

PROYECTO BASE DE DATOS NO RELACIONALES (NoSQL)

ANÁLISIS Y SCOUTING EN LIGAS DE FÚTBOL PROFESIONALES



ANTECEDENTES:

El fútbol no solo son equipos de fútbol seguidos por miles de fans, los cuales se enfrentan todas las semanas, sino que hay un trabajo detrás que involucra a la confección de estos. Desde el inicio del fútbol moderno, los principales clubs de Europa han dispuesto de ojeadores en los 4 continentes para poder captar a futuras perlas a precio de ganga y anticiparse a sus rivales.

Sin embargo, como ocurre en todos los ámbitos de nuestra sociedad, con la llegada de la tecnología, la relevancia de estos scouts ya no es tan importante, ya que todos los datos de los jugadores son recopilados en forma de big data y puede ser accesible para cualquier especialista y poder realizar un análisis más preciso de las características de los jugadores.

INTRODUCCIÓN:

Para la elaboración del proyecto, se ha simulado una colección de datos extraídos de algunas ligas europeas. Los datos han sido extraídos de las páginas especializadas “transfermarkt” y “Who Scored?”.



COLECCIONES:

La base de datos de este proyecto está conformada por 4 colecciones diferentes. En una se recogen los datos de los futbolistas, en otra su rendimiento, los datos de los equipos a los que pertenecen y sus patrocinadores.

*Colección jugador:

```
{
  "Cod": 1, (num)
  "nombre": "Jules Kounde", (str)
  "edad": 22 , (num)
  "posicion": "DEF", (str)
  "equipo": "SEV", (str)
  "finContrato": new Date("2024 - 06 - 30"), (date)
  "valorM€":50 , (num)
  "patrocinador":["001","A"]}, (array)
```

*Colección rendimiento:

```
{
  "_id":1 , (num)
  "goles":1 , (num)
  "PJ": 18 , (num)
  "minutosT":1950 , (num)
  "tarjA":2 , (num)
  "rating": 7.15, (num)
  "lesion":true} (boolean)
```

*Colección equipo:

```
{
  "liga": "ESP", (str)
  "nombreClub": "SEV", (str)
  "PJ": 38, (num)
  "golT": 59, (num)
  "valorTM€": 343.90, (num)
  "europa": true, (boolean)
  "patrocinador": [{      (array/nested documents)
    "CodP": "001",
    "tipo": "A"
  }, {
    "CodP": "007",
    "tipo": "A"
  }
  ],
}
```

*Colección patrocinador:

```
{
  "idPatrocinador": "001", (str)
  "nombre": "NIKE", (str)
  "contratoEquipo": [ (array/nested documents)
    {
      "tipoContrato": "A",
      "contratoM€": 10
    },
    {
      "tipoContrato": "B",
      "contratoM€": 6
    },
  ],
  "contratoJugador": [ (Array/nested Documents)
    {
      "tipoContrato": "A",
      "contratoM€": 2
    },
    {
      "tipoContrato": "B",
      "contratoM€": 1.5
    },
  ]
}
```

RELACIONES ENTRE LAS COLECCIONES:

Jugador:

Cod nombre edad posición equipo finContrato valorM€ patrocinador

Rendimiento:

_id goles PJ minutos tarjA tarjR rating lesión(boo)

