**PROYECTO SQL** **videoclub**, creamos una base de datos relacional para un videoclub para poder hacer consultas, antes haremos limpia.

**PASOS DADOS:**

* Abro en Excel mis tablas de datos para ver de qué datos dispongo, separados por columnas para visualizarlos mejor.
* Hago el diseño en una hoja de papel de mis relaciones entre columnas.
* Creamos una base de datos nueva ‘videoclub’.
* Desde MYSQL importo las tablas csv.
* Limpieza de tabla en MYSQL:

La tabla old-hdd relaciona las tablas: actor, category y film.

Pero hay columnas repetidas ‘first\_date’ y ‘last\_name de la tabla ‘actor’ y ‘title’ y ‘release\_year’ de la tabla ‘film’.

Compruebo que las columnas ‘first\_date’ y ‘last\_name’ de la tabla ‘old-hdd’ sean iguales que en la tabla principal: ‘actor’ con **select**.

‘SELECT actor.actor\_id,old\_hdd.first\_name,old\_hdd.last\_name

FROM actor

INNER JOIN old\_hdd ON actor.first\_name = old\_hdd.first\_name AND actor.last\_name = old\_hdd.last\_name;’

En la tabla ‘actor’ hay duplicado ‘SUSAN DAVIS’ con dos keys diferentes, borro un registro:

**select first\_name,last\_name,count(\*) from actor group by first\_name,last\_name having count(\*)>1;**

**delete from actor where actor\_id=110;**

Creo una nueva columna actor\_id (int) en la tabla ‘old\_hdd’ para relacionar ambas tablas, relleno los datos de la nueva columna y elimino las columnas duplicadas.

ALTER TABLE old\_hdd ADD COLUMN actor\_id int;

update old\_hdd

INNER JOIN actor ON actor.first\_name = old\_hdd.first\_name AND actor.last\_name = old\_hdd.last\_name

set old\_hdd.actor\_id=actor.actor\_id;

Creo otra columna ‘film\_id’ en la tabla old\_hdd (int) para relacionar ambas tablas y elimino las columnas repetidas. Las columnas que me quedan en la tabla old\_hdd **son foreign Key.**

Elimino la columna ‘original\_language\_id’ por tener todos los valores nulos.

Elimino la columna ’rating’ por tener datos que no se entienden.

Reemplazo las columnas ‘last\_update’ que están en formato ‘VACHAR’ y las dejo con formato fecha, creando columnas nuevas con el formato válido

Mi base de datos está limpia.

* Creamos la primary key y foreing keys de cada tabla desde esquema para que me haga las relaciones/diagrama.