

# **INFRAESTRUCTURES PER AL BIG DATA (17244106, 17254118, 17924106) HW1: The LIGA database**

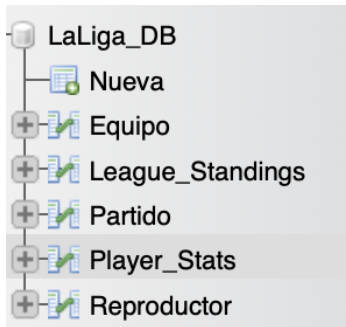
**Mauro Llorens Latorre  
Frances Carmona  
Vinyet Vendrell**



1. Create the Database

## Creation of the database and tables in MySQL

To begin, we created the database where we will store all the tables, called LaLiga\_DB, and selected it to work with it.



### TEAMS

This table stores information about the league teams.

The team\_id serves as the unique identifier for each team (it auto-increments).

#	Nombre	Tipo	Cotejamiento	Atributos	Nulo	Predeterminado	Comentarios	Extra
1	<b>team_id</b> 🔑	int			No	<i>Ninguna</i>		AUTO_INCREMENT
2	<b>team_name</b>	varchar(50)	utf8mb4_general_ci		Si	<i>NULL</i>		
3	<b>ciudad</b>	varchar(50)	utf8mb4_general_ci		Si	<i>NULL</i>		
4	<b>estadio</b>	varchar(50)	utf8mb4_general_ci		Si	<i>NULL</i>		
5	<b>coche</b>	varchar(50)	utf8mb4_general_ci		Si	<i>NULL</i>		

### MATCHES STATISTICS

This table stores statistics of the teams in the league.

#	Nombre	Tipo	Cotejamiento	Atributos	Nulo	Predeterminado
1	<b>team_id</b> 	int			No	<i>Ninguna</i>
2	<b>matches_played</b>	int			Sí	<i>NULL</i>
3	<b>WINS</b>	int			Sí	<i>NULL</i>
4	<b>Dibuja</b>	int			Sí	<i>NULL</i>
5	<b>Perdidas</b>	int			Sí	<i>NULL</i>
6	<b>Puntos</b>	int			Sí	<i>NULL</i>

## PLAYERS

This table stores information about the players in the league.

#	Nombre	Tipo	Cotejamiento	Atributos	Nulo	Predeterminado	Comentarios	Extra
1	<b>player_id</b> 	int			No	<i>Ninguna</i>		AUTO_INCREMENT
2	<b>first_name</b>	varchar(30)	utf8mb4_general_ci		Sí	<i>NULL</i>		
3	<b>last_name</b>	varchar(30)	utf8mb4_general_ci		Sí	<i>NULL</i>		
4	<b>posicion</b>	varchar(20)	utf8mb4_general_ci		Sí	<i>NULL</i>		
5	<b>birth_date</b>	date			Sí	<i>NULL</i>		
6	<b>nacionalidad</b>	varchar(30)	utf8mb4_general_ci		Sí	<i>NULL</i>		
7	<b>team_id</b> 	int			Sí	<i>NULL</i>		

## MATCHES

This table stores information about the matches played in the league.

#	Nombre	Tipo	Cotejamiento	Atributos	Nulo	Predeterminado	Comentarios	Extra
1	<b>match_id</b> 	int			No	<i>Ninguna</i>		AUTO_INCREMENT
2	<b>home_team</b> 	int			Sí	<i>NULL</i>		
3	<b>away_team</b> 	int			Sí	<i>NULL</i>		
4	<b>match_date</b>	date			Sí	<i>NULL</i>		
5	<b>estadio</b>	varchar(50)	utf8mb4_general_ci		Sí	<i>NULL</i>		
6	<b>home_goals</b>	int			Sí	<i>NULL</i>		
7	<b>away_goals</b>	int			Sí	<i>NULL</i>		

# PLAYERS STATISTICS

This table stores individual player statistics in matches.

