

Análisis estadístico de los Tiros Ejecutados en la Liga Uruguaya de Basketball 2019-2020

Ignacio Álvarez Andrés Sosa

Seminario del Instituto de Estadística - SIESTA

09 de Diciembre de 2020

1 Introducción

2 Proyecto: LUB 19-20

3 Exploración y visualización de los datos

4 Un modelo espacial

El proyecto se realizó de manera colectiva en el seminario *Estadística en el Deporte* entre Javier Rodríguez y Martín Grau (Licenciatura en Estadística), Lucas Barreiro (Maestría en Economía) y Ricardo Álvarez, Ignacio Álvarez y Andrés Sosa (docentes del DMMCC).

Sport analytics se refiere al análisis de datos y estadísticas avanzadas para medir el rendimiento en competencias deportivas.

En los últimos años se observa una **gran recopilación de información** en eventos deportivos y es posible establecer que la cantidad de información disponibles crece de manera exponencial.

Esta clase de análisis intenta **ser de utilidad** para jugadores, cuerpo técnico y demás personas vinculadas a equipos. El objetivo es **asesorar** sobre diferentes situaciones en competencia con el fin de realizar estrategias que mejoren el desempeño.

Objetivos fundamentales del área.

- Mejorar el rendimiento individual y colectivo de los jugadores.
- Analizar la salud del jugador y la prevención de lesiones.
- Mejorar el desempeño administrativo y financiero de la organización.
- Sistema de apuestas

La aplicación de diferentes ciencias es relevante:

- Medicina, Nutrición
- Psicología, Sociología
- Economía, Administración, Finanzas
- Matemática, Estadística

Dentro del área existe una gran cantidad de revistas científicas. Además de varias empresas que se dedica a la extracción y manejo de datos deportivos.

1 Introducción

2 Proyecto: LUB 19-20

3 Exploración y visualización de los datos

4 Un modelo espacial

El **basketball** es uno de los deportes más conocidos en el mundo. Su invención data del año 1891 cuando James Naismith ideó un deporte que se pudiera practicar bajo techo en Nueva Inglaterra.

Dentro de las particularidades del basketball se destaca **su rápida dinámica** lo cual lo convierte en atractivo para el espectador.

Esta característica permite analizar este deporte **desde muchas perspectivas y situaciones del partido**. Existe una extensa bibliografía científica en los últimos años debido al avance de datos disponibles en un sinfín de acciones.

Objetivo del proyecto

En el proyecto elegimos analizar **los tiros ejecutados** en la Liga Uruguay de Basketball 2019-2020 (LUB 19-20) desde una perspectiva estadística y de la visualización de datos.

Desde el análisis de la carta de tiro hasta la cantidad de los tiros y aciertos desde diferentes ópticas, se intenta proponer elementos que permiten analizar las fortalezas y las debilidades de cada equipos.

Los datos de cada partido se extraen de las estadísticas a tiempo real que provee FIBA (International Basketball Federation) desde su página web.

Hay **mucha información** disponible sobre cada partido:

- condiciones del partido
- carta de tiro
- estadísticas descriptivas de jugadores individuales y por equipos
- historia de eventos (estadísticas play by play)

The screenshot displays the FIBA website interface for a basketball game between Nacional and D Sporting. The score is 97-91, and the game is in the 2nd quarter with 00:00 remaining. The page provides a comprehensive overview of the game, including player statistics, team statistics, and event history.

FEDERACION URUGUAYA DE BASKETBALL - LIGA URUGUAYA DE BASKETBOL 2019-2020																	
NACIONAL					D SPORTING												
97					91												
00:00					Tiempo Reglamentario 2												
COMPETICIÓN Liga Uruguaya de Basketball 2019-2020 ESTADIO Anel Arena DETALLES DE JUEGO Salto inicial 14:15:00/11:00																	
Resumen Estadísticas completas Líneas Jugada a jugada Carta de tiro Estadísticas Penas																	
NACIONAL					D SPORTING					CNF		DSC					
Nº	JUGADOR	MIN	PTS	REB	AS	FP	OP	Nº	JUGADOR	MIN	PTS	REB	AS	FP	OP	PERFIL JUEGO	97-91
10	S. Mojica	48:11	17	3	1	0	9	1	C. Daniels	34:08	14	13	1	3	22	PERFIL PARTIDO	97-91
11	M. Rosero	37:56	15	5	6	4	26	1	A. Acosta	36:18	9	5	6	2	9	EE:04	S.A. Acosta, 30:00:00
20	M. Soderbergh	33:58	14	2	1	1	6	10	A. Boffi	30:24	11	2	1	3	9	EE:04	L.C. Daniels, 30:00:00
34	S. Alvarez	11:38	2	1	1	1	2	13	S. D'Amelio	4:42	2	1	0	0	2	EE:01	67:0-1
43	D. Morton	49:50	17	12	2	4	8	23	N. Gadsen	45:09	19	6	11	2	30	VER JUGADA A JUGADA COMPLETO	
VER ESTADÍSTICAS COMPLETAS																	
CNF				DSC				Puntos				Rebotes					
38/95 (40%)	LC	35/81 (43%)															
32/57 (56%)	3P%	25/48 (52%)															
6/28 (21%)	3P%	10/33 (30%)															
15/26 (57%)	3P%	11/12 (91%)															
55	REB	43															
15	As	24															
18	ST	7															
75/9																	

