



FACULTAD DE
CIENCIAS ECONÓMICAS
Y DE ADMINISTRACIÓN

IESTA INSTITUTO
DE ESTADÍSTICA



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY

Análisis por equipos de los tiros ejecutados en la Liga Uruguaya de Basketball temporada 2019-2020

Ignacio Alvarez-Castro

Lucas Barreiro

Martin Batista

Martin Grau

Andrés Sosa

Serie Documentos de Trabajo

Nº 09/21

Octubre, 2021

ISSN: 1688-6453

Universidad de la República.
Facultad de Ciencias Económicas y de Administración,
Instituto de Estadística (IESTA)

Montevideo, Uruguay.



Esta obra está bajo una Licencia Creative Commons Atribución - NoComercial - CompartirIgual 4.0 Internacional.

Ignacio Alvarez-Castro, Lucas Barreiro, Martín Batista, Martín Grau,
Andrés Sosa-Rodríguez (2021). *Análisis por equipos de los tiros ejecutados
en la Liga Uruguay de Basketball temporada 2019-2020*, DT (09/21).
Instituto de Estadística, Facultad de Ciencias Económicas y de
Administración, Universidad de la República, Uruguay.
<https://www.colibri.udelar.edu.uy/jspui/handle/20.500.12008/10518>

Análisis por equipos de los tiros ejecutados en la Liga Uruguaya de Basketball temporada 2019-2020

Ignacio Alvarez-Castro ¹ Lucas Barreiro ²; Martín Grau ³; Martín Batista ⁴; Andres Sosa-Rodríguez ⁵.

Departamento de Métodos Cuantitativos, Instituto de Estadística, Facultad de Ciencias Económicas y de Administración, Universidad de la República

Resumen

El objetivo del presente documento de trabajo es analizar desde la perspectiva estadística y la visualización de datos, los tiros ejecutados en la Liga Uruguaya de Basketball 2019-2020. En el trabajo se realiza un análisis descriptivo de los tiros en la Fase Regular y se hace un especial seguimiento de la serie de playoff disputada por los equipos de Nacional y Defensor Sporting en cuartos de final.

Palabras clave: Analítica de Deporte; Basketball, Modelo Espacial, Visualización de datos.

Clasificación MSC2010: 62-07; 62P99.

Abstract

This working paper shows an analysis of Liga Uruguaya de Basketball 2019-2020 shooting data from a statistical and data visualization point of view. We present a descriptive analysis for regular season shooting data and a one playoff series between Nacional and Defensor-Sporting teams.

Key words:

JEL CODES:

Mathematics Subject Classification MSC2010: 62-07; 62P99

¹email: ignacio.alvarez@fcea.edu.uy, ORCID: 0000-0003-1633-2432

²email: lucasbarreiro96@hotmail.com

³email: grauperezmartin@gmail.com

⁴email: mbatista@gmx.com

⁵email: andres.sosa@fcea.edu.uy, ORCID: 0000-0002-6007-4373

1. Introducción

La analítica de deporte o *Sport Analytics* es un área de conocimiento que integra diferentes ciencias básicas y aplicadas con el fin de evaluar el rendimiento de jugadores y equipos en distintas competencias deportivas. En los últimos años se extraen una gran cantidad de datos en eventos deportivos de diferente naturaleza. Esta disponibilidad ha colaborado en la creciente cantidad de trabajos de investigación que vinculan técnicas matemáticas, estadísticas y computacionales (Groll *et al.*, 2018). Los objetivos que se procuran con estos estudios son variados, sin embargo se destaca la utilidad para jugadores, cuerpo técnico y demás personas vinculadas a equipos con el fin de asesorar sobre diferentes estrategias que mejoren el desempeño competitivo.

El *basketball* es uno de los deportes más conocidos y practicados del mundo, sobre todo por el gran impacto de la National Basketball Association (NBA), además se destaca por ser uno de los deportes que más ingresos genera. Es posible establecer que en su dinámica existen varios tipos de eventos diferentes en cuestión de poco tiempo transcurrido. Estos eventos se combinan en diferentes secuencias de juego que son bien delimitadas por las posesiones de cada equipo. Esta dinámica torna relevante esta clase de análisis y en la actualidad se presentan un gran volumen de datos disponibles.

En la bibliografía académica se encuentran una gran cantidad de estudios en las últimas décadas que analizan diferentes *tipos de eventos* dentro de un partido de basketball. Sin embargo, uno de los tópicos más relevantes consiste en el estudio de los tiros al aro que es el objetivo de este documento de trabajo. En Yiugman (2018), se propone estudiar la eficiencia de los tiros utilizando métodos estadísticos, en particular la probabilidad de acierto de un tiro utilizando variables como la posición donde se realiza el lanzamiento, la diferencia en el tanteador entre los equipos al momento del tiro, el jugador que lo realiza y si es local o visitante. Utiliza distintas estrategias de identificación: una regresión logística, estimador naive de Bayes y bosques aleatorios. En Jaguszewski (2020) se propone otra clase de análisis al examinar el rol los tiros de tres puntos en la NBA a través del tiempo y se pregunta si es estadísticamente significativo su impacto en el resultado de los partidos antes y después de su inclusión. Por su parte, en Skinner (2012) se propone un modelo para resolver el problema de la selección de tiro. El trabajo se basa en que seleccionar o no un tiro depende de varias variables, entre ellas se estudian la probabilidad de acierto y la capacidad que va a tener el equipo de generar una situación mejor dada la restricción en el tiempo de posesión. El mismo autor en Skinner (2016), retoma el problema y agrega nuevos conceptos como definir la eficiencia de los lanzamientos de un jugador en base a una función y no sólo en un número. El trabajo se motiva en que un jugador puede tener muy buenos porcentajes de acierto desde cierto lugar de la cancha pero tomando muy pocos lanzamientos. Pero si aumentara la cantidad de tiros, es posible que tome decisiones de juego en las cuales las condiciones del tiro podrían cambiar (ejemplo: mayor nivel en la defensa) he inducir una disminución del porcentaje. El trabajo concluye que

en el óptimo todas las opciones de tiro debieran tener la misma eficiencia marginal. Para finalizar mediante un enfoque espacial, en Shortridge *et al.* (2014) se aborda el desafío de caracterizar y visualizar la efectividad relativa de los tiros mediante el desarrollo de ciertas métricas para evaluar su variabilidad.

