**activity1\_epV.pig**

inputFile = LOAD 'hdfs:///user/root/Project\_Activity1/episodeV\_dialouges.txt' AS (line:chararray);

words = FOREACH inputFile GENERATE FLATTEN(TOKENIZE(line)) AS word;

grouped = GROUP words BY word;

wordcount = FOREACH grouped GENERATE group, COUNT(words);

STORE wordcount INTO 'hdfs:///user/root/results2';

**pig command result:**

HadoopVersion PigVersion UserId StartedAt FinishedAt Features

3.3.1 0.17.0 root 2021-11-30 04:40:56 2021-11-30 04:43:41 GROUP\_BY

Success!

Job Stats (time in seconds):

JobId Maps Reduces MaxMapTime MinMapTime AvgMapTime MedianMapTime MaxReduceTime MinReduceTime AvgReduceTime MedianReducetime Alias Feature Outputs

job\_1638242045094\_0005 1 1 n/a n/a n/a n/a n/a n/a n/a n/a grouped,inputFile,wordcount,words GROUP\_BY,COMBINER hdfs:///user/root/results2,

Input(s):

Successfully read 0 records from: "hdfs:///user/root/Project\_Activity1/episodeV\_dialouges.txt"

Output(s):

Successfully stored 0 records in: "hdfs:///user/root/results2"