Hay (muchas) más situaciones para analizar!!

- **1777176**: Partido de Playoffs, Nacional vs Defensor Sporting, 3 de Noviembre de 2020.
- <http://www.fibalivestats.com/u/FUBB/1777176/>

Pasos necesarios:

- 1 validar identificación de partidos
- 2 descargar información para los partidos validos

Usamos herramientas de **web-scraping** para descargar y sistematizar la información en R.

Bibliotecas relevantes: `rvest`, `purrr`, `jsonlite`, `euRobasket`, `ballR`.

Formato del campeonato: Es posible separarlo en Fase Regular y Playoff.

La *Fase Regular* está formada por cuatro Rondas:

- Ronda Apertura y Ronda Clausura.
- Ronda Título y Ronda Permanencia

Los mejores ocho equipos clasifican a los *Playoff*:

- Cuartos de Final (a cinco partidos)
- Semifinal (a cinco partidos)
- Final (a siete partidos)

Hasta hoy se llevan disputados:

- 173 partidos de la Fase Regular (finalizada);
- 16 partidos de Cuartos de Final (finalizada)
- 2 partidos de Semifinales (suspendida)



1 Introducción

2 Proyecto: LUB 19-20

3 Exploración y visualización de los datos

4 Un modelo espacial

Análisis de la Fase Regular

Analizamos los 22640 tiros en la Fase Regular. En promedio corresponde a 65 tiros por equipo por partido.

EQUIPO	TIR. 2PT	ACI. 2PT	2PT (%)	TIR. 3PT	ACI. 3PT	3PT (%)
TOTALES	13947	7280	52.2	8693	3077	35.4
Aguada	937	516	55.1	796	311	39.1
Atenas	521	265	50.1	281	101	35.9
Biguá	1180	632	53.6	644	194	30.1
Capitol	982	498	50.7	712	232	32.6
Defensor Sporting	1114	632	56.7	719	262	36.4
Goes	1189	613	51.6	620	210	33.9
Hebraica y Macabi	1080	576	53.3	666	231	34.7
Malvín	1135	563	49.6	794	307	38.7
Nacional	1155	598	51.8	704	273	38.8
Olimpia	1167	597	51.2	683	245	35.9
Trouville	1231	621	50.4	735	268	36.5
Urunday Universitario	1125	612	54.4	720	250	34.7
Sayago	1131	557	49.2	619	193	31.2

Es importante destacar que los tiros libres fueron excluidos del análisis

En los Playoffs, se agrega un factor que al análisis es importante. Se repiten los equipos participantes al avanzar cada una de las series.

Es decir, resulta muy interesante analizar el desarrollo temporal de la serie al jugar contra el mismo rival.

- Malvín vs Urunday Universitario: Ganó Urunday en tres juegos
- Olimpia vs Aguada: Ganó Aguada en cuatro juegos
- **Nacional vs Defensor Sporting: Ganó Nacional en cinco juegos**
- Hebraica y Macabi vs Trouville: Ganó Trouville en cuatro juegos

