

Licenciatura en Ingeniería en Ciencias de la Computación
Base de datos 1
Sección: 10
Catedrático: Héctor Hurtarte



Análisis de datos para la apuesta a un equipo de fútbol

Javier Andres Chen - 22153
Pedro Pablo Guzman Mayen - 22111

26 de Febrero 2024

Introducción

El propósito de este trabajo es explorar y analizar el uso de una base de datos relacionada con el fútbol en diferentes aspectos del deporte, incluyendo la gestión de equipos, el análisis de rendimiento, estadísticas dentro de los partidos, analizar los mejores equipos bajo ciertos aspectos, probabilidades de la victoria de un equipo, etc. Nos sumergimos en la estructura y la funcionalidad de estas bases de datos, examinaremos cómo se recopilan y procesan los datos, y exploraremos cómo se utilizan para llegar a tomar la decisión de elegir al mejor equipo bajo ciertos parámetros y así poder hacer la apuesta más segura posible al equipo que tiene las mayores posibilidades para ganar.

Diccionario de la base de datos

Teams

- Name
- TeamID

Teamstats

- gameID
- teamID
- Season
- goals
- results

Shots

- GameID
- Minute
- shooterID
- AssisterID

Player

- Name
- PlayerID

Games

- gameID
- LeagueID
- Season
- homeTeamID
- awayTeamID

Appeareances

- gameID
- PlayerID
- leagueID
- goals
- assists
- shoots

Diagrama entidad relación de la base de datos

Link:

https://lucid.app/lucidchart/f678b74a-3117-4930-a4aa-0ad15d59bb2b/edit?viewport_loc=-1884%2C-682%2C5292%2C2368%2C0_0&invitationId=inv_3fab9f72-694d-4b18-a05a-112e2ed2ee7f

Resultados

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:

```
--Juegos disputados por cada equipo en cada temporada
select t."name", string_agg( distinct l."name"::text, ' ' ) as "Liga",
coalesce(sum(case when g."season" = 2014 then 1 else 0 end), 0) as "Juegos en 2014",
coalesce(sum(case when g."season" = 2015 then 1 else 0 end), 0) as "Juegos en 2015",
coalesce(sum(case when g."season" = 2016 then 1 else 0 end), 0) as "Juegos en 2016",
coalesce(sum(case when g."season" = 2017 then 1 else 0 end), 0) as "Juegos en 2017",
coalesce(sum(case when g."season" = 2018 then 1 else 0 end), 0) as "Juegos en 2018",
coalesce(sum(case when g."season" = 2019 then 1 else 0 end), 0) as "Juegos en 2019",
coalesce(sum(case when g."season" = 2020 then 1 else 0 end), 0) as "Juegos en 2020"
from teams t
left join
games g on t."teamID" = g."homeTeamID" or t."teamID" = g."awayTeamID"
join
leagues l on g."leagueID" = l."leagueID"
group by t."name" order by "Liga";
```

Query No.1 Cantidad de juegos disputados por cada equipo en cada temporada

Resultado:

name	Liga	Juegos en 2014	Juegos en 2015	Juegos en 2016	Juegos en 2017	Juegos en 2018	Juegos en 2019	Juegos en 2020
Hannover 96	Bundesliga	34	34	0	34	34	0	0
Borussia Dortmund	Bundesliga	34	34	34	34	34	34	34
Borussia M.Gladbach	Bundesliga	34	34	34	34	34	34	34
Union Berlin	Bundesliga	0	0	0	0	0	34	34
RasenBallsport Leipzig	Bundesliga	0	0	34	34	34	34	34
Paderborn	Bundesliga	34	0	0	0	0	34	0
Darmstadt	Bundesliga	0	34	34	0	0	0	0
Ingolstadt	Bundesliga	0	34	34	0	0	0	0
Wolfsburg	Bundesliga	34	34	34	34	34	34	34
Hertha Berlin	Bundesliga	34	34	34	34	34	34	34
Eintracht Frankfurt	Bundesliga	34	34	34	34	34	34	34
Augsburg	Bundesliga	34	34	34	34	34	34	34
Hoffenheim	Bundesliga	34	34	34	34	34	34	34
Bayer Leverkusen	Bundesliga	34	34	34	34	34	34	34
Nuernberg	Bundesliga	0	0	0	0	34	0	0
Werder Bremen	Bundesliga	34	34	34	34	34	34	34
FC Cologne	Bundesliga	34	34	34	34	0	34	34
Arminia Bielefeld	Bundesliga	0	0	0	0	0	0	34
Fortuna Duesseldorf	Bundesliga	0	0	0	0	34	34	0
Freiburg	Bundesliga	34	0	34	34	34	34	34
Schalke 04	Bundesliga	34	34	34	34	34	34	34
Mainz 05	Bundesliga	34	34	34	34	34	34	34
Bayern Munich	Bundesliga	34	34	34	34	34	34	34
VfB Stuttgart	Bundesliga	34	34	0	34	34	0	34

Tabla No.1 Juegos disputados por cada equipo en cada temporada de la Bundesliga

name	Liga	Juegos en 2014	Juegos en 2015	Juegos en 2016	Juegos en 2017	Juegos en 2018	Juegos en 2019	Juegos en 2020
Espanyol	La Liga	38	38	38	38	38	38	0
Mallorca	La Liga	0	0	0	0	0	38	0
Malaga	La Liga	38	38	38	38	0	0	0
Levante	La Liga	38	38	0	38	38	38	38
Getafe	La Liga	38	38	0	38	38	38	38
Girona	La Liga	0	0	0	38	38	0	0
Villarreal	La Liga	38	38	38	38	38	38	38
Granada	La Liga	38	38	38	0	0	38	38
Athletic Club	La Liga	38	38	38	38	38	38	38
Leganes	La Liga	0	0	38	38	38	38	0
Las Palmas	La Liga	0	38	38	38	0	0	0
Atletico Madrid	La Liga	38	38	38	38	38	38	38
Barcelona	La Liga	38	38	38	38	38	38	38
Valencia	La Liga	38	38	38	38	38	38	38
Alaves	La Liga	0	0	38	38	38	38	38
Sporting Gijon	La Liga	0	38	38	0	0	0	0
Almeria	La Liga	38	0	0	0	0	0	0
Sevilla	La Liga	38	38	38	38	38	38	38
SD Huesca	La Liga	0	0	0	0	38	0	38
Cadiz	La Liga	0	0	0	0	0	0	38
Real Valladolid	La Liga	0	0	0	0	38	38	38
Real Sociedad	La Liga	38	38	38	38	38	38	38
Celta Vigo	La Liga	38	38	38	38	38	38	38
Real Madrid	La Liga	38	38	38	38	38	38	38

Tabla No.2 Juegos disputados por cada equipo en cada temporada de La Liga Española

team name	Liga	Juegos en 2014	Juegos en 2015	Juegos en 2016	Juegos en 2017	Juegos en 2018	Juegos en 2019	Juegos en 2020
Strasbourg	Ligue 1	0	0	0	38	38	27	38
Amiens	Ligue 1	0	0	0	38	38	28	0
Angers	Ligue 1	0	38	38	38	38	28	38
Bordeaux	Ligue 1	38	38	38	38	38	28	38
Brest	Ligue 1	0	0	0	0	0	28	38
Caen	Ligue 1	38	38	38	38	38	0	0
Dijon	Ligue 1	0	0	38	38	38	28	38
Evian Thonon Gaillard	Ligue 1	38	0	0	0	0	0	0
GFC Ajaccio	Ligue 1	0	38	0	0	0	0	0
Guingamp	Ligue 1	38	38	38	38	38	0	0
Lens	Ligue 1	38	0	0	0	0	0	38
Lille	Ligue 1	38	38	38	38	38	28	38
Lorient	Ligue 1	38	38	38	0	0	0	38
Lyon	Ligue 1	38	38	37	38	38	28	38
Marseille	Ligue 1	38	38	38	38	38	28	38
Metz	Ligue 1	38	0	38	38	0	28	38
Monaco	Ligue 1	38	38	38	38	38	28	38
Montpellier	Ligue 1	38	38	38	38	38	28	38
Nancy	Ligue 1	0	0	38	0	0	0	0
Nantes	Ligue 1	38	38	38	38	38	28	38
Nice	Ligue 1	38	38	38	38	38	28	38
Nimes	Ligue 1	0	0	0	0	38	28	38
Paris Saint Germain	Ligue 1	38	38	38	38	38	27	38
Reims	Ligue 1	38	38	0	0	38	28	38

Tabla No.3 Juegos disputados en cada temporada por cada equipo de la Ligue 1

team name	Liga	Juegos en 2014	Juegos en 2015	Juegos en 2016	Juegos en 2017	Juegos en 2018	Juegos en 2019	Juegos en 2020
Tottenham	Premier League	38	38	38	38	38	38	38
Arsenal	Premier League	38	38	38	38	38	38	38
Middlesbrough	Premier League	0	0	38	0	0	0	0
Wolverhampton Wanderer	Premier League	0	0	0	0	38	38	38
Bournemouth	Premier League	0	38	38	38	38	38	0
Newcastle United	Premier League	38	38	0	38	38	38	38
Norwich	Premier League	0	38	0	0	0	38	0
Everton	Premier League	38	38	38	38	38	38	38
Crystal Palace	Premier League	38	38	38	38	38	38	38
Queens Park Rangers	Premier League	38	0	0	0	0	0	0
Chelsea	Premier League	38	38	38	38	38	38	38
Cardiff	Premier League	0	0	0	0	38	0	0
Watford	Premier League	0	38	38	38	38	38	0
Aston Villa	Premier League	38	38	0	0	0	38	38
West Bromwich Albion	Premier League	38	38	38	38	0	0	38
West Ham	Premier League	38	38	38	38	38	38	38
Burnley	Premier League	38	0	38	38	38	38	38
Brighton	Premier League	0	0	0	38	38	38	38
Sheffield United	Premier League	0	0	0	0	0	38	38
Southampton	Premier League	38	38	38	38	38	38	38
Stoke	Premier League	38	38	38	38	0	0	0
Hull	Premier League	38	0	38	0	0	0	0
Leeds	Premier League	0	0	0	0	0	0	38
Huddersfield	Premier League	0	0	0	38	38	0	0

Juegos disputados en cada temporada por los equipos de la Premier League

team name	Liga	Juegos en 2014	Juegos en 2015	Juegos en 2016	Juegos en 2017	Juegos en 2018	Juegos en 2019	Juegos en 2020
Sassuolo	Serie A	38	38	38	38	38	38	38
Lazio	Serie A	38	38	38	38	38	38	38
Cagliari	Serie A	38	0	38	38	38	38	38
Udinese	Serie A	38	38	38	38	38	38	38
Lecce	Serie A	0	0	0	0	0	38	0
Bologna	Serie A	0	38	38	38	38	38	38
Napoli	Serie A	38	38	38	38	38	38	38
Torino	Serie A	38	38	38	38	38	38	38
Verona	Serie A	38	38	0	38	0	38	38
Benevento	Serie A	0	0	0	38	0	0	38
Fiorentina	Serie A	38	38	38	38	38	38	38
Frosinone	Serie A	0	38	0	0	38	0	0
Empoli	Serie A	38	38	38	0	38	0	0
SPAL 2013	Serie A	0	0	0	38	38	38	0
Palermo	Serie A	38	38	38	0	0	0	0
Atalanta	Serie A	38	38	38	38	38	38	38
Parma Calcio 1913	Serie A	38	0	0	0	38	38	38
Pescara	Serie A	0	0	38	0	0	0	0
Spezia	Serie A	0	0	0	0	0	0	38
Crotone	Serie A	0	0	38	38	0	0	38
Chievo	Serie A	38	38	38	38	38	0	0
Brescia	Serie A	0	0	0	0	0	38	0
Cesena	Serie A	38	0	0	0	0	0	0
Carpi	Serie A	0	38	0	0	0	0	0

Tabla No.4 Juegos disputados en cada temporada por los equipos de la Serie A

Explicación del query: Primero se seleccionó la tabla teams (la cuál contiene el id y nombre de cada equipo), y hacemos la unimos a la izquierda (left join) con la tabla games (la cuál contiene información de los juegos que cada equipo ha disputado). El criterio que se usa para el join es que “teamID” de la tabla teams debe coincidir con “awayTeamID” o

“homeTeamID” de la tabla games, con esto se crea una nueva tabla en la cuál se obtiene a un equipo junto a todos los juegos que ha disputado a lo largo de las temporadas. Se decidió usar la opción left para hacer la unión ya que con esto se asegura de que estén presentes todos los equipos de la tabla teams. Una vez con la tabla formada por la unión entre estas dos tablas, se debe unir esta nueva tabla con la tabla leagues (la cuál contiene información de todas las ligas disponibles), el criterio para la unión debe ser que el campo “leagueID” de la tabla games debe ser igual a “leagueID” de la tabla leagues, con esto se asegura de que en la nueva tabla se indique a qué liga pertenece cada juego. Una vez con esta unión entre tablas, se agrupa la tabla con respecto al nombre del equipo, con esto se crea una partición en la tabla en donde los datos están agrupados con respecto al nombre del equipo. Una vez con eso, se selecciona el nombre del equipo y su correspondiente liga, para obtener la liga de cada equipo será pertinente usar distinct dentro de un string agg para asegurarnos de tener un único valor de liga. Ya con esos datos también se suma 1 cada vez que encontremos un juego en donde la temporada sea 2014, se realizará lo mismo con cada temporada disponible. Con esto se sumará la cantidad de juegos jugados por cada equipo en cada temporada. Finalmente se ordenará la tabla con respecto a las ligas de los equipos.

