**GDP Dataset**

hdfs dfs -put /home/cloudera/Desktop/Project/present.txt /user/cloudera/project/   
// moves the input text file named ‘present.txt’ to cloudera/project/ which was created earlier (containing GDP data set for years 2008-2015)

hdfs dfs -put /home/cloudera/Desktop/Project/past.txt /user/cloudera/project/   
// moves the input text file named ‘proj.txt’ to cloudera/project/ which was created earlier (containing GDP data set for years 2000-2008)

pig -x mapreduce

//change to grunt command line

**FOR GDP DATA 2008-2015**

dataset = LOAD '/user/cloudera/project/present.txt' using PigStorage(',') AS (State:chararray, D2008:double, D2009:double, D2010:double, D2011:double, D2012:double, D2013:double, D2014:double, D2015:double);

dataset\_low = FOREACH rows GENERATE LOWER(State)as Lstate, D2008, D2009, D2010, D2011, D2012, D2013, D2014, D2015;

//Converts the states to lower case so that all states are grouped without mismatch due to case insensitivity

Diff= FOREACH dataset\_low GENERATE $0, $1, $8-$1 as A;

Div= FOREACH Diff GENERATE $0, $1, $2/$1;

Percentage= FOREACH Div GENERATE $0, $1, $2, 100.0\*$2;

Dump Percentage;

**FOR GDP DATA 2000-2008**

dataset = LOAD '/user/cloudera/project/present.txt' using PigStorage(',') AS (State:chararray, D2000:double, D2001:double, D2002:double, D2003:double, D2004:double, D2005:double, D2006:double, D2007:double, D2008:double);

dataset\_low = FOREACH rows GENERATE LOWER(State)as Lstate, D2000, D2001, D2002, D2003, D2004, D2005, D2006, D2007, D2008;

Diff= FOREACH dataset\_low GENERATE $0, $1, $9-$1 as A;

Div= FOREACH Diff GENERATE $0, $1, $2/$1;

Percentage= FOREACH Div GENERATE $0, $1, $2, 100.0\*$2;

Dump Percentage;