

ANÁLISIS DEL RENDIMIENTO HISTÓRICO DEL REAL MADRID EN LA CHAMPIONS LEAGUE

Introducción

El tema de nuestro Exploratory Data Analysis (EDA) se centra en el fútbol, específicamente en la historia de la Champions League y el desempeño del Real Madrid, el equipo con más títulos en la historia de esta competición. Nos propusimos verificar si su éxito en cuanto a trofeos se corresponde también con el mejor rendimiento histórico en términos de métricas clave.

Para salirnos del molde, hemos estructurado el análisis de manera clara y progresiva, explorando múltiples hipótesis (manteniendo la hipótesis principal y cambiando otras hipótesis como si estuviésemos iterando para conseguir nuestros objetivos) y aspectos del torneo.

El primer paso fue preparar las herramientas necesarias para trabajar con los datos de manera eficiente.

- Se recopilaron datos de todas las temporadas de la Champions League desde 1955/1956 hasta 2023/2024.
- Se incluyeron las columnas Round, Date, Team 1, Team 2, Team 1 Goals, Team 2 Goals y Season.
- Los datos incluyen todas las rondas principales: Octavos, Cuartos, Semifinal y Final.
- El scraping respetó la estructura HTML de la página para asegurar la precisión en la extracción de los datos. Fuente de datos: www.bdfutbol.com
- El resultado arrojó 8240 filas en 7 columnas.

Dataset cargado con éxito: 8240 filas y 7 columnas

Dimensiones del dataset: (8240, 7)

Primeras 5 filas del dataset:

	Round	Date	Team 1	Team 1 Goals	Team 2 Goals	\
0	Octavos Ida	NaN	Sporting de Portugal	3	3	
1	Octavos Ida	NaN	Vörös Lobogó	6	3	
2	Octavos Ida	NaN	Servette	0	2	
3	Octavos Ida	NaN	Rot-Weiss Essen	0	4	
4	Octavos Ida	NaN	Djurgården	0	0	

	Team 2	Season
0	Partizan	1955-56
1	Anderlecht	1955-56
2	Real Madrid	1955-56
3	Hibernian	1955-56
4	Gwardia Warszawa	1955-56

Información general del dataset:

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 8240 entries, 0 to 8239

Data columns (total 7 columns):

#	Column	Non-Null Count	Dtype
0	Round	8240 non-null	object
1	Date	0 non-null	float64
2	Team 1	8240 non-null	object
3	Team 1 Goals	8240 non-null	int64
4	Team 2 Goals	8240 non-null	int64
5	Team 2	8240 non-null	object
6	Season	8240 non-null	object

dtypes: float64(1), int64(2), object(4)

memory usage: 450.8+ KB

Descripción estadística del dataset:

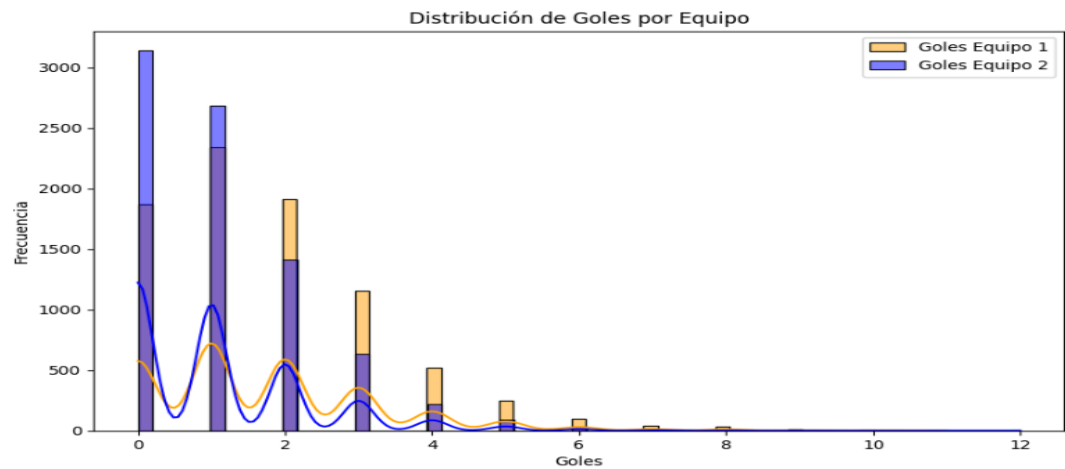
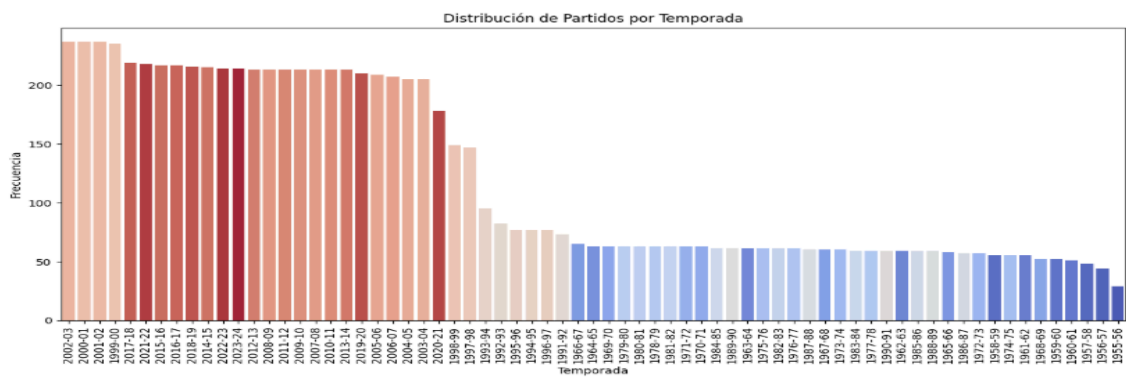
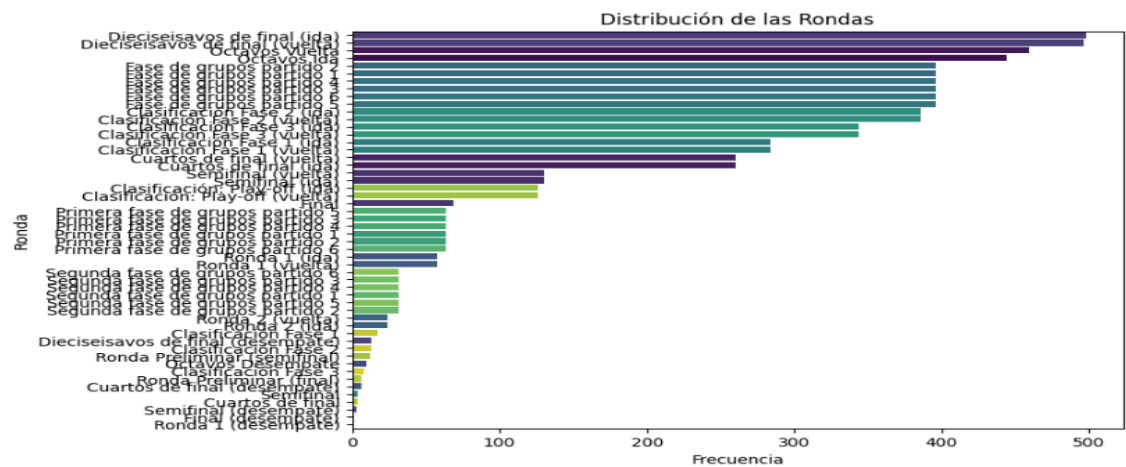
	Date	Team 1 Goals	Team 2 Goals
count	0.0	8240.000000	8240.000000
mean	NaN	1.734102	1.107039
std	NaN	1.545617	1.223609
min	NaN	0.000000	0.000000
25%	NaN	1.000000	0.000000
50%	NaN	1.000000	1.000000
75%	NaN	3.000000	2.000000