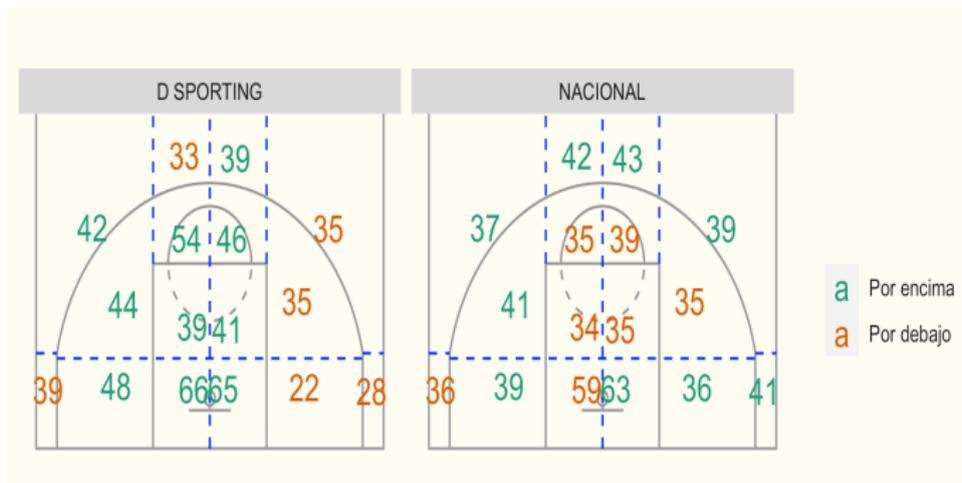
La propuesta es presentar diferentes visualizaciones sobre los tiros ejecutados en cada una de las series, para obtener conclusiones de los motivos del resultado final.

Resultado Final de la Serie: Nacional ganó los primeros dos (97/91 y 80/67), Defensor S. ganó el tercero y el cuarto (68/67 y 83/66). El quinto juego fue victoria de Nacional 77/66.

Análisis previo al comienzo de la serie

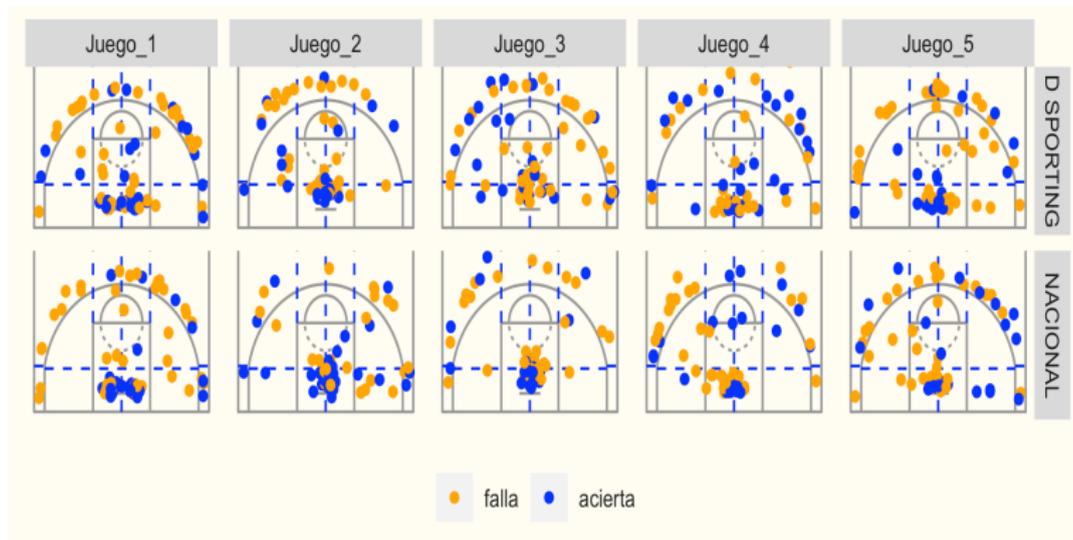
Es importante analizar el desempeño de ambos equipos en la *Fase Regular* para intentar **conocer algún patrón de juego** relevante.

Promedio de aciertos en las distintas zonas de la cancha



Posible conclusión: Defensor S. presenta mayor potencial en la zona de la Pintura y en cambio Nacional en las zonas de lanzamientos de Triples.

