	1	0	1	0	1	2
4	J. Pradilla	07:41	4	2/3	67%	0/1	0%	0/1	0%	0	0+0	0	0	1	0	0	0	0	1	1	-2	0
*5	J. Webb	14:50	0	0/1	0%	0/3	0%	0/0	0%	3	2+1	1	1	3	0	0	1	0	5	0	-4	-8
6	López-Arostegui	19:41	5	0/2	0%	1/5	20%	2/2	100%	4	4+0	1	1	1	0	0	1	0	2	2	-10	3
7	C. Jones	18:44	14	3/5	60%	2/5	40%	2/4	50%	2	2+0	2	0	2	0	0	0	0	1	2	-9	10
*12	J. Radebaugh	22:18	4	1/1	100%	0/3	0%	2/2	100%	4	2+2	0	1	0	0	0	0	0	2	2	-5	6
13	S. Evans	14:34	4	2/2	100%	0/1	0%	0/0	0%	1	1+0	2	1	0	0	0	0	0	1	0	-10	6
*14	B. Dubljevic	26:31	15	3/3	100%	2/6	33%	3/4	75%	3	1+2	2	3	1	0	0	0	0	3	5	-7	19
*24	M. Hermansson	19:03	2	1/4	25%	0/2	0%	0/0	0%	2	2+0	6	0	1	0	0	0	0	2	2	-1	4
41	J. Rivero	19:24	6	2/4	50%	0/1	0%	2/5	40%	4	2+2	1	1	2	0	1	0	0	1	4	-7	8
55	L. Mari																					
	Equipo									2	0+2		3	1					0	0		4
	Total	200:00	75	16/31	52%	10/35	29%	13/20	65%	33	22+11	16	12	13	2	1	3	0	20	20	-55	74
E	Álex Mumbrú																					
5f	J. Webb																					

ANTES



ANTES



¡R al rescate!



DESPUÉS

data_acb x																								
Filter																								
											REB	REB					TAP	TAP		FP	FP			
1	D	Nombre	Min	P	T2	T2 %	T3	T3 %	T1	T1 %	T	D+0	A	BR	BP	C	F	C	M	F	C	+/-	V	
2	1	V. Claver	21:55	14	0/2	0%	4/4	100%	2/2	100%	7	5+2	0	0	0	0	0	0	0	1	2	-1	20	
3	*2	J. Puerto	15:19	7	2/4	50%	1/4	25%	0/0	0%	1	1+0	1	1	1	2	0	1	0	1	0	1	2	
4	4	J. Pradilla	07:41	4	2/3	67%	0/1	0%	0/1	0%	0	0+0	0	0	1	0	0	0	0	1	1	-2	0	
5	*5	J. Webb	14:50	0	0/1	0%	0/3	0%	0/0	0%	3	2+1	1	1	3	0	0	1	0	5	0	-4	-8	
6	6	López-Arostegui	19:41	5	0/2	0%	1/5	20%	2/2	100%	4	4+0	1	1	1	0	0	1	0	2	2	-10	3	
7	7	C. Jones	18:44	14	3/5	60%	2/5	40%	2/4	50%	2	2+0	2	0	2	0	0	0	0	1	2	-9	10	
8	*12	J. Radebaugh	22:18	4	1/1	100%	0/3	0%	2/2	100%	4	2+2	0	1	0	0	0	0	0	2	2	-5	6	
9	13	S. Evans	14:34	4	2/2	100%	0/1	0%	0/0	0%	1	1+0	2	1	0	0	0	0	0	1	0	-10	6	
10	*14	B. Dubljevic	26:31	15	3/3	100%	2/6	33%	3/4	75%	3	1+2	2	3	1	0	0	0	0	3	5	-7	19	
11	*24	M. Hermansson	19:03	2	1/4	25%	0/2	0%	0/0	0%	2	2+0	6	0	1	0	0	0	0	2	2	-1	4	
12	41	J. Rivero	19:24	6	2/4	50%	0/1	0%	2/5	40%	4	2+2	1	1	2	0	1	0	0	1	4	-7	8	
13	55	L. Marí																						
14		Equipo									2	0+2		3	1					0	0		4	
15		Total	200:00	75	16/31	52%	10/35	29%	13/20	65%	33	22+11	16	12	13	2	1	3	0	20	20	-55	74	

DESPUÉS

	period	time_point	player	action	local_score	visitor_score	team
1	1C	00:00	Webb	Triple fallado	24	22	Valencia Basket
2	1C	00:03	Simanic	Tiro de 2 anotado	24	22	Casademont Zaragoza
3	1C	00:03	Ponitka	Asistencia	24	22	Casademont Zaragoza

DESPUÉS



Estadística descriptiva

Procesamiento de datos, visualización y análisis



1 **Estadísticas tradicionales.**

2 Datos jugada a jugada.

Nueva representación de las estadísticas tradicionales

<https://www.uv.es/vivigui/AppEuroACB.html>

- Desarrollo de la primera aplicación web para la visualización interactiva de las estadísticas del baloncesto europeo procedentes de los **datos de los partidos (box score data)**. Creada en 2020.

The screenshot displays the 'Baloncesto' application interface. It features a dark sidebar on the left with the following elements:

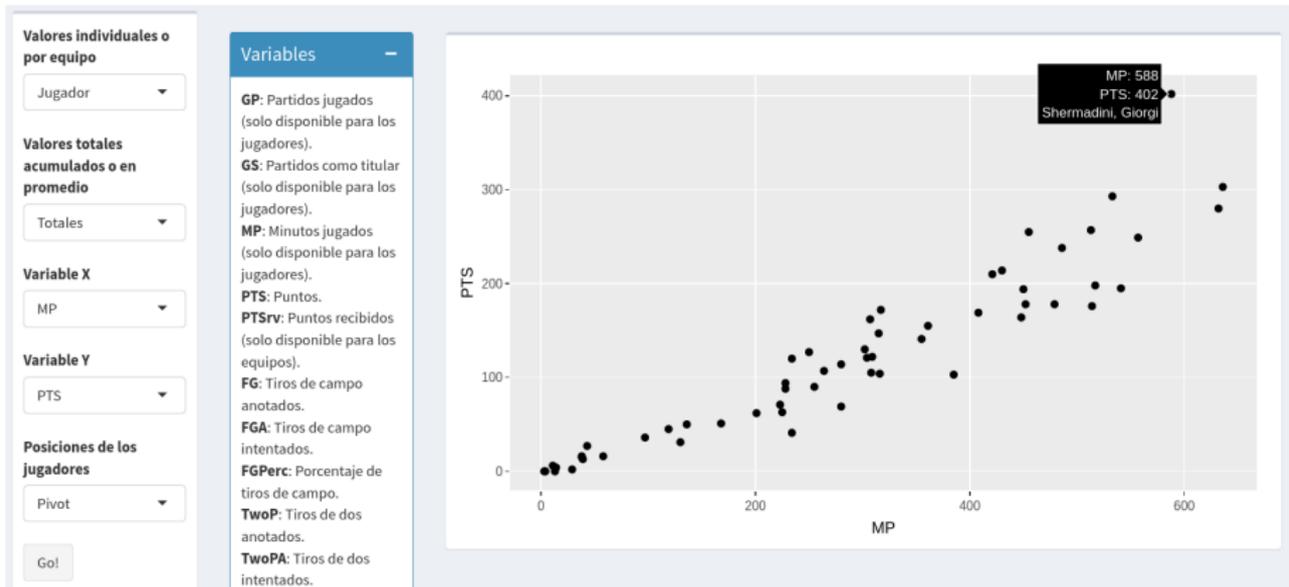
- Baloncesto** (Title)
- Language selection: Español, English
- Liga** (League): Dropdown menu showing 'ACB'
- Temporada** (Season): Dropdown menu showing '2019-2020'
- Fase** (Phase): Dropdown menu showing 'Temporada Regular'

The main content area on the right is light blue and contains:

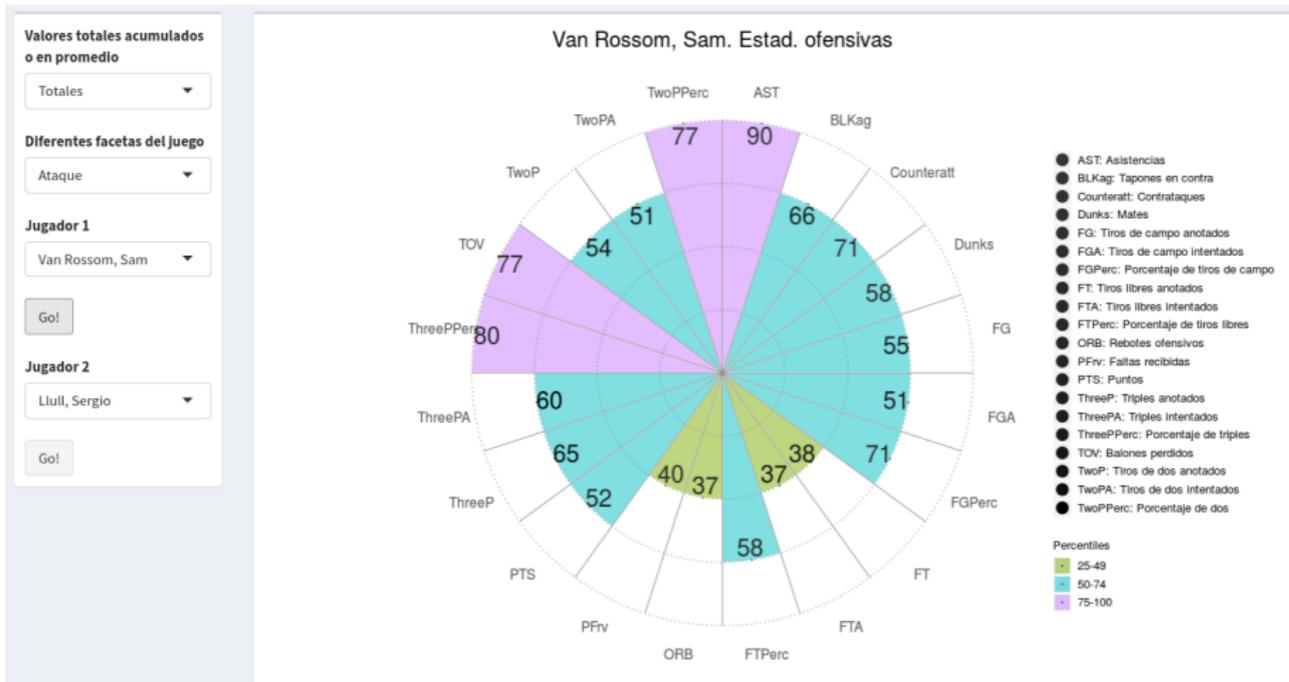
- A hamburger menu icon (☰)
- User information: Guillermo Vinue, guillermovinue@gmail.com, www.uv.es/vivigui
- A link: [FUNCIONAMIENTO DE LA WEB.](#)

- Los datos de los jugadores y equipos se pueden estudiar de manera muy intuitiva y cómoda.
- Incluye las estadísticas tradicionales de las tres competiciones más importantes de Europa:
 - **Liga ACB** (desde la temporada 1985-1986 a la 2019-2020 incluyendo la Copa del Rey y la Supercopa).
 - **Euroliga** (2000-2001 a 2019-2020).
 - **EuroCup** (2002-2003 a 2019-2020).
- Un total de 40 estadísticas disponibles y diferentes visualizaciones.

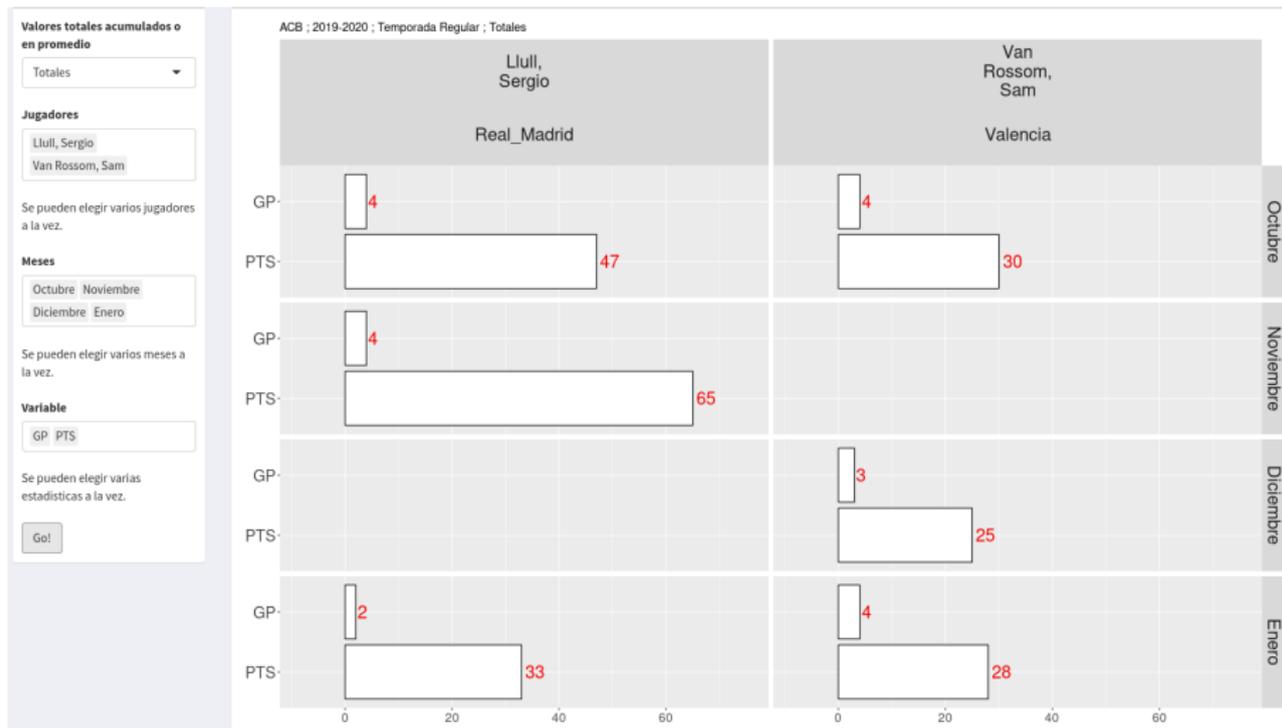
Correlación entre estadísticas. Comparación entre jugadores y equipos.