Sección 1: Análisis Univariado General

Introducción

El análisis univariado general tiene como objetivo proporcionar una vista inicial de las características más relevantes del dataset, sin entrar aún en el contraste de hipótesis. Esta primera aproximación permite identificar patrones, distribuciones y tendencias iniciales que guiarán los análisis posteriores y plantearán preguntas clave sobre los datos.

Para este propósito, se examinan variables generales del dataset como la distribución de rondas, temporadas y los goles por equipo. Estas métricas proporcionan un punto de partida sólido para entender el contexto histórico y la dinámica del torneo.



Conclusiones:

- **Distribución de las rondas:** Los octavos de final destacan como la etapa más representada en el dataset, lo que refleja el formato típico de la competición en las últimas décadas. Por otro lado, las rondas de desempate son extremadamente raras, sugiriendo su desaparición en temporadas recientes.
- **Distribución temporal:** Se observa un aumento en el número de partidos por temporada a partir del año 2000, coincidiendo con las expansiones en el formato de la Champions League. Las temporadas más recientes muestran una ligera reducción, atribuida probablemente a la pandemia de COVID-19.
- **Distribución de goles por equipo:** La mayoría de los equipos anotan entre 0 y 2 goles por partido, con picos menores en valores más altos. Esto destaca la naturaleza competitiva del torneo, donde las goleadas son poco frecuentes y los márgenes de victoria suelen ser ajustados.

Sección 2: Análisis Univariado con Hipótesis

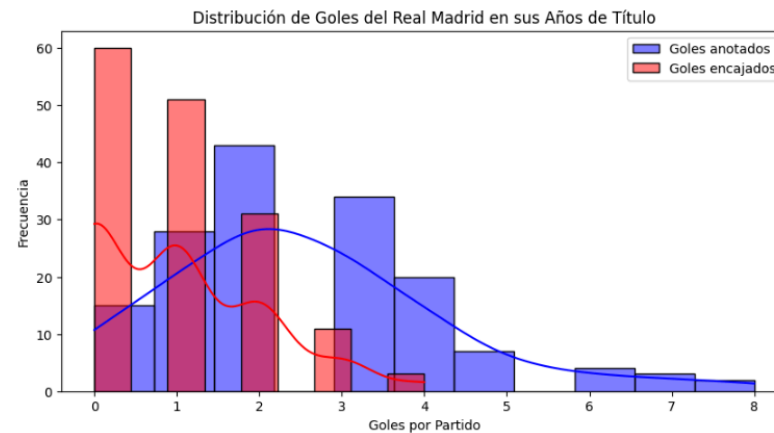
Hipótesis 1: Dominancia del Real Madrid en sus años de títulos

Introducción

El Real Madrid es el equipo más laureado de la historia de la Champions League. Estudiar sus años de títulos permite entender si su éxito se basa en un rendimiento consistentemente superior en métricas clave como goles anotados y encajados. Esto nos ayudará a identificar patrones que otros equipos podrían emular.

Conclusiones:

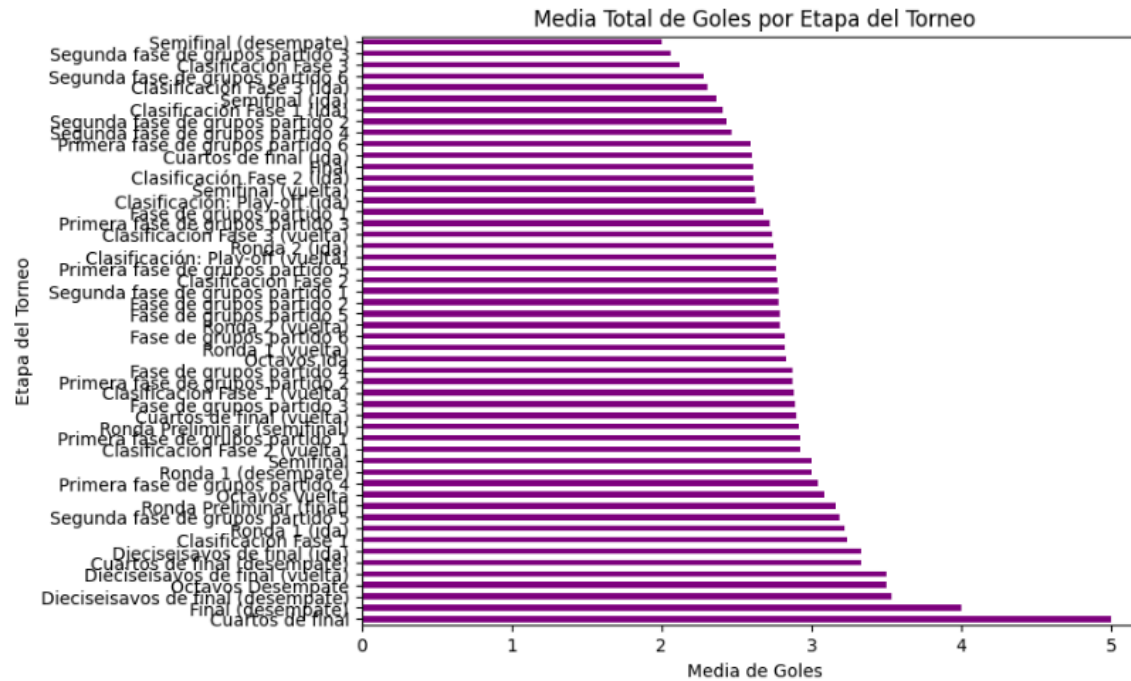
- Los goles más frecuentes son 1 y 2 por partido, reflejando una capacidad para ganar en partidos ajustados.
- Existe una marcada diferencia entre los goles anotados y encajados, lo que destaca su equilibrio táctico en años de éxito.



Hipótesis 2: Rendimiento por Etapa del Torneo

Introducción

Analizar las etapas del torneo nos permite identificar cómo cambian las dinámicas de los partidos a medida que se avanza en el campeonato. Esto es clave para entender la evolución táctica y la competitividad en fases decisivas como semifinales y finales.



