

UNIVERSIDAD DEL VALLE DE GUATEMALA

Facultad de Ingeniería



Proyecto No. 1 Desarrollo y consulta de Base de Datos

Gustavo Adolfo Cruz Bardales - 22779

(Licenciatura en: Ingeniería en Ciencias de la Computación)

Josué Emanuel Say Garcia - 22801

(Licenciatura en: Ingeniería en Ciencias de la Computación)

Guatemala, febrero 2023

Introducción

En la búsqueda de una estrategia efectiva para la selección de equipos en el ámbito de las apuestas deportivas, se implementó un criterio que evaluó la combinación de la probabilidad máxima en diversas casas de apuestas, la consistencia temporal de dicha probabilidad a lo largo de múltiples temporadas y la presencia exclusiva del equipo como único representante en su respectiva liga. Este enfoque se orientó a maximizar las posibilidades de éxito al considerar no solo el rendimiento histórico o la posición actual de los equipos, sino también la percepción de las casas de apuestas en términos de probabilidades.

En la aplicación de este criterio para la selección de equipos en apuestas deportivas, se han identificado dos equipos como candidatos sólidos para apostar en una liga específica: Paris Saint Germain en la "Ligue 1" o Bayern Munich en la "Bundesliga".

Análisis previo (Tablas)

Teams (Equipos)

Contiene el nombre e identificación del equipo.

Players (Jugadores)

Contiene el nombre e identificación de los jugadores.

Leagues (Ligas)

Contiene el nombre e identificador de ligas europeas.

Appearances (Apariciones)

Contiene la información relacionada al jugador en cada partido, tales como el número del partido, identificador del jugador, liga a la que pertenece goles a favor y contra, disparos, amonestaciones, etc.

Games (Partidos)

Contiene información sobre los partidos individuales, incluyendo los equipos en casa y fuera de casa, goles, probabilidades de las apuestas (como Bet365, Blue Square, Bet&Win, etc.), y los goles esperados (xGoals).

Shots (Disparos)

Registra detalles de los disparos durante los partidos, así como el tirador, asistente, situación previa antes de un resultado del disparo, la posición desde donde se realizó el disparo, minuto del disparo y el resultado en que finalizó dicho disparo.

TeamStats (Estadísticas del Equipo)

Contiene información relacionada al equipo en cada partido, tales como posición (local o visitante), identificador del equipo y partido, fecha, temporada, goles, tiros de esquina, amonestaciones, resultado del partido, etc.

Etapa 1

En la etapa 1 se encargará de procesar los archivos CSV proporcionados y levantarlos en una base de datos PostgreSQL donde pueda ejecutar sus queries para análisis. Para esto deberá:

1. Descargar los archivos CSV y verificar qué información se incluye en cada uno.
2. Crear la base de datos y las tablas en donde almacenará los datos.
3. Desarrollar un script de Python (o lenguaje a su elección) que pueda leer la información de los archivos CSV, conectarse a su base de datos y alimentar la información en las tablas creadas

[Enlace GitHub para observar el script en Python y Queries](#)

Etapa 2

En la etapa 2 usted ejecutará algunos queries que le permitan familiarizarse con el modelo de datos presentado. Para esto deberá obtener lo siguiente:

Según estadísticas:

Pregunta 1

La cantidad de juegos jugados en cada temporada por cada equipo, de cada liga (tome en cuenta que cada equipo puede jugar como visitante o como anfitrión).

Query:

```
SELECT
    g.leagueID,
    g.season,
    t.name,
    COUNT(*) AS partidos_jugados
FROM (
    SELECT leagueID, season, homeTeamID AS team
    FROM games
    UNION ALL
    SELECT leagueID, season, awayTeamID AS team
    FROM games
) AS g
JOIN teams t ON g.team = t.teamID
GROUP BY g.leagueID, g.season, t.name
ORDER BY g.leagueID, g.season, t.name;
```

Resultado:

	leagueid bigint	season bigint	name text	partidos_jugados bigint
1	1	2014	Arsenal	38
2	1	2014	Aston Villa	38
3	1	2014	Burnley	38
4	1	2014	Chelsea	38
5	1	2014	Crystal Palace	38
6	1	2014	Everton	38
7	1	2014	Hull	38
8	1	2014	Leicester	38
9	1	2014	Liverpool	38
10	1	2014	Manchester City	38

Conclusión:

Este query reflejar la cantidad de partidos jugados (como local y visitante) por cada liga y temporada.

Pregunta 2

¿Quién es el mejor equipo de todas las ligas y de todas las temporadas según las estadísticas de diferencia de goles?

Hint: Obtenga la cantidad de goles a favor, goles en contra y la diferencia entre las dos anteriores, esto por cada temporada y por cada equipo de cada liga.

Utilizando este mismo query, obtenga el ranking de los equipos por temporada y por liga, ordenados por ese ranking de manera descendente por diferencia (utilice la función Rank () over partition), para obtener el equipo ganador.

Query (Sin ranking):

```
WITH DiferenciaGoles AS (
    SELECT
        season,
        leagueid,
        hometeamid AS equipo,
        SUM(homegoals - awaygoals) AS diferencia_goles
    FROM
        games
    GROUP BY
        season,
        leagueid,
        hometeamid
    UNION ALL
    SELECT
        season,
        leagueid,
        awayteamid AS equipo,
        SUM(awaygoals - homegoals) AS diferencia_goles
    FROM
        games
    GROUP BY
        season,
        leagueid,
        awayteamid
)
SELECT
    season,
    leagueid,
    t.name
    equipo,
    SUM(diferencia_goles) AS goles_equipo_total
FROM
    DiferenciaGoles d, teams t
    where d.equipo = t.teamid
GROUP BY
    season,
    leagueid,
    equipo,
    t.name
ORDER BY
    4 DESC;
```

Resultado:

	season bigint	leagueid bigint	equipo text	goles_equipo_total numeric
1	2014	4	Barcelona	89
2	2015	4	Barcelona	83
3	2015	5	Paris Saint Germain	83
4	2014	4	Real Madrid	80
5	2017	1	Manchester City	79
6	2016	4	Barcelona	79
7	2017	5	Paris Saint Germain	79
8	2015	4	Real Madrid	76
9	2016	5	Monaco	76
10	2018	1	Manchester City	72

Conclusión:

En este query se muestra que el equipo con más goles por temporada y liga, sin embargo no se toma ningún ranking para generar un orden:

Query (Con ranking):

```
WITH DiferenciaGoles AS (
    SELECT
        season,
        leagueid,
        hometeamid AS equipo,
        SUM(homegoals - awaygoals) AS diferencia_goles
    FROM
        games
    GROUP BY
        season,
        leagueid,
        hometeamid
    UNION ALL
    SELECT
        season,
        leagueid,
        awayteamid AS equipo,
        SUM(awaygoals - homegoals) AS diferencia_goles
    FROM
        games
    GROUP BY
        season,
        leagueid,
        awayteamid
),
RankingEquipos AS (
    SELECT
        season,
        leagueid,
        equipo,
        SUM(diferencia_goles) AS diferencia_total,
        RANK() OVER (PARTITION BY season, leagueid ORDER BY SUM(diferencia_goles) DESC) AS ranking
    FROM
        DiferenciaGoles
    GROUP BY
        season,
        leagueid,
        equipo
)
SELECT
    season,
    leagueid,
    t.name,
    diferencia_total,
    ranking
FROM
    RankingEquipos r, teams t
    WHERE r.equipo = t.teamid
ORDER BY
    season,
    leagueid,
    ranking;
```

Resultado:

	season bigint	leagueid bigint	name text	diferencia_total numeric	ranking bigint
1	2014	1	Manchester City	45	1
2	2014	1	Chelsea	41	2
3	2014	1	Arsenal	35	3
4	2014	1	Manchester United	25	4
5	2014	1	Southampton	21	5
6	2014	1	Tottenham	5	6
7	2014	1	Liverpool	4	7
8	2014	1	Stoke	3	8
9	2014	1	Everton	-2	9
10	2014	1	Swansea	-3	10

Conclusión:

En este query ya se puede ver un orden con el ranking para determinar que el mejor equipo por la diferencia de goles (a favor menos en contra) de cada liga y temporada.

Pregunta 3

¿Quiénes son los jugadores que han realizado mayor cantidad de goles a través de todas las temporadas?

¿Cuáles son los jugadores con mayor cantidad de pases izquierdos y pases derechos que han hecho goles?

Query (jugadores con más goles):

```
SELECT a.playerid, p.name, SUM(a.goals) as total_goles
FROM appearances a
JOIN players p ON a.playerid = p.playerid
GROUP BY a.playerid, p.name
ORDER BY total_goles DESC;
```

Resultado:

	playerid bigint	name text	total_goles numeric
1	2097	Lionel Messi	231
2	2371	Cristiano Ronaldo	215
3	227	Robert Lewandowski	203
4	2098	Luis Suarez	168
5	647	Harry Kane	163
6	318	Pierre-Emerick Aubameyang	149
7	1209	Ciro Immobile	133
8	3294	Edinson Cavani	132
9	619	Sergio Aguero	132
10	1250	Mohamed Salah	130

Conclusión:

Se observa que el mejor jugador basado en más goles realizados en todas las temporadas es Lionel Messi con 231 goles.

Query (jugadores con más asistencias):

```

SELECT
    s.shooterid,
    p.name,
    COUNT(CASE WHEN s.shottotype = 'LeftFoot' AND s.shotresult = 'Goal' THEN 1 END) AS izquierdos_goles,
    COUNT(CASE WHEN s.shottotype = 'RightFoot' AND s.shotresult = 'Goal' THEN 1 END) AS derechos_goles,
    SUM(CASE WHEN s.shottotype = 'LeftFoot' AND s.shotresult = 'Goal' THEN 1 ELSE 0 END +
        CASE WHEN s.shottotype = 'RightFoot' AND s.shotresult = 'Goal' THEN 1 ELSE 0 END) as pases_totales
FROM shots s
JOIN players p ON s.shooterid = p.playerid
GROUP BY s.shooterid, p.name
ORDER BY izquierdos goles DESC, derechos goles DESC, pases totales DESC;

```

Resultado:

	shooterid bigint	name text	izquierdos_goles bigint	derechos_goles bigint	pases_totales bigint
1	2097	Lionel Messi	224	38	262
2	1250	Mohamed Salah	104	19	123
3	2270	Antoine Griezmann	102	16	118
4	2290	Iago Aspas	82	21	103
5	2120	Gerard Moreno	75	15	90
6	1294	Paulo Dybala	75	14	89
7	594	Romelu Lukaku	73	32	105
8	1612	Domenico Berardi	61	6	67
9	1426	Josip Ilicic	56	14	70
10	3667	Nabil Fekir	55	11	66

Conclusión:

Se observa que el mejor jugador basado en las asistencias que terminaron en goles en las temporadas fue Lionel Messi con 262 asistencias.

