# Resumen de los datos recopilados en los excel

Estos datos se encuentran en la carpeta “pruebas”>”csv”

Se separan los csv por liga y temporada. Por ejemplo los csv con los datos asociados a la liga de España en la temporada 2024, terminan por “…-primera-2024.csv”: Para esta liga y temporada se pueden encontrar por ejemplo “jugadores-primera-2024.csv”, “partidos-primera-2024.csv”… Los datos se agrupan en las siguientes entidades.

## Equipos

Se recogen los datos de los equipos en una temporada y liga en cada csv con la siguiente estructura:

A screenshot of a phone

Description automatically generated

## Partidos

Se recogen los datos de los partidos en una temporada y liga en cada csv con la siguiente estructura:

A screenshot of a table

Description automatically generated

## Jugadores

Se recogen los datos de los jugadores en una temporada y liga en cada csv con la siguiente estructura:

A screenshot of a computer

Description automatically generated

## DatosJugadoresPartidos

Se recogen los datos de los jugadores en cada partido en una temporada y liga en cada csv con la siguiente estructura:

A screenshot of a computer

Description automatically generated

Estos datos al ser más complejos se explican a continuación lo que se puede ver en la primera fila:

* Para el jugador con id 1773, en el partido 760, salió como titular, no recibió amarillas ni rojas, no se lesionó, no fue cambiado, no dió asistencias ni marcó goles y como no fue sustituido no hay un jugador que entrase por él.

## DatosPartidosJugados

Se recogen los datos de los partidos en una temporada y liga en cada csv con la siguiente estructura:

A screenshot of a computer

Description automatically generated

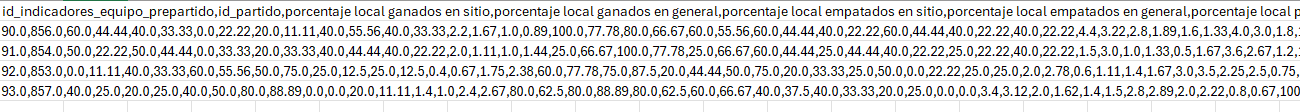
Las columnas son: id\_datos\_partido\_jugado,id\_partido,resultado\_local,resultado\_visitante,cambios\_local,cambios\_visitante,amarillas\_local,amarillas\_visitante,rojas\_local,rojas\_visitante,corners\_local,corners\_visitante,posesion\_local,posesion\_visitante,total\_tiros\_local,total\_tiros\_visitante

Estos datos al ser más complejos se explican a continuación lo que se puede ver en la primera fila:

* Para el partido 760, el local metió 0 goles, el visitante 2, el local hizo 5 cambios, el visitante hizo 5 cambios, el local recibió 1 amarilla, el visitante 3 amarillas…

## IndicadoresEquipoHistorico

Se recogen los datos de los indicadores de los partidos en una temporada y liga en cada csv con la siguiente estructura:



Las columnas son:

