



INTRODUCCIÓN

Consultas con HIVE

- Trabajaremos con los dos archivos que vienen con la tarea (**movies.dat** -*películas con sus géneros*- y **ratings.dat** -*con un millón de votaciones*-) con la siguiente estructura:

movies.dat - id, película y sus géneros

```
SCR5 / ANTONIO / DESKTOP / movies.dat
1*Toy Story (1995)*Animation|Children's|Comedy
2*Jumanji (1995)*Adventure|Children's|Fantasy
3*Grumpier Old Men (1995)*Comedy|Romance
4*Waiting to Exhale (1995)*Comedy|Drama
5*Father of the Bride Part II (1995)*Comedy
6*Heat (1995)*Action|Crime|Thriller
7*Sabrina (1995)*Comedy|Romance
```

rating.dat - id,usuario,voto,timestamp

```
1|1193|5|978300760
1|661|3|978302109
1|914|3|978301968
1|3408|4|978300275
1|2355|5|978824291
1|1197|3|978302268
```

- Para cada una de las consultas, mostrar su código y una captura con la salida de las mismas.

CONTENIDO

APARTADO A

- Crea las tablas HIVE adecuadas para cargar los datos de cada uno de los archivos. Para mayor eficiencia, crearemos ambas tablas con 5 *buckets* sobre el **id** de la película. Tener especial cuidado en el campo que almacene el género de las películas.
- Mostrar las consultas de creación y carga de las tablas con una captura que muestre los datos cargados.
- Mostrar una captura de las tablas en el *warehouse* de HIVE donde se vean los *buckets*.



CONTENIDO

APARTADO B

1. Mediante una consulta en HIVE, encontrar las cinco películas (código, título y media de votos) mejor valoradas, que hayan sido votadas al menos por 10 usuarios.

CONTENIDO

APARTADO C

1. Encontrar las cinco películas más antiguas con una valoración media por encima de 4 puntos.

CONTENIDO

APARTADO D

Investigando las funciones que nos ofrece HIVE para el manejo de cadenas....

1. Muestra los cinco años en que se editaron más películas indicando el número de ellas para cada año.

CONTENIDO

APARTADO E

Consultando el enlace de abajo:

<https://www.bigdatainrealworld.com/what-is-the-difference-between-explode-and-lateral-view-explode-in-hive/>

1. Muestra los diez géneros de películas más frecuentes.
2. Partiendo de la consulta del Apartado B de arriba. ¿Qué géneros son los más frecuentes en dichas películas?