Bases de Datos

**Practica 2: Gestión de la información de Productos Audiovisuales**

**Curso 2017-2018**

Integrantes del grupo: Ignacio Alonso, Patricia Briones y Jorge Fernández(coordinador)

e-mails en orden: [716344@unizar.es](mailto:716344@unizar.es) , [735576@unizar.es](mailto:735576@unizar.es) , [721529@unizar.es](mailto:721529@unizar.es)

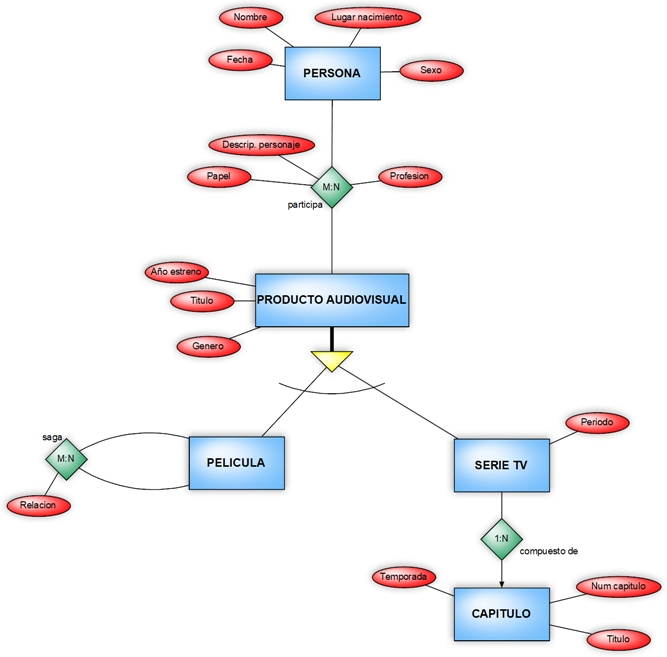
Grupo: Grupo 7 Tardes

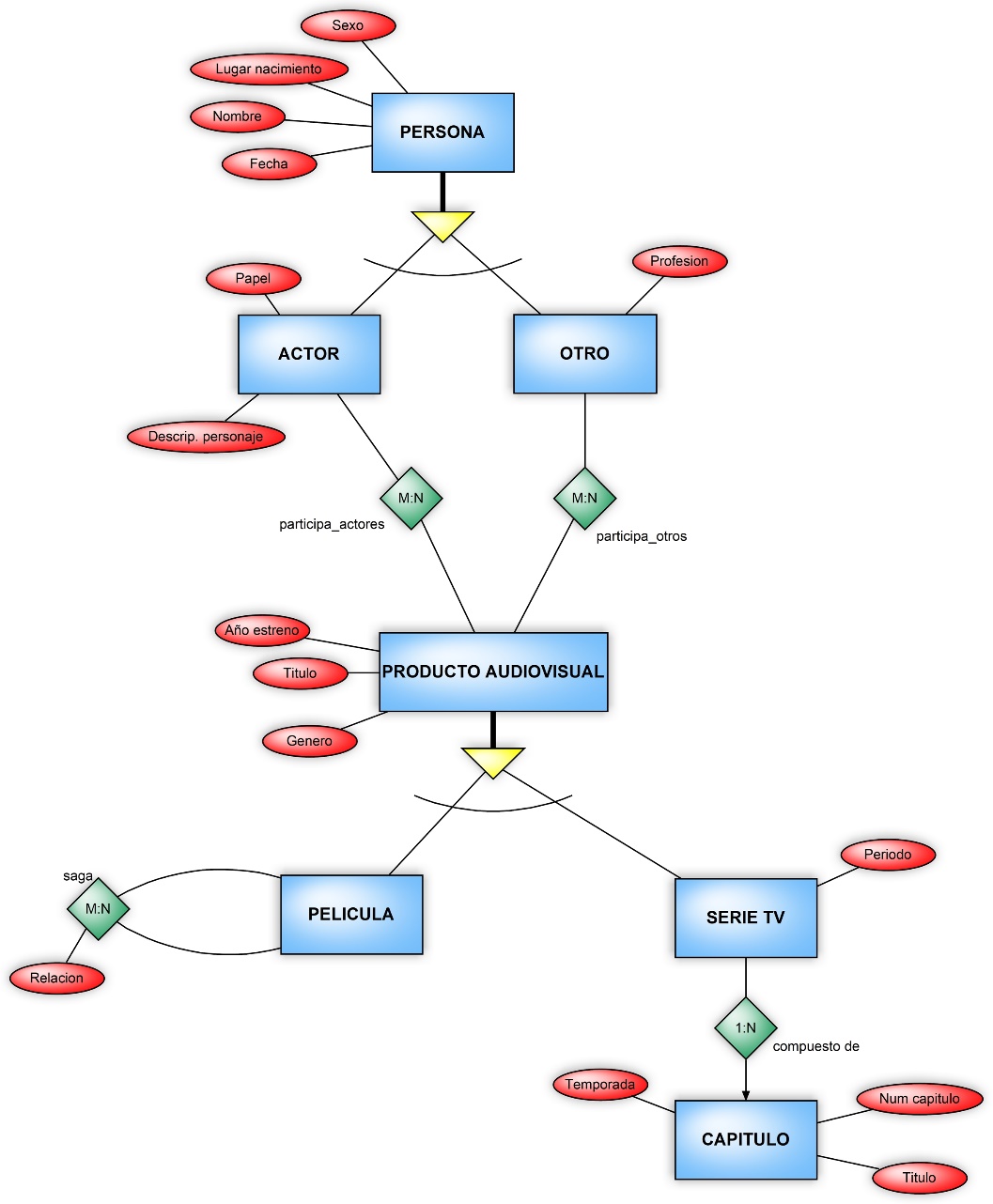
**Índice:**

1. *Esquema E/R ………………………………………………. Pág.3*
2. *Modelo Relacional Normalizado…………………. Pág.5*
3. *Traducción del modelo relacional a SQL……… Pág.7*
4. *Poblado de la base…………………………………….. Pág.8*
5. *Consultas………………………………………………….. Pág.9*
6. *Información sobre el grupo……………………… Pág.16*

**1.Esquema E/R:**

El esquema E/R sobre el que se ha trabajado es el siguiente:



****Otro E/R:

Decisiones más importantes:

-Se ha decidido crear una única relación participa entre Persona y Producto Audiovisual que contiene como atributos los datos relacionados a la profesión que ejercieron dentro de el Producto, en vez de crear dos relaciones entre Persona y Producto Audiovisual, una representando a las personas que trabajaron como actores en el Producto, y otra para representar a las personas que trabajaron en otras facetas.

Restricciones de nuestro modelo E/R:

* Una persona que haya participado en un producto audiovisual, en caso de haber realizado una misma profesión varias veces(véase, un actor que interpreta varios papeles en una misma serie/película), aparecerá únicamente con uno de estos.

**2.Modelo Relacional:**

Persona (ID, nombre, sexo, Lugar\_nacim, Fecha\_nacim)

Participa (ID\_persona, ID\_producto, profesion, papel, info)

Producto\_Audiovisual (ID, titulo, Anyo\_prod)