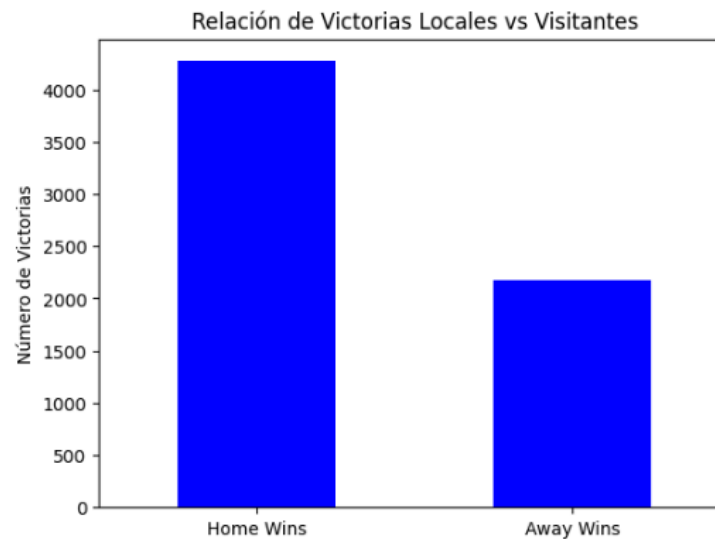
Conclusiones:

- Las fases iniciales presentan más goles debido a la desigualdad entre equipos.
- En semifinales y finales, la media de goles disminuye considerablemente, mostrando mayor paridad entre equipos.

Hipótesis 3: El Efecto Localía

Introducción

La localía ha sido tradicionalmente un factor influyente en el rendimiento de los equipos. Analizar su impacto en el número de goles anotados puede revelar si esta ventaja sigue siendo relevante en el fútbol moderno.



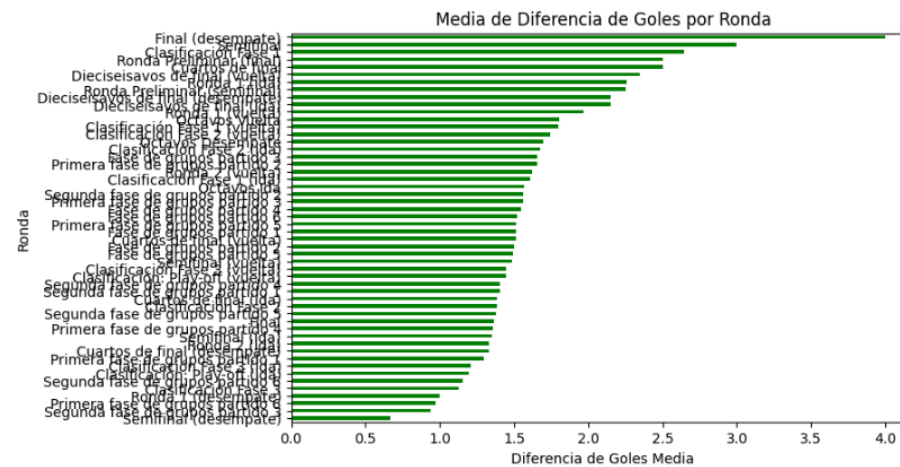
Conclusiones:

- Los equipos locales anotan más goles en promedio que los visitantes, aunque la diferencia ha disminuido en años recientes.
- Factores como la ausencia de público durante la pandemia pueden haber afectado esta tendencia.

Hipótesis 4: Rondas más Competitivas

Introducción

Entender qué rondas son históricamente más reñidas nos permite evaluar la evolución de la competitividad en el torneo. Esto es útil para analizar cómo los equipos gestionan partidos clave en etapas avanzadas.



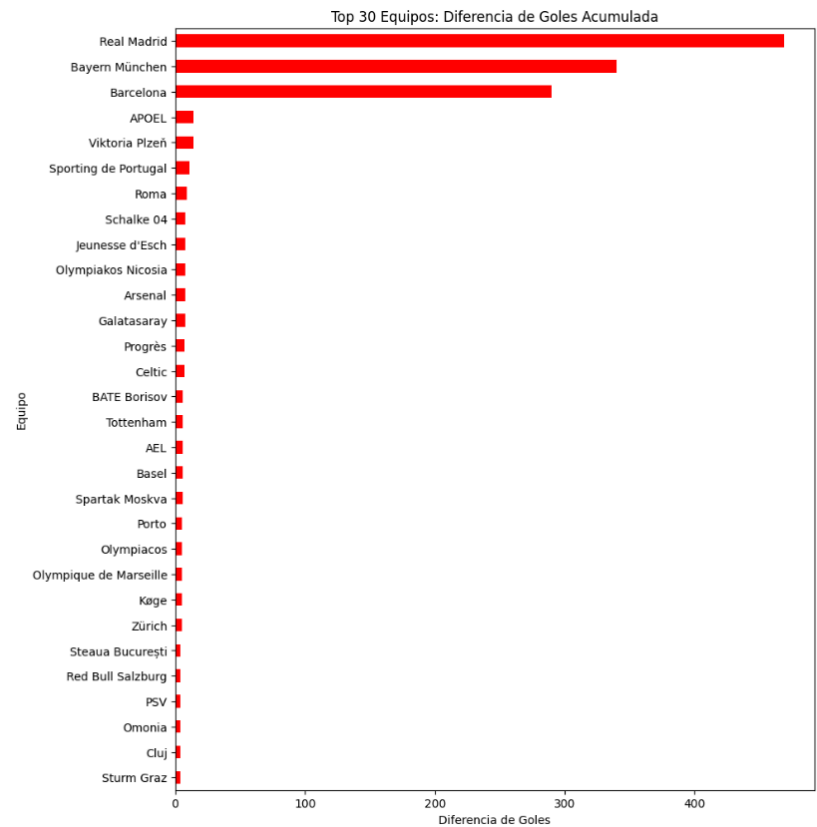
Conclusiones:

- Las semifinales y finales son las más reñidas, con diferencias promedio de goles más bajas.
- Las rondas iniciales presentan mayores diferencias, debido a enfrentamientos más desiguales en términos de calidad de equipos.

Hipótesis 5: Comparación entre Equipos Grandes

Introducción

Comparar equipos históricos como el Real Madrid, Barcelona y Bayern nos ayuda a entender qué factores han contribuido a su éxito sostenido y qué lecciones se pueden extraer para otros clubes.



