

Homework: Write your own MapReduce Job.

Prepare to do your homework by working through these practice exercises:

- Practice 2 Running And Writing Map Reduce
- Practice 3 Running MapReduce in Eclipse
- No, there's nothing to turn-in.

Homework assignment:

- Use the homework1 project in Eclipse on your VM. It contains stubs for your Mapper, Reducer and Driver.
- Write a program that outputs the number of words that begin with each letter.
 - The program should output a letter and a number, such as:
A, 25168
B, 73485
C, 8918
...
Z, 215*
(No, these aren't the "real" values)
 - Do not include punctuation marks or "white" characters - spaces, line-breaks, etc.
In other words, ignore all non-alphanumeric letters.
 - *Ignore case*: Convert all the letters to either upper or lower-case.
- Run your program from Eclipse and use the local shakespeare data in
 - **/home/cloudera/datasets/shakespeare**
 - Output your results wherever you please.
 - Remember, the results will be in a file named: **part-r-00000**

Turning in your homework.

- Login to your account on the official UCSC site at
<https://online.ucsc-extension.edu/xsl-portal>
- Go to **assignments** and submit the following:
 - Your java code for the driver, mapper and reducer.

- Your results: the part-r-00000 file.