**CSV File:** ep4\_activity2.csv

hive> show databases;

OK

default

Time taken: 0.298 seconds, Fetched: 1 row(s)

hive> create database episodeIV;

OK

Time taken: 0.062 seconds

hive> use episodeIV;

OK

Time taken: 0.149 seconds

hive> create table epV\_activity2

> (character STRING,dialog STRING)

> ROW FORMAT DELIMITED FIELDS TERMINATED BY ','

> TBLPROPERTIES ("skip.header.line.count"="1");

OK

Time taken: 0.129 seconds

hive> DESCRIBE epV\_activity2;

OK

character string

dialog string

Time taken: 0.105 seconds, Fetched: 2 row(s)

hive> load data local inpath

> '/ep4\_activity2.csv'

> into table epV\_activity2;

Loading data to table episodeiv.epv\_activity2

OK

Time taken: 1.858 seconds

**Activity 2**

hive> SELECT COUNT(\*) FROM epV\_activity2 WHERE INSTR(character,'Luke')>0;

Query ID = root\_20211130075135\_29286a1d-4781-47fd-9de3-367694259a1a

Total jobs = 1

Launching Job 1 out of 1

Number of reduce tasks determined at compile time: 1

In order to change the average load for a reducer (in bytes):

set hive.exec.reducers.bytes.per.reducer=<number>

In order to limit the maximum number of reducers:

set hive.exec.reducers.max=<number>

In order to set a constant number of reducers:

set mapreduce.job.reduces=<number>

Starting Job = job\_1638242045094\_0011, Tracking URL = http://4c77d6c6ae20:8088/proxy/application\_1638242045094\_0011/

Kill Command = /usr/local/hadoop/bin/mapred job -kill job\_1638242045094\_0011

Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1

2021-11-30 07:52:09,422 Stage-1 map = 0%, reduce = 0%

2021-11-30 07:52:33,491 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 14.02 sec

2021-11-30 07:52:43,435 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 19.47 sec

MapReduce Total cumulative CPU time: 19 seconds 470 msec

Ended Job = job\_1638242045094\_0011

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 19.47 sec HDFS Read: 64229 HDFS Write: 101 SUCCESS

Total MapReduce CPU Time Spent: 19 seconds 470 msec

OK

56

Time taken: 71.47 seconds, Fetched: 1 row(s)

Solution:

