**Episode 4**

#create table

hive> CREATE TABLE episodesIV\_counts

> (character STRING, dialouges STRING)

> ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t'

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

#load data

hive> LOAD DATA LOCAL INPATH

> '/shrikant/episodeIV\_dialouges.txt'

> INTO TABLE episodesIV\_counts;

#Exports to HDFS directory

hive> INSERT OVERWRITE DIRECTORY '/user/root/output/lukeLines'

> SELECT dialouges FROM episodesIV\_counts where dialouges rlike '.\*(Luke).\*';

#check results in HDFS

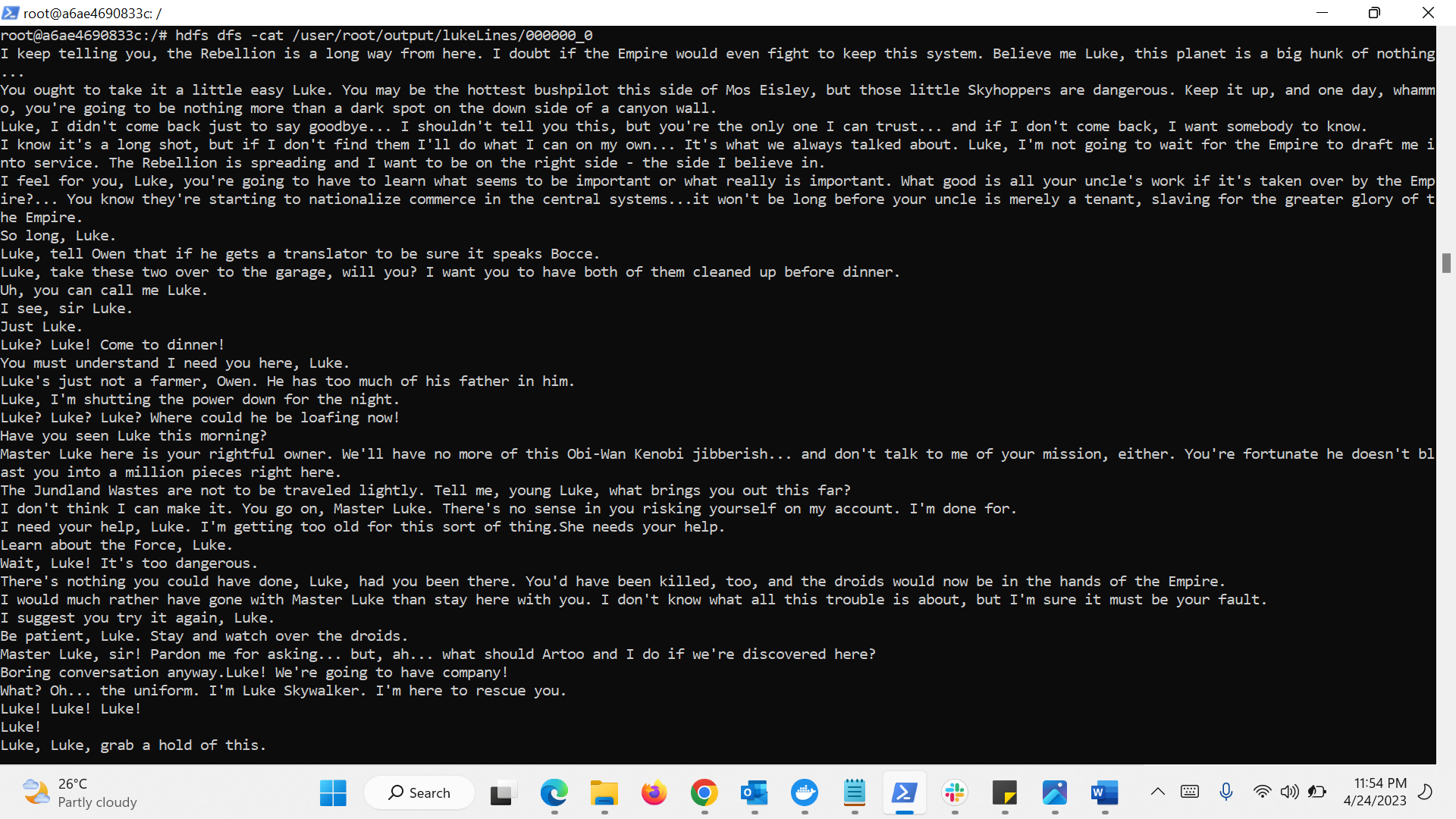
dfs -cat /user/root/output/lukeLines/000000\_0

#check results in local

hdfs dfs -get /user/root/output/lukeLines/000000\_0 /episodeIV/lukeLines

cat /episodeIV/lukeLines /000000\_0

Episode 4 line results with Luke HDFS



Episode 4 line results with Luke local

