

Unbetting Football, Capítulo II

La Caja de Pandora

La primera vez que leí un libro yo tenía 21 años. Sin embargo, ese primer libro lo leí por accidente porque un amigo lo dejó olvidado sobre la mesa. Aunque la portada de inmediato me llamó la atención porque reconocí a quien aparecía, confieso que lo miré durante un rato sin saber qué hacer. El acto de la lectura me era algo totalmente ajeno, ni siquiera entendía para qué o por qué lo hacían otras personas. Pero abrí ese libro y descubrí la lectura, y digo que la descubrí, así tal cual, porque para mí fue como descubrir algo de lo que no tenía ni la menor noción de su existencia.

Hace poco volví a leer ese mismo libro, casi 20 años después, y este capítulo es mi manera de dar las gracias.

En el año de la pandemia, en pleno 2020, con la práctica totalidad de las competiciones deportivas canceladas, de acuerdo con datos de Forbes, jugadores como Lionel Messi y Cristiano Ronaldo percibieron en conjunto una cantidad cercana a 5,000 millones de pesos mexicanos [1]. Sin embargo, si sumamos el ingreso de los 10 jugadores mejor pagados de ese mismo año, el monto alcanzaría casi los 12,000 millones de pesos y, aun así, nos quedaríamos cortos para igualar lo que ingresó una sola persona vinculada al fútbol: Denise Coates.

El Imperio del Tiempo

Hacia el inicio del año 2000 una empresa familiar solicitó un préstamo al *Royal Bank of Scotland* por un monto equivalente a casi 450 millones de pesos. No mucho después la empresa originaria de la ciudad Stoke-on-Trent, en el Reino Unido, realizó la compra de un dominio que al año siguiente sería lanzado como www.bet365.com. Para 2005 la empresa vendió sus 59 negocios de tragamonedas y casinos y decidió enfocarse en el negocio en línea [2].

En el 2020, el año de la pandemia, Bet365 llegó a ser una empresa cuyas perspectivas de crecimiento pueden plasmarse a través del siguiente dato: ese mismo año el ingreso (salario más bonos y participación en acciones de la empresa) de su CEO, Denise Coates, rebasó el equivalente a los 13,000 millones de pesos mexicanos [3]. Para ponerlo en perspectiva, ese mismo año el gobierno de México destinó, aproximadamente, esa misma cantidad en todo su Presupuesto de Cultura [4]. Dicho de otra forma, el ingreso bruto de Denise equivale a lo que todo un país de más de 120 millones de personas destina en su presupuesto cultural, lo que viene siendo en el caso de México, desde 2010, una constante que equivale a 0.2% de PIB, dicho de otra forma, si multiplicamos por 500 el ingreso de Denise tendríamos el equivalente al Producto Interno Bruto del décimo país más poblado del mundo.

En el año 2020 la pandemia obligó a que la mayoría de las competiciones deportivas del mundo fueran suspendidas, generando pérdidas de niveles históricos en el sector, sin embargo, y a pesar de ello el sector de las apuestas en línea ha dado un repunte notable hacia lo que todas las proyecciones concluyen será un crecimiento sostenido en los próximos años. De acuerdo con datos de Statista [5] el tamaño actual del mercado de apuestas en línea es superior a 1,180,000,000,000 de pesos y se espera que, para 2023, el tamaño del mercado casi se duplique. En general, existe un consenso en las proyecciones a futuro que sostienen que habrá un crecimiento en el sector de las apuestas en línea, pero lo más importante para este capítulo de Unbetting Football es que todo parece apuntar a que ese crecimiento se mantendrá **año con año**.

Para abrir tema, necesitamos entender por qué la casa siempre gana y, en especial, tenemos que entender qué hay detrás de las apuestas.

Detrás de las apuestas

En este artículo y en los posteriores se utilizará el formato decimal para las apuestas. En la siguiente tabla se muestra un escenario habitual para un partido de fútbol:

Victoria del Local	Empate	Victoria del Visitante
2	3.18	4

Si una persona apuesta \$100 a que el Equipo Visitante gana el partido, recibiría: $4 \times \$100 = \400 , pero, en caso de que el partido terminase con empate o victoria del equipo local, el apostador perdería su dinero y la casa se lo quedaría. Al número que multiplica a tu dinero se le llama **Momio**, sin embargo, un momio no es solo un multiplicador, pues también representa una probabilidad *implícita*. Para expresar un momio en términos de probabilidad se requiere hacer lo siguiente:

$$\text{Probabilidad} = \frac{1}{\text{Momio}}$$

Que, aplicado al ejemplo anterior, quedaría así:

Victoria del Local	Empate	Victoria del Visitante
2	3.18	4
$\frac{1}{2}$	$\frac{1}{3.18}$	$\frac{1}{4}$
Victoria del Local	Empate	Victoria del Visitante
0.5 (50%)	0.31 (31%)	0.25 (25%)