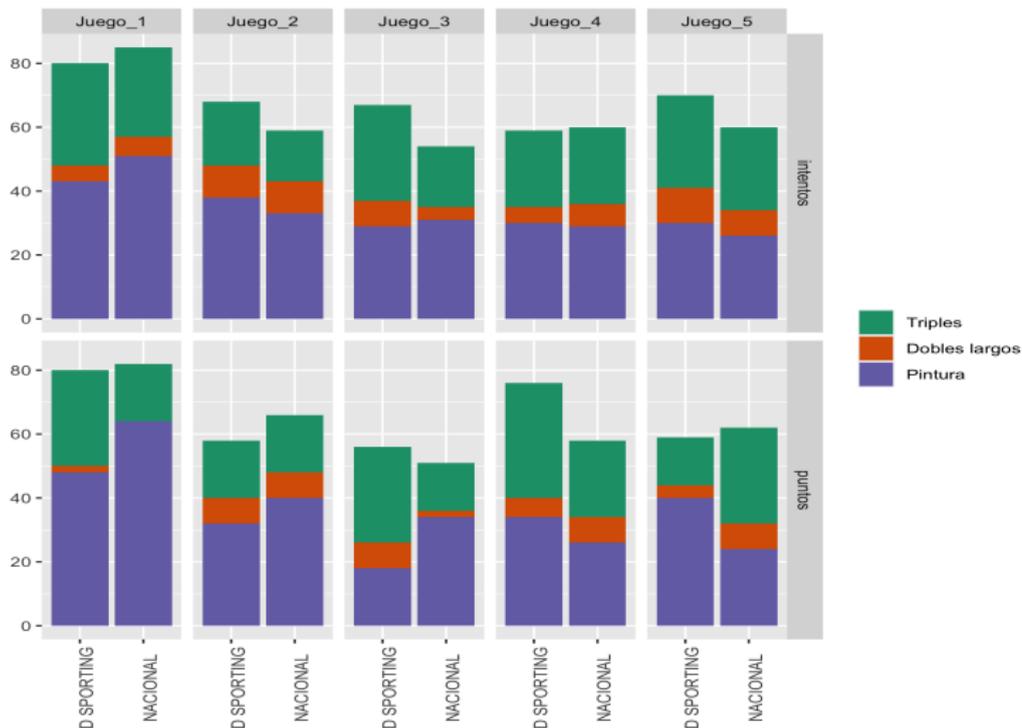
Carta de tiro: primera visualización de los datos



Si la idea es repetir el estudio de porcentajes de acierto en cada juego individual o en algún grupo de partidos nos encontramos **frente al problema** de analizar zonas con muy pocos (hasta ningún) datos.

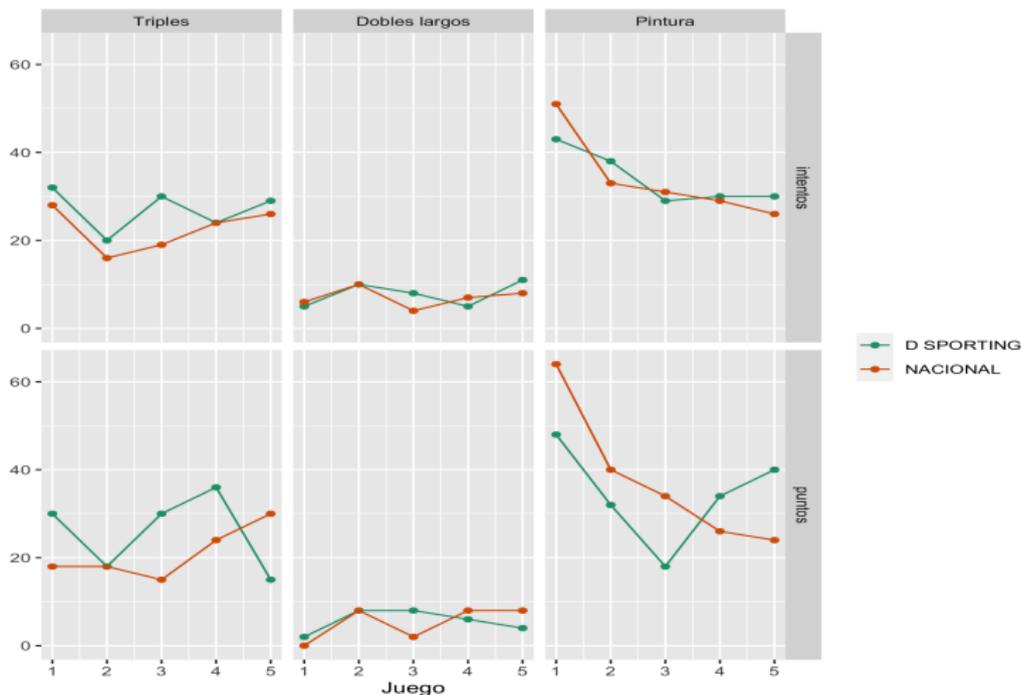
Intentos y aciertos en la serie: análisis de desempeño

Intentos y aciertos: foco en la distribución en cada juego



Intentos y aciertos en la serie: análisis de desempeño

Intentos y aciertos: foco en la evolución entre juegos



Esta propuesta de análisis no se encuentran en las estadísticas usuales. Se busca **evaluar la eficiencia en los tiros** de un equipo al compararlo con su historial de juego.

