



PRÁCTICA 2, FASE 1

Archivos y bases de datos

Joel López – joel.lopez.2015
Alex Pons – ls31155

Índice

1. Resumen del enunciado	2
2. Conceptos teóricos.....	2
3. Diagramas de la base de datos	3
4. Extracción de datos por línea de comandos.	4
En LOCAL:	4
En REMOTO:	5
5. Proceso de importación y exportación de la base de datos.	8
6. Dedicación.....	10
7. Conclusiones	10
8. Bibliografía	11

1. Resumen del enunciado

Tenemos que configurar una interface con Java capaz de conectarse con una base de datos que se encuentra subida en el servidor remoto Puigpedros de La Salle. También se nos solicita hacer un buscador de esta base de datos offline, por lo que tendremos que tener una copia de seguridad local de la base de datos.

2. Conceptos teóricos

OLTP (Procesamiento de Transacciones En Línea)

El OLTP es un tipo de procesamiento que facilita y administra aplicaciones transaccionales, usualmente para entrada de datos y recuperación y procesamiento de transacciones. Los paquetes de software para OLTP se basan en la arquitectura cliente-servidor ya que suelen ser utilizados por empresas con una red informática distribuida. El procesamiento de transacciones en línea tiene dos claros beneficios: la simplicidad y la eficiencia. Cuando hablamos de sus inconvenientes se puede decir que en su utilización hay algunas cuestiones en las que se debe pensar ya que pueden suponer un problema: la seguridad y los costes económicos o de tiempo.

OLAP (Procesamiento Analítico En Línea)

El OLAP es una solución utilizada en el campo de la llamada Inteligencia de negocios cuyo objetivo es agilizar la consulta de grandes cantidades de datos. Para ello utiliza estructuras de datos diversas, normalmente multidimensionales (Cubos OLAP), que contienen datos resumidos de grandes BBDD o Sistemas Transaccionales (OLTP). Se usa en informes de negocios de ventas, marketing, informes de dirección y áreas similares. La principal característica que potencia, es que es lo más rápido a la hora de ejecutar sentencias SQL de tipo SELECT, en contraposición con OLTP que es la mejor opción para operaciones de tipo INSERT, UPDATE Y DELETE.

Data Warehouse (almacenamiento de datos)

Data warehouse es como se denomina al almacén electrónico donde generalmente una empresa u organización mantiene una gran cantidad de información. Los datos de un *data warehouse* deben almacenarse de forma segura, fiable, fácil de recuperar y fácil de administrar.

Ventajas principales de usar *data warehouse*:

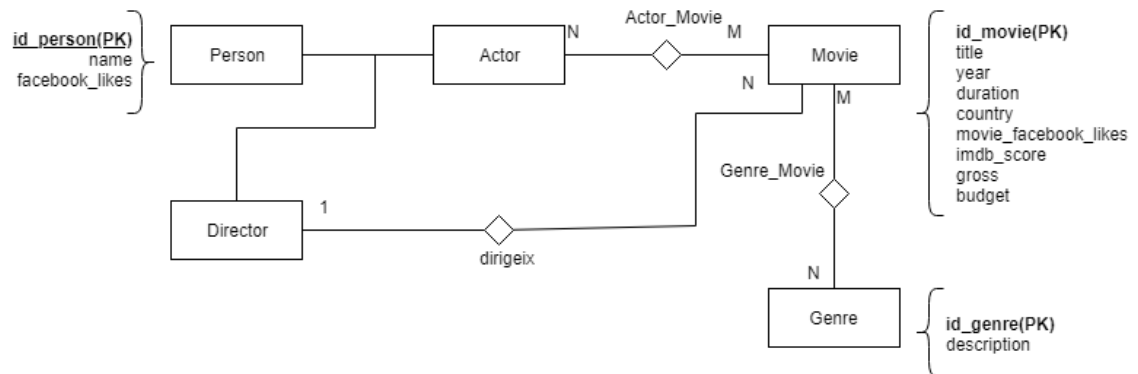
- Los almacenes de datos hacen más fácil el acceso a una gran variedad de datos a los usuarios finales
- Facilitan el funcionamiento de las aplicaciones de los sistemas de apoyo a la decisión tales como informes de tendencia, por ejemplo: obtener los ítems con la mayoría de las ventas en un área en particular dentro de los últimos dos años; informes de excepción, informes que muestran los resultados reales frente a los objetivos planteados a priori.