Ahora, si sumamos los porcentajes obtendremos: $50\% + 31\% + 25\% = 106\%$, por supuesto, esto no es un error. El 6% extra representa la ganancia o *Margen* de la Casa de Apuestas. Una manera sencilla de remover ese margen es dividir entre tres el margen:

$$6\% \div 3 = 2\%$$

Para después, solo restar ese 2% a cada resultado, así:

Victoria del Local	Empate	Victoria del Visitante
50%	31%	25%
-2%	-2%	-2%
Victoria del Local	Empate	Victoria del Visitante
48%	29%	23%

Ahora, por fin, cuando sumamos los porcentajes obtenemos: $48\% + 29\% + 23\% = 100\%$. Pero ¿por qué hacemos esto? Puede parecer poco, sin embargo, este proceso es necesario para hacer Probabilidad como los dioses mandan (recordemos que La Probabilidad debe estar entre 0% y 100%). Con este sencillo proceso podemos limpiar todas las apuestas deportivas del mundo y utilizarlas a nuestro favor. Lo que haremos a continuación será, tomando como base las probabilidades en las apuestas, construir una referencia para hacer **comparaciones** según los escenarios que ocurren en el fútbol. Si en una temporada un equipo anota 84 goles y otro 52, ¿cuál de los dos equipos tuvo un mejor desempeño y en qué medida lo fue? ¿El primer equipo tuvo un mejor desempeño solo por haber anotado más? Y si el primer equipo fuera el Manchester City y el segundo fuera el Wolverhampton, ¿sería lo mismo? Si un árbitro marcara 60 tarjetas amarillas en una temporada, ¿es eso mucho o poco? ¿sería diferente si el árbitro fuera inglés, español o mexicano? Si dos DTs dirigen a dos equipos distintos y en una temporada ambos equipos consiguen 75 puntos, ¿ambos tendrían el mismo desempeño? ¿y si el primer equipo hubiese sido el Real Madrid y el segundo el Betis, ¿sería lo mismo? Por supuesto que no sería lo mismo, por eso necesitamos construir nuestra medida para el fútbol.

La Caja de Pandora

Para elaborar el siguiente estudio y varios de los siguientes que publicaré pedí permiso a Joseph Buchdahl de Football-data.co.uk para emplear sus datos sobre resultados deportivos y apuestas vinculadas a ellos. Se trata de una base de datos de alrededor de 150,000 partidos de 28 ligas europeas, a lo que añadí resultados y datos de la Liga MX. Todos los datos presentados en este capítulo involucran solo resultados de partidos de liga o temporada regular y, para el caso particular de este capítulo, se incluyen resultados de 2012 a 2020. El tratamiento de datos fue hecho con Python 3.9.7.

El 24 de octubre de 2020 ocurrió el partido Barcelona vs Real Madrid y el escenario del partido, según las casas de apuestas, fue el siguiente:

Barcelona	Empate	Real Madrid
48%	26%	26%

En este momento debemos tener cuidado de no perder el sentido de lo que representa una Probabilidad, tengamos en cuenta que la Probabilidad no puede decirnos qué ocurrirá en un evento en específico, más bien:

La Probabilidad es una descripción de lo que tendería a ocurrir en el largo plazo

Continuando con la narración, supongamos que fuese posible que, en octubre de 2020, en ese momento en el tiempo, ese Barcelona y ese Real Madrid decidieran zanjar de una vez por todas quién es el mejor entre esos dos clubes, para lo que jugarían una serie de 100 partidos entre ellos. Si jugasen un partido por semana tardarían poco menos de 2 años en cumplir la centena de partidos. Pero ¿y si observamos qué ocurrió en los otros partidos que tenían el mismo escenario? Por **escenario** me refiero a los partidos que fueron clasificados por las casas de apuestas bajo el siguiente esquema:

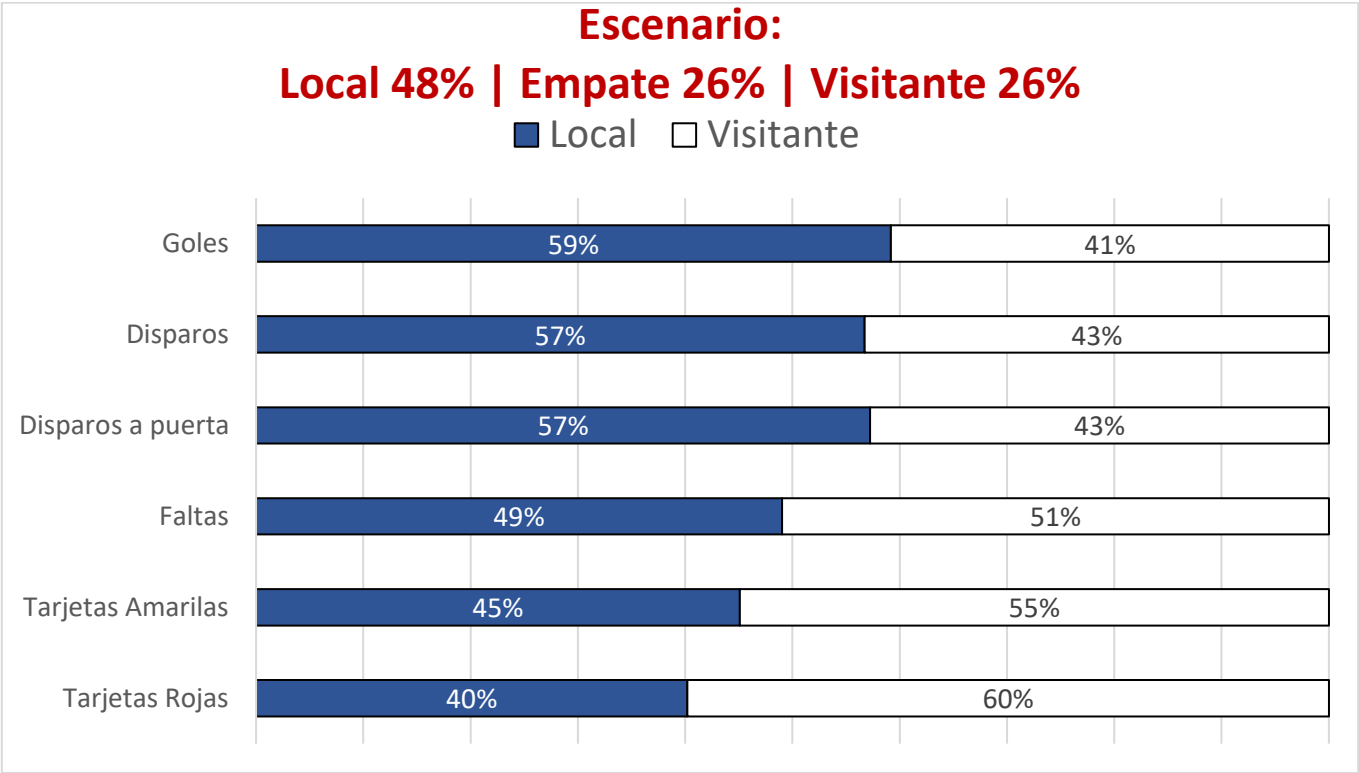
Victoria del Equipo Local	Empate	Victoria del Equipo Visitante
48%	26%	26%

En los siguientes párrafos llamaremos al escenario como Local48%, por practicidad. Una manera de interpretarlo podría ser la siguiente: considerando todos los partidos ocurridos en el periodo 2012 a 2021, de entre esos encuentros, observemos aquellos en los que la estimación de probabilidad cae en el escenario Local48%. Algo que distinguiría a todos esos partidos en ese escenario es que serían **equivalentes a escala**, y esto abre la puerta a varias cosas interesantes. Primero, supongamos que esa selección de partidos resultó ser, exactamente, de 100 encuentros. ¿Cómo se distribuirían los resultados de esos 100 partidos? ¿Cuántas veces ganarían los equipos locales? ¿cuántas ganarían los visitantes o cuántas empatarían? Las cantidades tenderían a ser las siguientes:

Victorias de los Equipos Locales	Partidos Empatados	Victorias de los Equipos Visitantes
48	26	26

Por supuesto, los partidos que caen en el escenario Local48% no son solo 100, en realidad fueron 2075, pero si tomásemos muestras aleatorias de 100 partidos de ese escenario, los promedios de los resultados tenderían a lo expresado en la tabla anterior. Reuniendo esos 2075 partidos se obtuvieron los promedios de las estadísticas registradas. Por ejemplo, se encontró que los equipos locales anotaron 1.55 goles en promedio por partido, y los visitantes promediaron 1.07 goles por partido, pero ¿qué nos dice esto? Una manera de interpretar los resultados podría ser pensarlos en términos de 100 partidos. Si dos equipos se enfrentasen en estas mismas circunstancias, al cabo

de 100 partidos entre ellos esperaríamos que el local tendiese a anotar 155 goles (59% del total), mientras que el visitante tendería a 107 (41% del total). Observemos el resumen de algunos datos presentados en la siguiente gráfica:



Una manera de interpretar el gráfico anterior es que, si pudiésemos extender en el tiempo los partidos, es decir, aumentar la cantidad de partidos entre el Barcelona y el Real Madrid del 24 de octubre de 2020, los resultados tenderían a lo que se muestra en la gráfica. Ahora bien, por supuesto que este tipo de competición (100 partidos seguidos contra el mismo rival) sería impracticable en casi cualquier deporte, en especial en el caso del fútbol en el que incluso una liga con pocos partidos como la Liga MX, donde cada equipo juega 17 partidos en fase regular, puede presentar hasta 17 distintos escenarios para cada equipo. Es necesario tomar un ángulo distinto: pensemos en un caso más simple y, después, veamos si a partir de ahí llegamos a algo interesante (y llegaremos).