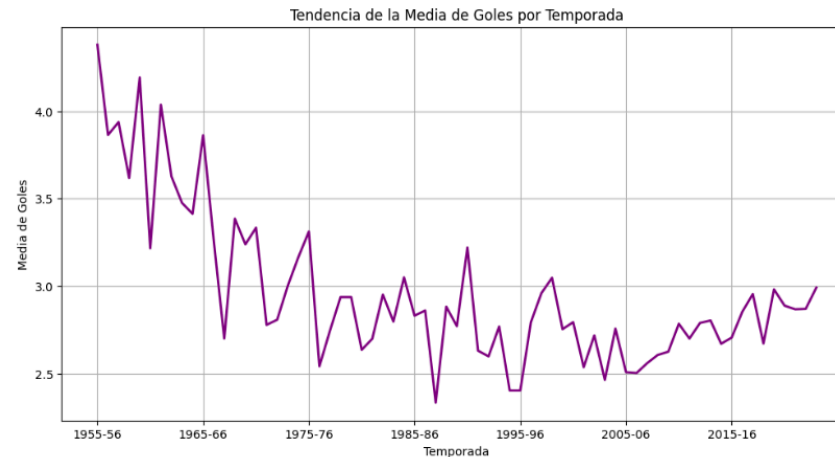
Conclusiones:

- Real Madrid, Bayern y Barcelona lideran en diferencia acumulada de goles, consolidando su estatus de equipos dominantes.
 - La brecha entre estos equipos y el resto refleja su consistencia histórica en la competición.
-

Hipótesis 6: Tendencias Históricas de Goles

Introducción

Explorar la evolución histórica de los goles por partido permite analizar cómo las reglas, tácticas y estilos de juego han influido en el desarrollo del torneo a lo largo de las décadas.



Conclusiones:

- Las primeras décadas mostraban más goles por partido, posiblemente debido a un estilo más ofensivo.
- Desde los años 70, la media de goles disminuyó, reflejando un enfoque más defensivo.

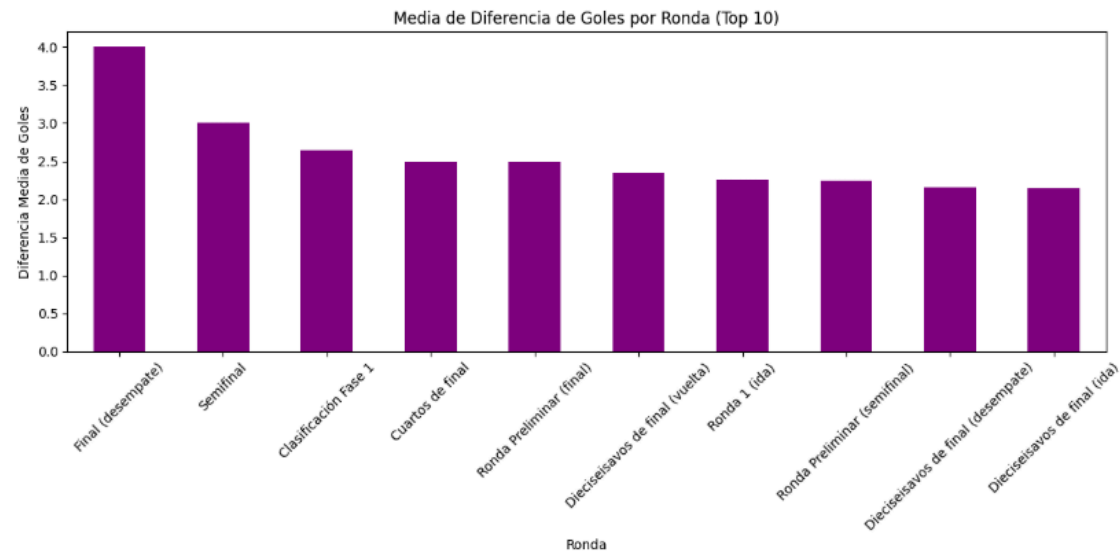
- En las últimas décadas, un repunte sugiere un regreso a estrategias más ofensivas, posiblemente impulsadas por cambios en las reglas.

Sección 3: Análisis Bivariado con Hipótesis

Hipótesis 1: Relación entre la etapa del torneo y la diferencia de goles

Introducción

Explorar cómo varía la diferencia de goles según la etapa del torneo nos permite entender en qué fases los partidos tienden a ser más reñidos y si las dinámicas competitivas cambian conforme avanza el campeonato.



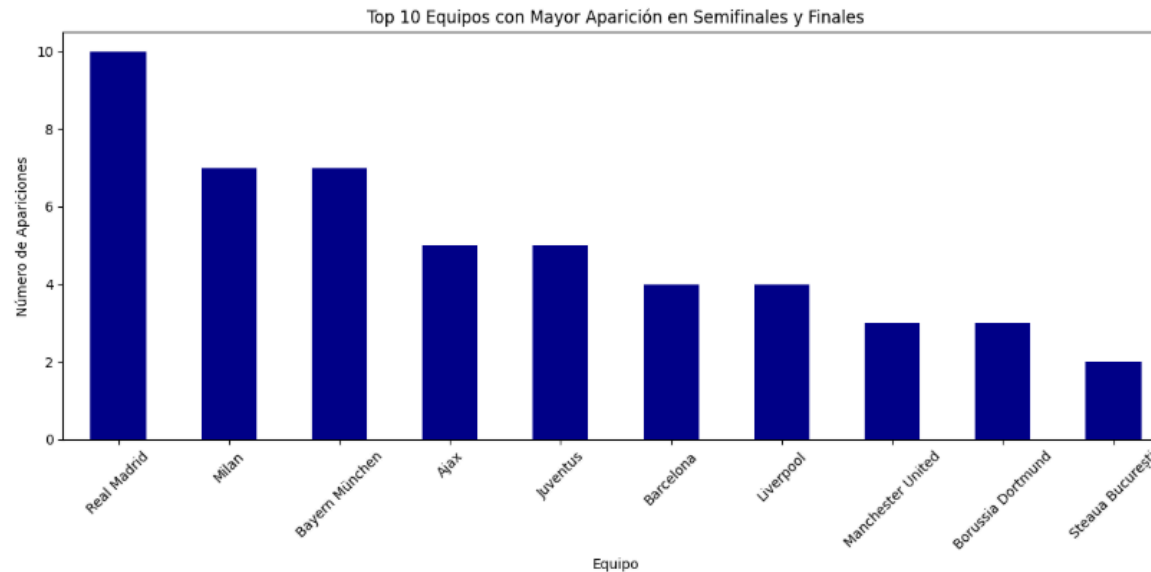
Conclusiones:

- Las semifinales y finales presentan las menores diferencias promedio de goles, reflejando un alto nivel de competitividad.
- Las rondas clasificatorias muestran mayores diferencias, debido a la disparidad de nivel entre equipos.

Hipótesis 2: Relación entre el equipo y la etapa del torneo alcanzada

Introducción

Identificar qué equipos alcanzan consistentemente las etapas avanzadas permite evaluar su dominio histórico y los factores que contribuyen a su éxito en estas fases.



