Your assignment is to create a linguistics and/or word processing application. For the first part of the assignment, you will discover the type of vocabulary used in a book/document by calculating the average length of the words in the text. You can find plain text copies of books/documents/speeches/articles/etc online. For example, one good place to find books in plain text is at www.gutenberg.org, although any way you can generate/find plain text files is fine.

For the rest of the assignment, think about interesting questions that you can ask about the texts or interesting operations that you can perform on the texts. For example, you could ask questions about style (does Jane Austen use longer sentences on average than James Joyce or are her novels on average shorter), vocabulary (word length, the number of distinct words used), how frequently certain words occur, counting compound sentences, counting sentences that begin with a pronoun, counting occurrences of specific phrases, etc. You could also perform word processing operations on the texts such as making the work gender-neutral or changing all female references to male ones, spelling out numerical references, doing find-and-replace according to the user’s wishes, etc. The purpose is up to you but think of applications that might be useful, interesting, or might tell you something about the linguistic style. Perhaps you can relate your program to another interest of yours (academic or otherwise), or something you’ve always wondered about your favorite story/book! One year a student did their project on the Harry Potter series of books, which they somehow acquired in plain text form...

Your program will start out by offering the user at least 2 menu options. A sample opening menu might look like this:

A. Calculate average word length of a book/document.
B. <A linguistics/word processing task you come up with.>
C. <Perhaps another linguistics/word processing task you come up with.>

At least one of the menu options should incorporate the task of calculating average word length. (The menu option in your program that incorporates this task could be something more elaborate, so your menu will not necessarily follow the format of the above sample menu). The rest of the program’s function and design is up to you. Just be sure to do something different in some way from what we’ve already done in class/lab, rather than simply re-purposing the code we’ve already written.

Your program should probably be general enough to be easily modified for other input files (not just the books you choose), though exceptions to this rule can certainly be made, for example if the questions you want to answer are very content-specific... please discuss with me if you think your idea should be an exception. Your program should include at least the following:

- Use of at least two texts
- Information that is gleaned from processing the entire texts
- Output that is clearly explained to the user
- Use of string functions
- Use of conditionals (if statements)
• Comments at the beginning of the module and throughout the code
• An interesting/useful/fun application in word processing or linguistics or literary analysis

Your submission folder should include the module (with the .py ending) along with any text files it uses. Your files should be placed in a folder called proj2_xxxxx, where xxxxx is your login name. Zip and upload this file to moodle (for later download and testing by the TAs when you go in for your demo).