**Follow the below blog and perform sentiment analysis using Mapreduce/pig**

**https://acadgild.com/blog/sentiment-analysis-on-tweets-using-afinn-dictionary/**

**Submit the screen shots of the final results with the Source code.**

**SOLUTION:**

*Screenshot is at the end.*

**1)Download the JARS:**

REGISTER '/home/acadgild/Downloads/elephant-bird-hadoop-compat-4.1.jar';

REGISTER '/home/acadgild/Downloads/elephant-bird-pig-4.1.jar';

REGISTER '/home/acadgild/Downloads/json-simple-1.1.1.jar';

**2)**

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

**3)**

**extract\_details = FOREACH load\_tweets GENERATE myMap#'id' as id,myMap#'text' as text;**  
  
describe extract\_details;

extract\_details: {id: bytearray,text: bytearray}

**4)**

**tokens = foreach extract\_details generate id,text, FLATTEN(TOKENIZE(text)) As word;**

dump tokens;

OUTPUT:  
Hadoop / Spark Conference Japan 2016 Tickets, Mon, Feb 8, 2016 at 10:00 AM | Eventbrite [https://t.](https://t./)…,|)  
(689095813549678593,RT @shimtoru: 参加登録始まりました！2月8日、参加費無料です |   
Hadoop / Spark Conference Japan 2016 Tickets, Mon, Feb 8, 2016 at 10:00 AM | Eventbrite [https://t.](https://t./)…,Eventbrite)  
(689095813549678593,RT @shimtoru: 参加登録始まりました！2月8日、参加費無料です |   
Hadoop / Spark Conference Japan 2016 Tickets, Mon, Feb 8, 2016 at 10:00 AM | Eventbrite [https://t.](https://t./)…,[https://t.](https://t./)…)  
(689098580037095424,#CustomerService #Job in #SãoPaulo: Customer Operations Engineer - Global at Cloudera <https://t.co/8yI45tXivr> #Hadoop #BigData #data #Jobs,#CustomerService)

**5)dictionary = load '/user/acadgild/AFINN.txt' using PigStorage('\t') AS(word:chararray,rating:int);**

describe dictionary;

output:  
dictionary: {word: chararray,rating: int}

**6)word\_rating = join tokens by word left outer, dictionary by word using 'replicated';**  
output:  
word\_rating: {tokens::id: bytearray,tokens::text: bytearray,tokens::word: chararray,dictionary::word: chararray,dictionary::rating: int}

**7)rating = foreach word\_rating generate tokens::id as id,tokens::text as text, dictionary::rating as rate;**

describe rating;  
  
rating: {id: bytearray,text: bytearray,rate: int}

**8)word\_group = group rating by (id,text);**

dump word\_group;

output:  
((689097535017865217,RT @MobinRanjbar: Cloudera Hue &gt; Apache Ambari   
@gethue   
@ApacheAmbari   
#hadoop  
#bigdata  
#hue  
#ambari),{(689097535017865217,RT @MobinRanjbar: Cloudera Hue &gt; Apache Ambari   
@gethue   
@ApacheAmbari   
#hadoop  
#bigdata  
#hue  
#ambari,),(689097535017865217,RT @MobinRanjbar: Cloudera Hue &gt; Apache Ambari   
@gethue   
@ApacheAmbari   
#hadoop  
#bigdata  
#hue  
#ambari,),(689097535017865217,RT @MobinRanjbar: Cloudera Hue &gt; Apache Ambari   
@gethue   
@ApacheAmbari   
#hadoop  
#bigdata  
#hue  
#ambari,),(689097535017865217,RT @MobinRanjbar: Cloudera Hue &gt; Apache Ambari   
@gethue   
@ApacheAmbari   
#hadoop  
#bigdata  
#hue  
#ambari,),(689097535017865217,RT @MobinRanjbar: Cloudera Hue &gt; Apache Ambari   
@gethue   
@ApacheAmbari   
#hadoop  
#bigdata  
#hue  
#ambari,),(689097535017865217,RT @MobinRanjbar: Cloudera Hue &gt; Apache Ambari   
@gethue   
@ApacheAmbari   
#hadoop  
#bigdata  
#hue  
#ambari,),(689097535017865217,RT @MobinRanjbar: Cloudera Hue &gt; Apache Ambari