Genero (Id\_PA, genero)

Pelicula (ID\_PA)

Serie\_TV (ID\_PA, periodo)

Saga (ID\_pelicula, ID\_pelicula\_2, relación)

Capitulo (ID\_capitulo, nombre, temporada, num\_capitulo, serie)

En la relación Participa, el atributo ID\_persona, se trata de una clave extranjera que hace referencia a un dato de la entidad Persona, así como ID\_producto hace lo propio con un dato de la entidad Producto Audiovisual.

En las relaciones Genero,Película, y Serie\_TV, el atributo ID\_PA hace referencia como clave extranjera al dato ID de un Producto Audiovisual.

En Saga, ID\_pelicula y ID\_pelicula\_2, también son clave extranjera que hacen referencia a un ID correspondiente a un Producto Audiovisual.

Por último, en Capitulo, el atributo serie es clave extranjera que hace referencia a el ID de un Producto Audiovisual.

A la hora de traducir el esquema E/R a modelo relacional, se ha optado por añadir códigos identificativos a las Personas, Productos Audiovisuales y Capitulos, dado que el único atributo que podría usarse como clave primaria, el nombre de estos, puede dar a confusión, debido a que existen Personas con el mismo nombre, así como Peliculas las cuales tienen nombre similar, o tienen el mismo nombre y la única diferencia se encuentra en que la última además de el nombre común, va seguida de un numero que la distingue de la anterior, al igual que puede ocurrir con los capítulos. En cuanto al modelo relacional, se puede observar una relación M:N entre Persona y Producto Audiovisual, que representa las personas que han trabajado en un determinado Producto Audiovisual, y contiene atributos que indican la profesión que ejerció en dicho Producto, asi como el personaje que interpretó y una breve descripción de este en caso de que participara como actor.