#	Nombre	Tipo	Cotejamiento	Atributos	Nulo	Predeterminado	Comentarios	Extra
1	stat_id 🔑	int			No	Ninguna		AUTO_INCREMENT
2	player_id 🔑	int			Sí	NULL		
3	match_id 🔑	int			Sí	NULL		
4	Metas	int			Sí	NULL		
5	Ayuda	int			Sí	NULL		
6	yellow_cards	int			Sí	NULL		
7	red_cards	int			Sí	NULL		

## 2. Insert Sample Data

# TEAMS

team_id	team_name	ciudad	estadio	coche
1	FC Barcelona	Barcelona	Spotify Camp Nou	Xavi Hernández
2	Real Madrid	Madrid	Santiago Bernabéu	Carlo Ancelotti
3	Atlético de Madrid	Madrid	Cívitas Metropolitano	Diego Simeone
4	Real Sociedad	San Sebastián	Reale Arena	Imanol Alguacil
5	Sevilla FC	Sevilla	Ramón Sánchez-Pizjuán	Quique Sánchez Flores
6	Villarreal CF	Villarreal	Estadio de la Cerámica	Marcelino García Toral

# PLAYERS

player_id	first_name	last_name	posicion ▾ 1	birth_date	nacionalidad	team_id
4	Mariona	Caldentey	Portera	2005-10-07	França	5
9	Leila	Ouahabi	Portera	2012-07-19	Brasil	5
1	Aitana	Bonmatí	Migcampista	1998-11-24	Espanya	1
5	Irene	Guerrero	Migcampista	2013-02-17	Espanya	5
6	Olga	Carmona	Migcampista	2008-04-13	Espanya	1
14	Laia	Codina	Migcampista	1997-10-14	Espanya	5
15	Marta	Torrejón	Migcampista	2006-08-30	Itàlia	5
17	Ona	Batlle	Migcampista	2014-05-30	França	5
18	Patricia	Gavilán	Migcampista	2007-07-25	Itàlia	2
20	Nahikari	García	Migcampista	2010-08-12	Països Baixos	2
2	Alexia	Putellas	Defensa	2007-05-28	Brasil	3
10	Mapi	León	Defensa	2000-11-16	Brasil	5
11	Athenea	Del Castillo	Defensa	2015-12-27	Itàlia	1
3	Claudia	Pina	Davantera	2015-07-17	Itàlia	6
7	Lucía	García	Davantera	2007-10-14	Itàlia	1
8	Eva	Navarro	Davantera	2011-12-31	Països Baixos	1
12	Salma	Paralluelo	Davantera	2009-03-06	Brasil	3
13	Teresa	Abelleira	Davantera	2002-01-17	Països Baixos	4
16	Andrea	Pereira	Davantera	2006-10-17	Països Baixos	3
19	Esther	González	Davantera	2002-06-11	Brasil	4

## MATCHES

match_id	home_team	away_team	match_date	estadio	home_goals	away_goals
1	1	3	2024-08-08	Spotify Camp Nou	0	2
2	2	4	2024-08-15	Santiago Bernabéu	1	1
3	3	4	2024-08-22	Cívitas Metropolitano	2	1
4	6	2	2024-08-29	Estadio de la Cerámica	3	3
5	2	5	2024-09-05	Santiago Bernabéu	0	4
6	1	3	2024-09-12	Spotify Camp Nou	4	2
7	4	5	2024-09-19	Reale Arena	4	0
8	6	1	2024-09-26	Estadio de la Cerámica	3	4
9	3	4	2024-10-03	Cívitas Metropolitano	4	4
10	1	3	2024-10-10	Spotify Camp Nou	2	3