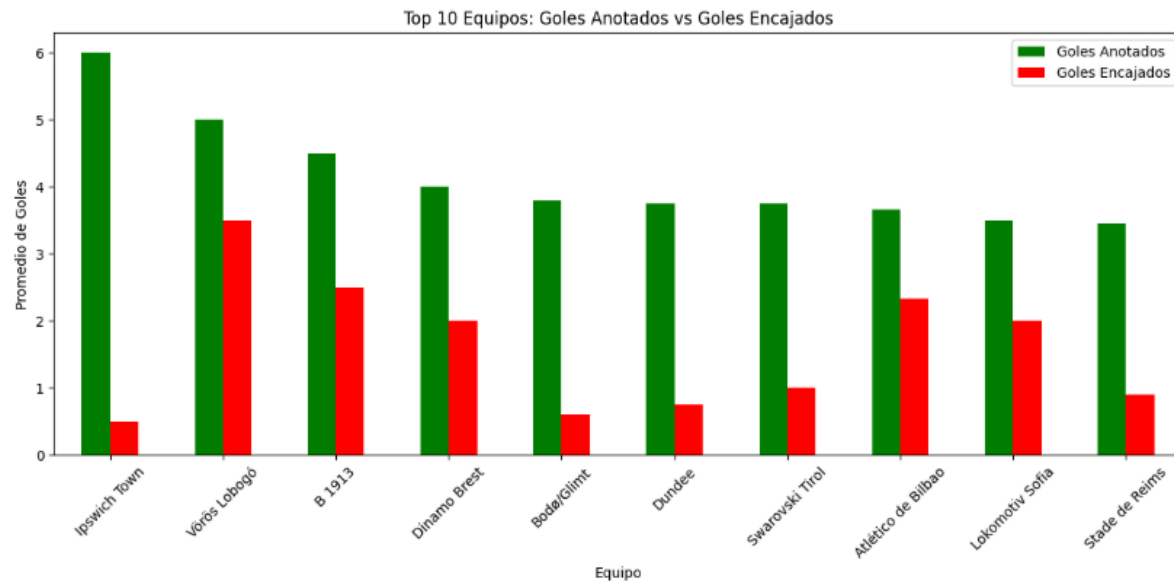
Conclusiones:

- Equipos como el Real Madrid, Bayern y Milan dominan en apariciones en semifinales y finales, reforzando su estatus de élite.
 - El Ajax y el Liverpool destacan como equipos que han tenido períodos específicos de alto rendimiento.
-

Hipótesis 3: Relación entre goles anotados y encajados por equipo

Introducción

Esta hipótesis estudia si los equipos más ofensivos también tienden a ser defensivamente sólidos o si, por el contrario, existe una compensación entre ambas métricas.



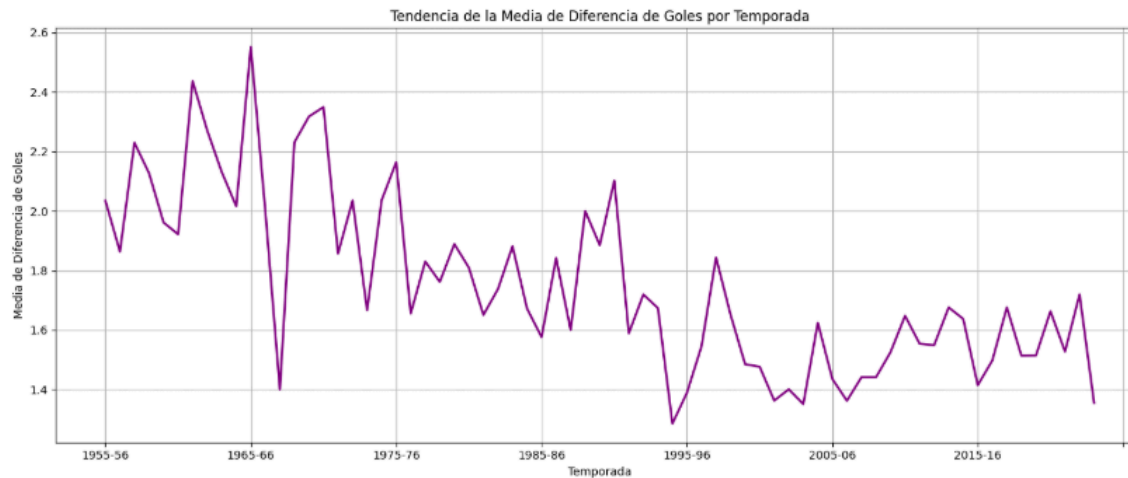
Conclusiones:

- Los equipos más ofensivos suelen encajar más goles, reflejando un estilo de juego abierto y arriesgado.
 - Los equipos con menos goles anotados y encajados probablemente representan a equipos que han participado en pocas temporadas.
-

Hipótesis 4: Relación entre temporada y diferencia de goles en partidos

Introducción

Analizar la diferencia promedio de goles por temporada nos permite entender si el torneo se ha vuelto más competitivo con el tiempo, reflejando una paridad creciente entre equipos.



Conclusiones:

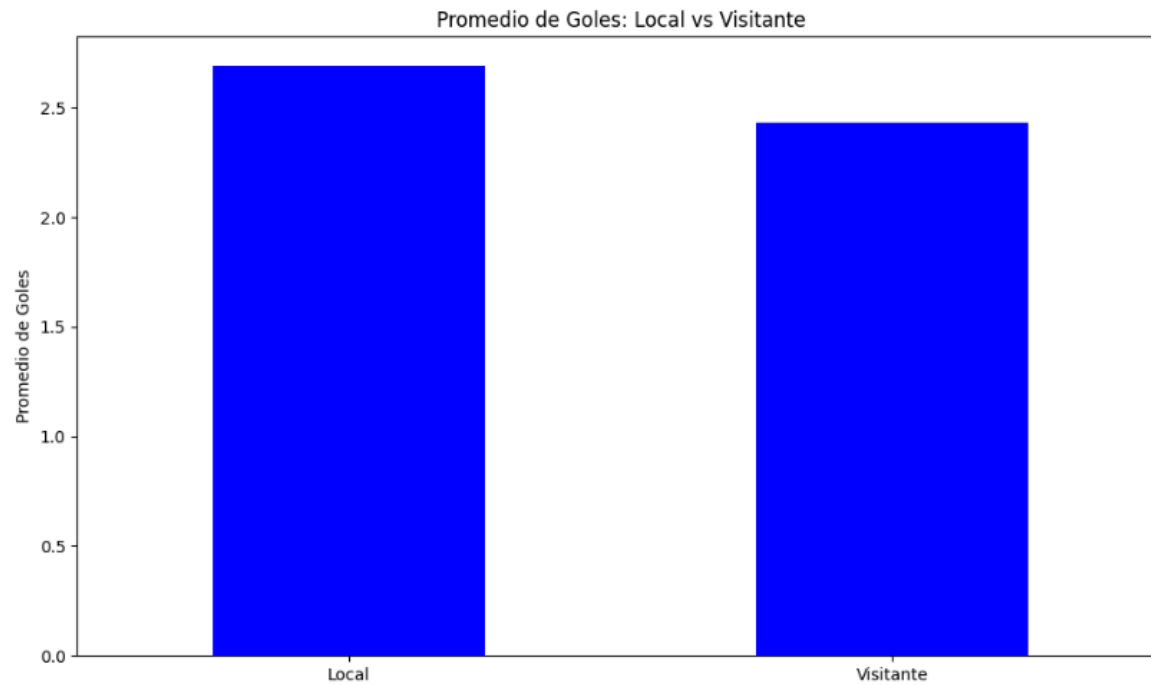
- Hay una tendencia descendente en la diferencia de goles a lo largo de las décadas, indicando mayor equilibrio competitivo.