Primeros inconvenientes:

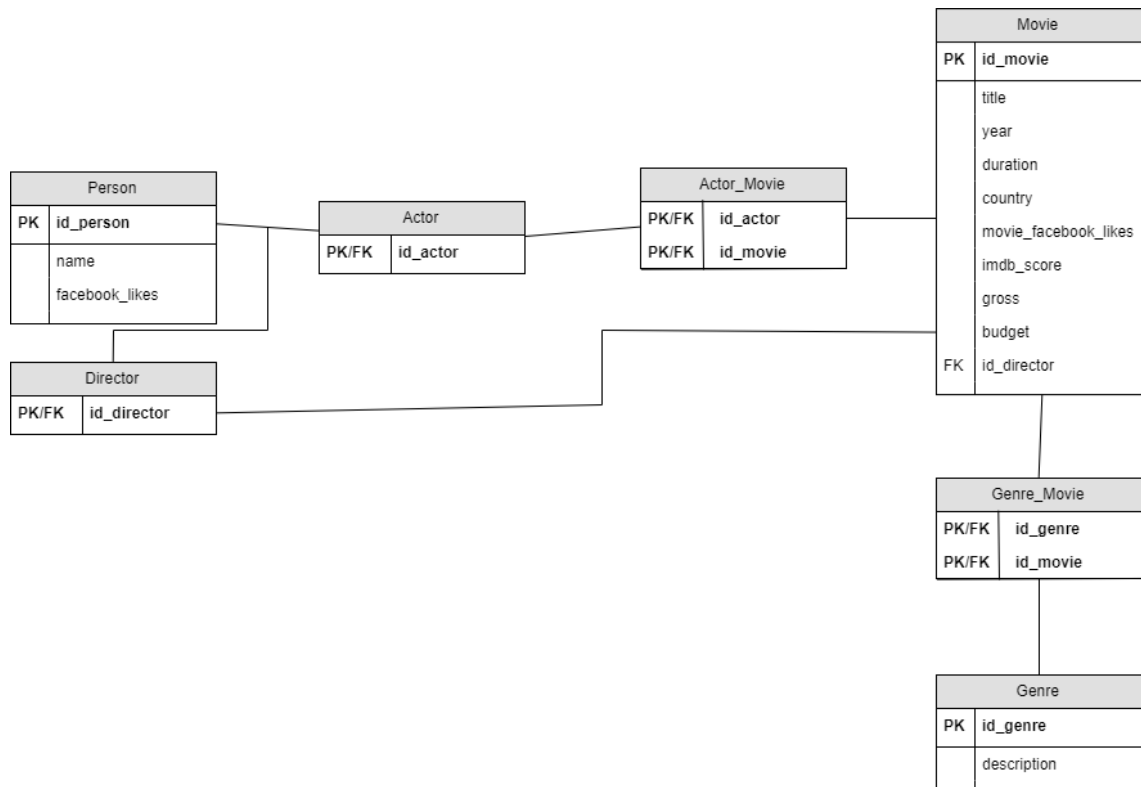
- A lo largo de su vida los almacenes de datos pueden suponer altos costos. El almacén de datos no suele ser estático. Los costos de mantenimiento son elevados.
- Los almacenes de datos se pueden quedar obsoletos relativamente pronto.
- A veces, ante una petición de información estos devuelven una información subóptima, que también supone una pérdida para la organización.