Query (Unión de jugadores con más goles y asistencias):

```
WITH GolesTotales AS (
    SELECT
        a.playerid,
        p.name,
        SUM(a.goals) AS total_goles
    FROM appearances a
    JOIN players p ON a.playerid = p.playerid
    GROUP BY a.playerid, p.name
),
PasesGoles AS (
    SELECT
        s.shooterid,
        p.name,
        COUNT(CASE WHEN s.shottype = 'LeftFoot' AND s.shotresult = 'Goal' THEN 1 END) AS pases_izquierdos_goles,
        COUNT(CASE WHEN s.shottype = 'RightFoot' AND s.shotresult = 'Goal' THEN 1 END) AS pases_derechos_goles,
        SUM(CASE WHEN s.shottype = 'LeftFoot' AND s.shotresult = 'Goal' THEN 1 ELSE 0 END +
            CASE WHEN s.shottype = 'RightFoot' AND s.shotresult = 'Goal' THEN 1 ELSE 0 END) AS pases_totales
    FROM shots s
    JOIN players p ON s.shooterid = p.playerid
    GROUP BY s.shooterid, p.name
)
SELECT
    GT.playerid,
    GT.name,
    GT.total_goles,
    PG.pases_izquierdos_goles,
    PG.pases_derechos_goles,
    PG.pases_totales
FROM GolesTotales GT
JOIN PasesGoles PG ON GT.playerid = PG.shooterid
ORDER BY GT.total_goles DESC, PG.pases_izquierdos_goles DESC, PG.pases_derechos_goles DESC, PG.pases_totales DESC;
```

Resultado:

	playerid bigint	name text	total_goles numeric	pases_izquierdos_goles bigint	pases_derechos_goles bigint	pases_totales bigint
1	2097	Lionel Messi	231	224	38	262
2	2371	Cristiano Ronaldo	215	47	138	185
3	227	Robert Lewandowski	203	31	141	172
4	2098	Luis Suarez	168	29	126	155
5	647	Harry Kane	163	35	103	138
6	318	Pierre-Emerick Aubameyang	149	29	111	140
7	1209	Ciro Immobile	133	16	104	120
8	619	Sergio Aguero	132	28	90	118
9	3294	Edinson Cavani	132	22	84	106
10	1250	Mohamed Salah	130	104	19	123

Conclusión:

El mejor jugador basado en más goles y asistencias que terminaron en gol en las temporadas fue Lionel Messi con 231 goles y 262 asistencias.

Compare contra los resultados del inciso 2 y determine de manera textual si dichos jugadores pertenecen a los equipos del inciso anterior.

Query (por goles):

```
WITH GolesTotales AS (
    SELECT
        a.playerid,
        p.name,
        SUM(a.goals) AS total_goles
    FROM appearances a
    JOIN players p ON a.playerid = p.playerid
    GROUP BY a.playerid, p.name
)
SELECT
    GT.playerid,
    GT.name,
    GT.total_goles,
    ARRAY(SELECT DISTINCT t."name" FROM teamstats ts JOIN teams t ON ts.teamid = t.teamid WHERE ts.gameid IN (SELECT gameid FROM appearances ap WHERE ap.playerid = GT.playerid)) AS equipos
FROM GolesTotales GT
ORDER BY GT.total_goles DESC;
```

Resultado:

playerid	name	total_goles	equipos
bigint	text	numeric	text[]
2097	Lionel Messi	231	{Alaves,Almeria,"Athletic Club",
2371	Cristiano Ronaldo	215	{"AC Milan",Alaves,Almeria,Atal
227	Robert Lewandowski	203	{"Arminia Bielefeld",Augsburg,E
2098	Luis Suarez	168	{Alaves,Almeria,"Athletic Club",
647	Harry Kane	163	{Arsenal,"Aston Villa",Bournemo
318	Pierre-Emerick Aubameyang	149	{Arsenal,"Aston Villa",Augsburg,
1209	Ciro Immobile	133	{"AC Milan",Atalanta,Augsburg,"
619	Sergio Aguero	132	{Arsenal,"Aston Villa",Bournemo
3294	Edinson Cavani	132	{Amiens,Angers,Arsenal,"Aston
1250	Mohamed Salah	130	{"AC Milan",Arsenal,"Aston Villa"

Query (por asistencias):

```
WITH AsistenciasTotales AS (
    SELECT
        s.shooterid,
        p.name,
        COUNT(CASE WHEN s.shottype = 'LeftFoot' AND s.shotresult = 'Goal' THEN 1 END) AS asistencias_izquierdas,
        COUNT(CASE WHEN s.shottype = 'RightFoot' AND s.shotresult = 'Goal' THEN 1 END) AS asistencias_derechas,
        SUM(CASE WHEN s.shottype = 'LeftFoot' AND s.shotresult = 'Goal' THEN 1 ELSE 0 END +
            CASE WHEN s.shottype = 'RightFoot' AND s.shotresult = 'Goal' THEN 1 ELSE 0 END) AS asistencias_totales
    FROM shots s
    JOIN players p ON s.shooterid = p.playerid
    GROUP BY s.shooterid, p.name
)
SELECT
    AT.shooterid,
    AT.name,
    AT.asistencias_izquierdas,
    AT.asistencias_derechas,
    AT.asistencias_totales,
    ARRAY(SELECT DISTINCT t."name" FROM teamstats ts JOIN teams t ON ts.teamid = t.teamid WHERE ts.gameid IN (SELECT gameid FROM shots s WHERE s.shooterid = AT.shooterid)) AS equipos
FROM AsistenciasTotales AT
ORDER BY AT.asistencias_totales DESC;
```

Resultado:

shooterid	name	asistencias_izquierdas	asistencias_derechas	asistencias_totales	equipos
2097	Lionel Messi	224	38	262	{Alaves,Almeria,"Athletic Club",
2371	Cristiano Ronaldo	47	138	185	{"AC Milan",Alaves,Almeria,Atala
227	Robert Lewandowski	31	141	172	{"Arminia Bielefeld",Augsburg,B
2098	Luis Suarez	29	126	155	{Alaves,Almeria,"Athletic Club",
318	Pierre-Emerick Aubameyang	29	111	140	{Arsenal,"Aston Villa",Augsburg,
647	Harry Kane	35	103	138	{Arsenal,"Aston Villa",Bournemo
2099	Neymar	37	87	124	{Alaves,Almeria,Amiens,Angers
1250	Mohamed Salah	104	19	123	{"AC Milan",Arsenal,"Aston Villa"
1209	Ciro Immobile	16	104	120	{"AC Milan",Atalanta,Augsburg,"
2270	Antoine Griezmann	102	16	118	{Alaves,Almeria,"Athletic Club",

Conclusión:

Al analizar los datos de los jugadores que destacaron en la cantidad de goles y asistencias durante las temporadas, se identificó a aquellos que lograron las mejores actuaciones. Sin embargo, debido a la falta de una tabla directa que relacione jugadores y equipos en partidos específicos, no es posible determinar con certeza a qué equipo específico pertenecen dichos jugadores en esos momentos particulares. A pesar de esta limitación, se pudo generar una lista detallada de los equipos en los que cada jugador registró goles y asistencias.

Es importante tener en cuenta que un jugador puede haber contribuido tanto como visitante como en calidad de jugador del equipo local en diferentes partidos. Por lo tanto, aunque no se pueda asignar a un jugador a un equipo específico en cada instancia, la lista proporcionada ofrece una visión completa de los equipos con los que estuvieron asociados a lo largo de sus destacadas actuaciones. En este contexto, se puede concluir que estos jugadores, en función de sus contribuciones a varios equipos, sí pertenecen a la categoría de equipos mencionada en el inciso 2.

Según apuestas:

Pregunta 4

Realice un comparativo de las probabilidades de todas las casas de apuesta por temporada, liga y equipo, eliminando aquellos equipos que no tienen estadísticas en ninguna casa de apuesta.

Tome en cuenta de que en la tabla de GAMES se representan los datos de probabilidades de que se gane el local, que gane el extranjero o que empate, según diferentes casas de apuestas como Bet365 (B365), Bet&Win (BW), Interwetten (IW), Ladbrokes (LB), William Hill (WH), VC Bet (VC), etc. Por tanto, escoja el valor más alto de estas columnas.

Luego de obtener las probabilidades correctas, escoja la que mejor le convenga para determinar qué equipo tiene la mayor probabilidad de ganar en qué liga de qué temporada.

Tome en cuenta que los valores que aparecen en las columnas (por ejemplo de B365H, B365D y B365A) no son en sí probabilidades porque no se encuentran en el rango entre 0 y 1. Por tanto, para obtener las probabilidades debe de realizar la división de 1 / b365h, por ejemplo.

Este es un mecanismo que usan las casas de apuesta para poder confundir al jugador.

Pre-Respuesta (ejemplo casa Bet365):

Utilizando de base el siguiente query para determinar la máxima probabilidad de apuesta por cada casa.

Se agregó un ordenamiento en base a la probabilidad máxima de la casa de apuesta y comparar los resultados por cada casa.

Query:

```

WITH RankedGames AS (
    SELECT
        leagueid,
        season,
        hometeamid,
        awayteamid,
        GREATEST(1 / b365h, 1 / b365d, 1 / b365a) AS max_probabilidad_b365,
        CASE
            WHEN 1 / b365h = GREATEST(1 / b365h, 1 / b365d, 1 / b365a) THEN 'Local'
            WHEN 1 / b365d = GREATEST(1 / b365h, 1 / b365d, 1 / b365a) THEN 'Empate'
            WHEN 1 / b365a = GREATEST(1 / b365h, 1 / b365d, 1 / b365a) THEN 'Visitante'
        END AS tipo_apuesta_b365,
        RANK() OVER (PARTITION BY leagueid, season ORDER BY GREATEST(1 / b365h, 1 / b365d, 1 / b365a) DESC) AS rank_b365
    FROM
        games
    WHERE
        b365h IS NOT NULL AND b365h <> 0 AND
        b365d IS NOT NULL AND b365d <> 0 AND
        b365a IS NOT NULL AND b365a <> 0
)
SELECT
    leagueid,
    season,
    hometeamid,
    awayteamid,
    max_probabilidad_b365,
    tipo_apuesta_b365
FROM
    RankedGames
WHERE
    rank_b365 = 1;

