|  |
| --- |
| Máster en Big Data |
| Tecnologías de Almacenamiento |
| 11. Hands-On: Hive |

Presentado por: Jose David Angulo y Albert Ripoll

Índice

[1. Introducción3](#_Toc508717060)

[2. Entorno3](#_Toc508717061)

[3. Creación de tablas3](#_Toc508717062)

[4. Consultas con Hive5](#_Toc508717063)

# Introducción

El objetivo de este Hands-On es familiarizarse con la utilización de Hive, tanto en la creación de tablas como en la realización de consultas.

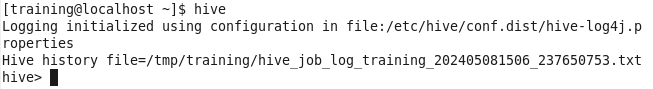
# Entorno

Para este Handos On, utilizaremos la máquina virtual desplegada en Hands-On anteriores llamada Developer\_Hadoop y todo será ejecutado vía Shell.

# Creación de tablas

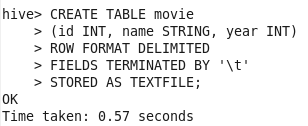
1. Ejecutar el Hive Shell

Para iniciar el Shell se hace con el comando hive

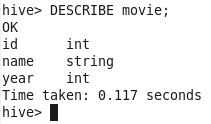


1. Crear la tabla movie basada en el archivo “movie” importado anteriormente

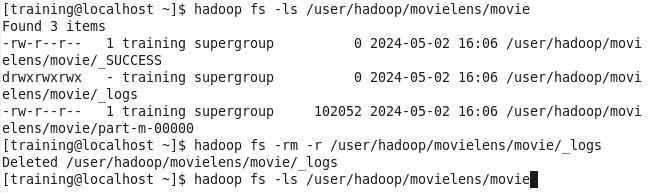
CREATE TABLE movie                        
    (id INT, name STRING, year INT)  
    ROW FORMAT DELIMITED                      
    FIELDS TERMINATED BY ';'  
    STORED AS TEXTFILE

;

DESCRIBE movie;