id\_indicadores\_equipo\_prepartido,id\_partido,porcentaje local ganados en sitio,porcentaje local ganados en general,porcentaje local empatados en sitio,porcentaje local empatados en general,porcentaje local perdidos en sitio,porcentaje local perdidos en general,porcentaje visitante ganados en sitio,porcentaje visitante ganados en general,porcentaje visitante empatados en sitio,porcentaje visitante empatados en general,porcentaje visitante perdidos en sitio,porcentaje visitante perdidos en general,proporcion local puntos en sitio,proporcion local puntos en general,proporcion visitante puntos en sitio,proporcion visitante puntos en general,"porcentaje local mas 1,5 en sitio","porcentaje local mas 1,5 en general","porcentaje visitante mas 1,5 en sitio","porcentaje visitante mas 1,5 en general","porcentaje local mas 2,5 en sitio","porcentaje local mas 2,5 en general","porcentaje visitante mas 2,5 en sitio","porcentaje visitante mas 2,5 en general","porcentaje local mas 3,5 en sitio","porcentaje local mas 3,5 en general","porcentaje visitante mas 3,5 en sitio","porcentaje visitante mas 3,5 en general","porcentaje local mas 4,5 en sitio","porcentaje local mas 4,5 en general","porcentaje visitante mas 4,5 en sitio","porcentaje visitante mas 4,5 en general",proporcion local goles totales en sitio,proporcion local goles totales en general,proporcion local goles marcados en sitio,proporcion local goles marcados en general,proporcion local goles encajados en sitio,proporcion local goles encajados en general,proporcion visitante goles totales en sitio,proporcion visitante goles totales en general,proporcion visitante goles marcados en sitio,proporcion visitante goles marcados en general,proporcion visitante goles encajados en sitio,proporcion visitante goles encajados en general,"porcentaje local mas 0,5 marcados en sitio","porcentaje local mas 0,5 marcados en general","porcentaje local mas 1,5 marcados en sitio","porcentaje local mas 1,5 marcados en general","porcentaje local mas 2,5 marcados en sitio","porcentaje local mas 2,5 marcados en general","porcentaje local mas 0,5 encajados en sitio","porcentaje local mas 0,5 encajados en general","porcentaje local mas 1,5 encajados en sitio","porcentaje local mas 1,5 encajados en general","porcentaje local mas 2,5 encajados en sitio","porcentaje local mas 2,5 encajados en general","porcentaje visitante mas 0,5 marcados en sitio","porcentaje visitante mas 0,5 marcados en general","porcentaje visitante mas 1,5 marcados en sitio","porcentaje visitante mas 1,5 marcados en general","porcentaje visitante mas 2,5 marcados en sitio","porcentaje visitante mas 2,5 marcados en general","porcentaje visitante mas 0,5 encajados en sitio","porcentaje visitante mas 0,5 encajados en general","porcentaje visitante mas 1,5 encajados en sitio","porcentaje visitante mas 1,5 encajados en general","porcentaje visitante mas 2,5 encajados en sitio","porcentaje visitante mas 2,5 encajados en general",proporcion local amarillas en sitio,proporcion local amarillas en general,proporcion visitante amarillas en sitio,proporcion visitante amarillas en general,proporcion local rojas en sitio,proporcion local rojas en general,proporcion visitante rojas en sitio,proporcion visitante rojas en general,proporcion local cambios en sitio,proporcion local cambios en general,proporcion visitante cambios en sitio,proporcion visitante cambios en general,proporcion local posesion en sitio,proporcion local posesion en general,proporcion visitante posesion en sitio,proporcion visitante posesion en general,proporcion local total tiros en sitio,proporcion local total tiros en general,proporcion visitante total tiros en sitio,proporcion visitante total tiros en general,proporcion local corners a favor en sitio,proporcion local corners a favor en general,proporcion visitante corners a favor en sitio,proporcion visitante corners a favor en general,proporcion local corners en contra en sitio,proporcion local corners en contra en general,proporcion visitante corners en contra en sitio,proporcion visitante corners en contra en general,proporcion local cambios lesionados sitio,proporcion local cambios lesionados en general,proporcion visitante cambios lesionados en sitio,proporcion visitante cambios lesionados en general,proporcion local cambios amarillas sitio,proporcion local cambios amarillas en general,proporcion visitante cambios amarillas en sitio,proporcion visitante cambios amarillas en general,proporcion local cambios goleadores sitio,proporcion local cambios goleadores en general,proporcion visitante cambios goleadores en sitio,proporcion visitante cambios goleadores en general,proporcion local cambios asistentes sitio,proporcion local cambios asistentes en general,proporcion visitante cambios asistentes en sitio,proporcion visitante cambios asistentes en general,media local cambios minutos sitio,media local cambios minutos en general,media visitante cambios minutos sitio,media visitante cambios minutos en general,proporcion local cambios delanteros a centrocampistas sitio,proporcion local cambios delanteros a centrocampistas en general,proporcion visitante cambios delanteros a centrocampistas en sitio,proporcion visitante cambios delanteros a centrocampistas en general,proporcion local cambios delanteros a defensas sitio,proporcion local cambios delanteros a defensas en general,proporcion visitante cambios delanteros a defensas en sitio,proporcion visitante cambios delanteros a defensas en general,proporcion local cambios centrocampistas a delanteros sitio,proporcion local cambios centrocampistas a delanteros en general,proporcion visitante cambios centrocampistas a delanteros en sitio,proporcion visitante cambios centrocampistas a delanteros en general,proporcion local cambios centrocampistas a defensas sitio,proporcion local cambios centrocampistas a defensas en general,proporcion visitante cambios centrocampistas a defensas en sitio,proporcion visitante cambios centrocampistas a defensas en general,proporcion local cambios defensas a delanteros sitio,proporcion local cambios defensas a delanteros en general,proporcion visitante cambios defensas a delanteros en sitio,proporcion visitante cambios defensas a delanteros en general,proporcion local cambios defensas a centrocampistas sitio,proporcion local cambios defensas a centrocampistas en general,proporcion visitante cambios defensas a centrocampistas en sitio,proporcion visitante cambios defensas a centrocampistas en general,proporcion local cambios antes descanso sitio,proporcion local cambios antes descanso en general,proporcion visitante cambios antes descanso en sitio,proporcion visitante cambios antes descanso en general,proporcion local cambios 45 a 60 sitio,proporcion local cambios 45 a 60 en general,proporcion visitante cambios 45 a 60 en sitio,proporcion visitante cambios 45 a 60 en general,proporcion local cambios 61 a 75 sitio,proporcion local cambios 61 a 75 en general,proporcion visitante cambios 61 a 75 en sitio,proporcion visitante cambios 61 a 75 en general,proporcion local cambios 76 a final sitio,proporcion local cambios 76 a final en general,proporcion visitante cambios 76 a final en sitio,proporcion visitante cambios 76 a final en general,proporcion local cambios alineacion defensa sitio,proporcion local cambios alineacion defensa en general,proporcion visitante cambios alineacion defensa en sitio,proporcion visitante cambios alineacion defensa en general,proporcion local cambios alineacion centrocampista sitio,proporcion local cambios alineacion centrocampista en general,proporcion visitante cambios alineacion centrocampista en sitio,proporcion visitante cambios alineacion centrocampista en general,proporcion local cambios alineacion delantero sitio,proporcion local cambios alineacion delantero en general,proporcion visitante cambios alineacion delantero en sitio,proporcion visitante cambios alineacion delantero en general