```

Resultado:

	leagueid bigint	season bigint	hometeamid bigint	awayteamid bigint	max_probabilidad_b365 double precision	tipo_apuesta_b365 text
1	1	2014	80	202	0.8547008547008548	Local
2	1	2015	83	71	0.8547008547008548	Local
3	1	2016	83	77	0.9174311926605504	Local
4	1	2017	88	81	0.9259259259259258	Local
5	1	2017	88	219	0.9259259259259258	Local
6	1	2018	88	227	0.9433962264150942	Local
7	1	2019	88	74	0.9345794392523364	Local
8	1	2020	88	76	0.9174311926605504	Local
9	2	2014	98	203	0.8849557522123894	Local
10	2	2015	105	112	0.9090909090909091	Local
11	2	2016	95	101	0.9433962264150942	Local
12	2	2017	98	222	0.9433962264150942	Local
13	2	2018	98	109	0.9009009009009008	Local
14	2	2019	98	243	0.8928571428571428	Local

Total rows: 48 of 48

Query complete 00:00:00.128

Casa de apuestas Bet365:

Query:

```
WITH RankedGames AS (
    SELECT
        leagueid,
        season,
        hometeamid,
        awayteamid,
        GREATEST(1 / b365h, 1 / b365d, 1 / b365a) AS max_probabilidad_b365,
        CASE
            WHEN 1 / b365h = GREATEST(1 / b365h, 1 / b365d, 1 / b365a) THEN 'Local'
            WHEN 1 / b365d = GREATEST(1 / b365h, 1 / b365d, 1 / b365a) THEN 'Empate'
            WHEN 1 / b365a = GREATEST(1 / b365h, 1 / b365d, 1 / b365a) THEN 'Visitante'
        END AS tipo_apuesta_b365,
        RANK() OVER (PARTITION BY leagueid, season ORDER BY GREATEST(1 / b365h, 1 / b365d, 1 / b365a) DESC) AS rank_b365
    FROM
        games
    WHERE
        b365h IS NOT NULL AND b365h <> 0 AND
        b365d IS NOT NULL AND b365d <> 0 AND
        b365a IS NOT NULL AND b365a <> 0
)
SELECT
    rg.leagueid,
    rg.season,
    rg.hometeamid,
    th.name AS hometeam_name,
    rg.awayteamid,
    ta.name AS awayteam_name,
    rg.max_probabilidad_b365,
    rg.tipo_apuesta_b365
FROM
    RankedGames rg
JOIN teams th ON rg.hometeamid = th.teamid
JOIN teams ta ON rg.awayteamid = ta.teamid
WHERE
    rg.rank_b365 = 1
ORDER BY rg.max_probabilidad_b365 DESC;
```

Resultado:

	leagueid bigint	season bigint	hometeamid bigint	hometeam_name text	awayteamid bigint	awayteam_name text	max_probabilidad_b365 double precision	tipo_apuesta_b365 text
1	4	2016	148	Barcelona	157	Osasuna	0.9803921568627451	Local
2	3	2018	117	Bayern Munich	128	Hannover 96	0.970873786407767	Local
3	5	2017	161	Paris Saint Germain	172	Troyes	0.9615384615384615	Local
4	4	2015	148	Barcelona	141	Espanyol	0.9615384615384615	Local
5	4	2015	148	Barcelona	149	Sporting Gijon	0.9615384615384615	Local
6	4	2014	148	Barcelona	142	Getafe	0.9615384615384615	Local
7	4	2017	150	Real Madrid	151	Levante	0.9523809523809523	Local
8	4	2017	148	Barcelona	151	Levante	0.9523809523809523	Local
9	5	2020	161	Paris Saint Germain	235	Nimes	0.9523809523809523	Local
10	3	2019	117	Bayern Munich	205	Paderborn	0.9523809523809523	Local

Conclusión:

En la casa de apuestas Bet365, en el enfrentamiento entre los equipos "Barcelona" y "Osasuna", se registró una probabilidad máxima de apuesta de 0.98, correspondiente al tipo de apuesta local.

Casa de apuestas Bet&Win:

Query:

```
WITH RankedGames AS (
    SELECT
        leagueid,
        season,
        hometeamid,
        awayteamid,
        GREATEST(1 / bwh, 1 / bwd, 1 / bwa) AS max_probabilidad_bw,
        CASE
            WHEN 1 / bwh = GREATEST(1 / bwh, 1 / bwd, 1 / bwa) THEN 'Local'
            WHEN 1 / bwd = GREATEST(1 / bwh, 1 / bwd, 1 / bwa) THEN 'Empate'
            WHEN 1 / bwa = GREATEST(1 / bwh, 1 / bwd, 1 / bwa) THEN 'Visitante'
        END AS tipo_apuesta_bw,
        RANK() OVER (PARTITION BY leagueid, season ORDER BY GREATEST(1 / bwh, 1 / bwd, 1 / bwa) DESC) AS rank_bw
    FROM
        games
    WHERE
        bwh IS NOT NULL AND bwh <> 0 AND
        bwd IS NOT NULL AND bwd <> 0 AND
        bwa IS NOT NULL AND bwa <> 0
)
SELECT
    rg.leagueid,
    rg.season,
    rg.hometeamid,
    th.name AS hometeam_name,
    rg.awayteamid,
    ta.name AS awayteam_name,
    rg.max_probabilidad_bw,
    rg.tipo_apuesta_bw
FROM
    RankedGames rg
JOIN teams th ON rg.hometeamid = th.teamid
JOIN teams ta ON rg.awayteamid = ta.teamid
WHERE
    rg.rank_bw = 1
ORDER BY rg.max_probabilidad_bw DESC;
```

Resultado:

	leagueid bigint	season bigint	hometeamid bigint	hometeam_name text	awayteamid bigint	awayteam_name text	max_probabilidad_bw double precision	tipo_apuesta_bw text
1	4	2015	148	Barcelona	141	Espanyol	0.970873786407767	Local
2	4	2016	148	Barcelona	157	Olasuna	0.970873786407767	Local
3	3	2018	117	Bayern Munich	128	Hannover 96	0.970873786407767	Local
4	5	2017	161	Paris Saint Germain	172	Troyes	0.9615384615384615	Local
5	1	2019	88	Manchester City	79	Norwich	0.9523809523809523	Local
6	4	2017	148	Barcelona	151	Levante	0.9523809523809523	Local
7	5	2020	161	Paris Saint Germain	181	Dijon	0.9523809523809523	Local
8	1	2018	87	Liverpool	219	Huddersfield	0.9523809523809523	Local
9	4	2014	148	Barcelona	151	Levante	0.9523809523809523	Local
10	4	2014	148	Barcelona	142	Getafe	0.9523809523809523	Local

Conclusión:

En la casa de apuestas Bet&Win, en el enfrentamiento entre los equipos "Barcelona" y "Espanyol", se registró una probabilidad máxima de apuesta de 0.97, correspondiente al tipo de apuesta local.

Casa de apuestas Interwetten:

Query:

```
WITH RankedGames AS (
    SELECT
        leagueid,
        season,
        hometeamid,
        awayteamid,
        GREATEST(1 / iwh, 1 / iwd, 1 / iwa) AS max_probabilidad_iw,
        CASE
            WHEN 1 / iwh = GREATEST(1 / iwh, 1 / iwd, 1 / iwa) THEN 'Local'
            WHEN 1 / iwd = GREATEST(1 / iwh, 1 / iwd, 1 / iwa) THEN 'Empate'
            WHEN 1 / iwa = GREATEST(1 / iwh, 1 / iwd, 1 / iwa) THEN 'Visitante'
        END AS tipo_apuesta_iw,
        RANK() OVER (PARTITION BY leagueid, season ORDER BY GREATEST(1 / iwh, 1 / iwd, 1 / iwa) DESC) AS rank_iw
    FROM
        games
    WHERE
        iwh IS NOT NULL AND iwh <> 0 AND
        iwd IS NOT NULL AND iwd <> 0 AND
        iwa IS NOT NULL AND iwa <> 0
)
SELECT
    rg.leagueid,
    rg.season,
    rg.hometeamid,
    th.name AS hometeam_name,
    rg.awayteamid,
    ta.name AS awayteam_name,
    rg.max_probabilidad_iw,
    rg.tipo_apuesta_iw
FROM
    RankedGames rg
JOIN teams th ON rg.hometeamid = th.teamid
JOIN teams ta ON rg.awayteamid = ta.teamid
WHERE
    rg.rank_iw = 1
ORDER BY rg.max_probabilidad_iw DESC;
```

Resultado:

	leagueid bigint	season bigint	hometeamid bigint	hometeam_name text	awayteamid bigint	awayteam_name text	max_probabilidad_iw double precision	tipo_apuesta_iw text
1	4	2016	148	Barcelona	157	Osasuna	0.970873786407767	Local
2	4	2014	148	Barcelona	142	Getafe	0.9523809523809523	Local
3	4	2014	150	Real Madrid	156	Eibar	0.9523809523809523	Local
4	4	2014	148	Barcelona	208	Almeria	0.9523809523809523	Local
5	4	2014	148	Barcelona	151	Levante	0.9523809523809523	Local
6	4	2015	150	Real Madrid	144	Las Palmas	0.9523809523809523	Local
7	4	2015	148	Barcelona	149	Sporting Gijon	0.9523809523809523	Local
8	4	2015	148	Barcelona	141	Espanyol	0.9523809523809523	Local
9	4	2015	150	Real Madrid	155	Granada	0.9523809523809523	Local
10	5	2017	161	Paris Saint Germain	172	Troyes	0.9523809523809523	Local

Conclusión:

En la casa de apuestas Interwetten, en el enfrentamiento entre los equipos "Barcelona" y "Osasuna", se registró una probabilidad máxima de apuesta de 0.97, correspondiente al tipo de apuesta local.

Casa de apuestas Pinnacle Sports:

Query:

```
WITH RankedGames AS (
    SELECT
        leagueid,
        season,
        hometeamid,
        awayteamid,
        GREATEST(1 / psh, 1 / psd, 1 / psa) AS max_probabilidad_ps,
        CASE
            WHEN 1 / psh = GREATEST(1 / psh, 1 / psd, 1 / psa) THEN 'Local'
            WHEN 1 / psd = GREATEST(1 / psh, 1 / psd, 1 / psa) THEN 'Empate'
            WHEN 1 / psa = GREATEST(1 / psh, 1 / psd, 1 / psa) THEN 'Visitante'
        END AS tipo_apuesta_ps,
        RANK() OVER (PARTITION BY leagueid, season ORDER BY GREATEST(1 / psh, 1 / psd, 1 / psa) DESC) AS rank_ps
    FROM
        games
    WHERE
        psh IS NOT NULL AND psh <> 0 AND
        psd IS NOT NULL AND psd <> 0 AND
        psa IS NOT NULL AND psa <> 0
)
SELECT
    rg.leagueid,
    rg.season,
    rg.hometeamid,
    th.name AS hometeam_name,
    rg.awayteamid,
    ta.name AS awayteam_name,
    rg.max_probabilidad_ps,
    rg.tipo_apuesta_ps
FROM
    RankedGames rg
JOIN teams th ON rg.hometeamid = th.teamid
JOIN teams ta ON rg.awayteamid = ta.teamid
WHERE
    rg.rank_ps = 1
ORDER BY rg.max_probabilidad_ps DESC;
```

Resultado:

	leagueid bigint	season bigint	hometeamid bigint	hometeam_name text	awayteamid bigint	awayteam_name text	max_probabilidad_ps double precision	tipo_apuesta_ps text
1	4	2015	148	Barcelona	141	Espanyol	0.9615384615384615	Local
2	4	2016	148	Barcelona	157	Olasuna	0.9523809523809523	Local
3	4	2017	148	Barcelona	151	Levante	0.9523809523809523	Local
4	4	2014	148	Barcelona	140	Real Sociedad	0.9433962264150942	Local
5	4	2014	148	Barcelona	142	Getafe	0.9433962264150942	Local
6	2	2017	98	Juventus	222	Benevento	0.9433962264150942	Local
7	5	2020	161	Paris Saint Germain	181	Dijon	0.9433962264150942	Local
8	3	2019	117	Bayern Munich	121	Augsburg	0.9345794392523364	Local
9	1	2018	88	Manchester City	227	Cardiff	0.9345794392523364	Local
10	3	2019	117	Bayern Munich	233	Fortuna Duesseldorf	0.9345794392523364	Local

Conclusión:

En la casa de apuestas Pinnacle Sports, en el enfrentamiento entre los equipos "Barcelona" y "Espanyol", se registró una probabilidad máxima de apuesta de 0.96, correspondiente al tipo de apuesta local.