En cuanto a el Producto Audiovisual, se ha decidido hacer uso de una especialización para diferenciar los distintos tipos de Productos Audiovisuales, así, en este caso, estos pueden ser Peliculas o Series. Después, para las películas, existe una relación M:N entre distintas Peliculas que identifica aquellas Peliculas que tienen algún tipo de relación, esta ultima está indicada mediante el atributo tipo de la relación. Por último, con respecto a las Series, además de almacenar el tiempo que estuvieron en emisión, estas tienen una relación 1:N con los Capitulos, que representan a Capitulos concretos de una Serie, con sus respectivos atributos.

Normalización:

1ª Forma normal: Hace referencia a atributos multievaluados, en nuestro caso, el único problema podrían ser los Productos Audiovisuales, que pueden tener varios géneros, pero en nuestro modelo E/R ya esta contemplado este problema, ya que se ha añadido una nueva tabla llamada Generos, para indicar todos los géneros de cada Producto Audiovisual.

2ª Forma normal: Solo es necesario tener en cuenta aquellas relaciones con claves compuestas. Ninguna clave de cada entidad se deduce a partir de los atributos.

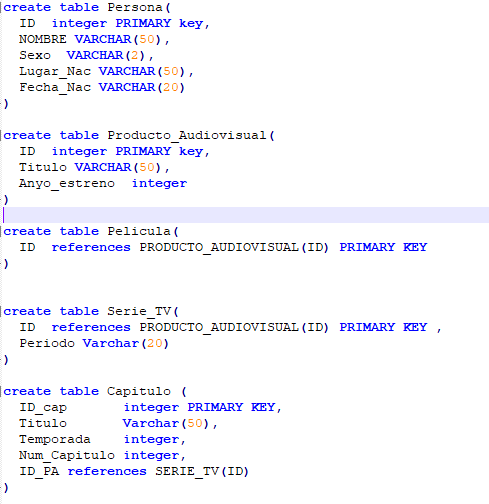
3ª Forma normal: Sí está en 3ª FN. No existe ninguna dependencia funcional transitiva entre los atributos

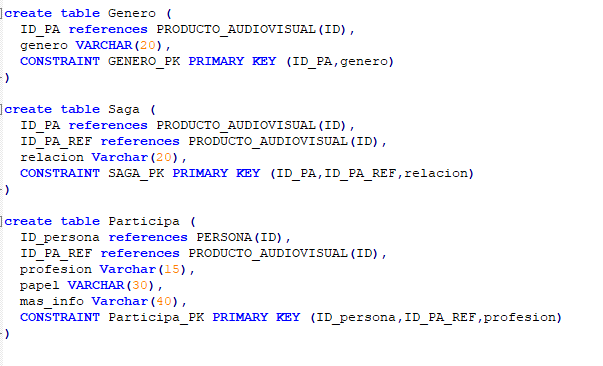
FN Boyce-Codd: Está en FN BC porque todas las dependencias dependen de la clave primaria o de claves candidatas.

4ª Forma normal: Todas las dependencias multievaluadas dependen de la clave primaria, por lo tanto, sí que está en 4ª FN.

**3.Traduccion del modelo relacional a SQL:**

Para pasar del modelo relacional a SQL, únicamente hemos traducido los datos del modelo relacional normalizado a SQL añadiendo las sentencias para crear las tablas:





**4.Poblado de la base:**

Para realizar el poblado de la base, hemos agrupado todos los datos en distintos ficheros csv, haciendo una serie de consultas sobre la base de datos MySql proporcionada y posteriormente se han ido introduciendo en la base de datos. En cuanto al poblado de la base, nos aparecieron problemas a la hora de insertar en la base de datos un gran número de tuplas a una misma tabla debido también al gran número de tuplas que contenían los distintos ficheros de datos. Para realizar el poblado de la base de datos, se hizo uso de la herramienta DataGrip.

Estadísticas de nuestras tablas:

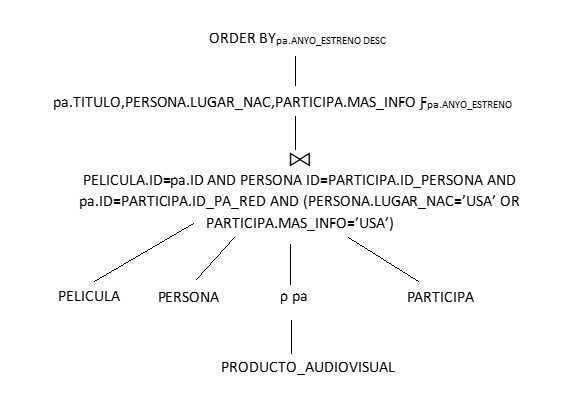
* Persona: 38978 tuplas, con un tamaño aproximado de 1480 KBytes
* Participa: 97910 tuplas con un tamaño de 3406 KBytes
* Producto\_Audiovisual: 5520 tuplas con un tamaño de 171 KBytes
* Pelicula: 3803 tuplas con un tamaño de 30 KBytes
* Serie\_TV: 1717 tuplas con tamaño de 29 KBytes
* Género: 22760 tuplas con un tamaño de 441 KBytes
* Saga: 180 tuplas con tamaño de 5 KBytes
* Capitulo: 162 tuplas con tamaño de 6 KBytes