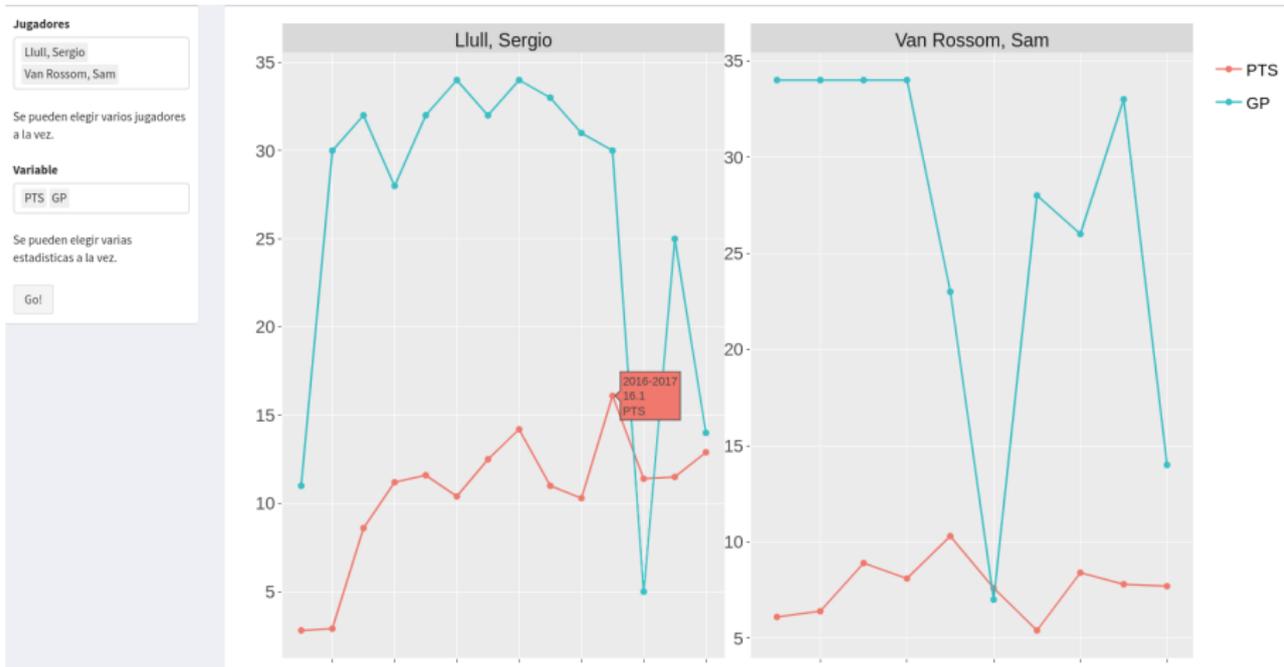
- Percentiles de las estadísticas (dónde se sitúa el jugador en esa estadística respecto a todos los demás). Comparación entre jugadores.



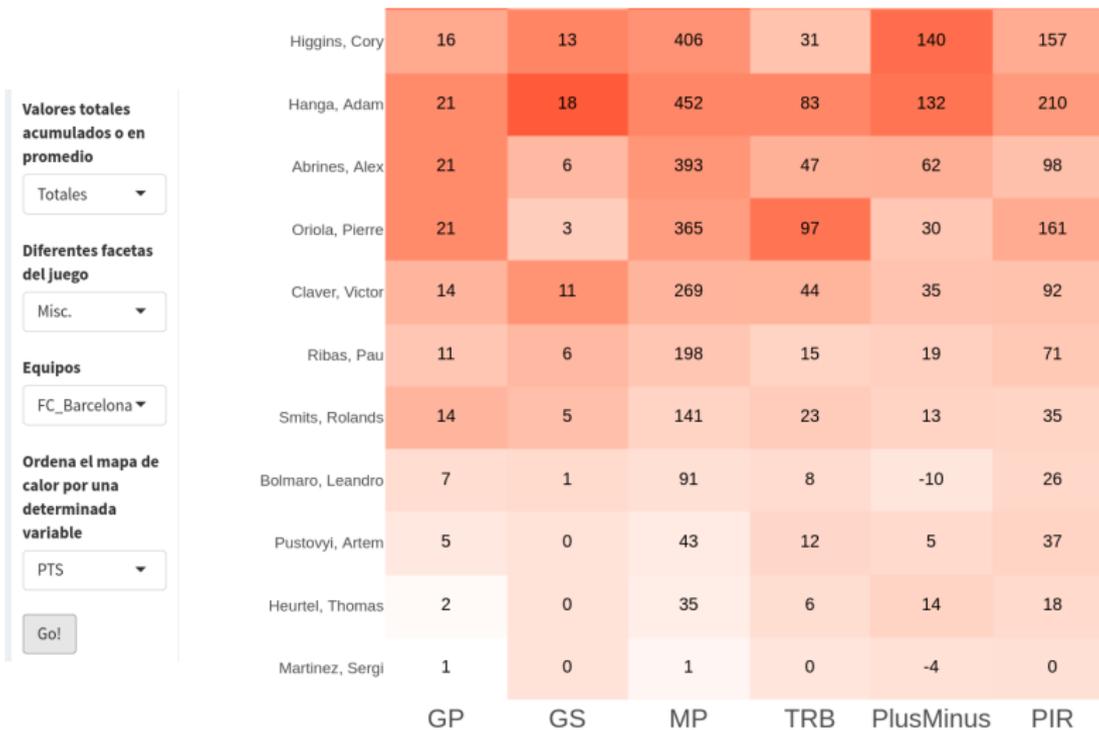
- Seguimiento de la evolución mensual de un jugador a lo largo de una temporada. Comparación entre jugadores.



- Seguimiento de la evolución anual de un jugador a lo largo de las temporadas. Comparación entre jugadores.



- Mapa de calor para comparar las estadísticas de los jugadores de un mismo equipo.



● Porcentajes de tiro de los jugadores de un mismo equipo.

Valores totales
acumulados o en
promedio

Totales

Equipos

Valencia

Tipo de tiro: tiros libres
(1), tiros de dos (2) o
tiros de tres (3)



Go!

Valencia ACB 2019-2020
Temporada Regular Totales

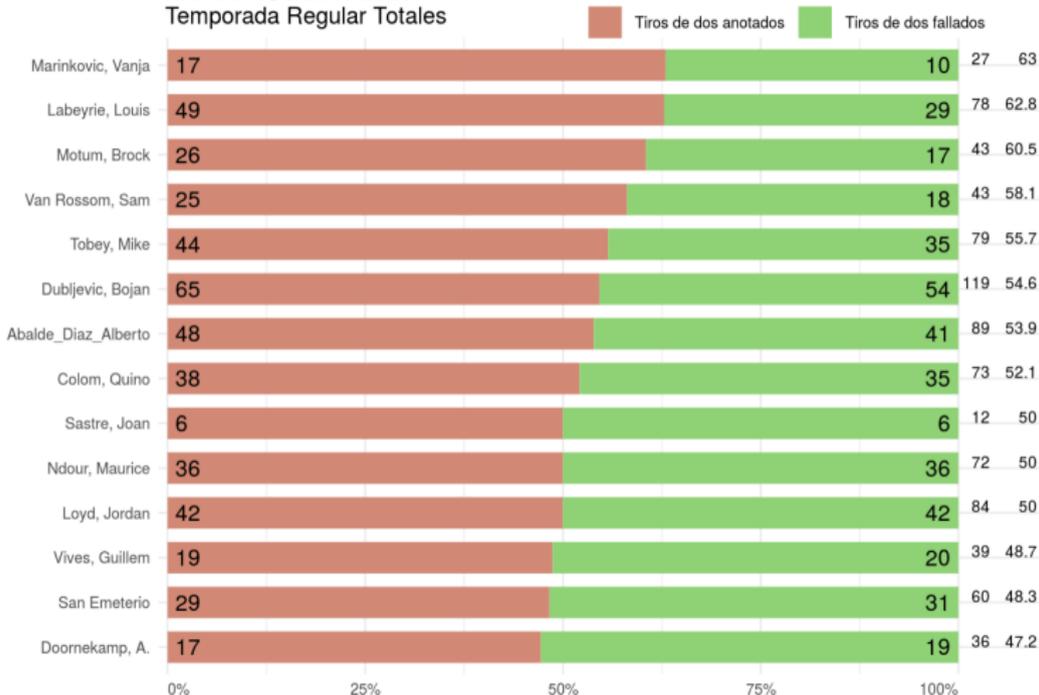
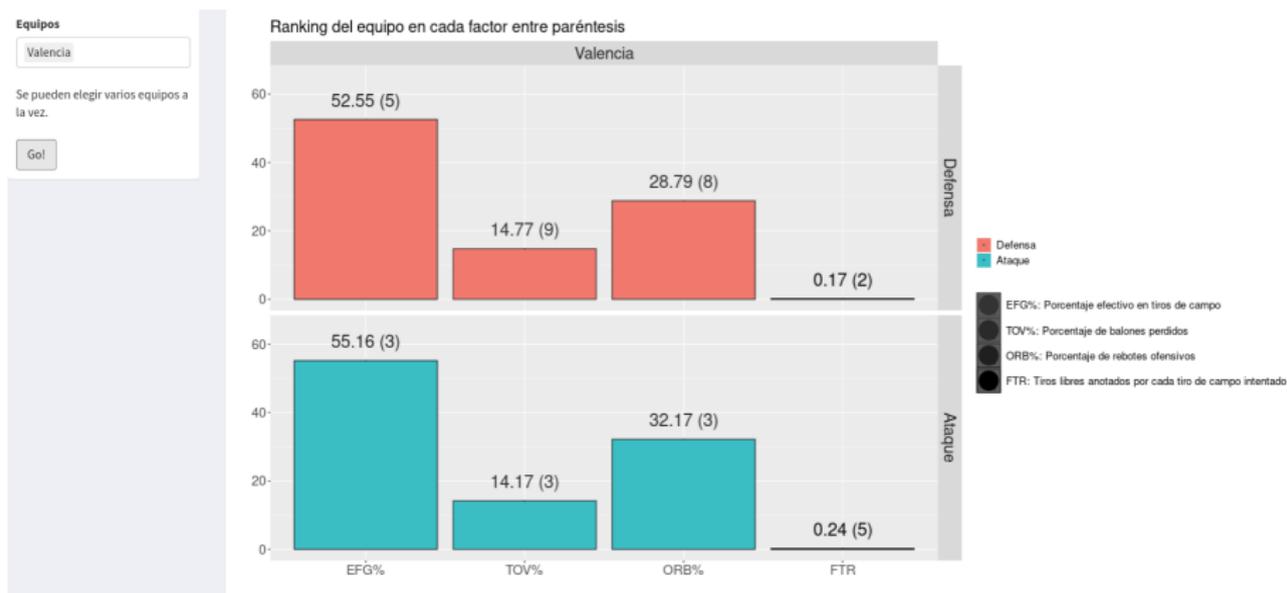


Gráfico de los cuatro factores. Comparación entre equipos.



Publicación

Big Data
Volume 8, Number 1, 2020
© Mary Ann Liebert, Inc.
DOI: 10.1089/big.2018.0124

A Web Application for Interactive Visualization of European Basketball Data

Guillermo Vinué*

Abstract

The statistical analysis of basketball games is a fast-growing field. Certainly, basketball data are scientifically relevant because an appropriate analysis provides a great deal of information about the performance of both players and teams. The number of games played each season generates a large amount of data worth analyzing. Basketball analytics is well established in U.S. leagues. In Europe, however, it has not been duly developed. This study focuses on the top three European team competitions: the EuroLeague, the EuroCup, and the Spanish ACB (Association of Basketball Clubs, acronym in Spanish) league. Their official websites provide access to game data for anyone who is interested, but they are only represented in a static tabular form. As a consequence, it is difficult to gain any valuable insights from them. This article presents a highly useful interactive tool, created with the free statistical software R, which makes it possible to visualize and explore basketball data from a large number of seasons. We will demonstrate its core functionality. An accompanying R package is presented in the Supplementary Data.

