

# Pig Latin script for episodeIV\_dialouges.txt:-

-- Load input file from HDFS

inputFile = LOAD 'hdfs:///user/dpanyang/episodeIV\_dialouges.txt' USING PigStorage('\t') AS (name:chararray, line:chararray);

-- Group data by name

grpd = GROUP inputFile BY name;

-- Count the occurence of each word (Reduce)

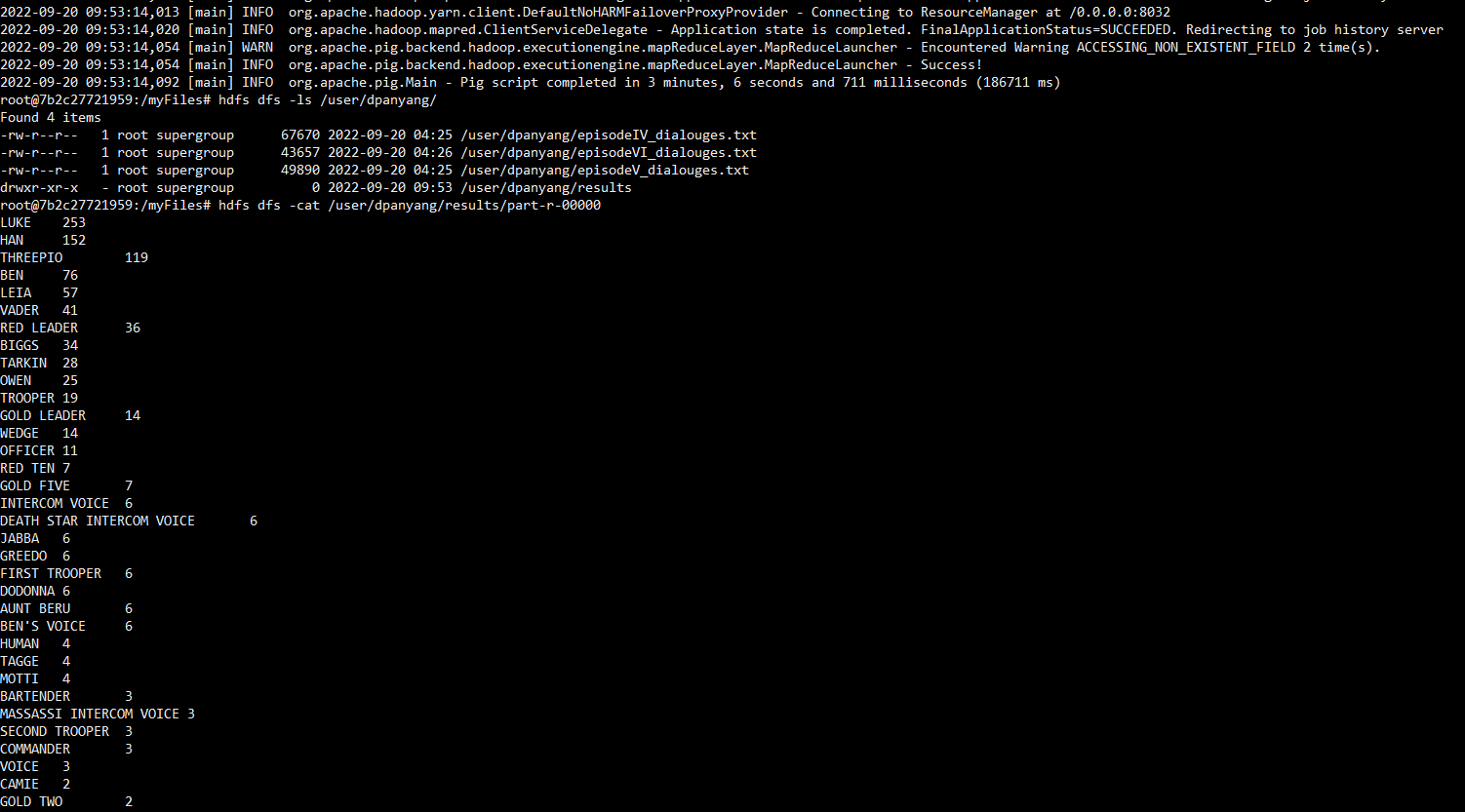
cntd = FOREACH grpd GENERATE group, COUNT(inputFile.name) AS cnt;

--Sort data

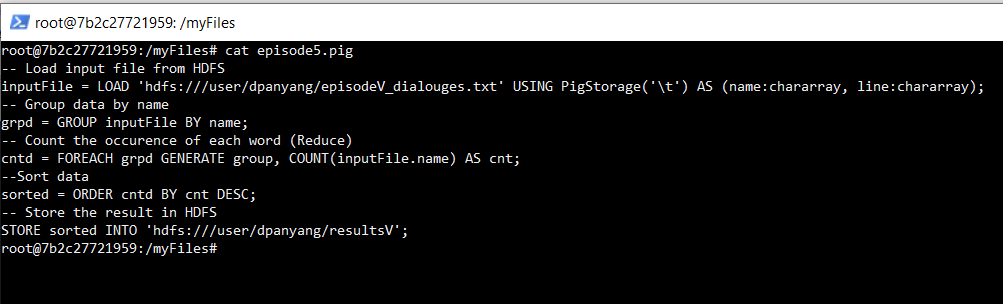
sorted = ORDER cntd BY cnt DESC;

