



ESTADISTICAS DEL FUTBOL MUNDIAL

Autor: Daniel Perco

Profesor: Camilo Andrés Redondo

Tutor: Emilio D Augero

Comision: 47365

Estadísticas de las Principales Ligas de Futbol Mundial 2023

Índice

- 1 -Introducción,
- 2 -Objetivo,
- 3 -Situación problemática,
- 4- Modelo de negocio,
- 5 -Diagrama de entidad relación
- 6 -Listado de tablas con descripción de estructura
- 7 -Scripts de creación de cada objeto de la base de datos,
- 8 - Scripts de inserción de datos,
- 9- Herramientas y tecnologías usadas,
- 10- Conclusiones
- 11– Reporte con PowerBI

1- Introducción:

Este proyecto intenta mostrar datos estadísticos, de las principales Ligas de futbol del mundo, en particular las ligas de los países: Argentina, Brasil, España, Italia, Inglaterra.

Los datos que intentaremos mostrar son los más básicos y generales, que suelen ver la mayoría de los aficionados a este deporte como posición de sus equipos, cantidad de puntos, partidos ganados, perdidos, empatados, cantidad de goles, tabla de goleadores, de asistidores, etc.

2- Objetivo:

El principal objetivo es procesar la gran diversidad y cantidad de datos, que se obtienen de la página: <https://footystats.org/es/download-stats-csv> , simplificar toda esta información y mostrarla de una manera que

cualquier aficionado lo pueda interpretar, sin grandes conocimientos técnicos.

3- Situación Problemática:

La problemática de este proyecto se presenta en la forma que se reciben los datos, son todas planillas .csv, con muchas columnas, con muchos errores de caracteres alfanuméricos, en los datos que tienen cargadas las diferentes celdas. Estos errores, se pueden detectar, además de verlos como una planilla en Excel, también se puede utilizar un editor de código o texto, en mi caso utilizo Vscod, para detectar y corregir los distintos caracteres erróneos.

Además, se deben adaptar los distintos tipos de datos recibidos, sin numéricos, float, fechas, etc y adaptarlos para el código SQL.

También, debemos adaptar e interpretar las distintas tablas, para que los usuarios puedan obtener información simple y clara, para esto selecciones y reducimos considerablemente, la cantidad de información que recibimos, para generar una base de datos más acorde, a lo que queremos mostrar.

