**Step 1: First load the tweet dataset in pig**

data= LOAD 'tweets.csv' USING PigStorage(',') as (sid:int,text:chararray,favorited:chararray,favoriteCount:int,replyToSN:chararray,created:chararray,truncated:chararray,replyToSID:int,id:int,replyToUID:int,statusSource:chararray,screenName:chararray,retweetCount:int,isRetweet:chararray,retweeted:chararray);

**Step 2: Extract id and text from the dataset**

extract\_details = FOREACH load\_tweets GENERATE $0 as id,$1 as text;

**Step 3: Now for the text we need to divide it into words to calculate sentiment analysis**

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

**Step 4:Now load the dictionary file in pig**

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

**Step 5: Perform a map side join by joining the tokens statement and the dictionary contents using this relation:**

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

**Step 6:Now we will extract the id,tweet text and word rating(from the dictionary) by using the below relation.**

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

**Step 7:Now, we will group the rating of all the words in a tweet by using the below relation:**

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

**Step 8:Now, let’s perform the Average operation on the rating of the words per each tweet.**

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

Now we have calculated the Average rating of the tweet using the rating of each word.

From the above relation, we will get all the tweets i.e., both positive and negative.

**Step 9: Now we will filter the positive tweets using the below statement:**

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

**Step 10:negative tweets:**

negative\_tweets = filter avg\_rate by tweet\_rating<0;

**Step 11: To know the count first we need to group the tweets individually.**

grp= group positive\_tweets all;

pos\_count= foreach grp generate COUNT(positive\_tweets);

same with negative tweets

neg= group negative\_tweets all;

neg\_count= foreach neg generate COUNT(negative\_tweets);