Project name: Use Flesch Score to measure readability

There are two relevant java files: Document.java and BasicDocument.java. Notice that Document is an abstract class. BasicDocument will implement the abstract methods in the Document class using the guidelines below.

Your Tasks:

1. Implement getNumSentences, getNumWords and getNumSyllables following the comments about how they are supposed to work that you will find in Document.java and BasicDocument.java. You'll need to implement countSyllables(String) in Document.java, to be called in getNumSyllables.

You must follow the definitions of what constitutes a syllable, word and sentence given in the documentation for each function exactly to pass the graders. We have provided you with several test cases in main as well as a helper method you can use to write your own tests.

Hints and additional information: Notice the helper method getTokens in the Document superclass. It takes a string which is a regular expression representing the regular expression of the “tokens” it’s trying to find. Consider passing a different pattern for each method. It is possible to write regex's to count the number of syllables directly, but this is not necessarily the best approach. If you find yourself doing mental gymnastics to come up with a single regex to count syllables directly, that's usually an indication that there's a simpler solution (hint: consider a loop over characters). Just because a piece of code (e.g. a regex) is shorter does not mean it is always better.

Under the rules for what defines a syllable, the words “the”, “fly”, “yes”, “cave” and “double” all have 1 syllable, but "segue" has two syllables. Notice that this isn’t exactly correct (“double” actually has 2 syllables), but it’s close enough for our purposes. Here are some more examples with the number of syllables your method should return to help you: "contiguous" (3 syllables), "sleepy" (2 syllables), "obvious" (2 syllables), "toga" (2 syllables). Notice that our rules get a lot wrong, especially when you have more than 2 vowels in a row, but these are the rules we will test you against.

Consider making a helper method in the Document class (i.e. the superclass) that counts the number of syllables in a single word. We have a stub method countSyllables(String) already in Document.java. Just use a loop to loop through the characters in the string and write your own logic for counting syllables.

2. Implement the getFleschScore method in Document.java. Fill in the method getFleschScore() in Document.java to calculate the Flesch Score for the text in the document. You should use the following formula, and make calls to the getNumSyllables, getNumWords, and getNumSentences you just implemented.

$$Flesch\ score = 206.835\ - 1.015\left(\frac{\#\ words}{\#\ sentences}\right) - 84.6\left(\frac{\#\ syllables}{\#\ words}\right)$$

You should test your code by calculating the Flesch score by hand on some very basic documents and then calling your method from main to make sure it's giving the same output.

What and how to submit

Create a zip file containing only Document.java and BasicDocument.java. You can find these files in the workspace directory you set up on your computer's filesystem. Please submit the zip file to Blackboard.