Keywords: European basketball; Big data mining; web application; interactive visualization; R

● Factor de Impacto: 2.128 ; Posición: 37/110 = 0.34 (Computer Science, Theory & Methods) ; Base: JCR (2020) ; Citas: 5 ; <https://doi.org/10.1089/big.2018.0124>

Publicación

<https://CRAN.R-project.org/package=BAwiR>

BAwiR: Analysis of Basketball Data

Collection of tools to work with basketball data. Functions available are related to friendly web scraping and visualization. Data were obtained from [<https://www.euroleague.net/>](https://www.euroleague.net/), [<https://www.eurocupbasketball.com/>](https://www.eurocupbasketball.com/) and [<https://www.acb.com/>](https://www.acb.com/), following the instructions of their respective robots.txt files, when available. Tools for visualization include a population pyramid, 2D plots, circular plots of players' percentiles, plots of players' monthly/yearly stats, team heatmaps, team shooting plots, team four factors plots, cross-tables with the results of regular season games and maps of nationalities. Please see Vinue (2020) [doi:10.1089/big.2018.0124](https://doi.org/10.1089/big.2018.0124).

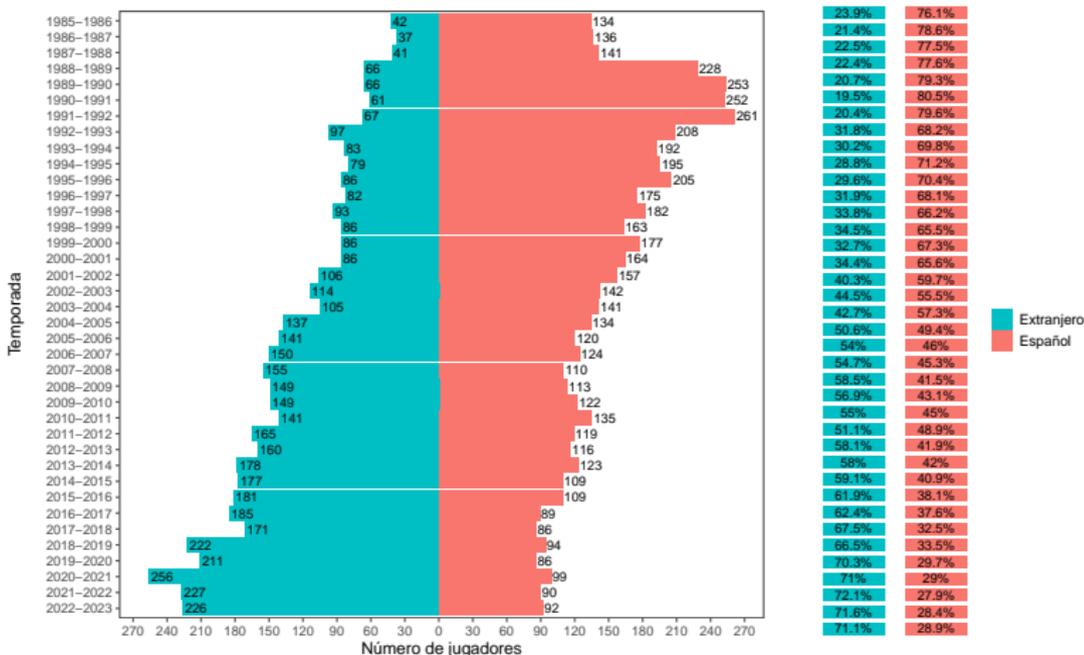
Version: 1.2.7
Depends: R (≥ 3.4.0)
Imports: [Anthropometry](#), [plyr](#), [dplyr](#), [ggplot2](#), [ggthemes](#), [grid](#), [httr](#), [lubridate](#), [magrittr](#), [purrr](#), [reshape2](#), [rvest](#), [rworldmap](#), [scales](#), [stringi](#), [stringr](#), [tibble](#), [tidyr](#), [xml2](#)
Suggests: [knitr](#), [markdown](#), [rmarkdown](#)
Published: 2021-07-19
Author: Guillermo Vinue
Maintainer: Guillermo Vinue <Guillermo.Vinue at uv.es>
License: [GPL-2](#) | [GPL-3](#) [expanded from: GPL (≥ 2)]
URL: <https://www.R-project.org>, <https://www.uv.es/vivigui/>, <https://www.uv.es/vivigui/AppEuroACB.html>
NeedsCompilation: no
Materials: [NEWS](#)
In views: [SportsAnalytics](#)
CRAN checks: [BAwiR results](#)

Documentation:

Reference manual: [BAwiR.pdf](#)
Vignettes: [Visualization of European basketball data](#)

Otros estudios: Análisis demográfico

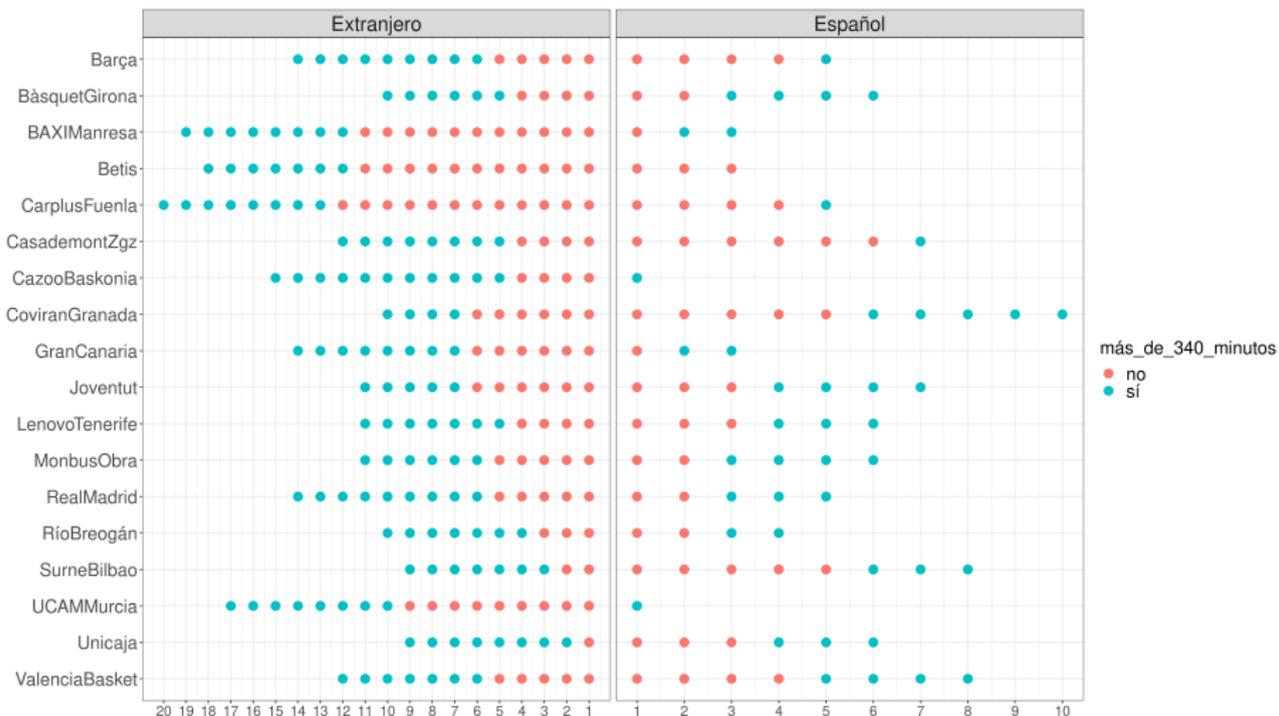
- Cada vez hay menos jugadores españoles en la ACB (menos del 30%).





- En los 80 y 90 había más españoles que extranjeros.
- Cambio de tendencia en la temporada 2004-2005.
- Desde entonces, la brecha no hace sino aumentar.

- Además, una gran parte de esos españoles es poco protagonista.



GIGANTES



- Gigantes del Basket, número 1522, 2022.

- 1 Estadísticas tradicionales.
- 2 Datos jugada a jugada.**

Nuevos análisis a partir de los datos jugada a jugada

- 1 Quintetos y otras combinaciones.
- 2 Desglose por cuartos y *clutch time*.
- 3 Faltas personales pormenorizadas.
- 4 Rebotes ofensivos que acaban en canasta.
- 5 Posesiones.
- 6 Eficacia ofensiva de los tiempos muertos.

Aclaración

- Los siguientes resultados se han calculado para los **cuatro cuartos** de cada partido de la **temporada regular**.
- Dicho de otro modo, **excepto para el cálculo del clutch time, no se han tenido en cuenta las prórrogas**. Así, cada partido tiene la misma duración y los resultados son exactamente comparables.
- De los 306 partidos de la liga regular, hubo 17 con prórroga: 15 con una y 2 con dos (5.6%).

- 1 **Quintetos y otras combinaciones.**
- 2 Desglose por cuartos y *clutch time*.
- 3 Faltas personales pormenorizadas.
- 4 Rebotes ofensivos que acaban en canasta.
- 5 Posesiones.
- 6 Eficacia ofensiva de los tiempos muertos.

Quintetos y otras combinaciones

- El baloncesto es un juego colectivo en el que se juega 5 contra 5.
- La estadística del **más / menos** indica la diferencia de puntos en el marcador entre un equipo y otro.

cuarto	minutos	equipo	quinteto en pista	marcador	más_menos
1	10:00 – 5:13	Valencia Basket	Dubljevic, Ferrando, Puerto, Radebaugh, Webb	7-8	-1
1	10:00 – 5:13	Casademont Zaragoza	Cruz, Hlinason, Jovic, Sant-Roos, Yusta	8-7	1

- Actualmente, el **más / menos** se ofrece únicamente para cada jugador de manera individual.

MÁS/MENOS			Ver lista ▶
1	NIKOLA MIROTIC	BARÇA	9,9
2	Aaron Doornekamp	Lenovo Tenerife	8,9
3	Sertac Sanli	Barça	8,6
4	Dzanan Musa	Real Madrid	8,4
5	Tomas Satoransky	Barça	8,3

- Sin embargo, es mucho más informativo conocer cómo interaccionan los jugadores entre ellos.
- A partir de los **quintetos**, se pueden extraer los **cuartetos**, los **tríos** y las **parejas** que forman parte de ellos.
- Teoría combinatoria: $C_{5,x} = \binom{5}{x} = \frac{5!}{x!(5-x)!}$, para $x = 4, 3, 2$
- **Información adicional para un entrenador**: saber qué combinaciones de jugadores funcionan mejor o peor a lo largo de la temporada.

Resultados

- Los tres mejores y peores QUINTETOS de la liga en términos del más / menos son los siguientes:

Mejores quintetos

equipo	quinteto	apariciones	tiempo	+ / -
Valencia	Dubljevic, Jones, Puerto, Radebaugh, Webb	15	58M 23S	52
Gran Canaria	Albicy, Balcerowski, Brussino, Shurna, Slaughter	52	2H 29M 44S	51
Murcia	Bellas, Diop, McFadden, Radovic, Rojas	17	54M 33S	45

- Nota:** De la 15 veces que han coincidido Dubljevic, Jones, Puerto, Radebaugh y Webb, solo una ha sido en el último cuarto.

Peores quintetos

equipo	quinteto	apariciones	tiempo	+ / -
Bilbao	Alonso, Kyser, Radicevic, Reyes, Sulejmanovic	11	26M 27S	-47
Granada	Bropleh, Caicedo, Costa, Moore, Ndoye	8	33M 2S	-45
Manresa	Harding, Pérez, Sagnia, Steinbergs, Tyson	6	9M 10S	-28

- Nota:** Si el número de apariciones o el tiempo jugado parece escaso para llegar a una conclusión relevante, se recomienda filtrar por un umbral que se considere adecuado.

