

Demostración de Algunos Factores que Afectan al Rendimiento de los equipos de Fútbol

Contreras Valeria, Guillén David, Zamora Samuel

Junio del 2024

Índice

Planteamiento del Problema	2
Interrogantes/Objetivos	2
Resumen Descriptivo de la BBDD	4
Goles Realizados por los Equipos Jugando como Local	4
Goles Realizados por los Equipos Jugando como Visitantes	4
Objetivo 1	5
¿Los Equipos Tienden a Ganar más Cuando Juegan en sus Estadios Locales?	9
Distribución de Frecuencias de los Resultados de los Partidos - Athletic Club Bilbao	9
Distribución de Frecuencias de los Resultados de los Partidos - Manchester City	10
Distribución de Densidad de los Resultados de los Partidos de los Equipos - Athletic y Manchester	11
Conclusión del Rendimiento por los Estadios Locales	12
¿Cuál fué el Efecto en la Consistencia Reflejada en la Cantidad de Goles en las Ligas Tras la Implementación del Video Assistant Referee (VAR)?	13
Dispersión de Goles - Bundesliga	13
Dispersión de Goles - Laliga	14
Conclusión Sobre la Consistencia con la Introducción del VAR	15
Bibliografía	16



Figura 1: Escuela de Estadística y Ciencias Actuariales

Planteamiento del Problema

El fútbol es un deporte a nivel profesional alto de once jugadores por equipo en el juego, involucra la cooperación mutua del equipo y se sabe que es un deporte que conlleva el contacto. Estos equipos tienen la misión de trasladar un balón utilizando cualquier parte del cuerpo exceptuando brazos y manos con la finalidad de encajarlo dentro de la portería o arco del conjunto rival, acción que se denomina gol. El equipo con más goles durante un partido de 90 minutos ganará el encuentro.

Pueden existir partidos en los que un equipo sea tácticamente muy superior al rival, o que cuente con jugadores de mejor calidad. Al final del partido puede influir en que dominen en estadísticas como mayor posesión de balón, mayor precisión en los pases, más disparos al arco rival, mayor cantidad de tiros libres y todo lo que pueda favorecer a un equipo para demostrar que es superior, pero lo único que sentencia un partido son los goles, y ganar como en cualquier deporte es la única manera de avanzar en un torneo o competición.

Ya que marcar goles para cualquier equipo es una necesidad, resulta interesante develar qué factores pueden influir en los equipos para que aumente o disminuya su capacidad goleadora.

Se tiene en estudio una base de datos relacionados a algunas ligas del fútbol europeo, entre ellas las 5 más importantes del mundo según la plataforma Global Football Rankings, que utiliza análisis de datos para evaluar y clasificar las ligas de fútbol en todo el mundo, esta plataforma concluye que en este top se encuentran la Premier League (Inglaterra), La Bundesliga (Alemania), LaLiga (España), Serie A (Italia) y Ligue 1 (Francia).

La base de datos cuenta con variables como los goles anotados por los equipos de cada liga tanto de local como de visitante, a lo largo de las temporadas que transcurrieron desde el año 2006 hasta el año 2021.

El objetivo del presente trabajo estadístico es exponer y demostrar algunos factores determinantes que tienen influencia en la cantidad de goles y victorias que se pueden llevar un equipo partido a partido. Todo esto en base a la secuencia de partidos de las ligas y divisiones que están registradas en la BBDD.

Trabajaremos con unas muestras de la BBDD la cual consta principalmente del equipo “Athletic Club” de Bilbao que juega en la primera división de “La Liga”, el campeonato nacional de España. Mientras avanzamos con el desarrollo del trabajo, también se expondrán otros equipos aleatorios a las pruebas del objetivo junto al principal anteriormente mencionado.

Interrogantes/Objetivos

- 1) ¿Los equipos tienden a ganar más cuando juegan en sus estadios locales?
- 2) ¿Influyeron los partidos sin público durante la pandemia en la capacidad de anotar goles por parte de los equipos que jugaron de visitante?
- 3) ¿Cuál fué el efecto en la consistencia reflejada en la cantidad de goles en las ligas tras la implementación del Video Assistant Referee (VAR)?

- 4) ¿Puede distinguir un cambio de entrenador y/o jugadores del equipo para un dominio del marcador en su mayoría de juegos?

Resumen Descriptivo de la BBDD

A continuación, un cuadro con detalles de la estadística descriptiva para la evaluación de las variables cuantitativas de la base de datos en estudio.

Variable	Total de datos	Moda	Media	Desv-Std	Min	Q1	Med	Q3	Max	CV
Goles Local	123404	1	1.48	1.25	0	1	1	2	10	84.61
Goles Visitante	123404	1	1.14	1.11	0	0	1	2	13	97.14

Goles Realizados por los Equipos Jugando como Local

- Hay un total de 123.404 datos o registros.
- La moda (número que más se repite) es 1 gol.
- El promedio de goles anotados de los equipos jugando como local es de 1,48.
- La desviación estándar nos dice que hay una dispersión de los datos de 1,25 aproximadamente.
- El valor mínimo presentado en goles anotados de los locales es de 0.
- Dado que el cuartil 1 es de 1, nos dice que el 25% de los datos son menores o igual a 1 gol por parte de los equipos locales.
- Dado que la mediana es de 1, nos dice que el 50% de los datos son menores o igual a 1 gol por parte de los equipos locales.
- Dado que el cuartil 3 es de 2, nos dice que el 25% de los datos son mayores a 2 goles por parte de los equipos locales.
- El valor máximo presentado en goles anotados de los locales es de 10.
- Obtenemos un coeficiente de variación de 84,61%, lo que nos dice que los goles locales presentan una variabilidad alta.