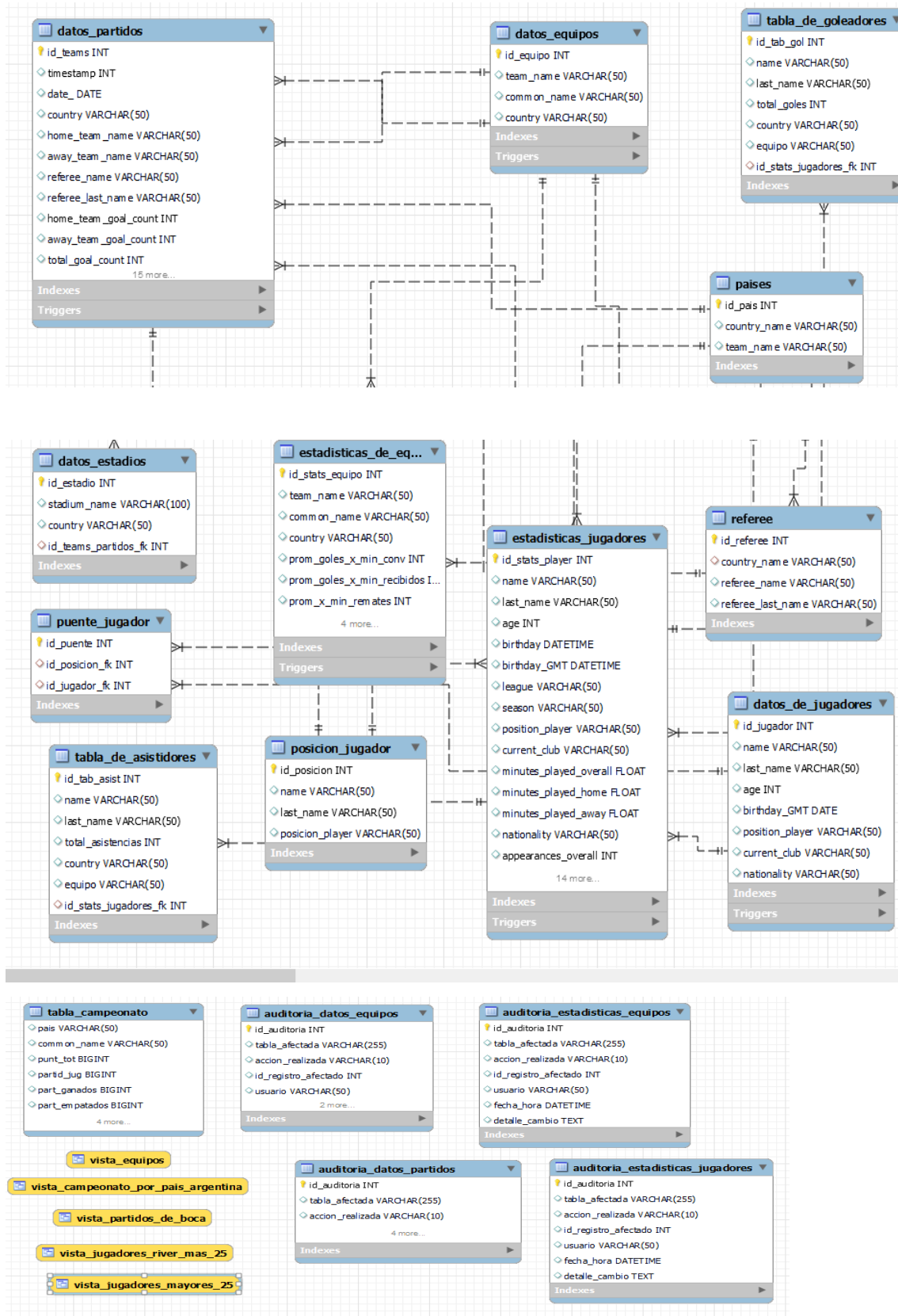
Hay ligas, donde faltan algunos datos , que no se brindan en los archivos.

4- Modelo de Negocio:

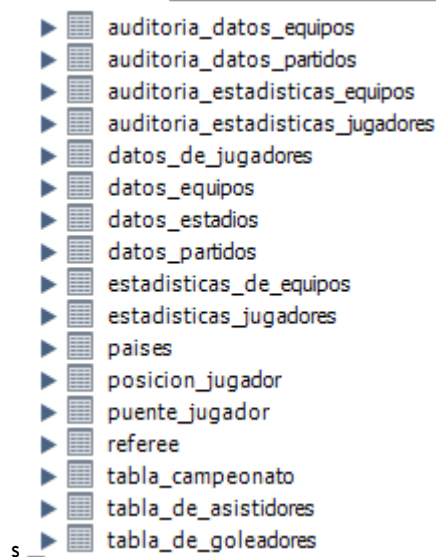
Este proyecto más allá de estar destinado principalmente ,para que aficionados al futbol ,puedan tener datos de sus equipos y de sus ligas preferidas , también se puede utilizar profesionalmente ,por ejemplo generando informes y métricas ,para los equipos técnicos de los distintos equipos, selecciones o profesionales del deporte ,ya que podemos generar métricas relacionadas con el desempeño de los jugadores , ya sea físicamente o en el desarrollo del juego , y también el desempeño de los equipos ,porque contamos con docenas de parámetros : metros recorridos por el jugador , cantidad de pases que dio, cuantos erro , cuantos minutos estuvo activo ,cuantos inmóvil , etc.

Simplemente, se deberá filtrar los datos que se requieran, para cada caso, en las planillas que se reciben de la página.

