**Session 2**

**Assignment 2.3**

Student Name: Karthik K

Course: Big Data Hadoop & Spark Training

**Assignment 2.3** –

***“Write a program to implement wordcount using Pig. Share the screenshots of the commands used with its associated output.***

***Here we will write a simple pig script for the word count program.”***

# PIG Commands – Word Count Example

## Load the data from HDFS

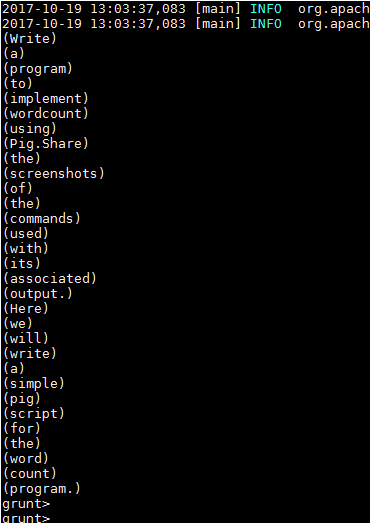
**“Word\_Count = LOAD '/home/acadgild/hadoop/wordcountpig.txt' USING PigStorage(',') AS (line:chararray);”**



## Convert the Sentence into words and Convert Column into Rows

**“Convert\_Words = FOREACH Word\_Count GENERATE FLATTEN(TOKENIZE(line,' ')) AS word;”**

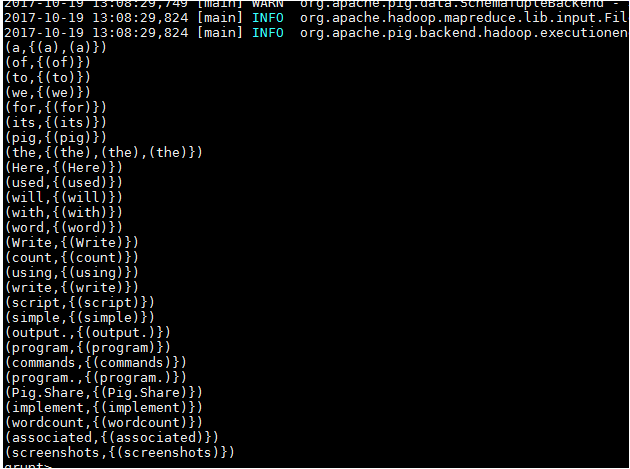




## Apply GROUP BY

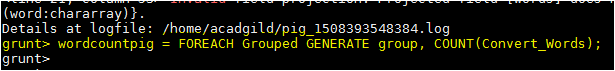
**“Grouped = GROUP Convert\_Words BY word;”**





## Generate Word Count

**“wordcountpig = FOREACH Grouped GENERATE group, COUNT(Convert\_Words);**



## Expected Output

**DUMP wordcountpig;**