En Uruguay, los datos sobre eventos de cada partido los recopila la FIBA (International Basketball Federation) y se encuentran disponibles en forma instantánea en la web *FIBA LiveStats* y quedan guardados además en el sitio web de la Federación Uruguaya de Basketball (FUBB). Sin embargo, a pesar de la cantidad de investigaciones a nivel internacional, la relevancia de la temática y la existencia de datos disponibles, no encontramos desarrollos basados en la información de la liga local. Por lo cual, nos proponemos analizar los tiros ejecutados en la Liga Uruguaya de Basketball 2019-2020 (LUB 19-20) desde una perspectiva de equipos. En el trabajo formulamos una serie de elementos que permitan analizar las fortalezas y las debilidades de los equipos (tanto en ataque como en defensa) desde el análisis de la denominada Carta de tiro hasta diferentes visualizaciones en el comportamiento de los tiros ejecutados. En el trabajo, todos los análisis estadísticos los realizamos en el programa **R** y la obtención de datos la realizamos desde la web mediante las herramientas de las bibliotecas *jsonlite* y *rvest* (Ooms, 2014; Wickham, 2020).

El resto del trabajo lo organizamos de la siguiente manera. En la sección 2 introducimos la forma de disputa del campeonato y establecemos la manera de trabajar con los datos. En esta sección, analizamos de manera descriptiva los tiros ejecutados en la denominada fase regular. En la sección 3, analizamos una serie de partidos de cuartos de final. En este caso, intentamos brindar diferentes manera de visualizar la información sobre los tiros ejecutados con el fin de brindar herramientas para su análisis. Terminamos la sección con una aplicación web que se encuentra disponible sobre el comportamiento de los tiros de cada equipo en comparación al promedio de la liga. En la sección 4, establecemos las conclusiones del trabajo y posibles extensiones en la temática .

2. Análisis descriptivo de los tiros ejecutados en la Fase Regular de la LUB 19-20

Para poner en contexto el trabajo contamos brevemente como se disputó la Liga Uruguaya de Basketball en su temporada 2019-2020. Esta temporada la disputaron 13 equipos, debido a que Bohemios había logrado el ascenso para participar del campeonato pero desistió de la competencia por razones económicas. Otro dato a destacar es que en el transcurso del campeonato se debió modificar el diagrama de partidos debido a que el equipo de Atenas abandonó la competición en plena fase regular.

Es posible separar el campeonato en dos partes bien diferentes: fase regular y fase playoffs. A su vez, la fase regular se divide en dos partes. La primera parte se disputa a dos ruedas en régimen todos contra todos denominada ronda apertura y ronda clausura.

Los primeros seis equipos clasifican a la ronda título y los restantes equipos a la ronda permanencia. La segunda parte es la disputa de estas dos nuevas rondas en formato todos contra todos a una sola rueda. Al finalizar ambas rondas, los equipos que disputaron la ronda título clasifican a la fase playoff según sus posiciones con los números 1 al 6 respectivamente y los equipos que alcancen la primera y la segunda posición en la ronda permanencia también clasifican a la fase playoff con los números 7 y 8. La fase playoff (de eliminación directa) se disputa en formato serie a cinco partidos entre los ocho equipos clasificados mediante los cruces habituales en esta clase de competencias.

Vale destacar que debido a la pandemia por COVID-19, el campeonato se suspendió el 14 de marzo de 2020 con la ronda permanencia terminada y faltando una fecha para finalizar la ronda título. Se reinició el 29 de octubre de 2020 con la disputa de los últimos tres partidos de la ronda título y la fase playoff comenzó el 03 de noviembre de 2020. En diciembre de 2020, nuevamente existió una suspensión del campeonato en la disputa de los playoffs (precisamente en el primer partido de cada serie de semifinal). Este momento, es el que finaliza la recolección de datos para el documento de trabajo. Para finalizar con el contexto, destacamos que en esta edición de la LUB 19-20 los equipos clasificados a playoffs según el orden fueron: Malvín (1), Olimpia (2), Nacional (3), Trouville (4), Hebraica y Macabi (5), Defensor Sporting (6), Aguada (7) y Urunday Universitario (8).

2.1. Obtención y manejo de los datos

Los datos de cada partido se extraen mediante técnicas de web-scraping de las estadísticas que provee FIBA desde su página web. A modo de ejemplo, el partido entre Nacional y Defensor Sporting, disputado el 3 de Noviembre de 2020 se encuentra en el enlace www.fibalivestats.com/u/FUBB/1777176/. La información de cada partido es muy amplia, entre ellos se destaca las estadísticas completas jugador a jugador, la carta de tiro y el jugada a jugada.

La efectividad en el acierto de tiros es posiblemente la principal medida de éxito en el basketball (Shortridge *et al.*, 2014). En este trabajo nos interesamos en las estadísticas derivadas de los tiros ejecutados desde la perspectiva del equipo (y no de jugadores individuales). Los tiros que analizamos son en situaciones de juego (es decir sin tiros libres) y la base de datos cuenta con las variables jugador, efectividad, posición de la cancha y tiempo de ejecución en la posesión.

La Tabla 1 presenta todos los tiros ejecutados en los 172 partidos de la fase regular. Se observa por equipos, el total de intentos, los aciertos y el porcentaje de acierto separado entre tiros de dos puntos (dobles) y tiros de tres puntos (triples). Se obtiene que el promedio de efectividad de todos los equipos es de 52.2% y 35.4% en dobles y triples respectivamente. Entre otros datos relevantes, Defensor Sporting es el equipo de mayor efectividad en dobles (56.7%) y Aguada el equipo de mayor efectividad en triples (39.1%).

| EQUIPO | TIR. 2PT | ACI. 2PT | POR. 2PT | TIR. 3PT | ACI. 3PT | POR. 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 |

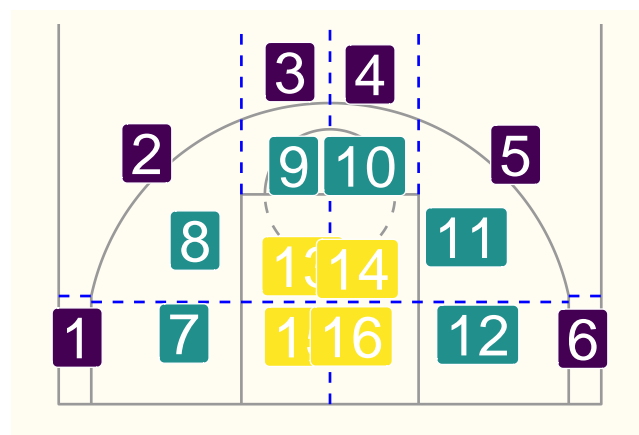
Tabla 1: Tiros, aciertos y porcentajes en la fase regular de LUB 19-20 según los equipos