5- DIAGRAMA ENTIDAD-RELACION



6-Tablas - descripción: estas tablas están construidas en relación a los datos que nos provee la pagina de donde extraemos la información.

A screenshot of a database interface showing a list of tables. The tables are listed in a column, each preceded by a small icon of a document with a grid. The tables are: auditoria_datos_equipos, auditoria_datos_partidos, auditoria_estadisticas_equipos, auditoria_estadisticas_jugadores, datos_de_jugadores, datos_equipos, datos_estadios, datos_partidos, estadisticas_de_equipos, estadisticas_jugadores, paises, posicion_jugador, puente_jugador, referee, tabla_campeonato, tabla_de_asistidores, and tabla_de_goleadores. The list is partially cut off on the left by the text 's _'.

auditoria_datos_equipos: es la tabla, que nos permite supervisar todas las transacciones que se realicen sobre la tabla datos equipos

como la carga de datos, modificación, etc, etc, tiene diferentes campos: un id para identificar la operación, uno que indica a que tabla se está auditando,

otro para indicar el usuario que realizo la operación, un campo que indica que tipo de operación, fecha, en síntesis, todos campos que me permitan

realizar una auditoría.

auditoria_datos_partidos: es la tabla, que nos permite supervisar todas las transacciones que se realicen sobre la tabla datos_partidos

como la carga de datos, modificación, etc, etc. tiene diferentes campos: un id para identificar la operación, uno que indica a que tabla se está auditando,

otro para indicar el usuario que realizo la operación, un campo que indica que tipo de operación, fecha, en síntesis, todos campos que me permitan

realizar una auditoría.

auditoria_estadisticas_equipos: es la tabla, que nos permite supervisar todas las transacciones que se realicen sobre la tabla estadísticas_equipos

como la carga de datos, modificación, etc, etc tiene diferentes campos: un id para identificar la operación, uno que indica a que tabla se está auditando,

otro para indicar el usuario que realizo la operación, un campo que indica que tipo de operación, fecha, en síntesis, todos campos que me permitan

realizar una auditoría.

auditoria_estadisticas_jugadores: es la tabla, que nos permite supervisar todas las transacciones que se realicen sobre la tabla estadisticas_jugadores

como la carga de datos, modificación, etc, etc tiene diferentes campos: un id para identificar la operación, uno que indica a que tabla se está auditando,

otro para indicar el usuario que realizo la operación, un campo que indica que tipo de operación, fecha, en síntesis, todos campos que me permitan

realizar una auditoría.

datos_de_jugadores: en esta tabla tenemos todos los datos relacionados a los jugadores, nombre, apellido, edad, fecha nacimiento, club donde juega, su nacionalidad

su posición en el campo de juego.

datos Equipos: esta tabla tiene como datos el nombre oficial del club y el nombre con el que se lo conoce tradicionalmente entre los aficionados, además contiene el país al que pertenece.

datos_partidos: en esta tenemos los datos relacionados a los partidos jugados en el primer semestre del 2023, figuran datos como los nombres de los equipos, el local y visitante, la fecha del

encuentro, el árbitro que dirigió el partido (hay varios partidos donde no se provee el nombre del árbitro, aparece como "no mame", en esos casos), el estadio donde se jugó,

el país donde se jugó, y algunas estadísticas como cantidad de goles convertidos y recibidos, de tarjetas amarillas y rojas, cantidad de remates al arco, cantidad de fules,

tiempo de posesión del balón.

datos_estadios: esta tabla tiene los nombres de los estadios y a que país pertenecen.

estadísticas_de Equipos: aquí tenemos las estadísticas que queremos mostrar para este proyecto, la fuente tiene mas de 50 o 60 indicadores, nosotros tomamos algunos, los que nos parecen mas

indicativos. La tabla cuenta con el nombre oficial del equipo, el nombre común con el que se lo reconoce, el país al que pertenece, promedio de goles convertidos y recibidos, cantidad de

fules, cantidad de remates al arco, cantidad de offside.

estadísticas_jugadores: esta tabla contiene también algunos de los datos estadísticos de los jugadores, los que seleccionamos de más de 100 indicadores que nos brinda la fuente.

