**ACD\_BDD\_Project1.1**

**1. Associated Data Files**

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

**2. Problem Statements**

Follow the below blog and perform sentiment analysis using MapReduce/Pig.

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

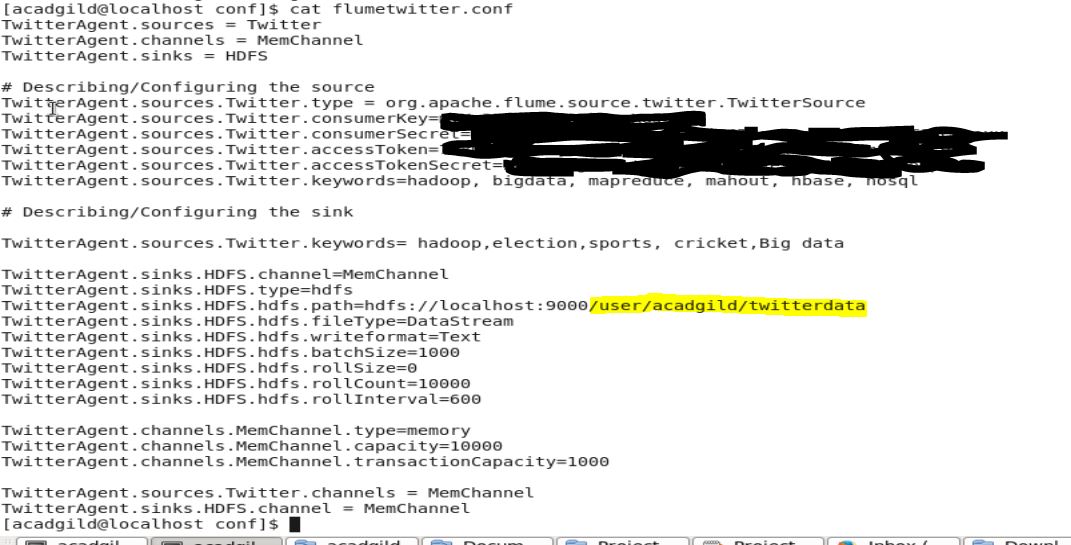
Submit the screen shots of the final results with the source code.

**Solution**

Created flume configuration file.

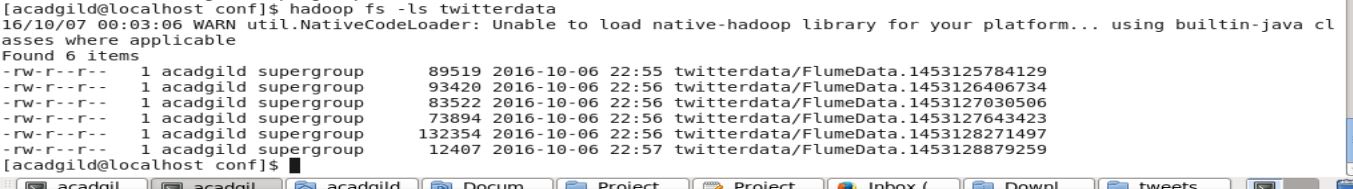


Configured flume conf file with consumer key and access token.



Stored streamed data in HDFS using below command.

**flume-ng agent -n TwitterAgent -f /usr/local/flume/conf/flumetwitter.conf**

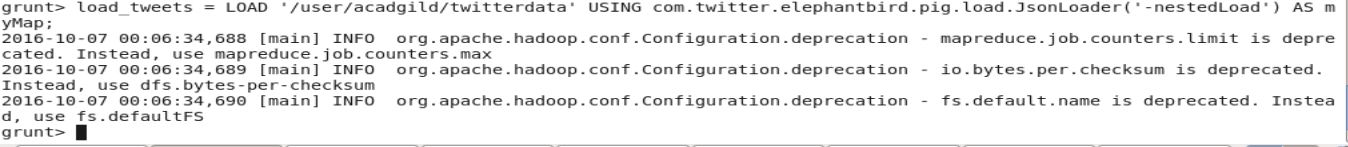


Stored tweets.



