

# ANÁLISIS ESTADÍSTICO DEL FC BARCELONA EN LAS TEMPORADAS 2020-2024

Jose Aguirre y Zadquiel Nieves

2025-10-22

## Planteamiento de Problema

El FC Barcelona es una de las instituciones deportivas más importantes del mundo. Su rendimiento en La Liga depende de muchos factores, desde el 2020 hasta 2024, el equipo ha cambiado mucho, se ha pasado de buscar entender qué pasa, a ver qué funciona mejor.

## Justificación del Problema

El fútbol actual depende cada vez más del análisis de datos, para un club como el Barcelona, con altas expectativas y un estilo de juego definido, usar estadísticas básicas es muy importante para:

- Detectar fortalezas y debilidades en diferentes situaciones, como partidos en casa o fuera.
- Mejorar la alineación, las estrategias según los rivales y las circunstancias.
- Medir cómo influyen los jugadores importantes en la creación de juego y en marcar goles, también defender bien.
- De cambio Entender el rendimiento en un momento tras la era Messi y en medio de problemas económicos.
- Este análisis ayudará a entender cómo ha evolucionado el equipo más allá de las opiniones personales, usando datos reales de La Liga.

## Objetivo General

Evaluar la evolución táctica y de rendimiento ofensivo del Fútbol Club Barcelona durante las temporadas 2020 a 2024, mediante el análisis estadístico y la visualización de métricas avanzadas como la Posesión (poss), Goles Esperados (xg), Goles esperados en contra (xga), Goles (gf), Goles en contra (ga), Resultado de los partidos en casa y afuera más Formaciones Tácticas. Y ver si esto ayudó al FC Barcelona a levantarse en una de sus peores temporadas en la historia (2020-2021)

## Objetivos Específicos

- Cuantificar la tendencia anual de las métricas clave de rendimiento (Goles a Favor, Goles en Contra, y Porcentaje de Posesión) para determinar la dirección de la evolución del modelo de juego del equipo.
- Establecer la relación de correlación (fuerza y dirección) entre el Porcentaje de Posesión y la Eficiencia de Oportunidades (poss, xg, xga) a nivel de partido, viendo los resultados por temporada.
- Determinar si existe una diferencia estadísticamente significativa en el Puntos por temporadas obtenidos por el club en función de las Formaciones Tácticas (formation) más utilizadas en cada temporada.

## Breve Historia

El FC Barcelona cambió mucho entre 2020 y 2024, después de que Messi se fue en 2021, el club pasó por estos cambios:

- Cambio en el entrenador: De Ronald Koeman a Xavi Hernández, si hubo cambios a nivel de juego.
- Promoción de jóvenes: Llegaron Pedri, Gavi, Lamine Yamal y fichajes como Lewandowski, Raphinha y Koundé entre otros.
- Cambio en la estrategia: Dejaron de usar solo el 3-5-2 y probaron formaciones más variadas como el 4-2-3-1 y el 4-3-3 con doble pivote.
- Resultados: Ganaron mucho en La Liga, pero tuvieron problemas en la Champions League en algunos momentos (todavía).

Este período explica los datos que vemos, Se trata de cambios y adaptaciones constantes.

## Marco Metodológico

### Fuente de Datos y Tiempo

Usamos el conjunto de datos completo `matches_full_la_liga.csv`. Este conjunto tiene registros de partidos de La Liga desde 2020-2021 hasta 2024-2025. Tiene 4,319 filas. Es una base de datos que muestra la competencia del fútbol español en ese periodo. Las variables incluyen datos cuantitativos y cualitativos. Algunas variables importantes son:

- Resultados: `gf` (goles a favor), `ga` (goles en contra), `result` (resultado)
- Rendimiento: `poss` (posesión), `sh` (tiros), `sot` (tiros a puerta), `xg` (goles esperados), `xga` (Goles esperado en contra) `dist` (distancia en total recorrida en el partido)
- Variables de contexto: `formación` (formación), `attendance` (asistencia), `venue` (localía)

### Variables y Análisis

El estudio tiene cuatro partes principales, cada una ayuda a entender el rendimiento del equipo: Rendimiento en Casa y Fuera: Comparamos el equipo cuando juega en su campo y cuando juega afuera, goles y goles en contra, esto muestra si el equipo actúa diferente según el lugar de juego. Posesión y Goles: Evaluamos la relación entre cuánto controla el balón, las oportunidades y los goles. Esto nos ayuda a entender si la posesión lleva a más goles esperados y si el equipo es eficiente en ataque. Formaciones y Tácticas: Analizamos cómo cambia la estrategia del equipo, revisión de formaciones y rotaciones del plantel en cada temporada, esto revela cómo evoluciona el estilo de juego.

## Infraestructura y Herramientas

Usamos tecnología para analizar los datos. Combinamos estadísticas con visualizaciones:

R, Rmarkdown: Aquí hicimos nuestro PDF. Respondimos las preguntas del planteamiento del problema. También presentamos nuestras investigaciones sobre la evolución del Barcelona de 2020 a 2024.

Python, Pandas, Plotly, Matplotlib: Desarrollamos una aplicación web usando Streamlit. La interfaz permite filtrar datos y ver patrones. También muestra gráficos interactivos relacionados con nuestro trabajo y nuestras investigaciones.

## Proceso de Trabajo

El trabajo tiene **cuatro fases principales**:

- **Fase 1: Preparación de datos:** Filtración y limpieza de los registros del FC Barcelona para el período 2020-2024.
- **Fase 2: Cálculos estadísticos descriptivos:** Análisis de tendencias, medidas de centralidad, dispersión, y distribución de variables clave.
- **Fase 3: Visualización avanzada:** Creación de gráficos y tablas para representar los hallazgos estadísticos.
- **Fase final: Conclusiones:** Síntesis de resultados y respuesta a los objetivos planteados.

Este método garantiza resultados confiables y útiles. Proporciona información clara para nosotros y también quienes toman decisiones en el club, desde técnicos hasta directivos.

Table 1: Tabla 1: Rendimiento General y Puntos por Temporada

season	partidos_jugados	victorias	empates	derrotas	puntos	efectividad
2020	38	25	7	6	82	71.9
2021	38	24	7	7	79	69.3
2022	38	21	10	7	73	64.0
2023	38	28	4	6	88	77.2
2024	38	26	7	5	85	74.6

Interpretación Tabla 1: El análisis del rendimiento general muestra cambios en la competencia entre 2020 y 2024. El equipo mejoró su capacidad para obtener puntos. La eficiencia aumentó del 64% en 2022 al 77,2% en 2023. La temporada 2023 fue la más exitosa. Se lograron 28 victorias y solo 6 derrotas en ese año. Además, se alcanzó un máximo de 88 puntos. Esto indica que el equipo pudo convertir su rendimiento en resultados en los últimos años. Esto ocurrió después de un período con menos consistencia.

Table 2: Tabla 2: Estadísticas Descriptivas de Goles y Posesión por Temporada

season	Prom. GF	SD GF	Prom. GC	SD GC	Dif. Goles	Prom. Poss (%)	SD Poss	promedio_xg	promedio_xga
2020	2.26	1.69	1.00	0.93	48	66.39	7.76	1.74	1.02
2021	2.24	1.55	1.00	0.93	47	65.53	8.49	1.93	1.08
2022	1.79	1.40	1.00	0.81	30	64.16	8.04	1.67	0.91
2023	1.84	1.33	0.53	0.89	50	64.26	7.47	1.99	0.87
2024	2.08	1.17	1.16	1.37	35	64.29	7.23	2.04	1.10