- Los tres mejores y peores CUARTETOS de la liga en términos del **más / menos** son los siguientes:

Mejores cuartetos				
equipo	quinteto	apariciones	tiempo	+ / -
Gran Canaria	Albicy, Brussino, Shurna, Slaughter	111	4H 38M 52S	88
Barcelona	Laprovittola, Mirotic, Sanli, Satoransky	37	1H 20M 24S	79
Valencia	Dubljevic, Jones, Puerto, Radebaugh	49	2H 46M 53S	79

Peores cuartetos				
equipo	quinteto	apariciones	tiempo	+ / -
Bilbao	Alonso, Kyser, Reyes, Sulejmanovic	34	1H 30M 46S	-69
Granada	Bropoleh, Costa, Moore, Ndoye	16	49M 15S	-57
Bilbao	Hakanson, Kyser, Smith, Sulejmanovic	51	1H 21M 49S	-52

- Los tres mejores y peores TRÍOS de la liga en términos del más / menos son los siguientes:

Mejores tríos

equipo	quinteto	apariciones	tiempo	+ / -
Barcelona	Laprovittola, Mirotic, Satoransky	112	4H 40M 32S	125
Barcelona	Laprovittola, Sanli, Satoransky	82	2H 53M 36S	119
Tenerife	Cook, Doornekamp, Salin	106	3H 4M 7S	114

Peores tríos

equipo	quinteto	apariciones	tiempo	+ / -
Bilbao	Alonso, Kyser, Sulejmanovic	49	1H 48M 15S	-83
Bilbao	Alonso, Kyser, Reyes	90	3H 25M 11S	-80
Bilbao	Kyser, Reyes, Sulejmanovic	88	2H 48M 20S	-80

- Las tres mejores y peores PAREJAS de la liga en términos del más / menos son las siguientes:

Mejores parejas				
equipo	quinteto	apariciones	tiempo	+ / -
Barcelona	Laprovittola, Satoransky	214	8H 8M 1S	186
Barcelona	Laprovittola, Sanli	156	4H 54M 16S	181
Tenerife	Doornekamp, Salin	249	6H 17M 54S	178

Peores parejas				
equipo	quinteto	apariciones	tiempo	+ / -
Bilbao	Kyser, Sulejmanovic	189	5H 21M 45S	-139
Fuenlabrada	Horton, Kromah	103	4H 13M 44S	-135
Bilbao	Kyser, Reyes	233	6H 57M 1S	-129

- Extensión con otras estadísticas. Mismo análisis que en la NBA.

<https://www.nba.com/stats/lineups/advanced>

SEASON	SEASON TYPE	TEAM	LINEUPS
2022-23	Regular Season	All Teams	5 Player Lineups
			5 Player Lineups
			4 Player Lineups
			3 Player Lineups
			2 Player Lineups

LINEUPS	TEAM	GP	MIN	OFFRTG	DEFRTG	NETRTG	AST%	AST/TO	AST RATIO	OREB%	DREB%	REB%	TO RATIO
H. Barnes - D. Sabonis - D. Fox - K. Huerter - K. Murray	SAC	63	900	118.3	116.1	2.2	62.0	2.30	19.4	25.8	76.4	50.6	0.1
C. Capela - D. Murray - J. Collins - T. Young - D. Hunter	ATL	44	738	115.9	110.6	5.3	56.9	2.25	18.4	28.8	74.2	51.1	0.1
K. Caldwell-Pope - A. Gordon - N. Jokic - J. Murray - M. Porter Jr.	DEN	41	706	124.3	111.2	13.1	69.8	2.66	23.2	28.3	75.1	52.5	0.1

- 1 Quintetos y otras combinaciones.
- 2 Desglose por cuartos y clutch time.**
- 3 Faltas personales pormenorizadas.
- 4 Rebotes ofensivos que acaban en canasta.
- 5 Posesiones.
- 6 Eficacia ofensiva de los tiempos muertos.

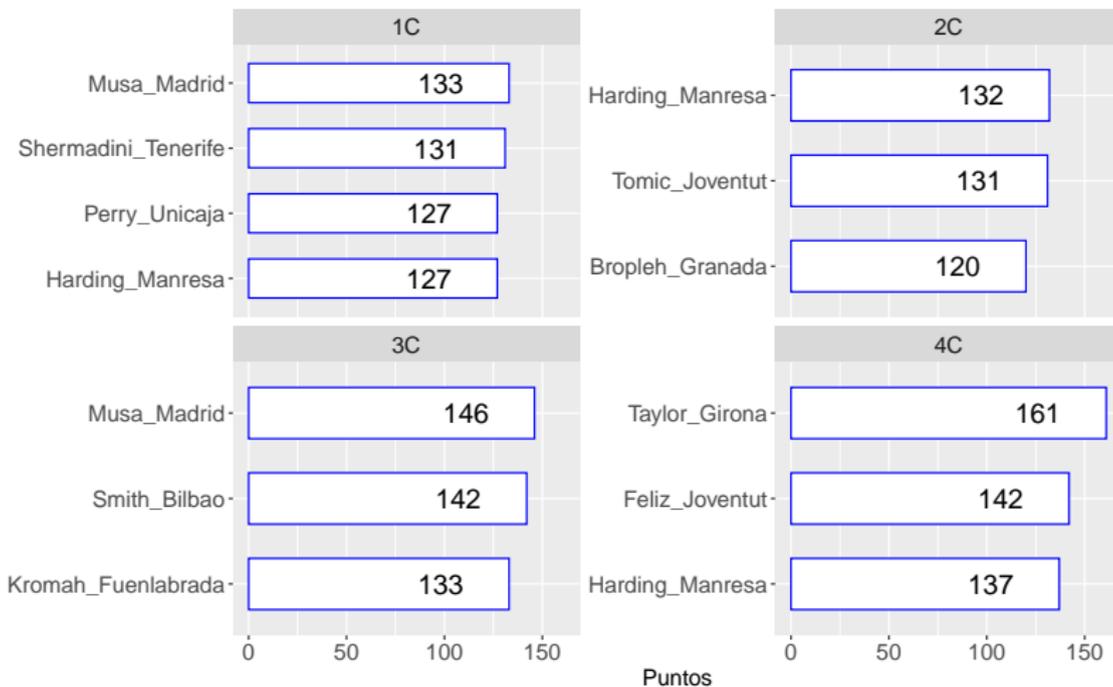
Desglose por cuartos

- Un partido de baloncesto en la ACB se divide en **cuatro cuartos de 10 minutos de duración** cada uno (**40 minutos** en total).
- Si el partido acaba en empate, hay sucesivas prórrogas de 5 minutos hasta que se deshace el empate al final de la correspondiente prórroga.
- Aparte de la actuación global durante todo el partido, es interesante saber **cómo se comporta el jugador en cada cuarto**.

- Por ejemplo, para saber...
- quién sale más enchufado a los partidos (primer o tercer cuarto) o...
- quién juega mejor cuándo el partido se decide (normalmente cerca del final), lo que en la NBA llaman *clutch time* (momento caliente o crítico del partido).

Resultados

- Top 3 de anotadores en cada cuarto:



- **Clutch time:** Según la NBA, momento del partido en el que quedan menos de 5 minutos de juego (del último cuarto o de la prórroga), siendo la diferencia máxima entre los dos equipos de 5 puntos.

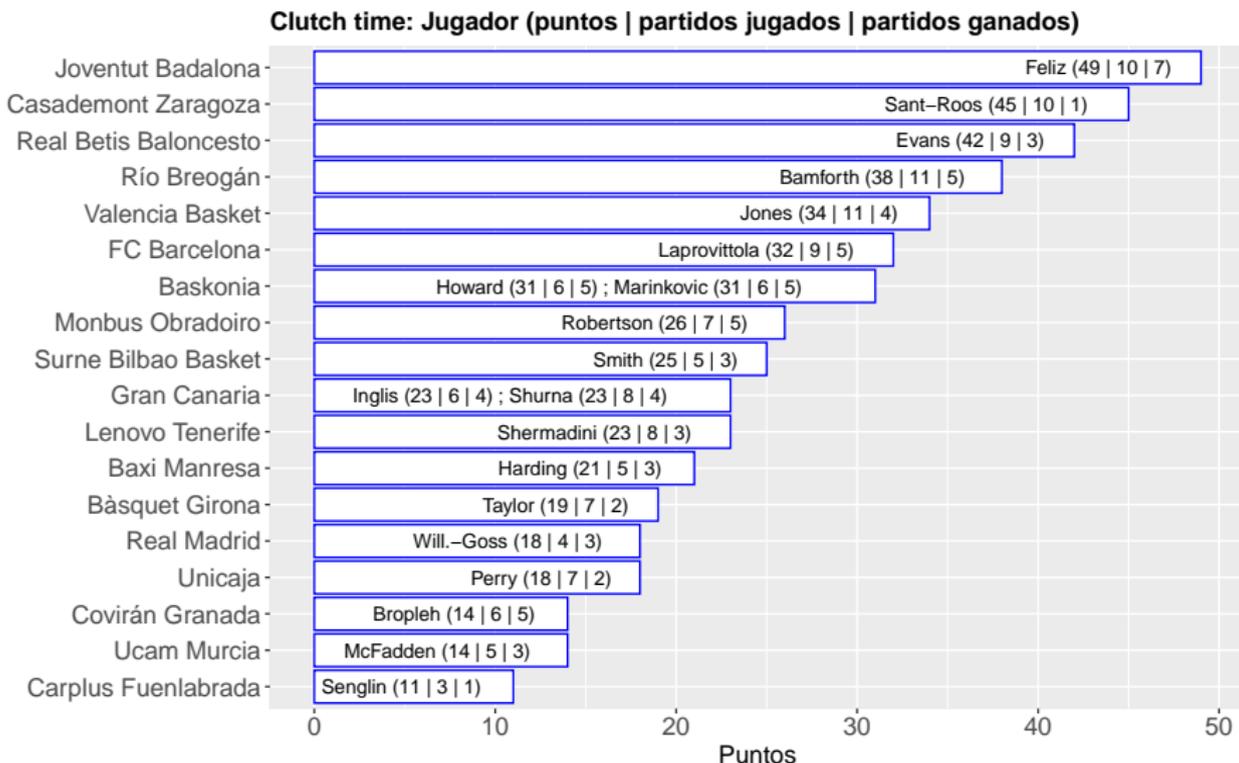


Resultados

- Ha habido 76 partidos de *clutch time* (aproximadamente el 25%, es decir, uno de cada cuatro).
- Top 3 de anotadores en el *clutch time*:

jugador	equipo	puntos	partidos jugados	partidos ganados	porcentaje victorias
Feliz	Joventut	49	10	7	70%
Sant-Roos	Zaragoza	45	10	1	10%
Evans	Betis	42	9	3	33.33%

- Máximo anotador de cada equipo en el *clutch time*:



- Extensión con otras estadísticas. Mismo análisis que en la NBA.

<https://www.nba.com/stats/players/clutch-traditional>

SEASON: 2022-23
 SEASON TYPE: Regular Season
 PER MODE: Totals
 SEASON SEGMENT: All Season Segments

Advanced Filters

Per Mode: Totals

GLOSSARY SHARE

451 Rows • Page 1 of 10

PLAYER	TEAM	AGE	GP	W	L	MIN	PTS	FGM	FGA	FG%	3PM	3PA	3P%	FTM	FTA	FT%	OREB	DREB	REB	AST	TOV	STL	BLK	PF
1 Reggie Bullock	DAL	32	48	21	27	162.8	39	11	35	31.4	9	30	30.0	8	13	61.5	3	16	19	6	3	3	3	21
2 Bam Adebayo	MIA	25	47	28	19	180.2	90	29	50	58.0	0	2	0.0	32	35	91.4	14	30	44	13	11	12	5	25
3 Anthony Edwards	MIN	21	45	26	19	162.7	121	44	104	42.3	13	39	33.3	20	28	71.4	5	23	28	11	16	6	10	16
3 Buddy Hield	IND	30	45	23	22	148.6	69	23	53	43.4	13	33	39.4	10	11	90.9	7	20	27	4	3	8	4	13

- 1 Quintetos y otras combinaciones.
- 2 Desglose por cuartos y *clutch time*.
- 3 **Faltas personales pormenorizadas.**
- 4 Rebotes ofensivos que acaban en canasta.
- 5 Posesiones.
- 6 Eficacia ofensiva de los tiempos muertos.

Faltas personales pormenorizadas

- En baloncesto, una falta personal consiste en un contacto físico con el contrincante de manera ilegal (no reglamentaria).
- Si en el momento en el que se realiza la falta, el jugador que ataca está tirando, se penaliza esa falta con tiros libres:
 - 1 tiro libre adicional si el tiro acaba en canasta.
 - 2 tiros libres si no hay canasta y la falta está dentro de la línea de tres.
 - 3 tiros libres si no hay canasta y la falta está fuera de la línea de tres.



- Aunque el jugador que recibe la falta no esté en acción de tiro, también se conceden dos tiros libres si el equipo rival está en *bonus* (ya ha hecho 5 faltas en un mismo cuarto).
- La estadística tradicional cuenta el número de faltas que comete y que recibe cada jugador.

- Sin embargo, se puede profundizar un poco más.
- **Idea de análisis:** averiguar cuántas **faltas de tiro** comete y recibe cada jugador.
- **Valor añadido para un jugador:** no solo fuerza faltas, sino que consigue que sean de tiro.
- **Descubrir una faceta negativa de un jugador:** no solo comete faltas, sino que además son de tiro para el jugador rival.

Aclaración

- Para algunos jugadores puede haber una ligera desviación entre el número de faltas totales calculadas en este análisis y el de la web de la ACB, debido a unas inconsistencias en la descripción de los datos, que no he podido solventar para todos los casos.