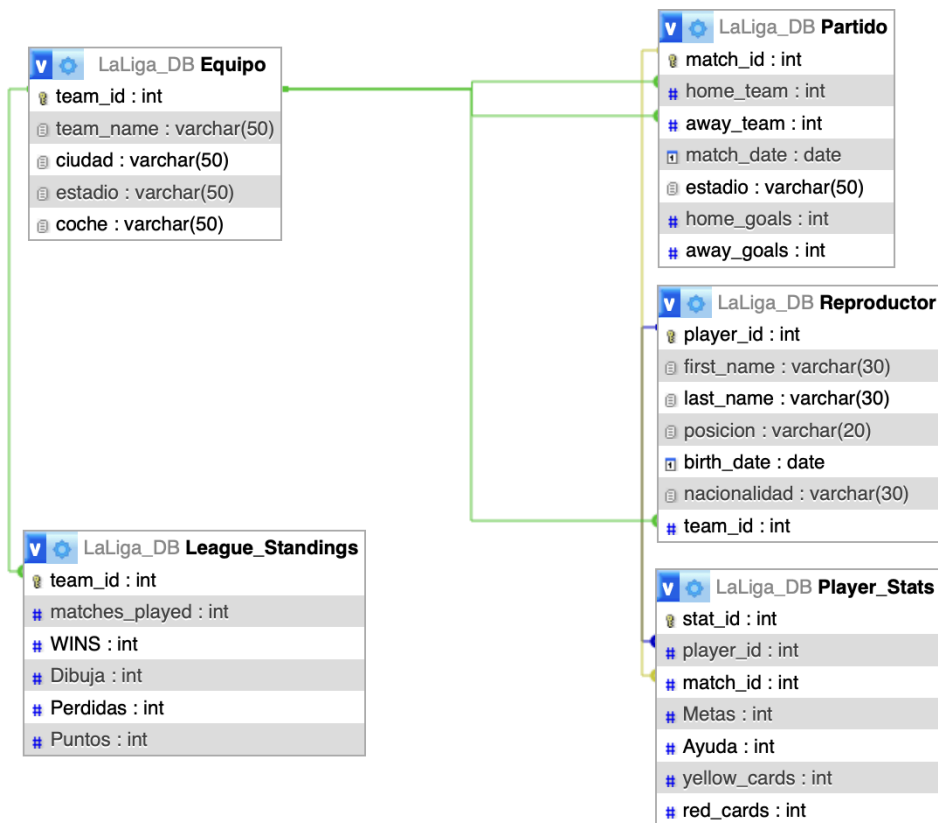
## PLAYER STATISTICS

stat_id	player_id	match_id	Metas	Ayuda	yellow_cards	red_cards
1	5	1	0	2	0	0
2	14	1	1	1	0	0
59	9	10	1	1	0	0
58	14	10	0	0	0	0
5	15	1	0	0	0	0
6	19	1	2	1	0	0
7	10	2	1	0	0	0
8	14	2	0	2	0	0
9	15	2	0	1	0	0
57	1	10	2	1	0	0
11	1	2	2	0	0	0
12	13	2	0	2	0	0
13	12	3	1	0	0	0
14	5	3	2	0	0	0
15	1	3	1	0	0	0
16	17	3	0	0	0	0
17	2	3	0	2	0	0
18	18	3	2	2	0	0
19	16	4	1	2	0	0
20	6	4	1	1	0	0
55	3	10	0	1	0	0
54	16	10	0	1	0	0
23	7	4	2	1	0	0
53	12	9	2	2	0	0
25	16	5	2	2	0	0

## MATCHES STATISTICS

team_id	matches_played	WINS	Dibuja	Perdidas	Puntos
1	10	3	2	5	11
2	10	4	4	2	16
3	10	4	3	3	15
4	10	3	4	3	13
5	10	3	4	3	13
6	10	5	3	2	18

## RELATIONSHIPS



## MATCHES

- Foreign Keys: home\_team and away\_team reference TEAMS.team\_id, ensuring valid team assignments.
- ON DELETE SET NULL: If a team is removed, its matches remain in the database but are marked with NULL in the corresponding home\_team or away\_team field, preserving match history.

Propiedades de la restricción	Columna	Restricción de clave foránea (INNODB)	
		Base de datos	Tabla
partido_ibfk_1	home_team	LaLiga_DB	Equipo
ON DELETE CASCADE	+ Añadir columna		
ON UPDATE RESTRICT			
partido_ibfk_2	away_team	LaLiga_DB	Equipo
ON DELETE CASCADE	+ Añadir columna		
ON UPDATE RESTRICT			

# PLAYER AND MATCHES STATS

- Foreign Key: team\_id references TEAMS, ensuring that each statistical entry belongs to a valid team.
- ON DELETE CASCADE: If a team is deleted from TEAMS, its related statistics in MATCHES\_STATISTICS will also be deleted to maintain data integrity.

Propiedades de la restricción	Columna ?	Restricción de clave foránea (INNODB)	
		Base de datos	Tabla
reproductor_ibfk_1	team_id	LaLiga_DB	Equipo
ON DELETE	CASCADE		
ON UPDATE	RESTRICT		

Propiedades de la restricción	Columna ?	Restricción de clave foránea (INNODB)	
		Base de datos	Tabla
reproductor_ibfk_1	team_id	LaLiga_DB	Equipo
ON DELETE	CASCADE		
ON UPDATE	RESTRICT		

# PLAYER STATS

- Foreign Keys:
  - player\_id references PLAYERS.player\_id, linking statistics to specific players.
  - match\_id references MATCHES.match\_id, linking statistics to specific matches.
- ON DELETE CASCADE: If a player is deleted, all their associated statistics are removed. Similarly, if a match is deleted, all statistics related to that match are also deleted, maintaining consistency.