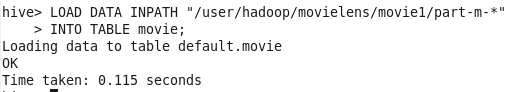


hadoop fs -rm -r /user/hadoop/movielens/movie/\_logs

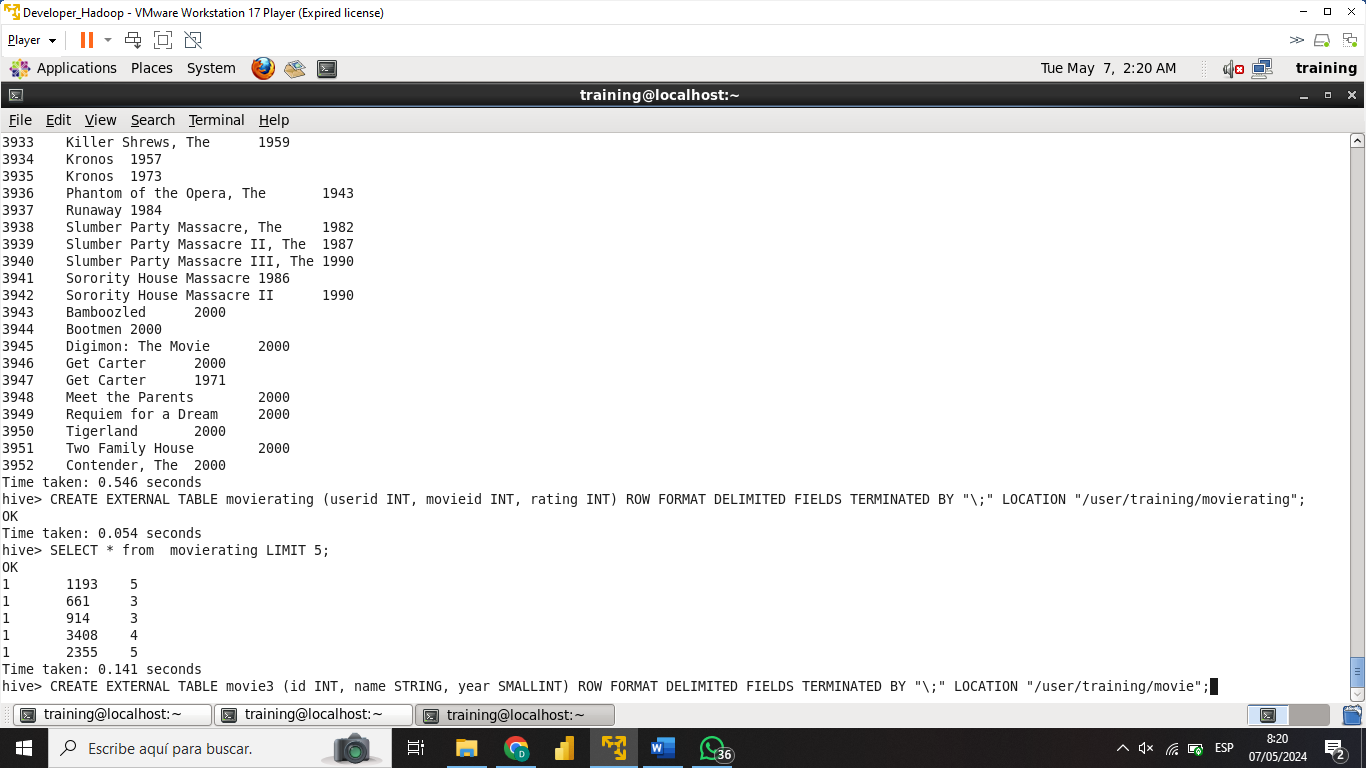


LOAD DATA INPATH "/user/hadoop/movielens/movie1/part-m-\*"

> INTO TABLE movie;



