
Lab Book

Bases de Datos Avanzadas

Mario Pérez Sánchez-Montañez

David Camuñas Sánchez

February 15, 2020

Contents

Jueves, 13 Febrero 2020	1
1 Obtención de los datos para el <i>DataWarehouse</i>	1
2 Cambio en diseño del <i>DataWarehouse</i>	1
Viernes, 14 Febrero 2020	2
1 Primer versión de la base de datos	2

Jueves, 13 Febrero 2020

1 Obtención de los datos para el DataWarehouse

Desarrollamos tres *scripts* en *Python* utilizando técnicas de *web scrapping* para obtener datos de jugadores y equipos de *Basketball Reference*

2 Cambio en diseño del DataWarehouse

Decidimos que vamos a tener una tabla con todos los jugadores, con los 30 equipos y de *rookies* (jugadores noveles). Todo ello estará relacionado en una tabla central, siguiendo el modelo de estrella.

Viernes, 14 Febrero 2020

1 Primer versión de la base de datos

Realizamos la construcción de la primera versión de la base de datos. Uno de los problemas que hemos encontrado es con los jugadores que han jugado en varios equipos en la misma temporada, ya que en la fuente que hemos usado se muestran datos en cada equipo donde han jugado y lo totales, identificando estos últimos como 'TOT' en la columna de equipo.

Hemos decido mantener esta estructura identificando las estadísticas totales con -1, para que en la base de datos no se utilice un id de un equipo existente.