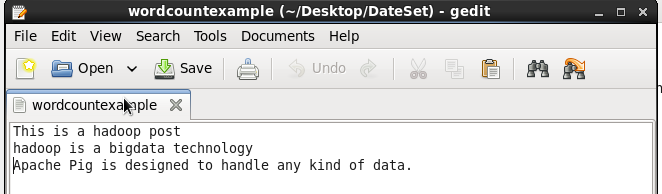
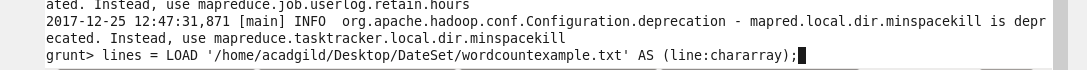
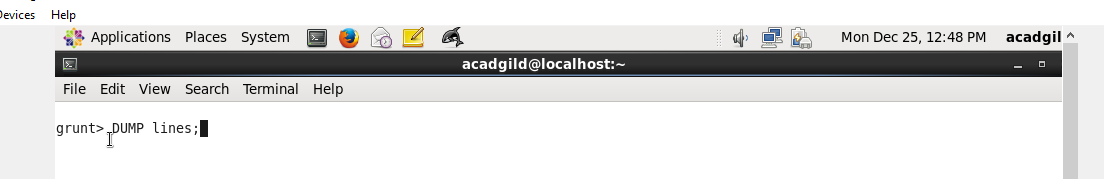
**Problem Statement: Write a program to implement wordcount using Pig.**

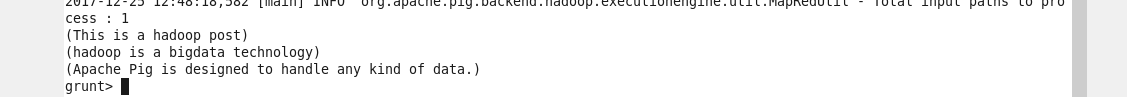
1. Create a input file for wordcount problem