2.2. División de la cancha en Zonas

En la planificación de un partido, es necesario conocer las fortalezas y las debilidades de los equipos participantes. Un dato relevante en el diseño de jugadas tanto de ataque como de defensa es conocer la zona en que se pretende culminar cierta clase de jugadas mediante un tiro ejecutado y su eficacia (del equipo o del jugador en particular) desde la posición. Para realizar estos análisis, lo usual es dividir la cancha mediante ciertas zonas. Existen varias propuestas para dividir la media cancha de ataque en zonas, nuestra propuesta de división se basa en 82games.com (2020) donde se divide la cancha en 14 zonas que se colapsan a 8 en el momento de estudiar patrones de tiro. En el trabajo, dividimos la mitad de cancha ofensiva en 16 zonas de ataque: hay seis zonas (numeradas del 1 al 6) que corresponden a los “triples”, seis zonas (numeradas del 7 a 12) que corresponde a los “dobles largos” y cuatro zonas (numeradas del 13 a 16) que corresponden a “la pintura”. Los tiros que se realizan desde la mitad de cancha defensiva (se deben a momentos muy cercanos al final de algún período del partido) se consideran por separado en el análisis.

A la izquierda de la Figura 1 se muestran las 16 zonas que utilizamos en el trabajo y a la derecha de la Figura 1 se detallan la cantidad de intentos y aciertos realizados en cada zona por todos los equipos en la fase regular. Al analizar los intentos de tiro, se observa que si bien existen tiros en cada zona la variabilidad entre ellas es notoria. De las 16 zonas, solamente 6 zonas tienen más de mil intentos en toda la fase regular, estas corresponden a las cuatro zonas de pintura (13 a 16) y las dos principales zonas de triple (2 y 5).

Con el fin de conocer la eficacia en cada zona de la cancha, se calculan los promedios de acierto. Los resultados se observan en la Figura 2. Es posible deducir que cuanto mayor



| Zona | Nombre | Tiros | Aciertos |
|------|-------------|-------|----------|
| 1 | Triple | 404 | 173 |
| 2 | Triple | 2788 | 987 |
| 3 | Triple | 894 | 301 |
| 4 | Triple | 882 | 305 |
| 5 | Triple | 3149 | 1118 |
| 6 | Triple | 485 | 192 |
| 7 | Doble Largo | 533 | 193 |
| 8 | Doble Largo | 629 | 234 |
| 9 | Doble Largo | 328 | 130 |
| 10 | Doble Largo | 393 | 163 |
| 11 | Doble Largo | 668 | 241 |
| 12 | Doble Largo | 503 | 177 |
| 13 | Pintura | 1172 | 448 |
| 14 | Pintura | 1247 | 470 |
| 15 | Pintura | 4082 | 2479 |
| 16 | Pintura | 4392 | 2745 |
| 17 | Defensiva | 91 | 1 |

Figura 1: División de la media cancha ofensiva en zonas

es la distancia al aro el porcentaje de acierto disminuye, sin embargo hay un hecho que balancea esa situación de pérdida de efectividad con la distancia al aro y es la posible situación de marca defensiva más agresiva en ciertas zonas cercanas al aro. Esto determina que la disminución no cumpliría una relación lineal, esencialmente las dos zonas de la pintura más cercanas al aro (zonas 15 y 16) presentan porcentajes de acierto mayores, mientras que el resto de las zonas tiene porcentajes de acierto similar entre ellos.

3. Visualización de datos

Al finalizar la fase regular, comienza la disputa de los playoffs momento en el cual se agrega un factor relevante para el análisis técnico y táctico de los equipos. Este factor es la repetición de partidos entre los mismos equipos en un breve período de tiempo (distancia promedio entre partido es de 4 días). Es decir, no solamente es posible analizar cómo fue el desempeño de cada equipo en la fase regular al disputar partidos con diferentes equipos sino que también es posible analizar el desarrollo temporal partido a partido al jugar contra el mismo rival. Esta clase de análisis permite ensayar diferentes tácticas de acuerdo a los resultados obtenidos. En la edición de la LUB 19-20, se jugaron las siguientes series de cuartos de final: Malvín (1) vs Urunday Universitario (8); Olimpia (2) vs Aguada (7); Nacional (3) vs Defensor Sporting (6) y Trouville (4) vs Hebraica y Macabi (5). Urunday Universitario ganó en tres partidos, Trouville y Aguada ganaron en cuatro partidos y la única serie en que se necesitó la disputa de los cinco partidos fue Nacional - Defensor Sporting. Por este motivo, se analiza el comportamiento de los tiros en esta

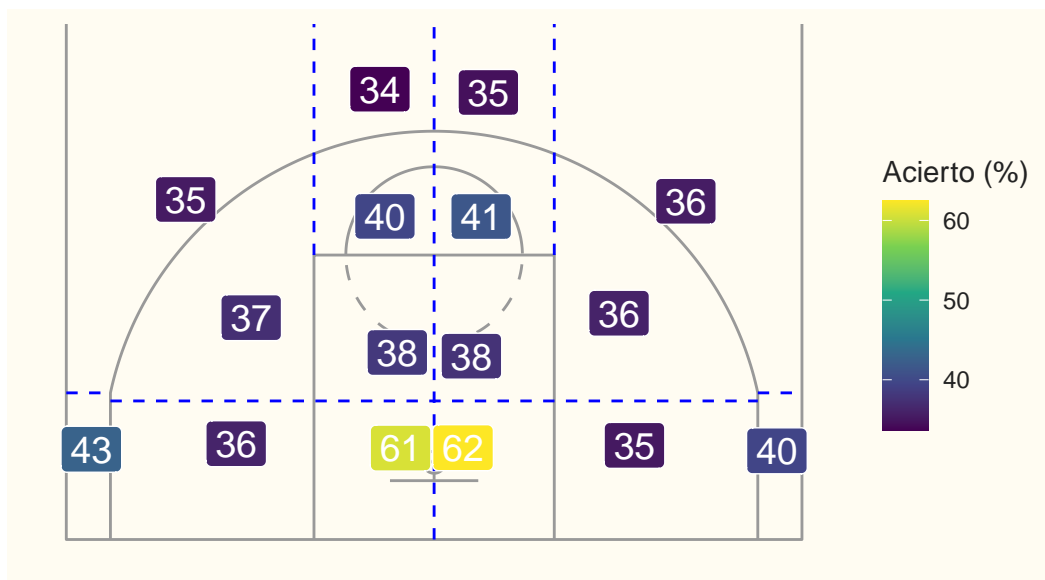


Figura 2: Porcentaje de acierto general en toda la fase regular

serie. Se presentan algunas estadísticas descriptivas de los equipos en estos partidos, pero fundamentalmente la idea es buscar distintas manera de visualizar la información con el fin de analizar el desempeño de los equipos. En esta serie, Nacional ganó los primeros dos (97/91 y 80/67), Defensor Sporting ganó el tercero y el cuarto (68/67 y 83/66) y el quinto juego fue victoria de Nacional (77/66).