3. Diagramas de la base de datos

Modelo Entidad-Relación



Modelo Relacional

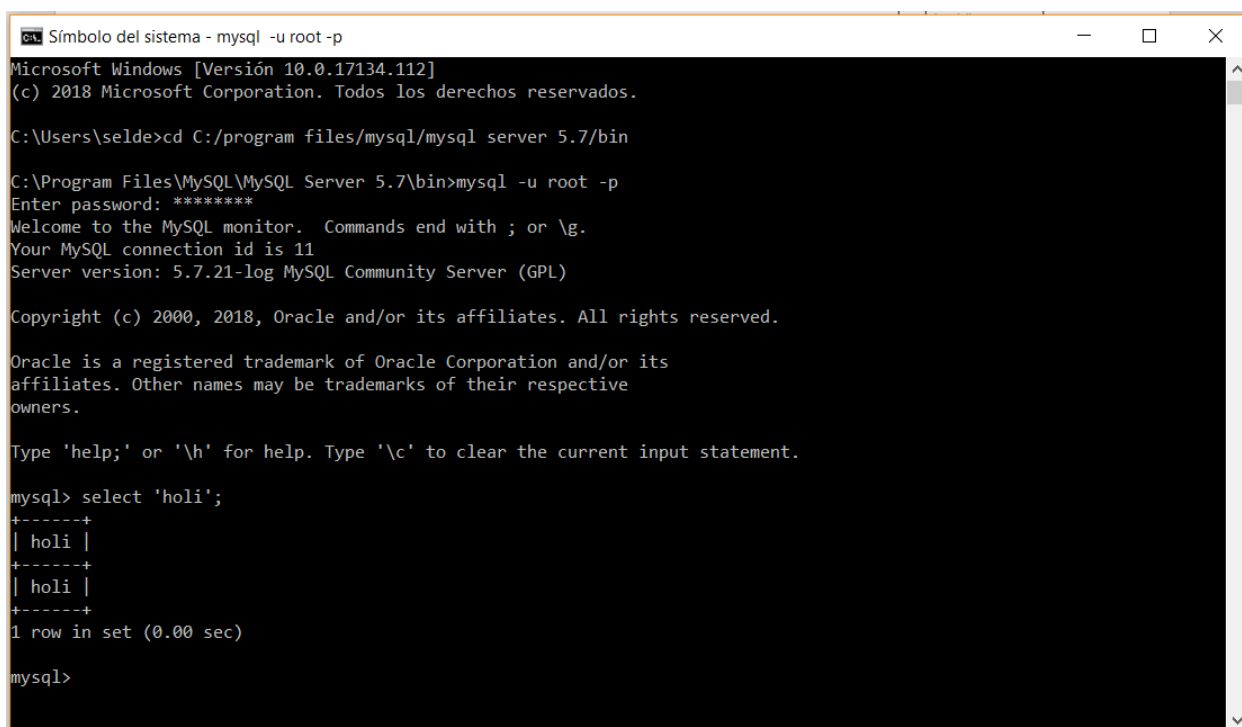


4. Extracción de datos por línea de comandos.

Para acceder a la información de la base de datos que debíamos importar debimos utilizar una serie de comandos básicos tras establecer la conexión en mysql desde el canal SSH, y para ver la información también de la base de datos local, donde íbamos a exportar los datos:

En LOCAL:

- **Para abrir el mysql por cmd:** ir hasta el directorio de mysql.server, y una vez allí escribir: `mysql -u <usuario> -p`
- *[luego pide contraseña, la introduces y has accedido al mysql desde la cmd]*



```
Símbolo del sistema - mysql -u root -p
Microsoft Windows [Versión 10.0.17134.112]
(c) 2018 Microsoft Corporation. Todos los derechos reservados.

C:\Users\selde>cd C:/program files/mysql/mysql server 5.7/bin

C:\Program Files\MySQL\MySQL Server 5.7\bin>mysql -u root -p
Enter password: *****
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 11
Server version: 5.7.21-log MySQL Community Server (GPL)

Copyright (c) 2000, 2018, Oracle and/or its affiliates. All rights reserved.

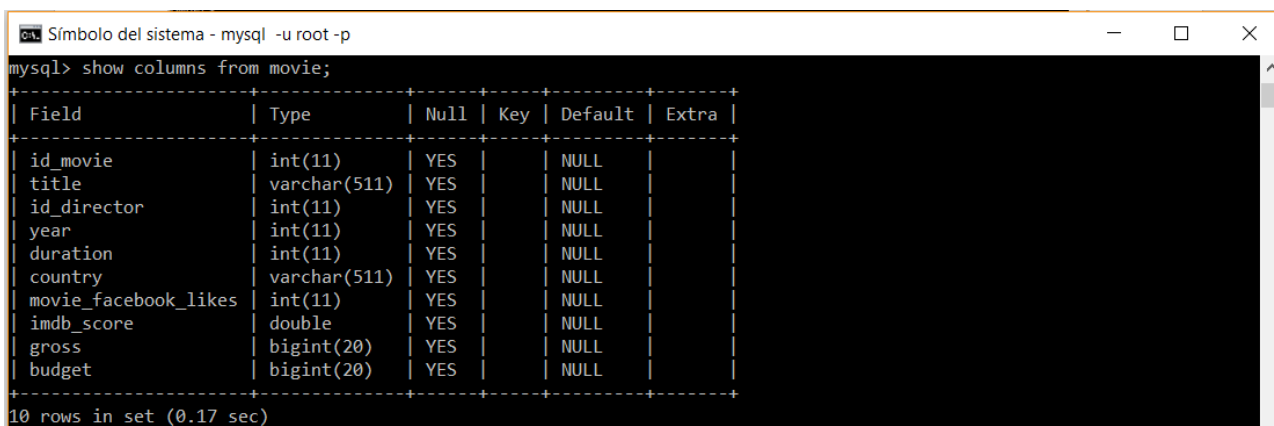
Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql> select 'holi';
+-----+
| holi |
+-----+
| holi |
+-----+
1 row in set (0.00 sec)

mysql>
```