- Las primeras décadas muestran diferencias significativamente mayores, probablemente debido a desigualdades en calidad entre equipos.
-

Hipótesis 5: Relación entre localía y cantidad de goles anotados

Introducción

Evaluar si la localía sigue siendo una ventaja significativa nos permite entender cómo el entorno puede influir en los resultados y qué estrategias podrían implementarse para contrarrestarla.



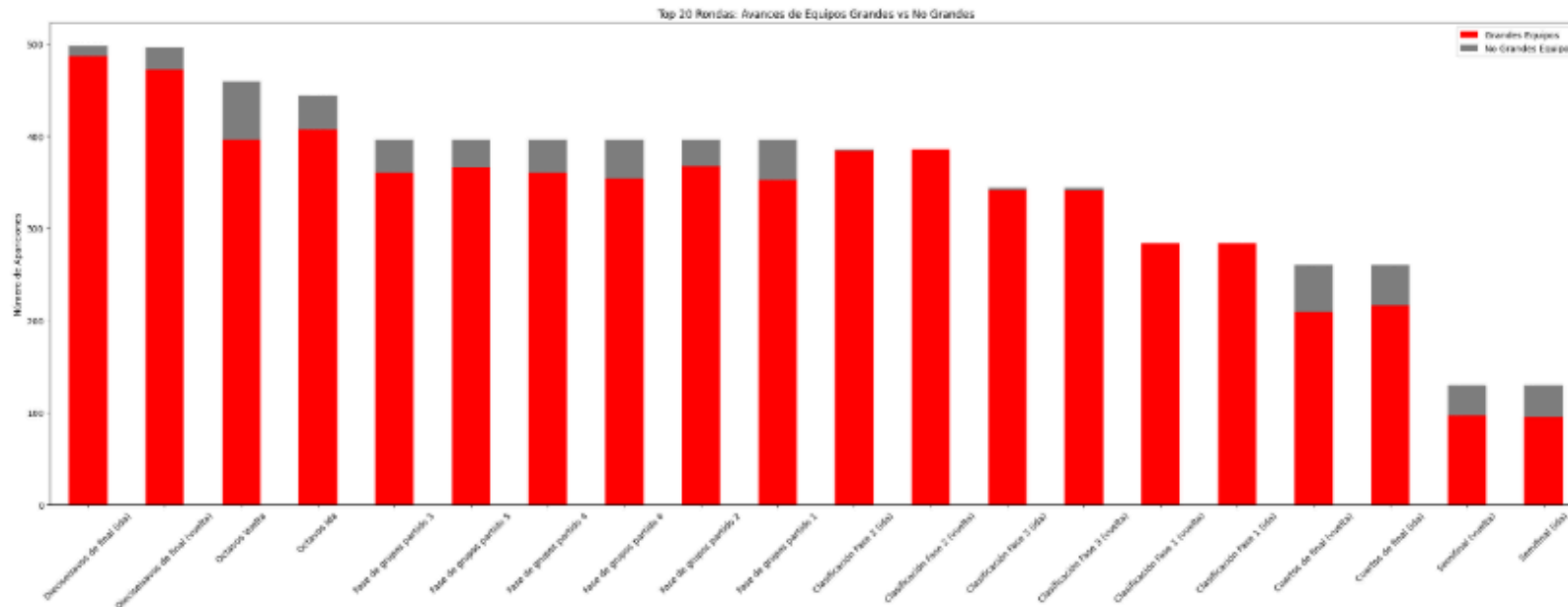
Conclusiones:

- Los equipos locales anotan más goles en promedio, pero la diferencia con los visitantes es relativamente pequeña.
- Esto sugiere que la ventaja local, aunque relevante, ha disminuido con el tiempo.

Hipótesis 6: Relación entre el historial de los equipos (grandes vs no grandes) y las etapas alcanzadas

Introducción

Comparar el desempeño de equipos grandes frente a no grandes en las etapas del torneo permite evaluar la consistencia y las sorpresas en la historia de la competición.



Conclusiones:

- Los equipos grandes dominan las apariciones en etapas avanzadas, pero los no grandes logran avanzar significativamente en rondas clasificatorias.
 - Esto resalta la importancia de las estrategias y los factores circunstanciales para los underdogs.
-

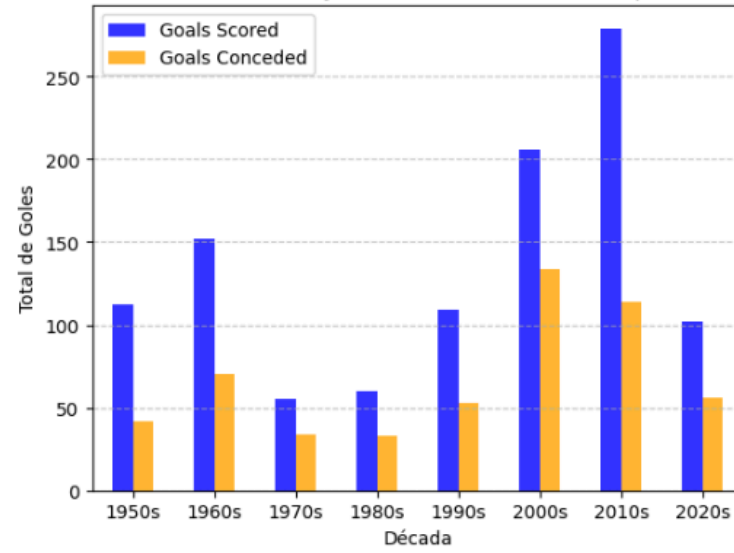
Sección 4: Análisis Multivariante con Hipótesis

Hipótesis 1: Rendimiento histórico del Real Madrid

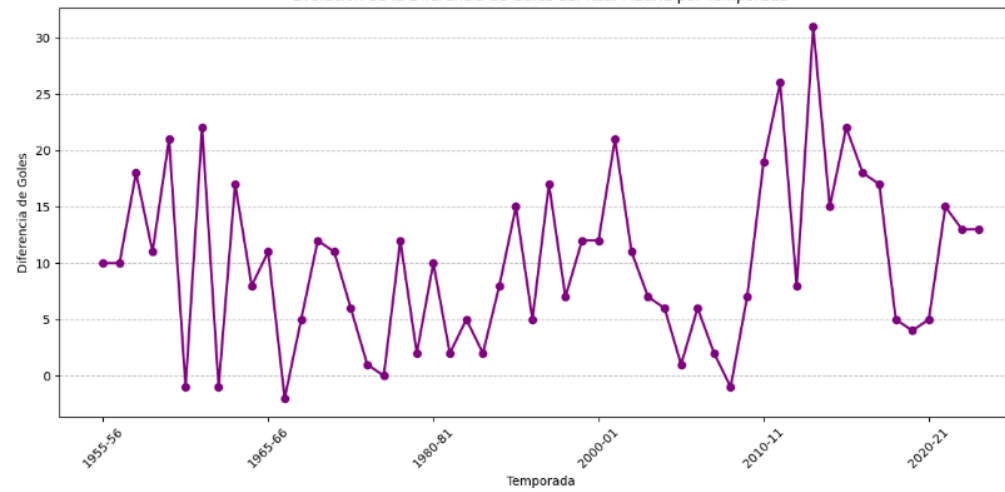
Introducción

El Real Madrid, como equipo con más títulos de Champions, se analiza para determinar si su dominio también se traduce en métricas clave como goles anotados, diferencia de goles y apariciones avanzadas.