Goles Realizados por los Equipos Jugando como Visitantes

- Hay un total de 123.404 datos o registros.
- La moda (número que más se repite) es 1 gol.
- El promedio de goles anotados de los equipos jugando como visitante es de 1,14.
- La desviación estándar nos dice que hay una dispersión de los datos de 1,11 aproximadamente.
- El valor mínimo presentado en goles anotados de los visitantes es de 0.
- Dado que el cuartil 1 es de 0, nos dice que el 25% de los datos son menores o igual a 0 goles por parte de los equipos visitantes.
- Dado que la mediana es de 1, nos dice que el 50% de los datos son mayores o igual a 1 gol por parte de los equipos visitantes.
- Dado que el cuartil 3 es de 2, nos dice que el 25% de los datos son mayores o iguales a 2 goles por parte de los equipos visitantes.
- El valor máximo presentado en goles anotados de los visitantes es de 13.
- Obtenemos un coeficiente de variación de 97,14%, lo que nos dice que los goles locales presentan una variabilidad alta.

Objetivo 1

```
par(mfcol = c(1, 2))

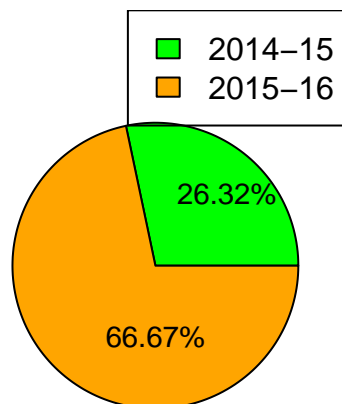
pie(x = c(Burnley_15$Porcentaje_Por_Partido, Burnley_16$Porcentaje_Por_Partido),
    labels = c("", ""),
    col = c("green", "orange"),
    main = "% Puntos Obtenidos (Burnley)")
text(0.4, 0.4, paste0(Burnley_15$Porcentaje_Por_Partido, "%"), col = "black", cex = 0.9)
text(0.0, -0.4, paste0(Burnley_16$Porcentaje_Por_Partido, "%"), col = "black", cex = 0.9)

legend("topright", legend = c("2014-15", "2015-16"), fill = c("green", "orange"))

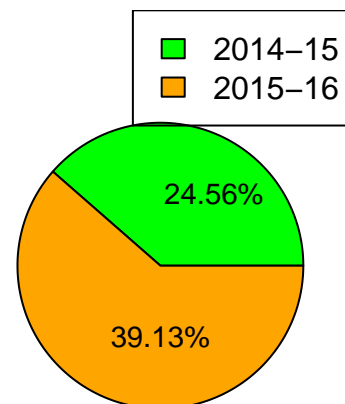
pie(x = c(QPR_15$Porcentaje_Por_Partido, QPR_16$Porcentaje_Por_Partido),
    labels = c("", ""),
    col = c("green", "orange"),
    main = "% Puntos Obtenidos (QPR)")
text(0.3, 0.4, paste0(QPR_15$Porcentaje_Por_Partido, "%"), col = "black", cex = 0.9)
text(0.0, -0.4, paste0(QPR_16$Porcentaje_Por_Partido, "%"), col = "black", cex = 0.9)

legend("topright", legend = c("2014-15", "2015-16"), fill = c("green", "orange"))
```

% Puntos Obtenidos (Burnley)



% Puntos Obtenidos (QPR)



```

pie(x = c(Hull_15$Porcentaje_Por_Partido, Hull_16$Porcentaje_Por_Partido),
    labels = c("", ""),
    col = c("green", "orange"),
    main = "% Puntos Obtenidos (Hull)")
text(0.3, 0.4, paste0(Hull_15$Porcentaje_Por_Partido, "%"), col = "black", cex = 0.9)
text(0.0, -0.4, paste0(Hull_16$Porcentaje_Por_Partido, "%"), col = "black", cex = 0.9)

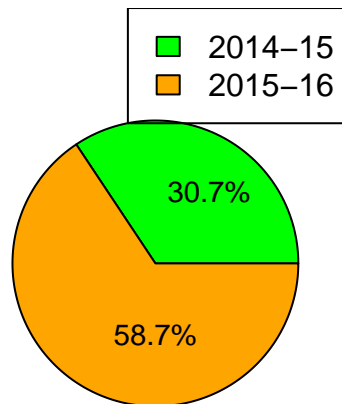
legend("topright", legend = c("2014-15", "2015-16"), fill = c("green", "orange"))

pie(x = c(Bournemouth_15$Porcentaje_Por_Partido, Bournemouth_16$Porcentaje_Por_Partido),
    labels = c("", " "),
    col = c("green", "orange"),
    main = "% Puntos Obtenidos (Bournemouth)")
text(-0.3, 0.2, paste0(Bournemouth_15$Porcentaje_Por_Partido, "%"), col = "black", cex = 0.9)
text(0.3, -0.4, paste0(Bournemouth_16$Porcentaje_Por_Partido, "%"), col = "black", cex = 0.9)

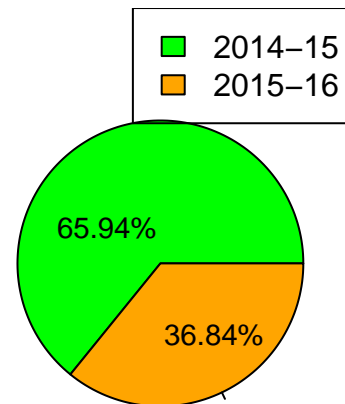
legend("topright", legend = c("2014-15", "2015-16"), fill = c("green", "orange"))

```

% Puntos Obtenidos (Hull)



% Puntos Obtenidos (Bournemouth)



```

pie(x = c(Norwich_15$Porcentaje_Por_Partido, Norwich_16$Porcentaje_Por_Partido),
    labels = c("", " "),
    col = c("green", "orange"),
    main = "% Puntos Obtenidos (Norwich)")
text(-0.3, 0.2, paste0(Norwich_15$Porcentaje_Por_Partido, "%"), col = "black", cex = 0.9)
text(0.4, -0.2, paste0(Norwich_16$Porcentaje_Por_Partido, "%"), col = "black", cex = 0.9)