589	2C	00:54	Abromaitis	Falta Personal (1TL)	39	33	2	103361	Lenovo Tenerife
590	2C	00:54	Bursac	Falta Recibida	39	33	2	103361	Bàsquet Girona
591	2C	00:46	Abromaitis	Falta Personal (2TL)	39	33	2	103361	Lenovo Tenerife
592	2C	00:46	Vila	Falta Recibida	39	33	2	103361	Bàsquet Girona
593	2C	00:46	Vila	Tiro Libre fallado	39	33	2	103361	Bàsquet Girona
594	2C	00:46	Vila	Tiro Libre anotado	39	34	2	103361	Bàsquet Girona

period	time_point	player	action	local_score	visitor_score	day	game_code	team
2C	07:50	Poirier	Falta Personal	24	21	1	103350	Real Madrid
2C	07:50	Gasol	Falta Recibida	24	21	1	103350	Bàsquet Girona
2C	07:49	Gasol	Tiro Libre anotado	25	21	1	103350	Bàsquet Girona
2C	07:49	Gasol	Tiro Libre anotado	26	21	1	103350	Bàsquet Girona

Resultados

- Top 3 de jugadores con **más faltas de tiros libres cometidas**:

equipo	jugador	nº de faltas de TL cometidas y porcentaje	nº de TL concedidos	tipo de TL	cantidad
Bilbao	Sulejmanovic	55 (58.5%)	97	1TL	15
				2TL	38
				3TL	2
Zaragoza	Hlinason	54 (57.4%)	97	1TL	12
				2TL	41
				3TL	1
Girona	Miletic	52 (66.7%)	88	1TL	16
				2TL	36
Breogán	Happ	52 (52.5%)	89	1TL	16
				2TL	35
				3TL	1

- Jugador(es) con **más faltas de un tiro libre, de dos o de tres cometidas:**

equipo	jugador	tipo de TL	cantidad
Breogán	Happ	1TL	16
Girona	Miletic	1TL	16
Murcia	Sakho	2TL	45
Baskonia	Costello	3TL	4
Valencia	Radebaugh	3TL	4
Gran Canaria	Inglis	3TL	4
Madrid	Causeur	3TL	4

- Top 3 de jugadores con **más faltas de tiros libres recibidas**:

equipo	jugador	nº de faltas de TL recibidas y porcentaje	nº de TL conseguidos	tipo de TL	cantidad
Tenerife	Shermadini	96 (63.2%)	173	1TL	19
				2TL	77
Madrid	Musa	87 (66.4%)	155	1TL	19
				2TL	68
Gran Canaria	Balcerowski	70 (72.9%)	124	1TL	16
				2TL	54

- Jugador(es) con **más faltas de un tiro libre, de dos o de tres recibidas:**

equipo	jugador	tipo de TL	cantidad
Madrid	Tavares	1TL	21
Tenerife	Shermadini	2TL	77
Baskonia	Howard	3TL	9
Bilbao	Andersson	3TL	9

- Análisis adicional con las **faltas de ataque**.
- Top 3 de jugadores con **más faltas de ataque cometidas**:

equipo	jugador	cantidad
Tenerife	Shermadini	20
Obradoiro	Blazevic	20
Unicaja	Ejim	15

- Top 3 de jugadores con **más faltas de ataque recibidas**:

equipo	jugador	cantidad
Gran Canaria	Albicy	26
Obradoiro	Westermann	20
Unicaja	Alberto Díaz	18

- Permite completar las estadísticas oficiales.

FALTAS RECIBIDAS			Ver lista ▶
1	GIORGI SHERMADINI	LENOVO TENERIFE	4,9
2	Ante Tomic	Joventut Badalona	4,3
3	Dzanan Musa	Real Madrid	4,1
4	Kassius Robertson	Monbus Obradoiro	4,0
5	Jasiel Rivero	Valencia Basket	4,0

FALTAS COMETIDAS			Ver lista ▶
1	MAREK BLAZEVIC	MONBUS OBRADOIRO	3,1
2	Juan Fernández	Carplus Fuenlabrada	3,1
3	Guillem Vives	Joventut Badalona	3,1
4	Ethan Happ	Río Breogán	2,9
5	Tryggvi Hlinason	Casademont Zaragoza	2,9

- 1 Quintetos y otras combinaciones.
- 2 Desglose por cuartos y *clutch time*.
- 3 Faltas personales pormenorizadas.
- 4 **Rebotes ofensivos que acaban en canasta.**
- 5 Posesiones.
- 6 Eficacia ofensiva de los tiempos muertos.

Rebotes ofensivos que acaban en canasta

- Cuando un equipo ataca y falla el tiro, puede seguir atacando si coge el rebote → **estadística de rebotes ofensivos**.
- Conseguir rebotes ofensivos se considera uno de los aspectos más importantes del juego.
- La estadística tradicional cuenta el número de rebotes ofensivos que ha cogido un jugador.

- Sin embargo, también aquí se puede profundizar un poco más.
- **Idea de análisis:** averiguar el resultado de ese rebote ofensivo, es decir, **cuántos rebotes ofensivos acaban en canasta.**
- **Valor añadido para un jugador:** no solo coge el rebote ofensivo, sino que lo anota.

Resultados

- Equipos ordenados por puntos anotados tras rebote ofensivo. En rojo, los tres con mejor porcentaje.

equipo	rebotes ofensivos	rebotes ofensivos anotados	porcentaje	puntos anotados
Barcelona	375	209	55.73	466
Manresa	362	212	58.56	461
Betis	405	210	51.85	457
Joventut	366	208	56.83	447
Murcia	408	203	49.75	436
Granada	389	202	51.93	429
Breogán	394	207	52.54	429
Valencia	345	194	56.23	427
Madrid	359	201	55.99	426

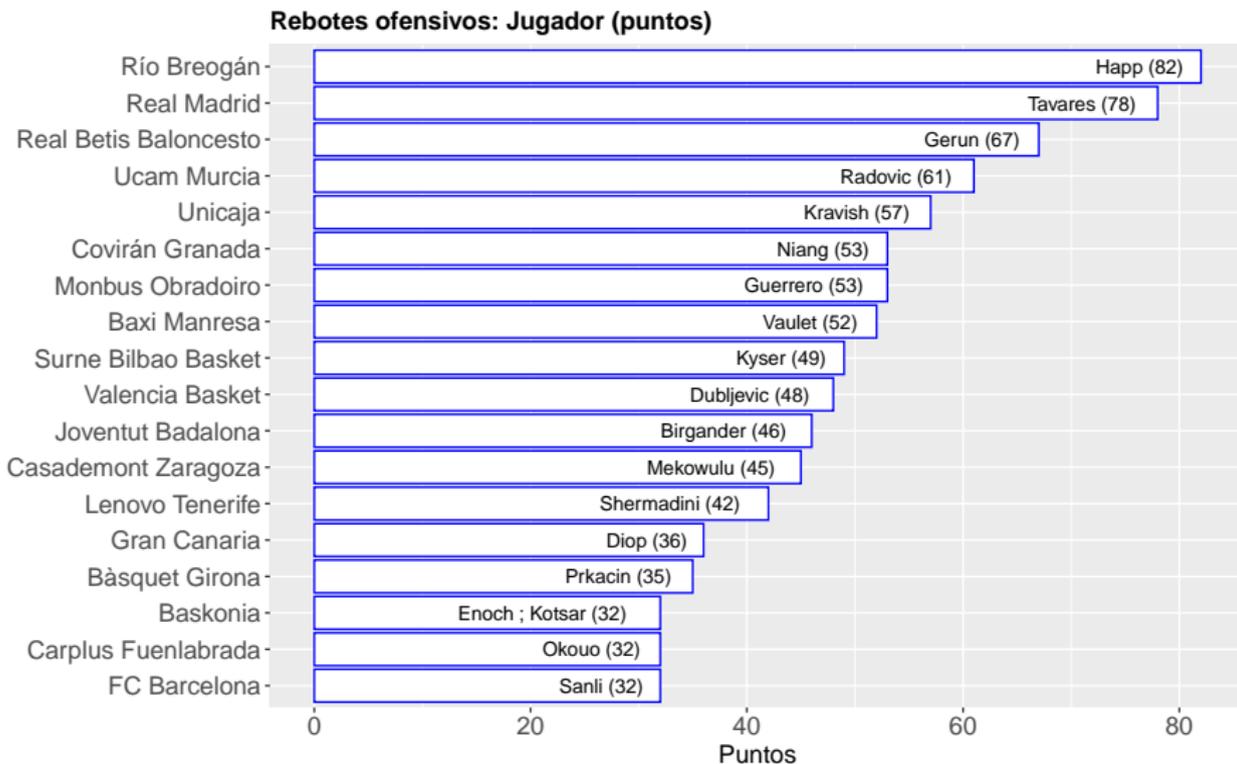
- Continúa en la siguiente diapositiva...

equipo	rebotes ofensivos	rebotes ofensivos anotados	porcentaje	puntos anotados
Gran Canaria	327	193	59.02	409
Unicaja	337	189	56.08	409
Obradoiro	336	174	51.79	382
Bilbao	355	182	51.27	375
Baskonia	323	163	50.46	363
Zaragoza	337	179	53.12	361
Girona	341	168	49.27	358
Fuenlabrada	330	167	50.61	346
Tenerife	304	160	52.63	345

- Permite completar las estadísticas oficiales.
- **Nota:** Hay una ligera desviación entre los datos totales de este análisis y los de la web de la ACB porque yo no estoy contando los rebotes de equipo.

		PAR	MIN	PT	TIROS DE 3			TIROS DE 2			TIROS LIBRES			REBOTES		
		Jug	Jug		Con	Int	%	Con	Int	%	Con	Int	%	Ofe	Def	Tot
1	Real Betis Baloncesto	34	6950:00	2.659	308	965	31,9%	628	1.266	49,6%	479	658	72,8%	413	814	1.227
2	UCAM Murcia	34	6825:00	2.714	318	1.000	31,8%	642	1.340	47,9%	476	622	76,5%	409	868	1.277
3	Río Breogán	34	6875:00	2.620	266	825	32,2%	722	1.402	51,5%	378	572	66,1%	397	790	1.187
4	Coviran Granada	34	6825:00	2.604	262	854	30,7%	678	1.324	51,2%	462	664	69,6%	392	849	1.241
5	Joventut Badalona	36	7250:00	2.948	285	838	34,0%	778	1.453	53,5%	537	689	77,9%	388	987	1.375
6	Barça	35	7000:00	2.979	342	922	37,1%	724	1.269	57,1%	505	655	77,1%	381	897	1.278
7	Real Madrid	35	7000:00	3.100	345	919	37,5%	744	1.321	56,3%	577	751	76,8%	365	968	1.333
8	BAXI Manresa	34	6850:00	2.853	287	861	33,3%	767	1.448	53,0%	458	618	74,1%	364	825	1.189
9	Valencia Basket	35	7050:00	2.916	351	1.011	34,7%	673	1.224	55,0%	517	665	77,7%	356	813	1.169
10	Surne Bilbao Basket	34	6800:00	2.587	288	869	33,1%	628	1.273	49,3%	467	630	74,1%	355	830	1.185
11	Unicaja	35	7050:00	3.041	355	990	35,9%	769	1.342	57,3%	438	616	71,1%	348	882	1.230
12	Casademont Zaragoza	34	6950:00	2.675	241	744	32,4%	737	1.401	52,6%	478	697	68,6%	348	928	1.276
13	Bàsquet Girona	34	6875:00	2.604	243	817	29,7%	718	1.387	51,8%	439	615	71,4%	345	891	1.236
14	Cazoo Baskonia	36	7275:00	3.295	444	1.098	40,4%	718	1.242	57,8%	527	698	75,5%	342	928	1.270
15	Dreamland Gran Canaria	35	7050:00	2.905	360	973	37,0%	654	1.203	54,4%	517	718	72,0%	340	838	1.178
16	Monbus Obradoiro	34	6850:00	2.694	345	953	36,2%	574	1.143	50,2%	511	670	76,3%	337	838	1.175
17	Carplus Fuenlabrada	34	6800:00	2.589	228	715	31,9%	711	1.388	51,2%	483	675	71,6%	330	781	1.111
18	Lenovo Tenerife	35	7075:00	2.893	325	908	35,8%	686	1.189	57,7%	546	695	78,6%	315	864	1.179

- Máximo anotador de cada equipo tras coger él mismo el rebote en ataque:



- Permite completar las estadísticas oficiales.

