Word count using Pig as below: (Ran as part of pig command sheet as well)

Grunt> lines = LOAD '/user/hadoop/HDFS\_File.txt' AS (line:chararray);

Grunt> words = FOREACH lines GENERATE FLATTEN(TOKENIZE(line)) as word;

Grunt> grouped = GROUP words BY word;

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

Grunt> **DUMP wordcount;**