lines = load '/user/test/acadgild/dataset\_1.txt' using PigStorage() as (line:chararray);

words = foreach lines generate FLATTEN(TOKENIZE(line)) as word;

grouped = group words by word;

wordcount = foreach grouped generate group,COUNT(words);