Para ver en más detalle una descripción sobre estos datos y como se calculan, ver excel “variables generadas.xlsx”

Estos datos al ser más complejos, cabe aclarar que cada fila se refiere a los datos de un partido. Estos datos describen los indicadores de los equipos que juegan el partido en base a los partidos previos. Por ejemplo, uno de estos indicadores es el porcentaje local ganados en sitio. Este porcentaje describe el porcentaje de partidos que el local ha ganado como local sobre el total de partidos que ha jugado como local.

## indicadoresEquipoHistoricoModelo

Se recogen los datos de los indicadores de los partidos en una temporada y liga junto con el ganador y los goles marcados para equipo filtrando los partidos posteriores a la jornada 9 de la temporada para que existan suficientes datos previos al partido para calcular estos indicadores en cada csv.

Las columnas son las mismas indicadoresEquipoHistorico añadiendo resultado\_local, resultado\_visitante y resultado\_partido.

Los datos de estos csv son los que se utilizan para entrenar y evaluar la precisión de los modelos.

## indicadoresEquipoPrepartido

Se recogen los datos de los indicadores de los partidos en una determinada fecha que estén por jugarse en una temporada y liga en cada csv con la misma estructura que indicadoresEquipoPrepartido. Estos se pueden utilizar para realizar una predicción con los modelos de en base a los indicadores del partido antes de jugarse, proporcionar la probabilidad que tiene cada equipo de ganar, empatar o perder o la probabilidad que tiene cada equipo de marcar un determinado número de goles.

# Análisis datos obtenidos

Con los datos generados previamente, se ejecuta un programa que extrae para cada equipo de cada temporada y liga, los valores de sus indicadores en su último partido jugado. Con esto se pretende evaluar como difieren las posiciones de los equipos según los indicadores. Por ejemplo, para cada equipo se extrae el valor de la “proporcion cambios defensa a centrocampista” en su último partido jugado y después de ordenan los equipos por este valor. De esta manera se obtiene una clasificación que se puede comparar con la clasificación verdadera por puntos para así ver si existen diferencias en las posiciones de los equipos según los indicadores. Por ejemplo, con esto se puede descubrir si los equipos que hacen muchos cambios de defensa a centrocampista tienen posiciones altas en la clasificación real por puntos o posiciones bajas y de esta manera evaluar si este tipo de cambios en bueno para ganar o no.

En la carpeta “pruebas”>”analisis atributos” se encuentra el excel “mapa de calor.xlsx” con los datos de los indicadores de los equipos y su diferencia con la clasificación por puntos en la temporada actual. Un ejemplo de un valor en esta tabla puede ser el siguiente. Si el Real Madrid en la columna de “proporción cambios centrocampista a delantero” tiene un -7, significa que en la clasificación ordenando por este indicador, tiene 7 posiciones menos que en la clasificación por puntos.

# Modelos entrenados con los datos

Con los datos generados previamente, se han entrenado diferentes modelos que se encuentran en la carpeta “pruebas”>”modelos”>”modelos\_futbol”. Se han juntado en un csv los datos obtenidos de diferentes ligas y temporadas hasta el momento de la entidad indicadoresEquipoHistoricoModelo y con ellos se han entrenado los modelos. De momento en estos modelos se ha puesto el foco en generar la probabilidad que hay para cada partido de que gane el local o el visitante o empaten, para los goles de cada equipo sería algo similar.

Uniendo los datos como comento, para la liga de España y liga de Inglaterra en los años 2022, 2023 y 2024, he obtenido 1730 registros que he introducido a los diferentes modelos. De momento los modelos generados por “Tensor.ipynb” y “Tensor1.ipynb” son los que más avanzados llevo.