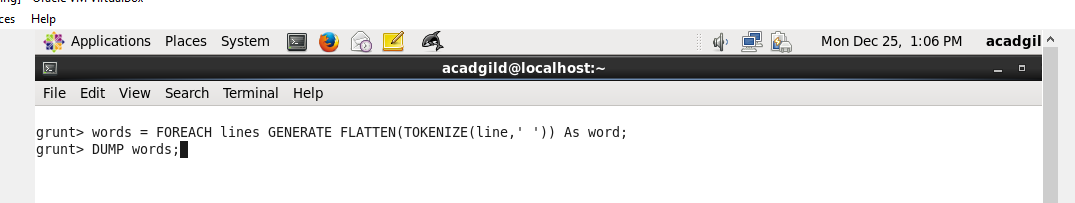
2 Load the data into Pig

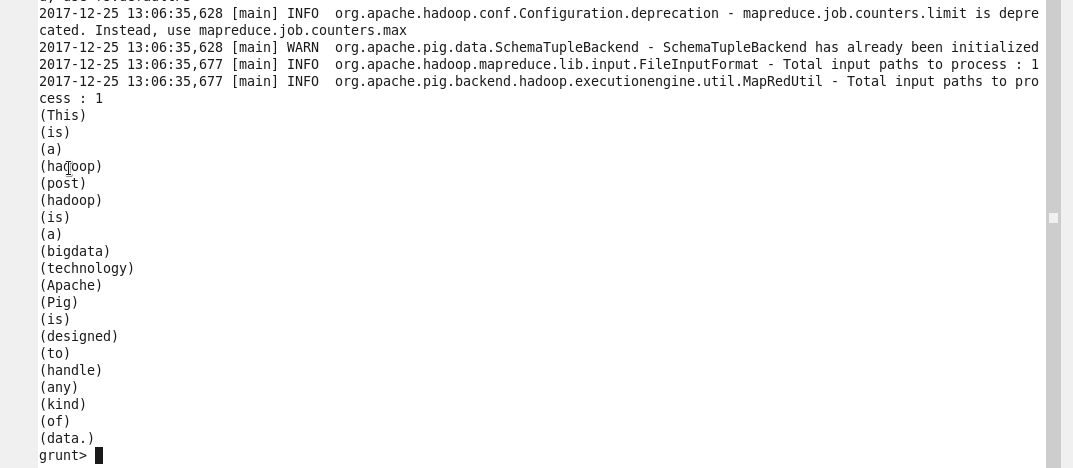






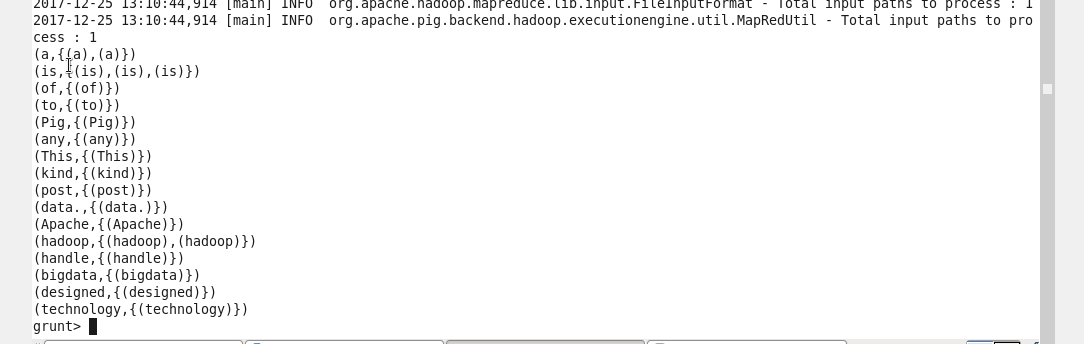
3 )first splits each line into words using the **TOKENIZE** operator. The tokenize function creates a bag of words. Using the **FLATTEN** function, the bag is converted into a tuple.



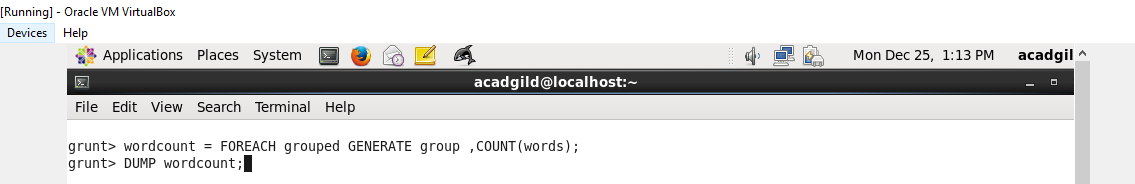


4) Now apply Group by for as we have to count each word occurrence.

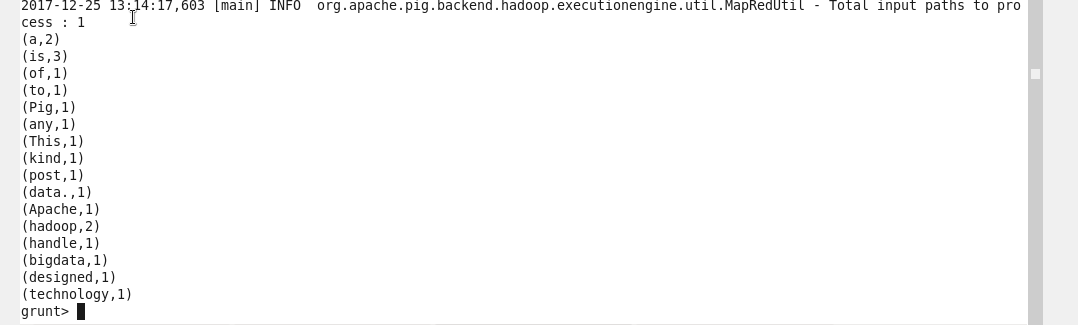




5) Now generate word count



Final output of word count



Overall pig script for word count example.

lines = LOAD '/home/acadgild/Desktop/DataSet/wordcountexample.txt' AS (line:chararray);

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

grouped = GROUP words BY word;

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

DUMP wordcount;