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

Query:

```
--Diferencia de goles
select "HomeGoals"."homeTeamID", t."name", l."name" as "Liga",
("HomeGoals"."homeGoals" + "AwayGoals"."awayGoals") as "Goles a favor",
("HomeGoals"."awayGoals" + "AwayGoals"."homeGoals") as "Goles en contra",
(("HomeGoals"."homeGoals" + "AwayGoals"."awayGoals") - ("HomeGoals"."awayGoals" + "AwayGoals"."homeGoals")) as "Diferencia de Gol",
rank() over (partition by l."name" order by
(("HomeGoals"."homeGoals" + "AwayGoals"."awayGoals") - ("HomeGoals"."awayGoals" + "AwayGoals"."homeGoals"))
desc) as ranking
from
(select "homeTeamID", sum("homeGoals") as "homeGoals", sum("awayGoals") as "awayGoals", string_agg(distinct "leagueID"::text, ' ') as "Liga"
from games group by "homeTeamID" order by "homeTeamID" desc)
as "HomeGoals"
join
(select "awayTeamID", sum("awayGoals") as "awayGoals", sum("homeGoals") as "homeGoals" from games group by "awayTeamID" order by "awayTeamID" desc)
as "AwayGoals"
on "HomeGoals"."homeTeamID" = "AwayGoals"."awayTeamID"
join teams t on "HomeGoals"."homeTeamID" = t."teamID"
join leagues l on "HomeGoals"."Liga"::int = l."leagueID" |
order by "Diferencia de Gol" desc;
```

Query No.2 Diferencia de goles de cada equipo

Resultado:

homeTeamID	ABC name	ABC Liga	Goles a favor	Goles en contra	Diferencia de Gol	ranking
117	Bayern Munich	Bundesliga	628	193	435	1
129	Borussia Dortmund	Bundesliga	505	294	211	2
136	RasenBallSport Leipzig	Bundesliga	327	190	137	3
119	Bayer Leverkusen	Bundesliga	412	311	101	4
130	Borussia M.Gladbach	Bundesliga	397	315	82	5
120	Hoffenheim	Bundesliga	393	353	40	6
131	Wolfsburg	Bundesliga	360	320	40	6
132	Eintracht Frankfurt	Bundesliga	359	363	-4	8
240	Union Berlin	Bundesliga	91	101	-10	9
262	Arminia Bielefeld	Bundesliga	26	52	-26	10
126	Ingolstadt	Bundesliga	69	99	-30	11
234	Nuernberg	Bundesliga	26	68	-42	12
233	Fortuna Duesseldorf	Bundesliga	85	132	-47	13
127	Darmstadt	Bundesliga	66	116	-50	14
122	Hertha Berlin	Bundesliga	302	355	-53	15
135	Freiburg	Bundesliga	256	323	-67	16
205	Paderborn	Bundesliga	68	139	-71	17
125	Mainz 05	Bundesliga	302	374	-72	18
124	Schalke 04	Bundesliga	291	365	-74	19
123	Werder Bremen	Bundesliga	334	409	-75	20
133	VfB Stuttgart	Bundesliga	216	296	-80	21
134	FC Cologne	Bundesliga	243	323	-80	21
118	Hamburger SV	Bundesliga	127	210	-83	23

Tabla No. 5 Ranking de la diferencia de gol de los equipos de la Bundesliga

homeTeamID	ABC name	ABC Liga	Goles a favor	Goles en contra	Diferencia de Gol	ranking
148	Barcelona	La Liga	698	228	470	1
150	Real Madrid	La Liga	628	256	372	2
143	Athletico Madrid	La Liga	431	177	254	3
138	Sevilla	La Liga	409	316	93	4
154	Villarreal	La Liga	377	300	77	5
146	Valencia	La Liga	384	324	60	6
140	Real Sociedad	La Liga	374	343	31	7
147	Athletic Club	La Liga	322	303	19	8
261	Cadiz	La Liga	36	58	-22	9
239	Mallorca	La Liga	40	65	-25	10
223	Girona	La Liga	87	112	-25	10
208	Almeria	La Liga	35	64	-29	12
232	SD Huesca	La Liga	77	118	-41	13
152	Celta Vigo	La Liga	355	400	-45	14
137	Malaga	La Liga	153	199	-46	15
206	Cordoba	La Liga	22	68	-46	15
207	Elche	La Liga	69	117	-48	17
142	Getafe	La Liga	231	279	-48	17
149	Sporting Gijon	La Liga	82	134	-52	19
231	Real Valladolid	La Liga	98	151	-53	20
153	Real Betis	La Liga	277	339	-62	21
159	Leganes	La Liga	137	200	-63	22
158	Alaves	La Liga	190	259	-69	23
145	Devo Vallecano	La Liga	120	211	-77	24

Tabla No.6 Ranking de la diferencia de gol de los equipos de La Liga Española

homeTeamID	ABC name	ABC Liga	Goles a favor	Goles en contra	Diferencia de Gol	ranking
161	Paris Saint Germain	Ligue 1	642	198	444	1
178	Lyon	Ligue 1	496	284	212	2
171	Monaco	Ligue 1	458	295	163	3
164	Marseille	Ligue 1	416	300	116	4
160	Lille	Ligue 1	330	266	64	5
170	Nice	Ligue 1	339	308	31	6
163	Rennes	Ligue 1	318	298	20	7
175	Saint-Etienne	Ligue 1	311	299	12	8
166	Montpellier	Ligue 1	327	323	4	9
176	Bordeaux	Ligue 1	319	324	-5	10
241	Brest	Ligue 1	84	103	-19	11
173	GFC Ajaccio	Ligue 1	37	58	-21	12
209	Evian Thonon Gaillard	Ligue 1	41	62	-21	12
225	Strasbourg	Ligue 1	183	205	-22	14
182	Nancy	Ligue 1	29	52	-23	15
210	Lens	Ligue 1	87	115	-28	16
177	Reims	Ligue 1	198	236	-38	17
162	SC Bastia	Ligue 1	102	142	-40	18
167	Angers	Ligue 1	234	279	-45	19
224	Amiens	Ligue 1	99	144	-45	19
235	Nimes	Ligue 1	126	173	-47	21
168	Nantes	Ligue 1	261	313	-52	22
179	Lorient	Ligue 1	185	246	-61	23
169	Guingamp	Ligue 1	210	291	-81	24
172	Troyes	Ligue 1	60	142	-82	25

Tabla No.7 Ranking de la diferencia de gol de los equipos de la Ligue 1

homeTeamID	ABC name	ABC Liga	Goles a favor	Goles en contra	Diferencia de Gol	ranking
88	Manchester City	Premier League	620	235	385	1
87	Liverpool	Premier League	519	275	244	2
82	Tottenham	Premier League	483	281	202	3
80	Chelsea	Premier League	469	285	184	4
89	Manchester United	Premier League	437	263	174	5
83	Arsenal	Premier League	471	305	166	6
75	Leicester	Premier League	404	353	51	7
245	Leeds	Premier League	62	54	8	8
72	Everton	Premier League	358	357	1	9
229	Wolverhampton Wanderers	Premier League	134	138	-4	10
93	Middlesbrough	Premier League	27	53	-26	11
81	West Ham	Premier League	367	394	-27	12
202	Queens Park Rangers	Premier League	42	73	-31	13
227	Cardiff	Premier League	34	69	-35	14
74	Southampton	Premier League	334	371	-37	15
238	Sheffield United	Premier League	59	102	-43	16
85	Stoke	Premier League	165	224	-59	17
91	Hull	Premier League	70	131	-61	18
84	Swansea	Premier League	161	227	-66	19
220	Brighton	Premier League	148	214	-66	19
228	Fulham	Premier League	61	134	-73	21
77	Sunderland	Premier League	108	184	-76	22
79	Norwich	Premier League	65	142	-77	23
219	Huddersfield	Premier League	50	134	-84	24
78	Crystal Palace	Premier League	304	389	-85	25

Tabla No.8 Ranking de la diferencia de gol de los equipos de la Premier League

homeTeamID	name	Liga	Goles a favor	Goles en contra	Diferencia de Gol	ranking
98	Juventus	Serie A	533	206	327	1
105	Napoli	Serie A	542	281	261	2
95	Roma	Serie A	499	292	207	3
106	Inter	Serie A	474	269	205	4
96	Lazio	Serie A	482	333	149	5
107	Atalanta	Serie A	463	325	138	6
111	AC Milan	Serie A	410	303	107	7
110	Fiorentina	Serie A	383	343	40	8
113	Torino	Serie A	373	386	-13	9
260	Spezia	Serie A	52	72	-20	10
103	Carpi	Serie A	37	57	-20	10
104	Sassuolo	Serie A	373	396	-23	12
102	Sampdoria	Serie A	361	388	-27	13
243	Lecce	Serie A	52	85	-33	14
203	Cesena	Serie A	36	73	-37	15
242	Brescia	Serie A	35	79	-44	16
114	Pescara	Serie A	35	83	-48	17
108	Empoli	Serie A	166	232	-66	18
100	Palermo	Serie A	124	197	-73	19
97	Bologna	Serie A	264	341	-77	20
101	Genoa	Serie A	311	390	-79	21
112	Frosinone	Serie A	64	145	-81	22
221	SPAL 2013	Serie A	110	192	-82	23
222	Benevento	Serie A	73	159	-86	24
115	Crotone	Serie A	119	216	-97	25

Tabla No.9 Ranking de la diferencia de gol de los equipos de la Serie A

homeTeamID	name	Liga	Goles a favor	Goles en contra	Diferencia de Gol
148	Barcelona	La Liga	698	228	470
161	Paris Saint Germain	Ligue 1	642	198	444
117	Bayern Munich	Bundesliga	628	193	435
88	Manchester City	Premier League	620	235	385
150	Real Madrid	La Liga	628	256	372
98	Juventus	Serie A	533	206	327
105	Napoli	Serie A	542	281	261
143	Atletico Madrid	La Liga	431	177	254
87	Liverpool	Premier League	519	275	244
178	Lyon	Ligue 1	496	284	212
129	Borussia Dortmund	Bundesliga	505	294	211
95	Roma	Serie A	499	292	207
106	Inter	Serie A	474	269	205
82	Tottenham	Premier League	483	281	202
80	Chelsea	Premier League	469	285	184
89	Manchester United	Premier League	437	263	174
83	Arsenal	Premier League	471	305	166
171	Monaco	Ligue 1	458	295	163
96	Lazio	Serie A	482	333	149
107	Atalanta	Serie A	463	325	138
136	RasenBallsport Leipzig	Bundesliga	327	190	137
164	Marseille	Ligue 1	416	300	116
111	AC Milan	Serie A	410	303	107

Tabla No.10 Equipos ordenados en base a la diferencia de gol

Explicación: Se crea una tabla llamada “Home goals” a partir de agrupar los datos de la tabla games mediante la columna “homeTeamID”. Una vez que los datos estén agrupados mediante “homeTeamID”, se sumarán los valores de todos los goles de la columna “homeGoals” de cada equipo, este mostrará los goles a favor como local para cada equipo, también se sumarán todos los valores de la columna “awayGoals” para obtener todos los goles en contra cuando un equipo juega como visitante. Luego de eso se formará una tabla llamada “Away Goals” la cuál obtiene los goles a favor como visitante y los goles en contra como local usando el mismo mecanismo para formar la tabla anterior solamente cambiará el

valor para hacer la agrupación el cuál ahora será “awayTeamID”. Una vez con las tablas “Home goals” y “Away goals” formadas, estas son unidas mediante el criterio de que el valor “homeTeamID” de “Home Goals” debe ser igual al valor “awayTeamID” de “Away goals”. Ya con la unión de estas tablas, también se deberá hacer una nueva unión entre esta nueva tabla y la tabla leagues para determinar el nombre de la liga a la que pertenece un equipo. Para poder hacer esta unión, se debió de aplicar un string_agg() para seleccionar los diferentes valores de “leagueID” en la tabla games, esto pudo haberse hecho tanto en la tabla “Home goals” como en la tabla “Away Goals”. Ya con el valor de la liga para cada equipo, solamente queda seleccionar el nombre del equipo, su liga, sus goles a favor seleccionando la suma de “HomeFGoals” y “AwayGoals”, sus goles en contra seleccionando la suma de “HomeCGoals” y “AwayCoals” y un ranking en el cuál se rankea la diferencia de gol de cada equipo particionando los valores con respecto a la liga de cada equipo. Con esto se obtiene un query que muestra cuáles son los equipos con mejor diferencia de gol en cada liga.

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

Querys:

```
select p."name" as "Jugador", sum(a.goals) as "Goles totales" from appearances a join players p on a."playerID" = p."playerID" group by p."name" order by "Goles totales" desc;
```

Query No.3 Máximos goleadores en cada liga

Explicación: se selecciona la tabla appearances y se une con la tabla players en las tuplas en las que el valor “playerID” de players es igual al valor “playerID” de la tabla appearances, luego se agrupa por nombre de jugador y se selecciona el nombre del jugador y la suma de todos sus goles, finalmente se ordena de mayor a menor en base a la cantidad de goles de cada jugador

Resultado:

ABC Jugador	127 Goles totales
Lionel Messi	231
Cristiano Ronaldo	215
Robert Lewandowski	203
Luis Suarez	173
Harry Kane	163
Pierre-Emerick Aubameyang	149
Ciro Immobile	133
Edinson Cavani	132
Sergio Aguero	132
Mohamed Salah	130
Romelu Lukaku	128
Alexandre Lacazette	126
Mauro Icardi	121
Karim Benzema	120
Jamie Vardy	118
Antoine Griezmann	116
Neymar	114
Gonzalo Higuain	113
Kylian Mbappe-Lottin	107
Wissam Ben Yedder	107
Iago Aspas	105
Fabio Quagliarella	102
Zlatan Ibrahimovic	99
Andrea Belotti	97

Tabla No.11 Máximos goleadores de todas las ligas

```
--Jugadores con más goles y pases derechos
select p."name" as "Jugador", count(*) as "Pases a la derecha", sum(case when "shotResult" = 'Goal' then 1 else 0 end) as "Goles"
from shots s join players p on p."playerID" = s."shooterID"
where s."lastAction" = 'Pass' and "shotType" = 'RightFoot'
group by p."name" order by count(*) desc ;
```

Query No.4 jugadores con más pases a la derecha que meten goles

Explicación: se selecciona la tabla shots y se une con la tabla players en las tuplas en las que el valor "playerID" de players y "shooterID" de shots es igual. Luego de eso se aplican filtros que obtengan sólo aquellas tuplas en las que contengan una acción de tipo 'Pass' y que el tipo de disparo haya sido con la pierna derecha. Una vez hecho esto se agrupa por nombre de jugador y se selecciona el nombre del jugador, la cantidad de pases (la cantidad de veces en las que aparece en la tabla) y la cantidad de goles que anota.

