

## Práctica 2: Importación de datos con Sqoop

Para trabajar con la consola de la base de datos mySQL hay que situarse en **node1**:

```
ssh root@node1
mysql -u root -p
contraseña: hadoop123
```

### Consulta de bases de datos, tablas y campos

```
mysql> show databases;
mysql> use movielens;
mysql> show tables;
mysql> select * from movie limit 10;
```

```
+-----+-----+-----+-----+
| id | name                               | year |
+-----+-----+-----+-----+
| 1 | Toy Story                         | 1995 |
| 2 | Jumanji                          | 1995 |
| 3 | Grumpier Old Men                  | 1995 |
| 4 | Waiting to Exhale                 | 1995 |
| 5 | Father of the Bride Part II      | 1995 |
| 6 | Heat                              | 1995 |
| 7 | Sabrina                           | 1995 |
| 8 | Tom and Huck                      | 1995 |
| 9 | Sudden Death                     | 1995 |
| 10 | GoldenEye                        | 1995 |
+-----+-----+-----+-----+
```

10 rows in set (0.02 sec)

```
mysql> select count(*) from movie;
```

```
+-----+
| count(*) |
+-----+
|      3881 |
+-----+
```

1 row in set (0.00 sec)

```
mysql> describe movie;
```

```
+-----+-----+-----+-----+-----+-----+
| Field | Type          | Null | Key | Default | Extra |
+-----+-----+-----+-----+-----+-----+
| id    | int(11)       | NO   | PRI | 0        |       |
| name  | char(75)      | YES  |     | NULL     |       |
| year  | smallint(6)   | YES  |     | NULL     |       |
+-----+-----+-----+-----+-----+-----+
```

3 rows in set (0.44 sec)

Salimos de la consola de mySQL con **quit**;

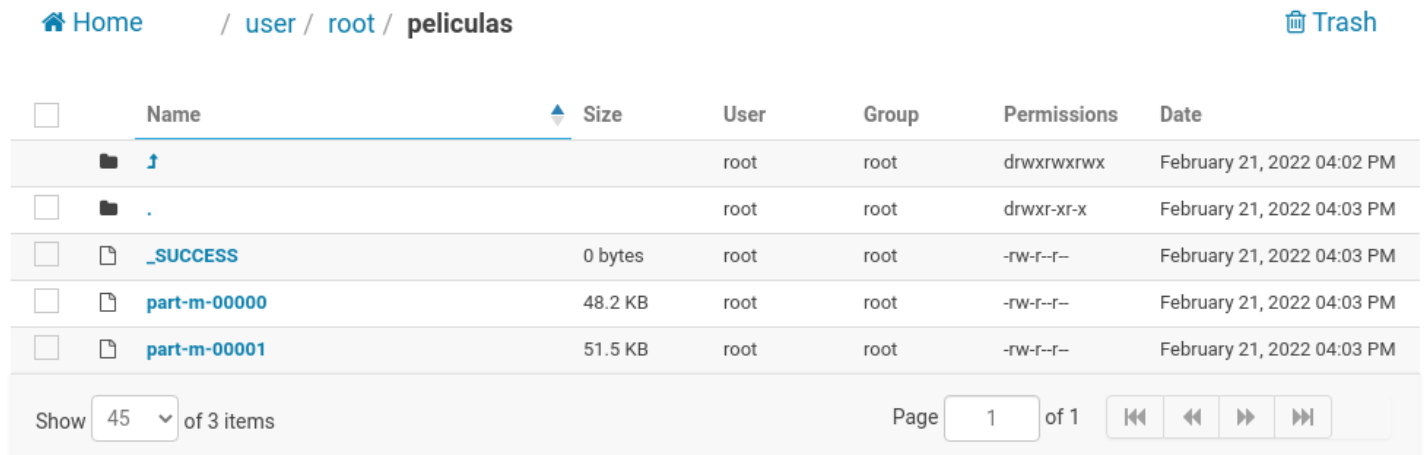
## Importar la tabla con Sqoop






Desde node1 ejecutamos la siguiente sentencia:

```
sqoop import --connect jdbc:mysql://node1/movielens --username root \
--password hadoop123 --target-dir /user/root/peliculas --table movie \
--fields-terminated-by ',' -m 1 --driver com.mysql.jdbc.Driver
```

Aparece como resultado en la última línea:

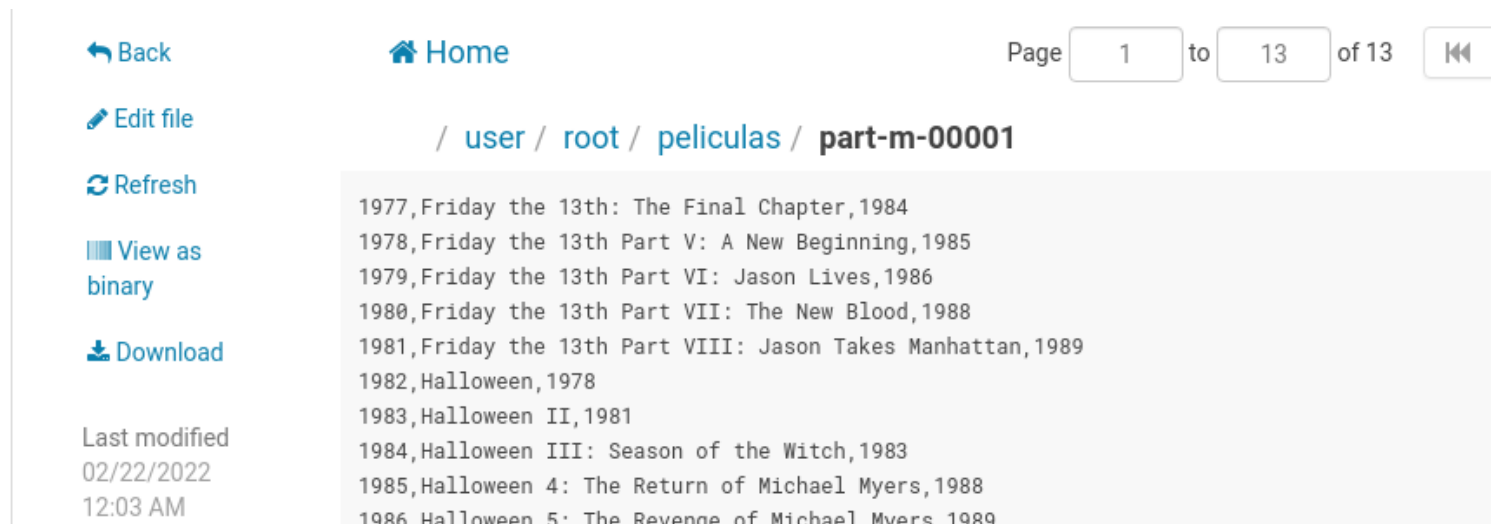
INFO mapreduce.ImportJobBase: Retrieved 3881 records.



<input type="checkbox"/>	Name	Size	User	Group	Permissions	Date
<input type="checkbox"/>			root	root	drwxrwxrwx	February 21, 2022 04:02 PM
<input type="checkbox"/>			root	root	drwxr-xr-x	February 21, 2022 04:03 PM
<input type="checkbox"/>	 <b>_SUCCESS</b>	0 bytes	root	root	-rw-r--r--	February 21, 2022 04:03 PM
<input type="checkbox"/>	 <b>part-m-00000</b>	48.2 KB	root	root	-rw-r--r--	February 21, 2022 04:03 PM
<input type="checkbox"/>	 <b>part-m-00001</b>	51.5 KB	root	root	-rw-r--r--	February 21, 2022 04:03 PM

Show 45 of 3 items Page 1 of 1

Figure 1: Directorio peliculas tras la importación



	Content
1777, Friday the 13th: The Final Chapter, 1984	
1978, Friday the 13th Part V: A New Beginning, 1985	
1979, Friday the 13th Part VI: Jason Lives, 1986	
1980, Friday the 13th Part VII: The New Blood, 1988	
1981, Friday the 13th Part VIII: Jason Takes Manhattan, 1989	
1982, Halloween, 1978	
1983, Halloween II, 1981	
1984, Halloween III: Season of the Witch, 1983	
1985, Halloween 4: The Return of Michael Myers, 1988	
1986, Halloween 5: The Revenge of Michael Myers, 1989	

Figure 2: Contenido de los archivos importados

Los archivos obtenidos son de texto delimitados por comas en este caso.

Estando ya en HDFS pueden ser utilizados por el resto de aplicaciones de tratamiento de datos.

**Recoge en un pantallazo el contenido de *peliculas* como muestra de que has hecho la práctica. Recuerda que se debe ver tu nombre en la imagen.**