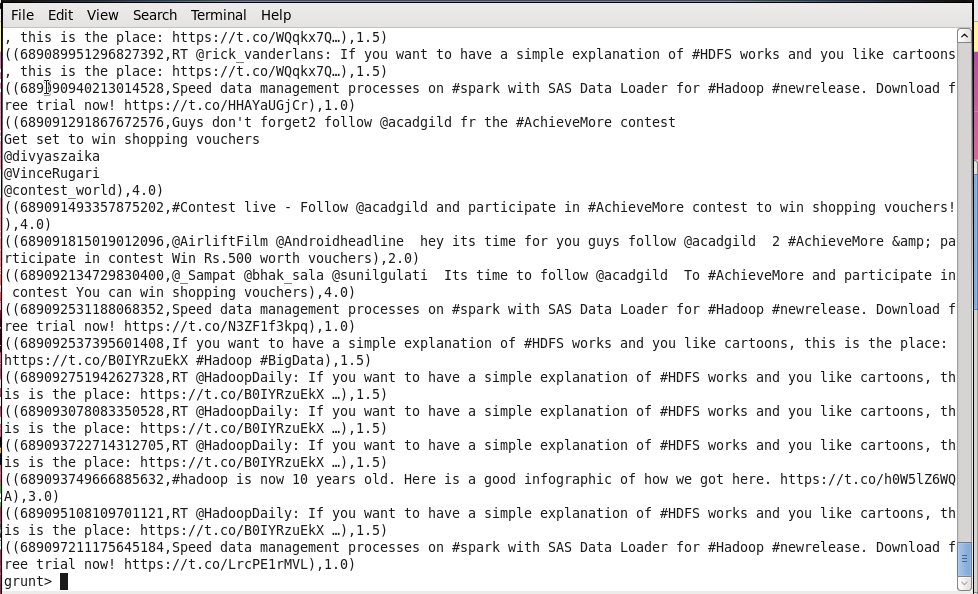
**9)avg\_rate = foreach word\_group generate group, AVG(rating.rate) as tweet\_rating;**  
  