3.1. Exploración gráfica de los datos de tiros en la serie de playoff: Nacional - Defensor

Para el análisis de los equipos previo al comienzo de la serie, es importante realizar un análisis similar al de la Figura 2. Esto nos permite conocer el comportamiento y si existe algún patrón de juego relevante de cada equipo en la fase regular. La Figura 3 muestra el porcentaje de acierto de los equipos en la fase regular. En verde están representadas las zonas que tienen porcentajes de acierto de tiros por encima del promedio general de la LUB 19-20 y en rojo las zonas que tienen porcentaje de acierto de tiros por debajo. La Figura 3 sugiere que Defensor Sporting se destacó durante la fase regular en la zona de la pintura y en cambio Nacional en las zonas de triple.

Para el análisis durante la disputa de la serie, es posible obtener estrategias que pueden adoptar cada equipo. La primera visualización que realizamos es la información de los tiros ejecutados en cada partido mediante la denominada Carta de Tiro. Estos datos son importantes comprenderlos para analizar las zonas de ataque, lo cual brinda una apro-

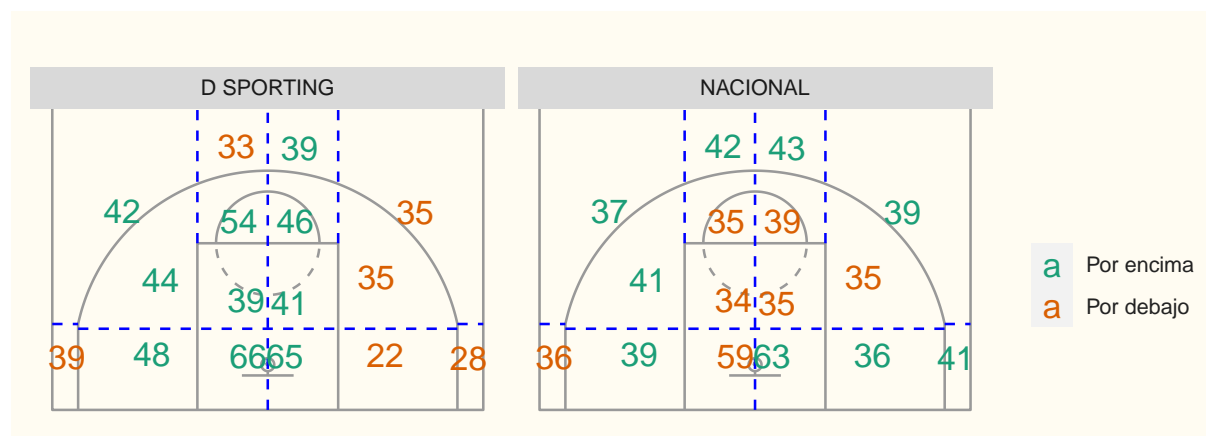


Figura 3: Promedio fase regular por Equipos.

ximación de las características ofensivas de cada equipo. Por supuesto, estos resultados también están relacionados a la estrategia en defensa del equipo rival. En la Figura 4, se observan los tiros de cada uno de los equipos en cada partido.

Si bien esta visualización es muy importante, para realizar análisis cuantitativos es relevante incluir las zonas de la cancha determinadas en la sección 2. Esto se debe a que salvo ciertas excepciones, los tiros se realizan desde diferentes posiciones. Podría ser de interés el estudio de porcentajes de acierto por zonas en cada partido o grupo de partidos como se realiza en la Figura 3 para la fase regular. Sin embargo, como se analiza un único partido se presenta la dificultad que existen zonas con muy poca información, ya que en el promedio de la LUB19-20 cada equipo ejecutó aproximadamente 65 tiros por partido. Por esta razón, para realizar análisis adicionales dentro de una serie de pocos partidos decidimos colapsar las 16 zonas iniciales en 3 zonas bien definidas. Se seleccionan los “triple”, “doble largo” y “pintura”. Cabe destacar, que la manera que utilizamos en dividir la cancha, nos permite varias formas de colapsar las zonas para obtener distintas clases de conclusiones como puede ser el criterio “izquierda”, “central” y “derecha”.

Debido a que buscamos encontrar visualizaciones de la información disponible de los tiros ejecutados, proponemos en la Figura 5 dos clases de visualizaciones que contienen la cantidad de intentos y de puntos realizados en cada zona de la cancha. A la izquierda de la Figura 5, en los paneles de la fila superior se establecen cantidad de intentos mientras que en los paneles de la fila inferior corresponden a cantidad de puntos. Cada columna representa cada partido de la serie y en cada panel se observa un gráfico de barras apiladas que se identifican mediante los colores por zonas de la cancha. A la derecha de la Figura 5, se representa la misma información desde una perspectiva diferentes a través de un gráfico de líneas donde las facetos se construyen combinando zonas de cancha y en cada panel se establece la cantidad de puntos o intentos por partido mediante colores por equipo.

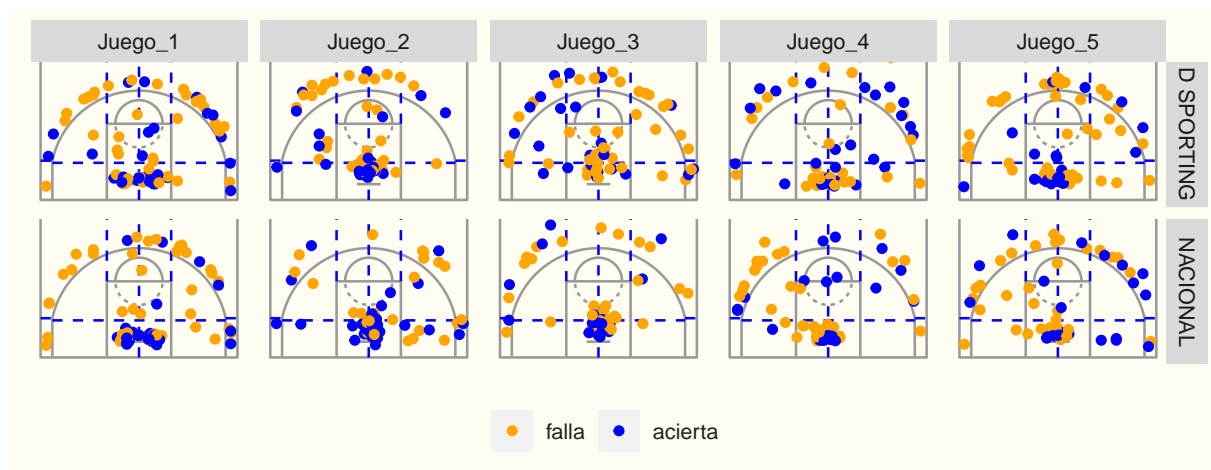


Figura 4: Carta de tiro en cada juego de la serie.