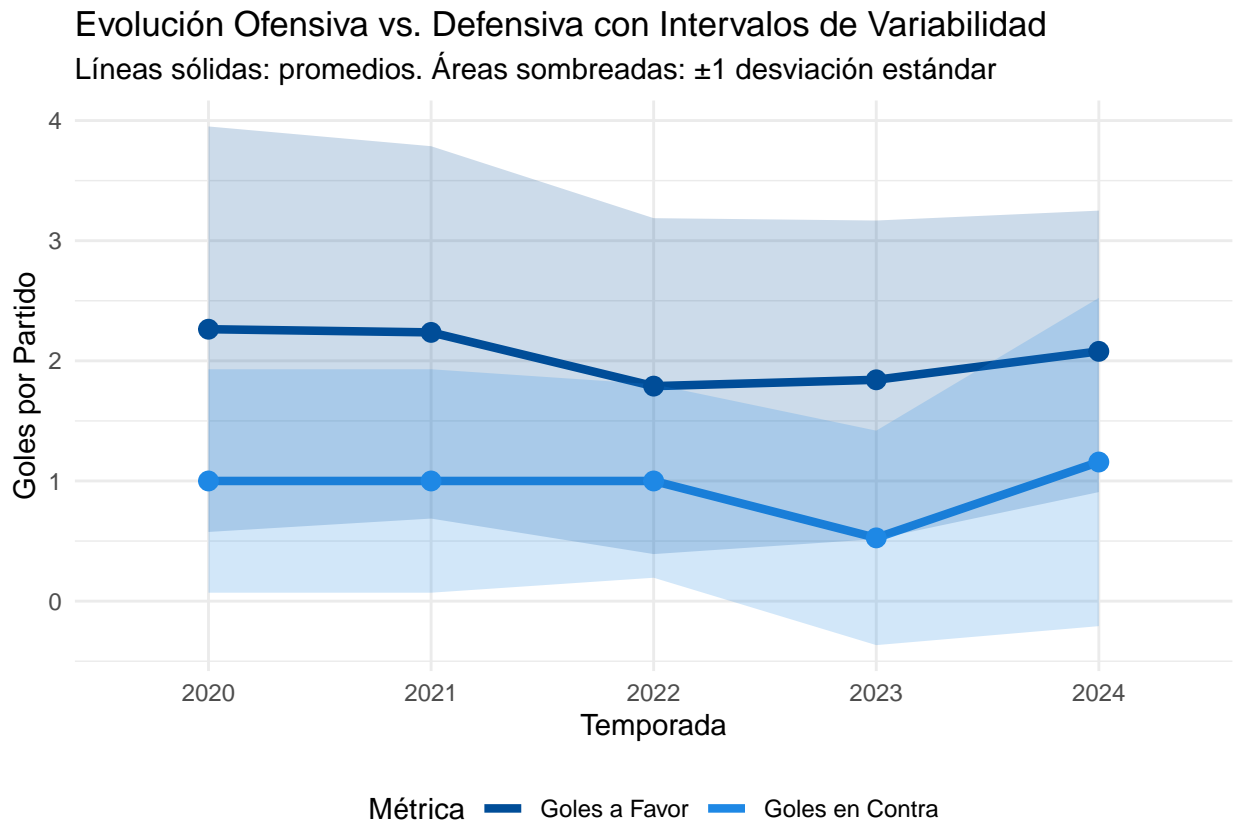
Interpretación Tabla 2: Los datos confirman una fuerte adherencia al modelo de juego de control durante todo el período, con el promedio de posesión (Prom. Poss) manteniéndose constantemente alto, en un rango estrecho de 64.16% a 66.39% y una baja desviación estándar (SD). No obstante, la métrica defensiva mostró la mejora más dramática y eficiente en la temporada 2023, donde el promedio de Goles en Contra (Prom. GC) se redujo a 0.53, resultando en la mayor diferencia de goles (+50) para el equipo en las cinco temporadas. Si bien la producción ofensiva (Prom. GF) fue más alta en 2020 y 2021, la combinación de una defensa sólida y una ofensiva eficiente en 2023 fue la clave del éxito en puntos.

Table 3: Tabla 3: Estadísticas Descriptivas Globales (2020-2024)