```

```

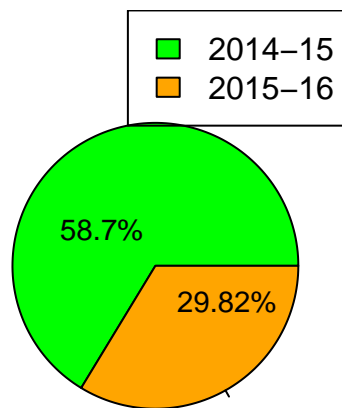
legend("topright", legend = c("2014-15", "2015-16"), fill = c("green", "orange"))

pie(x = c(Watford_15$Porcentaje_Por_Partido, Watford_16$Porcentaje_Por_Partido),
    labels = c("", " "),
    col = c("green", "orange"),
    main = "% Puntos Obtenidos (Watford)")
text(0.0, 0.3, paste0(Watford_15$Porcentaje_Por_Partido, "%"), col = "black", cex = 0.9)
text(0.3, -0.4, paste0(Watford_16$Porcentaje_Por_Partido, "%"), col = "black", cex = 0.9)

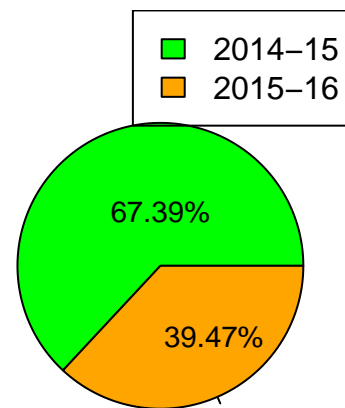
legend("topright", legend = c("2014-15", "2015-16"), fill = c("green", "orange"))

```

% Puntos Obtenidos (Norwich)



% Puntos Obtenidos (Watford)



```

par(mfcol = c(1, 1))

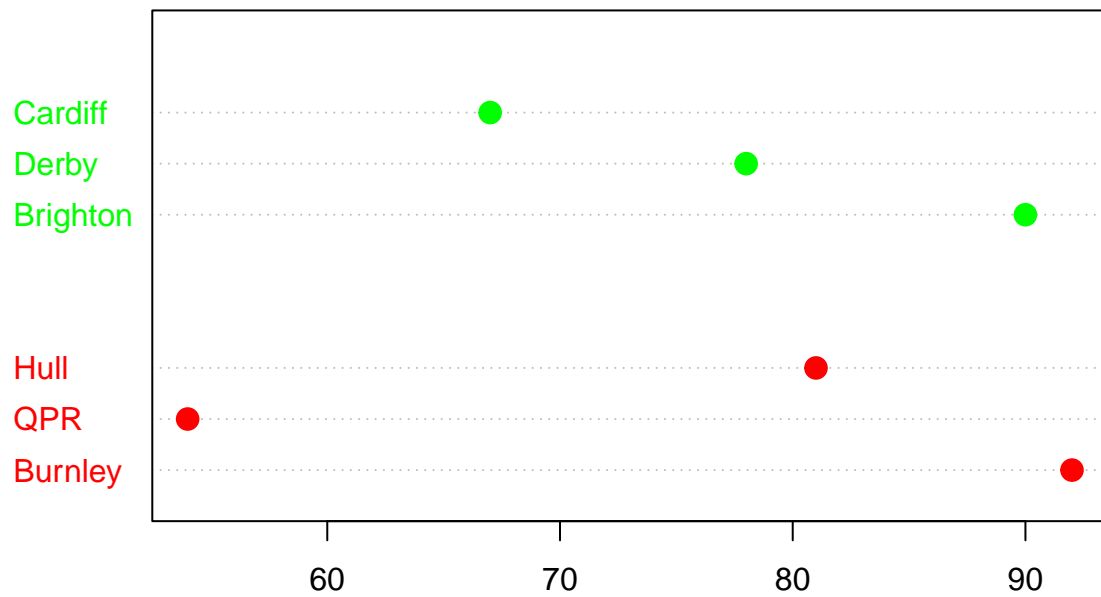
```

```

# Crear un vector de colores
colores <- rep("black", length(Equipos))
colores[VS == 1] <- "red"
colores[VS == 2] <- "green"

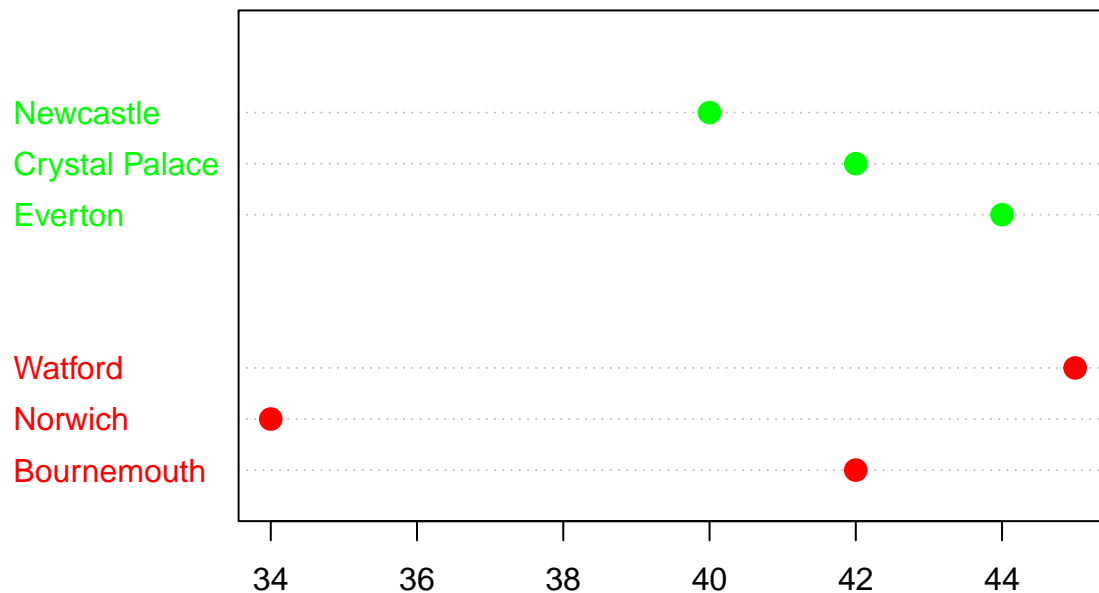
# Crear el gráfico de puntos
dotchart(Datos1$Total_Puntos, labels = Datos1$Equipos, pch = 19,
         pt.cex = 1.5, groups = rev(Datos1$VS), color = colores)