**Pig commands:**

REGISTER '/home/acadgild/Documents/Project1.1/elephant-bird-hadoop-compat-4.1.jar';  
REGISTER '/home/acadgild/Documents/Project1.1/elephant-bird-pig-4.1.jar';  
REGISTER '/home/acadgild/Documents/Project1.1/json-simple-1.1.1.jar';  
  
load\_tweets = LOAD '/user/acadgild/twitterdata' USING com.twitter.elephantbird.pig.load.JsonLoader('-nestedLoad') AS myMap;



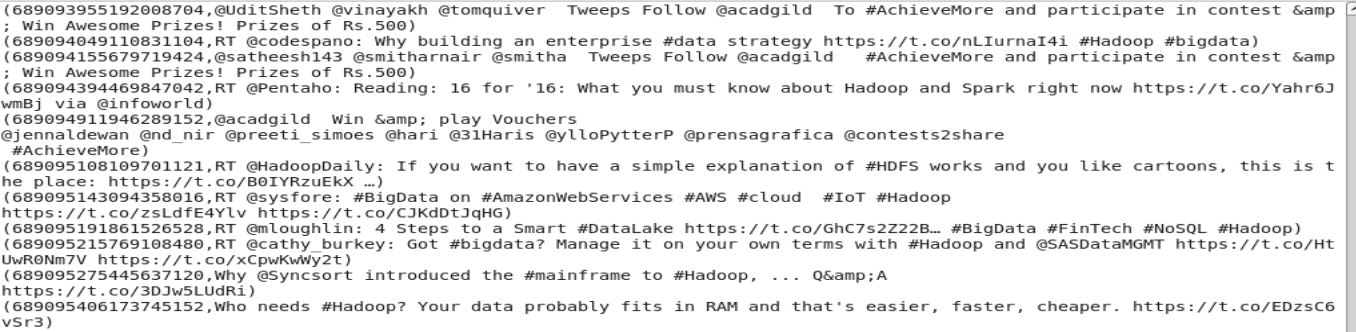
dump load\_tweets;



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

C:\Users\RAKESH\Desktop\Screenshot\Project1.1\S8.JPG

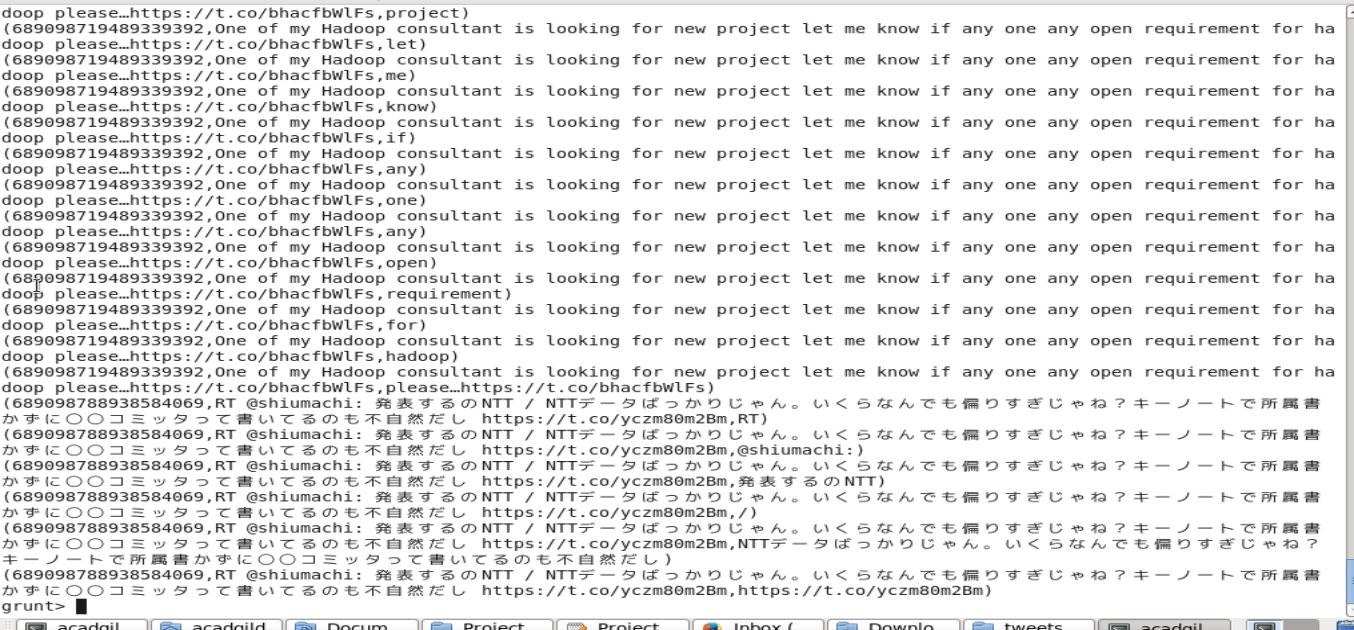
dump extract\_details;