-- Store the result in HDFS

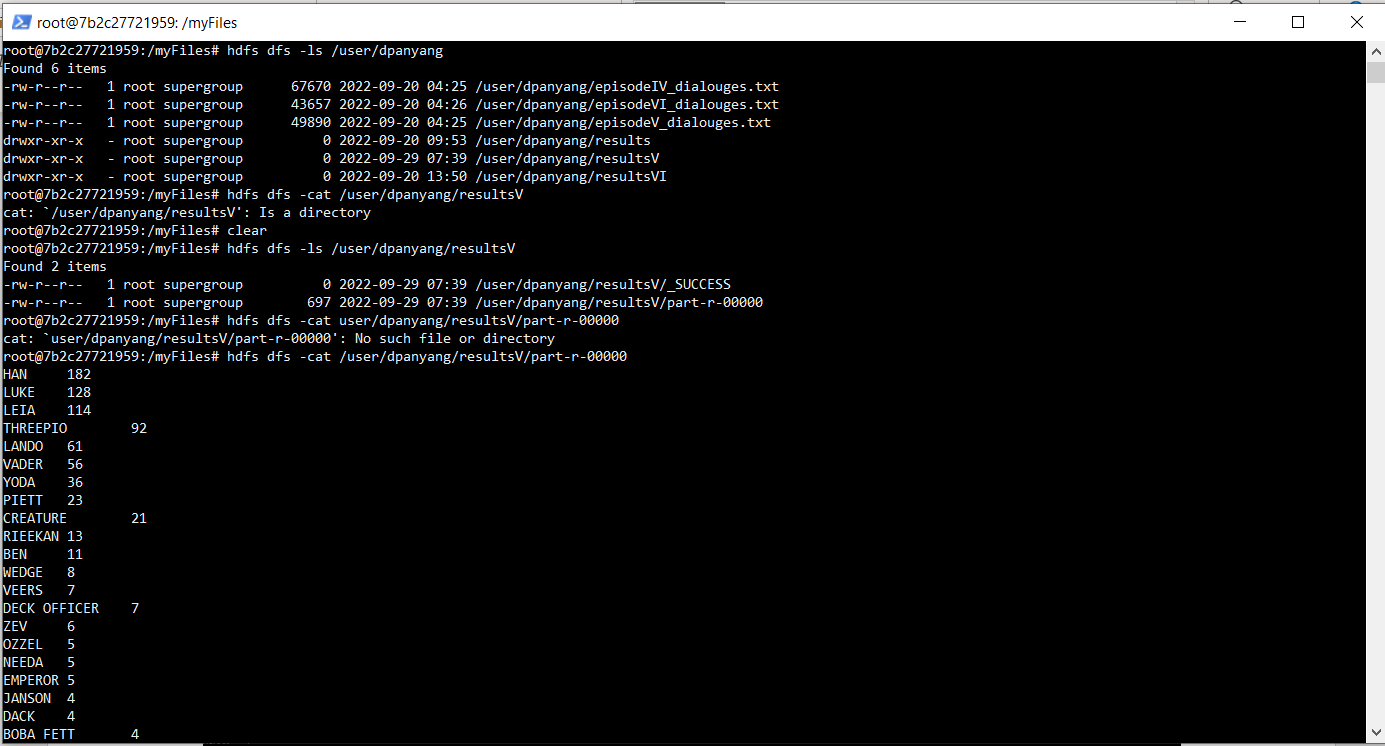
STORE sorted INTO 'hdfs:///user/dpanyang/results';



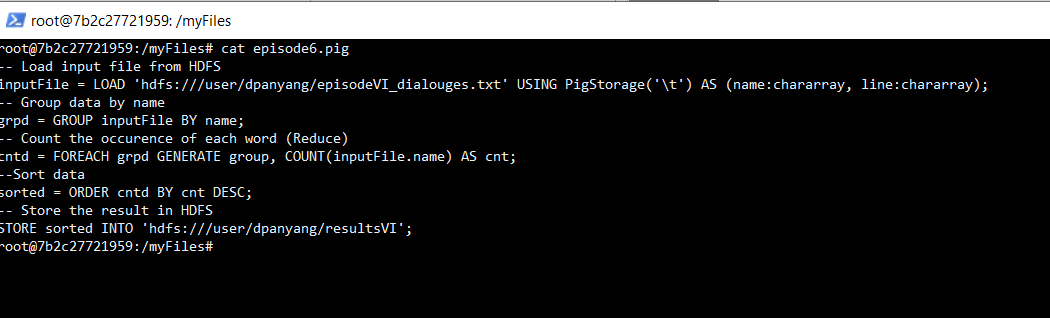
# Pig Latin script for episodeV\_dialouges.txt:-



Results after execute program :



# Pig Latin script for episodeVI\_dialouges.txt:-



Result after execute pig script :