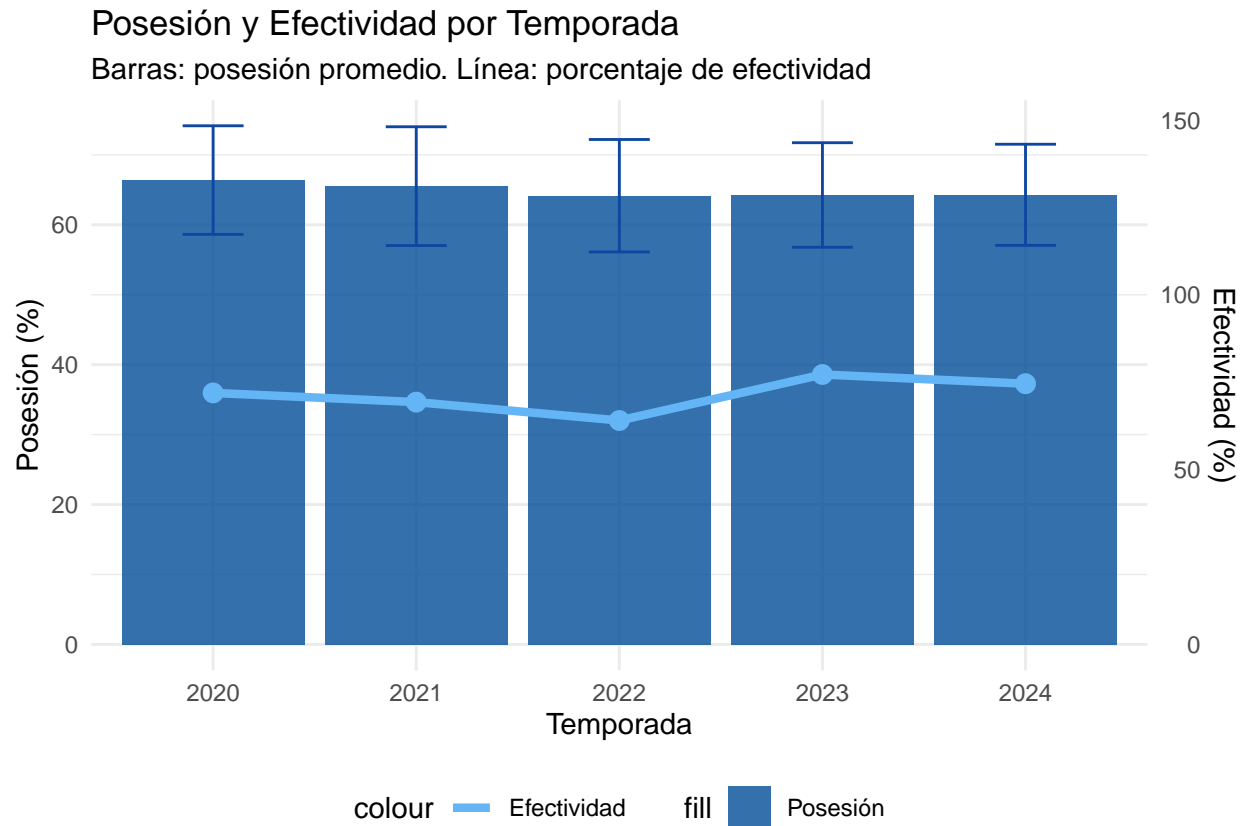
Métrica	Media	Mediana	SD	Mínimo	Máximo
Goles a Favor	2.04	2	1.44	0	6
Goles en Contra	0.94	1	1.02	0	5

Métrica	Media	Mediana	SD	Mínimo	Máximo
Posesión	64.93	65	7.78	45	82

Interpretación Tabla 3: A nivel global del período, el FC Barcelona mantuvo un promedio de posesión muy alto del 64.93%. El promedio de Goles a Favor (GF) fue de 2.04 por partido. La desviación estándar (SD) de GF fue de 1.44. Esto significa que la cantidad de goles en ofensiva varió entre partidos. El menor número de goles fue 0 y el mayor fue 6. El promedio de Goles en Contra (GC) fue de 0.94. Esto indica que la defensa fue fuerte. Algunos partidos tuvieron hasta 5 goles recibidos.