Para poder identificar a cada jugador, tenemos su nombre, apellido, nacionalidad, fecha de nacimiento, equipo donde juega y en que posición juega.

En cuanto a los datos estadísticos que seleccionamos: minutos jugadores por partido de local y de visitante, remates realizados, goles, tarjetas amarillas y rojas recibidas,

Asistencias para goles realizadas. penales convertidos y errados, y pases concretados

países: esta contiene la información de a qué país pertenece cada equipo.

posición jugadora: esta tabla cuenta con la información de la posición en la que juega cada jugador, con su nombre apellido y la posición de juego.

punteo jugador: esta tabla puente, entre la tabla datos jugadores y posiciones jugado
, los que nos permite optimizar las relaciones entre muchos registros entre distintas tablas

referee: esta tiene los nombre y apellido de los referee que dirigieron partidos y su nacionalidad.

tabla campeonato: esta tabla contiene los datos de las posiciones de cada campeonato, se crea posterior a la carga

de los datos de las tablas anteriores, ya que las columnas toman datos de las otras tablas

y varias columnas son resultante de varias operaciones matemáticas que se realiza, como cantidad de partidos ganados, diferencia de gol, cantidad de puntos totales.

además, contiene el nombre del club, país donde pertenece, cantidad de partidos jugados , cantidad de goles realizados, cantidad recibidos, la posición que ocupa en la tabla.

tabla asistentes: esta tabla contiene la cantidad de asistencias que realizo el jugador, que terminan en gol. Tiene

nombre y apellido del jugador, nacionalidad, equipo donde juega y el total de asistencias.

tabla_goleadores: esta contiene la cantidad de goles que convirtió cada jugador, contiene nombre y apellido, nacionalidad, equipo donde juega, y total de goles.

- 7- Objetos de la Base de DATOS

-Views:

Hay 5 vistas:

vista_campeonato_por_pais_argentina: se obtiene la vista de la tabla de posiciones del campeonato argentino

vista_equipos: se obtiene la vista de todos los equipos que figuran en la base de datos

vista_jugadores_mayores_25: se obtiene la vista de todos los jugadores mayores a 25 años

vista_jugadores_river_mas_25: se obtiene la vista de todos los jugadores de River mayores a 25 años

vista_partidos_de_boca: se obtienen todos donde jugo el equipo de Boca Juniors

-Stored-Procedure:

Hay 3:

BuscarJugadoresPorCriterio: este procedure tienen una consulta dinámica, que devuelve el listado de jugadores de un club determinado, según la columna que pasemos como parámetro: puede ser edad, equipo, nacionalidad, etc

EquiposCon PromedioMinutosPorGol: nos devuelve el listado de promedio de edad de los jugadores de cada equipo de acuerdo al promedio de minutos de gol que se ingresa

ObtenerJugadoresPorEquipo: devuelve el listado de jugadores de un equipo

-Functions:

Hay 2:

BuscarJugadoresPorNombre: busca jugadores por un nombre o apellido

contarAsistenciasPorEquipo: nos muestra la cantidad Total de asistencias de gol de un equipo

-Triggers

Hay 7:

before_insert_jugador: registra que usuario está iniciando la carga de un nuevo jugador

Before_insert_equipo: registra que usuario está empezando a cargar datos de un nuevo equipo

tr_datos_equipos_after_insert: registra quien realizo modificaciones en los datos de un jugador

After_update_equipo: registra quien realizo actualizaciones de datos de equipos

tr_datos_partidos_after_insert: registra que usuario está cargando nuevos datos de partidos

tr_estadisticas_equipos_after_update: registra que usuario realizo modificaciones en las estadísticas de un equipo

tr_estadisticas_jugadores_after_update: registra que usuario realizo cambios de datos en las estadísticas de jugadores

-Data Control Lenguaje: en este script se crea un usuario con permisos de administrador y otro con permisos de solo lectura.