- ➔ A partir de aquí, las comandas a utilizar son las mismas que si estuviéramos en MySQL (*use Movies;* para ver la base de datos; *show tables;* para ver las tablas de la base de datos que estamos utilizando; ***show columns from <tabla>;*** para ver todos los atributos de una tabla y sus respectivos tipos, además de otra información no tan relevante... -esto fue útil para ver cómo son las tablas de la base de datos remota!).

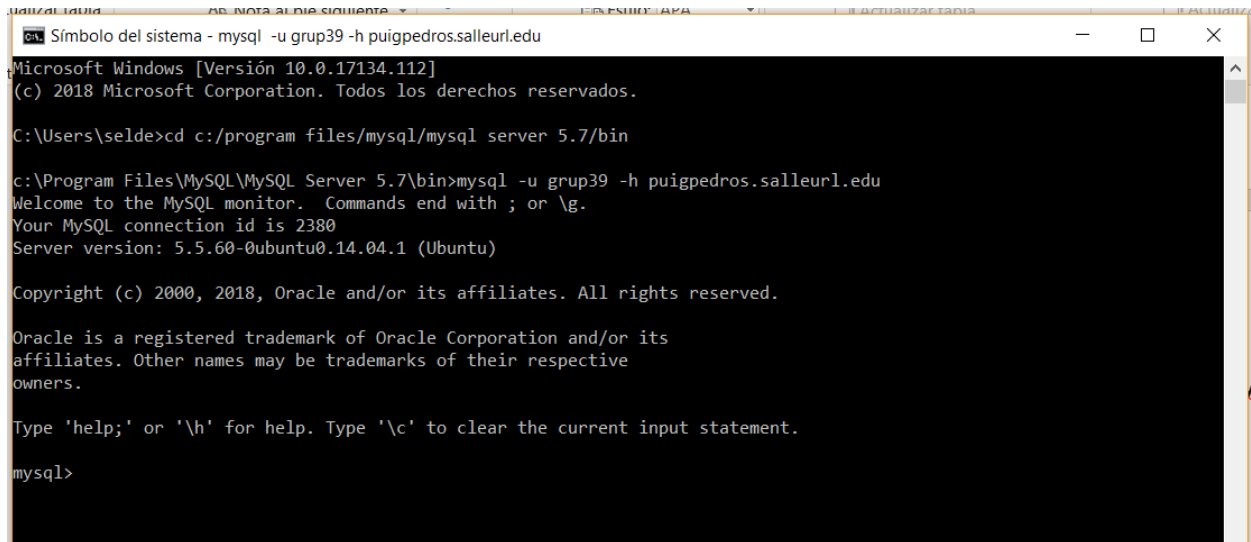


```
Símbolo del sistema - mysql -u root -p
mysql> show columns from movie;
+-----+-----+-----+-----+-----+-----+
| Field          | Type          | Null | Key | Default | Extra |
+-----+-----+-----+-----+-----+-----+
| id_movie       | int(11)       | YES  |     | NULL    |       |
| title          | varchar(511)  | YES  |     | NULL    |       |
| id_director    | int(11)       | YES  |     | NULL    |       |
| year           | int(11)       | YES  |     | NULL    |       |
| duration       | int(11)       | YES  |     | NULL    |       |
| country        | varchar(511)  | YES  |     | NULL    |       |
| movie_facebook_likes | int(11)       | YES  |     | NULL    |       |
| imdb_score     | double        | YES  |     | NULL    |       |
| gross          | bigint(20)    | YES  |     | NULL    |       |
| budget         | bigint(20)    | YES  |     | NULL    |       |
+-----+-----+-----+-----+-----+-----+
10 rows in set (0.17 sec)
```

En REMOTO:

Podemos ir a la carpeta de mysql server, como antes, o abrir directamente desde el **mysql Shell**, y escribir:

- **Para conectarnos al servidor remoto de puigpedros por línea de comandos:**
`mysql -u <usuario1> -h puigpedros.salleurl.edu`
- En caso de pedirnos contraseña (no la pide, pero bueno), al no tener ninguna simplemente pulsamos enter y ya estaríamos dentro:



```
Símbolo del sistema - mysql -u grup39 -h puigpedros.salleurl.edu
Microsoft Windows [Versión 10.0.17134.112]
(c) 2018 Microsoft Corporation. Todos los derechos reservados.

C:\Users\selde>cd c:/program files/mysql/mysql server 5.7/bin

c:\Program Files\MySQL\MySQL Server 5.7\bin>mysql -u grup39 -h puigpedros.salleurl.edu
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 2380
Server version: 5.5.60-0ubuntu0.14.04.1 (Ubuntu)

Copyright (c) 2000, 2018, Oracle and/or its affiliates. All rights reserved.

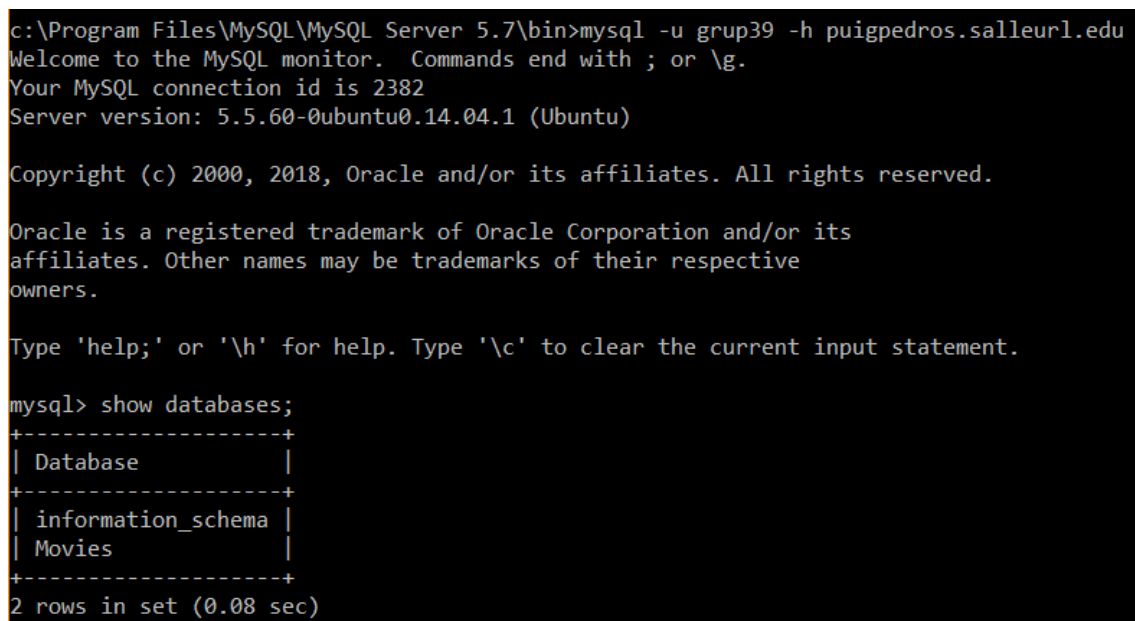
Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql>
```

Ahora que ya estamos conectados al servidor, podemos utilizar las mismas comandas que en la conexión local para encontrar información sobre la base de datos a importar. Nuestras comandas más utilizadas para esto han sido (en orden de uso):

➔ `show databases;` (para ver qué bases de datos había en el servidor remoto)



```
c:\Program Files\MySQL\MySQL Server 5.7\bin>mysql -u grup39 -h puigpedros.salleurl.edu
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 2382
Server version: 5.5.60-0ubuntu0.14.04.1 (Ubuntu)

Copyright (c) 2000, 2018, Oracle and/or its affiliates. All rights reserved.

Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql> show databases;
+-----+
| Database                |
+-----+
| information_schema      |
| Movies                  |
+-----+
2 rows in set (0.08 sec)
```

¹ En nuestro caso, 'grup39'.