Propiedades de la restricción	Columna	Restricción de clave foránea (INNODB)	
		Base de datos	Tabla
player_stats_ibfk_1	player_id	LaLiga_DB	Reproductor
ON DELETE: CASCADE	+ Añadir columna		
ON UPDATE: RESTRICT			
player_stats_ibfk_2	match_id	LaLiga_DB	Partido
ON DELETE: CASCADE	+ Añadir columna		
ON UPDATE: RESTRICT			

## 4. Data Analysis Using SQL Queries

a) Recupera a todos los jugadores que sean más jóvenes que la edad media de todos los jugadores de la liga.

```
SELECT * FROM Reproductor WHERE birth_date > ( SELECT
AVG(birth_date) FROM Reproductor );
```

player_id	first_name	last_name	posicion	birth_date	nacionalidad	team_id
3	Claudia	Pina	Davantera	2015-07-17	Itàlia	6
5	Irene	Guerrero	Migcampista	2013-02-17	Espanya	5
6	Olga	Carmona	Migcampista	2008-04-13	Espanya	1
8	Eva	Navarro	Davantera	2011-12-31	Països Baixos	1
9	Leila	Ouahabi	Portera	2012-07-19	Brasil	5
11	Athenea	Del Castillo	Defensa	2015-12-27	Itàlia	1
12	Salma	Paralluelo	Davantera	2009-03-06	Brasil	3
17	Ona	Batlle	Migcampista	2014-05-30	França	5
20	Nahikari	García	Migcampista	2010-08-12	Països Baixos	2

los elementos que están marcados:



Editar



Copiar



Borrar



Exportar

b) Muestra todos los partidos, incluidos los equipos locales y visitantes, las fechas de los partidos y los resultados.

```
SELECT p.match_id, e1.team_name AS home_team, e2.team_name
AS away_team, p.match_date, p.home_goals, p.away_goals FROM
Partido p JOIN Equipo e1 ON p.home_team = e1.team_id JOIN
Equipo e2 ON p.away_team = e2.team_id;
```

match_id	home_team	away_team	match_date	home_goals	away_goals
1	FC Barcelona	Atlético de Madrid	2024-08-08	0	2
2	Real Madrid	Real Sociedad	2024-08-15	1	1
3	Atlético de Madrid	Real Sociedad	2024-08-22	2	1
4	Villarreal CF	Real Madrid	2024-08-29	3	3
5	Real Madrid	Sevilla FC	2024-09-05	0	4
6	FC Barcelona	Atlético de Madrid	2024-09-12	4	2
7	Real Sociedad	Sevilla FC	2024-09-19	4	0
8	Villarreal CF	FC Barcelona	2024-09-26	3	4
9	Atlético de Madrid	Real Sociedad	2024-10-03	4	4
10	FC Barcelona	Atlético de Madrid	2024-10-10	2	3

**c) Encuentra los 3 mejores equipos con más puntos en la clasificación.**

```
SELECT e.team_name, l.Puntos FROM League_Standings l JOIN
Equipo e ON l.team_id = e.team_id ORDER BY l.Puntos DESC
LIMIT 3;
```

team_name	Puntos ▾ 1
Villarreal CF	18
Real Madrid	16
Atlético de Madrid	15

**d) Enumera a todos los jugadores junto con sus respectivos equipos.**

```
SELECT r.first_name, r.last_name, e.team_name FROM Reproductor r
JOIN Equipo e ON r.team_id = e.team_id;
```

first_name	last_name	team_name
Aitana	Bonmatí	FC Barcelona
Olga	Carmona	FC Barcelona
Lucía	García	FC Barcelona
Eva	Navarro	FC Barcelona
Athenea	Del Castillo	FC Barcelona
Patricia	Gavilán	Real Madrid
Nahikari	García	Real Madrid
Alexia	Putellas	Atlético de Madrid
Salma	Paralluelo	Atlético de Madrid
Andrea	Pereira	Atlético de Madrid
Teresa	Abelleira	Real Sociedad
Esther	González	Real Sociedad
Mariona	Caldentey	Sevilla FC
Irene	Guerrero	Sevilla FC
Leila	Ouahabi	Sevilla FC
Mapi	León	Sevilla FC
Laia	Codina	Sevilla FC
Marta	Torrejón	Sevilla FC
Ona	Batlle	Sevilla FC
Claudia	Pina	Villarreal CF

**e) Identifica a los jugadores que han marcado más de 3 goles en total.**

```
SELECT r.first_name, r.last_name, SUM(p.Metas) AS total_goals FROM
Player_Stats p JOIN Reproductor r ON p.player_id = r.player_id GROUP
BY r.player_id HAVING total_goals > 3;
```

first_name	last_name	total_goals
Aitana	Bonmatí	8
Alexia	Putellas	6
Olga	Carmona	5
Lucía	García	4
Mapi	León	4
Laia	Codina	4
Esther	González	4