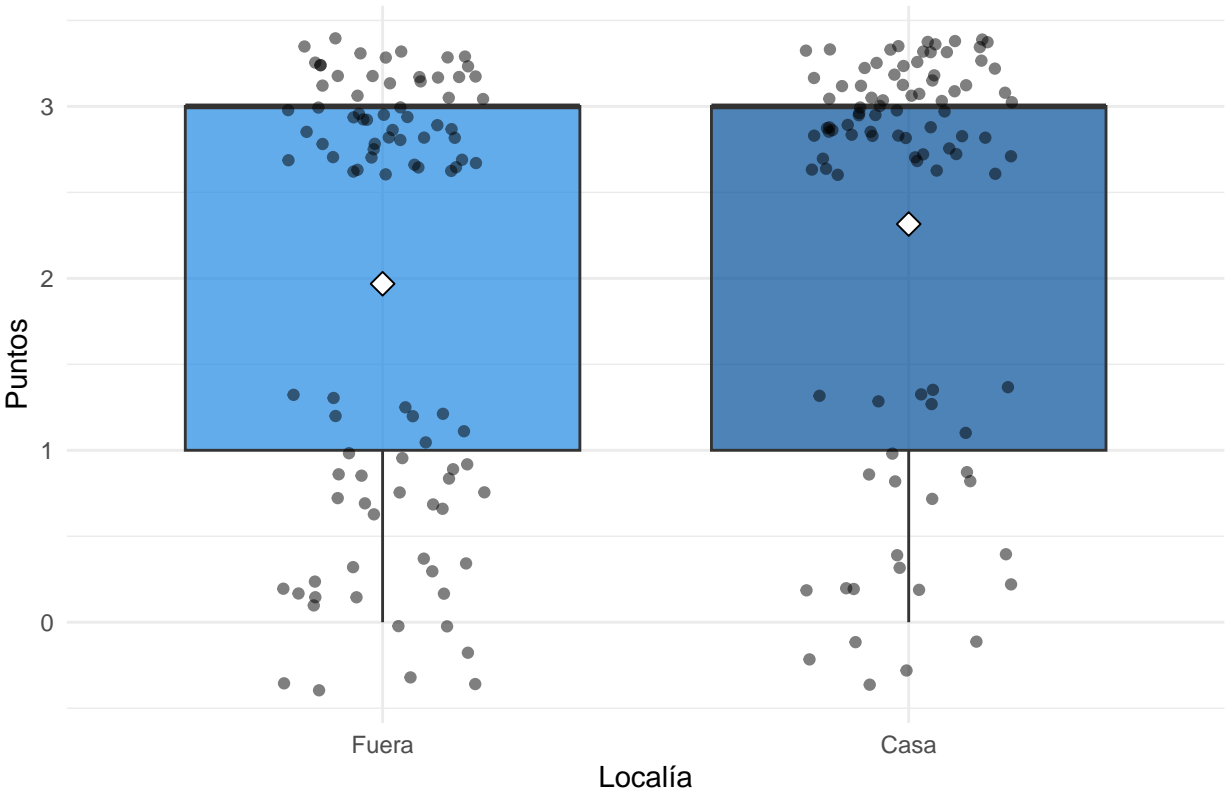


Interpretación Gráfico 1: Este gráfico muestra el rendimiento del equipo. La línea de Goles a Favor se mantuvo cerca de 2 goles por partido. En 2022, esta cantidad bajó ligeramente. La línea de Goles en Contra mejoró con el tiempo. Después de una temporada, se mantuvo en un nivel más bajo. Las áreas sombreadas alrededor de esta línea se hicieron más estrechas entre 2022 y 2024. Esto significa que la defensa del equipo fue más constante en las últimas temporadas.