```



```
# Crear un vector de colores
colores <- rep("black", length(Equipos))
colores[VS == 1] <- "red"
colores[VS == 2] <- "green"

# Crear el gráfico de puntos
dotchart(Datos2$Total_Puntos, labels = Datos2$Equipos, pch = 19,
         pt.cex = 1.5, groups = rev(Datos2$VS), color = colores)
```

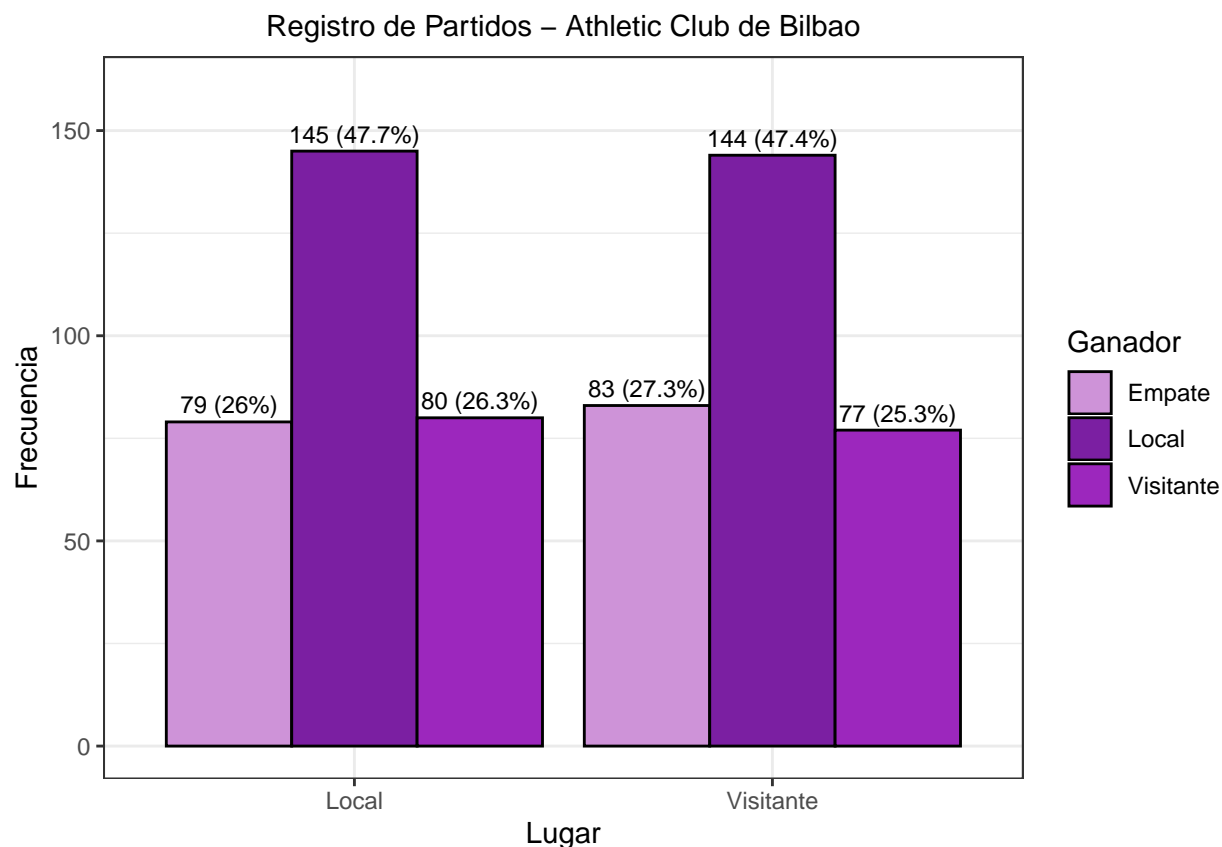
¿Los Equipos Tienden a Ganar más Cuando Juegan en sus Estadios Locales?

Se procede a evaluar los resultados de todos los juegos de la BBDD donde “Athletic Club” y el “Manchester City” juegan de local y serán comparados con los resultados donde juegan de visitante, de esta manera se visualizará en que posición se han llevado más victorias.

También es importante mencionar que son todos sus partidos en sus ligas desde la temporada 2005-06 hasta la 2020-21.

Distribución de Frecuencias de los Resultados de los Partidos - Athletic Club Bilbao

A continuación un histograma de frecuencias que nos indican los resultados de los partidos totales del Athletic Club en La Liga.