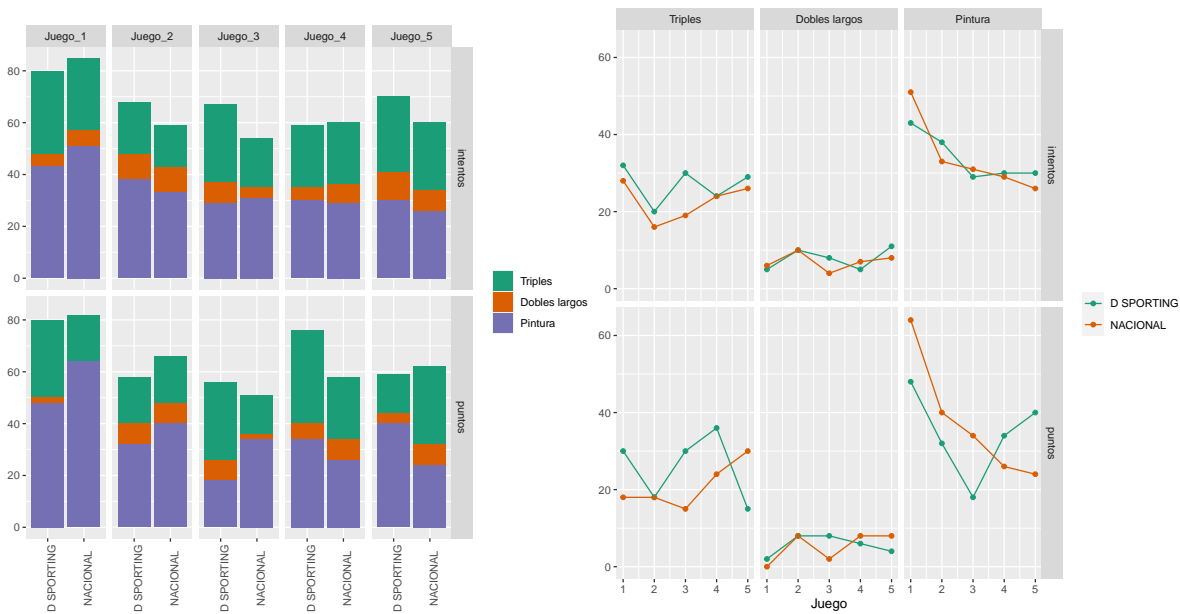


Figura 5: Visualización de intentos y aciertos en la serie

En cuanto a posibles conclusiones tácticas de la Figura 5 (izquierda) se puede establecer que en el primer partido existió una fuerte presencia en la zona de la pintura por parte de ambos equipos, siendo Nacional más eficaz. En cambio, desde la Figura 5 (derecha) es posible observar como Defensor Sporting en toda la serie presentó más intentos de triples y únicamente en el quinto juego realizó menos puntos que Nacional mediante esta zona. Es decir, esta situación puede ser una de las posibles causas de perder el juego definitivo. Por supuesto, este es un breve análisis que debe ser complementado con otra clase de análisis con el fin de comprender la situación de partido.

En esta clase de visualizaciones pretendemos obtener conclusiones sobre los partidos de manera individual o sobre la evolución de la serie en su totalidad. Si analizamos la efectividad que presenta cada equipo por zona obtenemos que Defensor Sporting desde abajo del aro (zonas 15 y 16), tiene efectividades del 62 % y 58 % respectivamente, y en el primer partido tuvo 21 y 14 intentos respectivamente, desde esas zonas. Nacional, con efectividades de 61 % en ambas zonas indicadas, en el primer partido tuvo 26 y 21 intentos respectivamente. En ese primer partido queda claro que ambos equipos intentaron atacar por la vía de las zonas de mayor efectividad. En los posteriores partidos se observa que la distribución de tiros cambia, al menos referido a esas zonas de alta efectividad, posiblemente debido al análisis realizado por los cuerpos técnicos de esta situación. Además, en la conjunción de figuras se observa que porcentaje de los puntos convertidos son realizados desde el triple, o por doble largo, o en la pintura. En el caso de Nacional es claro ver cómo aumentó el caudal de puntos (siempre como referencia de porcentaje sobre el total), en base a triples, a lo largo de los cinco juegos, pilar fundamental de su juego en la fase regular. En el caso de Defensor Sporting fue más irregular la manera de realizar el ataque a lo largo de los juegos, pero en particular en el quinto juego es muy clara la baja efectividad en los tiros de tres puntos.

3.2. Puntos esperados

Una herramienta que nos parece muy valiosa analizar y no se encuentra en las estadísticas usuales de los partidos es conocer si los equipos tuvieron la eficacia en un partido de acuerdo a su historial de tiro. Nuestra propuesta es calcular la variable denominada *puntos esperados* que se calcula en base a los intentos, los tiros acertados y la media de tiro de cada equipo en la fase regular en cada zona de la cancha. La ecuación (1) define la variable denominada puntos esperados PE_{ejz} por el equipo e en el juego j y la zona z .

$$PE_{ejz} = v_z \times a_{ez} \times n_{ejz} \quad (1)$$

donde v_z es el valor de los puntos obtenidos en cada zona z (2 o 3 puntos según la zona), a_{ez} es la probabilidad de acierto del equipo e en zona z estimado el porcentaje de acierto histórico y n_{ejz} es la cantidad de intentos que realizó el equipo e en el juego j desde la zona z .

