hdfs dfs -put /home/cloudera/unemp\_rate.txt /user/cloudera/pjct/input   
// moves the input text file named ‘unemp\_rate.txt’ to cloudera/pjct/input which was created erlier.

pig -x mapreduce

//change to grunt command line

rows = LOAD '/user/cloudera/pjct/input/unemp\_rate.txt' using PigStorage(',') AS (Year:int,State:chararray,Rate:float);

//Loads the file data using ‘,’(comma) as delimiter with schema Year,State,Rate

row\_low = FOREACH rows GENERATE Year,LOWER(State)as lstate,Rate;

//Converts the states to lower case so that all states are grouped without mismatch due to case insensitivity  
  
row\_l\_2009 = FILTER row\_low BY Year==2009;

//Filters the unemployment rates by year=2009  
  
row\_l\_2010 = FILTER row\_low BY Year==2010;  
//Filters the unemployment rates by year=2010  
  
 row\_l\_2011 = FILTER row\_low BY Year==2011;  
//Filters the unemployment rates by year=2011  
  
row\_l\_2012 = FILTER row\_low BY Year==2012;  
//Filters the unemployment rates by year=2012  
  
row\_l\_2013 = FILTER row\_low BY Year==2013;  
//Filters the unemployment rates by year=2013  
  
 row\_l\_2014 = FILTER row\_low BY Year==2014;  
//Filters the unemployment rates by year=2014  
  
row\_l\_2015 = FILTER row\_low BY Year==2015;  
//Filters the unemployment rates by year=2015  
  
row\_l\_2016 = FILTER row\_low BY Year==2016;  
//Filters the unemployment rates by year=2016

The following code compares the difference in unemployment rates from previous year.

 Joined = JOIN row\_l\_2009 by $1, row\_l\_2010 by $1, row\_l\_2011 by $1, row\_l\_2012 by $1, row\_l\_2013 by $1, row\_l\_2014 by $1, row\_l\_2015 by $1 ,row\_l\_2016 by $1;

//Joins the data by using state as common parameter  
  
 diff = FOREACH Joined Generate $1 as State,$5-$2 as d\_09\_10, $8-$5 as d\_10\_11, $11-$8 as d\_11\_12, $14-$11 as d\_12\_13, $17-$14 as d\_13\_14, $20-$17 as d\_14\_15, $23-$20 as d\_15\_16 ;

//produces the difference in unemployment from previous year.