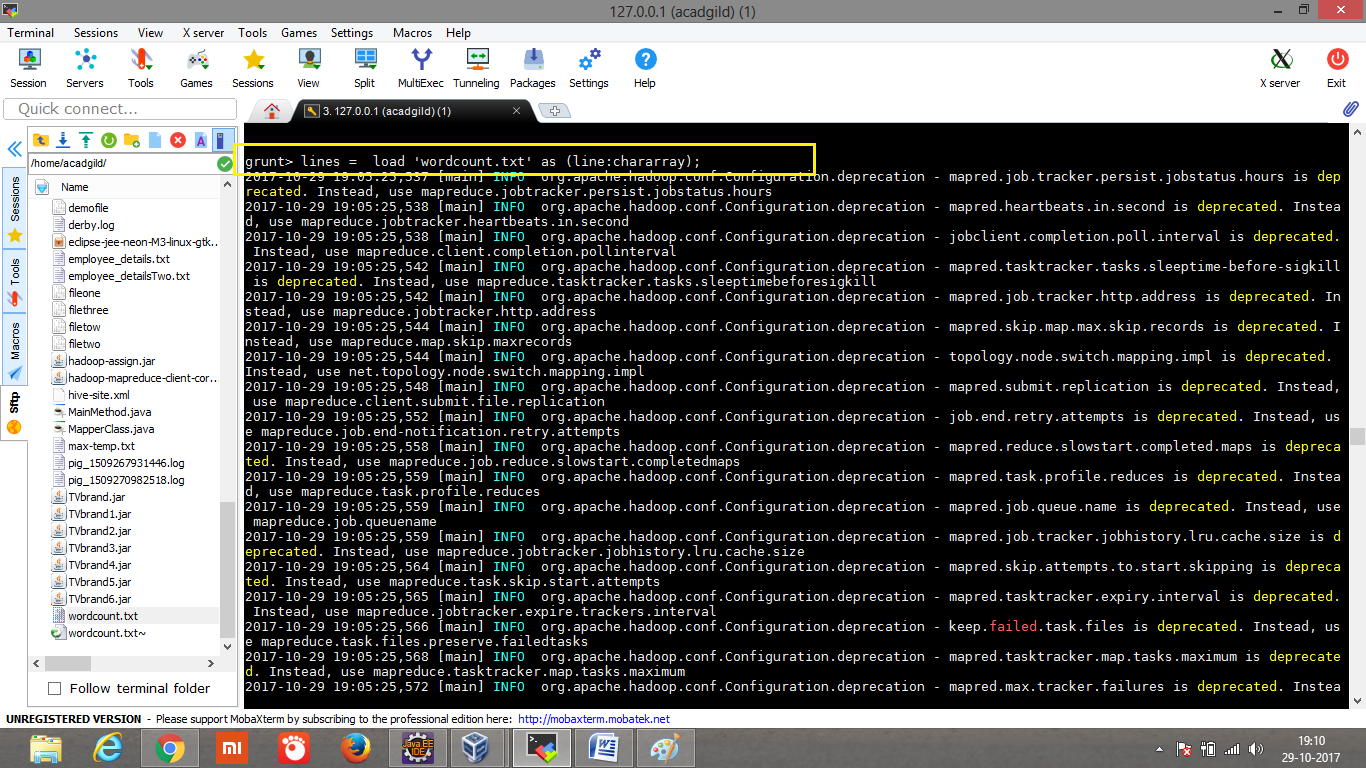
Assignment 4.3: Write a program to implement wordcount using Pig.

Created the sample datafile called wordcount.txt.

Load the file into PIG GRUNT shell using load command;

**lines = LOAD 'wordcount.txt' AS (line:chararray);** (in mobax it took file name directly, in vm we needd to give /home/acadgild/wordcount.txt)