Podemos observar en la gráfica que Athletic jugando de local:

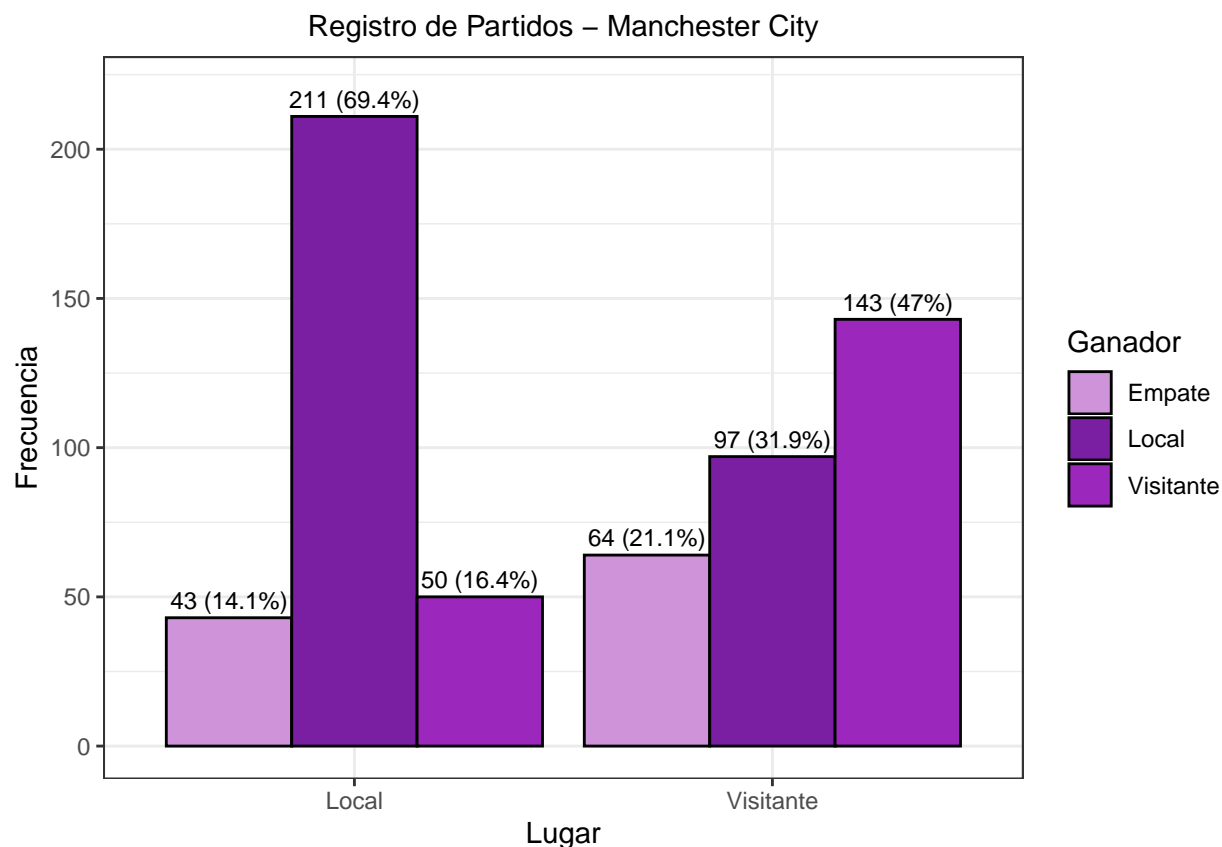
- Ha ganado el 47,7% de sus partidos.
- Ha perdido el 26,3% de sus partidos.
- Ha empatado el 26% de sus partidos.

Por otra parte, Athletic jugando como visitante:

- Ha ganado el 25,3% de sus partidos.
- Ha perdido el 47,4% de sus partidos.
- Ha empatado el 27,3% de sus partidos.

Distribución de Frecuencias de los Resultados de los Partidos - Manchester City

A continuación un histograma de frecuencias que nos indican los resultados de los partidos totales del Manchester City en la Premier Ligue.



Podemos observar en la gráfica que el Manchester City jugando de local:

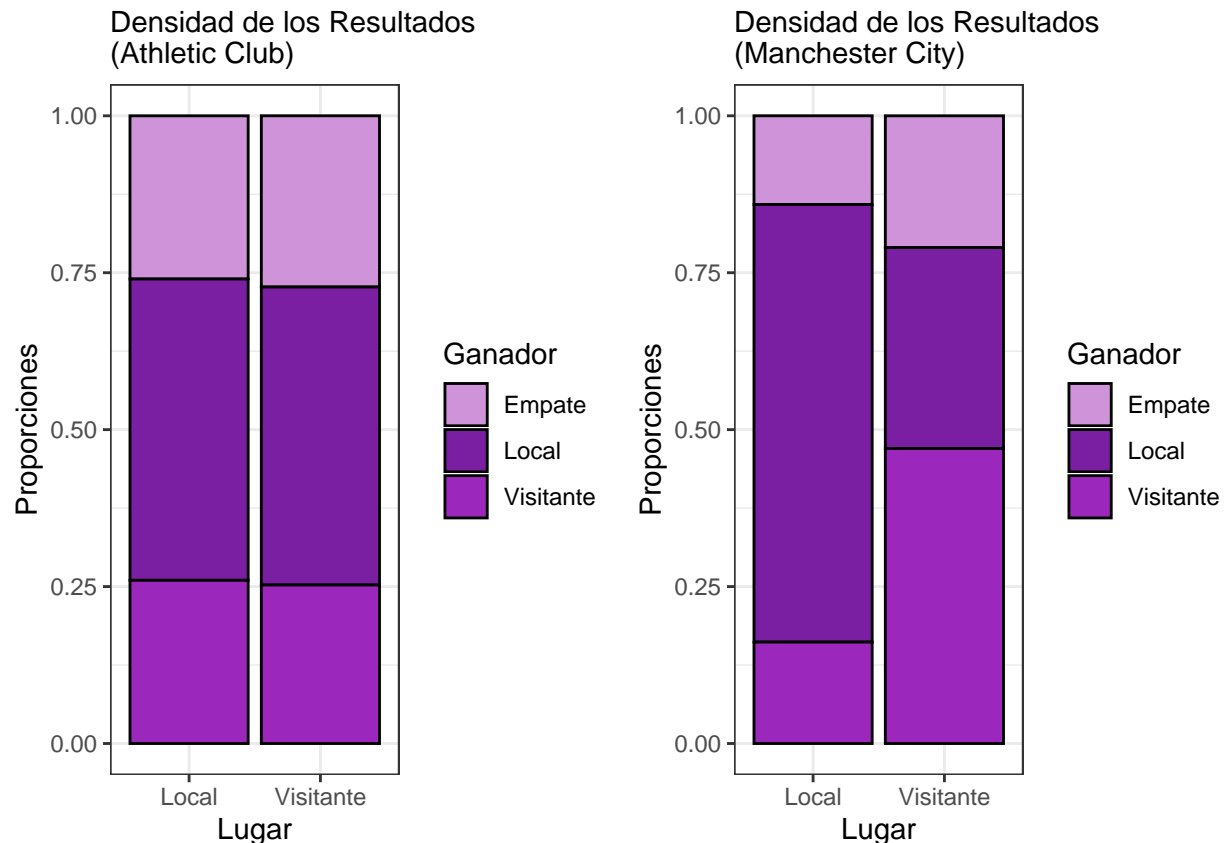
- Ha ganado el 69,4% de sus partidos.
- Ha perdido el 16,5% de sus partidos.
- Ha empatado el 14,1% de sus partidos.