Casa de apuestas William Hill:

Query:

```
WITH RankedGames AS (
    SELECT
        leagueid,
        season,
        hometeamid,
        awayteamid,
        GREATEST(1 / whh, 1 / whd, 1 / wha) AS max_probabilidad_wh,
        CASE
            WHEN 1 / whh = GREATEST(1 / whh, 1 / whd, 1 / wha) THEN 'Local'
            WHEN 1 / whd = GREATEST(1 / whh, 1 / whd, 1 / wha) THEN 'Empate'
            WHEN 1 / wha = GREATEST(1 / whh, 1 / whd, 1 / wha) THEN 'Visitante'
        END AS tipo_apuesta_wh,
        RANK() OVER (PARTITION BY leagueid, season ORDER BY GREATEST(1 / whh, 1 / whd, 1 / wha) DESC) AS rank_wh
    FROM
        games
    WHERE
        whh IS NOT NULL AND whh <> 0 AND
        whd IS NOT NULL AND whd <> 0 AND
        wha IS NOT NULL AND wha <> 0
)
SELECT
    rg.leagueid,
    rg.season,
    rg.hometeamid,
    th.name AS hometeam_name,
    rg.awayteamid,
    ta.name AS awayteam_name,
    rg.max_probabilidad_wh,
    rg.tipo_apuesta_wh
FROM
    RankedGames rg
JOIN teams th ON rg.hometeamid = th.teamid
JOIN teams ta ON rg.awayteamid = ta.teamid
WHERE
    rg.rank_wh = 1
ORDER BY rg.max_probabilidad_wh DESC;
```

Resultado:

	leagueid bigint	season bigint	hometeamid bigint	hometeam_name text	awayteamid bigint	awayteam_name text	max_probabilidad_wh double precision	tipo_apuesta_wh text
1	3	2018	117	Bayern Munich	128	Hannover 96	0.9803921568627451	Local
2	4	2015	148	Barcelona	155	Granada	0.9803921568627451	Local
3	5	2020	161	Paris Saint Germain	181	Dijon	0.970873786407767	Local
4	5	2017	161	Paris Saint Germain	172	Troyes	0.9615384615384615	Local
5	4	2016	148	Barcelona	157	Olasuna	0.9615384615384615	Local
6	2	2018	98	Juventus	109	Chievo	0.9615384615384615	Local
7	4	2014	148	Barcelona	142	Getafe	0.9523809523809523	Local
8	4	2018	148	Barcelona	232	SD Huesca	0.9523809523809523	Local
9	2	2017	98	Juventus	222	Benevento	0.9523809523809523	Local
10	5	2018	161	Paris Saint Germain	167	Angers	0.9523809523809523	Local

Conclusión:

En la casa de apuestas William Hill, en el enfrentamiento entre los equipos "Bayern Munich" y "Hannover 96", se registró una probabilidad máxima de apuesta de 0.98, correspondiente al tipo de apuesta local.

Casa de apuestas VC Bet:

Query:

```
WITH RankedGames AS (
    SELECT
        leagueid,
        season,
        hometeamid,
        awayteamid,
        GREATEST(1 / vch, 1 / vcd, 1 / vca) AS max_probabilidad_vc,
        CASE
            WHEN 1 / vch = GREATEST(1 / vch, 1 / vcd, 1 / vca) THEN 'Local'
            WHEN 1 / vcd = GREATEST(1 / vch, 1 / vcd, 1 / vca) THEN 'Empate'
            WHEN 1 / vca = GREATEST(1 / vch, 1 / vcd, 1 / vca) THEN 'Visitante'
        END AS tipo_apuesta_vc,
        RANK() OVER (PARTITION BY leagueid, season ORDER BY GREATEST(1 / vch, 1 / vcd, 1 / vca) DESC) AS rank_vc
    FROM
        games
    WHERE
        vch IS NOT NULL AND vch <> 0 AND
        vcd IS NOT NULL AND vcd <> 0 AND
        vca IS NOT NULL AND vca <> 0
)
SELECT
    rg.leagueid,
    rg.season,
    rg.hometeamid,
    th.name AS hometeam_name,
    rg.awayteamid,
    ta.name AS awayteam_name,
    rg.max_probabilidad_vc,
    rg.tipo_apuesta_vc
FROM
    RankedGames rg
JOIN teams th ON rg.hometeamid = th.teamid
JOIN teams ta ON rg.awayteamid = ta.teamid
WHERE
    rg.rank_vc = 1
ORDER BY rg.max_probabilidad_vc DESC;
```

Resultado:

	leagueid bigint	season bigint	hometeamid bigint	hometeam_name text	awayteamid bigint	awayteam_name text	max_probabilidad_vc double precision	tipo_apuesta_vc text
1	3	2018	117	Bayern Munich	128	Hannover 96	0.970873786407767	Local
2	2	2017	98	Juventus	222	Benevento	0.970873786407767	Local
3	4	2015	148	Barcelona	141	Espanyol	0.970873786407767	Local
4	4	2014	148	Barcelona	142	Getafe	0.970873786407767	Local
5	5	2020	161	Paris Saint Germain	181	Dijon	0.9615384615384615	Local
6	2	2016	95	Roma	101	Genoa	0.9615384615384615	Local
7	5	2017	161	Paris Saint Germain	172	Troyes	0.9615384615384615	Local
8	4	2017	148	Barcelona	151	Levante	0.9615384615384615	Local
9	4	2016	148	Barcelona	156	Eibar	0.9523809523809523	Local
10	4	2016	150	Real Madrid	149	Sporting Gijon	0.9523809523809523	Local

Conclusión:

En la casa de apuestas VC Bet, en el enfrentamiento entre los equipos "Bayern Munich" y "Hannover 96", se registró una probabilidad máxima de apuesta de 0.97, correspondiente al tipo de apuesta local.

Pregunta 5

¿Cuál es el mejor equipo de todas las ligas y de todas las temporadas según las apuestas?

Hint: Apóyese o complemente el query del inciso anterior para obtener este.

Casa de apuestas Bet365:

Resultado:

	leagueid bigint	season bigint	hometeamid bigint	hometeam_name text	awayteamid bigint	awayteam_name text	max_probabilidad_b365 double precision	tipo_apuesta_b365 text
1	4	2016	148	Barcelona	157	Osasuna	0.9803921568627451	Local

Casa de apuestas Bet&Win:

Resultado:

	leagueid bigint	season bigint	hometeamid bigint	hometeam_name text	awayteamid bigint	awayteam_name text	max_probabilidad_bw double precision	tipo_apuesta_bw text
1	4	2015	148	Barcelona	141	Espanyol	0.970873786407767	Local

Casa de apuestas Interwetten:

Resultado:

	leagueid bigint	season bigint	hometeamid bigint	hometeam_name text	awayteamid bigint	awayteam_name text	max_probabilidad_iw double precision	tipo_apuesta_iw text
1	4	2016	148	Barcelona	157	Osasuna	0.970873786407767	Local

Casa de apuestas Pinnacle Sports:

Resultado:

	leagueid bigint	season bigint	hometeamid bigint	hometeam_name text	awayteamid bigint	awayteam_name text	max_probabilidad_ps double precision	tipo_apuesta_ps text
1	4	2015	148	Barcelona	141	Espanyol	0.9615384615384615	Local

Casa de apuestas William Hill:

Resultado:

	leagueid bigint	season bigint	hometeamid bigint	hometeam_name text	awayteamid bigint	awayteam_name text	max_probabilidad_wh double precision	tipo_apuesta_wh text
1	3	2018	117	Bayern Munich	128	Hannover 96	0.9803921568627451	Local

Casa de apuestas VC Bet:

Resultado:

	leagueid bigint	season bigint	hometeamid bigint	hometeam_name text	awayteamid bigint	awayteam_name text	max_probabilidad_vc double precision	tipo_apuesta_vc text
1		3	2018	117	Bayern Munich	128	Hannover 96	0.970873786407767 Local

Conclusión:

En base al resultado del inciso 4 y obtener la primera tupla de cada casa se puede determinar que tanto el equipo “Barcelona” como “Bayer Munich” son los mejores equipos en cuanto probabilidades máxima de apuestas (0.98) de las ligas y temporadas de las casas de apuestas.

Otros:

Pregunta 6

¿Quiénes son los jugadores de cada liga y cada temporada que tienen los mejores atributos – características de juego -pases, goles, etc.? ¿De acuerdo a este inciso, y comparándolo con el inciso 2 y 5 anteriores, alguno de los jugadores más valiosos se encuentra dentro del mejor equipo?

Resultado:

Conclusión:

Los resultados de los jugadores más valiosos concuerdan con los resultados obtenidos tanto en la pregunta 2 como en la pregunta 5. Con la pregunta 5, encontramos que tanto el Barcelona como el Bayern Munich poseen jugadores con características muy buenas, haciendo que estos equipos sean considerados como los mejores de todas las temporadas y ligas. Por otra parte, con respecto a la pregunta 2, vemos que hay equipos que se repiten, como el Barcelona, Real Madrid, Paris Saint Germain, Bayern Munich, siendo en ambos casos, los resultados bastante similares.

Pregunta 7

Obtenga el rendimiento de los equipos en promedio, comparando goles metidos contra la expectativa de goles, determinando qué equipo era quien tenía más expectativa de goles contra quien fue en realidad el que acertó más goles (goals vs expected goals, xgoals) en general, pero también es necesario que lo muestre si dichos equipos jugaron como locales o como extranjeros.