Resultado:

ABC Jugador	127 Pases a la derecha	127 Goles
Lorenzo Insigne	425	25
Cristiano Ronaldo	405	44
Philippe Coutinho	299	32
Dries Mertens	294	29
Harry Kane	265	37
Gonzalo Higuain	247	39
Karim Benzema	240	39
Ciro Immobile	240	27
Luis Suarez	238	55
Hakan Calhanoglu	237	10
Sergio Aguero	219	26
Kevin De Bruyne	218	16
Pierre-Emerick Aubameyang	213	47
Rodrigo de Paul	213	15
Alejandro Gomez	210	15
Robert Lewandowski	209	41
Paul Pogba	209	12
Alexis Sanchez	202	19
Radja Nainggolan	199	14
Alassane Plea	198	23
Luis Alberto	197	17
Timo Werner	197	27
Antonio Candreva	193	9
Neymar	190	28

Tabla No.12 Jugadores con más pases derechos que meten goles

```
--Jugadores con más pases izquierdos que meten goles
select p."name" as "Jugador", count(*) as "Pases a la derecha", sum(case when "shotResult" = 'Goal' then 1 else 0 end) as "Goles"
from shots s join players p on p."playerID" = s."shooterID"
where s."lastAction" = 'Pass' and "shotType" = 'LeftFoot'
group by p."name" order by count(*) desc ;
```

Query No.5 Jugadores con más pases izquierdos que meten goles

Explicación: tiene el mismo funcionamiento que el query anterior, solamente que este suma los pases que se dan con la pierna izquierda

```
--Jugadores con más pases izquierdos y derechos que meten goles
select p."name" as "Jugador", count(*) as "Pases", sum(case when "shotResult" = 'Goal' then 1 else 0 end) as "Goles"
from shots s join players p on p."playerID" = s."shooterID"
where s."lastAction" = 'Pass'
group by p."name" order by count(*) desc ;
```

Query No.6 Jugadores con más pases izquierdos y derechos que meten goles

Explicación: funciona como los queries anteriores, solamente que en este no se aplica ningún tipo de filtro sobre la pierna con la que se da el pase

Resultado:

	asc Jugador	123 Pases a la derecha	123 Goles
1	Cristiano Ronaldo	600	68
2	Lionel Messi	588	118
3	Lorenzo Insigne	476	28
4	Harry Kane	352	48
5	Gonzalo Higuain	334	48
6	Philippe Coutinho	332	35
7	Dries Mertens	328	35
8	Karim Benzema	325	44
9	Alejandro Gomez	320	22
10	Luis Suarez	315	67
11	Paulo Dybala	307	37
12	Sergio Aguero	305	38
13	Kevin De Bruyne	305	23
14	Hakan Calhanoglu	294	11
15	Josip Ilcic	291	25
16	Robert Lewandowski	282	50
17	Mohamed Salah	281	46
18	Christian Eriksen	281	21
19	Pierre-Emerick Aubameyang	279	63
20	Neymar	279	45
21	Antoine Griezmann	278	47
22	Domenico Berardi	276	18
23	Ciro Immobile	273	32
24	Nabil Fekir	273	25

Tabla No.13 Jugadores con más pases izquierdos y derechos que meten goles

Según la información de la página transfermarkt (Transfermarkt, sf) los 24 jugadores con más goles pertenecen a los equipos con mejor diferencia de gol tanto en su liga como en general. Por ejemplo, el FC Barcelona es el equipo con mejor diferencia de gol tanto de su liga como en general, este equipo cuenta con Lionel Messi y Antoine Griezmann, el primero es el máximo anotador en general y es el jugador con más pases izquierdos y que mete goles, el segundo se encuentra en la posición 16 de los máximos goleadores de todas las ligas y también es de los jugadores que mete goles con más pases en general. Podemos asumir que tener a uno de los mejores goleadores en el equipo ayuda a tener más goles a favor y mejorar la diferencia de gol.

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.

Query:

```
--Comparativo de las probabilidades
select t."name" as "Equipo",
greatest(max(1/"B365H"), max (1/"BWH"), max(1/"IWH"), max(1/"PSH"), max(1/"WHH"), max(1/"VCH"), max(1/"PSCH")) as "Max Home probability",
greatest(max(1/"B365A"), max (1/"BWA"), max(1/"IWA"), max(1/"PSA"), max(1/"WHA"), max(1/"VCA"), max(1/"PSCA")) as "Max away probability",
greatest(max(1/"B365D"), max (1/"BWD"), max(1/"IWD"), max(1/"PSD"), max(1/"WHD"), max(1/"VCD"), max(1/"PSCD")) as "Max draw probability"
from games g join teams t on g."homeTeamID" = t."teamID" or g."awayTeamID" = t."teamID"
where g."B365H" != 0 and g."B365D" !=0 and g."B365A" !=0 and g."BWH" !=0 and g."BWD" != 0 and g."BWA" !=0
and g."IWH" != 0 and g."IWD" !=0 and g."IWA" !=0 and g."PSH" !=0 and g."PSD" !=0 and g."PSA" !=0 and g."WHH" !=0 and g."WHD" !=0
and g."WHA" !=0 and g."VCH" !=0 and g."VCD" !=0 and g."VCA" != 0 and g."PSCH" != 0 and g."PSCD" != 0 and g."PSCA" !=0
group by t."name";
```

Query No.7 Comparativo de probabilidades

Explicación: se selecciona la tabla games y se une con la tabla teams para obtener los nombres de cada equipo y sus respectivos juegos, luego se filtra para aquellas tuplas en donde la probabilidad de las casas de apuestas no es 0. Después se agrupa en base al nombre del equipo y finalmente se selecciona el valor más grande entre los máximos de cada probabilidad, mostrando así las mejores probabilidades que ha tenido cada equipo en cada categoría.

Resultado:

	ABC Equipo	123 Max Home probability	123 Max away probability	123 Max draw probability
1	Nantes	0,9433962264	0,8196721311	0,3636363636
2	Verona	0,9090909091	0,8403361345	0,3472222222
3	Bournemouth	0,9345794393	0,8620689655	0,3333333333
4	Empoli	0,8849557522	0,8	0,3448275862
5	Wolverhampton V	0,9009009009	0,8	0,3703703704
6	Cordoba	0,9615384615	0,9259259259	0,3448275862
7	Palermo	0,8928571429	0,8	0,3389830508
8	Las Palmas	0,9523809524	0,9259259259	0,3333333333
9	Crotone	0,8928571429	0,8403361345	0,3367003367
10	Metz	0,9345794393	0,9090909091	0,3703703704
11	Rennes	0,8333333333	0,8333333333	0,3636363636
12	Leeds	0,7936507937	0,78125	0,3333333333
13	Cesena	0,9009009009	0,8196721311	0,3389830508
14	Dijon	0,9708737864	0,8928571429	0,3484320557
15	Alaves	0,9433962264	0,8547008547	0,3773584906
16	Nancy	0,8849557522	0,7518796992	0,3571428571
17	Real Madrid	0,9615384615	0,9174311927	0,3267973856
18	Lyon	0,8928571429	0,7692307692	0,3389830508
19	Manchester City	0,9615384615	0,9174311927	0,3333333333
20	Troyes	0,9433962264	0,8547008547	0,3472222222
21	Levante	0,9615384615	0,8771929825	0,3448275862
22	Hoffenheim	0,9259259259	0,8333333333	0,3333333333
23	Augsburg	0,9345794393	0,8849557522	0,3333333333
24	SPAL 2013	0,9090909091	0,8196721311	0,3703703704
25	Stoke	0,8771929825	0,8196721311	0,3333333333

Tabla No.14 Comparativo de probabilidad de cada equipo

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

Query:

```
--Mejores equipos según las apuestas
select subquery."name" as "Equipo", AVG("Prob") as "Probabilidad" from
(select t."name",
AVG(1.0 / "B365H" + 1.0 / "BWH" + 1.0 / "IWH" + 1.0 / "PSH" + 1.0 / "MHH" + 1.0 / "VCH" +
1.0 / "PSCH") as "Prob"
from games g join teams t on t."teamID" = g."homeTeamID" where g."B365H" != 0 and g."B365D" != 0 and g."B365A" != 0 and g."BWH" != 0 and g."BWD" != 0 and g."BWA" != 0
and g."IWH" != 0 and g."IWD" != 0 and g."IWA" != 0 and g."PSH" != 0 and g."PSD" != 0 and g."PSA" != 0 and g."MHH" != 0 and g."MHD" != 0
and g."MHA" != 0 and g."VCH" != 0 and g."VCD" != 0 and g."VCA" != 0 and g."PSCH" != 0 and g."PSCD" != 0 and g."PSCA" != 0 group by t."name") as subquery
order by "Probabilidad" desc;
```

Query No.8 Mejores equipos según las casas de apuestas

Explicación: Se unen las tablas teams y games para determinar el nombre de los equipos, una vez con esto se filtra para aquellas tuplas en donde los valores de las casas de apuestas no son igual a cero. De la tabla creada, se seleccionan el promedio de la suma de todas las probabilidades de las casas de apuesta, finalmente se selecciona el promedio de la suma de tales valores.

Resultado:

select subquery."name" as "Equipo", AVG("Prob") as "Prob" Enter		
Grilla	Equipo	Probabilidad
1	Bayern Munich	5,8327801218
2	Barcelona	5,768613134
3	Paris Saint Germain	5,7454090259
4	Real Madrid	5,4678941829
5	Manchester City	5,3524533712
6	Juventus	5,1123641221
7	Borussia Dortmund	4,8336963917
8	Napoli	4,8304923739
9	Liverpool	4,7354055958
10	Atletico Madrid	4,7207303654
11	Chelsea	4,6399517342
12	Lyon	4,5071100188
13	Inter	4,4179040962
14	Roma	4,3964135928
15	Manchester United	4,3769261797
16	RasenBallsport Leipzig	4,3010852096
17	Arsenal	4,2956455394
18	Tottenham	4,1677173014
19	Lazio	4,1455711548
20	Monaco	4,1314732864
21	Sevilla	4,1150053374
22	Bayer Leverkusen	4,0905616737
23	Marseille	4,0179838957
24	AC Milan	3,9211783208
25	Atalanta	3,7654947518

Tabla No.14 Mejores equipos según las casas de apuestas

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?

Query:

```
--Mejores jugadores
select p."name" as "Jugador", sum("goals") as "Goles", sum("shots") as "Tiros", sum("xGoalsChain") as "Goal Chain", sum("xGoalsBuildup") as "GoalBuildup",
sum("assists") as "Asistencias", sum("keyPasses") as "Pases clave"
from appearances a join players p on p."playerID" = a."playerID" group by p."name"
order by sum("goals") desc, sum("shots") desc, sum("xGoalsChain") desc, sum("xGoalsBuildup") desc,
sum("assists") desc, sum("keyPasses") desc;
```

Query No.9 Jugadores con mejores atributos

Explicación: se selecciona la tabla appearances y se une con la tabla players bajo el criterio de que el valor "playerID" de players debe ser igual al valor "playerID" de appearances. Luego de eso se agrupan los datos en base al nombre del jugador y se hace una suma de las características más importantes del juego para cada jugador (pases clave, asistencias, goles, participación en jugadas, etc), finalmente se ordena de mayor a menor en base a esas características

Resultados:

	Reb Jugador	128 Goles	128 Tiros	128 Goal Chain	128 GoalBuildup	128 Asistencias	128 Pases clave
1	Lionel Messi	231	1.244	298,7463631444	119,2948233867	97	596
2	Cristiano Ronaldo	215	1.342	222,6676497152	58,8338882262	54	333
3	Robert Lewandowski	203	943	232,7946682665	50,8270333484	32	221
4	Luis Suarez	173	795	233,2485771631	52,9548434103	73	347
5	Harry Kane	163	886	154,5790643534	38,1700885026	34	252
6	Pierre-Emerick Aubameyang	149	670	176,2004773747	30,9985196684	31	210
7	Ciro Immobile	133	725	133,3265505098	30,2381694037	36	275
8	Sergio Aguero	132	714	145,3888693061	35,3724603904	31	190
9	Edinson Cavani	132	560	150,0957663693	30,1150887571	25	138
10	Mohamed Salah	130	721	187,8850783221	54,7318522893	53	410
11	Romelu Lukaku	128	670	155,6484987093	40,58544371	37	295
12	Alexandre Lacazette	126	521	137,5770062497	43,0238113794	30	260
13	Mauro Icardi	121	535	127,5323174596	22,5158159081	29	200
14	Karim Benzema	120	650	186,151882872	54,8810807848	55	314
15	Jamie Vardy	118	536	128,809199607	15,8362164684	38	239
16	Antoine Griezmann	116	569	151,0535237771	53,3554154988	42	324
17	Neymar	114	610	212,1432900615	81,4472775497	61	454
18	Gonzalo Higuain	113	707	132,8749198429	31,0968894325	23	241
19	Wissam Ben Yedder	107	490	132,5188148683	35,8011766514	35	271
20	Kylian Mbappe-Lottin	107	453	144,059039576	33,9091687538	36	207
21	Iago Aspas	105	472	134,8082176913	47,3531957995	35	354
22	Fabio Quagliarella	102	768	97,8898637667	14,8994466835	26	214
23	Zlatan Ibrahimovic	99	507	104,3784500565	33,9177624453	31	178
24	Andrea Belotti	97	663	106,4137499854	22,2538148947	29	252
25	Sadio Mane	95	531	152,0159342214	46,5938344849	36	338

Tabla No 16. Jugadores con mejores atributos de juego

- 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.

Query:

```

select "Home"."name" as "Team", "Home"."Home expected goals" as "Home xGoals"
, "Home"."Home goals" as "Home goals",
"Away"."Away expected goals" as "Away xGoals",
"Away"."Away goals" as "Away Goals",
("Home"."Home expected goals" + "Away"."Away expected goals") as "Total xGoals",
("Home"."Home goals" + "Away"."Away goals") as "Total Goals"
from
(select t."name", sum("xGoals") as "Home expected goals", sum("goals") as "Home goals" from teamstats ts join teams t on t."teamID" =
ts."teamID" where "location" = 'h'
group by t."name" order by
sum("goals") desc) as "Home" join
(select t."name", sum("xGoals") as "Away expected goals", sum("goals") as "Away goals" from teamstats ts join teams t on t."teamID" =
ts."teamID" where "location" = 'a' group by t."name" order by
sum("goals") desc) as "Away" on "Home"."name" = "Away"."name" order by "Total xGoals" desc;

```

Query No.10 Goles esperados vs Expected goals de cada equipo

Explicación: Se crea una tabla llamada home uniendo la tabla teamstats y teams bajo el criterio de que "teamID" de team debe ser igual que "teamID" de teamstats una vez con eso se aplican filtros para obtener solamente aquellos juegos en donde el equipo fue local. Luego de eso se agrupa mediante el nombre del equipo y se selecciona la suma de todos los expected goals del equipo y los goles reales del equipo jugando como local. Se hace el mismo procedimiento para formar la tabla away, solamente aplicando criterios para obtener los expected goals y los goles reales del equipo como visitante. Una vez con las tablas away y home, se unen con el criterio de que el nombre del equipo de la tabla home debe ser igual al nombre del equipo en la tabla away. Ya con eso, se selecciona el nombre del equipo, la suma de away expected goals y home expected goals, la suma de home goals y away goals, los goles esperados como visitante, los goles esperados como local, los goles anotados como local, los goles anotados como visitante. Finalmente se ordena en base al total de goles esperados de mayor a menor.