→ use Movies; (para utilizar la base de datos en cuestión)

```
mysql> show databases;
+-----+
| Database |
+-----+
| information_schema |
| Movies |
+-----+
2 rows in set (0.08 sec)

mysql> use Movies;
Database changed
mysql>
```

→ show tables; (para ver qué tablas había en la base de datos a exportar –‘Movies’)

```
mysql> use Movies;
Database changed
mysql> show tables;
+-----+
| Tables_in_Movies |
+-----+
| Actor |
| Actor_movie |
| Director |
| Genre |
| Genre_Movie |
| Movie |
| Person |
+-----+
7 rows in set (0.08 sec)

mysql>
```

→ show columns from <tabla>; (muestra todas las columnas y tipos de cada atributo, en el ejemplo de la tabla ‘Movie’)

```
mysql> show columns from Movie;
+-----+-----+-----+-----+-----+-----+
| Field | Type | Null | Key | Default | Extra |
+-----+-----+-----+-----+-----+-----+
| id_movie | int(11) | NO | PRI | NULL | auto_increment |
| title | varchar(511) | YES | | NULL | |
| id_director | int(11) | YES | MUL | NULL | |
| year | int(11) | YES | | NULL | |
| duration | int(11) | YES | | NULL | |
| country | varchar(511) | YES | | NULL | |
| movie_facebook_likes | int(11) | YES | | NULL | |
| imdb_score | double | YES | | NULL | |
| gross | bigint(20) | YES | | NULL | |
| budget | bigint(20) | YES | | NULL | |
+-----+-----+-----+-----+-----+-----+
10 rows in set (0.08 sec)
```

MODO ALTERNATIVO: tras conectarnos al servidor remoto del mismo modo, una vez hemos accedido al sistema de input por mysql, utilizar el *information_schema*:

→ use information_schema; (para utilizar esta base de datos)

```
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 2387
Server version: 5.5.60-0ubuntu0.14.04.1 (Ubuntu)

Copyright (c) 2000, 2018, Oracle and/or its affiliates. All rights reserved.

Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql> use information_schema;
Reading table information for completion of table and column names
You can turn off this feature to get a quicker startup with -A

Database changed
mysql>
```

→ show tables; (para ver de qué tablas disponemos para obtener la información que necesitamos):

Entonces podemos ver la tabla COLUMNS, que puede contener información relevante; además de poder hacer una query sencilla para sacar más información:

```
Database changed
mysql> show tables;
+-----+
| Tables_in_information_schema |
+-----+
| CHARACTER_SETS               |
| COLLATIONS                   |
| COLLATION_CHARACTER_SET_APPLICABILITY |
| COLUMNS                     |
| COLUMN_PRIVILEGES            |
| ENGINES                      |
| EVENTS                       |
| FILES                        |
| GLOBAL_STATUS                |
| GLOBAL_VARIABLES             |
| KEY_COLUMN_USAGE             |
| PARAMETERS                   |
| PARTITIONS                   |
| PLUGINS                      |
| PROCESSLIST                  |
| PROFILING                    |
| REFERENTIAL_CONSTRAINTS     |
| ROUTINES                     |
| SCHEMATA                     |
| SCHEMA_PRIVILEGES            |
| SESSION_STATUS               |
| SESSION_VARIABLES            |
| STATISTICS                   |
| TABLES                      |
| TABLESPACES                 |
| TABLE_CONSTRAINTS           |
| TABLE_PRIVILEGES            |
| TRIGGERS                     |
| USER_PRIVILEGES              |
| VIEWS                        |
| INNODB_BUFFER_PAGE           |
| INNODB_TRX                   |
| INNODB_BUFFER_POOL_STATS     |
| INNODB_LOCK_WAITS            |
| INNODB_CMPMEM                |
| INNODB_CMP                   |
| INNODB_LOCKS                 |
| INNODB_CMPMEM_RESET         |
| INNODB_CMP_RESET             |
| INNODB_BUFFER_PAGE_LRU       |
+-----+
40 rows in set (0.00 sec)
```


→ select * from TABLES where TABLE_SCHEMA not like 'information_schema' group by TABLE_NAME;

```
mysql> select * from TABLES where TABLE_SCHEMA not like 'information_schema' group by TABLE_NAME;
```