Ejemplos de tuplas:

* Persona :31, "'El Francés', José", m, "Montpellier, Hérault, France",1971
* Producto Audiovisual:961, 003 y medio, 1979
* Saga: 175875, 175888, follows
* Serie TV: 961, 1979-1980
* Capitulo: 196409, Buscando un Valium desesperadamente, 2, 4, 196397

1. **Consultas**

1ª Consulta: Busca las películas relacionadas con una población del mundo:



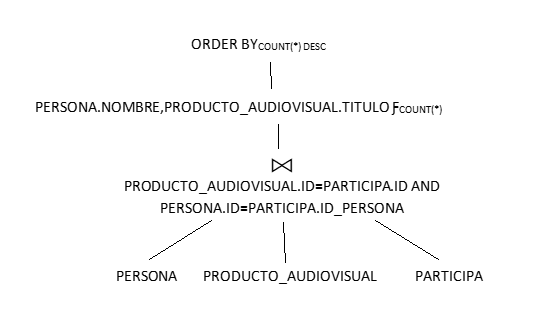
En esta consulta, buscamos el nombre de las películas, asi como su año de estreno, que están relacionadas con una población, en este caso esa población es USA.

Los resultados son los siguientes:

In the Soup 1992 USA

"Amore, piombo e furore" 1978 USA

2ªConsulta: Personas con más cargos en una película o serie:



En esta consulta, buscamos el número de apariciones de una persona en la tabla Participa, y ordenamos estos datos según el nombre del actor y las películas o series en las que participó.

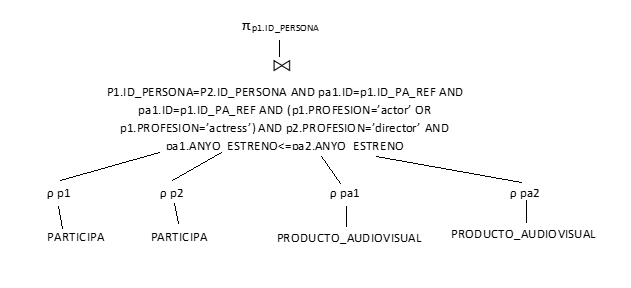
Resultados de la Consulta:

4 "Abad, Diego"

4 "Aguilar Vicente, Antonio"

4 "Almodóvar, Pedro"

3ª Consulta: Directores que empezaron como actores:



En esta consulta, buscamos dentro de la tabla Participa aquellas personas que aparecen tanto con el papel de actor como con el de director en esta tabla, y nos aseguramos además de que aparezca primero como actor, haciendo uso del atributo de los Productos audiovisuales, anyo\_estreno.

Resultados:

1160

1188

20256

20573

20682

21709

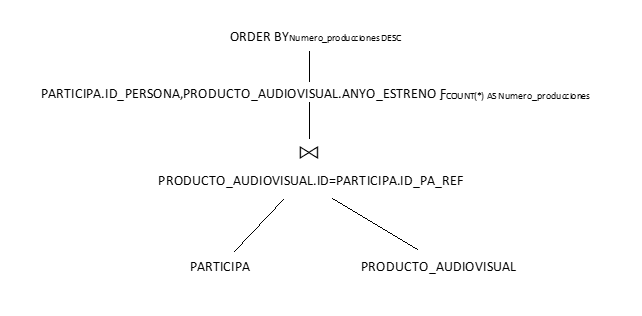
21954

22123

23427

30839

1ª Consulta realizada por nosotros: Mayor numero de producciones estrenadas en un año por una persona:



En esta consulta, buscamos en la tabla Participa, y realizamos la ordenación según el año de estreno del Producto Audiovisual.