Resultado:

name	location	goles	goles Esperados	diferencia
Bayern Munich	h	3.1428571428571429	2.655358411764707	0.4874987310924359
Barcelona	h	2.9699248120300752	2.680157150375938	0.289767616541376
Paris Saint Germain	h	2.921875000000000	2.6153895390625004	0.3064854609374996
Manchester City	h	2.6015037593984962	2.4543660601503756	0.1471376992481206
Real Madrid	h	2.5714285714285714	2.317695932330827	0.2537326390977448
Borussia Dortmund	h	2.4705682352941176	2.299110294117648	0.17147794117646997
Napoli	h	2.2857142857142857	2.0712432481203007	0.2144710375939849
Barcelona	a	2.2781954887218045	2.130852819548872	0.1473426691729327
Juventus	h	2.2556390977443609	2.0065923233082708	0.24904677443609025
Liverpool	h	2.1729323308270677	1.9656568270676689	0.20727550375939896

Conclusión:

El equipo que tuvo la mayor expectativa de goles fue el Barcelona con un 2.68 en promedio y una diferencia contra los goles metidos de 0.28, jugaba de local. Por otro lado, el equipo que más goles metió fue el Bayern Munich, quien tenía un 3.14 de goles metidos, con una diferencia de 0.48 contra la expectativa de goles., la cual era de 2.65.

Pregunta 8

¿Cuáles son las características/atributos de los equipos que han sido los líderes de sus ligas en las distintas temporadas? ¿Sus comportamientos son similares?

Resultado:

liga	equipo	goles	tiros	tiro_al_arco	temporada	victorias
Bundesliga	Bayern Munich	80	585	235	2014	25
Bundesliga	Bayern Munich	80	627	263	2015	28
Bundesliga	Bayern Munich	89	622	230	2016	25
Bundesliga	Bayern Munich	92	601	237	2017	27
Bundesliga	Bayern Munich	88	635	255	2018	24
Bundesliga	Bayern Munich	100	615	254	2019	26
Bundesliga	Bayern Munich	99	576	234	2020	24
La Liga	Real Madrid	118	686	283	2014	30
La Liga	Barcelona	112	605	266	2015	29
La Liga	Real Madrid	106	663	265	2016	29
La Liga	Barcelona	99	580	272	2017	28
La Liga	Barcelona	90	562	254	2018	26
La Liga	Real Madrid	70	563	236	2019	26
La Liga	Atletico Madrid	67	460	186	2020	26
Ligue 1	Paris Saint Germain	83	481	219	2014	24
Ligue 1	Paris Saint Germain	102	565	252	2015	30
Ligue 1	Monaco	107	555	244	2016	30
Ligue 1	Paris Saint Germain	108	626	253	2017	29
Ligue 1	Paris Saint Germain	105	564	230	2018	29
Ligue 1	Paris Saint Germain	75	448	195	2019	22
Ligue 1	Paris Saint Germain	86	568	215	2020	26
Premier League	Chelsea	73	564	210	2014	26
Premier League	Leicester	68	522	179	2015	23
Premier League	Chelsea	85	580	204	2016	30
Premier League	Manchester City	106	663	264	2017	32
Premier League	Manchester City	95	683	260	2018	32
Premier League	Liverpool	85	591	231	2019	32
Premier League	Manchester City	83	600	219	2020	27
Serie A	Juventus	72	602	201	2014	26
Serie A	Juventus	75	599	215	2015	29
Serie A	Juventus	77	583	220	2016	29
Serie A	Juventus	86	555	207	2017	30
Serie A	Juventus	70	609	200	2018	28
Serie A	Juventus	76	679	243	2019	26
Serie A	Inter	89	553	203	2020	28

Conclusión:

Las características seleccionadas fueron los goles, tiros y tiros al arco, con lo que se observa que todos los equipos tienen una gran cantidad de los mismos, lo cual indica dominancia en el campo. Además, los equipos en sus respectivas ligas, son constantes pues todas las temporadas tienen resultados similares con un gran número de victorias.

Pregunta 9

¿Según la casa de apuesta Beat365 (tome la mejor probabilidad de las 3 medidas), cuales deberían de ser los equipos que tenían la mayor probabilidad de ganar en cada una de las temporadas (seasons)?

Query:

```
WITH RankedGames AS (
    SELECT
        season,
        hometeamid,
        awayteamid,
        GREATEST(1 / b365h, 1 / b365d, 1 / b365a) AS max_probabilidad_b365,
        CASE
            WHEN 1 / b365h = GREATEST(1 / b365h, 1 / b365d, 1 / b365a) THEN 'Local'
            WHEN 1 / b365d = GREATEST(1 / b365h, 1 / b365d, 1 / b365a) THEN 'Empate'
            WHEN 1 / b365a = GREATEST(1 / b365h, 1 / b365d, 1 / b365a) THEN 'Visitante'
        END AS tipo_apuesta_b365,
        RANK() OVER (PARTITION BY season ORDER BY GREATEST(1 / b365h, 1 / b365d, 1 / b365a) DESC) AS rank_b365,
        RANK() OVER (PARTITION BY season ORDER BY GREATEST(1 / b365h, 1 / b365d, 1 / b365a) DESC) AS rank_season
    FROM
        games
    WHERE
        b365h IS NOT NULL AND b365h <> 0 AND
        b365d IS NOT NULL AND b365d <> 0 AND
        b365a IS NOT NULL AND b365a <> 0
)
SELECT
    rg.season,
    rg.hometeamid,
    th.name AS hometeam_name,
    rg.awayteamid,
    ta.name AS awayteam_name,
    rg.max_probabilidad_b365,
    rg.tipo_apuesta_b365
FROM
    RankedGames rg
JOIN teams th ON rg.hometeamid = th.teamid
JOIN teams ta ON rg.awayteamid = ta.teamid
WHERE
    rg.rank_season = 1
ORDER BY 1
```

Resultado:

	season bigint	hometeamid bigint	hometeam_name text	awayteamid bigint	awayteam_name text	max_probabilidad_b365 double precision	tipo_apuesta_b365 text
1	2014	148	Barcelona	142	Getafe	0.9615384615384615	Local
2	2015	148	Barcelona	141	Espanyol	0.9615384615384615	Local
3	2015	148	Barcelona	149	Sporting Gijon	0.9615384615384615	Local
4	2016	148	Barcelona	157	Osasuna	0.9803921568627451	Local
5	2017	161	Paris Saint Germain	172	Troyes	0.9615384615384615	Local
6	2018	117	Bayern Munich	128	Hannover 96	0.970873786407767	Local
7	2019	117	Bayern Munich	205	Paderborn	0.9523809523809523	Local
8	2020	161	Paris Saint Germain	235	Nimes	0.9523809523809523	Local

Conclusión:

Según Beat365, en la temporada 2014 a 2016 el equipo con mayor probabilidad a ganar fue Barcelona, en la temporada 2017 fue Paris Saint Germain, de la temporada 2018 a 2019 fue Bayer Munich y en el 2020 fue Paris Saint Germain.

Pregunta 10

Obtenga el top 10 de estadísticas de los equipos más limpios en jugar (mejor faltas, menos tarjetas amarillas, menos tarjetas rojas) y también el top 10 de los equipos más sucios.

Equipos más limpios:

nombre_equipo	faltas	tarjetas_amarillas	tarjetas_rojas
Nuernberg	376	59	4
Cardiff	384	65	1
Arminia Bielefeld	391	51	1
Leeds	427	61	1
Cadiz	435	80	3
Queens Park Rangers	447	75	3
Lecce	463	99	5
Middlesbrough	478	76	1
Cordoba	494	93	9
Brescia	494	95	6
(10 filas)			

Conclusión:

Como se observa en el query, los equipos con menos faltas son el Nuernberg, Cardiff, Arminia Bielefeld, Leeds, Cadiz, Queens Park Rangers, Lecce, Middlesbrough, Córdoba y Brescia.

Equipos más sucios:

nombre_equipo	faltas	tarjetas_amarillas	tarjetas_rojas
Torino	4131	612	29
Genoa	4074	640	47
Atalanta	3897	558	38
Sevilla	3854	710	26
Fiorentina	3757	637	33
Getafe	3722	714	34
Eibar	3717	625	27
Lazio	3705	663	40
Sassuolo	3664	636	39
AC Milan	3643	596	50
(10 filas)			

Conclusión:

Los equipos que cometieron más faltas fueron Torino, Genoa, Atalanta, Sevilla, Fiorentina, Getafe, Eibar, Lazio, Sassuolo y el AC Milan.

Etapa 3

Criterio

¿Cuál es el equipo más adecuado para realizar apuestas, considerando la combinación de la probabilidad máxima de apuesta en diferentes casas de apuestas, la consistencia de esta probabilidad a lo largo de múltiples temporadas y la presencia del equipo como único representante en su respectiva liga?

Sub-Criterios

- **Probabilidad Máxima:** Seleccionar equipos que tienen la probabilidad máxima de apuesta en las diferentes casas de apuestas.
- **Consistencia Temporal:** Evaluar equipos que mantienen esta probabilidad máxima a lo largo de múltiples temporadas, indicando cierta consistencia en su rendimiento percibido.
- **Presencia Única en la Liga:** Seleccionar equipos que son el único representante de su liga en una temporada específica, lo que sugiere una mayor probabilidad de que ese equipo resulte ganador en esa liga.

Explicación

Para desarrollar una estrategia efectiva en la selección de equipos para apuestas, se ha seguido un criterio que aborda distintos aspectos clave. Inicialmente, se calculó la probabilidad máxima de apuesta para cada tipo (local, empate, visitante) en diversas casas de apuestas. Un detalle esencial en este análisis fue la identificación de la probabilidad máxima para el tipo de apuesta "local" en cada partido, según lo revelado en la *pregunta 4, etapa 2*.

Justificación

Al abordar la decisión de apostar por un equipo, se ha optado por un enfoque específico basado en la percepción de las casas de apuestas. A diferencia de considerar únicamente el rendimiento histórico o la posición actual en la temporada, las cuotas de las casas de apuestas ofrecen una visión más precisa de la probabilidad percibida de cada resultado. Esta decisión se fundamenta en la exhaustividad del análisis y la experiencia de los corredores para colocar las cuotas/probabilidades en cada casa de apuestas.

Es importante destacar que, al centrarnos en ligas específicas en lugar de abordar toda una temporada, buscamos evitar la dilución de la variabilidad en el rendimiento de los equipos. Apostar por toda una temporada podría oscurecer patrones consistentes y factores clave que podrían influir en los resultados.

