

Presentación del conjunto de datos para el proyecto final *Una liga cualquiera*:

- Categoría: fútbol 7 alevín (S12), tercera división
- Periodo: 01/10/23 al 04/02/24, primera vuelta.
- Deporte: fútbol base catalán.

Características generales del conjunto de datos:

- Tipología: datos de competición deportiva.
- Sector: deporte base - fútbol.
- Tipo de datos: numéricos (posición, puntos, nº de partidos y goles) y categóricos (grupo).
- Fuente: Federación Catalana de Fútbol: <https://www.fcf.cat/competicio>.
- Contexto: clasificación de equipos en 29 grupos.

Contenido:

405 filas y 14 columnas.

- Grupo: subdivisión de la liga.
- Posición: rango en la tabla de clasificación.
- Nombre: identificación del equipo.
- Puntos: total de puntos acumulados.
- JC: partidos jugados en casa (local).
- GC: ganados en casa.
- EC: empates en casa.
- PC: perdidos en casa.
- JF: partidos jugados fuera (visitante).
- GF: ganados fuera.
- EF: empates fuera.
- PF: perdidos fuera.
- GF: goles a favor totales.
- GC: goles en contra totales.

Objetivos del proyecto:

1. Predecir la clasificación obtenida de los equipos en base al número de victorias, empates y derrotas, como local y como visitante.
2. Clasificar a los equipos en base a todas las variables disponibles.

Modelos utilizados:

1. Dado que las relaciones entre la variable dependiente y las independientes será probablemente lineal, utilizaremos un modelo de *Regresión Lineal Múltiple*.
2. Para clasificar a los equipos utilizaremos uno de estos dos modelos de la librería sklearn: *DecisionTreeClassifier* o *KNeighborsClassifier*.