		PAR	MIN	PT	TIROS DE 3			TIROS DE 2			TIROS LIBRES			REBOTES		
		Jug	Jug		Con	Int	%	Con	Int	%	Con	Int	%	Ofe	Def	Tot
1	Ethan Happ	34	828	427	0	0	0,0%	188	337	55,8%	51	93	54,8%	106	162	268
2	Volodymyr Gerun	29	695	242	0	2	0,0%	89	153	58,2%	64	109	58,7%	95	96	191
3	Nemanja Radovic	33	746	344	9	25	36,0%	138	258	53,5%	41	54	75,9%	74	116	190
4	Bojan Dubljevic	30	673	312	42	117	35,9%	60	111	54,1%	66	76	86,8%	72	129	201
5	Tyson Pérez	28	698	354	25	82	30,5%	101	170	59,4%	77	100	77,0%	69	146	215
6	Edy Tavares	29	591	303	1	1	100,0%	122	176	69,3%	56	77	72,7%	67	109	176
7	Emir Sulejmanovic	34	707	248	24	71	33,8%	75	139	54,0%	26	49	53,1%	66	140	206
8	Vincent Poirier	26	472	219	3	4	75,0%	78	149	52,3%	54	73	74,0%	66	59	125
9	Tryggvi Hlinason	33	642	243	0	1	0,0%	97	124	78,2%	49	86	57,0%	65	100	165
10	Khalifa Diop	34	684	263	0	2	0,0%	110	187	58,8%	43	70	61,4%	63	107	170
11	David Kravish	34	655	285	10	26	38,5%	117	202	57,9%	21	35	60,0%	63	124	187
12	Simon Birgander	33	560	220	0	1	0,0%	89	165	53,9%	42	50	84,0%	62	102	164
13	Ante Tomic	34	786	443	1	3	33,3%	174	323	53,9%	92	128	71,9%	61	149	210
14	Thomas Scrubb	33	974	383	41	108	38,0%	107	195	54,9%	46	61	75,4%	61	121	182
15	Michale Kyser	34	674	260	4	18	22,2%	99	180	55,0%	50	69	72,5%	60	65	125

- 1 Quintetos y otras combinaciones.
- 2 Desglose por cuartos y *clutch time*.
- 3 Faltas personales pormenorizadas.
- 4 Rebotes ofensivos que acaban en canasta.
- 5 **Poseiones.**
- 6 Eficacia ofensiva de los tiempos muertos.

- En baloncesto, se define el concepto de **posesión** como el tiempo en el que un equipo tiene el balón en su poder y ataca.
- La posesión finaliza cuando la defensa recupera el balón (y entonces comienza su posesión).
- **Importancia:** Las variables se normalizan en base a un criterio común.
- Sirve para comparar ligas, partidos, **equipos** o jugadores.

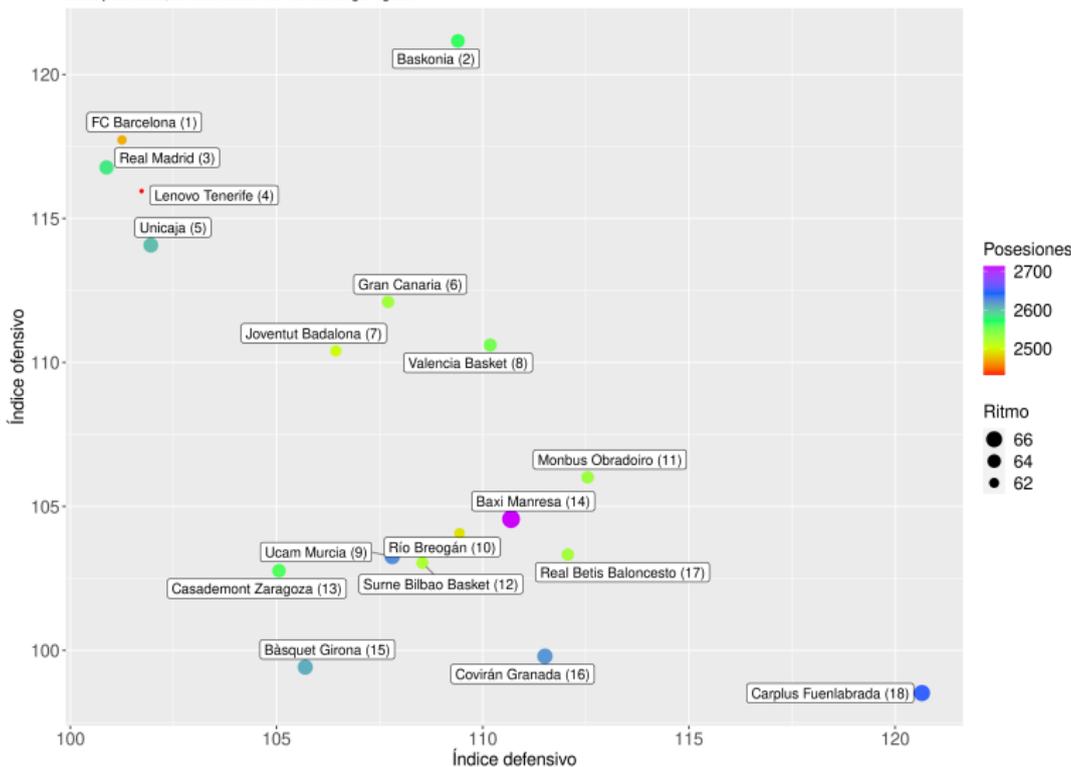
Variables de equipo relacionadas con la posesión

- **Índice ofensivo:** puntos anotados / n° de posesiones atacando.
- **Índice defensivo:** puntos encajados / n° de posesiones defendiendo.
- **Ritmo:** n° de posesiones en 40 minutos.

Resultados

Estudio de las posesiones de cada equipo.

Entre paréntesis, su clasificación al final de la liga regular.



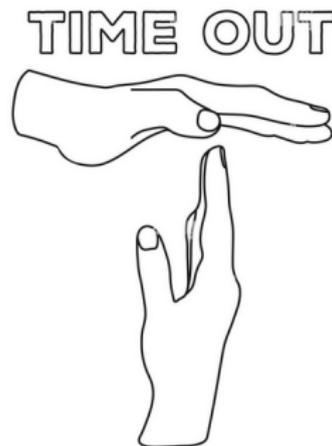
- 1 Quintetos y otras combinaciones.
- 2 Desglose por cuartos y *clutch time*.
- 3 Faltas personales pormenorizadas.
- 4 Rebotes ofensivos que acaban en canasta.
- 5 Posesiones.
- 6 **Eficacia ofensiva de los tiempos muertos.**

Eficacia ofensiva de los tiempos muertos

- La labor de los entrenadores es más difícil de analizar fríamente.
- Conclusión habitual: si el equipo **gana**, los **jugadores** son unos fenómenos. Si el equipo **pierde**, el **entrenador** es muy malo.



- Para añadir objetividad a la opinión sobre los entrenadores podemos intentar estudiar cómo intervienen durante los partidos.
- Una intervención directa es a través de los tiempos muertos.



- Los entrenadores utilizan los tiempos muertos para orientar a sus jugadores tanto en ataque como en defensa.
- Cuando un entrenador pide un tiempo muerto es porque en la siguiente jugada le toca atacar.
- **Idea de análisis:** averiguar si la jugada de ataque que propone el entrenador se traduce en anotar (al menos un punto).

Resultados

- Entrenadores ordenados por la mayor cantidad de puntos anotados tras tiempo muerto. En rojo, los tres con mejor porcentaje.

equipo	entrenador	partidos	tiempos muertos solicitados	tiempos muertos exitosos	porcentaje	puntos anotados
Murcia	S. Alonso	34	141	60	42.55	133
Valencia	Á. Mumbrú	33	108	55	50.93	128
Baskonia	J. Peñarroya	34	111	56	50.45	124
Granada	P. Pin	34	137	56	40.88	122
Obradoiro	M. Fernández	34	126	56	44.44	119
Breogán	V. Mirsic	34	139	52	37.41	107
Barcelona	S. Jasikevicius	34	107	44	41.12	101
Joventut	C. Duran	34	110	46	41.82	97
Betis	L. Casimiro	34	108	44	40.74	95
Zaragoza	P. Fisac	30	104	43	41.35	91
Bilbao	J. Ponsarnau	34	104	35	33.65	82

- Continúa en la siguiente diapositiva...

equipo	entrenador	partidos	tiempos muertos solicitados	tiempos muertos exitosos	porcentaje	puntos anotados
Gran Canaria	J. Lakovic	30	91	33	36.26	77
Manresa	P. Martínez	33	97	34	35.05	74
Girona	A.G. Reneses	34	119	35	29.41	74
Madrid	C. Mateo	34	80	31	38.75	71
Tenerife	T. Vidorreta	31	71	30	42.25	65
Unicaja	I. Navarro	33	87	32	36.78	64
Fuenlabrada	J.L Pichel	14	57	21	36.84	45
Fuenlabrada	Ó. Quintana	17	76	20	26.32	45
Fuenlabrada	J.M. Raventós	3	13	4	30.77	9
Zaragoza	M. Schiller	4	17	5	29.41	9
Gran Canaria	V. García	4	10	3	30	9

- **Curiosidad:** El entrenador del Tenerife, Txus Vidorreta, estuvo dos partidos seguidos sin pedir un tiempo muerto (jornadas 8 y 9).

Entrenadores que no pidieron ningún tiempo muerto

jornada	equipo	entrenador	partido	resultado
6	Unicaja	I. Navarro	Unicaja-Betis	106-60
8	Tenerife	T. Vidorreta	Tenerife-Valencia	94-78
9	Tenerife	T. Vidorreta	Breogán-Tenerife	61-86
17	Tenerife	T. Vidorreta	Tenerife-Betis	88-64
33	Valencia	Á. Mumbrú	Fuenlabrada-Valencia	75-93
34	Manresa	P. Martínez	Manresa-Gran Canaria	105-75

Estadística predictiva

Pronóstico del rendimiento futuro de los jugadores



- El objetivo del estudio es proponer predicciones del rendimiento futuro de los jugadores, en base a su actividad pasada.
- El enfoque estadístico se basa en combinar el análisis de **arquetipoides** con **datos funcionales escasos** (*sparse functional data*, en inglés).

- En el deporte profesional es muy habitual que los datos sean **escasos** e **irregulares**.
 - Son **escasos** porque la mayoría de jugadores no juegan mucho tiempo en la misma liga.
 - Son **irregulares** porque cada jugador juega un número diferente de temporadas.

Ejemplo de los datos disponibles (datos temporales):

- Filas: Jugadores.
- Columnas: Edad (años).
- Celdas: Valor de la estadística de interés.

	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
Claver, V.	NA	NA	NA	1.9	6.5	10.6	9.7	13	7.3	NA	NA	NA	NA	6.8	6.8	NA	NA	NA	NA	NA	NA
Gasol, P.	NA	NA	NA	4.0	5.7	14.5	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Perasovic, V.	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	24.2	18.1	17.9	17.4	13.7	NA	20.6	19	21.2
Rubio, R.	2.6	8.3	13.9	14.7	11.5	9.0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Sabonis, A.	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	26.7	24.7	31.1	25.3	28.6	34.2	NA	NA	NA	NA	NA

Metodología (I): Arquetipoides

- El análisis de **arquetipos** permite identificar en un conjunto de datos aquellos **patrones de actividad que destacan notablemente sobre el resto**.
- Estos patrones reciben el nombre de **arquetipos**, puesto que son modelos que sirven para ejemplificar un determinado comportamiento.

- El análisis de **arquetipoides** es una **extensión del análisis de arquetipos** en la que los arquetipos **se corresponden con individuos concretos de la base de datos** → arquetipoides.
- En baloncesto (y en cualquier deporte), estos arquetipoides son jugadores con un rendimiento muy específico (para bien o para mal).
- Cada **individuo** de la base de datos se representa como una **combinación de los arquetipoides**.