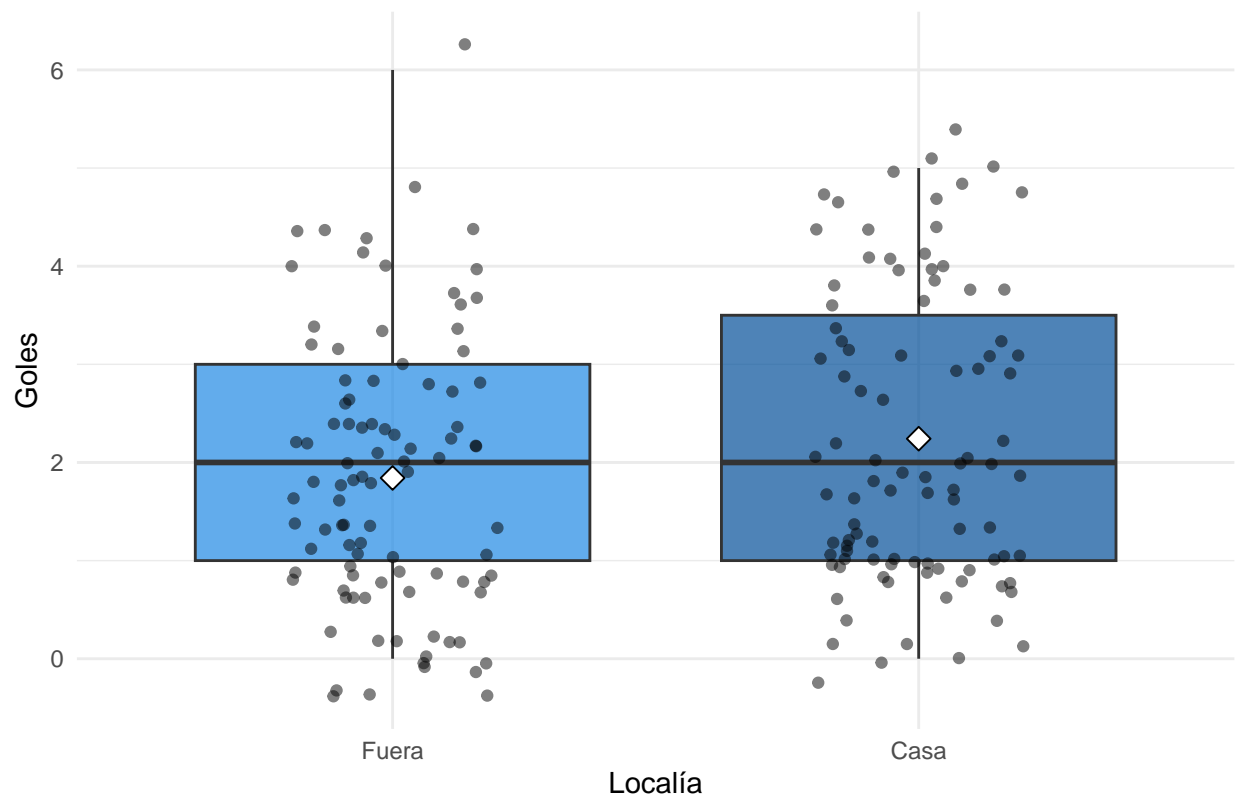


Interpretación Gráfico 2: El gráfico muestra que tener la posesión del balón no indica cuán efectivo es un equipo para anotar puntos. La barra de posesión se mantuvo alta y estable, entre 64% y 67%. La línea de efectividad tuvo cambios durante la temporada. En 2023, el equipo tuvo su mejor rendimiento. En ese año, lograron la máxima cantidad de puntos por partido. La posesión del balón en 2023 fue similar a la de otras temporadas. Esto indica que los mejores resultados no dependieron solo del control del balón. En cambio, mejoraron en cómo usaron esa posesión para anotar. La eficiencia en las jugadas y en finalizar las acciones fue mayor en 2023. Por lo tanto, la clave para mejorar fue la estrategia y la ejecución, no solo tener más posesión.