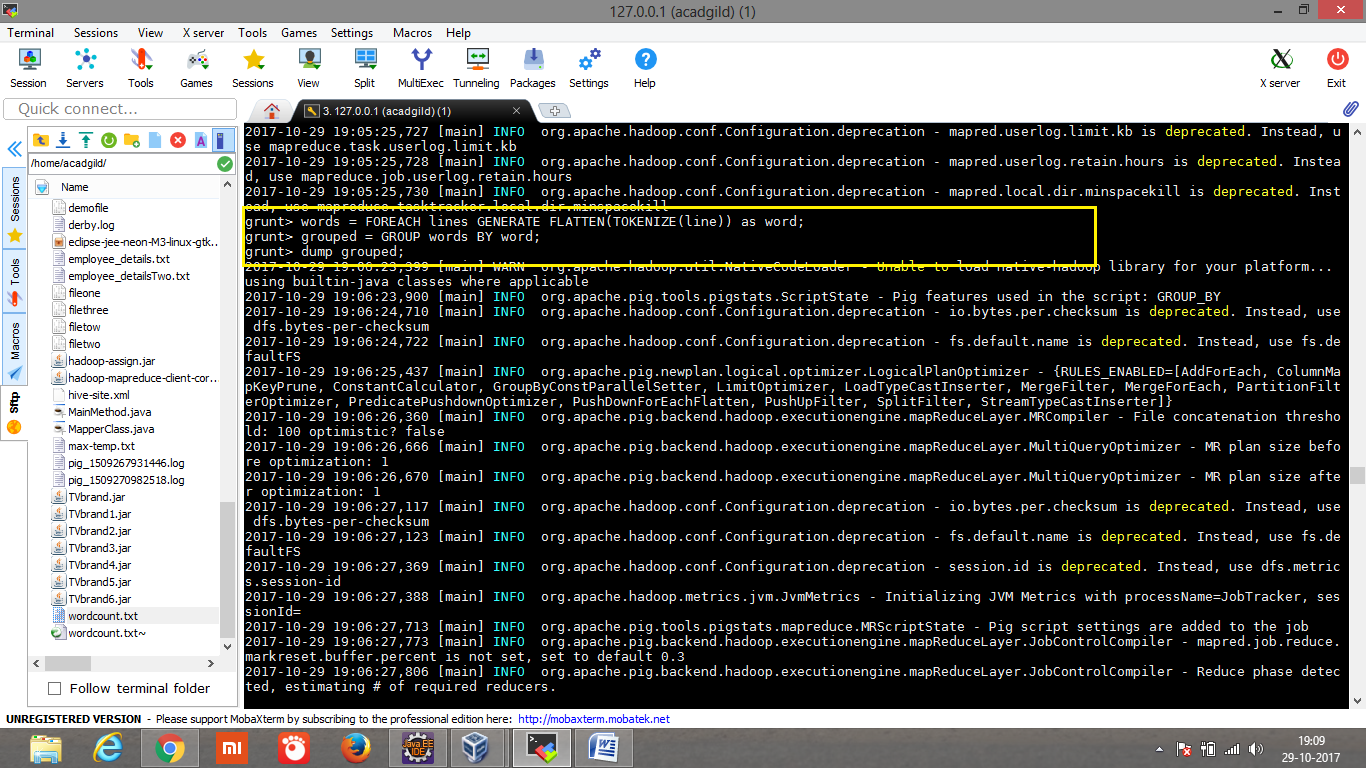


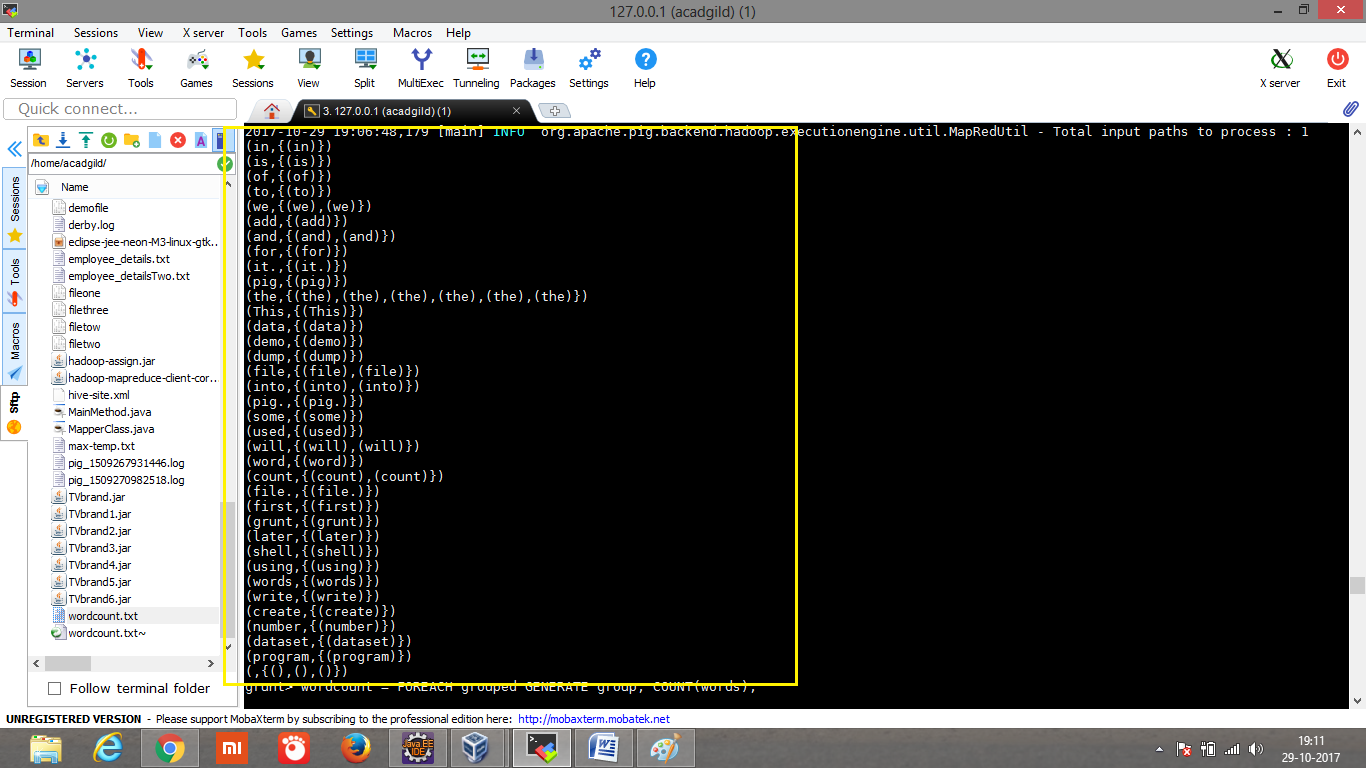
**words = FOREACH lines GENERATE FLATTEN(TOKENIZE(line)) as word;**