Por otra parte, el Manchester City jugando como visitante:

- Ha ganado el 47% de sus partidos.
- Ha perdido el 31,9% de sus partidos.
- Ha empatado el 21,1% de sus partidos.

Distribución de Densidad de los Resultados de los Partidos de los Equipos - Athletic y Manchester

Adentrándonos en estas frecuencias y porcentajes, podemos estudiar su composición de proporciones, lo que nos ayuda a medir en una misma escala todo, y notar a simple vista lo que ocurre en comparación entre los equipos.



Se puede apreciar que el Athletic Club tiene un mejor rendimiento si juega como local, con una mayor proporción de victorias acumuladas delante las otras entre las temporadas 2005-06 a la la 2020-21.

Al igual que el equipo anterior, el Manchester City tiene una mayor tasa de victorias jugando como local, pero también se aprecia que a su vez ha ganado la mayor parte de la proporción de partidos jugando como visitante. Entonces se puede decir que es un destacable equipo de la premier league, pero tiene un mejor rendimiento jugando de local.

Como dato, el que un equipo tenga muchas victorias jugando de local puede tener más detalles a considerar, como el sentido de comodidad por jugar en su “hogar” del equipo, una mayor cantidad de afición apoyandolos en los asientos a comparación del equipo rival, entre otros.

Conclusión del Rendimiento por los Estadios Locales

Si, los equipos jugando de manera local suelen tener una ventaja de victoria delante su rival. Presentan una mayor proporción de ganar delante los otros dos posibles resultados (empatar o perder) en la muestra, y por regularidad estadística, si se toma una muestra de un espacio muestral, se va a comportar igual al espacio muestra.

Se respalda esta información con el artículo que hizo Fatima Juan Belmonte, una reportera de prensa deportiva.

Mediante hechos con respecto a un equipo español, demostró la importancia de la afición a su equipo, la fuente de motivación que son para el mismo e incluso su relevancia a la hora de nuevos fichajes.

También Carlos Cueva, profesor de Economía de la Universidad de Alicante, realizó un estudio sobre 41 ligas (antes y después de que comenzó la pandemia), y comprobó que la afición local genera una presión al equipo visitante, recopilando resultados, se notó que sin la afición de ninguna de las partes, los equipos visitantes aumenta su tasa de victoria hasta un 12% por temporada, pero nunca es mayor a la tasa de victorias local.

¿Cuál fué el Efecto en la Consistencia Reflejada en la Cantidad de Goles en las Ligas Tras la Implementación del Video Assistant Referee (VAR)?

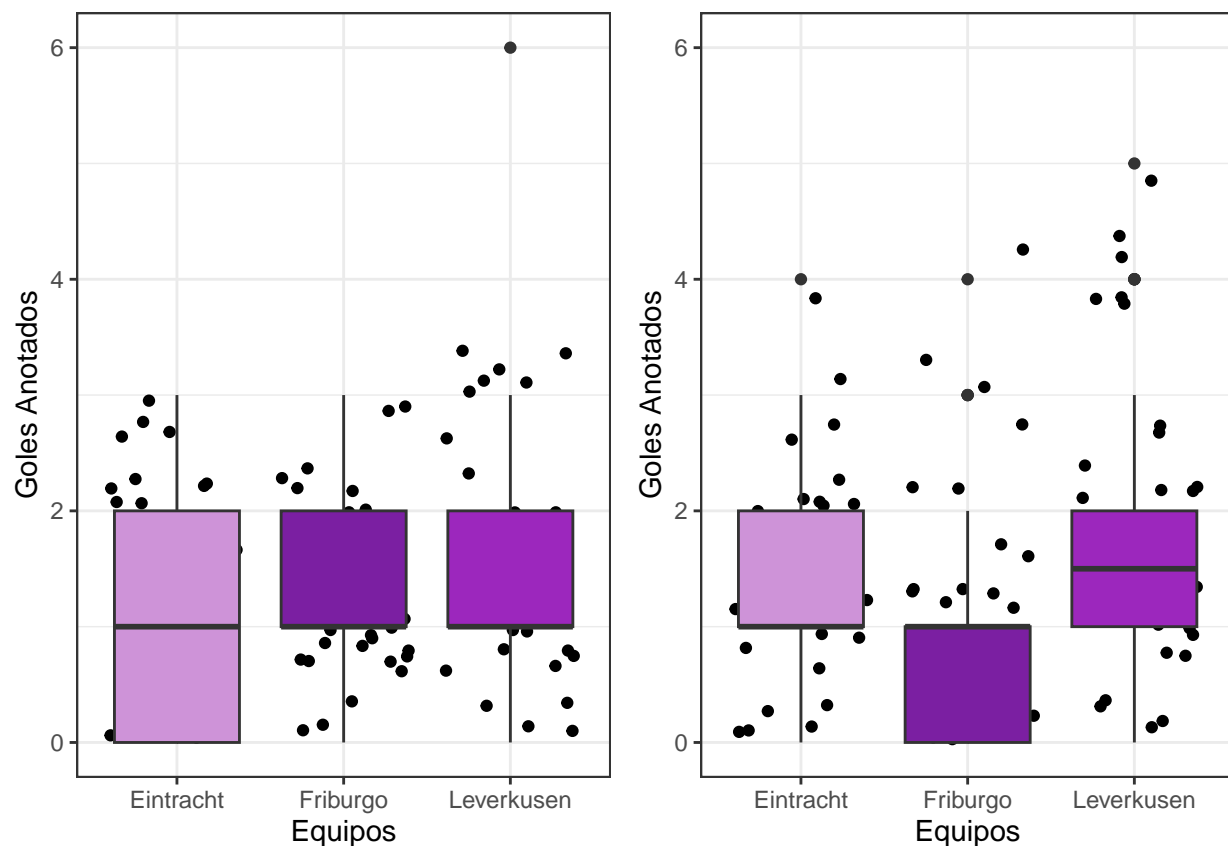
Se procede a evaluar los goles anotados por 3 equipos de las ligas “Bundesliga” y “Laliga”, y comparar sus cambios una temporada antes y después la implementación oficial del VAR en sus respectivas divisiones de 1era. Un dato importante antes de continuar, es que se tomaron 2 equipo regulares de las ligas y 1 equipo por trayectoria muy bueno de cada liga.