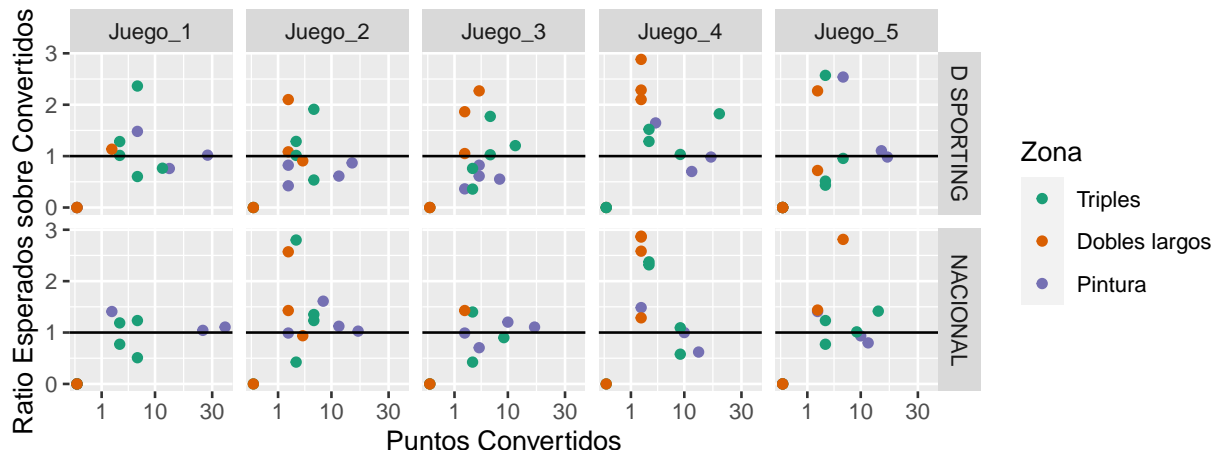


Figura 6: Ratio de Puntos Convertidos sobre Puntos Esperados

En la disputa de cada partido es posible comparar los puntos convertidos con los puntos esperados en cada zona de la cancha. En la Figura 6 proponemos visualizar el ratio entre los puntos convertidos y los puntos esperados en función de los puntos convertidos (el color representa cada zona de la cancha). La idea general en esta visualización es que los puntos que se encuentran encima de la línea horizontal $y = 1$ representan zonas en las cuales el equipo en un partido ejecutó de mejor manera que su promedio histórico. Por ejemplo, es posible observar que el ganador en cada uno de los partidos de la serie realizó tiros de mejor manera que su promedio en la zona de mayor puntos convertidos en el partido.

3.3. Modelo espacial

Una manera alternativa de generalizar la información que brinda la carta de tiro o es mediante la construcción de un modelo estadístico que estime la probabilidad de acierto como función de las coordenadas espaciales de la cancha en cada tiro ejecutado. Esta manera de abordar el problema evita el colapso de ciertas zonas de áreas de interés.

Mediante la ecuación (2) se propone un modelo logístico en el cual la variable de respuesta es la variable indicadora r_{ei} (toma el valor $r_{ei} = 1$ cuando el tiro i realizado por el equipo e es acertado y $r_{ei} = 0$ en caso contrario), la probabilidad de acierto para cada equipo se modela exclusivamente dependiendo de la posición de la cancha (coordenadas (x, y)) en que se efectuó el tiro. El modelo queda determinado

$$\log \left(\frac{p_e}{1-p_e} \right) = f_e(x, y); \quad r_{ei} \stackrel{indep}{\sim} Ber(p_e); \quad (2)$$

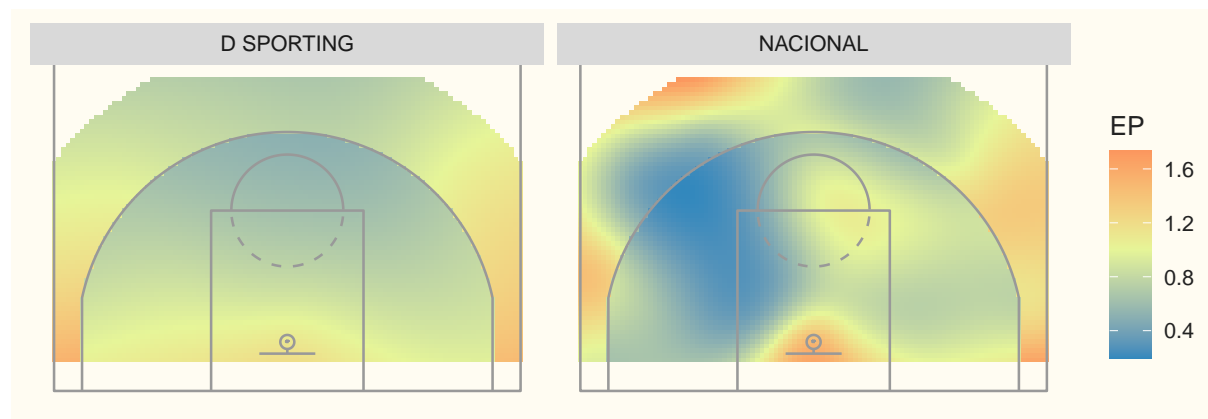


Figura 7: Patrón espacial de puntos esperados.

Vale destacar que en este modelo no se asume una forma funcional concreta para $f_e()$ (por ejemplo, polinómica) lo que brinda una mayor flexibilidad al modelo. El principal resultado que obtenemos es una descripción del patrón de tiro de un equipo en todas las posiciones de la cancha.

El modelo propuesto es simple de interpretar pero contiene algunas simplificaciones que en trabajos posteriores se deben generalizar. En primer lugar, se supone que todos los tiros del equipo son independientes cuando en realidad no lo son (ejemplo: hay los tiros ejecutados por el mismo jugador). Además, es claro que hay otros factores que afectan la probabilidad de acierto como ser el tiempo transcurrido dentro de la posesión, el tiempo transcurrido de partido, o más en general la situación del partido en que fue efectuado el tiro.

Con el fin de obtener las estimaciones de *puntos esperados* en cada posición de la cancha, suplantamos el porcentaje de acierto en cada zona por la probabilidad de acierto estimada por el modelo en la ecuación (2). Con la información de los tiros en cada partido de la serie, se estiman los puntos esperados de los equipos que están en semifinales. La estimación del modelo en **R** se realiza mediante la utilización de la biblioteca *mgcv* (ver Wood (2017)).

En la Figura 7 se observa el resultado de aplicar el modelo para ambos equipos de la serie, las zonas azules representan lugares de la cancha en donde el equipo tiene baja efectividad y en las zonas naranjas es donde la efectividad es alta.

