**Source Codes:**

grunt> REGISTER '/home/acadgild/flume/pig-jars/elephant-bird-hadoop-compat-4.1.jar';

grunt> REGISTER '/home/ acadgild/flume/pig-jars/elephant-bird-pig-4.1.jar';

grunt> REGISTER '/home/ acadgild/flume/pig-jars /json-simple-1.1.1.jar';

1.grunt> load\_tweets = LOAD '/user/flume/tweets/' USING com.twitter.elephantbird.pig.load.JsonLoader('-nestedLoad') AS myMap;

2.grunt> extract\_details = FOREACH load\_tweets GENERATE myMap#'id' as id,myMap#'text' as text;

3.grunt> tokens = foreach extract\_details generate id,text, FLATTEN(TOKENIZE(text)) As word;

4.grunt> dictionary = load '/AFINN.txt' using PigStorage('\t') AS(word:chararray,rating:int);

5.grunt> word\_rating = join tokens by word left outer, dictionary by word using 'replicated';

6.grunt> rating = foreach word\_rating generate tokens::id as id,tokens::text as text, dictionary::rating as rate;

7.grunt> word\_group = group rating by (id,text);

8.grunt> avg\_rate = foreach word\_group generate group, AVG(rating.rate) as tweet\_rating

9.grunt> positive\_tweets = filter avg\_rate by tweet\_rating>=0;

10.grunt> dump positive tweets;

|  |  |
| --- | --- |
| **Output** |  |
|  |  |