Los equipos a evaluar de la Bundesliga en las temporadas 2016-17 y 2017-18 son: Eintracht Fráncfort, Bayer 04 Leverkusen, SC Friburgo.

Los equipos a evaluar de Laliga en las temporadas 2017-18 y 2018-19 son: FC Barcelona, Sevilla FC, Athletic Club.

Dispersión de Goles - Bundesliga

A continuación gráficos de caja y bigote que ayuda a visualizar la dispersión de los goles anotados por los equipos de la Bundesliga en estudio.



Equipo	Min	Q1	Med	Q3	Max	Media(antes)	Desv-Std(antes)	Media(después)	Desv-Std(después)
Eintracht	0	0	1	2	3	1.06	1.10	1.32	1.04
Friburgo	0	1	1	2	3	1.24	0.82	0.94	1.10
Leverkusen	0	1	1	2	6	1.56	1.31	1.71	1.45

Específicamente hablando de la dispersión de los datos, independientemente de su mejoría o desmejoría a nivel de goles. Notemos que en estos equipos hubieron 2 de 3 (Friburgo y el Leverkusen) que aumentaron su desviación estándar con respecto a su media.

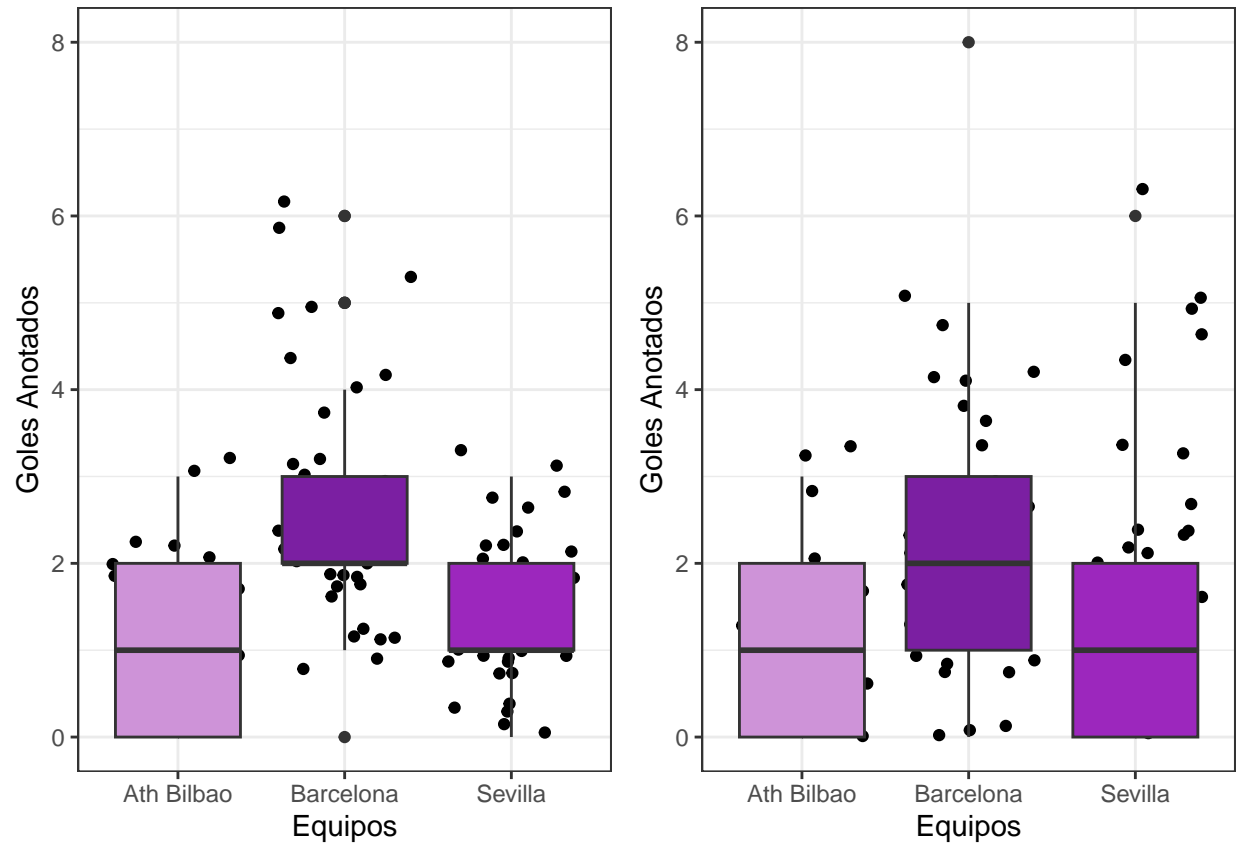
- En el caso de Friburgo en pocas palabras, su concentración de datos del 50% bajó dos unidades, pasó a encontrarse entre los dos más bajos (0 y 1 goles). En definitiva empeoró su capacidad goleadora.
- En el caso de Leverkusen, su concentración de datos del 50% se mantuvo entre 1 y 2 goles, pero su mediana aumentó ligeramente.

