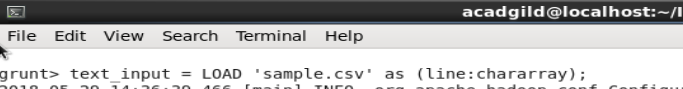
Bigdata assignment 2.3

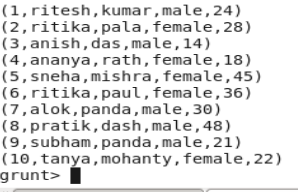
Write a program to implement wordcount using Pig.

* Entered into grunt shell by using the command – **pig -x local**



* Loaded the data into an alias - **text\_input = LOAD 'sample.csv' as (line:chararray);**



* The content of the alias - **dump text\_input;**
* Then we separate the words using TOKENIZE , after that we flatten it by using FLATTEN as foreach generates bags of tokens.

**words = foreach text\_input generate FLATTEN(TOKENIZE(line,',')) as word;**

**dump words;**

* **words** contains all the tokens.

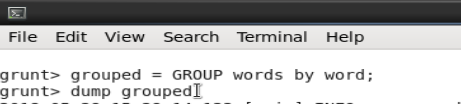




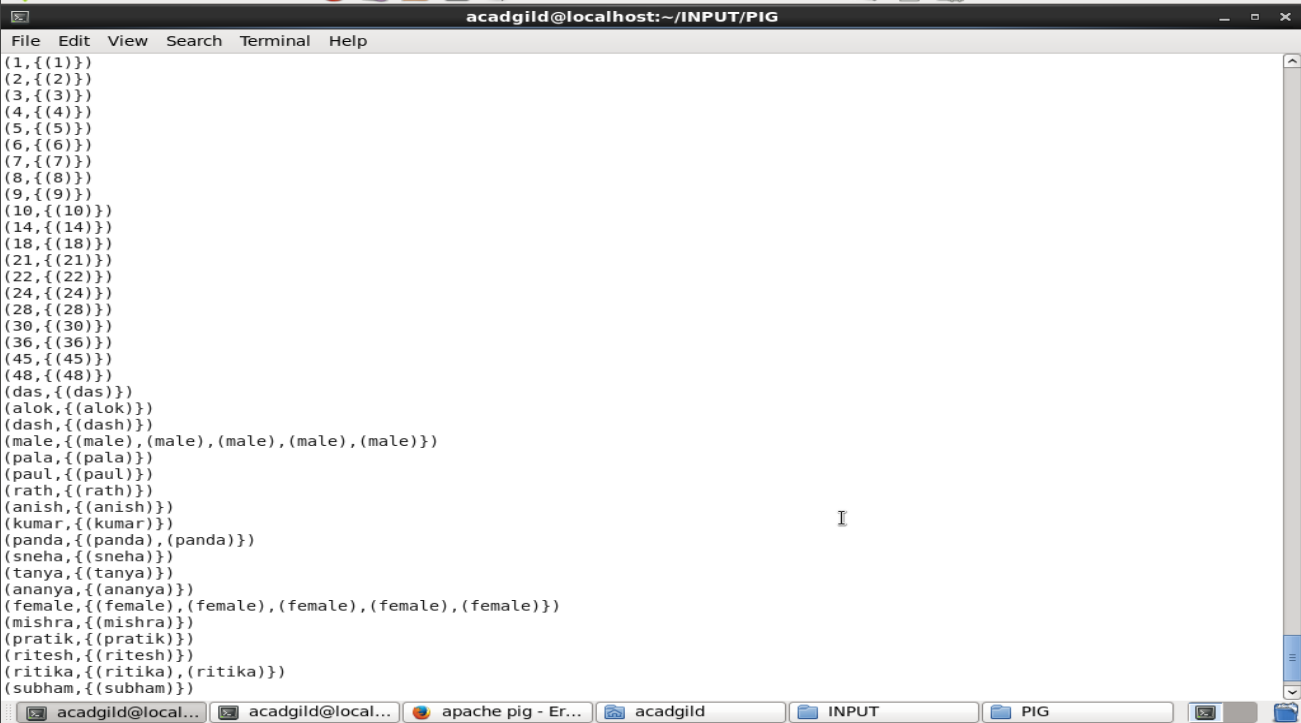
* Then we grouped the words .

**grouped = GROUP words by word;**

**dump grouped;**



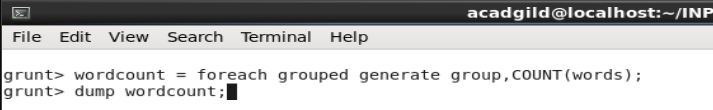
* grouped contains all the groups of words.



* After grouping the words , we used COUNT function to calculate the total count of the word.

**wordcount = foreach grouped generate group, COUNT(words);**

**dump wordcount;**



* OUTPUT- The final output wordcount shows the count of each word in the file.