1. Crear la tabla movierating basada en el archivo “movierating” importado anteriormente

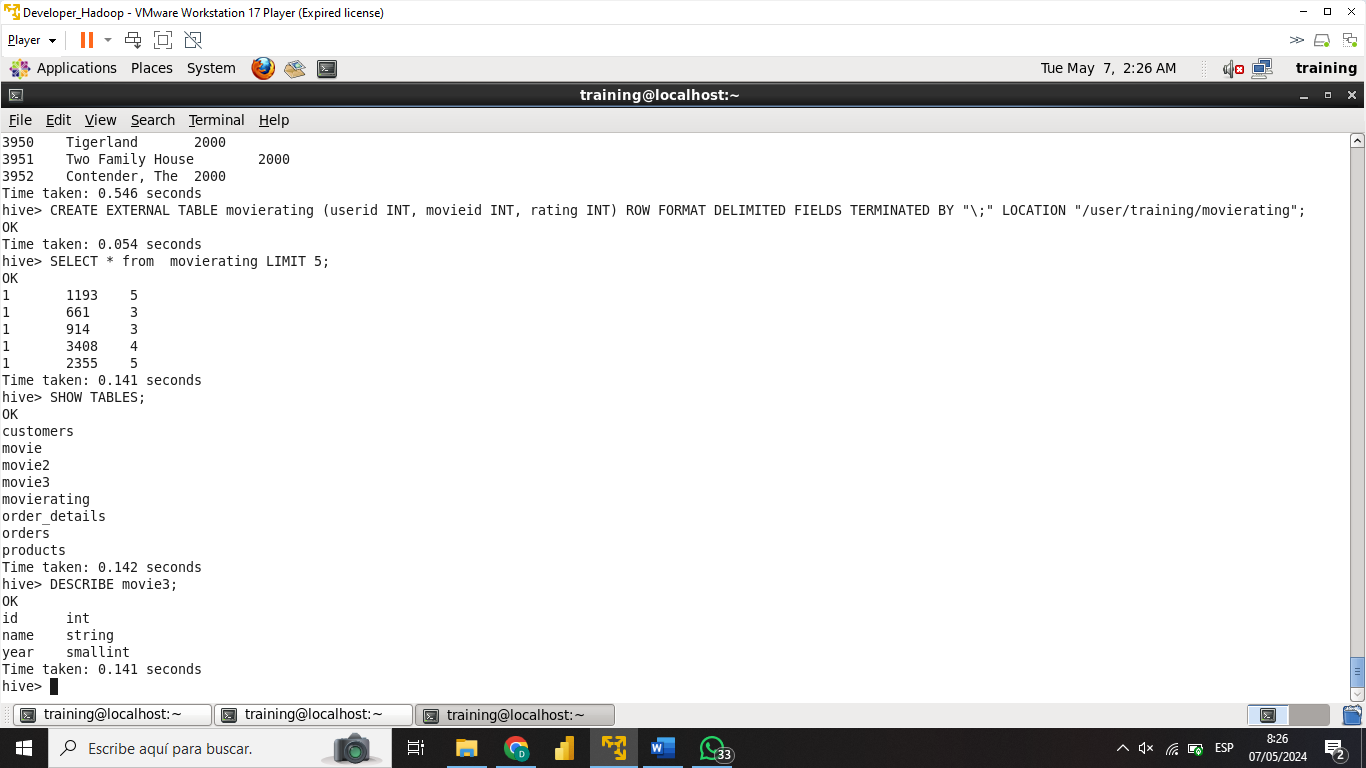


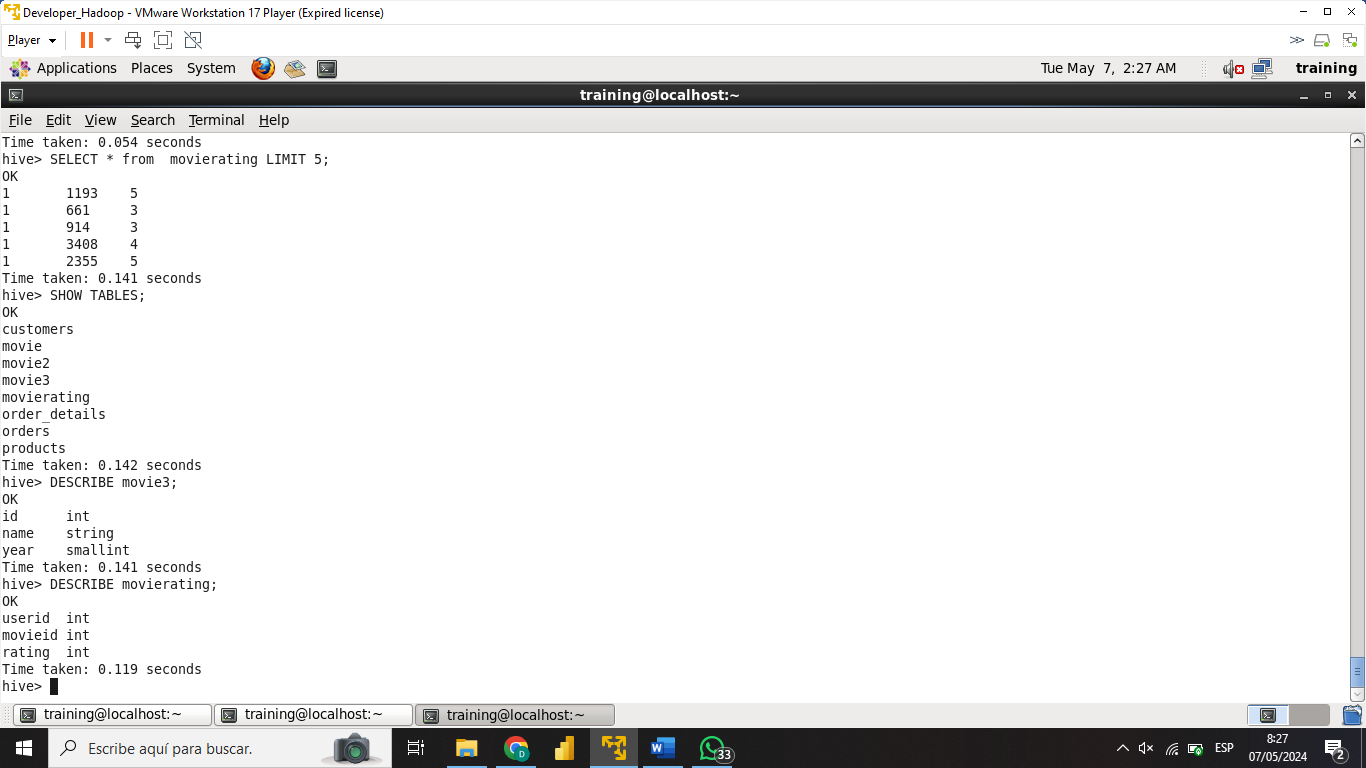
