Session8: Pig

Assignment 1

# Problem Statement

Write a wordcount program to implement wordcount using Pig.

**File Data:**



**Script:**

A = load ' /home/hadoopuser/JIGYASA\_15Mar16/assignment1.txt';

B = foreach A generate flatten(TOKENIZE((chararray)$0)) as word;

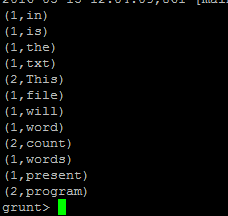
C = filter B by word matches '\\w+';

D = group C by word;

E = foreach D generate COUNT(C), group;

store E into '/home/hadoopuser/JIGYASA\_15Mar16/assignment1\_wordcount';

**Output:**



**Or**

**Script:**

lines =LOAD '/home/hadoopuser/JIGYASA\_15Mar16/assignment1.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;

**Output :**