Definición matemática

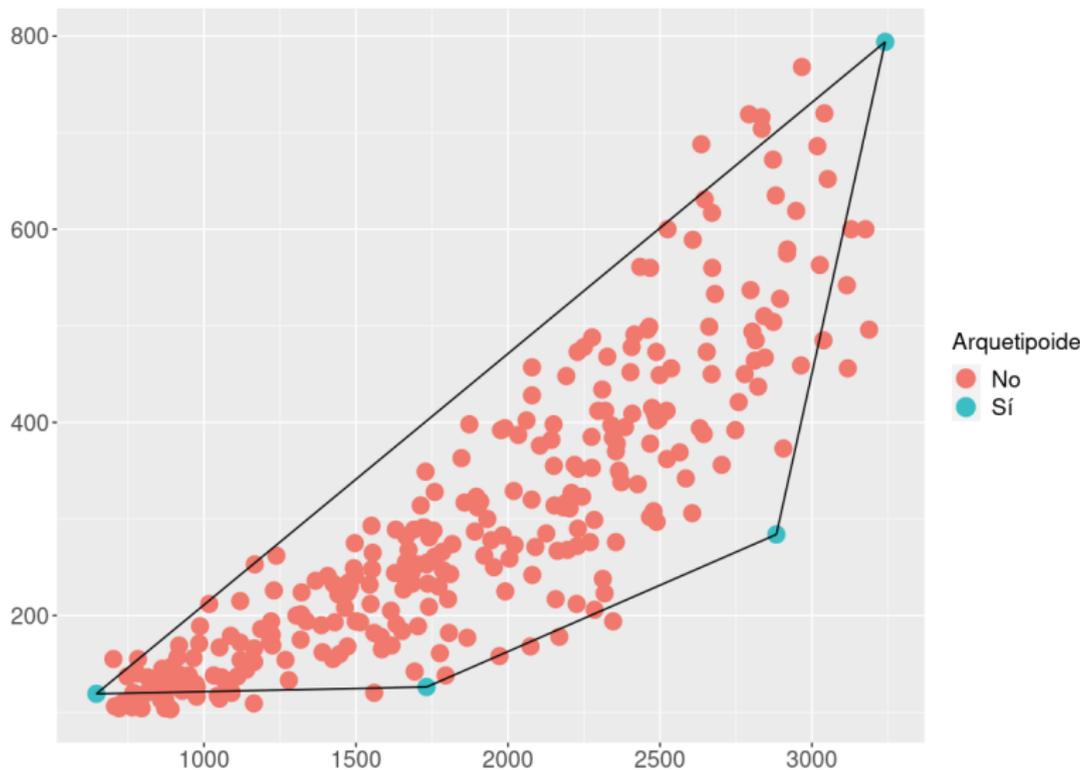
- Sea una matriz de datos multivariante $\mathbf{X}_{n \times m}$ (n individuos, m variables).
- El análisis de arquetipoides tiene por objetivo identificar k arquetipoides mediante el cálculo de las matrices α y β ($n \times k$) que minimicen la suma de residuos cuadrados (RSS) siguiente:

$$RSS = \sum_{i=1}^n \left\| \mathbf{x}_i - \sum_{j=1}^k \alpha_{ij} \mathbf{z}_j \right\|^2 = \sum_{i=1}^n \left\| \mathbf{x}_i - \sum_{j=1}^k \alpha_{ij} \sum_{l=1}^n \beta_{jl} \mathbf{x}_l \right\|^2$$

bajo las restricciones:

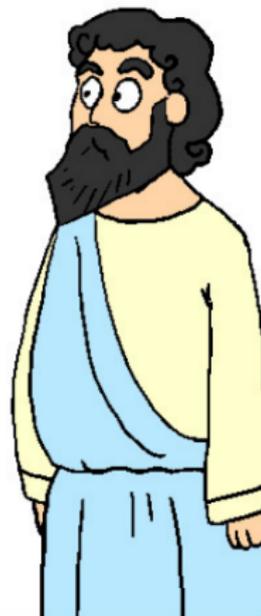
- 1) $\sum_{l=1}^n \beta_{jl} = 1$ con $\sum_{l=1}^n \beta_{jl} = 1$ y $\boxed{\beta_{jl} \in \{0, 1\}}$ $\implies \mathbf{z}_j = \sum_{l=1}^n \beta_{jl} \mathbf{x}_l$
- 2) $\sum_{j=1}^k \alpha_{ij} = 1$ con $\alpha_{ij} \geq 0 \implies \hat{\mathbf{x}}_i = \sum_{j=1}^k \alpha_{ij} \mathbf{z}_j$

- Idealmente, los arquetipoides se sitúan en las fronteras de la nube de datos.



Aproximación filosófica a los arquetipoides, o...

LA HISTORIA DE DOS ENTRENADORES LLAMADOS PLATÓN Y ARISTÓTELES



- Platón (427-347 a. C.) ya planteó la idea de arquetipo de manera abstracta al definirlo como modelo original de cualquier manifestación de la realidad.
- ¿Qué es la belleza?



¿Qué es la belleza?

¿Es un ideal de perfección y simetría?

¿Es un paradigma de brillo y esplendor?

- Aristóteles (384-322 a. C.) era más pragmático.
- ¿Que qué es la belleza? Déjate de líos...



La belleza es...



UN ARQUETIPO PARA PLATÓN:



UN ARQUETIPOIDE PARA ARISTÓTELES:



Método de los análogos

- Para afinar las predicciones, utilizamos el llamado *método de los análogos*.
- La idea es identificar jugadores con una actividad similar a la del jugador del que queremos realizar un pronóstico.

- Como sabemos cómo han actuado esos jugadores, podemos utilizar sus datos para realizar predicciones para nuestro jugador de interés.
- Para encontrar esos jugadores similares, aplicamos el análisis de arquetipoides.

Metodología (II): Datos funcionales

- Una serie temporal es una secuencia de datos ordenados en el tiempo.
- El análisis de datos funcionales sirve para analizar series temporales.
- La idea de este trabajo es considerar la **secuencia de datos de cada jugador como una función continua**, en lugar de como observaciones individuales.

- Cada dato es una función, $x_i(t)$, que representa el valor de la estadística de estudio para el jugador i en la edad t .

- El análisis de arquetipoides fue adaptado al contexto de los datos funcionales por Irene Epifanio.

- Para obtener las predicciones, utilizamos el método llamado *Regularized Optimization for Prediction and Estimation with Sparse data* (**ROPES**), desarrollado por Alexander Dokumentov y Rob Hyndman.
- Realiza predicciones con datos funcionales escasos bidimensionales.
- Proporciona intervalos de predicción, que son muy importantes porque expresan la incertidumbre asociada a las predicciones.

Resultados: Luka Doncic

- Los jugadores extranjeros tienen cada vez mayor impacto en la NBA (Nikola Jokic, Giannis Antetokounmpo, Luka Doncic...).
- La identificación de futuras estrellas internacionales ha cobrado gran importancia.

- Luka Doncic jugó en la ACB para el Real Madrid entre las temporadas 2014-2015 y 2017-2018 (entre los 15 y los 18 años).
- En el draft de la NBA de 2018, fue elegido número 3, lo que indica las altas expectativas que había con él (y que se están cumpliendo).

- Planteamos este estudio en 2018.
- Nuestra propuesta fue utilizar esta metodología para pronosticar la evolución de Doncic en futuras temporadas, en términos de su valoración ACB.
- Este pronóstico se podía usar como aproximación de su posible rendimiento en la NBA.

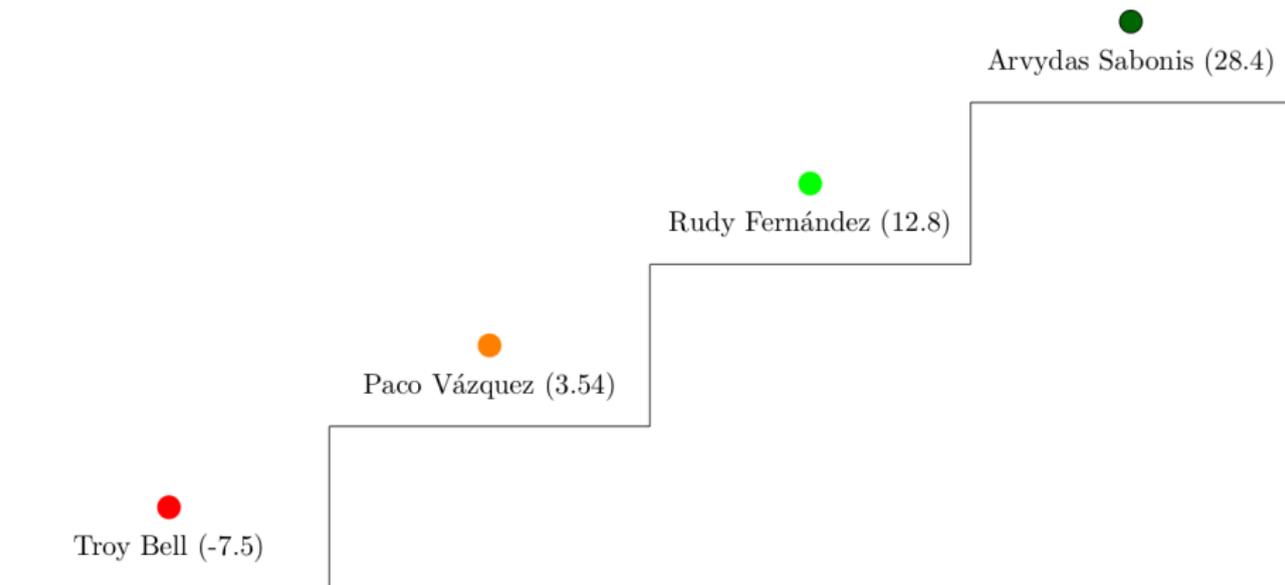
Estadística de estudio

- En la ACB y en otras ligas europeas, la valoración (o *Performance Index Rating (PIR)*, en inglés) se utiliza para evaluar el rendimiento de los jugadores.
- Es un valor único para cada jugador que resulta de sumar sus acciones positivas y restar las negativas:
 - **Suma:** puntos + rebotes + asistencias + robos + tapones a favor + faltas recibidas.
 - **Resta:** tiros de campo fallados - tiros libres fallados - pérdidas - tapones en contra - faltas cometidas.

Procedimiento

- **Base de datos:** Valoraciones promedio de todos los jugadores de la ACB desde la temporada 1985-1986 a la **2017-2018**.
- ① Expresar los datos de las valoraciones de todos los jugadores como funciones en una base ortonormal.
- ② Aplicar el análisis de arquetipoides a los coeficientes de esa base.

- Se obtuvieron 4 jugadores (arquetipoides) representativos de diferentes tipos de rendimiento (valoración entre paréntesis):



- **Troy Bell:** Representa aquellos jugadores con un rendimiento insignificante o con una escasa presencia en la liga.
- **Paco Vázquez:** Representa aquellos jugadores de nivel medio con una presencia habitual en sus equipos.
- **Rudy Fernández:** Representa aquellos jugadores de nivel muy alto.
- **Arvydas Sabonis:** Representa aquellos jugadores catalogados como súper estrellas, el nivel más alto.

- 3 Obtener los jugadores análogos a Doncic utilizando el método de los análogos.

Arquetipoides y porcentaje de similitud				
	P. Vázquez	A. Sabonis	T. Bell	R. Fernández
L. Doncic	0	54%	0	46%



54%

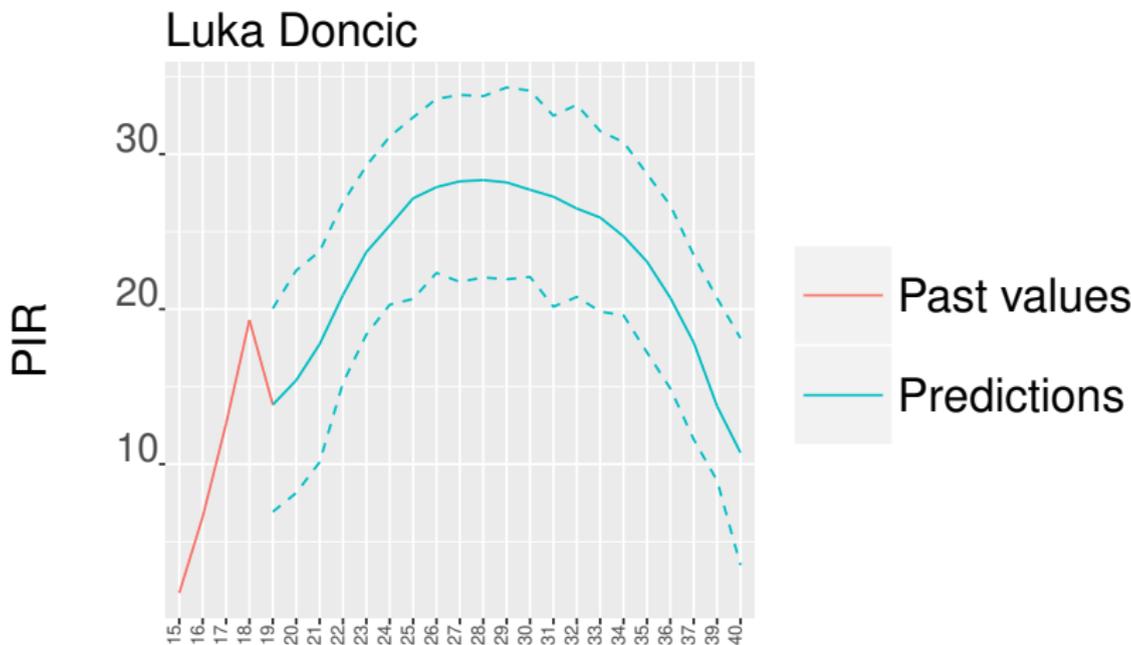


46%



- Los jugadores a los que se puede parecer Doncic son los que tienen un porcentaje de similitud con Sabonis mayor que el de Doncic.