Si esbozamos unas conclusiones desde la visualización, en general el patrón de tiro de Defensor Sporting en los cinco partidos no presenta sorpresas. Se observa que las zonas de mayor puntos esperados son los triples de las esquinas (zonas 1 y 6) lugar que en cada intento se espera obtener casi 1.5 puntos, luego la pintura cercana al aro (zonas 15 y 16) y los triples de las zonas 2 y 5 muestran aproximadamente un punto esperado por cada intento. Finalmente dobles largos y triples desde el centro muestran representaciones de menor efectividad. A su vez el panel de la derecha en la Figura 7 muestra el

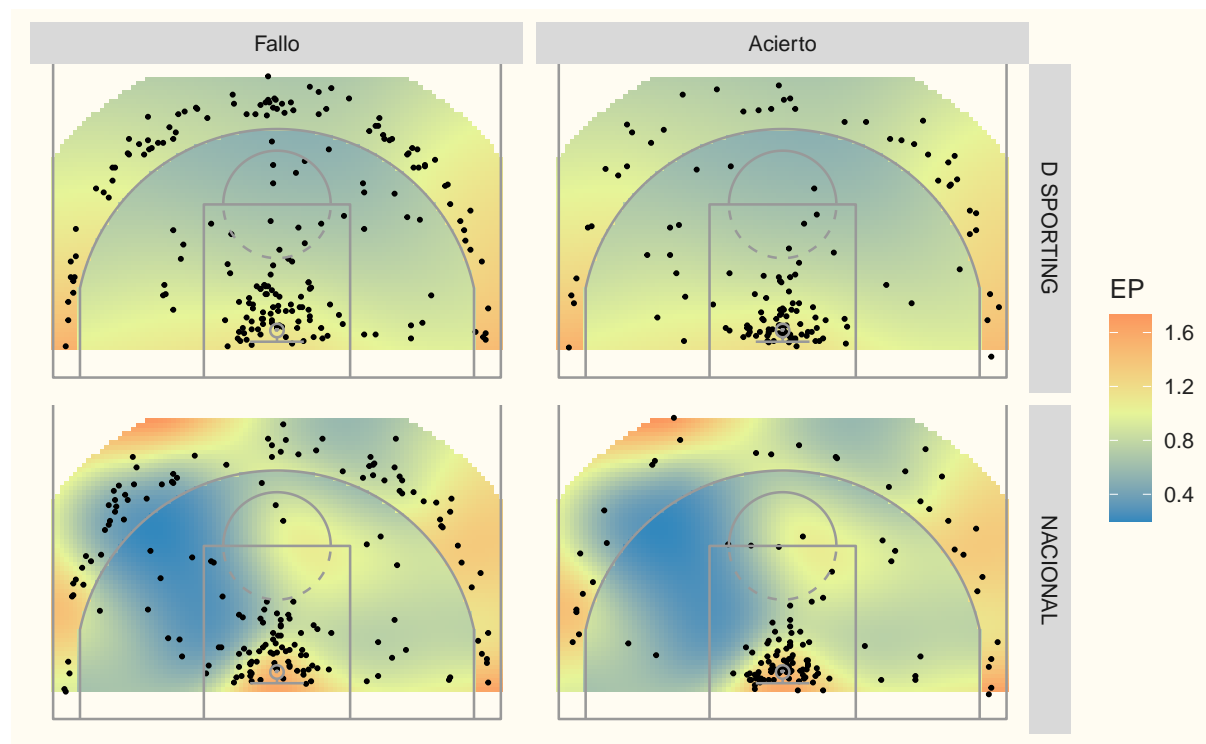


Figura 8: Tiros observados (aciertos y fallas) sobre patrón espacial de puntos esperados.

modelo estimado para Nacional, donde se observan varias regiones de alta efectividad en distintos lugares de la cancha. Se destaca visualmente una zona de muy mal ejecución de tiros en la zona derecha del ataque (tanto en dobles como intentos de tres puntos) pero se continúa por una pequeña zona de triples a mucha distancia con casi 1.5 puntos esperados por intento.

Con el objetivo de comprender de mejor manera los resultados obtenidos en el modelo, agregamos las posiciones de los tiros observados al separar aciertos y fallos. La Figura 8 presenta los tiros ejecutados separados en fallos y aciertos que son los datos utilizados para la estimación del modelo. Por ejemplo, se destaca que en Nacional, las posiciones de la cancha donde el modelo establece los puntos esperados más bajos aparecen varios tiros ejecutados y todos fallados.

3.4. Visualización interactiva

Debido a lo relevante de la modelización espacial propuesta en la sección anterior decidimos elaborar una visualización de los patrones de tiros de cada equipo en sus partidos de fase regular.

Esta aplicación permite al lector visualizar el patrón espacial de cada equipo en base al modelo (2) y compararlo con el patrón observado por todos los equipos en conjunto. Para su creación empleamos la biblioteca *shiny* donde generamos una aplicación web la cual se encuentra disponible en IESTA-Basket.

La aplicación tiene dos pestañas, en primer lugar se presentan tres gráficos que describimos brevemente y en la segunda pestaña se muestra la descripción de los datos utilizados junto a sus variables.

La interacción que permite la aplicación es básica, el usuario elige el equipo de la liga que quiera analizar. En el equipo seleccionado se presentan tres gráficos de la mitad de cancha ofensiva. En el panel izquierdo superior de la Figura 9 se encuentra el resultado del modelo espacial que se construye con los tiros de todos los equipos de la liga. Esta figura sugiere una relación no lineal entre puntos esperados y distancia al aro, en particular se puede identificar una zona de muy baja efectividad en la zona pegada a la pintura baja (zonas 7,12,13,14) y luego la efectividad mejora cuando el tiro se realiza de mayor distancia, aún en la región de tiros de 2 puntos. Por otro lado las zonas de mayor puntos esperados corresponde a la pintura baja (15 y 16) y los tiros de 3 puntos. Al interior de las zonas de triples (1 al 6) la figura sugiere un descenso lineal en la efectividad con la distancia al aro. El panel derecho superior de la Figura 9 también muestra el resultado del modelo espacial de tiros pero al utilizar únicamente los tiros del equipo seleccionado. Se puede observar en este caso que la zona de triple del lado derecho del ataque y las zonas debajo del aro de Defensor Sporting son las que tienen mayor cantidad de puntos esperados.

Para facilitar la comparación de estos dos patrones espaciales se muestra en el panel inferior de la Figura 9 el cociente entre las dos estimaciones previas. Los valores por encima de 1 indican que Defensor Sporting tiene mayores puntos esperados que el promedio de toda la liga y viceversa.

4. Comentarios Finales

El área de investigación **Sport Analytics** aplicada al basketball cuenta con mucha literatura a nivel mundial y creciente actividad en los años recientes. La liga profesional en Uruguay cuenta con una base de datos pública que ofrece muchas variables de cada partido disputado. Sin embargo, a pesar de la creciente literatura y la disponibilidad de datos públicos, en esta área no se encuentran investigaciones desde la perspectiva académica.

