Group HW 1- Due date is 2nd March, 11.59PM

1. Use the file, which I have uploaded on blackboard as an input.
2. Arrange the chapters with most common words in ascending order.
3. Get rid of all the punctuations, i.e., any character other than a to z and space. It will be helpful to use a regular expression to replace them with empty strings.
4. Also, get rid of “a”, “the”, etc. and do not count them.
5. Develop the Java/python Mapper/Reducer code to implement
6. You should submit two files: Mapper.java and Reducer.java . Please print them out and upload on the due date one per group.
7. Also, submit the output file of the final result with the words and their count along with chapter numbers.
8. Plus a description.txt file with a short description about your map/reduce procedure.