-Transacciones Control Lenguaje: se crean una transaccione, como crear y borrar datos en una de las tablas

8 - Scripts de inserción de datos

Para la inserción de datos, debemos ir ejecutando los archivos .SQL, que se encuentran dentro de estas carpetas, respetando el orden

- 1_Creacion_base_datos_y_tablas
- 2_carga_datos
- 3_tabla_campeonato
- 4_forend_key
- 5_views
- 6_stored_procedure
- 7_functions
- 8_Trigger
- 9_Document_Control_lenguaje
- 10_transaction_Control_lenguaje
- 11_backup

9 - Herramientas y tecnologías usadas: las herramientas utilizadas Workbench, SQL, Excel, Word, Vscod, PowerBI

10- Conclusiones:

Podemos concluir, que, dada la cantidad de información disponible, en la plataforma, de donde obtenemos los datos, al ser una plataforma que maneja distintos deportes, podemos generar

Para distintos deportes, una variada cantidad de proyectos ya sea de forma informativa o profesional.

Se podrá realizar una simple base de datos, para generar dashboard o paginas para los usuarios en general, como informes para profesionales, donde se pueda tomar decisiones, por ejemplo: características

de un jugador, el cual podrá ser adquirido por algún equipo.

11 – Reporte con PowerBI

Se genera utilizando PowerBI, algunos reportes utilizando esta base de datos:

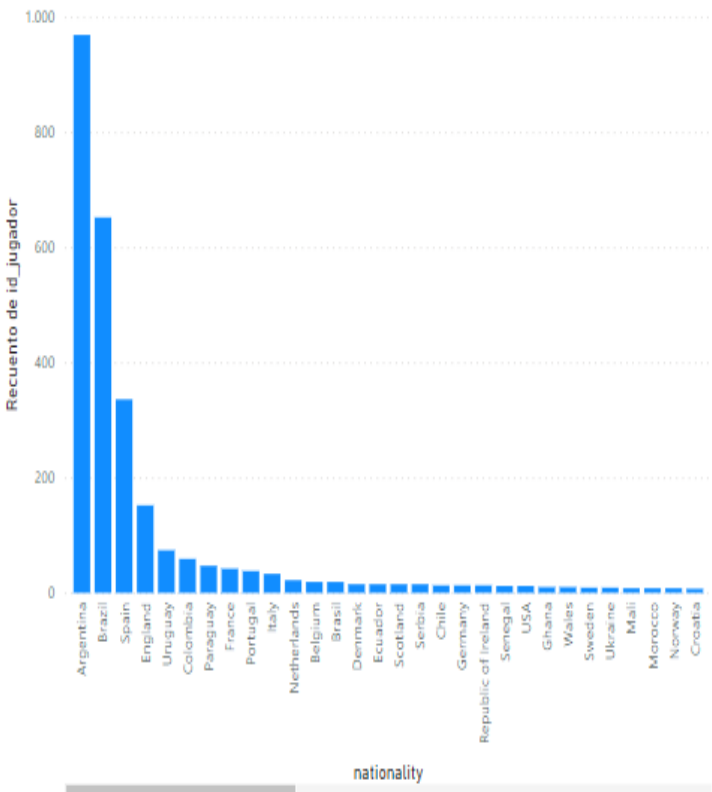
El generador de reportes del proyecto se encuentra en este enlace, hay que descargarlo y ejecutarlo:

https://github.com/dperco/estadisticas_futbol_2023/blob/main/Documentacion/informes_estad_futbol_2023.pbix

Reporte cantidad de jugadores por nacionalidad

Recuento de id_jugador

por nationality



nationality	Jugadores por Nacionalidad
Afghanistan	1
Albania	3
Algeria	4
Argentina	969
Australia	1
Austria	2
Belgium	19
Bolivia	2
Brasil	19
Brazil	652
Bulgaria	1
Burkina Faso	2
Cameroon	4
Canada	1
Cape Verde	1
Chile	13
Colombia	59
Congo DR	2
Costa Rica	1
Croatia	7
Cyprus	1
Czech Republic	3
Denmark	15
Ecuador	15
Total	2786