Con respecto al otro equipo (Eintracht), apenas y hubo una muy ligera disminución de su desviación estándar con respecto a su media.

Lo curioso de este caso es que 2 de los 3 equipos (Eintracht y el Leverkusen) evaluados en términos de goles han aumentado (en goles) su rango de concentración de datos del 50% y su media, se dice que después de la inclusión del VAR les fue mejor anotando goles “válidos”.
 Por otra parte el otro equipo (Friburgo) se vió afectado y disminuyó estas estadísticas, tras la inclusión del VAR le fue peor anotando goles “válidos”.

Dispersión de Goles - Laliga

A continuación gráficos de caja y bigote que ayuda a visualizar la dispersión de los goles anotados por los equipos de Laliga en estudio.



Equipo	Min	Q1	Med	Q3	Max	Media(antes)	Desv-Std(antes)	Media(después)	Desv-Std(después)
Ath Bilbao	0	0	1	2	3	1.08	0.91	1.08	0.91

Barcelona	0	2	2	3	6	2.61	1.44	2.37	1.60
Sevilla	0	1	1	2	3	1.29	0.98	1.63	1.65

Podemos notar que un equipo (Ath Bilbao) se mantuvo por completo igual, la realidad tras esto, es que apenas y cambió su orden univariante. Mientras que los otros 2 equipos (Barcelona y el Sevilla) aumentaron su rango de concentración de datos una unidad menos, y sus medias por separado, disminuyeron para uno y aumentó para otro. Aquí es donde entra la dispersión nos da respuestas claras.

Analizando la dispersión de los datos, independientemente de su mejoría o desmejoría a nivel de goles. Estos dos equipos (Barcelona y el Sevilla) sí tuvieron tras la inclusión del VAR un aumento de su desviación estándar.

- En el Sevilla su concentración del 50% de los datos ahora se encuentra en entre los 2 valores más bajos, pero tiene una cantidad considerable de datos lejanos positivamente a esta concentración, por eso su media y desviación estándar aumentaron.
- El Barcelona por aumentar su rango de concentración de los datos del 50% una unidad hacia abajo, hizo que su media disminuyera, además, de por sí su media era muy alta. Pero su desviación se ve superada porque tiene un dato atípico muy alto en 8 goles.

Conclusión Sobre la Consistencia con la Introducción del VAR

Casi todos los equipos vieron un aumento de su desviación típica con respecto a su media hablando de goles, es un fenómeno que viene de mano a que en la mayoría de los casos su rango del 50% central de los datos disminuyó una unidad. Esto nos dice que los equipos que solían anotar “muchos” goles en el marcador bajaron esa cantidad.

Por otro lado equipos que no son de anotar muchos goles mantuvieron esa dispersión con respecto a su media goleadora, lo que nos dice que no se vieron afectados por ese lado, o en el mejor o peor de sus casos estabilizaron ese rango de concentración del 50% de goles que suelen anotar a uno más o menos.

Bibliografía

- Belmonte, F. (2017, noviembre 19). La importancia de la afición. VAVEL. Recuperado de <https://www.vavel.com/es/futbol/2017/11/19/rayo-vallecano/849777-la-importancia-de-la-aficion.html>
- Guerrero, S. (2021, febrero 18). ¿Cuál es la influencia de las hinchadas locales en el fútbol?. PRIMICIAS. Recuperado de <https://www.primicias.ec/noticias/jugada/peso-influencia-hinchada-equipos-locales-estadios-futbol/>
- CAMBIOS DE ENTRENADORES DURANTE LA TEMPORADA. (s.f.). transfer markt. Recuperado de https://www.transfermarkt.es/laliga/trainerwechsel/wettbewerb/ES1/plus/?saison_id=2018
- CAMBIOS DE ENTRENADORES DURANTE LA TEMPORADA. (s.f.). transfer markt. Recuperado de https://www.transfermarkt.es/laliga/trainerwechsel/wettbewerb/ES1/plus/?saison_id=2017
- IMPACTO DE ENTRENADOR. (s.f.). transfer markt. Recuperado de https://www.transfermarkt.es/laliga/trainereffekt/wettbewerb/ES1/plus/?saison_id=2018&spiele=15
- IMPACTO DE ENTRENADOR. (s.f.). transfer markt. Recuperado de https://www.transfermarkt.es/laliga/trainereffekt/wettbewerb/ES1/plus/?saison_id=2017&spiele=15
- EQUIPO ATHLETIC CLUB - TEMPORADA 18/19. (s.f.). transfer markt. Recuperado de https://www.transfermarkt.es/athletic-club/kader/verein/621/plus/0/galerie/0?saison_id=2018
- EQUIPO ATHLETIC CLUB - TEMPORADA 17/18. (s.f.). transfer markt. Recuperado de https://www.transfermarkt.es/athletic-club/kader/verein/621/plus/0/galerie/0?saison_id=2017
- EQUIPO RC CELTA DE VIGO - TEMPORADA 18/19. (s.f.). transfer markt. Recuperado de https://www.transfermarkt.es/rc-celta-de-vigo/kader/verein/940/plus/0/galerie/0?saison_id=2018
- EQUIPO RC CELTA DE VIGO - TEMPORADA 17/18. (s.f.). transfer markt. Recuperado de https://www.transfermarkt.es/rc-celta-de-vigo/kader/verein/940/plus/0/galerie/0?saison_id=2017