Distribución de Puntos por Partido según Localía



### Distribución de Goles por Partido según Localía

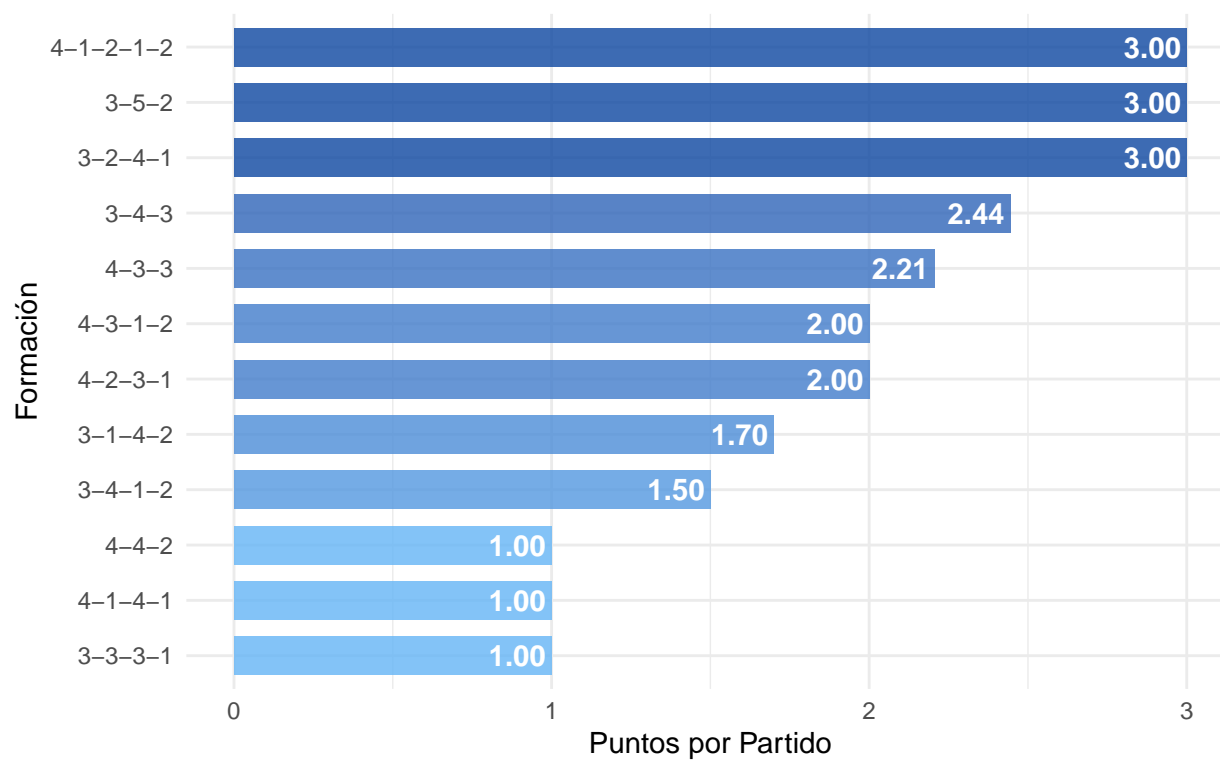


Interpretación Gráfico 3 (Puntos): El boxplot muestra que la mediana de puntos por partido es 3 en casa y fuera. Esto indica que la mayoría de los resultados en ambos casos son victorias. La dispersión de los resultados es mayor cuando el equipo juega fuera. Como visitante, el equipo obtiene más resultados de 1 punto, que son empates. Además, hay más partidos con 0 puntos, que son derrotas o bajo rendimiento.

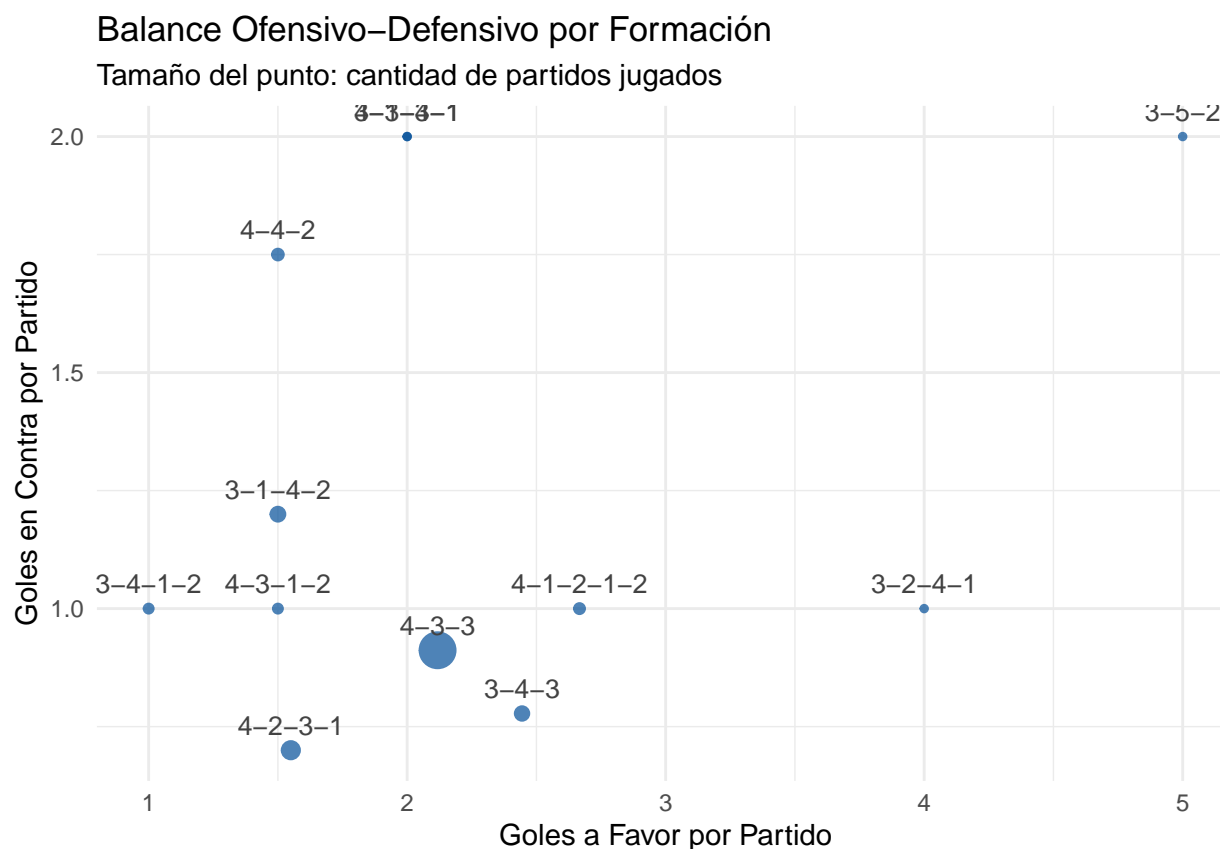
Interpretación Gráfico 4 (Goles): El rendimiento ofensivo del Barcelona es mejor cuando juega en casa. La cantidad de goles por partido tiene una mediana similar en ambos casos, aproximadamente 2 goles. El boxplot de “Casa” es más compacto. Esto muestra que el equipo tiene una producción goleadora más constante en casa. Fuera de casa, los resultados varían más. Hay más partidos con poca o ninguna producción ofensiva. Esto indica que jugar en casa ayuda a que el ataque sea más eficiente.