Imaginemos que hay una liga de fútbol en la que un equipo, al que llamaremos La Jaiba Brava (saludos a Tampico), se enfrentará solamente a tres rivales a los que solo recibirá como local y, además, deberá enfrentarse 10 veces a cada uno, dando un total de 30 partidos. Cada uno de los enfrentamientos cae en los escenarios planteados a continuación:

La Jaiba Brava	Empate	Rival 1
58%	25%	17%

La Jaiba Brava	Empate	Rival 2
48%	26%	26%

La Jaiba Brava	Empate	Rival 3
38%	29%	33%

Como seguro estarás intuyendo, existe una nutrida muestra disponible para cada uno de esos 3 escenarios (algunos miles, de hecho, entre 2012 y 2021). Para simplificar el análisis inicial, pensemos solo en términos de otra estadística, consideremos primero los **Puntos Promedio** que, para esos 3 escenarios, serían los siguientes:

Escenario	Puntos Promedio del Equipo Local	Puntos Promedio del Equipo Visitante
Local 58%	1.9	0.8
Local 48%	1.7	1.1
Local 38%	1.4	1.3

Si consideramos que en cada escenario se jugarían 10 partidos, podríamos solo **multiplicar por 10 cada uno** de los promedios y sacar sus totales sumándolos, para encontrar los **Puntos Esperados**, como se muestra aquí:

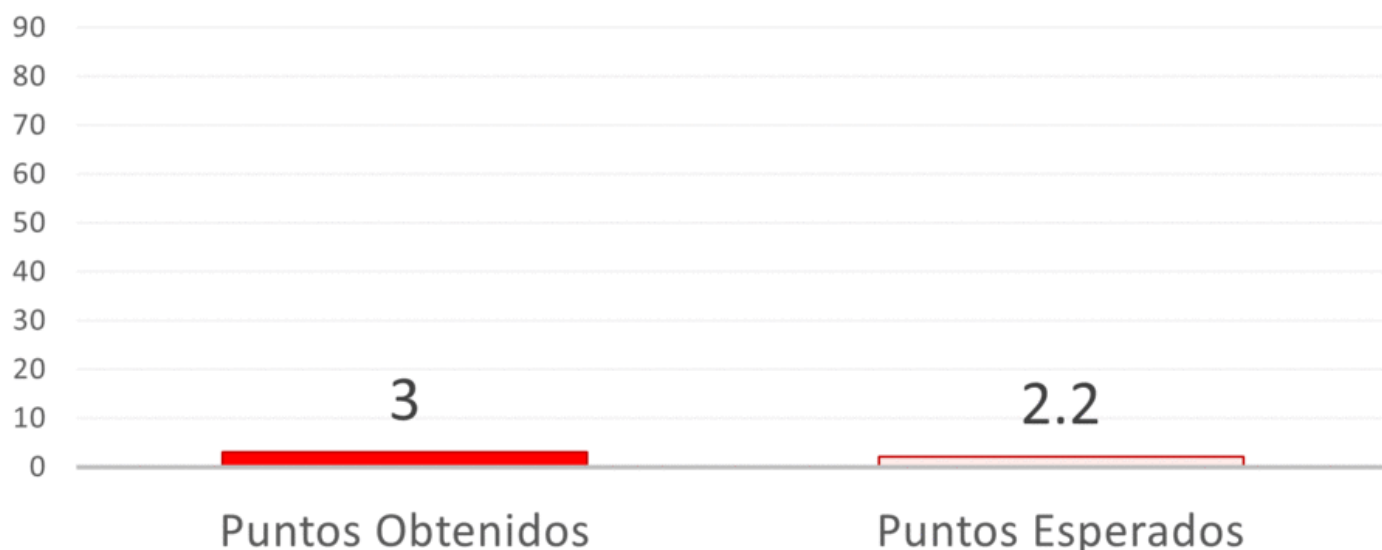
Escenario	Puntos Esperados del Equipo Local	Puntos Esperados del Equipo Visitante
Local 58%	19	8
Local 48%	17	11
Local 38%	14	13
TOTALES	50	32

