CS 5103 Project Requirements

* Script must be able to count how many times each unique word appears in a document
  + Capitalization shouldn’t matter (i.e. Test and test aren’t two different words)
* Script should allow for spaces, newline characters, and tabs as separators between words
* Script must be able to count the number of lines in the document
  + Lines in this case are defined as each newline character in the document, although empty lines are ignored
* Script must be able to count the number of characters in the document
  + This ignores spaces, newline characters, and tab characters, and counts the letters and any punctuation (i.e. ‘.’, ‘/’, ‘:’, etc.)
* Script should have unit testing to ensure that it works properly
  + There is one unit test for each requirement
  + One to check that the word count is correct, one for the line count, and one for the character count
  + There is also one to ensure that the document is not empty

**Assumptions:**

* The document is not empty (there is a test case to check but the script itself will not return an error if document is empty)
* Punctuation and other non-letter characters are acceptable for the character count
* In the word count, words that are followed by special characters are counted separately from the same word not attached to such character (i.e. test: and test are counted separately)
* The word, line, and character counts are known beforehand for the test cases

**Note:** Ideally the test\_cases.py script would call the Counts class from the project2.py script to be utilized for test cases. I couldn’t get this to work, so instead I just copy and pasted the project2.py script into the test\_cases code; any changes to project2.py must also be made to test\_cases.py (I understand this isn’t ideal).

Repo Link: https://github.com/hjhill99/CS5103