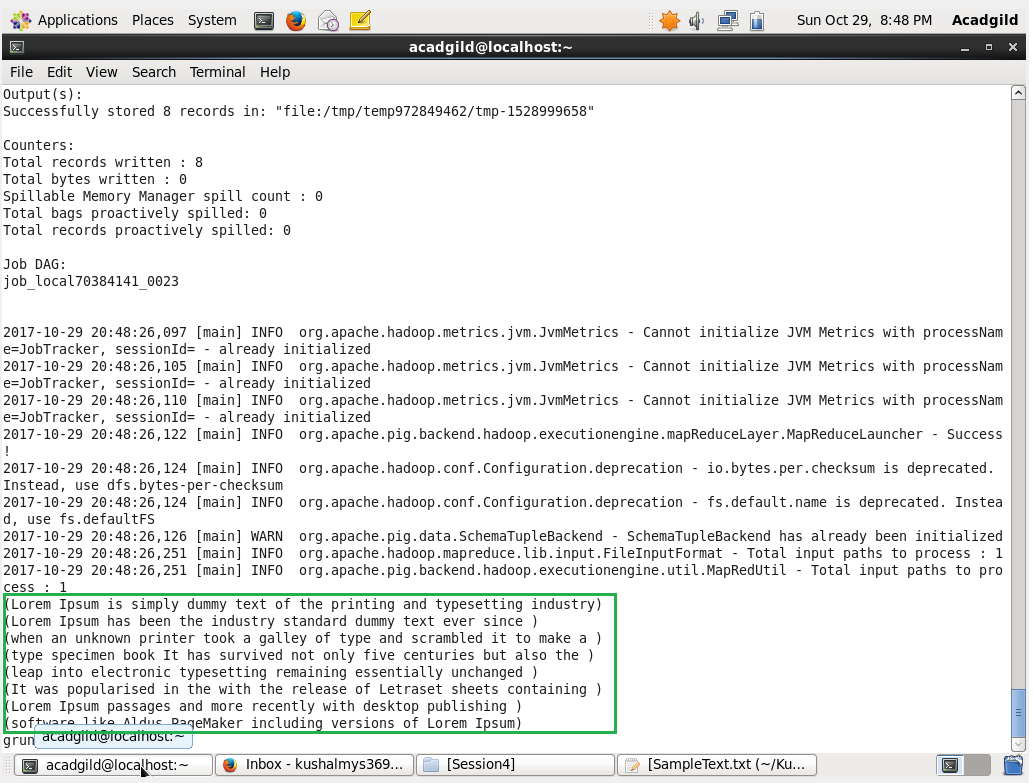
Session 4 Assignment 3 – Word Count using PIG

Loading Sample Text Data -

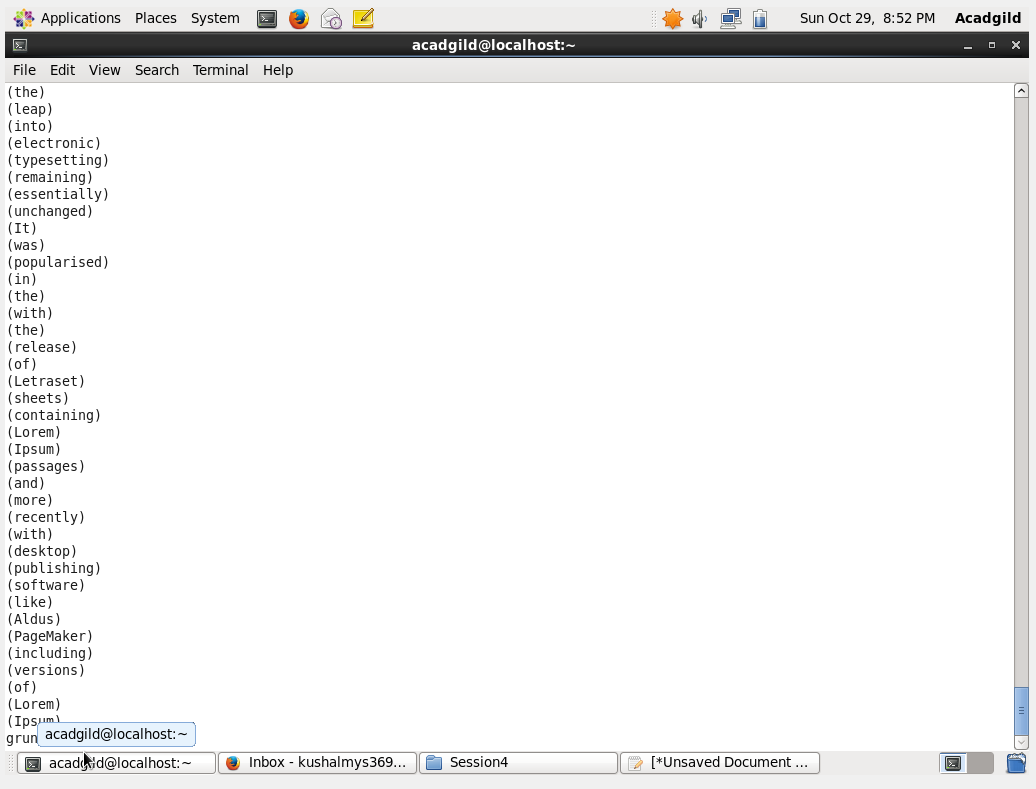
* sampleLines = LOAD ‘/home/acadgild/Kushal/Session4/SampleText.txt’ AS (sample\_Line:chararray);



Boxed is the loaded Sample data of text.