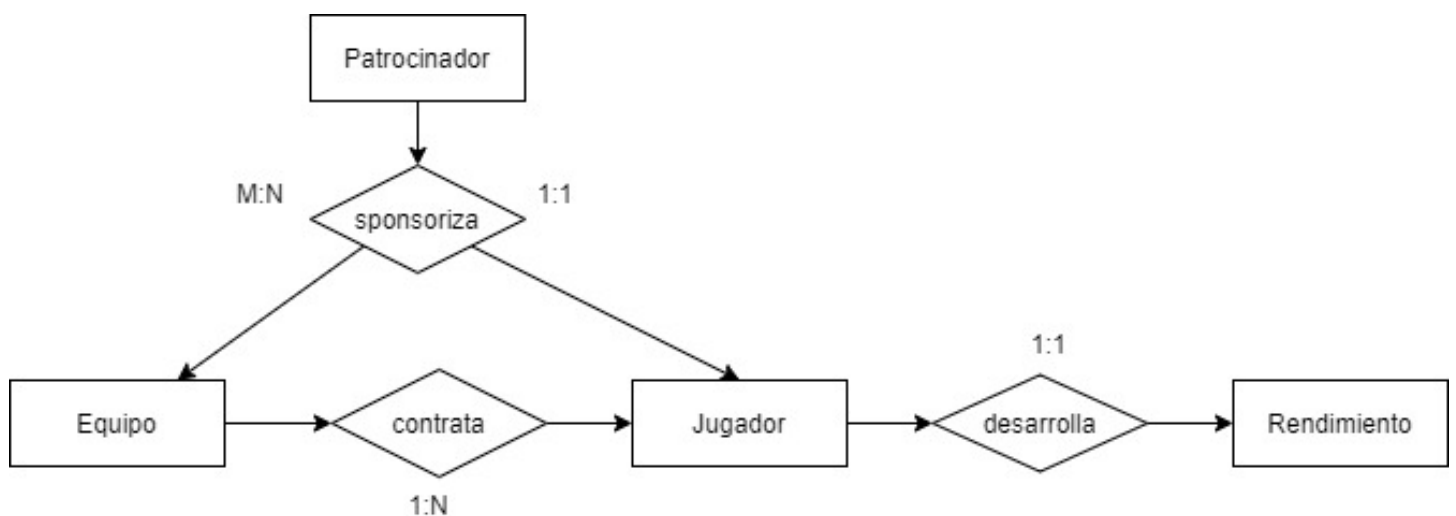
Equipo:

Lig nombreClub PJ golT valorTM€ Europa (boo) patrocinador

Patrocinador:

idPatrocinador nombre contratoEquipo contratoJugador

DIAGRAMA ERR SIMPLIFICADO:



TRATAMIENTO DE LOS DATOS: MÉTODO AGGREGATE ():

Las operaciones de agregación procesan registros de datos y devuelven resultados calculados. Las operaciones de agregación agrupan valores de varios documentos y pueden realizar una variedad de operaciones en los datos agrupados para devolver un solo resultado

En este proyecto se han usado:

- **\$match:** Empareja los datos con la instrucción introducida.
- **\$expr:** Permite el uso de expresiones de agregación dentro de un lenguaje de consulta.
- **\$group:** Agrupa los registros con el mismo valor dependiendo del campo elegido.
- **\$lookup:** Realiza una unión externa izquierda a una colección no fragmentada en la misma base de datos para filtrar los documentos de la colección "unida" para su procesamiento.
- **\$unwind:** Deconstruye un campo de matriz a partir de los documentos de entrada para generar un documento para cada elemento.
- **\$project:** Se usa como filtro para las siguientes etapas.
- **\$sum:** Realiza el sumatorio de los valores de los campos seleccionados.
- **\$avg:** Realiza la media de los valores de los campos seleccionados.
- **\$divide:** Realiza la división de los valores de los campos seleccionados.
- **\$multiply:** Realiza la multiplicación de los valores de los campos seleccionados.
- **\$push:** Permite rescatar campos que comparte la misma clave de grupo en forma de array; Solo disponible con \$group.
- **\$bucket:** Es un método de agrupación estableciendo límites o rangos. Cada grupo devuelto o "bucket" tiene un `_id`
- **\$facet:** Permite crear agregaciones multifacéticas que caracterizan los datos en múltiples dimensiones dentro de una sola etapa de agregación.
- **\$map:** Operador de arrays. Permite trabajar con campos tipo array.
- **\$zip:** Permite la manipulación de arrays.

IMAGEN EXTRAIDA CON \$MERGE Y EXCEL