Rendimiento Ofensivo y Defensivo del Real Madrid por Década



Evolución de la Diferencia de Goles del Real Madrid por Temporada



Conclusiones:

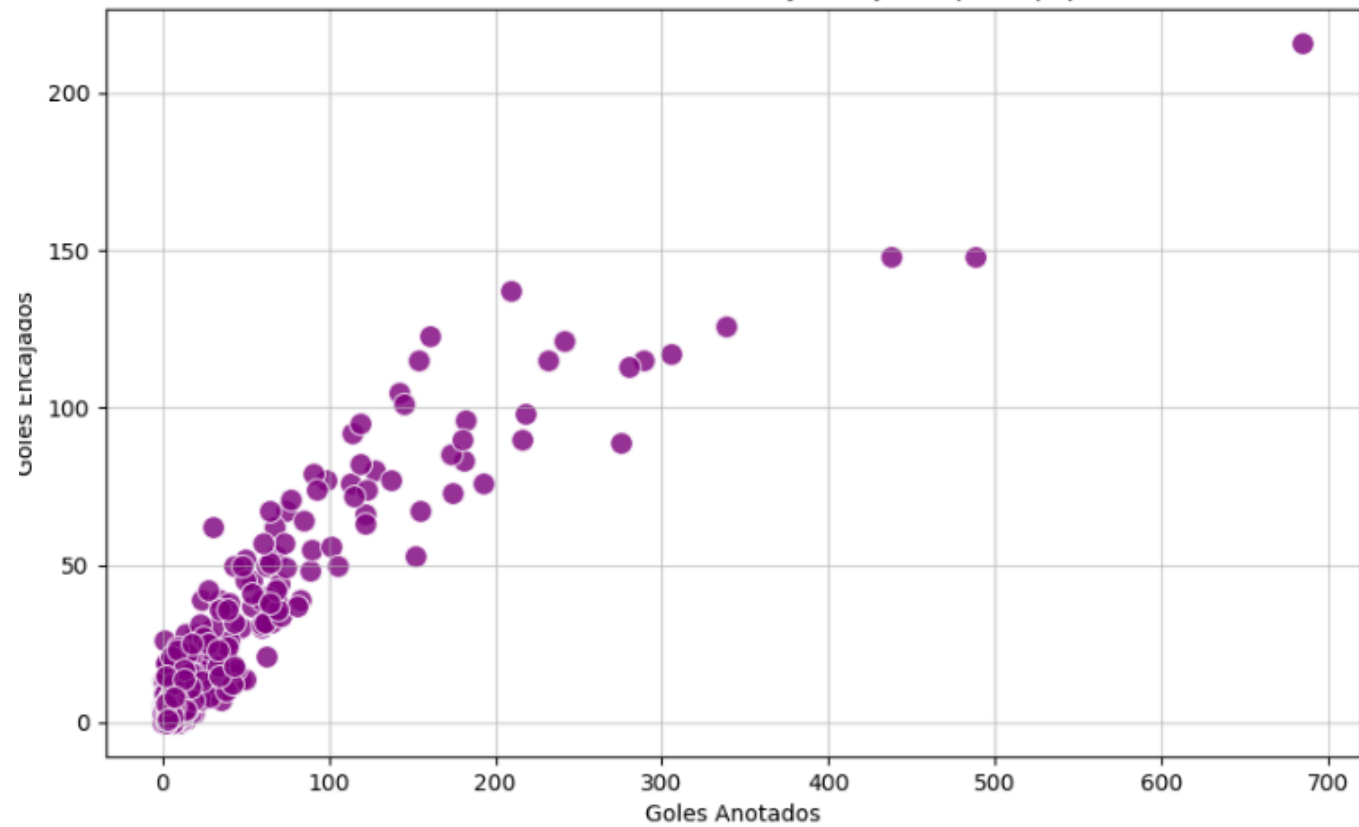
- La década de 2010 destaca por un rendimiento ofensivo excepcional, con un equilibrio positivo en goles anotados y encajados.
 - Las fluctuaciones en la diferencia de goles reflejan ciclos de éxito y transición, coincidiendo con cambios en plantillas y estrategias.
-

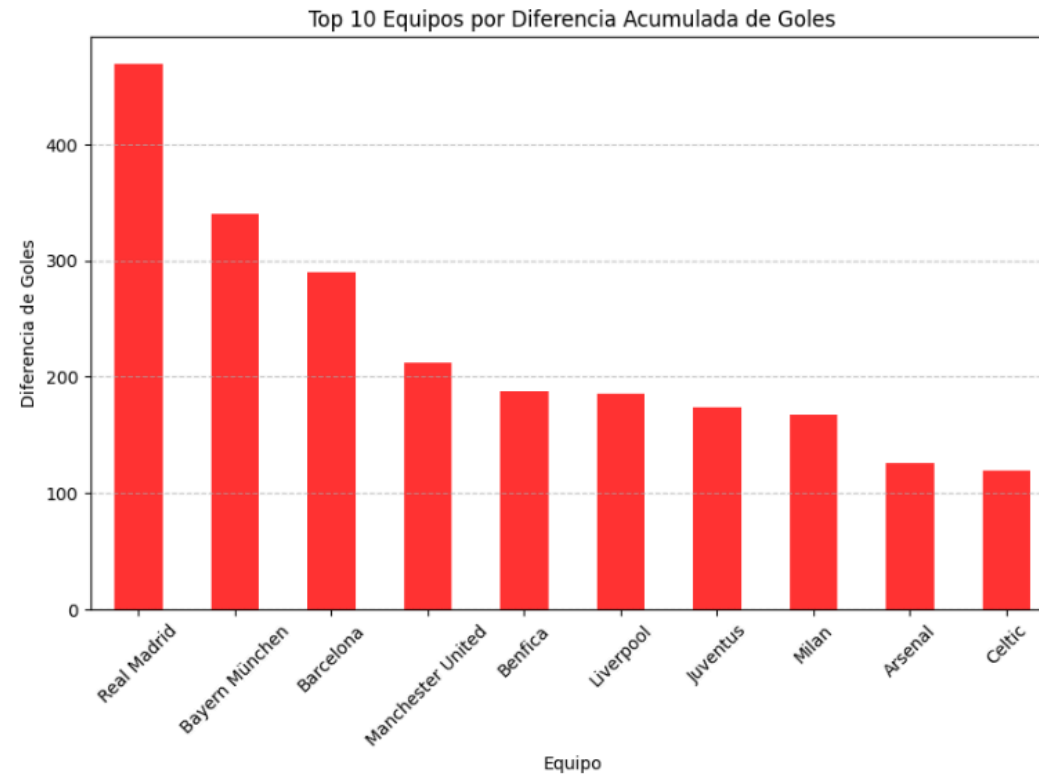
Hipótesis 2: Éxito vinculado a ofensiva y defensiva

Introducción

Explorar cómo el éxito de un equipo depende tanto de su ataque como de su defensa ayuda a identificar patrones comunes en los equipos más exitosos.

