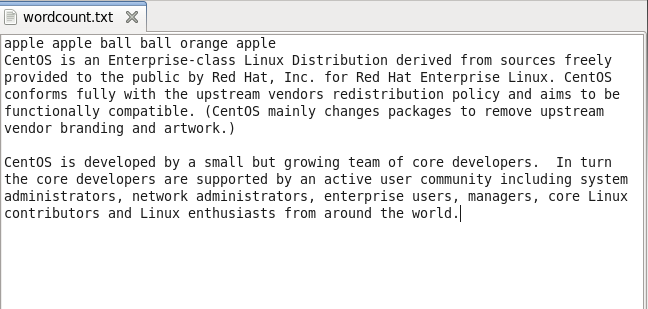
**Question:** Write a wordcount program to implement wordcount using Pig.

**Input Dataset:**

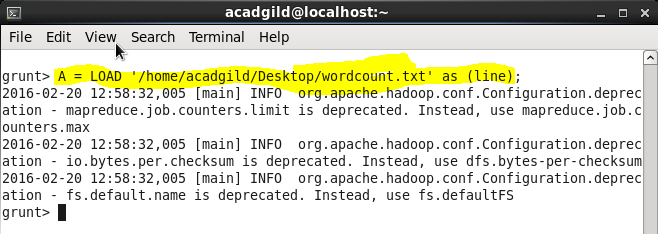


**Pig Command:**

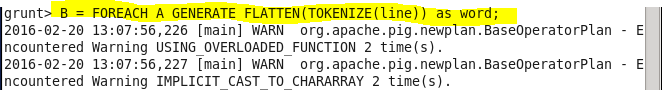
**grunt>** **set debug off;**  
**grunt> A = LOAD ''/home/acadgild/Desktop/wordcount.txt' as (line);**  
**grunt> B = FOREACH A GENERATE FLATTEN(TOKENIZE(line)) as word;**  
**grunt>** **G = GROUP B by word;  
grunt>** **Describe G;  
grunt>CNT = FOREACH G GENERATE by group, COUNT(B);  
grunt>** **Describe CNT;  
grunt>** **Dump CNT;  
grunt>** **Store CNT into '/home/acadgild/Desktop/Wordcount\_Result.txt' ;**

**Word Program Execution :**

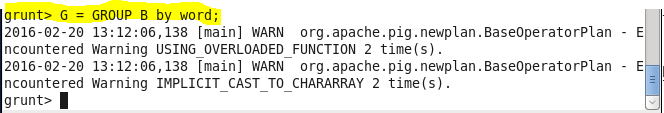
* Load input dataset to temporary alias as schema line



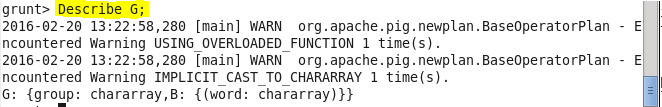
* TOKENIZE splits the line into a field for each word.
* FLATTEN will the collection of records returned by TOKENIZE and produce separate record for each one.



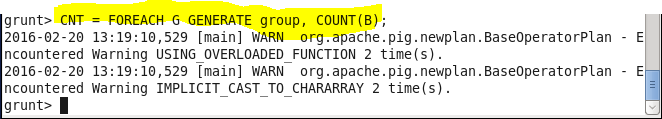
* Group them together by word



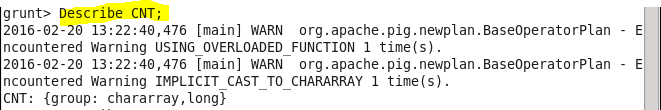
* Describe group (G)



* Count the occurrences



* Describe CNT

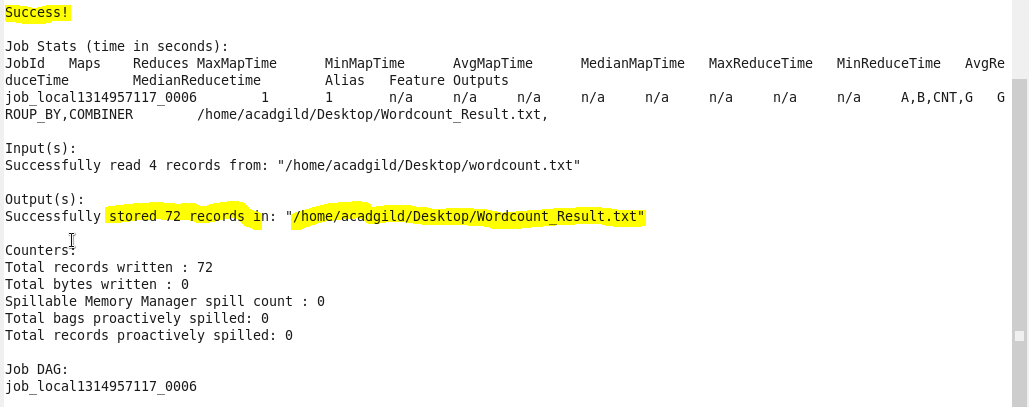


* Print Results on screen



* Store results into a folder called Wordcount\_Result.txt





**Output :**

* Two files are created in Wordcount\_Results.txt
  + Part-r-00000 (Reducer output)
  + \_SUCCESS