1. Listar todas las tablas de Hive

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

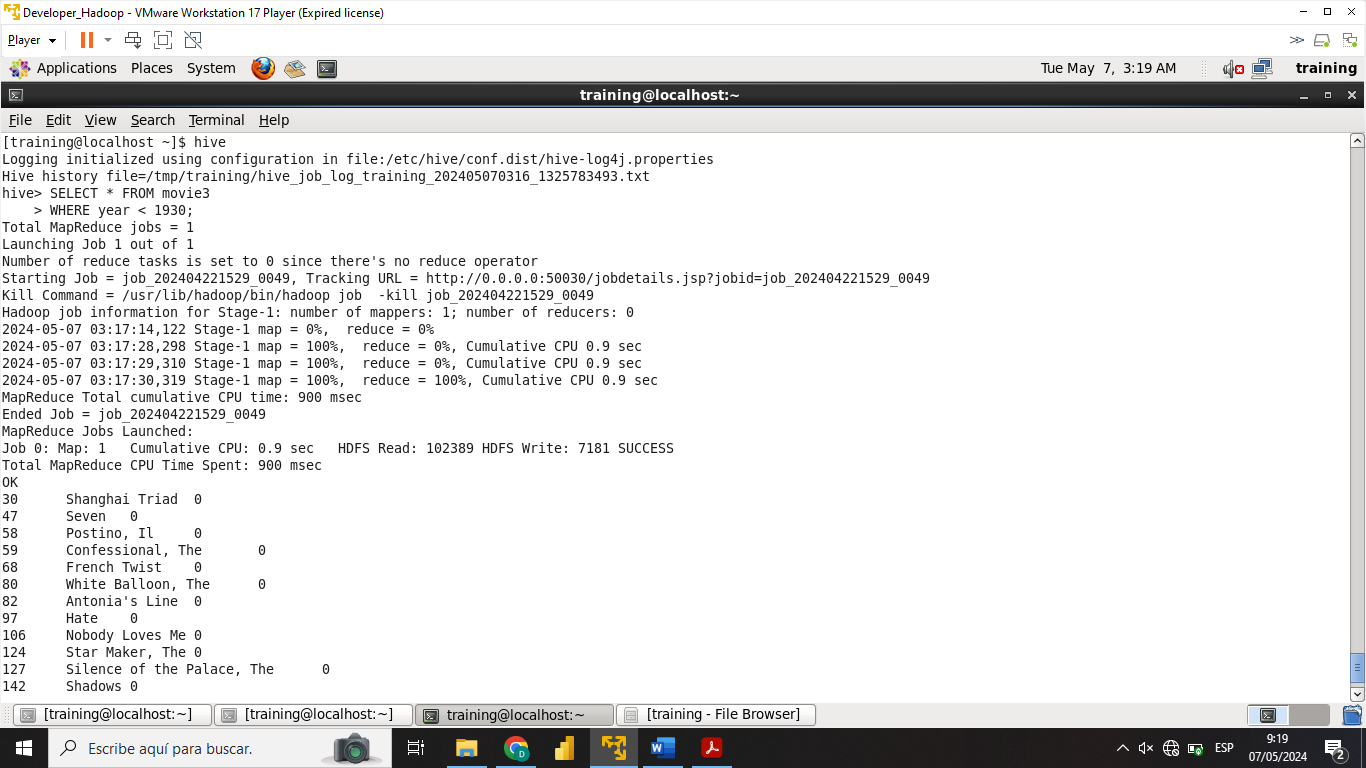
1. Ver la metainformación de las tablas movie y movierating