Resultado:

	Team	Home xGoals	Home goals	Away xGoals	Away Goals	Total xGoals	Total Goals
1	Barcelona	356,460901	395	283,403425	303	639,864326	698
2	Paris Saint Germain	334,769861	374	257,587214	268	592,357075	642
3	Real Madrid	308,253559	342	266,0034	286	574,256959	628
4	Bayern Munich	315,987651	374	238,7419142	254	554,7295652	628
5	Manchester City	326,430686	346	260,508708	274	586,939394	620
6	Napoli	275,475352	304	213,6078303	238	489,0831823	542
7	Juventus	266,876779	300	204,98377	233	471,860549	533
8	Liverpool	261,432358	289	215,606635	230	477,038993	519
9	Borussia Dortmund	273,594125	294	202,894162	211	476,488287	505
10	Roma	271,68328	277	211,040628	222	482,723908	499
11	Lyon	245,857186	263	207,732079	233	453,589265	496
12	Tottenham	232,30143	258	189,925328	225	422,226758	483
13	Lazio	244,272012	275	192,338271	207	436,610283	482
14	Inter	243,799488	263	216,032802	211	459,83229	474
15	Arsenal	255,187595	267	191,861487	204	447,049082	471
16	Chelsea	254,930109	253	198,001891	216	452,932	469
17	Atalanta	247,966516	246	207,341742	217	455,308258	463
18	Monaco	232,736605	254	185,8485292	204	418,5851342	458
19	Manchester United	228,922971	243	185,299269	194	414,22224	437
20	Atletico Madrid	223,621405	246	167,540241	185	391,161646	431
21	Marseille	212,5447	228	170,7975543	188	383,3422543	416
22	Bayer Leverkusen	212,526614	224	170,273733	188	382,800347	412
23	AC Milan	220,508725	208	183,497447	202	404,006172	410
24	Sevilla	251,534766	238	181,728157	171	433,262923	409
25	Leicester	200,9067677	212	182,2766477	192	383,1834154	404

Tabla No.17 Expected goals vs goals de cada equipo

Queries alternativos

```

select "Home"."name" as "Team", "Home"."Home expected goals" as "Home xGoals"
, "Home"."Home goals" as "Home goals",
"Away"."Away expected goals" as "Away xGoals",
"Away"."Away goals" as "Away Goals",
("Home"."Home expected goals" + "Away"."Away expected goals") as "Total xGoals",
("Home"."Home goals" + "Away"."Away goals") as "Total Goals"
from
(select t."name", sum("xGoals") as "Home expected goals", sum("goals") as "Home goals" from teamstats ts join teams t on t."teamID" =
ts."teamID" where "location" = 'h'
group by t."name" order by
sum("goals") desc) as "Home" join
(select t."name", sum("xGoals") as "Away expected goals", sum("goals") as "Away goals" from teamstats ts join teams t on t."teamID" =
ts."teamID" where "location" = 'a' group by t."name" order by
sum("goals") desc) as "Away" on "Home"."name" = "Away"."name" order by "Total xGoals" desc limit 1;

```

Query No.11 Equipo con mayores expected goals

Resultado:

Team	Home xGoals	Home goals	Away xGoals	Away Goals	Total xGoals	Total Goals
Barcelona	356,460901	395	283,403425	303	639,864326	698

Tabla No.18 Equipo con mayor número de goles esperados

```

select "Home"."name" as "Team", "Home"."Home expected goals" as "Home xGoals"
, "Home"."Home goals" as "Home goals",
"Away"."Away expected goals" as "Away xGoals",
"Away"."Away goals" as "Away Goals",
("Home"."Home expected goals" + "Away"."Away expected goals") as "Total xGoals",
("Home"."Home goals" + "Away"."Away goals") as "Total Goals"
from
(select t."name", sum("xGoals") as "Home expected goals", sum("goals") as "Home goals" from teamstats ts join teams t on t."teamID" =
ts."teamID" where "location" = 'h'
group by t."name" order by
sum("goals") desc) as "Home" join
(select t."name", sum("xGoals") as "Away expected goals", sum("goals") as "Away goals" from teamstats ts join teams t on t."teamID" =
ts."teamID" where "location" = 'a' group by t."name" order by
sum("goals") desc) as "Away" on "Home"."name" = "Away"."name" order by "Total Goals" desc limit 1;

```

Query No.12 Equipo con más goles anotados

Resultado:

Team	Home xGoals	Home goals	Away xGoals	Away Goals	Total xGoals	Total Goals
Barcelona	356,460901	395	283,403425	303	639,864326	698

Tabla No.19 Equipo con más goles anotados

Se obtuvo que el FC Barcelona es tanto el equipo del que más se esperan goles como también el equipo que más goles anota.

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

Query:

```

select t."name" as "Equipo",g."season" as "Temporada", l."name" as "Liga",
((sum(case when t."teamID" = g."homeTeamID" then g."homeGoals" else 0 end) +
sum(case when t."teamID" = g."awayTeamID" then g."awayGoals" else 0 end))) as "Goles",
sum(case when t."teamID" = g."homeTeamID" then g."awayGoals" else 0 end ) +
sum(case when t."teamID" = g."awayTeamID" then g."homeGoals" else 0 end) as "Goles en contra",
((sum(case when t."teamID" = g."homeTeamID" then g."homeGoals" else 0 end) +
sum(case when t."teamID" = g."awayTeamID" then g."awayGoals" else 0 end))) -
(sum(case when t."teamID" = g."homeTeamID" then g."awayGoals" else 0 end) +
sum(case when t."teamID" = g."awayTeamID" then g."homeGoals" else 0 end)) as "Diferencia de gol",
sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" > g."awayGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" > g."homeGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end ) +
sum(case when t."teamID" = g."homeTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) as "Puntos",
rank() over (partition by l."name", g."season" order by
sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" > g."awayGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" > g."homeGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end ) +
sum(case when t."teamID" = g."homeTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end)
desc ) as "Posicion", count(*)
from games g join teams t on t."teamID" = g."homeTeamID" or t."teamID" = g."awayTeamID" join leagues l on g."leagueID" = l."leagueID"
group by t."name", g."season", l."name"
order by l."name", g."season", "Posicion";

```

Query No.13 Características de los equipos que han sido campeones

Explicación:

Se selecciona la tabla games y se une con la tabla teams para obtener a cada equipo y sus respectivos juegos, también se une con la tabla leagues para obtener los nombres de cada liga, luego se agrupa por nombre de liga, temporada y equipo, esto para distinguir las características de cada equipo en temporada y ligas en específico. Luego se selecciona la suma de los goles como visitante y como local, para lograr eso se debe evaluarse el id del equipo local en la tabla juegos es igual al Id del equipo, si eso se cumple se agrega a la suma la cantidad de goles que marcó el equipo local, se hace el mismo procedimiento para los goles en contra. También se hace una suma de 3 si el equipo gana y una suma de 1 si el equipo empató. Luego se usa la función rank para hacer un ranking en base a la liga y la temporada, también se selecciona el valor count para ver los juegos totales que disputo el equipo.

Resultado:

	Asc Equipo	123 Temporada	Asc Liga	123 Goles	123 Goles en contra	123 Diferencia de gol	123 Puntos	123 Posicion	123 count
1	Bayern Munich	2.014	Bundesliga	80	18	62	79	1	34
2	Bayern Munich	2.015	Bundesliga	80	17	63	88	1	34
3	Bayern Munich	2.016	Bundesliga	89	22	67	82	1	34
4	Bayern Munich	2.017	Bundesliga	92	28	64	84	1	34
5	Bayern Munich	2.018	Bundesliga	88	32	56	78	1	34
6	Bayern Munich	2.019	Bundesliga	100	32	68	82	1	34
7	Bayern Munich	2.020	Bundesliga	99	44	55	78	1	34
8	Barcelona	2.014	La Liga	110	21	89	94	1	38
9	Barcelona	2.015	La Liga	112	29	83	91	1	38
10	Real Madrid	2.016	La Liga	106	41	65	93	1	38
11	Barcelona	2.017	La Liga	99	29	70	93	1	38
12	Barcelona	2.018	La Liga	90	36	54	87	1	38
13	Real Madrid	2.019	La Liga	70	25	45	87	1	38
14	Atletico Madrid	2.020	La Liga	67	25	42	86	1	38
15	Paris Saint Germain	2.014	Ligue 1	83	36	47	83	1	38
16	Paris Saint Germain	2.015	Ligue 1	102	19	83	96	1	38
17	Monaco	2.016	Ligue 1	107	31	76	95	1	38
18	Paris Saint Germain	2.017	Ligue 1	108	29	79	93	1	38
19	Paris Saint Germain	2.018	Ligue 1	105	35	70	91	1	38
20	Paris Saint Germain	2.019	Ligue 1	75	24	51	68	1	27
21	Lille	2.020	Ligue 1	64	23	41	83	1	38
22	Chelsea	2.014	Premier League	73	32	41	87	1	38
23	Leicester	2.015	Premier League	68	36	32	81	1	38
24	Chelsea	2.016	Premier League	85	33	52	93	1	38
25	Manchester City	2.017	Premier League	106	27	79	100	1	38

Tabla No.20 Características de los equipos que han sido campeones

Se observa que los equipos que ganan tienen más goles anotados que partidos, además de que la cantidad de goles en contra es inferior o ligeramente superior a la cantidad de

partidos, también se observa que la diferencia de gol de los equipos ganadores es bastante grande, superando la cantidad de partidos jugados.

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)?

```

1 SELECT "teamID", mayor_probabilidad, season, name
2 FROM (
3     SELECT ts."teamID",
4           GREATEST(MAX(g."B365H"), MAX(g."B365D"), MAX(g."B365A")) AS mayor_probabilidad,
5           g.season,
6           t.name,
7           ROW_NUMBER() OVER (PARTITION BY g.season ORDER BY GREATEST(MAX(g."B365H"), MAX(g."B365D"), MAX(g."B365A")) DESC) AS row_num
8     FROM teamstats ts
9    JOIN games g ON ts."gameID" = g."gameID"
10   JOIN teams t ON ts."teamID" = t."teamID"
11   GROUP BY g.season, ts."teamID", t.name
12 ) AS ranked_data
13 WHERE row_num = 1
14 ORDER BY season;

```

	teamID bigint	mayor_probabilidad double precision	season bigint	name text
1	148	41	2014	Barcelona
2	148	41	2015	Barcelona
3	148	41	2016	Barcelona
4	98	34	2017	Juventus
5	87	41	2018	Liverpool
6	117	34	2019	Bayern Munich
7	117	28	2020	Ravenn Munich

Query No.14 equipos que tenían la mayor probabilidad de ganar en cada una de las temporadas

Explicación: Seleccionamos teamID, mayor_probabilidad, season y name(equipo) como los datos que vamos a mostrar, usamos la función Greatest para encontrar el valor máximo entre las tres probabilidades de Bet365 (B365H, B365D y B365A) y llamamos a toda esta función mayor_probabilidad, utilizamos Row_number para asignar un número de fila a cada combinación de temporada y equipo, usamos Joins para unir la tabla teamstats y team y agrupamos teamID, season y name(equipo) y ordenamos por season.

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.

```

1 SELECT ts."teamID", t.name,
2       SUM(ts."fouls") AS total_faltas,
3       SUM(ts."yellowCards") AS total_tarjetas_amarillas,
4       SUM(ts."redCards") AS total_tarjetas_rojas
5 FROM teamstats ts
6 JOIN teams t ON ts."teamID" = t."teamID"
7 GROUP BY ts."teamID", t.name
8 ORDER BY total_faltas ASC, total_tarjetas_amarillas ASC, total_tarjetas_rojas ASC
9 LIMIT 10;
10

```

Data Output Messages Notifications

	teamID bigint	name text	total_faltas numeric	total_tarjetas_amarillas numeric	total_tarjetas_rojas numeric
1	234	Nuernberg	376	59	4
2	227	Cardiff	384	65	1
3	262	Arminia Bielefeld	391	51	1
4	245	Leeds	427	61	1
5	261	Cadiz	435	80	3
6	202	Queens Park Rangers	447	75	3
7	243	Lecce	463	99	5
8	93	Middlesbrough	478	76	1
9	206	Cordoba	494	93	9
10	242	Brescia	494	95	6

Query No.15 Equipos más limpios

```

1 SELECT ts."teamID", t.name,
2       SUM(ts."fouls") AS total_faltas,
3       SUM(ts."yellowCards") AS total_tarjetas_amarillas,
4       SUM(ts."redCards") AS total_tarjetas_rojas
5 FROM teamstats ts
6 JOIN teams t ON ts."teamID" = t."teamID"
7 GROUP BY ts."teamID", t.name
8 ORDER BY total_faltas DESC, total_tarjetas_amarillas DESC, total_tarjetas_rojas DESC
9 LIMIT 10;
10

```

Data Output Messages Notifications

	teamID bigint	name text	total_faltas numeric	total_tarjetas_amarillas numeric	total_tarjetas_rojas numeric
1	113	Torino	4131	612	29
2	101	Genoa	4074	640	47
3	107	Atalanta	3897	558	38
4	138	Sevilla	3854	710	26
5	110	Florentina	3757	637	33
6	142	Getafe	3722	714	34
7	156	Eibar	3717	625	27
8	96	Lazio	3705	663	40
9	104	Sassuolo	3664	636	39
10	111	AC Milan	3643	596	50

Query No.16 Equipos más sucios

Explicación: seleccionamos teamID de la tabla teamstats y el nombre del equipo (name) de la tabla teams. Además, calculamos la suma de las faltas, tarjetas amarillas y tarjetas rojas para cada equipo utilizando la función de agregación SUM(). Renombramos estas sumas

como total_faltas, total_tarjetas_amarillas y total_tarjetas_rojas respectivamente. Hacemos un JOIN para combinar las tablas por medio de teamID agrupamos los resultados por el identificador del equipo y el nombre del equipo, ordenamos los resultados en función del número total de faltas, tarjetas amarillas y tarjetas rojas en orden descendente con DESC si queremos saber los equipos que juegan más sucio y con ASC para ordenar de forma ascendente para los equipos que juegan más limpio y limitamos los resultados a 10 con LIMIT.

11. ¿Cuáles son los equipos que regularmente ganan cuando su probabilidad de victoria es baja?