**f) Recupera los detalles de los partidos en los que ganó el equipo local.**

```
SELECT match_id, match_date, team_name AS local_winner FROM
PartidoJOIN Equipo WHERE home_goals > away_goals;
```

match_id	match_date	local_winner
7	2024-09-19	FC Barcelona
6	2024-09-12	FC Barcelona
3	2024-08-22	FC Barcelona
7	2024-09-19	Real Madrid
6	2024-09-12	Real Madrid
3	2024-08-22	Real Madrid
7	2024-09-19	Atlético de Madrid
6	2024-09-12	Atlético de Madrid
3	2024-08-22	Atlético de Madrid
7	2024-09-19	Real Sociedad
6	2024-09-12	Real Sociedad
3	2024-08-22	Real Sociedad
7	2024-09-19	Sevilla FC
6	2024-09-12	Sevilla FC
3	2024-08-22	Sevilla FC
7	2024-09-19	Villarreal CF
6	2024-09-12	Villarreal CF
3	2024-08-22	Villarreal CF

**g) Calcula el promedio de goles marcados por partido para cada equipo.**

```
SELECT e.team_name, ROUND(AVG(CASE WHEN p.home_team =
e.team_id THEN p.home_goals ELSE p.away_goals END), 2) AS
avg_goals FROM Partido p JOIN Equipo e ON p.home_team =
e.team_id OR p.away_team = e.team_id GROUP BY e.team_id;
```

team_name	avg_goals
FC Barcelona	2.50
Real Madrid	1.33
Atlético de Madrid	2.60
Real Sociedad	2.50
Sevilla FC	2.00
Villarreal CF	3.00

**h) Identifica al jugador con más asistencias.**

```
SELECT r.first_name, r.last_name, SUM(p.Ayuda) AS
total_assists FROM Player_Stats p JOIN Reproductor r ON
p.player_id = r.player_id GROUP BY r.player_id ORDER BY
total_assists DESC LIMIT 1;
```

first_name	last_name	total_assists
Laia	Codina	6

**i) Encuentra el equipo con más victorias.**

```
SELECT e.team_name, I.WINS FROM League_Standings I JOIN
Equipo e ON I.team_id = e.team_id ORDER BY I.WINS DESC
LIMIT 1;
```

team_name	WINS
Villarreal CF	5

**j) Recupera todos los partidos jugados en un estadio específico (por ejemplo, 'Camp Nou').**

SELECT \* FROM Partido WHERE estadio = 'Spotify Camp Nou';

match_id	home_team	away_team	match_date
1	1	3	2024-08-08
6	1	3	2024-09-12
10	1	3	2024-10-10

k) Identifica al jugador que ha recibido más tarjetas amarillas sin recibir una tarjeta roja.

SELECT r.first\_name, r.last\_name, SUM(p.yellow\_cards) AS yellows  
FROM Player\_Stats p JOIN Reproductor r ON p.player\_id = r.player\_id  
GROUP BY r.player\_id HAVING SUM(p.red\_cards) = 0 ORDER BY  
yellows DESC LIMIT 1;

first_name	last_name	yellows
Aitana	Bonmatí	1

l) Busca partidos en los que ambos equipos hayan marcado el mismo número de goles (es decir, empates).

SELECT \* FROM Partido WHERE home\_goals = away\_goals;

match_id	home_team	away_team	match_date
2	2	4	2024-08-15
4	6	2	2024-08-29
9	3	4	2024-10-03

**m) Determina el jugador que ha jugado en más partidos.**

```
SELECT r.first_name, r.last_name, COUNT(p.match_id) AS
games FROM Player_Stats p JOIN Reproductor r ON
p.player_id = r.player_id GROUP BY r.player_id ORDER BY
games DESC LIMIT 1;
```

first_name	last_name	games
Aitana	Bonmatí	5

**n) Identifica al equipo que ha marcado más goles solo en partidos fuera de casa.**

```
SELECT e.team_name, SUM(p.away_goals) AS away_goals FROM
Partido p JOIN Equipo e ON p.away_team = e.team_id GROUP BY
e.team_id ORDER BY away_goals DESC LIMIT 1;
```

team_name	away_goals
Atlético de Madrid	7

**o) Recupere los detalles de los jugadores que han jugado para más de un equipo en el conjunto de datos.**

```
SELECT r.first_name, r.last_name, COUNT(DISTINCT r.team_id) AS
equipos FROM Reproductor r GROUP BY r.player_id HAVING equipos
> 1;
```