## Efectividad de Formaciones Tácticas

Puntos promedio por partido por formación (mínimo 5 partidos)







Interpretación Gráfico 5: Este gráfico muestra la eficiencia de las formaciones en puntos por partido. Algunas formaciones poco comunes, como el 4-1-2-1-2, el 3-5-2 y el 3-2-4-1, lograron un promedio de 3.00 puntos por partido. La formación 4-3-3 es la más usada. Tiene un promedio de 2.21 puntos por partido. Esta formación es la más frecuente entre las formaciones habituales. La elección sugiere que el cuerpo técnico busca un equilibrio en los resultados a largo plazo. Las formaciones con tres defensas, como el 3-5-2 y el 3-4-3, tienen un promedio de 2.44 puntos por partido. Son menos efectivas que un sistema con cuatro defensas, que es más confiable para obtener puntos.

Interpretación Gráfico 6: El Balance Ofensivo-Defensivo revela la calidad de cada táctica. El 4-3-3 se destaca como la formación más popular y equilibrada, con un alto número de goles a favor y pocos en contra. La concentración de las formaciones más utilizadas en la zona de alto rendimiento sugiere que el equipo técnico está aprendiendo y ajustando constantemente su estrategia, eligiendo esquemas que realmente marcan la diferencia en el juego. Además, el 4-2-3-1 también se presenta como una opción sólida en términos de defensa.

Table 4: Tabla 4: Análisis Comparativo de Formaciones Tácticas

Formación	Partidos	Frecuencia (%)	Puntos/Partido	GF/Partido	GC/Partido	Posesión (%)
3-2-4-1	1	0.53	3.00	4.00	1.00	58.00
3-5-2	1	0.53	3.00	5.00	2.00	80.00
4-1-2-1-2	3	1.58	3.00	2.67	1.00	67.33
3-4-3	9	4.74	2.44	2.44	0.78	69.00
4-3-3	136	71.58	2.21	2.12	0.91	64.57
4-2-3-1	20	10.53	2.00	1.55	0.70	66.05
4-3-1-2	2	1.05	2.00	1.50	1.00	73.00
3-1-4-2	10	5.26	1.70	1.50	1.20	66.00

Formación	Partidos	Frecuencia (%)	Puntos/Partido	GF/Partido	GC/Partido	Posesión (%)
3-4-1-2	2	1.05	1.50	1.00	1.00	51.00
3-3-3-1	1	0.53	1.00	2.00	2.00	61.00
4-1-4-1	1	0.53	1.00	2.00	2.00	72.00
4-4-2	4	2.11	1.00	1.50	1.75	58.00

Table 5: Tabla 5: Conclusiones del Análisis Estadístico Descriptivo

Objetivo	Hallazgo Principal	Valor	Implicación
Evolución del modelo de juego	Mantenimiento del modelo de posesión alto	64.9 %	Confirmación identidad de juego
Relación posesión-eficiencia	Posesión no garantiza mejor creación	0.09	Validación modelo táctico
Efectividad de formaciones	Formaciones con alta efectividad identificadas	3	Base decisiones técnicas
Rendimiento por localía	Rendimiento similar en ambos escenarios	0.35	Mejorar rendimiento visitante
Consistencia ofensiva	Irregularidad en producción ofensiva	1.43	Estabilidad en rendimiento