HOMEWORK ASSIGNMENT: WEEK 1

Assignment 01:

Complete the following four tasks:

1. Create hdfs directory /inputs. Under this directory create following directories
   a. mr_inputs
   b. pig_inputs
   c. project_inputs

2. Create a simple text file (Let's say write some names, one per line) and save it in /inputs/mr_inputs hdfs directory

3. Try to run in-built word count program with the input you created in Task 2 of this assignment

4. Validate your result through command line

Assignment 02:

Complete the following three tasks:

1. Create a file with some data (about 3 to 5 sentences per line) and copy it to a hdfs location of your choice

2. Write a MapReduce program to count the number of lines in the file created in the previous task (task 1 of this assignment)

3. Write a MapReduce program to remove the records for which the length of the record is less than 15