Query (Ejemplo para la casa Bet365):

```

WITH RankedGames AS (
    SELECT
        leagueid,
        season,
        hometeamid,
        awayteamid,
        GREATEST(1 / b365h, 1 / b365d, 1 / b365a) AS max_probabilidad_b365,
        CASE
            WHEN 1 / b365h = GREATEST(1 / b365h, 1 / b365d, 1 / b365a) THEN 'Local'
            WHEN 1 / b365d = GREATEST(1 / b365h, 1 / b365d, 1 / b365a) THEN 'Empate'
            WHEN 1 / b365a = GREATEST(1 / b365h, 1 / b365d, 1 / b365a) THEN 'Visitante'
        END AS tipo_apuesta_b365,
        RANK() OVER (PARTITION BY leagueid, season ORDER BY GREATEST(1 / b365h, 1 / b365d, 1 / b365a) DESC) AS rank_b365
    FROM
        games
    WHERE
        b365h IS NOT NULL AND b365h <> 0 AND
        b365d IS NOT NULL AND b365d <> 0 AND
        b365a IS NOT NULL AND b365a <> 0
)
SELECT
    rk.leagueid,
    ls.name AS nombre_liga,
    rk.hometeamid AS nombre_equipo,
    tn.name,
    rk.tipo_apuesta_b365,
    count(*) AS presencia
FROM
    RankedGames rk
    JOIN teams tn ON rk.hometeamid = tn.teamid
    JOIN leagues ls ON rk.leagueid = ls.leagueid
WHERE
    rank_b365 = 1
GROUP BY rk.leagueid,
    ls.name,
    rk.hometeamid,
    tn.name,
    rk.tipo_apuesta_b365
ORDER BY 1,5 DESC;

```

Resultado (Bet365):

	leagueid bigint	nombre_liga text	nombre_equipo bigint	name text	tipo_apuesta_b365 text	presencia bigint
1	1	Premier League	80	Chelsea	Local	1
2	1	Premier League	83	Arsenal	Local	2
3	1	Premier League	88	Manchester City	Local	5
4	2	Serie A	95	Roma	Local	1
5	2	Serie A	98	Juventus	Local	4
6	2	Serie A	105	Napoli	Local	1
7	2	Serie A	107	Atalanta	Local	1
8	3	Bundesliga	117	Bayern Munich	Local	11
9	4	La Liga	148	Barcelona	Local	11
10	4	La Liga	150	Real Madrid	Local	2
11	5	Ligue 1	161	Paris Saint Germain	Local	9

Resultado (Betway):

	leagueid bigint	nombre_liga text	nombre_equipo bigint	name text	tipo_apuesta_bw text	presencia bigint
1	1	Premier League	80	Chelsea	Local	2
2	1	Premier League	83	Arsenal	Local	2
3	1	Premier League	87	Liverpool	Local	1
4	1	Premier League	88	Manchester City	Local	5
5	2	Serie A	95	Roma	Local	1
6	2	Serie A	98	Juventus	Local	5
7	2	Serie A	105	Napoli	Local	1
8	3	Bundesliga	117	Bayern Munich	Local	9
9	4	La Liga	148	Barcelona	Local	9
10	5	Ligue 1	161	Paris Saint Germain	Local	8

Resultado (Interwetten):

	leagueid bigint	nombre_liga text	nombre_equipo bigint	name text	tipo_apuesta_iw text	presencia bigint
1	1	Premier League	80	Chelsea	Local	2
2	1	Premier League	83	Arsenal	Local	2
3	1	Premier League	87	Liverpool	Local	1
4	1	Premier League	88	Manchester City	Local	9
5	2	Serie A	95	Roma	Local	1
6	2	Serie A	98	Juventus	Local	5
7	2	Serie A	105	Napoli	Local	2
8	2	Serie A	107	Atalanta	Local	3
9	3	Bundesliga	117	Bayern Munich	Local	8
10	4	La Liga	148	Barcelona	Local	13
11	4	La Liga	150	Real Madrid	Local	4
12	5	Ligue 1	161	Paris Saint Germain	Local	14

Resultado (Pinnacle Sports):

	leagueid bigint	nombre_liga text	nombre_equipo bigint	name text	tipo_apuesta_ps text	presencia bigint
1	1	Premier League	83	Arsenal	Local	2
2	1	Premier League	88	Manchester City	Local	7
3	2	Serie A	98	Juventus	Local	4
4	2	Serie A	105	Napoli	Local	2
5	2	Serie A	107	Atalanta	Local	1
6	3	Bundesliga	117	Bayern Munich	Local	9
7	4	La Liga	148	Barcelona	Local	9
8	5	Ligue 1	161	Paris Saint Germain	Local	8

Resultado (William Hill):

	leagueid bigint	nombre_liga text	nombre_equipo bigint	name text	tipo_apuesta_wh text	presencia bigint
1	1	Premier League	80	Chelsea	Local	1
2	1	Premier League	83	Arsenal	Local	2
3	1	Premier League	87	Liverpool	Local	2
4	1	Premier League	88	Manchester City	Local	7
5	2	Serie A	95	Roma	Local	1
6	2	Serie A	98	Juventus	Local	5
7	2	Serie A	105	Napoli	Local	1
8	2	Serie A	107	Atalanta	Local	1
9	3	Bundesliga	117	Bayern Munich	Local	14
10	4	La Liga	148	Barcelona	Local	11
11	4	La Liga	150	Real Madrid	Local	1
12	5	Ligue 1	161	Paris Saint Germain	Local	11

Resultado (VC Bet):

	leagueid bigint	nombre_liga text	nombre_equipo bigint	name text	tipo_apuesta_vc text	presencia bigint
1	1	Premier League	83	Arsenal	Local	2
2	1	Premier League	87	Liverpool	Local	1
3	1	Premier League	88	Manchester City	Local	5
4	2	Serie A	95	Roma	Local	1
5	2	Serie A	98	Juventus	Local	4
6	2	Serie A	105	Napoli	Local	2
7	2	Serie A	107	Atalanta	Local	2
8	3	Bundesliga	117	Bayern Munich	Local	10
9	4	La Liga	148	Barcelona	Local	10
10	4	La Liga	150	Real Madrid	Local	2
11	5	Ligue 1	161	Paris Saint Germain	Local	9

Resultados

Cuadro 1. Probabilidad de Apuesta Máxima por Temporada en Be56.

Temporada	Equipo Local	Probabilidad de Apuesta
2014	Barcelona	0.961538462
2015	Barcelona	0.961538462
2015	Barcelona	0.961538462
2016	Barcelona	0.980392157
2017	Paris Saint Germain	0.961538462
2018	Bayern Munich	0.970873786
2019	Bayern Munich	0.952380952
2020	Paris Saint Germain	0.952380952

Fuente: Elaboración propia, Datos: Say y Cruz, 2024.

Cuadro 2. Participación de Equipos con Probabilidad de Apuestas
Máximas por Liga en Be56.

Liga	Equipo	Participación
Premier League	Chelsea	1
Premier League	Arsenal	2
Premier League	Manchester City	5
Serie A	Roma	1
Serie A	Juventus	4
Serie A	Napoli	1
Serie A	Atalanta	1
Bundesliga	Bayern Munich	11
La Liga	Barcelona	11
La Liga	Real Madrid	2
Ligue 1	Paris Saint Germain	9

Fuente: Elaboración propia, Datos: Say y Cruz, 2024.

Cuadro 3. Participación de Equipos con Probabilidad de Apuestas MÁXIMAS por Liga en Betway.

Liga	Equipo	Participación
Premier League	Chelsea	2
Premier League	Arsenal	2
Premier League	Liverpool	1
Premier League	Manchester City	5
Serie A	Roma	1
Serie A	Juventus	5
Serie A	Napoli	1
Bundesliga	Bayern Munich	9
La Liga	Barcelona	9
Ligue 1	Paris Saint Germain	8

Fuente: Elaboración propia, Datos: Say y Cruz, 2024.

Cuadro 4. Participación de Equipos con Probabilidad de Apuestas Máximas por Liga en Interwetten.

Liga	Equipo	Participación
Premier League	Chelsea	2
Premier League	Arsenal	2
Premier League	Liverpool	1
Premier League	Manchester City	9
Serie A	Roma	1
Serie A	Juventus	5
Serie A	Napoli	2
Serie A	Atalanta	3
Bundesliga	Bayern Munich	8
La Liga	Barcelona	13
La Liga	Real Madrid	4
Ligue 1	Paris Saint Germain	14

Fuente: Elaboración propia, Datos: Say y Cruz, 2024.

Cuadro 5. Participación de Equipos con Probabilidad de Apuestas Máximas por Liga en Pinnacle Sports.

Liga	Equipo	Participación
Premier League	Arsenal	2
Premier League	Manchester City	7
Serie A	Juventus	4
Serie A	Napoli	2
Serie A	Atalanta	1
Bundesliga	Bayern Munich	9
La Liga	Barcelona	9
Ligue 1	Paris Saint Germain	8

Fuente: Elaboración propia, Datos: Say y Cruz, 2024.

Cuadro 6. Participación de Equipos con Probabilidad de Apuestas Máximas por Liga en William Hill.

Liga	Equipo	Participación
Premier League	Chelsea	1
Premier League	Arsenal	2
Premier League	Liverpool	2
Premier League	Manchester City	7
Serie A	Roma	1
Serie A	Juventus	5
Serie A	Napoli	1
Serie A	Atalanta	1
Bundesliga	Bayern Munich	14
La Liga	Barcelona	11
La Liga	Real Madrid	1
Ligue 1	Paris Saint Germain	11

Fuente: Elaboración propia, Datos: Say y Cruz, 2024.

Cuadro 7. Participación de Equipos con Probabilidad de Apuestas Máximas por Liga en VC Bet.