Query:

```
select "Less home probability"."name",
("Less home probability"."sumHome" + "Less away probability"."sumAway") as "Partidos en donde el contexto no era favorable"
from
(select t."name", count(*) as "sumHome" from games g join teams t on t."teamID" = g."homeTeamID"
where "homeProbability" < "awayProbability" and "homeGoals" > "awayGoals"
group by t."name" order by count(*) desc) as "Less home probability"
join
(select t."name", count(*) as "sumAway" from games g join teams t on t."teamID" = g."awayTeamID"
where "awayProbability" < "homeProbability" and "awayGoals" > "homeGoals"
group by t."name" order by count(*) desc) as "Less away probability"
on "Less home probability"."name" = "Less away probability"."name"
order by "Partidos en donde el contexto no era favorable" desc ;
```

Query No.16 Equipos que más ganan en contextos no favorables

Resultado:

	abc Equipo ▼	123 Partidos en donde el contexto no era favorable ▼	
1	Nice	32	
2	Genoa	28	
3	Burnley	28	
4	Levante	27	
5	Manchester United	27	
6	Newcastle United	25	
7	Valencia	24	
8	West Ham	24	
9	Torino	24	
10	Saint-Etienne	24	
11	Werder Bremen	23	
12	Rennes	23	
13	Juventus	23	
14	Atletico Madrid	22	
15	AC Milan	22	
16	Borussia M.Gladbach	22	
17	Hoffenheim	22	
18	Eintracht Frankfurt	22	
19	Athletic Club	22	
20	Sampdoria	22	
21	Villarreal	21	
22	Real Betis	21	
23	Celta Vigo	21	
24	Schalke 04	20	

Tabla 21. Equipos con más victorias en contextos poco favorables

Explicación: Se forma una tabla llamada “Less home probability” uniendo la tabla games y la tabla teams con el criterio “homeTeamID” = “teamID”, una vez con la unión se aplican filtros los cuáles permiten obtener las tuplas en donde el valor “homeProbability” era menor que “awayProbability” y el resultado fue una victoria, ya con las tuplas, se agrupa la tabla mediante el nombre de cada equipo. Finalmente se selecciona el nombre de cada equipo y la cantidad de veces que aparece en la tabla. Se realiza un proceso similar para obtener aquellos partidos en donde el equipo jugó de visitante. Ya con 2 tablas, una que nos indica los partidos de local en donde se cumple que hubo una victoria a pesar de tener una probabilidad de victoria, mientras la otra tabla indica los partidos en donde se cumplieron las condiciones cuando un equipo jugó de local. Estas 2 tablas resultantes se unen con el criterio: “Less home probability”.name = “Less away probability”.name, con la tabla producto de la unión, se suma la cantidad de partidos en donde se cumplieron las condiciones, tanto de local como visitante.

12. ¿Cuáles son los equipos que más pierden cuando su probabilidad de victoria es mayor?

Query:


```

select "Most home probability"."name" as "Equipo",
("Most home probability"."sumHome" + "Most away probability"."sumAway") as "Partidos en donde el contexto era favorable"
from
(select t."name", count(*) as "sumHome" from games g join teams t on t."teamID" = g."homeTeamID"
where "homeProbability" > "awayProbability" and "homeGoals" < "awayGoals"
group by t."name" order by count(*) desc) as "Most home probability"
join
(select t."name", count(*) as "sumAway" from games g join teams t on t."teamID" = g."awayTeamID"
where "awayProbability" > "homeProbability" and "awayGoals" < "homeGoals"
group by t."name" order by count(*) desc) as "Most away probability"
on "Most home probability"."name" = "Most away probability"."name"
order by "Partidos en donde el contexto era favorable" desc ;

```

Query No.17 Equipos que pierden en contextos favorables

Explicación: se usa el query del inciso anterior, pero esta vez se buscan aquellas tuplas en donde la probabilidad de victoria es mayor que la probabilidad de derrota pero el resultado fue una derrota

Resultado:

	Asc Equipo	123 Partidos en donde el contexto era favorable
1	Villarreal	30
2	Genoa	30
3	Southampton	30
4	Udinese	29
5	Crystal Palace	25
6	Napoli	24
7	Nantes	24
8	Chelsea	24
9	Borussia M.Gladbach	23
10	Leicester	23
11	VfB Stuttgart	23
12	Eibar	23
13	Marseille	23
14	Toulouse	23
15	Valencia	22
16	Angers	22
17	Atalanta	22
18	Sevilla	22
19	Lyon	22
20	Manchester City	22
21	Rennes	21
22	Saint-Etienne	21
23	Fiorentina	21
24	Burnley	21
25	Bordeaux	20

Tabla No.22 Equipos que más pierden en contextos favorables

13. ¿Quiénes son los jugadores con mejor promedio goleador?

Query:

```
--Jugadores con mejor promedio goleador
select p."name", sum("goals")/count(*) as "PromGoleador" from appearances a join players p
on a."playerID" = p."playerID"
group by p."name" having count(*) > 35 order by sum("goals")/count(*) desc ;
```

Query No.18: Promedio goleador de los jugadores

Resultado:

	ABC name	123 PromGoleador
1	Cristiano Ronaldo	0,9598214286
2	Lionel Messi	0,950617284
3	Erling Haaland	0,9302325581
4	Robert Lewandowski	0,9269406393
5	Zlatan Ibrahimovic	0,792
6	Kylian Mbappe-Lottin	0,722972973
7	Harry Kane	0,7056277056
8	Sergio Aguero	0,7021276596
9	Luis Suarez	0,692
10	Neymar	0,6826347305
11	Edinson Cavani	0,6734693878
12	Pierre-Emerick Aubam	0,6592920354
13	Ciro Immobile	0,5964125561
14	Mauro Icardi	0,5873786408
15	Mohamed Salah	0,5676855895
16	Alexandre Lacazette	0,56
17	Andre-Pierre Gignac	0,5526315789
18	Alexander Meier	0,552238806
19	Gonzalo Higuain	0,5485436893
20	Cheick Diabate	0,5483870968
21	Karim Benzema	0,5357142857
22	Wout Weghorst	0,53
23	Romelu Lukaku	0,5161290323
24	Iago Aspas	0,4906542056

Tabla No.23 Jugadores con mejor promedio de gol

Explicación: Se seleccionan las tablas appearances y players y se unen con el criterio de que el valor "playerID" de player debe ser igual al valor "playerID" de appearances, esto permite obtener a cada jugador y sus respectivas apariciones en cada partido que ha disputado. Ya con esta tabla formada, se agrupa la tabla por el nombre del jugador creando una partición en la tabla formada anteriormente, finalmente se selecciona el nombre de cada jugador y el cociente entre su cantidad de goles y la cantidad de veces en las que aparece en la tabla, se aplica el filtro de que la cantidad de veces que el jugador aparece en la tabla debe ser mayor a 35, pues puede ser que en la tabla se encuentren jugadores que solamente jugaron un partido y anotaron un solo gol, lo cuál afecta el orden de la tabla pues estos jugadores serían los que poseen promedios más altos, sin embargo si se aplica el filtro se asegura la inclusión a jugadores que solamente han jugado algunas temporadas pero tienen una gran proyección a futuro ya que están siendo importantes para su equipo.

14. ¿Cuáles son los equipos que más han mejorado en la última temporada con respecto a la primera temporada?

Query:

```

select "Puntos_2014"."Equipo", "Puntos_2014"."Liga", "Puntos_2020"."Puntos" - "Puntos_2014"."Puntos" as "Diferencia de puntos"
from
(
  (select "Equipo", "Liga", "Temporada", "Puntos" from
  (select t."name" as "Equipo", g."season" as "Temporada", l."name" as "Liga",
  ((sum(case when t."teamID" = g."homeTeamID" then g."homeGoals" else 0 end) +
  sum(case when t."teamID" = g."awayTeamID" then g."awayGoals" else 0 end))) as "Goles",
  sum(case when t."teamID" = g."homeTeamID" then g."awayGoals" else 0 end) +
  sum(case when t."teamID" = g."awayTeamID" then g."homeGoals" else 0 end) as "Goles en contra",
  ((sum(case when t."teamID" = g."homeTeamID" then g."homeGoals" else 0 end) +
  sum(case when t."teamID" = g."awayTeamID" then g."awayGoals" else 0 end))) -
  (sum(case when t."teamID" = g."homeTeamID" then g."awayGoals" else 0 end) +
  sum(case when t."teamID" = g."awayTeamID" then g."homeGoals" else 0 end)) as "Diferencia de gol",
  sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" > g."awayGoals" then 3 else 0 end) +
  sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" > g."homeGoals" then 3 else 0 end) +
  sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) +
  sum(case when t."teamID" = g."homeTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) as "Puntos",
  rank() over (partition by l."name", g."season" order by
  sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" > g."awayGoals" then 3 else 0 end) +
  sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" > g."homeGoals" then 3 else 0 end) +
  sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) +
  sum(case when t."teamID" = g."homeTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end)
  desc ) as "Posicion", count(*)
  from games g join teams t on t."teamID" = g."homeTeamID" or t."teamID" = g."awayTeamID" join leagues l on g."leagueID" = l."leagueID"
  group by t."name", g."season", l."name"
  order by l."name", g."season", "Posicion")
  where "Temporada" = 2014) as "Puntos_2014"

join
(select "Equipo", "Liga", "Temporada", "Puntos" from
  (select t."name" as "Equipo", g."season" as "Temporada", l."name" as "Liga",
  ((sum(case when t."teamID" = g."homeTeamID" then g."homeGoals" else 0 end) +
  sum(case when t."teamID" = g."awayTeamID" then g."awayGoals" else 0 end))) as "Goles",
  sum(case when t."teamID" = g."homeTeamID" then g."awayGoals" else 0 end) +
  sum(case when t."teamID" = g."awayTeamID" then g."homeGoals" else 0 end) as "Goles en contra",
  ((sum(case when t."teamID" = g."homeTeamID" then g."homeGoals" else 0 end) +
  sum(case when t."teamID" = g."awayTeamID" then g."awayGoals" else 0 end))) -
  (sum(case when t."teamID" = g."homeTeamID" then g."awayGoals" else 0 end) +
  sum(case when t."teamID" = g."awayTeamID" then g."homeGoals" else 0 end)) as "Diferencia de gol",
  sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" > g."awayGoals" then 3 else 0 end) +
  sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" > g."homeGoals" then 3 else 0 end) +
  sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) +
  sum(case when t."teamID" = g."homeTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) as "Puntos",
  rank() over (partition by l."name", g."season" order by
  sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" > g."awayGoals" then 3 else 0 end) +
  sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" > g."homeGoals" then 3 else 0 end) +
  sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) +
  sum(case when t."teamID" = g."homeTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end)
  desc ) as "Posicion", count(*)
  from games g join teams t on t."teamID" = g."homeTeamID" or t."teamID" = g."awayTeamID" join leagues l on g."leagueID" = l."leagueID"
  group by t."name", g."season", l."name"
  order by l."name", g."season", "Posicion")
  where "Temporada" = 2020) as "Puntos_2020"
on "Puntos_2014"."Equipo" = "Puntos_2020"."Equipo"
) order by "Diferencia de puntos" desc;

```

Query No.19 Equipos que han mejorado con respecto a la primera temporada

Explicación: se crean 2 nuevas tablas llamadas “puntos en 2014” y “puntos en 2020” respectivamente, estas tablas se crean siguiendo la estructura del Query No.13, pero con una variación, solamente se selecciona la primera y la última temporada. Las tablas se unen y se encuentra la diferencia entre los puntos de la última temporada y la primera temporada.

Resultado:

	ABC Equipo	ABC Liga	123 Diferencia de puntos
1	Atalanta	Serie A	41
2	Inter	Serie A	36
3	Lens	Ligue 1	28
4	AC Milan	Serie A	27
5	Lille	Ligue 1	27
6	Leicester	Premier League	25
7	Borussia Dortmund	Bundesliga	18
8	West Ham	Premier League	18
9	Eintracht Frankfurt	Bundesliga	17
10	Aston Villa	Premier League	17
11	Metz	Ligue 1	17
12	Real Sociedad	La Liga	16
13	Napoli	Serie A	14
14	Sassuolo	Serie A	13
15	Everton	Premier League	12
16	Granada	La Liga	11
17	Freiburg	Bundesliga	11
18	VfB Stuttgart	Bundesliga	9
19	Atletico Madrid	La Liga	8
20	Rennes	Ligue 1	8
21	Manchester City	Premier League	7
22	Liverpool	Premier League	7
23	Monaco	Ligue 1	7
24	Burnley	Premier League	6

Tabla No 24 equipos que han mejorado su rendimiento

Se puede observar que el equipo Atalanta de Bérgamo es el equipo que más ha mejorado, pues tiene una diferencia de 41 puntos con respecto a la primera temporada. El AC Milán y el Inter son también un equipo que ha mejorado bastante.

15. ¿Cuáles son los equipos que han empeorado?

Query:

```
select "Puntos_2014"."Equipo", "Puntos_2014"."Liga", "Puntos_2020"."Puntos" - "Puntos_2014"."Puntos" as "Diferencia de puntos"
from
(
select "Equipo", "Liga", "Temporada", "Puntos" from
(select t."name" as "Equipo", g."season" as "Temporada", l."name" as "Liga",
((sum(case when t."teamID" = g."homeTeamID" then g."homeGoals" else 0 end) +
sum(case when t."teamID" = g."awayTeamID" then g."awayGoals" else 0 end))) as "Goles",
sum(case when t."teamID" = g."homeTeamID" then g."awayGoals" else 0 end) +
sum(case when t."teamID" = g."awayTeamID" then g."homeGoals" else 0 end) as "Goles en contra",
((sum(case when t."teamID" = g."homeTeamID" then g."homeGoals" else 0 end) +
sum(case when t."teamID" = g."awayTeamID" then g."awayGoals" else 0 end))) -
(sum(case when t."teamID" = g."homeTeamID" then g."awayGoals" else 0 end) +
sum(case when t."teamID" = g."awayTeamID" then g."homeGoals" else 0 end)) as "Diferencia de gol",
sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" > g."awayGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" > g."homeGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) +
sum(case when t."teamID" = g."homeTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) as "Puntos",
rank() over (partition by l."name", g."season" order by
sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" > g."awayGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" > g."homeGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) +
sum(case when t."teamID" = g."homeTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end)
desc ) as "Posicion", count(*)
from games g join teams t on t."teamID" = g."homeTeamID" or t."teamID" = g."awayTeamID" join leagues l on g."leagueID" = l."leagueID"
group by t."name", g."season", l."name"
order by l."name", g."season", "Posicion")
where "Temporada" = 2014) as "Puntos_2014"
```

```

join
(select "Equipo", "Liga", "Temporada", "Puntos" from
(select t."name" as "Equipo", g."season" as "Temporada", l."name" as "Liga",
((sum(case when t."teamID" = g."homeTeamID" then g."homeGoals" else 0 end) +
sum(case when t."teamID" = g."awayTeamID" then g."awayGoals" else 0 end))) as "Goles",
sum(case when t."teamID" = g."homeTeamID" then g."awayGoals" else 0 end) +
sum(case when t."teamID" = g."awayTeamID" then g."homeGoals" else 0 end) as "Goles en contra",
((sum(case when t."teamID" = g."homeTeamID" then g."homeGoals" else 0 end) +
sum(case when t."teamID" = g."awayTeamID" then g."awayGoals" else 0 end))) -
(sum(case when t."teamID" = g."homeTeamID" then g."awayGoals" else 0 end) +
sum(case when t."teamID" = g."awayTeamID" then g."homeGoals" else 0 end)) as "Diferencia de gol",
sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" > g."awayGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" > g."homeGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) +
sum(case when t."teamID" = g."homeTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) as "Puntos",
rank() over (partition by l."name", g."season" order by
sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" > g."awayGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" > g."homeGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) +
sum(case when t."teamID" = g."homeTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end)
desc ) as "Posicion", count(*)
from games g join teams t on t."teamID" = g."homeTeamID" or t."teamID" = g."awayTeamID" join leagues l on g."leagueID" = l."leagueID"
group by t."name", g."season", l."name"
order by l."name", g."season", "Posicion")
where "Temporada" = 2020) as "Puntos 2020"
on "Puntos 2014"."Equipo" = "Puntos 2020"."Equipo"
) order by "Diferencia de puntos" asc;

```

Query No.20. Equipos que han empeorado su rendimiento con respecto a la primera temporada

Explicación: se usa la misma estructura que el query anterior, con la variación de que ahora se ordena de manera ascendente

Resultado:

	ABC Equipo	ABC Liga	123 Diferencia de puntos
1	Valencia	La Liga	-34
2	Schalke 04	Bundesliga	-32
3	Fiorentina	Serie A	-24
4	Saint-Etienne	Ligue 1	-23
5	Chelsea	Premier League	-20
6	West Bromwich Albion	Premier League	-18
7	Bordeaux	Ligue 1	-18
8	Borussia M.Gladbach	Bundesliga	-17
9	Genoa	Serie A	-17
10	Torino	Serie A	-17
11	Southampton	Premier League	-17
12	Barcelona	La Liga	-15
13	Arsenal	Premier League	-14
14	Augsburg	Bundesliga	-13
15	Werder Bremen	Bundesliga	-12
16	Bayer Leverkusen	Bundesliga	-9
17	Athletic Club	La Liga	-9
18	Marseille	Ligue 1	-9
19	Juventus	Serie A	-9
20	Wolfsburg	Bundesliga	-8
21	Real Madrid	La Liga	-8
22	Roma	Serie A	-7
23	FC Cologne	Bundesliga	-7
24	Parma Calcio 1913	Serie A	-6

Tabla No.25 Equipos que han empeorado su rendimiento

Se observa que el Valencia es un equipo que ha estado empeorando en los últimos años, el Chelsea también es un equipo que posiblemente tenga peor rendimiento.