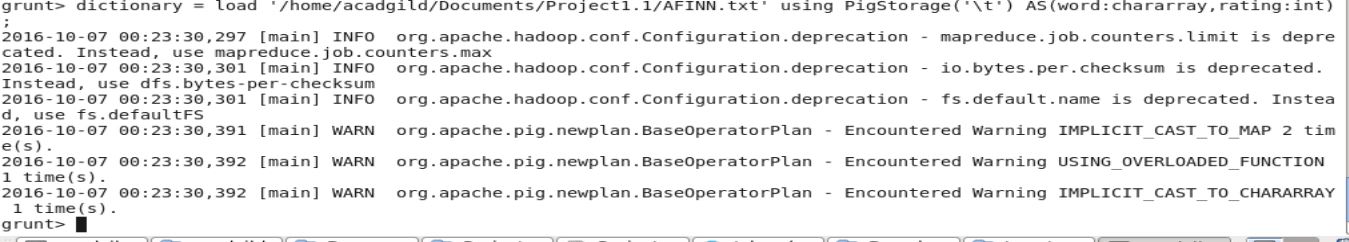
tokens = foreach extract\_details generate id,text, FLATTEN(TOKENIZE(text)) As word;



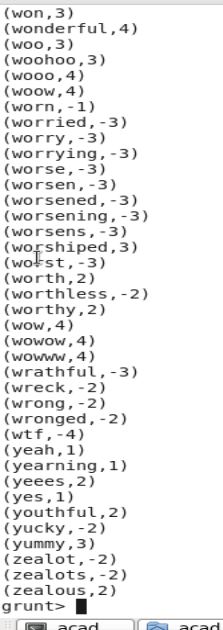
dump tokens;



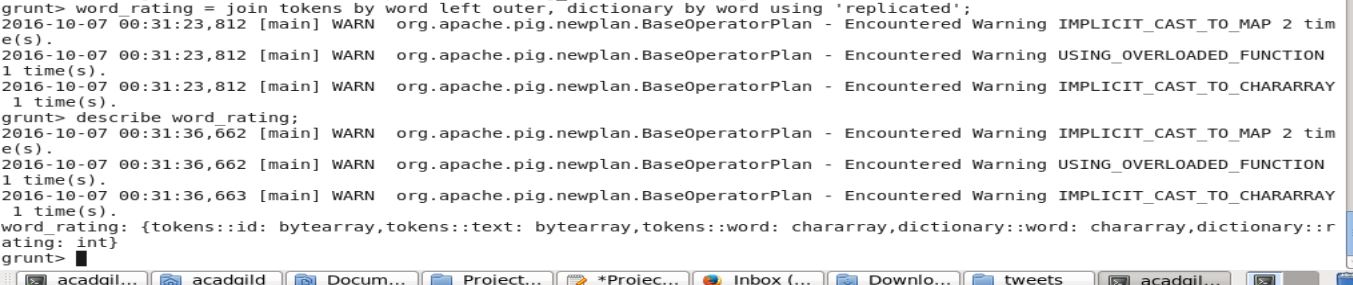
dictionary = load '/user/acadgild/afinn/AFINN.txt' USING PigStorage('\t') AS (word:chararray,rating:int);



dump dictionary;



word\_rating = join tokens by word left outer, dictionary by word using 'replicated';