Cuando llegase el final del torneo La Jaiba Brava podría medir su desempeño en cuanto a los puntos que obtuvo en el torneo mediante una simple división; si el equipo obtuvo, por ejemplo, 55 puntos y se esperaba que obtuviera 50, su desempeño estaría 10% por arriba de lo esperado. Lo mismo se puede hacer para establecer una referencia de desempeño para cualquier otra estadística (por ejemplo, si La Jaiba hubiese cometido un total de 360 faltas y se esperaba que cometiese 300, tendría 20% más faltas de lo esperado). Por supuesto, este sencillo ejercicio: **medir las variaciones respecto a lo esperado**, podemos replicarlo en cualquier liga e, incluso, desde diferentes perspectivas, lo único que necesitaríamos sería completar nuestra anterior tablita para incluir más escenarios y los promedios de los datos que se necesite (goles, disparos, faltas, tarjetas, fueras de juego, etcétera).

Registrando los puntos esperados para cada equipo, jornada a jornada podemos construir el desempeño de los equipos mediante pequeñas sumas y, llegado una cantidad suficiente de partidos (más bien, temporadas completas), empezar a hacer observaciones. Pero mejor veamos algunos ejemplos. Empecemos con el Atlético de Madrid Campeón de **LaLiga en la temporada 2020-2021**, en el que resultaron campeones con la suma de 86 puntos en 38 partidos de liga. Pero ¿es esto mucho o poco? ¿En qué medida cumplieron con su expectativa?