Reportes Tablas de posiciones por ligas

pais

Argentina

ENGLAND

Italy

common_name	Suma de punt_tot	Suma de partid_jug	Suma de part_ganados	Suma de part_empatados	Suma de part_perdidos	Suma de goles_a_f
River Plate	105	62	27	24	11	
Talleres Cordoba	100	61	24	28	9	
San Lorenzo	88	61	18	34	9	
Defensa y Justicia	83	61	17	32	12	
Estudiantes	80	61	17	29	15	
Boca Juniors	79	59	19	22	18	
Lanus	79	62	15	34	13	
Argentinos Juniors	77	62	16	29	17	
Godoy Cruz	77	62	15	32	15	
Racing Club	76	61	15	31	15	
Atletico Tucuman	73	62	12	37	13	
Belgrano	73	62	15	28	19	
Rosario Central	72	60	12	36	12	
Tigre	72	62	15	27	20	
Newell's Old Boys	71	62	12	35	15	
Barracas Central	68	61	12	32	17	
Instituto	68	62	13	29	20	
Sarmiento	68	62	12	32	18	
Union Santa Fe	68	62	10	38	14	
Ranfield	67	63	12	31	20	
Total	2013	1722	385	858	479	

Informe Estadísticas Furbol 2023

Cantidad Jugadores x nacionalidad

Tablas de posiciones

+

pais

Argentina

ENGLAND

Italy

common_name	Suma de punt_tot	Suma de partid_jug	Suma de part_ganados	Suma de part_empatados	Suma de part_perdidos	Suma de goles_a_f
AC Milan	36	26	5	21	0	
Juventus	34	26	4	22	0	
Inter Milan	33	25	4	21	0	
Napoli	31	26	3	22	1	
Atalanta	30	26	3	21	2	
Hellas Verona	30	26	3	21	2	
Lecce	29	25	2	23	0	
Florentina	28	26	2	22	2	
Genoa	27	26	2	21	3	
Lazio	27	26	2	21	3	
Bologna	26	25	1	23	1	
Frosinone	26	25	1	23	1	
Torino	25	25	1	22	2	
Udinese	25	26	0	25	1	
Sassuolo	24	25	1	21	3	
Cagliari	23	26	0	23	3	
Monza	23	25	1	20	4	
Salernitana	23	25	0	23	2	
Roma	22	25	0	22	3	
Empoli	21	25	0	21	4	
Total	543	510	35	438	37	

Informe Estadísticas Furbol 2023

Cantidad Jugadores x nacionalidad

Tablas de posiciones

+

pais

☐ Argentina

☐ ENGLAND

☒ España

☐ Italy

common_name	Suma de punt_tot	Suma de partid_jug	Suma de part_ganados	Suma de part_empatados	Suma de part_perdi
Real Madrid	12	4	4	0	
FC Barcelona	10	4	3	1	
Girona FC	10	4	3	1	
Athletic Club Bilbao	7	4	2	1	
Cadiz	7	4	2	1	
Real Betis	7	4	2	1	
CA Osasuna	6	4	2	0	
Deportivo Alaves	6	4	2	0	
Rayo Vallecano	6	4	2	0	
Real Sociedad	6	4	1	3	
Valencia CF	6	4	2	0	
Atetico Madrid	4	2	1	1	
Atletico Madrid	4	2	1	1	
Celta de Vigo	4	4	1	1	
Getafe CF	4	4	1	1	
Granada CF	3	4	1	0	
Villarreal	3	4	1	0	
RCD Mallorca	2	4	0	2	
UD Las Palmas	2	4	0	2	
Almeria	1	4	0	1	
Sevilla FC	1	4	0	1	
Total	111	80	31	18	



Informe Estadísticas Furbol 2023

Cantidad Jugadores x nacionalidad

Tablas de posiciones