Liga	Equipo	Participación
Premier League	Arsenal	2
Premier League	Liverpool	1
Premier League	Manchester City	5
Serie A	Roma	1
Serie A	Juventus	4
Serie A	Napoli	2
Serie A	Atalanta	2
Bundesliga	Bayern Munich	10
La Liga	Barcelona	10
La Liga	Real Madrid	2
Ligue 1	Paris Saint Germain	9

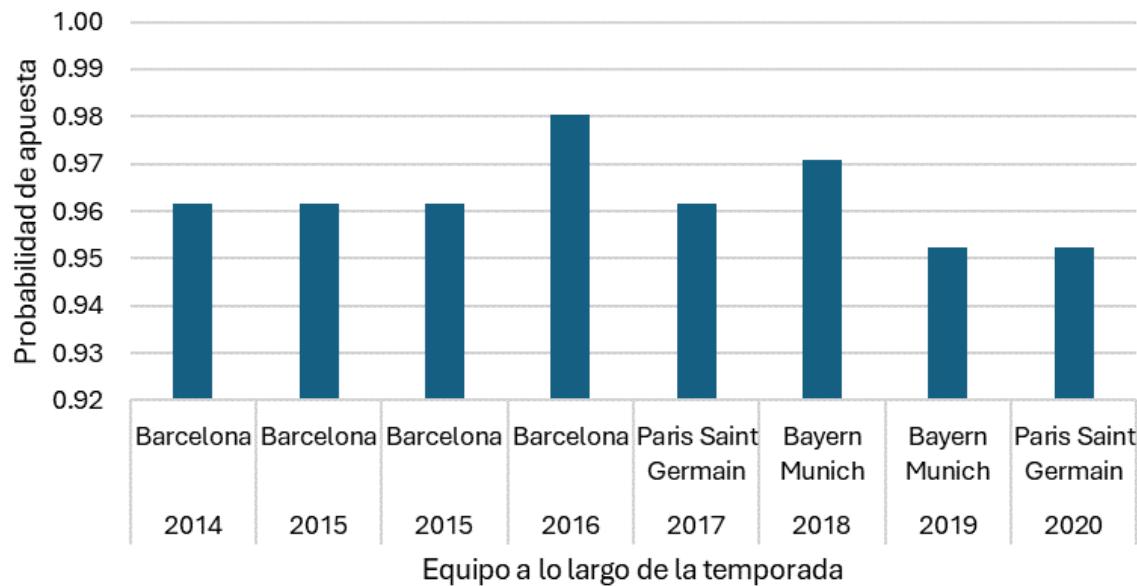
Fuente: Elaboración propia, Datos: Say y Cruz, 2024.

Cuadro 8. Equipos Candidatos para Apuestas basada en Probabilidades Máximas de Todas las Ligas y Casas de Apuestas.

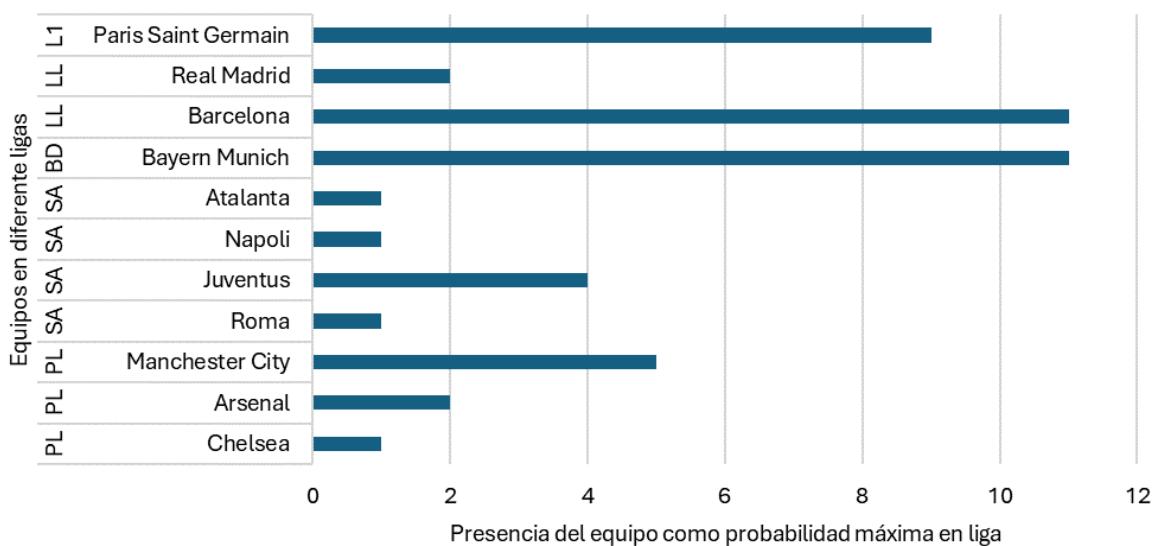
Liga	Equipo	Participación
Ligue 1	Paris Saint Germain	14
Bundesliga	Bayern Munich	14
Bundesliga	Bayern Munich	11
Ligue 1	Paris Saint Germain	11
Bundesliga	Bayern Munich	10
Ligue 1	Paris Saint Germain	9
Bundesliga	Bayern Munich	9
La Liga	Barcelona	9
Bundesliga	Bayern Munich	9
La Liga	Barcelona	9
Ligue 1	Paris Saint Germain	9
Ligue 1	Paris Saint Germain	8
Bundesliga	Bayern Munich	8

Fuente: Elaboración propia, Datos: Say y Cruz, 2024.

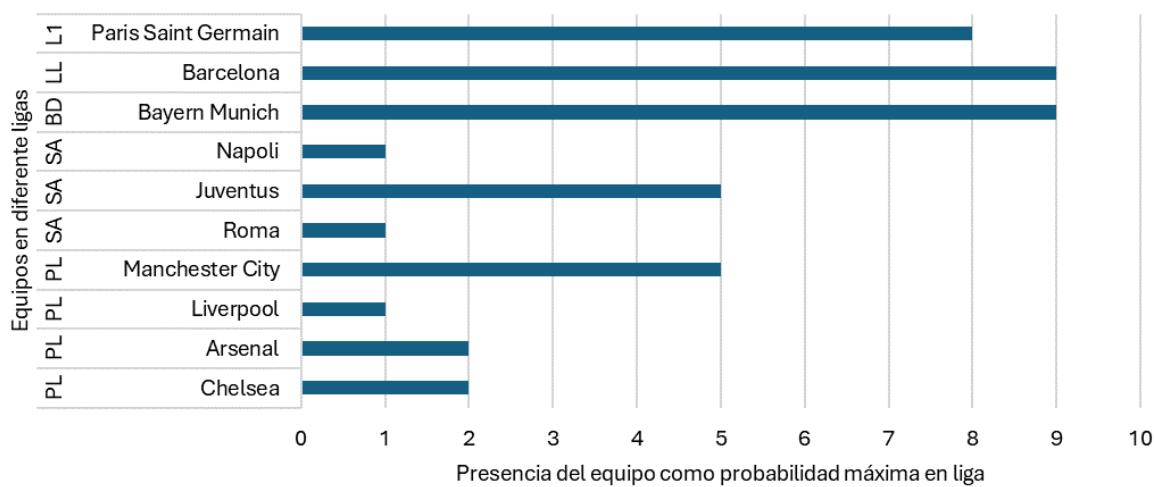
Gráfica 1. Apuestas Máximas por Temporada en Be56.



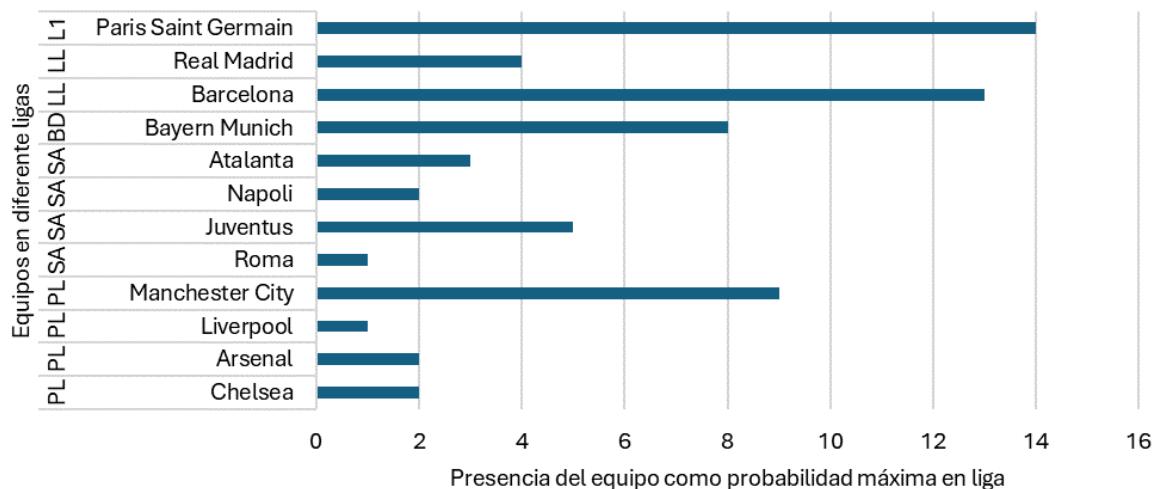
Gráfica 2. Participación de Equipos en Apuestas Máximas por Liga en Be56.



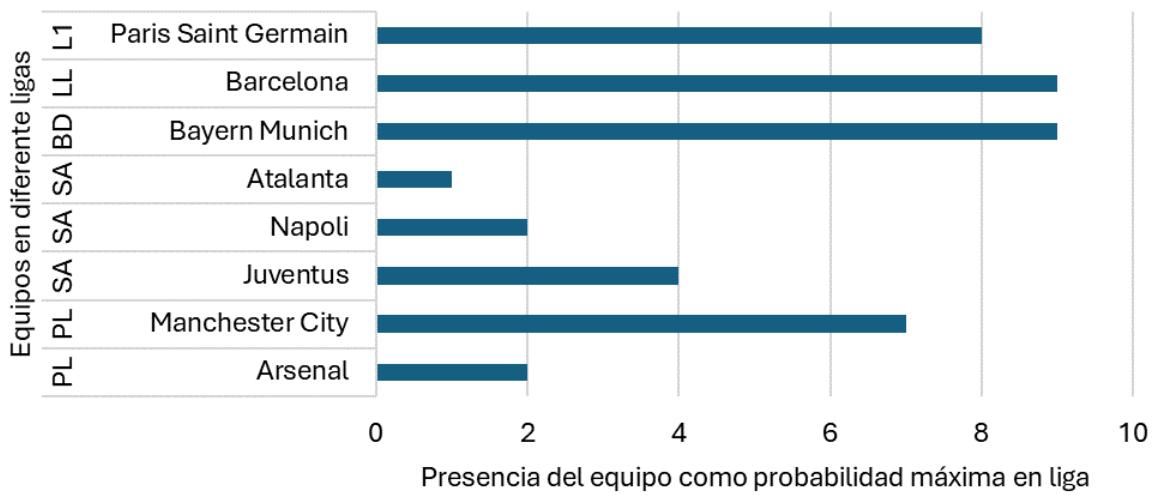
Gráfica 3. Participación de Equipos en Apuestas Máximas por Liga en Betway.