Atlético de Madrid Jornada 1

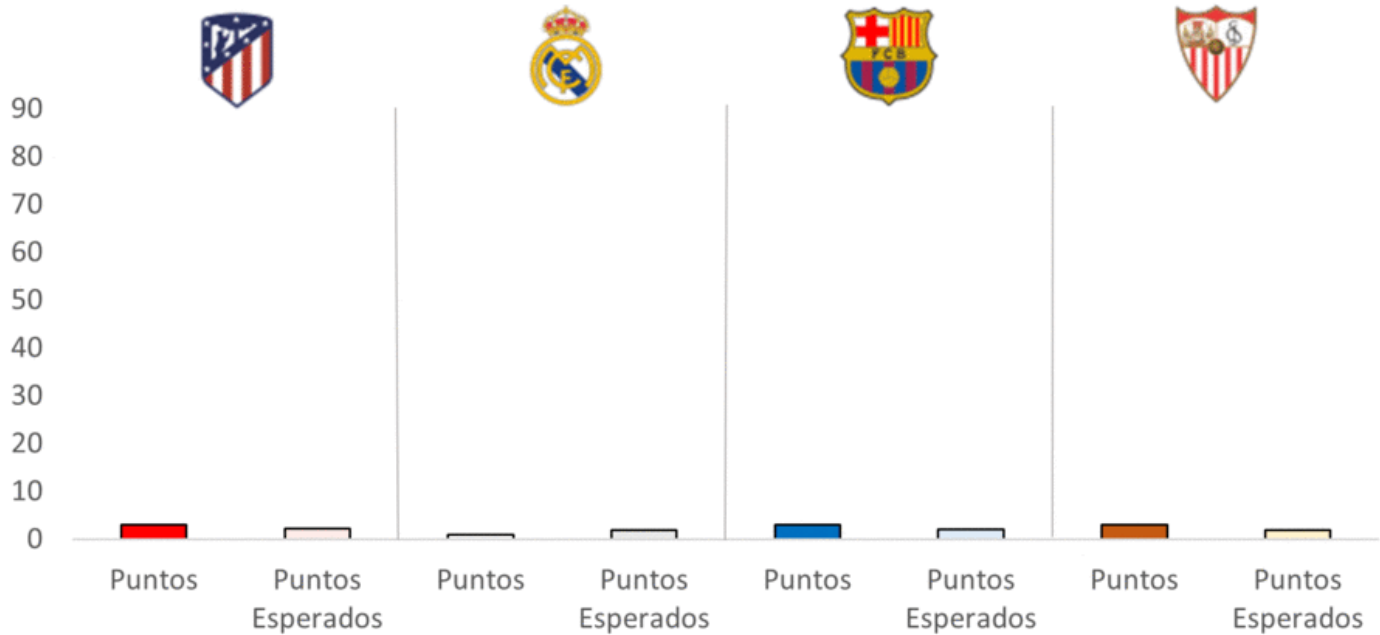


Atlético de Madrid Jornada 38

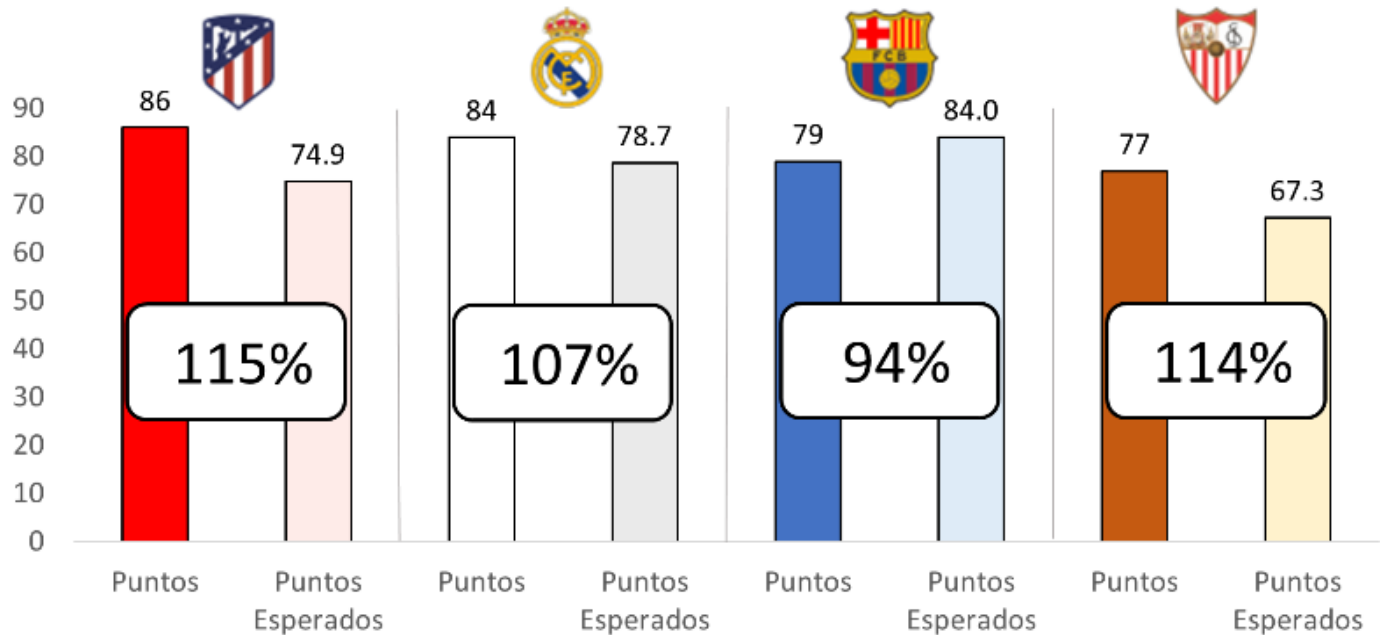


Observando la diferencia entre los Puntos Obtenidos y los Esperados, vemos que el Atlético de Madrid obtuvo 11.1 puntos más de lo que se esperaba, es decir, 15% más de los puntos que se esperaban de ese equipo en ese torneo (86 de 74.9). Sin embargo, si consideramos que a inicio del campeonato los principales favoritos eran Real Madrid y Barcelona, ¿qué tuvo que pasar para que el Atlético de Madrid ganase LaLiga? ¿solo fue mejor el Atlético de lo esperado o hubo una combinación de variaciones en los desempeños? Veamos la siguiente gráfica del Desempeño en Puntos de los primeros cuatro lugares de la temporada 2020-2021 de LaLiga:

Desempeño en Puntos 2020-2021 Jornada 1



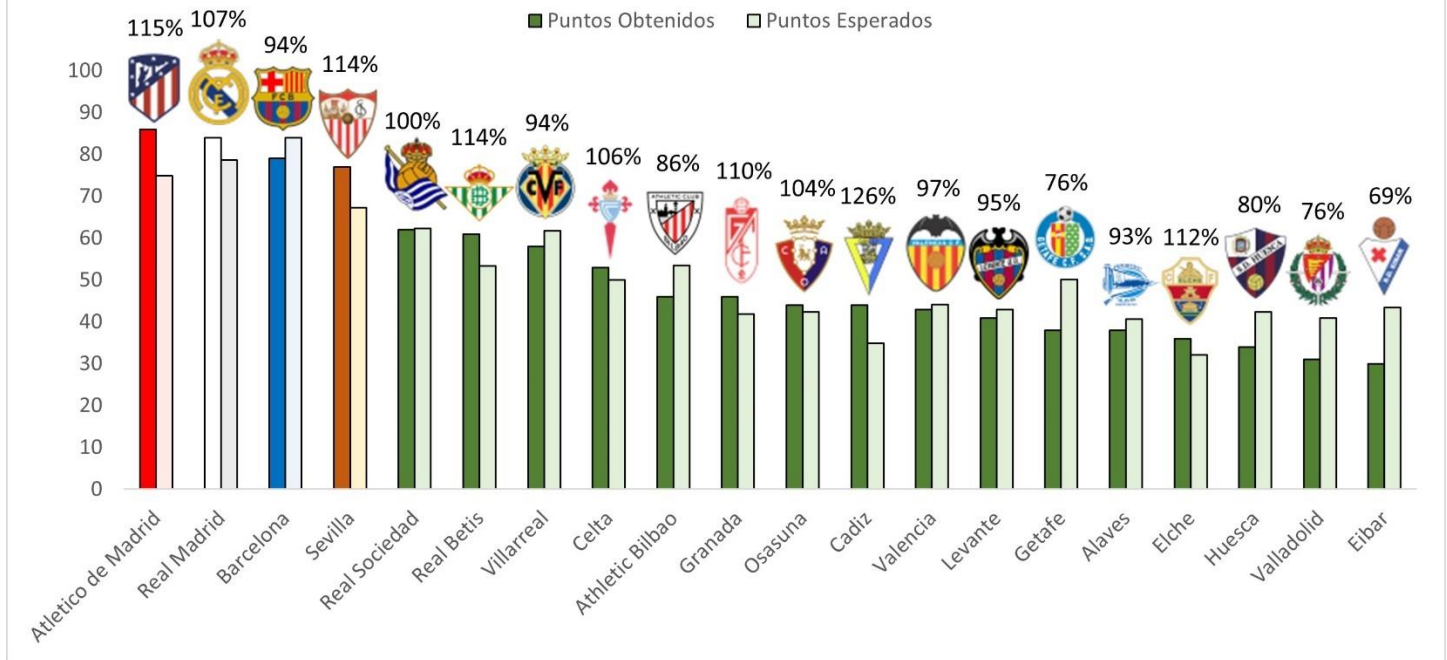
Desempeño en Puntos 2020-2021 Jornada 38



Resulta muy ilustrativo que Sevilla y Atlético de Madrid tuviesen casi el mismo desempeño y, sin embargo, uno fue cuarto lugar y el otro campeón, esto se debe a que **el desempeño depende de lo que se espera de cada equipo en los escenarios a los que se enfrenta**. Veamos la siguiente tabla para LaLiga en la misma temporada 2020-2021 que incluye a todos los equipos.

LaLiga temporada 2020-2021

Desempeño en puntos



Resulta curioso que el equipo con mejor desempeño en el torneo no solo no fue el campeón, sino que estuvo lejos de serlo, Cádiz obtuvo 126% (44 puntos de 35 esperados) alcanzando apenas el 12vo puesto en la tabla general. Mientras que Barcelona, a pesar de tener un desempeño de apenas 94% (79 puntos de 84 esperados), le alcanzó para ser tercer lugar general. Observemos que Getafe, a pesar de tener el segundo peor desempeño del campeonato, logró salvar la categoría.

Las expectativas sobre un equipo dependen del propio equipo (con todos sus actores) y del entorno en el que competirá. La métrica presentada en este capítulo es precisamente eso: una cuantificación de las expectativas y su relación con el entorno. Por ejemplo, solo observando el número de puntos obtenidos por un equipo, ¿en qué medida es sensato exigir al cuerpo técnico de un equipo que obtenga tal o cual cantidad? ¿sería sensato exigir al Osasuna que obtenga 60 puntos, si sus puntos esperados terminan en 43? ¿Podríamos considerar como bueno el torneo de un equipo que solo alcanza 86% de su desempeño (quedándose a más de 7 puntos de lo esperado), como fue el caso del Athletic de Bilbao? ¿o habrá algún factor añadido, como una desventaja, que influya de manera sostenida en el rendimiento del Athletic de Bilbao? En páginas posteriores pretendo hacer un comentario sobre las muestras de equipos con restricciones en sus pools de talento, me basaré en lo que expresa Jared Diamond en su libro *Armas, Gérmenes y Acero* (recomendación al final), pero en el caso del fútbol le llamaré: El Problema del Volumen.

Por otro lado, ¿qué papel juegan los DTs y sus cuerpos técnicos en todo esto? Una pista: el DT no interfiere en el qué (puntos), pero sí en el cómo se llega (es decir, en todo lo demás).

¿Y los árbitros? ¿se podrá saber en qué medida y en qué aspectos los silbantes tienen injerencia en los partidos? ¿cómo reconocer a los mejores árbitros? Una pista: los mejores son los que más pasan desapercibidos (menor diferencia sobre lo esperado).

¿Se podrá usar este tipo de métrica para otros deportes? Una pista: sí, pero el grado y tipo de aplicación depende de las reglas del deporte (necesitamos llegar al capítulo en el que hablaremos sobre La Ergodicidad).

¿Se podrá usar esta métrica para construir una estrategia de apuestas? Una pista: sí.

Mucho que explorar en futuros capítulos.