Getting Words from each line -

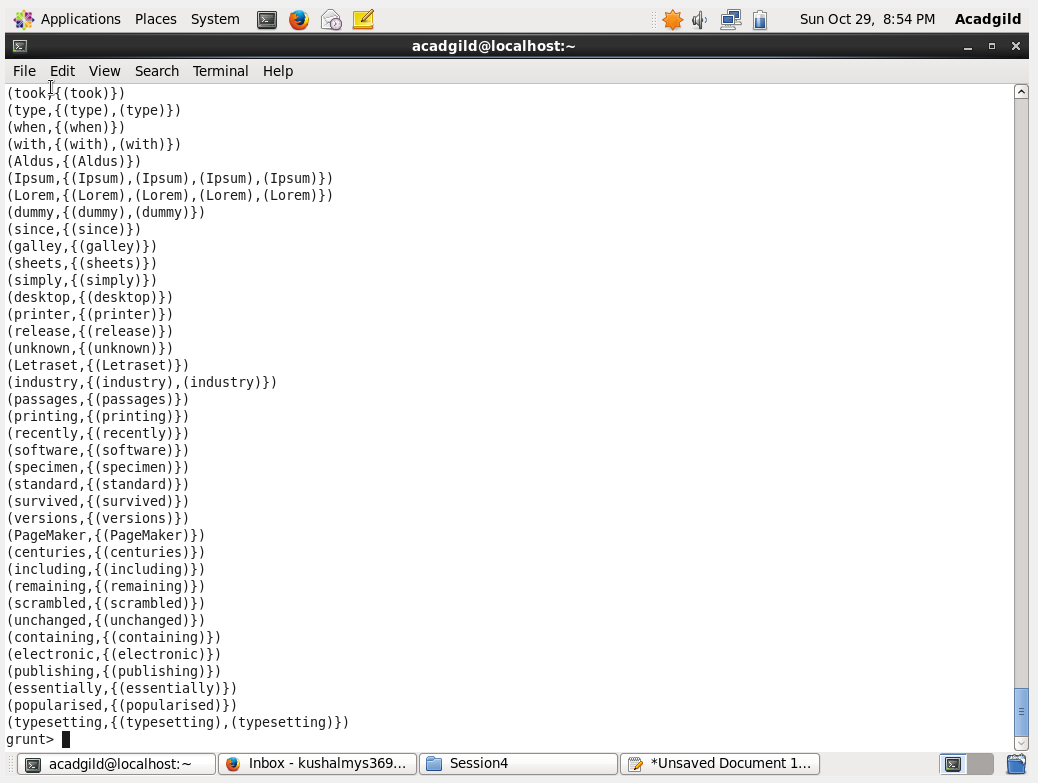
* sampleWords = FOREACH sampleLines GENERATE FLATTEN(TOKENIZE(sample\_Line)) as word;



Above is the Tokenized Words.

Grouping of words -

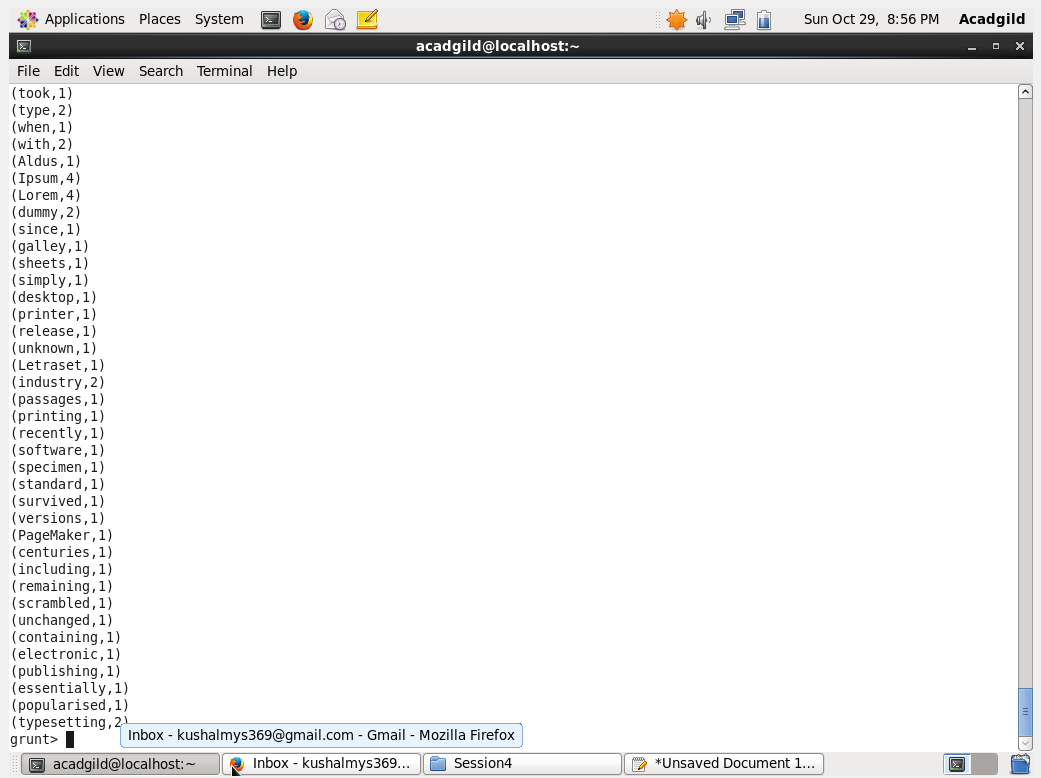
* sampleWordGroup = GROUP sampleWords BY word;



Above is the grouping of same words.

Storing the Word count with word -

* wordCount = FOREACH sampleWordGroup GENERATE group, COUNT(sampleWords);



Above is the count of corresponding words in sample Text File.