Relación entre Goles Anotados y Encajados por Equipo





Conclusiones:

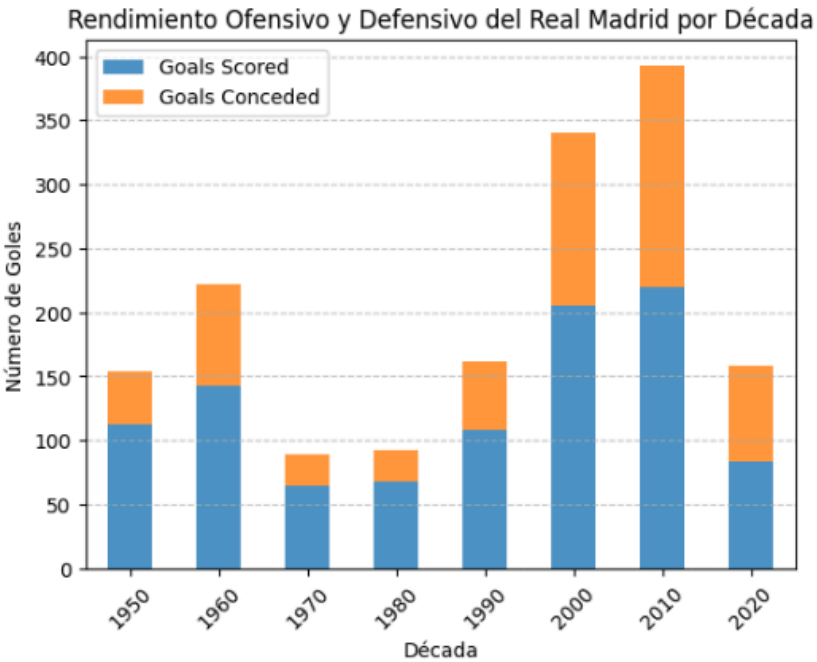
- Los equipos con un alto equilibrio entre ofensiva y defensiva tienden a tener un desempeño superior en el torneo.
- El Real Madrid lidera la diferencia acumulada de goles, reflejando su consistencia histórica.

Sección 5: Insights Adicionales

Insight 1: Rendimiento por década del Real Madrid

Introducción

Analizar cómo el rendimiento del Real Madrid ha evolucionado por décadas ofrece una perspectiva temporal sobre sus fortalezas y áreas de mejora.



Promedio de goles necesarios para ganar un título: 3.78

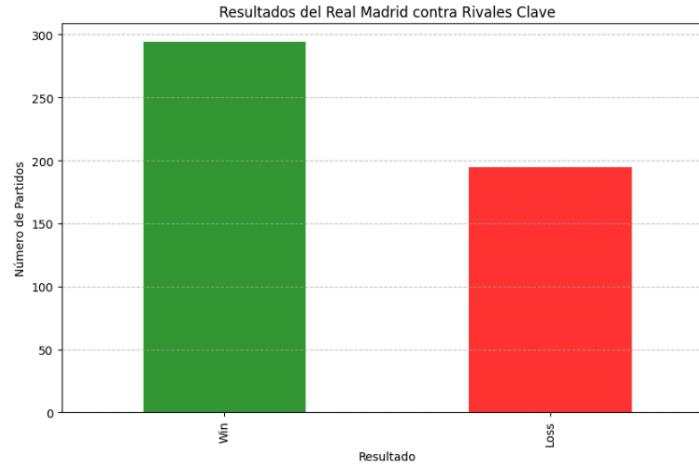
Conclusiones:

- Las décadas de 2000 y 2010 destacan por un mayor número de goles anotados, coincidiendo con períodos de éxito internacional.
 - Las décadas anteriores muestran un menor rendimiento ofensivo pero mayor estabilidad defensiva.
-

Insight 2: Resultados contra rivales clave

Introducción

Evaluar el desempeño del Real Madrid contra rivales históricos permite entender cómo afronta partidos de alta exigencia en las etapas avanzadas.



Conclusiones:

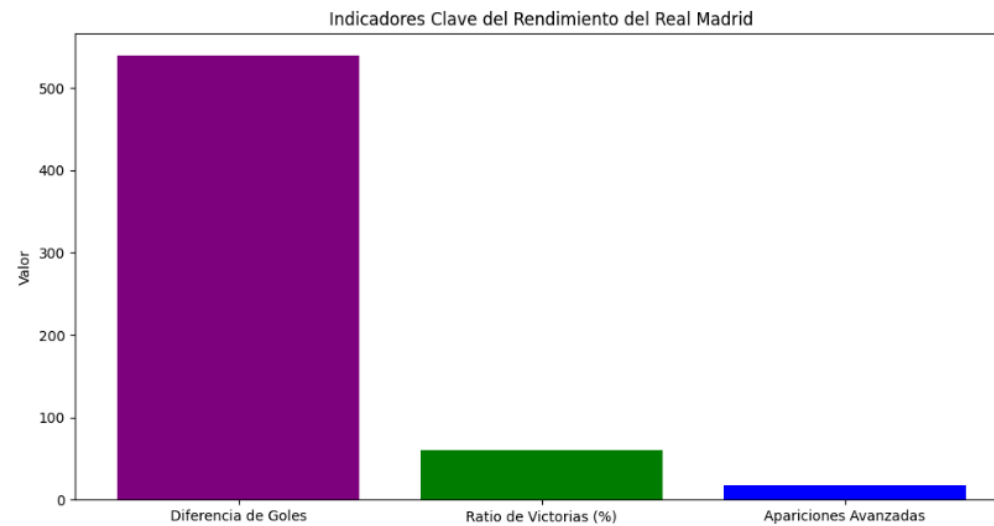
- El balance positivo contra rivales como Bayern y Barcelona refuerza la capacidad del equipo para brillar en contextos de alta presión.
- Identificar patrones en las derrotas podría ofrecer oportunidades para ajustes estratégicos futuros.

Sección 6: Contraste con Hipótesis

Contraste de Hipótesis

Introducción

Se realiza un contraste final para validar si las hipótesis planteadas encuentran respaldo en los datos analizados, con un enfoque en las métricas clave del Real Madrid.



Conclusiones:

- El Real Madrid lidera en diferencia de goles acumulada, con un ratio de victorias significativo y frecuentes apariciones en etapas avanzadas.
- Las métricas confirman su posición dominante en la historia del torneo, aunque su rendimiento no siempre garantiza victorias en partidos cruciales.

Sección 7: Dashboard interactivo

El análisis realizado se complementa con un dashboard interactivo desarrollado en Streamlit, que permite explorar en profundidad el rendimiento de equipos en la Champions League mediante un selector.

[Ver dashboard](#)