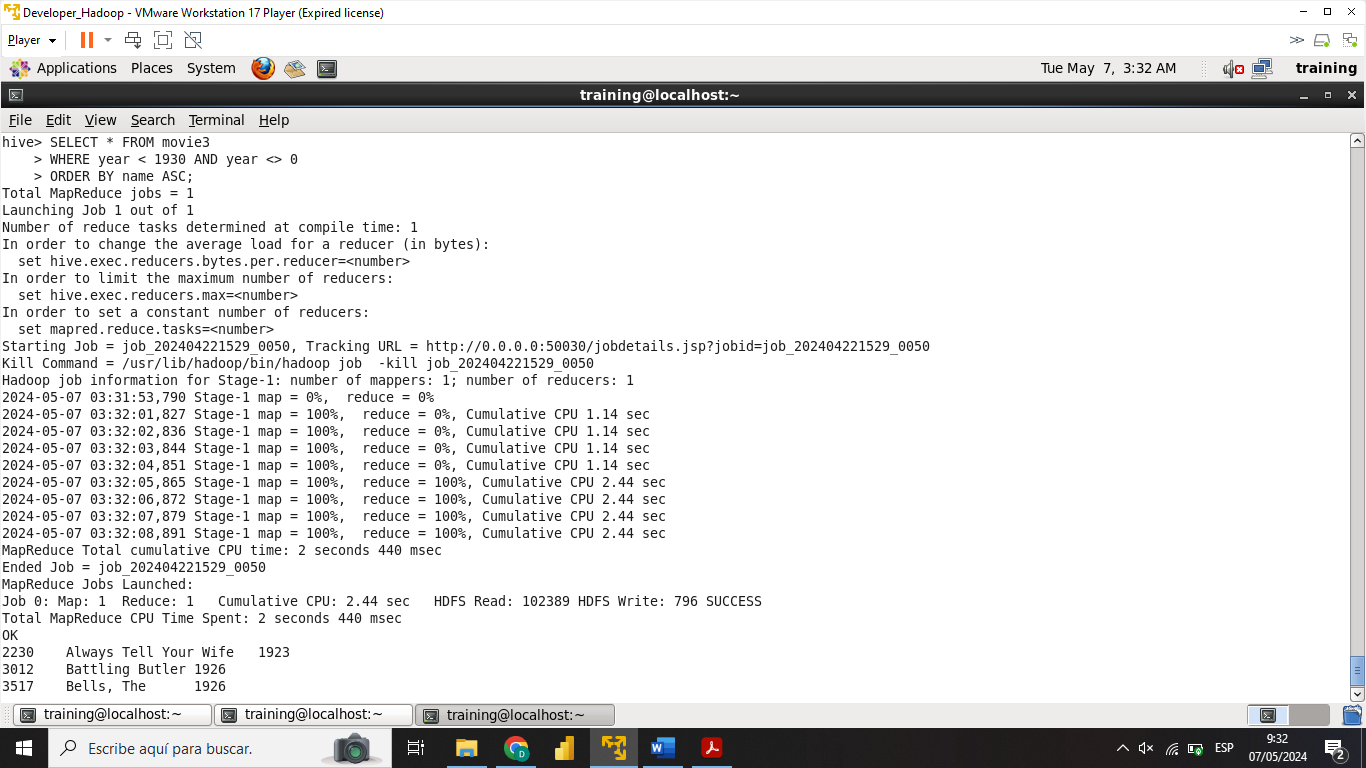


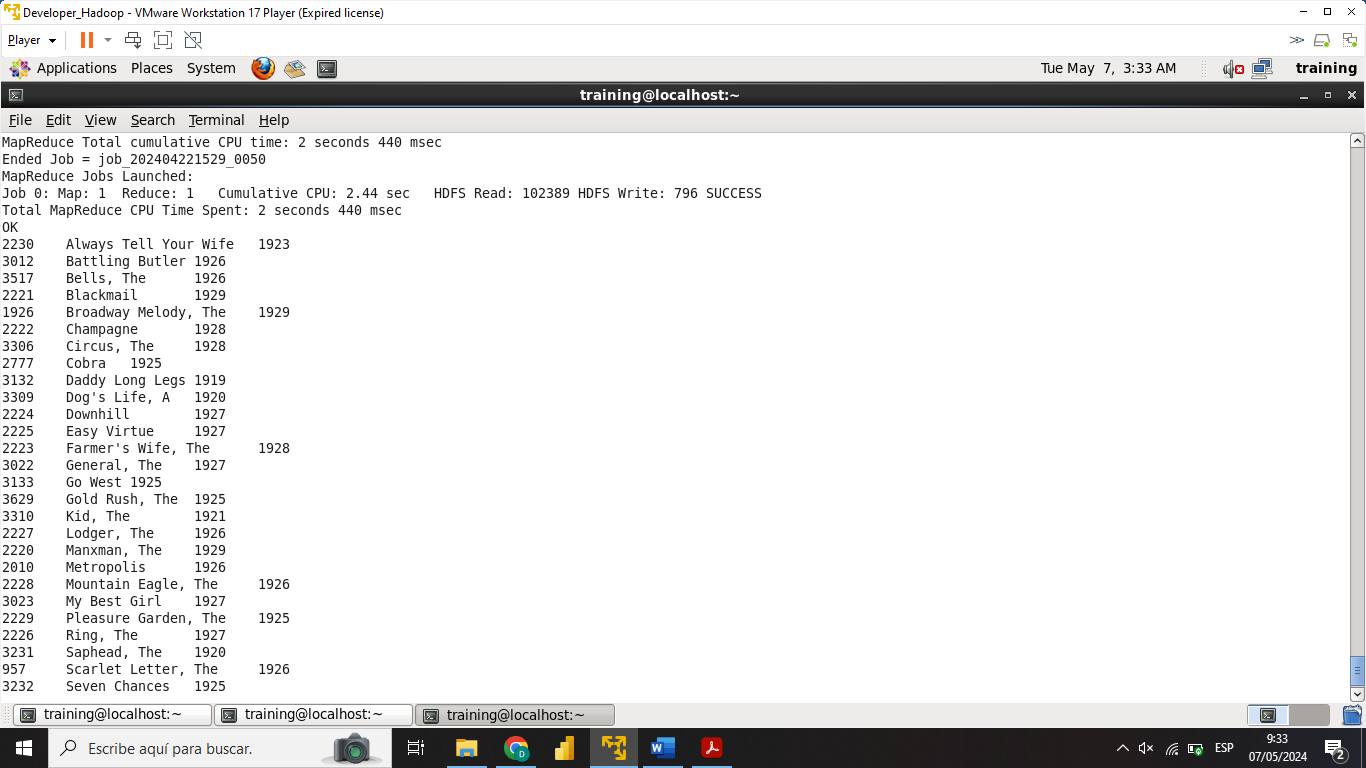
# Consultas con Hive

1. Listar todas las películas lanzadas antes de 1930

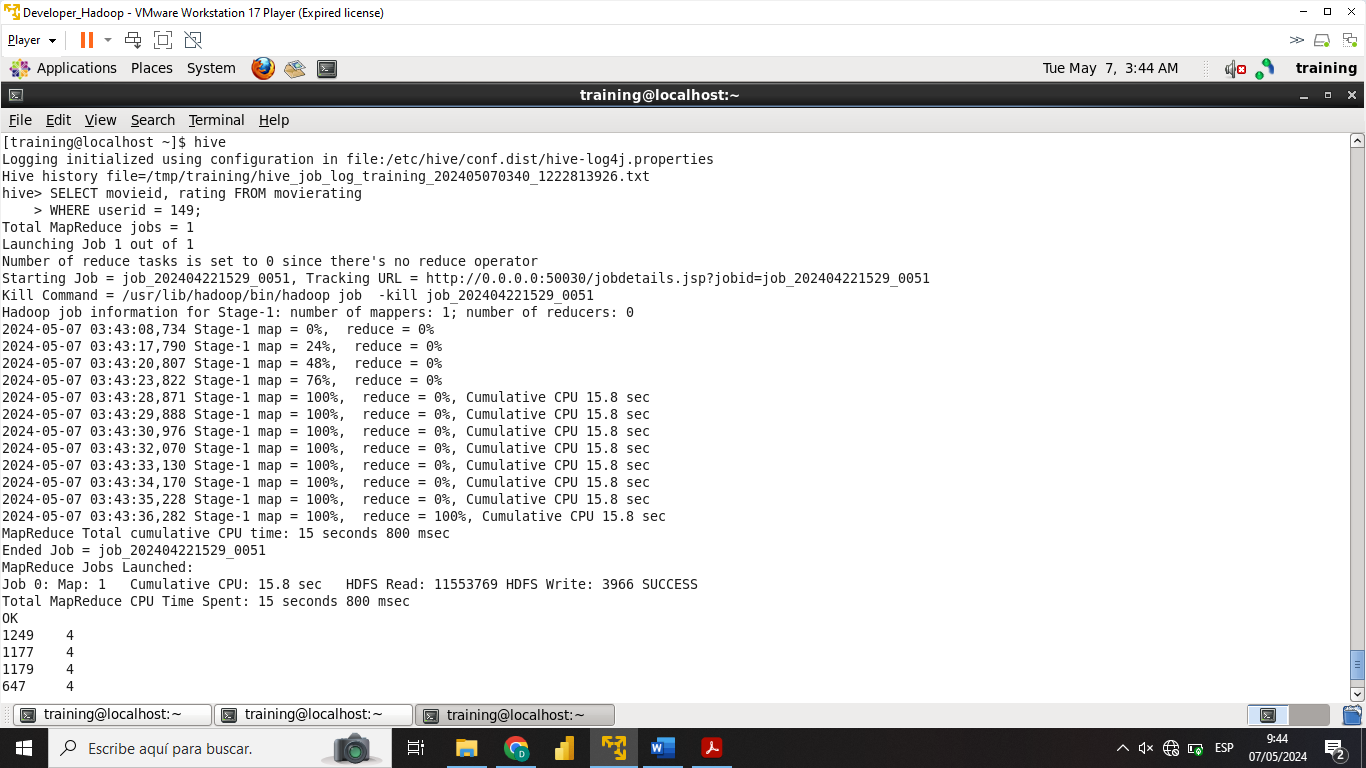


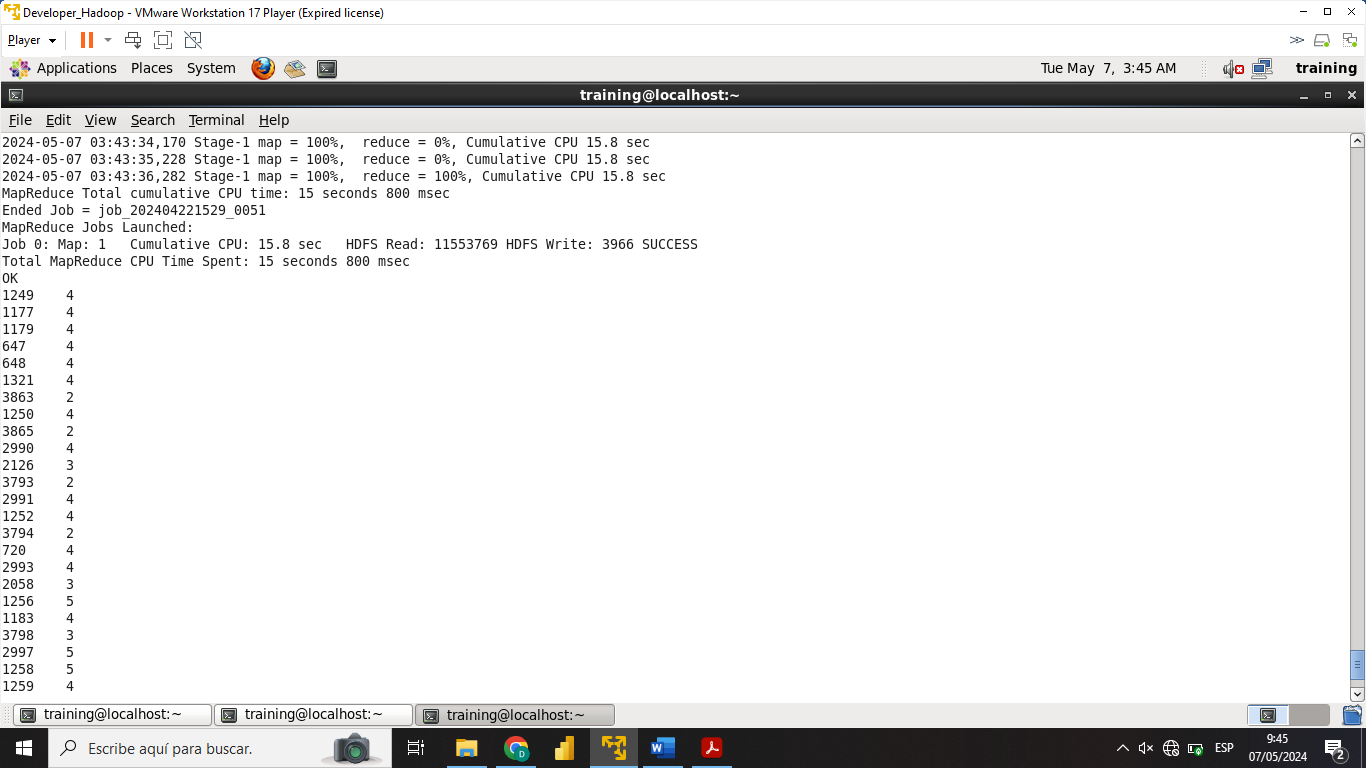
1. Listar todas las películas lanzadas antes de 1930. Descarta todas aquellas que no tengan año conocido (year=0) y ordenalas por nombre





1. Selecciona todas las películas valoradas por el usuario con id = 149 (Muestra solamente los campos relativos al id de película y su valoración)





1. Utiliza información de las dos tablas, por ejemplo, incluyendo el nombre de la película en la lista generada en el apartado anterior

Interfaz de usuario gráfica, Texto

Descripción generada automáticamente

Interfaz de usuario gráfica, Texto

Descripción generada automáticamente

1. Calcula el promedio con el que el usuario 149 califica las pel·lícules

Interfaz de usuario gráfica, Texto

Descripción generada automáticamente

Interfaz de usuario gráfica, Texto

Descripción generada automáticamenteInterfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. Lista cada usuario que ha valorado películas, el número de películas que ha valorado y el promedio de valoración que ha proporcionado

Interfaz de usuario gráfica, Texto

Descripción generada automáticamente

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. Inserta toda la información del apartado anterior en una nueva tabla llamada “userrating”.

Interfaz de usuario gráfica, Texto

Descripción generada automáticamente

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. Exporta la tabla a MySQL con Sqoop.

sqoop export \

--connect jdbc:mysql://localhost:3306/Albert \

--username training \

--password training \

--table mobie\_table \

--export-dir /user/hive/movielens/movie\_table \

--input-fields-terminated-by '\t' \

-m 1