Los resultados son:

11 2486003 1975

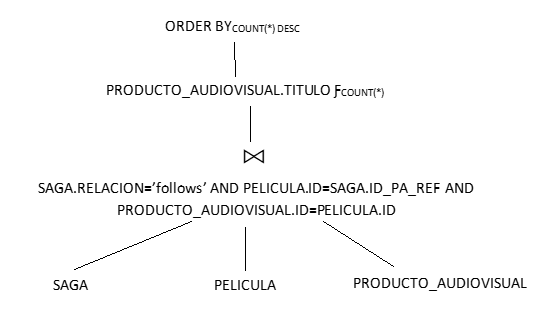
10 90447 1982

9 71922 1976

9 71922 1977

9 91086 1977

2ª Consulta propuesta: Peliculas con más secuelas:



En esta consulta, se realiza el select sobre la tabla Saga, buscando las veces que aparece una misma película

Los resultados son:

4 La maldición de la bestia

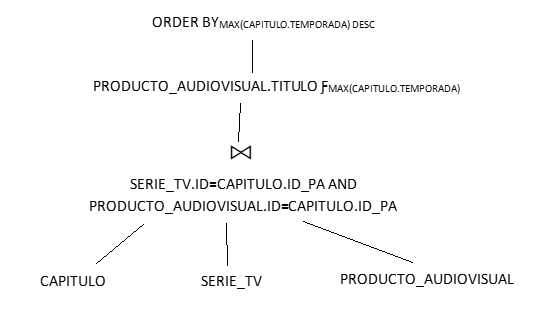
3 El retorno del Hombre-Lobo

3 La guerra de los niños

2 Bodas de sangre

2 La segunda guerra de los niños

3ª Consulta propuesta: Series con más temporadas:



En esta consulta, realizamos la operación count sobre la tabla Series\_TV y la ordenamos por serie.

Como resultados nos aparecieron:

6 Canguros

4 Eurocops

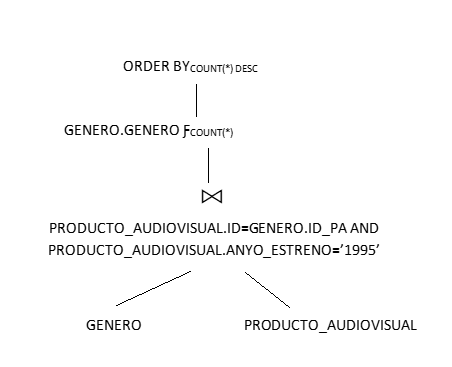
3 Menudo es mi padre

2 Dark Justice

2 Función de noche

Debido al problema antes nombrado de la introducción de los partidos, estos resultados pueden sufrir modificaciones en caso de haber podido introducir todos los partidos correctamente.

4ª Consulta propuesta: Genero más popular:



Resultados de la consulta:

5 based-on-novel

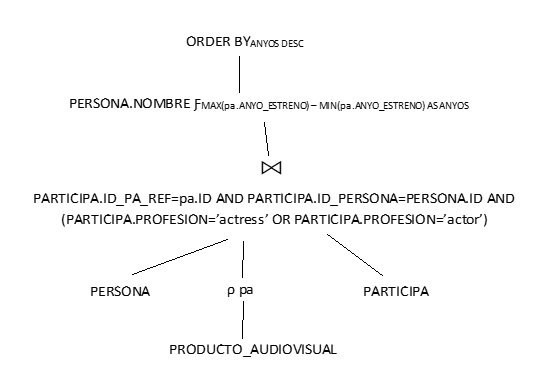
4 black-comedy

4 family-relationships

4 non-fiction

4 children

5ª Consulta propuesta: Actores con más años de vida profesional:



Resultados de la consulta:

"Abril, Victoria" 24

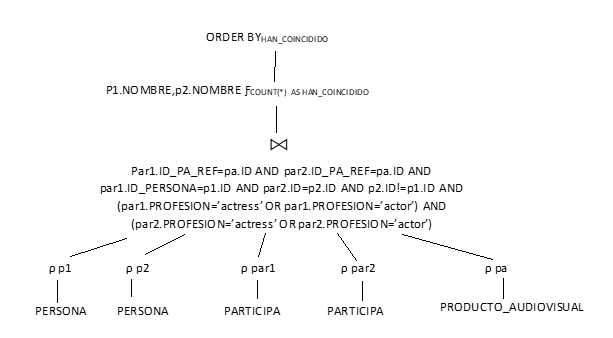
"Acaso, Félix" 23

"Acevedo, Mari Ángeles" 23

"Adánez, María" 21

"Adriani, Patricia" 20

6ª Consulta propuesta: Coincidencias de una pareja de actores:



Resultados de la consulta:

"Abeledo, Danny" "Abeledo, Danny" 2

"'El Francés', José" "Abdallah, Mouzahem" 1

"Abascal, Silvia" "Adánez, María" 1

"Abbu, Anwuar" "Adams, Maud" 1

1. **Información sobre el grupo:**

HORAS:

Patricia Briones Yus: 2h – E/R, 1h - sacar datos, 1h poblar, 2h – consultas, 1.30h – álgebra relacional.