**Exercise: Basic Pig Operators for Reading, Filtering, and Writing Data**

# Using Basic Pig Operators for Reading, Filtering, and Writing Data.

In the following exercises, you create Pig scripts to read data from the **census\_2010.csv** file in HDFS. Write the results to the **/user/student/DIHPS/output** location. The **census\_2010.csv** file is ***tab delimited***. It contains two columns, **COUNTY** (string) and **POPULATION** (integer). Below are the specifics of the exercise:

* + Copy the **/home/student/DIHPS/data** directory to **/user/student/DIHPS** in HDFS.
  + Load the **census\_2010.csv** file and write the **POPULATION** records to **c03ex00\_load**. Partial Output of **/user/student/DIHPS/output/c03ex00\_load**
  + Load the **census\_2010.csv** file, and write the results to **c03ex01\_filter**, where **POPULATION** is greater than *50000*.

54684

183216

27336

22879

57344

Partial Output of **/user/student/DIHPS/output/c03ex01\_filter**

54684

183216

57344

118443

50179