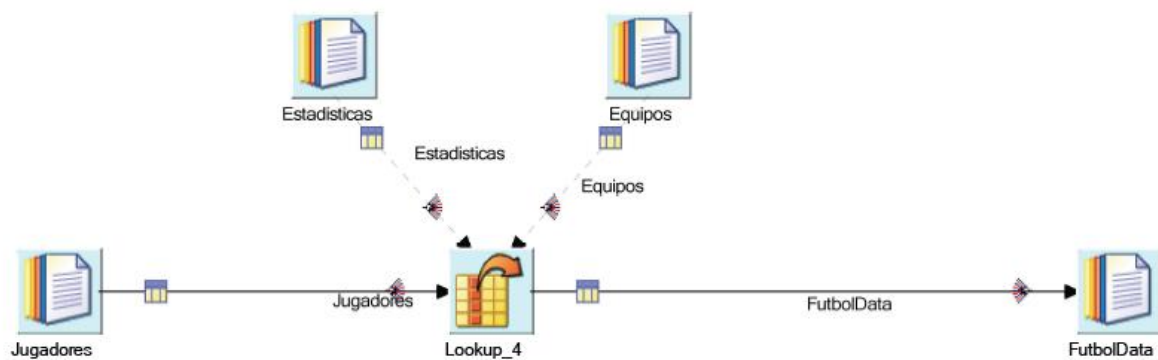


## Proyecto de bases de datos con *Datastage*

Para este proyecto se requería de un o varios ingenieros de Datastage para la gestión de datos de una liga profesional. Nos preveían de una serie de ficheros con datos de Clubes.txt, Jugadores.txt y estadísticas.txt. Están subidos en el repositorio el fichero con todos los job realizados y las opciones seleccionadas dentro de cada job.

De cara a facilitar futuras peticiones por parte de la liga se procede a juntar los datos de los 3 ficheros txt en uno solo. Una vez cargado los datos y los metadatos de cada fichero se proceden a la construcción del job. Ya que hay columnas en común la función Lookup es ideal para leer y cargar en un fichero secuencial los datos agregados sin repetir ninguna columna y añadiendo las columnas únicas de cada fichero. Se pueden observar que los 3 ficheros se particionan en el look up para volver a cargarse en el fichero secuencial final. A continuación, se observa el job de preparación:



Finalmente, se guardan los metadatos de el fichero target para la construcción de los próximos Jobs.

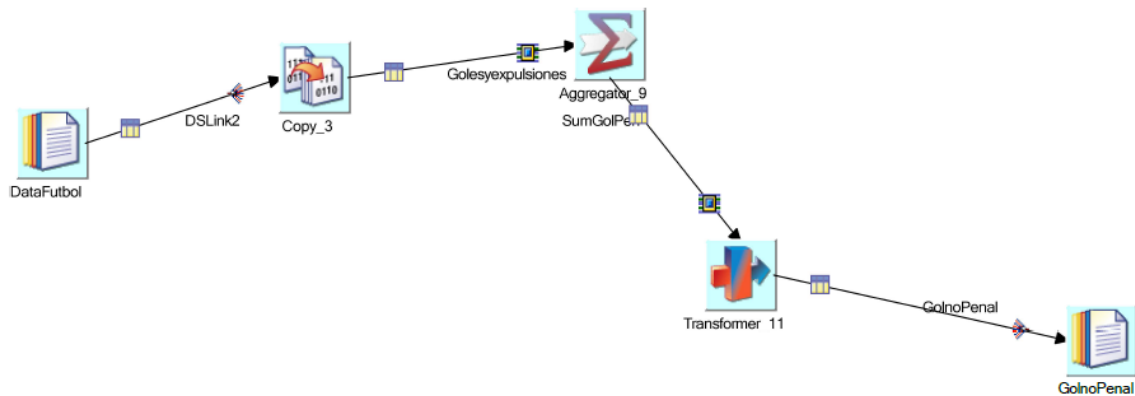
1. La primera tarea era calcular la suma de goles y de expulsiones agrupándolo por federaciones.

Una vez cargado el fichero secuencial de preparación, hecho en el job pasado y sus metadatos, hacemos a construir un nuevo job atendiendo a las características que nos piden. La función copy nos servirá para seleccionar las columnas que nos interesen y la función aggregator nos permitirá sumar los goles y expulsiones, además de añadir la opción group by de las "federaciones". A continuación se visualiza el job:



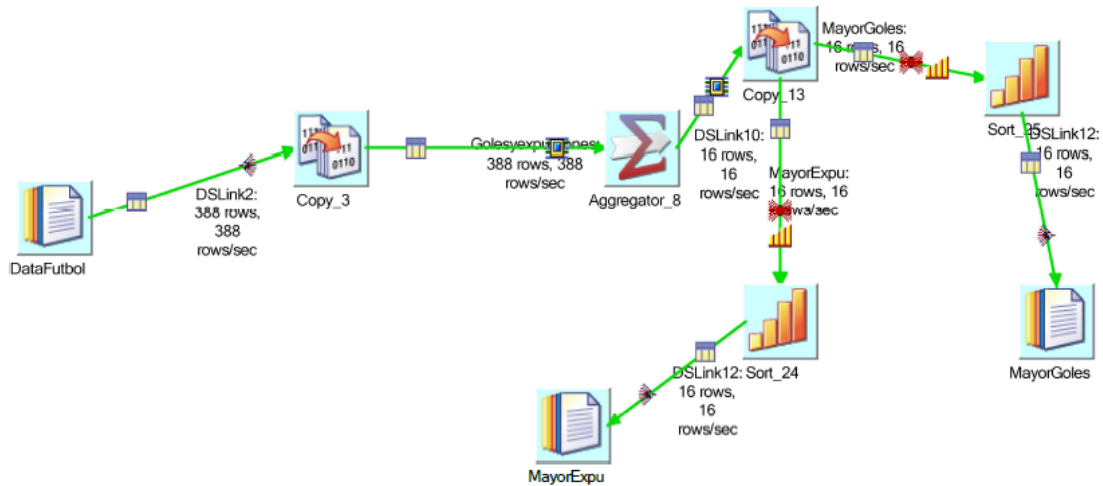
- Posteriormente se pide que mostremos el número de goles no conseguidos de penalti y que esté agrupado por Clubes.

Utilizando el mismo fichero secuencia source lo particionamos utilizando un copy sacando las columnas que nos interesa. Metemos un agregador para la suma de goles y un transformer para filtrarlo por goles de no penalti:



- Se pide el nombre del club en el que juega el máximo goleador y el del jugador con más expulsiones.

Para este yo se particiona con el copy seleccionando del fichero source las columnas necesarias. Se utiliza un agregador para hacer un group by por clubes y se subdivide en 2 sorts para crear dos ficheros ordenados de forma descendente, uno ordenado por expulsiones y el otro ordenado por goles de mayor a menor.



4. Por último, se pide que se cree una tabla en DB2 que incluya el Nombre de jugador, club, numero de partidos, goles y expulsiones y la fecha que se inserta este registro.

Para este job se reutiliza el mismo fichero source, se hace un copy para seleccionar las columnas mencionadas y se transcribe a un conector db2. Tiene que haber una instancia creada.