- 4. Aplicar ROPES al conjunto de esos jugadores análogos (incluyendo a Doncic).
- Pronóstico realizado en 2018 sobre el rendimiento futuro de Doncic:

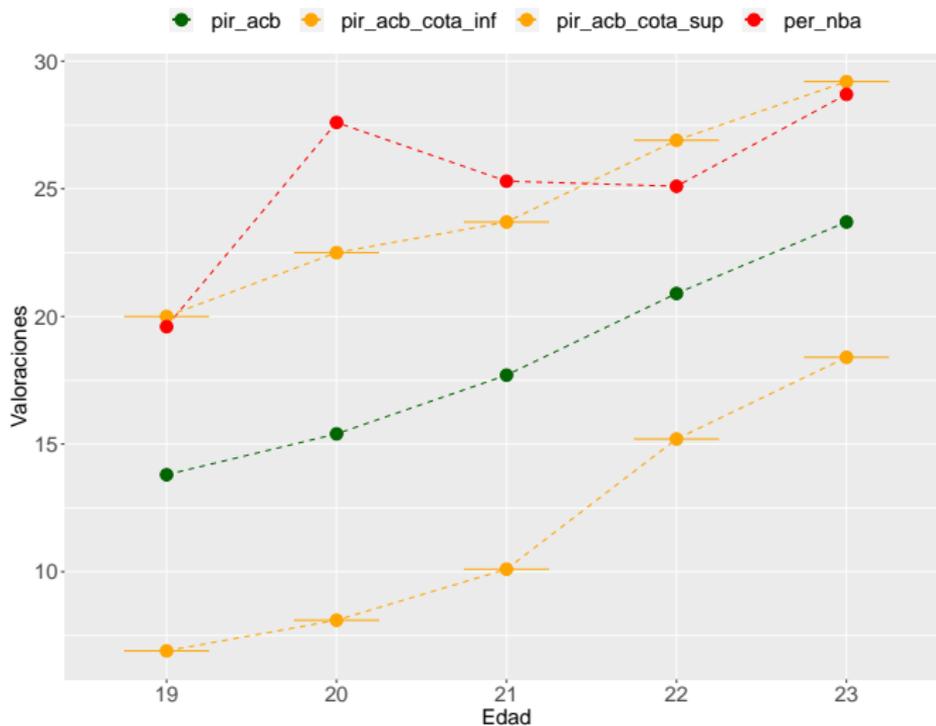


- Pasados los años, podemos comparar las predicciones con las observaciones. Doncic lleva jugando en la NBA desde los 19 años.
- La estadística de la NBA que podemos considerar más próxima a la valoración de la ACB (*Performance Index Rating (PIR)*) es el índice de eficiencia del jugador (*Player Efficiency Rating (PER)*)⁴.

⁴https://en.wikipedia.org/wiki/Player_efficiency_rating

- PER, como PIR, intenta reducir a un único número todas las contribuciones de un jugador.
- Pero no usan la misma fórmula. Por tanto, esta comparación sólo la podemos hacer de modo **aproximado**.

Luka Doncic (predicciones ACB y observaciones NBA)



- Verde: predicción PIR puntual.
- Naranja: cota superior e inferior de los intervalos de predicción PIR.
- Rojo: valor observado PER.

- El PER obtenido por Doncic es parecido a la cota superior de predicción.
- En tres de cinco casos, está dentro del intervalo de predicción. En los otros dos, lo supera.

PERO VOLVAMOS A NUESTROS ENTRENADORES...



Querido Aristóteles,
me has convencido.

He aplicado el análisis de
arquetipoides y me ha salido
como mejor jugador...





Hay que ficharlo.
Habla con su representante

Pero, Platón, ¿este
es Lebron James!
¡No tenemos dinero
para él!





No vamos a salir
nunca de tercera
regional...

Pero que no cunda el pánico,
¡gracias a los arquetipoides igual
podemos encontrar a alguien que
se le parezca y sea más barato!...





Querido Aristóteles,
qué haríamos sin
ti...

El análisis de arquetipoides aplicado al baloncesto nos permite:

El análisis de arquetipoides aplicado al baloncesto nos permite:

- 1 Identificar patrones representativos de rendimiento diferentes al resto, personificados en jugadores concretos (arquetipoides).

El análisis de arquetipoides aplicado al baloncesto nos permite:

- 1 Identificar patrones representativos de rendimiento diferentes al resto, personificados en jugadores concretos (arquetipoides).
- 2 Relacionar a los demás jugadores con los arquetipoides en base a un rendimiento parecido, y, en consecuencia, ayudar en la búsqueda de jugadores con un alto rendimiento, pero más asequibles en cuanto a exigencias económicas.

El análisis de arquetipoides aplicado al baloncesto nos permite:

- 1 Identificar patrones representativos de rendimiento diferentes al resto, personificados en jugadores concretos (arquetipoides).
- 2 Relacionar a los demás jugadores con los arquetipoides en base a un rendimiento parecido, y, en consecuencia, ayudar en la búsqueda de jugadores con un alto rendimiento, pero más asequibles en cuanto a exigencias económicas.
- 3 Combinándolo con datos funcionales, dar un pronóstico del desarrollo de los jugadores.

Received: 11 January 2018

Revised: 30 July 2019

Accepted: 5 August 2019

DOI: 10.1002/sam.11436

RESEARCH ARTICLE

WILEY

Forecasting basketball players' performance using sparse functional data

Guillermo Vinué¹  | Irene Epifanio²

¹Department of Statistics and O.R.,
University of Valencia, Burjassot, Spain

²Departamento de Matemàtiques and Institut
de Matemàtiques i Aplicacions de Castelló,
Universitat Jaume I, Castelló, Spain

Correspondence

Guillermo Vinué, Department of Statistics
and O.R., University of Valencia, 46100
Burjassot, Spain.
Email: guillermo.vinue@uv.es

Funding information

Ministerio de Ciencia, Innovación y
Universidades (AEI/FEDER, UE), Grant
DPI2017-87333-R; Universitat Jaume I,
UJI-B2017-13

Abstract

Statistics and analytic methods are becoming increasingly important in basketball. In particular, predicting players' performance using past observations is a considerable challenge. The purpose of this study is to forecast the future behavior of basketball players. The available data are sparse functional data, which are very common in sports. So far, however, no forecasting method designed for sparse functional data has been used in sports. A methodology based on two methods to handle sparse and irregular data, together with the analogous method and functional archetypoid analysis is proposed. Results in comparison with traditional methods show that our approach is competitive and additionally provides prediction intervals. The methodology can also be used in other sports when sparse longitudinal data are available.

KEYWORDS

archetypal analysis, basketball, forecasting, functional data analysis, functional sparse data

● Factor de Impacto: 1.396 ; Posición: 50/124 = 0.4 (Statistics & Probability) ; Base: JCR (2019) ; Citas: 17 ;
<https://doi.org/10.1089/big.2018.0124>

Conclusiones



- Estudios pioneros en el baloncesto europeo y español. Colocar al baloncesto español a la altura de la NBA también en la vanguardia del análisis estadístico.
- **Tratamiento de datos:** Propuesta de mejora mediante **visualizaciones** de las **estadísticas tradicionales** y realización de **nuevos análisis con los datos jugada a jugada**, no disponibles hasta ahora en el baloncesto español.
- **Línea de investigación:** Publicaciones en revistas científicas y de divulgación. Contribución al software de código libre.

- **Complementar, no sustituir**, el conocimiento técnico de directores deportivos y entrenadores.
- Extensión a cualquier deporte, tanto masculino como femenino.
- Ampliación de la cobertura periodística. Más información para cualquier persona interesada en el baloncesto.

- Importancia real de disponer de todo el conocimiento posible. Caso Kevin de Bruyne.



Alberto P. Sierra  albersierra

Actualizado a 17 de abril de 2021 11:22 CEST

INGLATERRA

Kevin De Bruyne lidera la revolución de los contratos

Kevin de Bruyne usó un estudio de Big Data sobre su juego en el Manchester City para demostrar su valor al club de manera tangible. Ha renovado por un 30% más.

Trabajo en curso

Spanish

English

Spanish



BALONCESTO PLATAFORMA DE ANÁLISIS DE DATOS

CUADRO DE MANDOS PARA LOS DATOS DE LA CARTA DE LOS PARTIDOS

¡Vamos!

Descripción

Análisis de los datos tradicionales mostrados en tablas al finalizar el partido.

Plataforma creada con el software estadístico R.
<https://www.r-project.org/>

CUADRO DE MANDOS PARA LOS DATOS DE JUGADA A JUGADA

¡Vamos!

Descripción

Análisis de los datos de jugada a jugada, que se recogen durante el transcurso del partido.

Nota: Ambos cuadros de mando pueden tardar unos segundos en iniciarse hasta que se cargan todos los paquetes R necesarios.

www.uv.es/vivigui

España Español

Temporada

2022-2023



Panel de inicio



Operaciones



Quintetos



Posesiones



Tiempos muertos



Periodos



Faltas



Rebotes ofensivos

Cuadro de mandos para los datos jugada a jugada.

Este cuadro de mandos permite visualizar y analizar los datos jugada a jugada de los partidos de la liga regular de la liga española de baloncesto ACB.

Acceso directo a cada ventana:

[Quintetos](#)

[Posesiones](#)

[Tiempos muertos](#)

[Periodos](#)

[Faltas](#)

[Rebotes ofensivos](#)

[Software](#)

Autor:

Guillermo Vinue

guillermovinue@gmail.com

CRAN.R-project.org/package=BAwiR

www.uv.es/vivigui

Gracias por utilizar esta aplicación.

Equipo

Valencia Basket

Jugador de interés

Alexander

Posición para combinar

Base

Jugador(es) para combinar

Evans Harper

Combinación

Pareja

Periodo

Todos

Rival(es)

Todos

Obtener resultados

Partidos

Resumen

Jugadores

Show 5 entries

Search:

jornada	partido	periodo	combinación	segundos	más/menos	promedio	marcador	resultado
2	Valencia Basket - Joventut Badalona	1C	Alexander, Harper	64	1	0.016	101 - 97	victoria
2	Valencia Basket - Joventut Badalona	1C-2C	Alexander, Harper	190	2	0.011	101 - 97	victoria
2	Valencia Basket - Joventut Badalona	2C	Alexander, Harper	179	4	0.022	101 - 97	victoria
2	Valencia Basket - Joventut Badalona	4C	Alexander, Harper	137	1	0.007	101 - 97	victoria
2	Valencia Basket - Joventut Badalona	4C	Alexander, Harper	51	-1	-0.02	101 - 97	victoria

Showing 1 to 5 of 109 entries

Previous

1

2

3

4

5

...

22

Next

Basketball Big Data Platform for Box-Score and Play-By-Play Data

G. Vinué, PhD

Spain

Abstract

This is the second part of a research diptych devoted to improve the basketball data management in Spain. The Spanish ACB (Association of Basketball Clubs, acronym in Spanish for the male professional league) is the top European national competition. It attracts most of the best foreign players out of the NBA (National Basketball Association, acronym for the male professional league in USA) and accelerates the development of Spanish players that ultimately contribute to the successes of the Spanish national team. However, this excellence in sportive terms is not reciprocated by an advanced treatment of the data generated by teams and players, the so called *statistics*. On the contrary, the statistics management is still very rudimentary. A first paper published in this journal introduced in 2020 the first open web application for interactive visualization of the box-score data, those data provided once the game is finished. By following the same inspiration, this new research wishes to present the work done with one-step further advanced data, namely, play-by-play data, that is provided as the game develops. This type of data allows us to extract a deeper knowledge of basketball performance. A novel data platform encompassing the visualization of both box-score and play-by-play data is available at.... The R package presented in the first paper is updated with new functions.

Referencias I



A. Cutler and L. Breiman.

Archetypal Analysis.

Technometrics, 36(4):338–347, 1994.

<http://dx.doi.org/10.2307/1269949>.



A. Dokumentov and R. J. Hyndman.

Low-dimensional decomposition, smoothing and forecasting of sparse functional data.

<http://robjhyndman.com/papers/ROPES.pdf>. Working paper, 1-31, 2016.



I. Epifanio.

Functional archetype and archetypoid analysis.

Computational Statistics & Data Analysis, 104:24–34, 2016.

<https://doi.org/10.1016/j.csda.2016.06.007>.



S. Munzert, C. Rubba, P. Meissner, and D. Nyhuis.

Automated Data Collection with R. A Practical Guide to Web Scraping and Text Mining.

John Wiley & Sons Ltd, 2015.



G. Vinué.

A Web Application for Interactive Visualization of European Basketball Data.

Big Data, 8(1):70–86, 2020.

<https://doi.org/10.1089/big.2018.0124>.

Referencias II



G. Vinué and I. Epifanio.

Archetypoid Analysis for Sports Analytics.

Data Mining and Knowledge Discovery, 31(6):1643–1677, 2017.

<https://doi.org/10.1007/s10618-017-0514-1>.



G. Vinué and I. Epifanio.

Forecasting basketball players' performance using sparse functional data.

Statistical Analysis and Data Mining: The ASA Data Sci Journal, 12(6):534–547, 2019.

<https://doi.org/10.1002/sam.11436>.



G. Vinué, I. Epifanio, and S. Alemany.

Archetypoids: A new approach to define representative archetypal data.

Computational Statistics & Data Analysis, 87:102–15, 2015.

<https://doi.org/10.1016/j.csda.2015.01.018>.



H. Wickham et al.

Welcome to the tidyverse.

Journal of Open Source Software, 4(43):1686, 2019.

[10.21105/joss.01686](https://doi.org/10.21105/joss.01686).

GRACIAS POR LA ATENCIÓN

guillermovinue@gmail.com

www.uv.es/vivigui

www.sportsandvisualdata.com