

# Java Programming Test

Please write, in Java, an API to read the contents of a plain text file and enable the display of the total number of words, the average word length, the most frequently occurring word length, and a list of the number of words of each length.

Submit your code, along with unit tests, and maven POM file, in a zip file or Github repository, via email along with instructions for its installation and use, along with any assumptions that you have made about what defines a word (using as a basis the rules that can be deduced from the example below).

For example given a file that contains the following text:

Hello world & good morning. The date is 18/05/2016

We would expect the following output:

Word count = 9

Average word length = 4.556

Number of words of length 1 is 1

Number of words of length 2 is 1

Number of words of length 3 is 1

Number of words of length 4 is 2

Number of words of length 5 is 2

Number of words of length 7 is 1

Number of words of length 10 is 1

The most frequently occurring word length is 2, for word lengths of 4 & 5

Note that we will test code submissions on a number of files of various lengths ranging from the above example through to the large books such as the Bible:

([http://janelwashere.com/files/bible\\_daily.txt](http://janelwashere.com/files/bible_daily.txt)).

Clue – Do not forget to consider formatted numbers.

On the code test:

- This is as much about reading and understanding the problem statement as it is about coding.
- We are looking for eloquent, efficient, machine efficient code.
- Use advanced coding features to help with the above.
- The corollary of the above - do not over complicate the solution.
- Make sure the completed version works!

**We prefer submissions via GitHub.**

Once completed, either upload Java files to github and send me the link, or attach your files to email via <https://www.fromsmash.com> or <https://wettransfer.com>