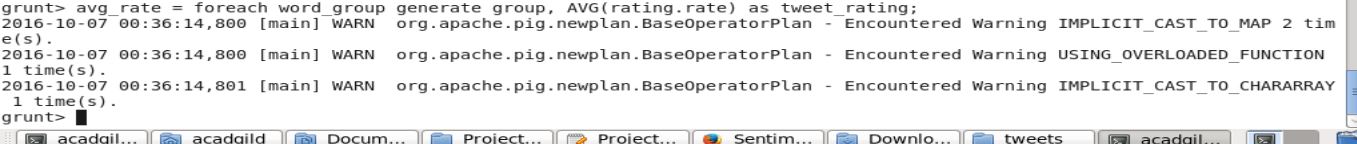
rating = foreach word\_rating generate tokens::id as id,tokens::text as text, dictionary::rating as rate;



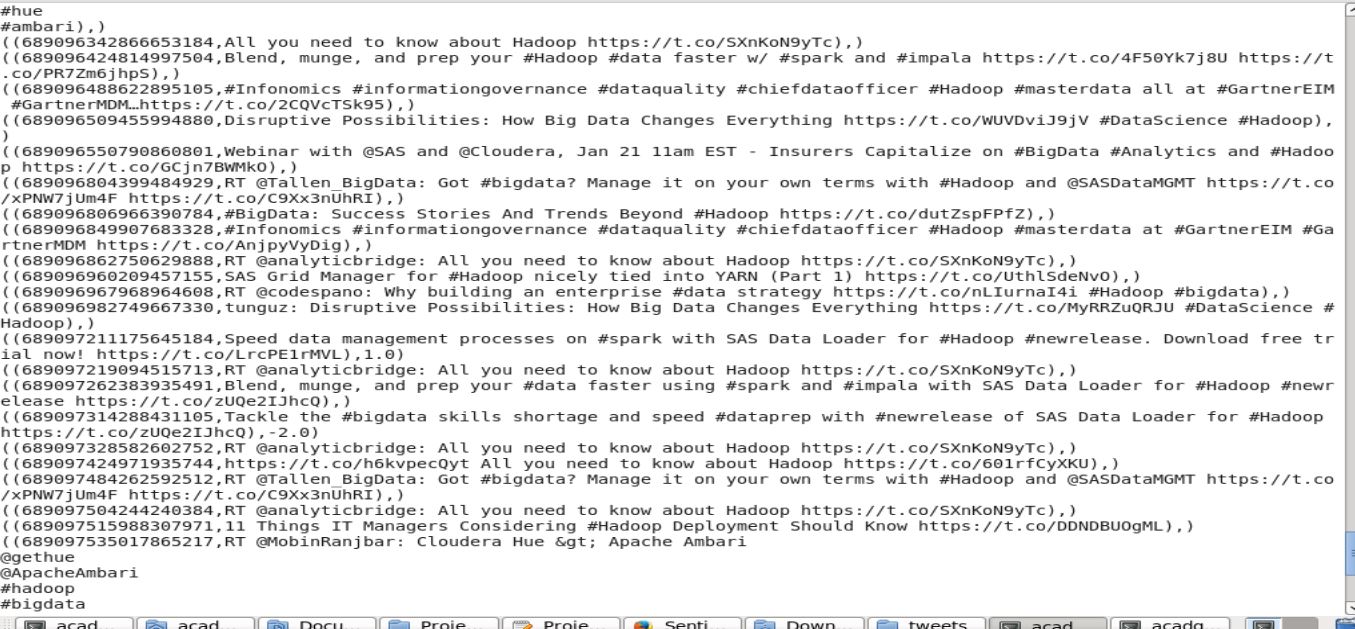
word\_group = group rating by (id,text);



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



dump avg\_rate;



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



dump positive\_tweets;