TABLE_CATALOG	TABLE_SCHEMA	TABLE_NAME	TABLE_TYPE	ENGINE	VERSION	ROW_FORMAT	TABLE_ROWS	AVG_ROW_LENGTH	DATA_LENGTH	MAX_DATA_LENGTH	INDEX_LENGTH	DATA_FREE	AUTO_INCREMENT
def	Movies	Actor	BASE TABLE	InnoDB	10	Compact	6338	28	180224	0	0	4194304	
def	Movies	Actor_movie	BASE TABLE	InnoDB	10	Compact	12721	124	1589248	0	294912	4194304	
def	Movies	Director	BASE TABLE	InnoDB	10	Compact	1776	55	98304	0	0	4194304	
def	Movies	Genre	BASE TABLE	InnoDB	10	Compact	27	606	16384	0	0	4194304	
def	Movies	Genre_movie	BASE TABLE	InnoDB	10	Compact	19725	80	1589248	0	376832	4194304	
def	Movies	Movie	BASE TABLE	InnoDB	10	Compact	5155	95	491520	0	147456	4194304	
def	Movies	Person	BASE TABLE	InnoDB	10	Compact	8767	46	409600	0	0	4194304	

7 rows in set (0.01 sec)

→ También podríamos sacar más información de esa tabla COLUMNS, haciendo referencia a los atributos, como el tipo y los privilegios.

5. Proceso de importación y exportación de la base de datos.

A grandes pasos, el proceso de importación y exportación de la base Movies es el siguiente:

- Establecer una conexión con ambas bases de datos (la remota, desde donde importaremos los datos; y la local, donde los exportaremos).
- Al utilizar un sistema que utiliza dos variables (ambas del mismo tipo, un puente SSH), una para la remota y otra para la local, podemos pasar la información de una a otra sin necesidad de almacenar los datos en la aplicación. Dicho esto, abrimos las conexiones.

```
// 1. Inicialitzem les connexions
// 1.1. Crea classe que connecta amb la base remota
SSHConnector remoteCon = new SSHConnector( USR: "grup" + remoteNGroup, pass: "", db: "Movies", port: 3306, dest: "puigpedros");
// 1.2. Crea classe que connecta amb la base local
SSHConnector localCon = new SSHConnector( USR: "root", localPass, db: "Movies", port: 3306, dest: "localhost");

// 2. Connecta amb ambdues classes
connected = remoteCon.connect() && localCon.connect();
```

- [NO NECESARIO] Mostramos por pantalla las tablas que contiene dicha base de datos.

```
public void mostraTaules(SSHConnector remoteCon) {
    ResultSet remoteTables;
    // *. Mostrem les taules a importar:
    System.out.println("\nAquestes són les taules a importar de la base de dades remota:");
    remoteTables = remoteCon.selectQuery("show tables;");
    String result = "";
    try {
        while (remoteTables.next()) {
            result = remoteTables.getString( columnLabel: "Tables_in_movies");
            System.out.println("\t\t- " + result);
        }
        System.out.println("");
    } catch (SQLException e) {
        System.out.println("Error de SQL.");
    }
}
```

- Si las conexiones se han establecido correctamente, comenzamos a importar: aunque podríamos ir tabla a tabla importando y exportando reutilizando una misma variable, al no ser un requisito, hicimos una variable de tipo ResultSet para cada tabla, de forma que queda más intuitivo: importamos, con una query, toda la información (*select * from <tabla>;*) en dicha variable ResultSet.

```
// 2. Connecta amb ambdues classes
connected = remoteCon.connect() && localCon.connect();

if (connected) {
    System.out.println("\t-> Hem connectat amb ambdues bases de dades, ara passarem a importar la informació!");
    remoteCon.mostraTaules(remoteCon);

    // 3. Preparem tots els ResultSet's (tot i que podriem optimitzar-ho, així es mes intuitiu)
    ResultSet person = remoteCon.selectQuery("select * from Person;");
    ResultSet director = remoteCon.selectQuery("select * from Director;");
    ResultSet actor = remoteCon.selectQuery("select * from Actor;");
    ResultSet actor_movie = remoteCon.selectQuery("select * from Actor_movie;");
    ResultSet movie = remoteCon.selectQuery("select * from Movie;");
    ResultSet genre_movie = remoteCon.selectQuery("select * from Genre_Movie;");
    ResultSet genre = remoteCon.selectQuery("select * from Genre;");
}
```