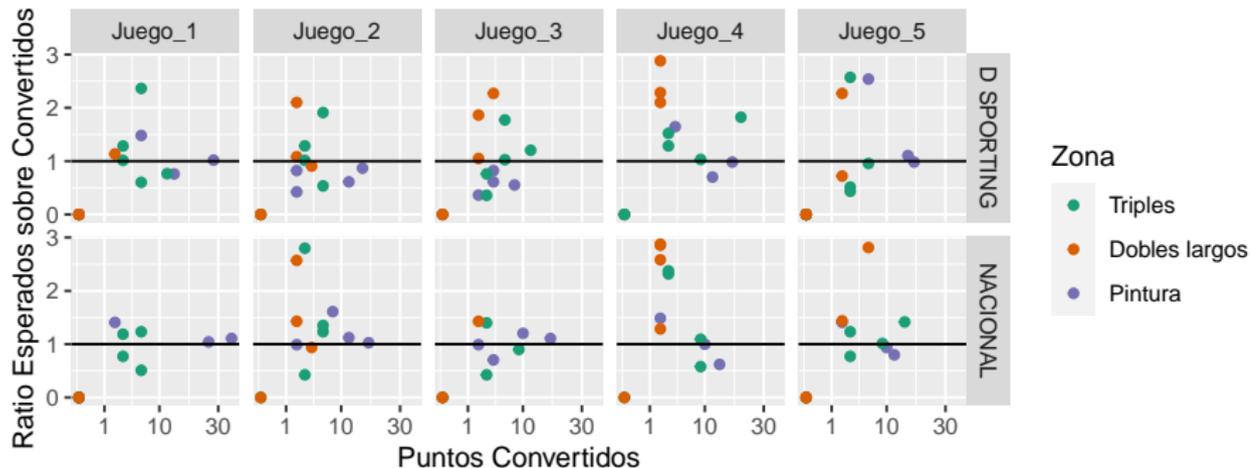
Definición de Puntos Esperados:

$$PE_{ejz} = v_z \times a_{ez} \times n_{ejz}$$

- v_z : valor de la zona z (2 o 3 puntos)
- $\times a_{ez}$: porcentaje de acierto histórico del equipo e en zona z
- $\times n_{ejz}$: cantidad de intentos por el equipo e en el juego j desde la zona z .

En la disputa de un partido es posible comparar Puntos Convertidos y Puntos Esperados en cada zona de la cancha.

Ratio entre Puntos Convertidos y Puntos Esperados contra Puntos Convertidos



Visualización: Arriba de $y = 1$ se observan las zonas donde el desempeño fue mejor que lo esperado

Posible conclusión: El ganador de cada partido tira mejor que su promedio en la zona que marcó mayor puntuación.

1 Introducción

2 Proyecto: LUB 19-20

3 Exploración y visualización de los datos

4 Un modelo espacial

En este caso, la idea es **plantear un modelo** con el fin de obtener estimaciones de *puntos esperados* para cada posición en la cancha.

$$r_{xy} \overset{\text{indep}}{\sim} \text{Ber}(p_{xy})$$

Algunas posibilidades para modelar p_{xy} :

Linear, efecto simple $\log\left(\frac{p_{xy}}{1-p_{xy}}\right) = \beta_0 + \beta_1 x + \beta_2 y$

No-parametrico $\log\left(\frac{p_{xy}}{1-p_{xy}}\right) = f(x, y)$

Una manera de estimar el modelo es:

$$r_{xy} \overset{\text{indep}}{\sim} \text{Ber}(p_{xy})$$

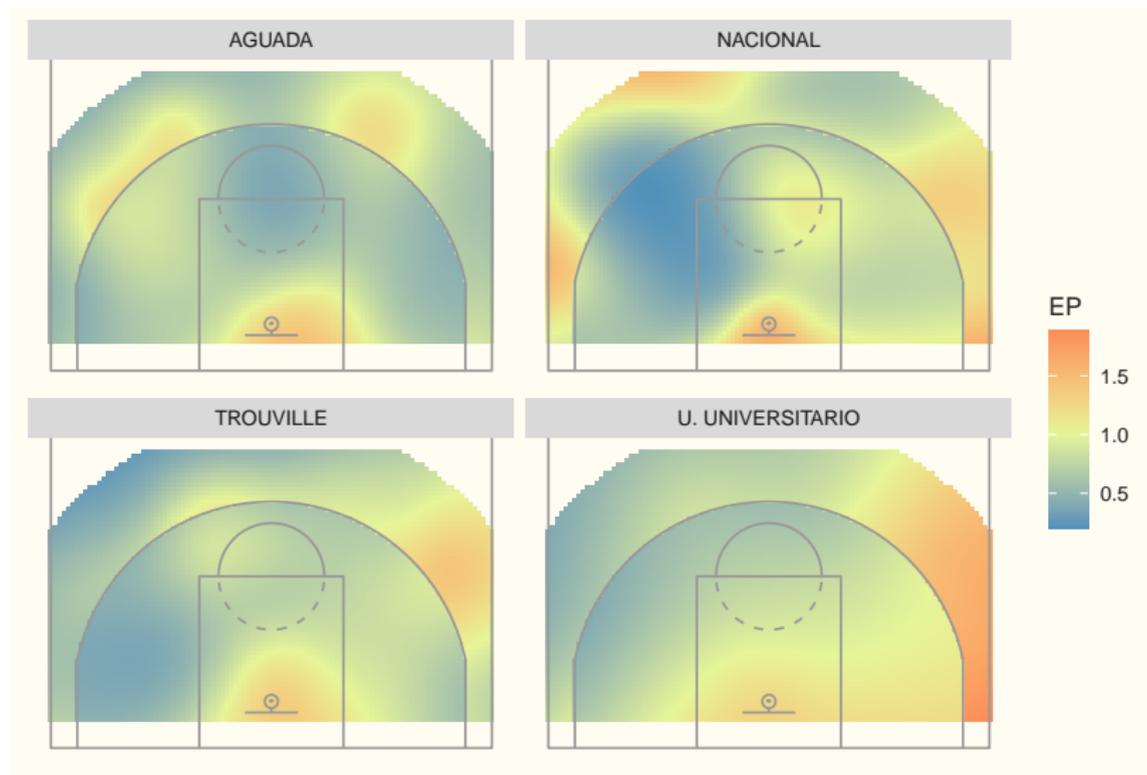
$$f(x, y) = \sum_i \phi(x, y) b_i$$

$$b \sim N(0, \sigma^2 S^{-1} / \lambda)$$

Utilizamos la función `mgcv::gam`

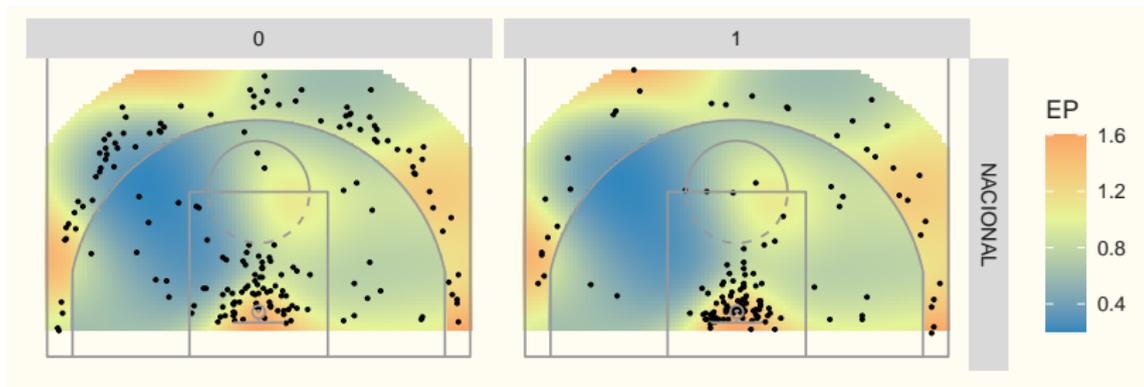
Palpitando las semifinales!

Con los datos obtenidos de todos los partidos de los Cuartos de Final, se estiman los puntos esperados de los equipo que están en Semifinales



Análisis detallado del modelo: Desempeño de Nacional (CdF)

Con el fin de comprender en mayor medida los resultados, se observan los tiros ejecutados por Nacional en Cuartos de Final (datos utilizados para estimar el modelo).



A la izquierda se encuentran los tiros fallados y a la derecha los tiros acertados.

Muchas gracias por su atención

nachalca@iesta.edu.uy

asosa@iesta.edu.uy