avg\_rate: {group: (id: bytearray,text: bytearray),tweet\_rating: double}

**10)**

**positive\_tweets = filter avg\_rate by tweet\_rating>=0;**

positive\_tweets: {group: (id: bytearray,text: bytearray),tweet\_rating: double}

((689085590822891521,@filmfan hey its time for you guys follow @acadgild  To #AchieveMore and participate in contest Win Rs.500 worth vouchers),2.0)  
((689085611639205888,@sujitjohn Follow @acadgild  To #AchieveMore &amp; participate in contest Hurry up! &amp; win Flipkart vouchers),4.0)  
((689085636456923138,@sujitlalwani Hey tweeps want to win Flipkart vouchers so follow @acadgild  To #AchieveMore contest Prizes of Rs.500),2.5)  
((689085663334035456,@im\_bicky Hey Friends &amp; Tweeples Please Follow @acadgild  To #AchieveMore and participate in contest Be lucky get Rs.500 vouchers),3.0)  
((689085686390063104,@ShreyVithalani You don't wanna miss this Go follow @acadgild  To #AchieveMore and play in contest &amp; win Flipkart vouchers),1.0)  
((689085710079537154,@SANGEETAAGRAWA Tweethearts! Follow @acadgild  To #AchieveMore and participate in contest and win Flipkart vouchers! Aye!),4.0)  
((689085776731213824,@AryanSarath Contestest freaks! Follow @acadgild  To #AchieveMore and play in contest &amp; win shopping vouchers Wohooo!),4.0)  
((689085820674945025,@RaviThakkar Its time to follow @acadgild  To #AchieveMore and participate in contest You can win shopping vouchers),4.0)  
((689085845538762752,@iwant\_us Follow @acadgild  To #AchieveMore and participate in contest &amp; Win Flipkart vouchers! Win Rs.500 worth vouchers),2.0)  
((689085871522496518,@DeEp\_ Its Crackling! The Contest will blow your mind! follow @acadgild  To #AchieveMore &amp; win Flipkart vouchers),4.0)  
((689085896730263553,@sutharsweta MASSIVE contest FOLLOW @acadgild  To #AchieveMore and play in contest &amp; win Flipkart vouchers Win Rs.500 worth vouchers),3.0)  
((689085921854124032,@ygeshsni Tweethearts!! Follow @acadgild  To #AchieveMore and participate in contest Hurry Up! &amp; win Flipkart vouchers),4.0)  
((689085951352680448,@adi\_shah Tweethearts! Wanna beat your day blues! Follow @acadgild  To #AchieveMore &amp; play in contest &amp; win Flipkart vouchers),4.0)  
((689087505174540288,If you want to have a simple explanation of #HDFS works and you like cartoons, this is the place: <https://t.co/WQqkx7QpP3> #Hadoop #BigData),1.5)  
((689088647493189632,RT @rick\_vanderlans: If you want to have a simple explanation of #HDFS works and you like cartoons, this is the place: <https://t.co/WQqkx7Q>…),1.5)  
((689089909542531072,RT @rick\_vanderlans: If you want to have a simple explanation of #HDFS works and you like cartoons, this is the place: <https://t.co/WQqkx7Q>…),1.5)  
((689089951296827392,RT @rick\_vanderlans: If you want to have a simple explanation of #HDFS works and you like cartoons, this is the place: <https://t.co/WQqkx7Q>…),1.5)  
((689090940213014528,Speed data management processes on #spark with SAS Data Loader for #Hadoop #newrelease. Download free trial now! <https://t.co/HHAYaUGjCr>),1.0)  
((689091291867672576,Guys don't forget2 follow @acadgild fr the #AchieveMore contest  
Get set to win shopping vouchers  
@divyaszaika   
@VinceRugari   
@contest\_world),4.0)  
((689091493357875202,#Contest live - Follow @acadgild and participate in #AchieveMore contest to win shopping vouchers!),4.0)  
((689091815019012096,@AirliftFilm @Androidheadline  hey its time for you guys follow @acadgild  2 #AchieveMore &amp; participate in contest Win Rs.500 worth vouchers),2.0)  
((689092134729830400,@\_Sampat @bhak\_sala @sunilgulati  Its time to follow @acadgild  To #AchieveMore and participate in contest You can win shopping vouchers),4.0)  
((689092531188068352,Speed data management processes on #spark with SAS Data Loader for #Hadoop #newrelease. Download free trial now! <https://t.co/N3ZF1f3kpq>),1.0)  
((689092537395601408,If you want to have a simple explanation of #HDFS works and you like cartoons, this is the place: <https://t.co/B0IYRzuEkX> #Hadoop #BigData),1.5)  
((689092751942627328,RT @HadoopDaily: If you want to have a simple explanation of #HDFS works and you like cartoons, this is the place: <https://t.co/B0IYRzuEkX> …),1.5)  
((689093078083350528,RT @HadoopDaily: If you want to have a simple explanation of #HDFS works and you like cartoons, this is the place: <https://t.co/B0IYRzuEkX> …),1.5)  
((689093722714312705,RT @HadoopDaily: If you want to have a simple explanation of #HDFS works and you like cartoons, this is the place: <https://t.co/B0IYRzuEkX> …),1.5)  
((689093749666885632,#hadoop is now 10 years old. Here is a good infographic of how we got here. <https://t.co/h0W5lZ6WQA>),3.0)  
((689095108109701121,RT @HadoopDaily: If you want to have a simple explanation of #HDFS works and you like cartoons, this is the place: <https://t.co/B0IYRzuEkX> …),1.5)  
((689097211175645184,Speed data management processes on #spark with SAS Data Loader for #Hadoop #newrelease. Download free trial now! <https://t.co/LrcPE1rMVL),1.0>)

****