1. For transferring csv data to hdfs-

hdfs dfs put /home/training/trend\_training/data/Book1.csv /user/training/Book\_input

1. For seeing it-

Hdfs dfs –ls /

3.for putting mapper code into hdfs-

Hdfs dfs –pt /home/training/trend\_training/code/mapper.py

1. Same as the reducer code
2. Make the book\_output folder in home-training
3. Run the code by-

hs mapper.py reducer.py /user/training.Book\_input /home/training/book\_output

to view the output-

hdfs dfs –ls .home/training/book\_output

hdfs dfs –cat /home/training/book\_output/part-00000

open terminal in code folder

Feature 1-word count run and extracting ip

Feature 2-Clusteing of browser in python

Feature 3-k means in size in Hadoop or python