Technical Assignment

QUESTION 1

Please develop a console / GUI application that is used to determine the frequency of each word that appears in a text data file.

I wrote a program in Python 2.7, Technical\_Ass.py, and used ‘romeo-full.txt’ as the data file, to produce the necessary output, see results in “Technical\_Ass\_output.docx”.

Internet sources used:

1) SCRABBLE: https://stackoverflow.com/questions/32235033/creating-a-python-scrabble-function-that-takes-a-string-as-input-and-returns-a-s

2) to test for a numeric value in the text: <https://stackoverflow.com/questions/23183868/why-isnt-isnumeric-working>

Output requested:

Most frequent word: {word} occurred {x} times

Most frequent 7-character word: {word} occurred {x} times

Highest scoring word(s) (according to Scrabble): {word} with a score of {x}

My OUTPUT:

Most frequent word: ' I ' occurred 60 times.

Most frequent 7-character word: ' blessed ' occurred 2 times.

Highest scoring word(s) (according to Scrabble): ' lazypacing ' with a score of 27

QUESTION 2

Please assess the performance of your application using the usual measures. Please write a paragraph or two describing the performance achieved and possible ways you could improve it.

I realized that with more complex words, it may give problems. I ran it with “clown6.txt”, showing when reading more complex words like an internet address, and splitting up a word containing “/” (OR, like typical in an article).