- Ahora, para cada tabla:
 - o Pasamos la información del ResultSet a la tabla en cuestión con un *insert into <tabla> values (<datos_del_resultset>);* [el ejemplo es solo de la tabla **Person** i **Movie**]

```
// 4.1. Person
try {
    System.out.print("\t-> Comencem a exportar a la taula 'Person'...\t");
    while (person.next()) {
        localCon.insertQuery("insert into person values (" +
            person.getInt( columnLabel: "id_person") + ", " +
            "" + person.getString( columnLabel: "name").replace( target: "'", replacement: " ") + ", " +
            person.getInt( columnLabel: "facebook_likes") + ");");
    }
    System.out.println("'Person' exportada amb éxit!");
} catch (SQLException ex) {
    System.out.println("\t-> Problema de SQL!");
}
```

```
// 4.5. Movie
try {
    System.out.print("\t-> Comencem a exportar a la taula 'Movie'...\t");
    while (movie.next()) {
        localCon.insertQuery("insert into movie values (" +
            movie.getInt( columnLabel: "id_movie") + ", " +
            "" + movie.getString( columnLabel: "title").replace( target: "'", replacement: " ") + ", " +
            movie.getInt( columnLabel: "id_director") + ", " +
            movie.getInt( columnLabel: "year") + ", " +
            movie.getInt( columnLabel: "duration") + ", " +
            "" + movie.getString( columnLabel: "country").replace( target: "'", replacement: " ") + ", " +
            movie.getInt( columnLabel: "movie_facebook_likes") + ", " +
            movie.getDouble( columnLabel: "imdb_score") + ", " +
            movie.getLong( columnLabel: "gross") + ", " +
            movie.getLong( columnLabel: "budget") + ");");
    }
    System.out.println("'Movie' exportada amb éxit!");
} catch (SQLException ex) {
    System.out.println("\t-> Problema de SQL!");
}
```

- Desconectamos ambas bases de datos.

```
// x. Tanquem les sessions i sortim del sistema:
System.out.println("\t-> Sortim de les sessions obertes de cada connexió...");
remoteCon.disconnect();
System.out.println("\t\t o Hem sortit de la connexió remota!");
localCon.disconnect();
System.out.println("\t\t o Hem sortit de la connexió local!");
System.out.println("\n >> Sortint del sistema... <<");
```

RESULTAT:

```
Escriu el nombre de grup assignat (només el número!): 
Escriu la contrassenya per connectar amb la base local: 
Connexió a base de dades jdbc:mysql://puigpedros:3306/Movies ... Ok
Connexió a base de dades jdbc:mysql://localhost:3306/Movies ... Ok
-> Hem connectat amb ambdues bases de dades, ara passarem a importar la informació!

Aquestes són les taules a importar de la base de dades remota:
- Actor
- Actor_movie
- Director
- Genre
- Genre_Movie
- Movie
- Person

-> Comencem a exportar a la taula 'Person'... 'Person' exportada amb éxit!
-> Comencem a exportar a la taula 'Director'... 'Director' exportada amb éxit!
-> Comencem a exportar a la taula 'Actor'... 'Actor' exportada amb éxit!
-> Comencem a exportar a la taula 'Actor_Movie'... 'Actor_Movie' exportada amb éxit!
-> Comencem a exportar a la taula 'Movie'... 'Movie' exportada amb éxit!
-> Comencem a exportar a la taula 'Movie_Genre'... 'Movie_Genre' exportada amb éxit!
-> Comencem a exportar a la taula 'Genre'... 'Genre' exportada amb éxit!

-> Sortim de les sessions obertes de cada connexió...
    o Hem sortit de la connexió remota!
    o Hem sortit de la connexió local!

>> Sortint del sistema... <<

Process finished with exit code 0
```

6. Dedicación

Aproximadamente hemos dedicado toda la semana antes de la entrega, que llega a ser unas 15h.

7. Conclusiones

Las conclusiones que hemos podido llegar de momento con la fase 1 de esta práctica 2 es que ya sabemos cómo conectarnos desde una interfaz en java a un servidor donde luego, en este mismo servidor, se encuentra la base de datos en la cual estamos interesados. Ahora tenemos muchos más conocimientos en relación con el lenguaje de Java y las comandas que se utilizan para conectarse y obtener información de las tablas de la base de datos, además de entender

cómo podemos utilizar las comandas que conocemos, tanto de bases de datos como las de la consola de comandos del sistema para exportar e importar información.

8. Bibliografía

<https://es.wikipedia.org/wiki/Wikipedia:Portada>

<https://www.techopedia.com/definition/24436/online-transaction-processing-oltp>

<https://www.powerdata.es/data-warehouse>