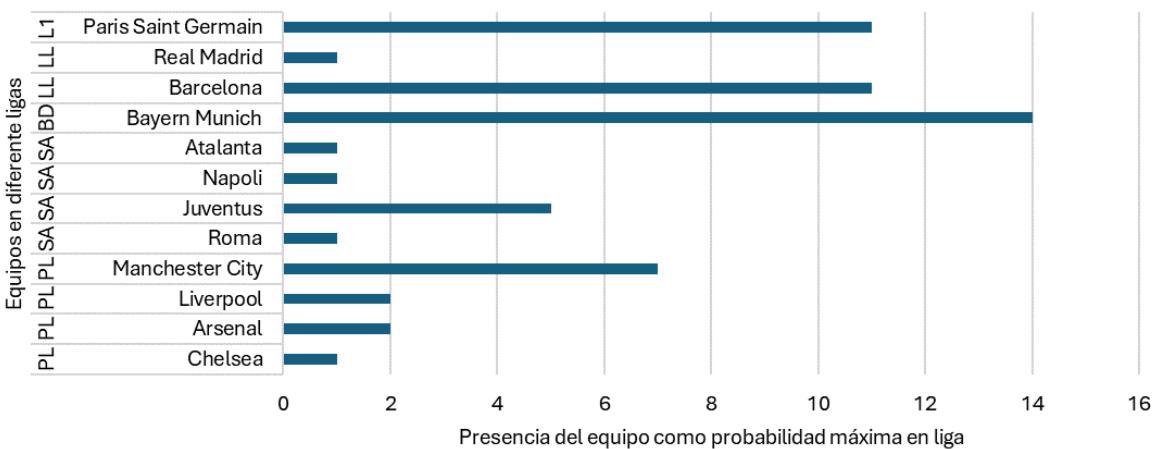
Gráfica 4. Participación de Equipos en Apuestas Máximas por Liga en Interwetten .



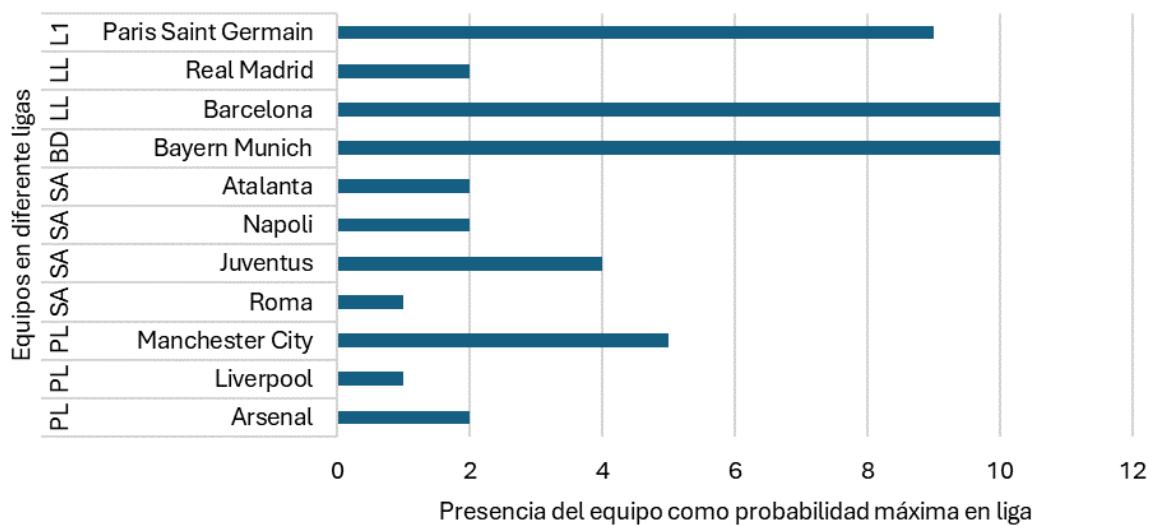
Gráfica 5. Participación de Equipos en Apuestas Máximas por Liga en Pinnacle Sports.



Gráfica 5. Participación de Equipos en Apuestas Máximas por Liga en William Hill.



Gráfica 7. Participación de Equipos en Apuestas Máximas por Liga en VC Bet.



Discusión

La búsqueda de la probabilidad máxima por temporada, realizada en la *pregunta 9 de la etapa 2*, reveló la importancia de analizar las ligas de manera más específica para identificar patrones más claros y consistentes (*cuadro 1 y gráfica 1*). Este enfoque fue respaldado por la observación de que un equipo, como el Barcelona, podía destacar con probabilidad máxima en varias temporadas, pero al considerar la liga en su conjunto, otros equipos podían destacar por su constancia en su probabilidad máxima *pregunta 4 de la etapa 2*.

Este criterio permitió centrarse en equipos que no solo tenían un rendimiento destacado en términos de probabilidades, sino que también demostraban constancia y singularidad en el contexto de sus respectivas ligas, maximizando así las posibilidades de éxito en las apuestas deportivas.

La combinación de estos factores: probabilidad máxima, consistencia temporal, énfasis en apuestas locales y presencia única en la liga, se utilizó como criterio para seleccionar equipos candidatos para las apuestas. Esta estrategia integral buscaba maximizar la probabilidad de éxito al considerar no solo las cifras de las casas de apuestas, sino también la estabilidad y singularidad de los equipos en sus respectivas ligas. En base a este enfoque, se identificaron equipos como Paris Saint Germain en la "Ligue 1" o Bayern Munich en la "Bundesliga" como candidatos sólidos para apuestas (*cuadro 8*).

Referencias

Skrill. (22 de noviembre de 2023). Claves para Entender las Cuotas de Apuestas Deportivas y Calcular Apuestas de Valor. <https://acortar.link/fjMUib>

Anexos

Say, J. y Cruz G. (2024). ApuestasEnFutbol. [Github](#)

```

SELECT liga, equipo, goles, tiros, tiro_al_arco, temporada, victorias,
juegos_como_local, victorias_local, juegos_como_visitante,
victorias_como_visitante FROM (SELECT leagues.name AS liga,
teams.name AS equipo, SUM(teamstats.goals) AS goles,
SUM(teamstats.shots) AS tiros, SUM(teamstats.shotsontarget) AS
tiro_al_arco, teamstats.season AS temporada, COUNT(*) FILTER (WHERE
teamstats.result = 'W') AS victorias, COUNT(*) FILTER (WHERE
games.hometeamid = teams.teamid) AS juegos_como_local, COUNT(*) FILTER
(WHERE games.awayteamid = teams.teamid) AS juegos_como_visitante, COUNT(*) FILTER
(WHERE teamstats.result = 'W' AND games.hometeamid = teams.teamid) AS
victorias_local, COUNT(*) FILTER (WHERE teamstats.result = 'W' AND
games.awayteamid = teams.teamid) AS victorias_como_visitante, ROW_NUMBER()
OVER (PARTITION BY leagues.name, teamstats.season ORDER BY COUNT(*)
DESC) AS row_num FROM teamstats JOIN games ON teamstats.gameid = games.gameid JOIN
leagues ON leagues.leagueid = games.leagueid JOIN teams ON
teams.teamid = teamstats.teamid GROUP BY leagues.name, teams.name,
teamstats.season) AS ranked WHERE row_num = 1;

```

row_num = 1;	liga	equipo	goles	tiros	tiro_al_arco	temporada	victorias	juegos_como_local	victorias_local	juegos_como_visitante	victorias_como_visitante	porcentaje_victorias_local	row_num
	Bundesliga	Bayern Munich	80	585	323	2014	25	17	14	17	11	50.00000000000000	1
	Bundesliga	Bayern Munich	80	577	263	2015	28	17	16	17	13	53.57142857142857	1
	Bundesliga	Bayern Munich	89	622	230	2016	25	17	13	17	12	52.00000000000000	1
	Bundesliga	Bayern Munich	92	601	237	2017	27	17	14	17	13	51.85185185185185	200
	Bundesliga	Bayern Munich	88	635	255	2018	24	17	13	17	11	54.16666666666666	700
	Bundesliga	Bayern Munich	100	615	254	2019	26	17	13	17	13	50.00000000000000	1
	Bundesliga	Bayern Munich	91	576	234	2020	24	17	13	17	14	54.16666666666666	700
	La Liga	Real Madrid	118	600	323	2014	38	19	16	19	14	53.42857142857142	1
	La Liga	Barcelona	112	645	266	2015	29	19	16	19	13	55.472373314344827680	1
	La Liga	Real Madrid	106	663	265	2016	29	19	14	19	15	48.2758620689551724100	1
	La Liga	Barcelona	99	586	272	2017	28	19	16	19	12	57.142857142857142857	1
	La Liga	Barcelona	98	562	254	2018	26	19	15	19	11	57.6923076923076923000	1
	La Liga	Real Madrid	70	563	236	2019	26	19	15	19	11	57.6923076923076923000	1
	La Liga	Atletico Madrid	67	480	200	2020	20	19	15	19	11	57.6923076923076923000	1
	Ligue 1	Paris Saint Germain	83	581	219	2014	24	19	15	19	9	62.50000000000000	1
	Ligue 1	Paris Saint Germain	102	565	252	2015	30	19	15	19	15	50.00000000000000	1
	Ligue 1	Monaco	107	555	244	2016	30	19	17	19	13	56.66666666666666	700
	Ligue 1	Paris Saint Germain	108	626	253	2017	29	19	17	19	12	58.6268965517241379300	1
	Ligue 1	Paris Saint Germain	105	564	230	2018	29	19	17	19	12	58.6268965517241379300	1
	Ligue 1	Paris Saint Germain	75	448	195	2019	27	14	12	19	10	54.16666666666666	700
	Ligue 1	Paris Saint Germain	96	510	215	2020	26	19	15	19	13	50.00000000000000	1
	Premier League	Chelsea	73	564	210	2014	26	19	15	19	11	57.6923076923076923000	1
	Premier League	Leicester	68	522	179	2015	23	19	12	19	11	52.17391304347826687000	1
	Premier League	Chelsea	85	588	204	2016	30	19	17	19	13	56.66666666666666	700
	Premier League	Manchester City	106	663	264	2017	32	19	16	19	16	50.00000000000000	1
	Premier League	Manchester City	93	683	268	2018	32	19	18	19	14	50.00000000000000	1
	Premier League	Leicester	85	531	231	2019	32	19	18	19	14	56.25000000000000	1
	Premier League	Manchester City	93	600	220	2020	27	19	13	19	14	48.54814014814814914000	1
Serie A	Juventus	72	602	201	2014	26	19	16	19	10	61.53846153846153846200	1	
Serie A	Juventus	75	599	215	2015	29	19	16	19	13	55.172379313434827680	1	
Serie A	Juventus	77	583	220	2016	29	19	18	19	11	62.06896551724137931000	1	
Serie A	Juventus	86	555	207	2017	30	19	16	19	14	53.33333333333333333300	1	
Serie A	Juventus	70	609	209	2018	28	19	15	19	13	53.33333333333333333300	1	
Serie A	Juventus	76	579	243	2019	26	19	16	19	10	61.53846153846153846200	1	
Serie A	Inter	89	553	203	2020	28	19	17	19	11	60.71428571428571428600	1	
(35 filas)													

Conclusión: Se realizó un query, con base al query realizado en la etapa 2 pregunta 8, donde ahora se contrastó el total de victorias de los mejores equipos como locales y como visitantes, donde en la mayoría de los casos, salvo por 2 casos atípicos, los equipos que jugaban como locales resultaban ganadores.