16. Según la casa de apuestas Bet365, ¿cuáles son los equipos que más veces ganan cuando la probabilidad de victoria es menor?

Query:

```
--Equipos que ganan cuando no tienen un contexto favorable según las apuestas
select "Less home probability"."name" as "Equipo",
("Less home probability"."sumHome" + "Less away probability"."sumAway") as "Partidos en donde el contexto no era favorable según Bet365"
from
(select t."name", count(*) as "sumHome" from games g join teams t on t."teamID" = g."homeTeamID"
where 1/"B365H" < 1/"B365A" and "homeGoals" > "awayGoals" and "B365H" >0 and "B365A" >0
group by t."name" order by count(*) desc) as "Less home probability"
join
(select t."name", count(*) as "sumAway" from games g join teams t on t."teamID" = g."awayTeamID"
where 1/"B365A" < 1/"B365H" and "awayGoals" > "homeGoals" and "B365H" >0 and "B365A" >0
group by t."name" order by count(*) desc) as "Less away probability"
on "Less home probability"."name" = "Less away probability"."name"
order by "Partidos en donde el contexto no era favorable según Bet365" desc ;
```

Query No.21 Cantidad de veces en las que un equipo gana cuando las apuestas no lo favorecen

Explicación: sigue la estructura del query del inciso 11, con la variación de que en este query los criterios "AwayProbability" y "HomeProbability" son reemplazados por la probabilidad de que un equipo gane o pierda según la casa de apuestas Bet365.

Resultado:

	ABC Equipo	123 Partidos en donde el contexto no era favorable según Bet365
1	Burnley	48
2	Nice	48
3	West Ham	45
4	Sampdoria	43
5	Crystal Palace	43
6	Sassuolo	41
7	Newcastle United	39
8	Freiburg	37
9	Augsburg	37
10	Levante	37
11	Werder Bremen	36
12	Alaves	36
13	Mainz 05	36
14	Leicester	35
15	Eibar	35
16	Nantes	35
17	Genoa	35
18	Bordeaux	34
19	Real Betis	33
20	Eintracht Frankfurt	33
21	Montpellier	32
22	Udinese	31
23	Hertha Berlin	31
24	Cagliari	30

Tabla No.26 Partidos en donde un equipo ganó a pesar de que sus probabilidades según Bet365 no eran favorables

17. Según la casa de apuestas Bet365, ¿cuáles son los equipos que más veces ganan cuando la probabilidad de victoria es menor?

Query:

```
--Equipos que pierden cuando sus probabilidades de ganar son altas
select "Most home probability"."name" as "Equipo",
("Most home probability"."sumHome" + "Most away probability"."sumAway") as "Partidos en donde el contexto era favorable seegún Bet365"
from
(select t."name", count(*) as "sumHome" from games g join teams t on t."teamID" = g."homeTeamID"
where "homeProbability" > "awayProbability" and "homeGoals" < "awayGoals"
group by t."name" order by count(*) desc) as "Most home probability"
join
(select t."name", count(*) as "sumAway" from games g join teams t on t."teamID" = g."awayTeamID"
where "awayProbability" > "homeProbability" and "awayGoals" < "homeGoals"
group by t."name" order by count(*) desc) as "Most away probability"
on "Most home probability"."name" = "Most away probability"."name"
order by "Partidos en donde el contexto era favorable" desc ;

--Jugadores con mejor promedio goleador
select p."name", sum("goals")/count(*) as "PromGoleador" from appeareances a join players p
on a."playerID" = p."playerID"
group by p."name" having count(*) > 35 order by sum("goals")/count(*) desc ;
```

Explicación: posee el mismo mecanismo del query del inciso anterior con la variación de que ahora se buscan las tuplas en donde la probabilidad de victoria según Bet365 fue mayor pero el equipo obtuvo una derrota.

Resultado:

	ABC Equipo	123 Partidos en donde el contexto era favorable
1	Lyon	45
2	Bayer Leverkusen	45
3	Chelsea	43
4	Villarreal	42
5	Arsenal	42
6	Borussia Dortmund	41
7	Southampton	40
8	Real Sociedad	40
9	Napoli	39
10	Everton	38
11	Monaco	38
12	Manchester City	38
13	Inter	38
14	Manchester United	37
15	Lazio	37
16	Borussia M.Gladbach	36
17	Tottenham	36
18	Roma	34
19	Schalke 04	34
20	Rennes	34
21	Real Madrid	34
22	Bordeaux	33
23	Lille	33
24	Marseille	33
25	AC Milan	33

Tabla No.27 equipos con más derrotas en contextos favorables.

18. ¿Qué factores determinan la posición de un equipo al finalizar la temporada?

Query:

```
select t."name" as "Equipo", g."season" as "Temporada", l."name" as "Liga",
((sum(case when t."teamID" = g."homeTeamID" then g."homeGoals" else 0 end) +
sum(case when t."teamID" = g."awayTeamID" then g."awayGoals" else 0 end))) as "GF",
sum(case when t."teamID" = g."homeTeamID" then g."awayGoals" else 0 end) +
sum(case when t."teamID" = g."awayTeamID" then g."homeGoals" else 0 end) as "GC",
((sum(case when t."teamID" = g."homeTeamID" then g."homeGoals" else 0 end) +
sum(case when t."teamID" = g."awayTeamID" then g."awayGoals" else 0 end))) -
(sum(case when t."teamID" = g."homeTeamID" then g."awayGoals" else 0 end) +
sum(case when t."teamID" = g."awayTeamID" then g."homeGoals" else 0 end)) as "GD",
sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" > g."awayGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" > g."homeGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) +
sum(case when t."teamID" = g."homeTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) as "Puntos",
rank() over (partition by l."name", g."season" order by
sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" > g."awayGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" > g."homeGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) +
sum(case when t."teamID" = g."homeTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end)
desc ) as "Posicion", count(*) as "Partidos jugados",
sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" > g."awayGoals" then 1 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" > g."homeGoals" then 1 else 0 end) as "W",
sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" < g."awayGoals" then 1 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" < g."homeGoals" then 1 else 0 end) as "L",
sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" = g."awayGoals" then 1 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) as "D"
from games g join teams t on t."teamID" = g."homeTeamID" or t."teamID" = g."awayTeamID" join leagues l on g."leagueID" = l."leagueID"
group by t."name", g."season", l."name"
order by l."name", g."season", "Posicion"
```

Query No. 22 Posición, cantidad de goles y puntos obtenidos por cad equipo en cada temporada

Explicación: Se usa la estructura del query No.13, pero en esta ocasión se obtienen los puntos, goles y posiciones de todos los equipos en cada temporada y en cada liga para determinar si los goles anotados en una temporada determinan la posición final del equipo al final. Con estos datos es posible generar un diagrama de dispersión para aplicar una regresión lineal y determinar si el número de goles puede ser determinante para la posición final de un equipo al final de la temporada. Para poder hacer los diagramas de dispersión se hizo un script de python el cuál trasladaba los datos del query a un archivo de excel, este script está adjuntado en la entrega de este proyecto.

Resultados:

	abc Equipo	123 Temporada	abc Liga	123 Goles	123 Puntos	123 Posicion	123 count
1	Bayern Munich	2.014	Bundesliga	80	79	1	34
2	Wolfsburg	2.014	Bundesliga	72	69	2	34
3	Borussia M.Gladbach	2.014	Bundesliga	53	66	3	34
4	Bayer Leverkusen	2.014	Bundesliga	62	61	4	34
5	Augsburg	2.014	Bundesliga	43	49	5	34
6	Schalke 04	2.014	Bundesliga	42	48	6	34
7	Borussia Dortmund	2.014	Bundesliga	47	46	7	34
8	Hoffenheim	2.014	Bundesliga	49	44	8	34
9	Werder Bremen	2.014	Bundesliga	50	43	9	34
10	Eintracht Frankfurt	2.014	Bundesliga	56	43	9	34
11	Mainz 05	2.014	Bundesliga	45	40	11	34
12	FC Cologne	2.014	Bundesliga	34	40	11	34
13	Hannover 96	2.014	Bundesliga	40	37	13	34
14	VfB Stuttgart	2.014	Bundesliga	42	36	14	34
15	Hertha Berlin	2.014	Bundesliga	36	35	15	34
16	Hamburger SV	2.014	Bundesliga	25	35	15	34
17	Freiburg	2.014	Bundesliga	36	34	17	34
18	Paderborn	2.014	Bundesliga	31	31	18	34
19	Bayern Munich	2.015	Bundesliga	80	88	1	34
20	Borussia Dortmund	2.015	Bundesliga	82	78	2	34
21	Bayer Leverkusen	2.015	Bundesliga	56	60	3	34
22	Borussia M.Gladbach	2.015	Bundesliga	67	55	4	34
23	Schalke 04	2.015	Bundesliga	51	52	5	34
24	Hertha Berlin	2.015	Bundesliga	42	50	6	34
25	Mainz 05	2.015	Bundesliga	46	50	6	34

Tabla No. 28 resultados y posiciones de los equipos en cada temporada

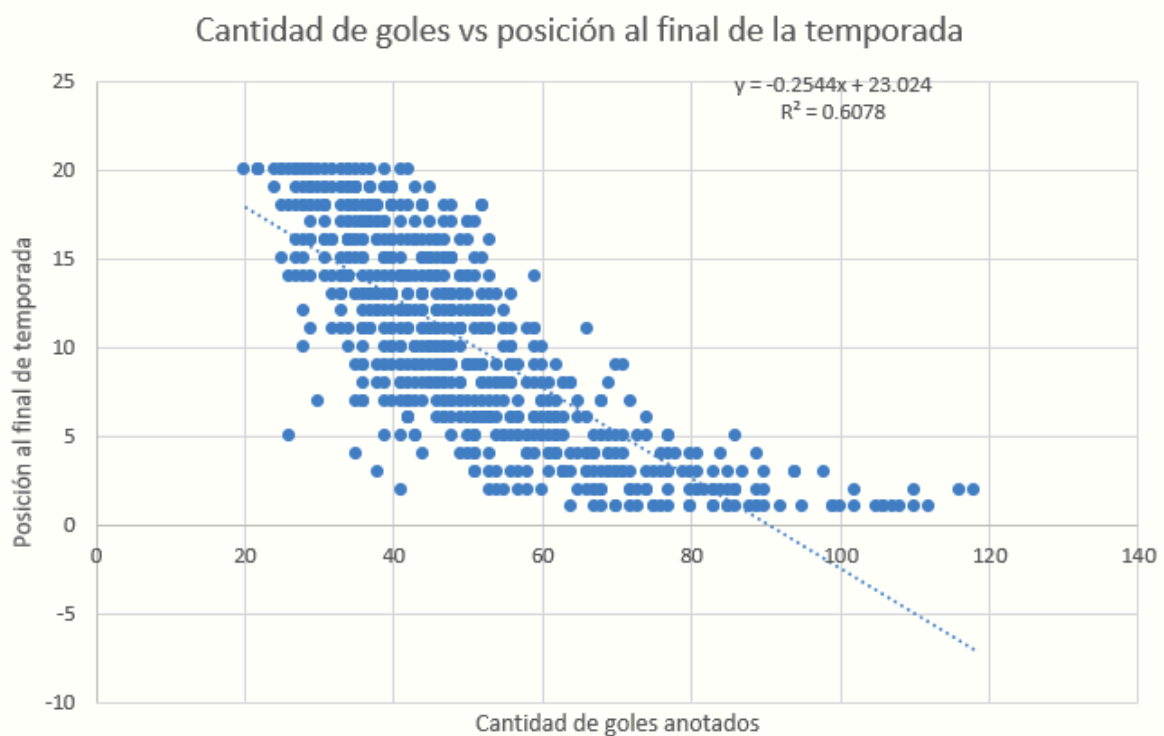


Figura No.1 Regresión lineal entre la cantidad de goles anotados y la posición al final de la temporada

En base a los resultados obtenidos se puede decir que en la mayor parte de los casos, la cantidad de goles anotados por un equipo puede llegar a determinar su posición final al terminar la temporada regular, sin embargo hay una cantidad considerable de casos en la que la cantidad de goles no determina la posición del equipo al finalizar la temporada, por eso el coeficiente R^2 no tiene un valor más elevado.



Figura No.2 Regresión lineal entre la cantidad de victorias y la posición al final de la temporada

Se puede afirmar que la cantidad de victorias es un factor muy importante a la hora de determinar la posición final en la temporada, pues el coeficiente de correlación es bastante alto e indica que la mayor parte de las veces los equipos con más victorias suelen terminar en los puestos más altos.