En este trabajo, procuramos realizar un análisis descriptivo de los tiros en la fase regular, al dividir la cancha en 17 zonas de ataque. Identificamos zonas de mayor incidencia de cada equipo al analizar la eficacia en cada una durante todo el campeonato. Adicionalmente, proponemos varias visualizaciones para el análisis táctico de una serie de playoff, donde

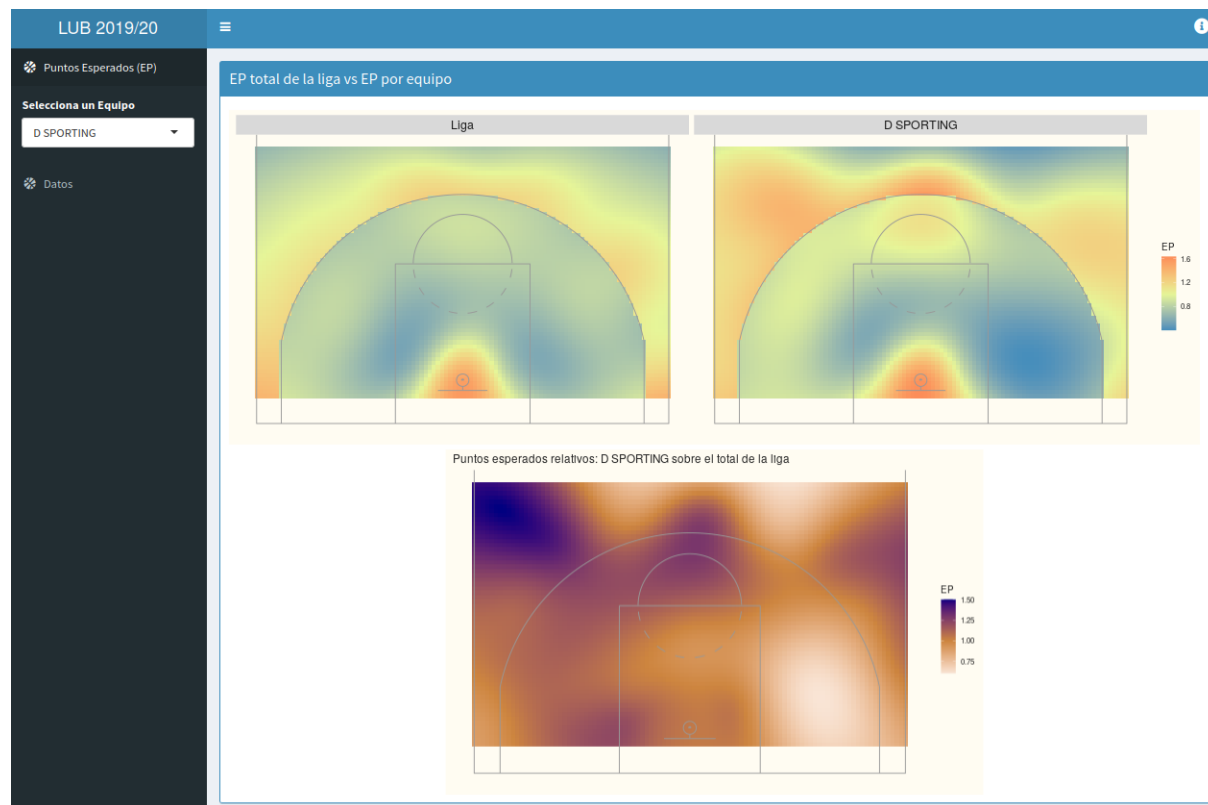


Figura 9: Visualización del shiny para Defensor Sporting.

las condiciones de los partidos puede cambiar pero los rivales son los mismos. Con el fin de complementar estas estadísticas, proponemos un modelo para la probabilidad de acierto de un tiro que depende únicamente de la posición en la cancha que el tiro se realizó. El modelo propuesto es no-paramétrico y nos ofrece como principal resultado una superficie continua para la probabilidad de acierto de cada equipo. Con el fin de difundir los resultados de nuestro modelo, elaboramos una aplicación web con visualización interactiva de los patrones de tiro de cada equipo durante la temporada 2019-2020.

Este documento de trabajo nos representa un inicio auspicioso a todo el estudio que es posible realizar con los datos disponible de la LUB. Tenemos varias líneas de trabajo para continuar la investigación. Por un lado, creemos relevante generalizar este análisis desde la perspectiva jugador o grupos de jugadores. Eventualmente, podemos generar rankings de jugadores según su habilidad en distintas zonas de la cancha. Por otro lado, queremos realizar otro tipo de análisis con los datos disponibles que proporcionan mayor detalle de los eventos que suceden dentro del partido. Estos datos nos permitirían obtener estadísticas de los equipos por posesión de juego, con el objetivo de analizar la eficiencia en diferentes rubros. Por ejemplo, podríamos generalizar el modelo de la probabilidad de acierto mediante otros factores como lo es el tiempo de juego.

Referencias Bibliográficas

82games.com (2020). Locations.

Groll, A., Manisera, M., Schauburger, G., y Zuccolotto, P. (2018). Guest editorial ‘statistical modelling for sports analytics’. *Statistical Modelling*, 18(5-6):385–387.

Jaguszewski, M. (2020). Increasing role of three-point field goals in national basketball association.

Ooms, J. (2014). The jsonlite package: A practical and consistent mapping between json data and r objects. *arXiv:1403.2805 [stat.CO]*.

Shortridge, A., Goldsberry, K., y Adams, M. (2014). Creating space to shoot: quantifying spatial relative field goal efficiency in basketball. *Journal of Quantitative Analysis in Sports*, 10(3):303–313.

Skinner, B. (2012). The problem of shot selection in basketball. *PloS one*, 7(1):e30776.

Skinner, B. (2016). Handbook of statistical methods and analyses in sports. capítulo Optimal Strategy in Basketball. CRC Press.

Wickham, H. (2020). *rvest: Easily Harvest (Scrape) Web Pages*. R package version 0.3.6.

Wood, S. (2017). *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC.

Yiugman, C. (2018). Predictive analysis of successful basketball shots?: The euroleague case. Tesis de maestría, Master Tesis of Science in Business Analytics. Sabanci University.

Instituto de Estadística

Serie Documentos de Trabajo



FACULTAD DE
CIENCIAS ECONÓMICAS
Y DE ADMINISTRACIÓN

IESTA INSTITUTO
DE ESTADÍSTICA



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY

Gonzalo Ramirez 1926, Piso 1, Oficina 23 - C.P. 11200 -
Montevideo, Uruguay
Teléfono: (598) 2410 2564
<https://iesta.fcea.udelar.edu.uy/>
Área Publicaciones

Octubre, 2021
DT (09/21)