A manera de conclusión

Las siguientes secciones y capítulos tendrán por objetivo mostrar cómo puede usarse la tabla de valores a la que, por cierto, me tomaré la libertad de llamarla *Metron* (plural de “medida” en griego [6]).

También te contaré la historia del primer libro que leí.

Ing. Jaime G. Meléndez

Twitter: [@JamesGMelendez](https://twitter.com/JamesGMelendez)

Web: <https://academiathi.com/>

LinkedIn: [jaimemelendezthi](https://www.linkedin.com/in/jaimemelendezthi)

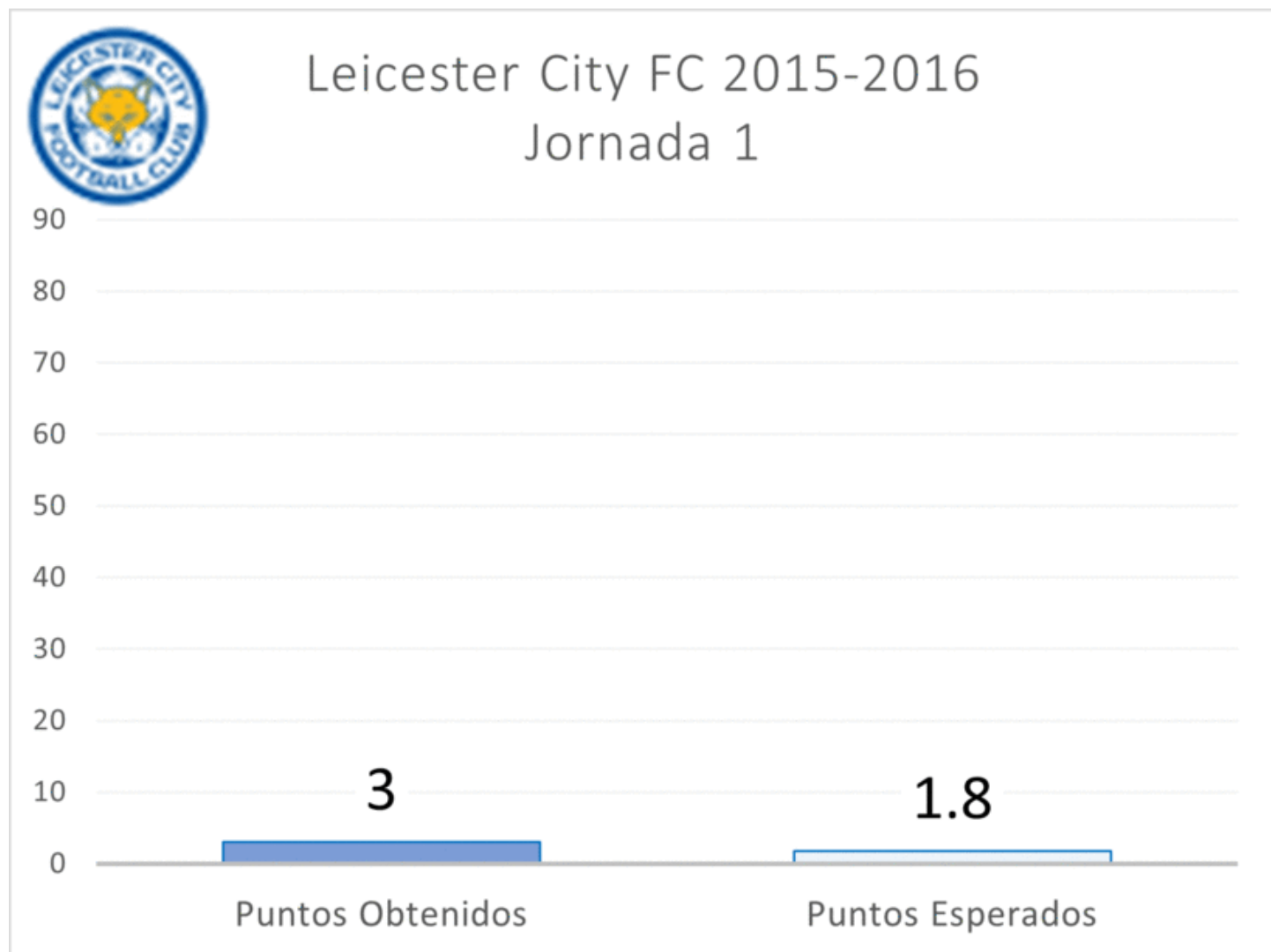
Síguenos en Enrachados Deportes

Podcast: [Enrachados jueves 21:00 horas](#)

Twitter e Instagram: [@Enrachados_deportes](#)

Apéndice

Leicester City FC, EPL 2015-2016. ¿De qué magnitud fue la hazaña del Leicester?





Leicester City FC 2015-2016

Jornada 38