Splits each line into words using the TOKENIZE operator. The tokenize function creates a bag of words. Using the FLATTEN function, the bag is converted into a tuple.

**grouped = GROUP words BY word;**

The words are grouped together so that the count can be computed

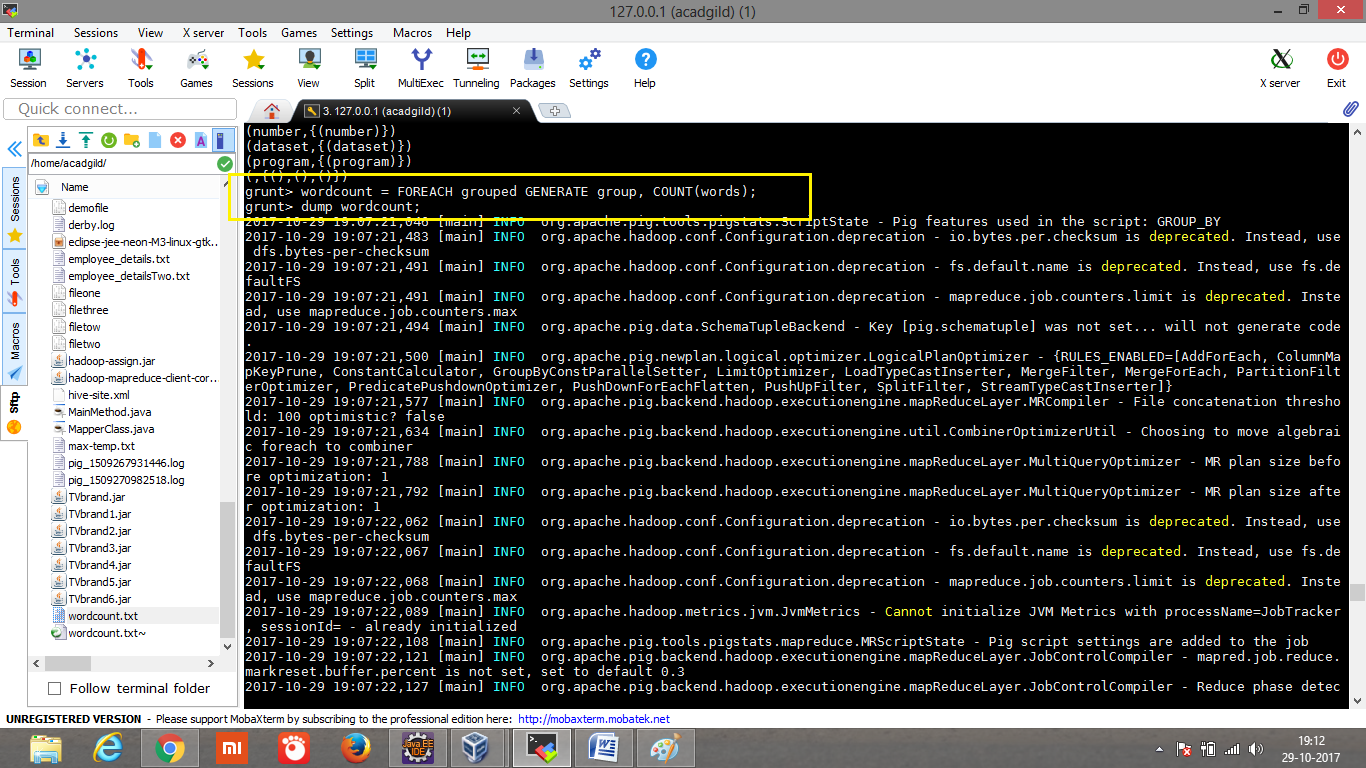




**wordcount = FOREACH grouped GENERATE group, COUNT(words);**

**DUMP wordcount;**

Count the words present in each group



**Output:**