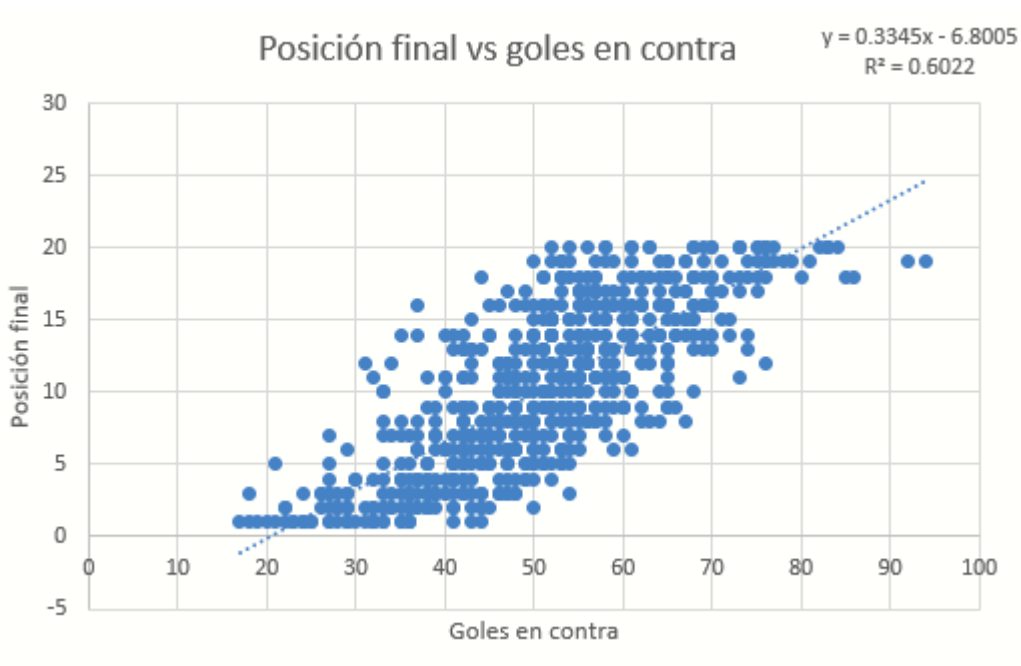


Figura No.3 Regresión lineal entre la cantidad de victorias y la posición al final de la temporada

Al igual que los goles a favor, la mayor parte el tiempo los equipos con más goles en contr suelen quedar de último hay una considerable cantidad de casos en donde esto no es así.



Figura No.4 Regresión lineal entre la cantidad de derrotas y la posición al final de la temporada

Ya que el coeficiente de correlación es bastante elevado, la cantidad de derrotas juega un papel clave a la hora de determinar si un equipo queda en las primeras posiciones.

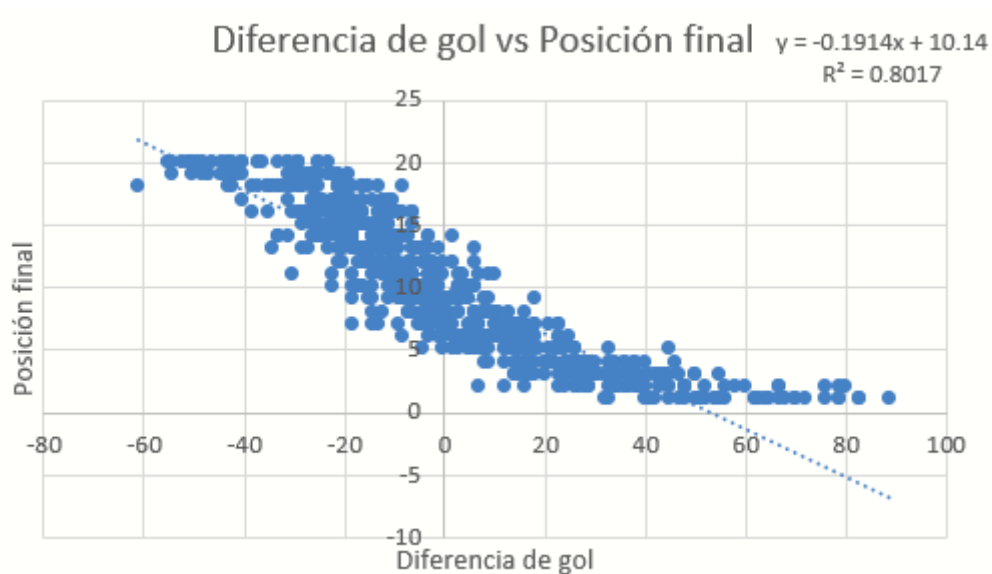


Figura No.5 Regresión lineal entre la cantidad de derrotas y la posición al final de la temporada

A diferencia de otros factores goleadores como la cantidad de goles, la diferencia de gol es fundamental para determinar si un equipo puede estar entre las primeras posiciones de su respectiva competición, los equipos con diferencia de gol más elevada son los que pueden competir por los mejores puestos.

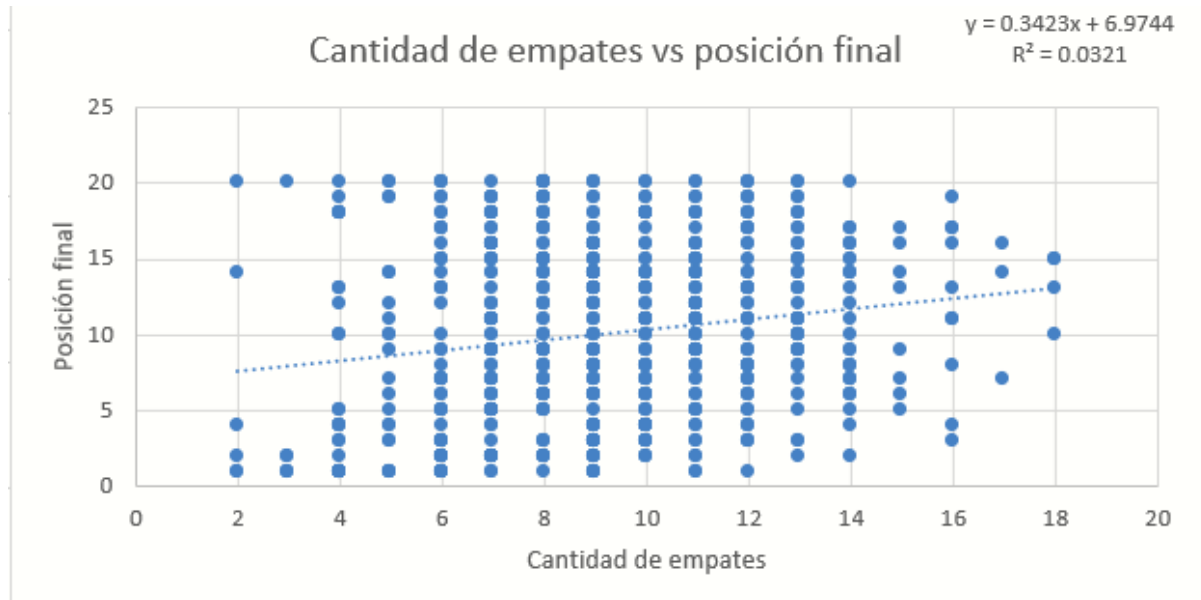


Figura No.6 Regresión lineal entre la cantidad de empates y la posición al final de la temporada

La cantidad de empates es un factor irrelevante a la hora de determinar qué equipo acaba en las primeras posiciones, los equipos que han sido campeones han tenido una cantidad de empate bastante distinta.

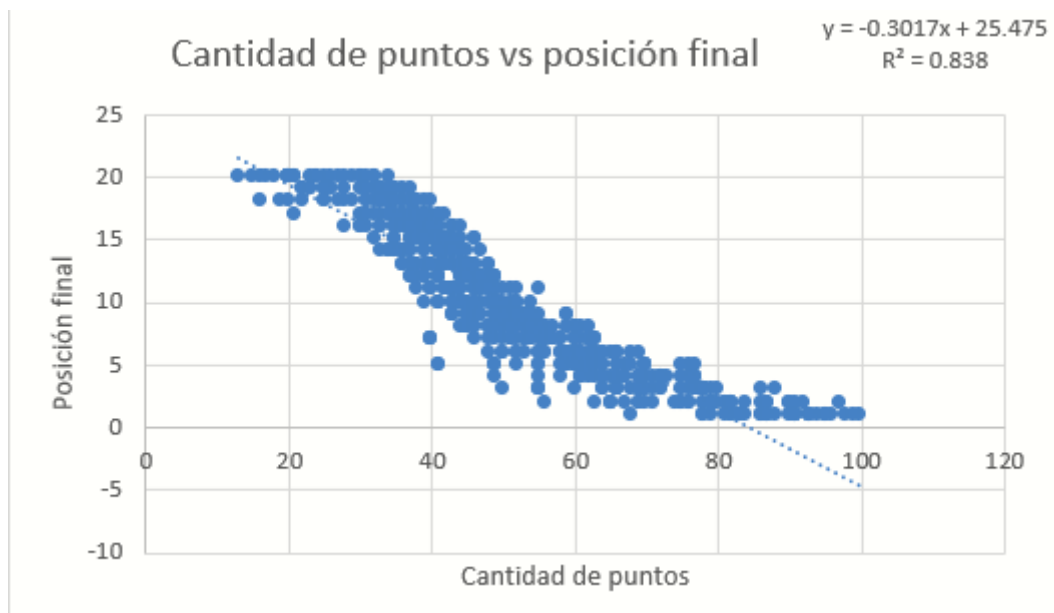


Figura No.7 Regresión lineal entre la cantidad de puntos y la posición al final de la temporada

Es bastante evidente que a mientras más puntos consiga un equipo, mejor va a ser su posición al final de la temporada, sin embargo también es relevante tener en cuenta cuáles son las cantidades de puntos que regularmente te acercan a las primeras posiciones.

19. ¿En qué posición acaba un equipo en promedio al final de la temporada?
 ¿Cuántas victorias, derrotas y empates tiene cada equipo en promedio?
 ¿Cuántos goles anotó un equipo en promedio?

Query:

```
select "Stats"."Equipo", avg("Stats"."W") as "WP",
avg("Stats"."D") as "DP",
avg("Stats"."GF") as "GFP",
avg("Stats"."GC") as "GCP",
avg("Stats"."GD") as "GDP",
avg("Stats"."L") as "LP",
avg("Stats"."Puntos") as "Puntos promedio",
avg("Stats"."Posicion") as "Poscion promedio"
from
(select t."name" as "Equipo", g."season" as "Temporada", l."name" as "Liga",
((sum(case when t."teamID" = g."homeTeamID" then g."homeGoals" else 0 end) +
sum(case when t."teamID" = g."awayTeamID" then g."awayGoals" else 0 end))) as "GF",
sum(case when t."teamID" = g."homeTeamID" then g."awayGoals" else 0 end) +
sum(case when t."teamID" = g."awayTeamID" then g."homeGoals" else 0 end) as "GC",
((sum(case when t."teamID" = g."homeTeamID" then g."homeGoals" else 0 end) +
sum(case when t."teamID" = g."awayTeamID" then g."awayGoals" else 0 end))) -
(sum(case when t."teamID" = g."homeTeamID" then g."awayGoals" else 0 end) +
sum(case when t."teamID" = g."awayTeamID" then g."homeGoals" else 0 end)) as "GD",
sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" > g."awayGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" > g."homeGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) +
sum(case when t."teamID" = g."homeTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) as "Puntos",
rank() over (partition by l."name", g."season" order by
sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" > g."awayGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" > g."homeGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) +
sum(case when t."teamID" = g."homeTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end)
desc ) as "Posicion", count(*) as "Partidos jugados",
sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" > g."awayGoals" then 1 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" > g."homeGoals" then 1 else 0 end) as "W",
sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" < g."awayGoals" then 1 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" < g."homeGoals" then 1 else 0 end) as "L",
sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" = g."awayGoals" then 1 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) as "D"
from games g join teams t on t."teamID" = g."homeTeamID" or t."teamID" = g."awayTeamID" join leagues l on g."leagueID" = l."leagueID"
group by t."name", g."season", l."name"
order by l."name", g."season", "Posicion") as "Stats"
group by "Stats"."Equipo";
```

Query No.23 Promedios de equipos

Explicación: se selecciona la tabla vista en anteriores queries la cuál tiene información de las posiciones, victorias, empates, goles, derrotas, etc, de todos los equipos. Sobre esa tabla se agrupa mediante el nombre del equipo y con la partición creada se pueden obtener los promedios de cada equipo.

Resultado:

	Equipo	123 WP	123 DP	123 GFP	123 GCP	123 GDP	123 LP	123 Puntos promedio	123 Poscion promedio
1	Nantes	12	9,8571428571	37,2857142857	44,7142857143	-7,4285714286	14,7142857143	45,8571428571	12,1428571429
2	Verona	9	11,2	40,6	61	-20,4	17,8	38,2	14,2
3	Bournemouth	11,2	8,6	48,2	66	-17,8	18,2	42,2	12,8
4	Empoli	9,5	11	41,5	58	-16,5	17,5	39,5	15
5	Wolverhampton V14	10,6666666667	10,6666666667	44,6666666667	46	-1,3333333333	13	53,6666666667	8,3333333333
6	Cordoba	3	11	22	68	-46	24	20	20
7	Palermo	9,3333333333	10	41,3333333333	65,6666666667	-24,3333333333	18,6666666667	38	15,3333333333
8	Crotone	8	6,6666666667	39,6666666667	72	-32,3333333333	23,3333333333	30,6666666667	18
9	Las Palmas	9	8	40,6666666667	67	-26,3333333333	21	35	14,6666666667
10	Metz	8,8	9,6	35	58,4	-23,4	17,6	36	15,4
11	Rennes	14	10,8571428571	45,4285714286	42,5714285714	2,8571428571	11,7142857143	52,8571428571	7
12	Leeds	18	5	62	54	8	15	59	9
13	Cesena	4	12	36	73	-37	22	24	20
14	Dijon	8,2	9,4	36,8	60,2	-23,4	18,4	34	16,2
15	Alaves	12,2	9,2	38	51,8	-13,8	16,6	45,8	12,6
16	Nancy	9	8	29	52	-23	21	35	19
17	Real Madrid	25,8571428571	6,7142857143	89,7142857143	36,5714285714	53,1428571429	5,4285714286	84,2857142857	2
18	Lyon	19,7142857143	8	70,8571428571	40,5714285714	30,2857142857	8,7142857143	67,1428571429	3,5714285714
19	Manchester City	26,1428571429	5,5714285714	88,5714285714	33,5714285714	55	6,2857142857	84	2
20	Troyes	6	7,5	30	71	-41	24,5	25,5	19,5
21	Levante	10,3333333333	10,5	44,5	61,8333333333	-17,3333333333	17,1666666667	41,5	14,8333333333
22	Hoffenheim	13	10,1428571429	56,1428571429	50,4285714286	5,7142857143	10,8571428571	49,1428571429	8
23	Augsburg	10	8,5714285714	42,1428571429	54,2857142857	-12,1428571429	15,4285714286	38,5714285714	12
24	SPAL 2013	8	9,3333333333	36,6666666667	64	-27,3333333333	20,6666666667	33,3333333333	16,6666666667

Tabla No.29 Promedios de equipos

20. ¿Cuántos puntos obtiene un equipo que queda en el top 5 en promedio y cuáles son los valores mínimos y máximos de puntos de un equipo que está en estas posiciones?

¿Cuántos puntos obtiene un equipo que queda en entre la posición 6 y 10?

¿Cuántos puntos obtienen en promedio los equipos que quedan en la posición 11 y 15?

¿Cuántos puntos obtienen en promedio los equipos que quedan en los últimos puestos?

```
--Promedio de puntos por posición
select avg("Puntos") from
(select t."name" as "Equipo",g."season" as "Temporada", l."name" as "Liga",
((sum(case when t."teamID" = g."homeTeamID" then g."homeGoals" else 0 end) +
sum(case when t."teamID" = g."awayTeamID" then g."awayGoals" else 0 end))/count(*) as "Goles",
sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" > g."awayGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" > g."homeGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end ) +
sum(case when t."teamID" = g."homeTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) as "Puntos",
rank() over (partition by l."name", g."season" order by
sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" > g."awayGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" > g."homeGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end ) +
sum(case when t."teamID" = g."homeTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end)
desc ) as "Posicion", count(*)
from games g join teams t on t."teamID" = g."homeTeamID" or t."teamID" = g."awayTeamID" join leagues l on g."leagueID" = l."leagueID"
group by t."name", g."season", l."name"
order by l."name", g."season", "Posicion") as "Primeras posiciones"
where "Posicion" between 1 and 5;
```

Query No.26 Puntos de los primeros 5 de la tabla en cada temporada

```
--Posiciones entre 6 y 10
select max("Puntos") as "Máximo", avg("Puntos") as "Promedio", min("Puntos") as "Minimo" from
(select t."name" as "Equipo",g."season" as "Temporada", l."name" as "Liga",
((sum(case when t."teamID" = g."homeTeamID" then g."homeGoals" else 0 end) +
sum(case when t."teamID" = g."awayTeamID" then g."awayGoals" else 0 end))/count(*) as "Goles",
sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" > g."awayGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" > g."homeGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) +
sum(case when t."teamID" = g."homeTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) as "Puntos",
rank() over (partition by l."name", g."season" order by
sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" > g."awayGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" > g."homeGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) +
sum(case when t."teamID" = g."homeTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end)
desc ) as "Posicion", count(*)
from games g join teams t on t."teamID" = g."homeTeamID" or t."teamID" = g."awayTeamID" join leagues l on g."leagueID" = l."leagueID"
group by t."name", g."season", l."name"
order by l."name", g."season", "Posicion") as "Posiciones medias"
where "Posicion" between 6 and 10;
```

Query No.27 Puntos de los equipos de media tabla

```
--Posiciones 11 y 15
select max("Puntos") as "Máximo", avg("Puntos") as "Promedio", min("Puntos") as "Minimo" from
(select t."name" as "Equipo",g."season" as "Temporada", l."name" as "Liga",
((sum(case when t."teamID" = g."homeTeamID" then g."homeGoals" else 0 end) +
sum(case when t."teamID" = g."awayTeamID" then g."awayGoals" else 0 end))/count(*) as "Goles",
sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" > g."awayGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" > g."homeGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) +
sum(case when t."teamID" = g."homeTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) as "Puntos",
rank() over (partition by l."name", g."season" order by
sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" > g."awayGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" > g."homeGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) +
sum(case when t."teamID" = g."homeTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end)
desc ) as "Posicion", count(*)
from games g join teams t on t."teamID" = g."homeTeamID" or t."teamID" = g."awayTeamID" join leagues l on g."leagueID" = l."leagueID"
group by t."name", g."season", l."name"
order by l."name", g."season", "Posicion") as "Primeras posiciones"
where "Posicion" between 11 and 15;
```

Query No.28.Puntos de los equipos entre las posiciones 10 y 15

```
--Posiciones entre más bajas
select max("Puntos") as "Máximo", avg("Puntos") as "Promedio", min("Puntos") as "Minimo" from
(select t."name" as "Equipo",g."season" as "Temporada", l."name" as "Liga",
((sum(case when t."teamID" = g."homeTeamID" then g."homeGoals" else 0 end) +
sum(case when t."teamID" = g."awayTeamID" then g."awayGoals" else 0 end))/count(*) as "Goles",
sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" > g."awayGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" > g."homeGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) +
sum(case when t."teamID" = g."homeTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) as "Puntos",
rank() over (partition by l."name", g."season" order by
sum(case when t."teamID" = g."homeTeamID" and g."homeGoals" > g."awayGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" > g."homeGoals" then 3 else 0 end) +
sum(case when t."teamID" = g."awayTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end) +
sum(case when t."teamID" = g."homeTeamID" and g."awayGoals" = g."homeGoals" then 1 else 0 end)
desc ) as "Posicion", count(*)
from games g join teams t on t."teamID" = g."homeTeamID" or t."teamID" = g."awayTeamID" join leagues l on g."leagueID" = l."leagueID"
group by t."name", g."season", l."name"
order by l."name", g."season", "Posicion") as "Primeras posiciones"
where "Posicion" > 15;
```

Query No. 29. Puntos de los equipos en las últimas posiciones

Explicación: se forma una tabla que muestra los puntos y goles de cada equipo por temporada siguiendo la estructura del Query No.13, de esta tabla se seleccionan los tuplas en donde los valores de posición están entre 1 y 5, finalmente se selecciona el promedio y los valores máximo y mínimo de puntos. Se hace el mismo procedimiento para obtener los valores para las posiciones entre 6 y 10, 11 y 15, 16 y 20.

Resultados:

	123 Máximo	123 Promedio	123 Minimo
1	100	73,2067039106	41

Tabla No.30 Puntos máximos y promedio de puntos de los equipos que acaban en las primeras posiciones

123 Máximo ▼	123 Promedio ▼	123 Minimo ▼
69	52,75	39

Tabla No.31 Puntos máximos y promedio de puntos de los equipos que acaban en posiciones de media tabla

123 Máximo ▼	123 Promedio ▼	123 Minimo ▼
55	42,3295454545	32

Tabla No.32 Puntos máximos y promedio de puntos de los equipos que acaban entre la posición 11 y 15

	123 Máximo ▼	123 Promedio ▼	123 Minimo ▼
1	44	31,9139072848	13

Tabla No.33 Puntos máximos y promedio de puntos de los equipos que acaban en las últimas posiciones

Discusión

Para seleccionar al mejor equipo para apostar, se debe tener en cuenta de que al equipo al que se le debe apostar sea un equipo consistente en su liga. En el query No.1 se puede observar que hay bastantes equipos que juegan una cantidad de partidos menor, se sabe que en la Ligue 1 esto pasó debido a que en la pandemia del año 2020 esta competición fue suspendida (Marca, 2020), sin embargo hay equipos que en algunas temporadas han llegado a jugar 0 partidos, esto puede deberse a que estos equipos descienden de categoría constantemente. A pesar de que los equipos más débiles suelen ser mejores para apostar, no es una buena idea apostar por equipos que regularmente descienden, pues los equipos que regularmente descienden son los que menos victorias tienen y menos puntos pueden hacer en una temporada regular. A pesar de eso es buena idea apostar por un equipo que, a pesar de haber descendido en temporadas anteriores, en las temporadas más actuales no haya tenido ningún descenso y tenga una cantidad de puntos decente .

Otro punto importante es el promedio goleador de los jugadores, en la tabla 3 se observan a los máximos goleadores de todas las ligas, es una buena estadística a tener en cuenta pues es importante que un equipo cuente con jugadores con la capacidad de meter una gran cantidad de goles, sin embargo esto a veces puede resultar engañoso ya que hay jugadores que participan en más partidos, por lo tanto tienen ventaja para poder anotar más goles, es por eso que el promedio goleador entre en juego ya que este nos indica la cantidad de goles por partido que un jugador puede llegar a hacer, además también puede mostrarnos a jugadores que son bastante prometedores, pues a pesar de no contar con la misma cantidad de partidos que los veteranos, estos pueden llegar ser muy importantes. Otro aspecto importante es tener en cuenta si la cantidad de goles anotados define la posición del equipo al final de la temporada, en la Figura No.1 se observa que a pesar de que a simple vista parece que la cantidad de goles anotados es importante y un factor determinante para la posición en la que acabará algún equipo en verdad no es tan importante, si bien en la mayor parte de casos anotar la mayor cantidad de goles ayuda a los equipos a conseguir mejores posiciones en sus competiciones, hay una gran cantidad de casos en los que la cantidad de goles no ha sido lo más importante para un equipo. En la figura No.2 también se observa que recibir menos goles tampoco es un factor muy importante a la hora de ganar o determinar la posición en la que terminará un equipo. En cambio, la diferencia de gol es uno de los factores más importantes a la hora de decidir si un equipo, el coeficiente de correlación que se obtuvo (Ver Figura No.5) fue bastante elevado y observando los resultados , esto nos dice que a pesar de que un equipo no anote muchos goles, si este no recibe muchos goles posiblemente el equipo acabó en las mejores posiciones, de igual manera con los equipos que reciben muchos goles y tengan la capacidad de anotar muchos. Es de gran interés que un equipo tenga bastantes puntos y acabe en las mejores posiciones ya que siempre será una apuesta segura y en muy pocas ocasiones se perderá dinero con este equipo. Sin embargo, también se quiere tener a un equipo que de vez en cuando pueda dar victorias sorpresivas y que apostar por él en partidos donde su contexto no es muy favorable pueda brindarnos una gran cantidad de dinero que ayude a mitigar las pérdidas que se sufren cuando no se hace una buena apuesta. Para determinar esto, se sabe cuáles son los equipos con mayor capacidad de sorpresa (Ver tablas 16 y 21). Un buen candidato para apostar es el Nice de la Ligue 1, es un equipo con una gran capacidad sorpresiva, tanto en casas de apuestas como en capacidad de ganar a pesar de que las probabilidades no estén a su favor, a su vez, este

equipo acaba normalmente en 6ta posición y acaba siempre con un promedio de 56 puntos cada temporada. Según el modelo planteado en la figura 2, se espera que el Niza obtenga 18.57 victorias por temporada si este queda regularmente en la 6ta posición, este dato no es muy favorable para el equipo, pues regularmente juega 38 partidos y en casi la mitad de estos se espera que el Niza empate o pierda. Otro punto a favor del Niza es que no es un equipo que pierda fácilmente cuando las cosas no están de su lado, pues no está entre los 25 equipos que más veces pierden cuando la probabilidad de su victoria es alta. Un equipo interesante puede ser el Burnley ya que según Bet365 es el equipo que más veces gana cuando las probabilidades en esta casa de apuesta no están a su favor, sin embargo el Burnley es un equipo que en promedio queda en la posición 13 y tiene una diferencia de gol negativa, lo que nos indica que el equipo es muy propenso a no ganar una gran cantidad de partidos.

Equipos que destacan en las listas de equipos que ganan en contextos desfavorables son el AC Milan y el Atlético de Madrid, estos equipos están entre los 15 mejores equipos en esta categoría, un punto en contra de estos equipos es que no son de los 25 mejores en ganar partidos desfavorables según bet365, el Milán es incluso un equipo que pierde cuando sus probabilidades en de ganar en Bet365 son altas. Sin embargo estos equipos están entre los 25 equipos con mejor diferencia de gol y hacen un promedio de puntos por temporada que los posicionan entre los 5 primeros o muy cerca de esta posición. El Atlético también destaca en tener a uno de los jugadores con mejores características de gol, Luis Suárez. Otro punto a favor del Atlético es que según Bet365, este equipo no pierde muchos partidos donde su contexto es favorable. Atlético de Madrid es el equipo con las características ideales para apostar, pues constantemente quede entra los primeros 5 de su liga, tiene talento diferencial en la delantera, poseen una gran defensa pues a pesar de no tener los números de gol de Barcelona y Real Madrid, este equipo es de los mejores en cuanto a diferencia de gol. Por lo tanto, el mejor equipo para apostar es el Atlético de Madrid, pues es un equipo que tiene una gran capacidad de generar sorpresas en contextos desfavorables, es competitivo en una liga en donde compite contra 2 de los 5 mejores equipos estadísticamente hablando y asegura una gran cantidad de victorias y sorpresas en una temporada.

	ABC Equipo	123 WP	123 DP	123 GFP	123 GCP	123 GDP	123 LP	123 Puntos promedio	123 Poscion promedio
1	AC Milan	18	10,1428571429	58,5714285714	43,2857142857	15,2857142857	9,8571428571	64,1428571429	6
2	Atletico Madrid	23,2857142857	9,4285714286	61,5714285714	25,2857142857	36,2857142857	5,2857142857	79,2857142857	2,4285714286
3	Burnley	11,3333333333	9,3333333333	37,3333333333	53,3333333333	-16	17,3333333333	43,3333333333	13,8333333333
4	Inter	20,7142857143	9	67,7142857143	38,4285714286	29,2857142857	8,2857142857	71,1428571429	4,1428571429
5	Newcastle United	11	9,3333333333	41,5	57,1666666667	-15,6666666667	17,6666666667	42,3333333333	13,5
6	Nice	15,5714285714	9,2857142857	48,4285714286	44	4,4285714286	11,7142857143	56	6,7142857143

Tabla No.36. Estadísticas de equipos de interés.

ABC Equipo	123 Partidos en donde el contexto era favorable seegún Bet365
Atletico Madrid	6

Tabla No.37. Partidos en donde el atlético

Conclusiones

- Los equipos que tienen una capacidad de ganar en contextos desfavorables y que además tienen una buena cantidad de victorias son los mejores para apostar, pues a pesar de que en gran parte de los partidos se pagará poco por su victoria, acumular pequeñas cantidades de dinero y de vez en cuando optar a grandes cantidades es mejor que estar constantemente perdiendo dinero.
- La diferencia de gol es uno de los factores más determinantes a la hora de decidir la posición final de un equipo
- Los equipos que compiten con equipos más competitivos tienen mejor oportunidad de ganar en contextos desfavorables
- Anotar goles no necesariamente asegura una buena posición en unas ligas, lo importante es tener una gran capacidad para tener mejores resultados defensivos y ofensivos a la vez

Anexos

1. Cálculos sobre el modelo de posición y número de puntos

Sea 6.13 la posición promedio del club Niza, entonces una ecuación para hallar su número de victorias esperado es:

$$6.13 = -0.851x + 21.935$$

$$\frac{-21.935 + 6.13}{-0.851} = x$$

$$x = 18.57$$

Referencias

1. Fichajes de fútbol, rumores, valores de mercado y noticias. (s/f). Transfermarkt.es. Recuperado el 26 de febrero de 2024, de <https://www.transfermarkt.es/>
2. Redacción. (2020, abril 30). Oficial: Francia declara campeón al PSG de la Ligue 1 2019-2020. *Marca*.
<https://www.marca.com/futbol/liga-francesa/2020/04/30/5eaabc3922601dc3168b45f8.html>