**HW 5B**

**Learning Experience**

Part 1:

The first part of the assignment was a tricky program. The program was about to read a text file .txt and then process the file and determine the total number of sentences, words, and characters. To do so, I just needed to open the file, read the file, and store the whole text in a string.

Graphical user interface

Description automatically generated

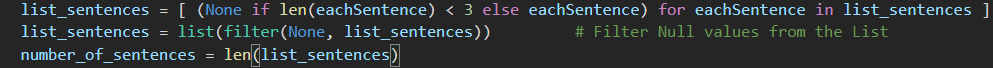
Once I have the whole text in a string, I needed to break each sentence up and store it in a list; however, the text was compound with symbol characters that makes a new sentence . Therefore, what I needed was to split the sentences with multiple delimiters and to do that I had to import “re” from python library which allows me to split the text with multiple delimiters.



After the code above, I was able to get whole sentences, but some of those sentences had leading spaces and trailing spaces and also had some \n as a string inside the sentences. So, I implemented the built-in function strip() to get rid of leading and trailing spaces, and the replace() function to replace \n strings to a empty space. To illustrate, the following code does apply:



However, there were still more issues with formatting the sentences. Since I broke the sentences up with the delimiters (‘? . :’), there were some words that had these delimiters and cause to be split on. For example, words like “http://www.py4e.com/code/”, or “.py” have dots in it, so my program split them into sentences. I could not find any solution to only split when the dot is at the end of the sentences, but what I did was to remove all those sentences that have only 2 words since sentences cannot be composed of 2 words.



For the list of words, I implemented the same way as the previous assignment last week. Some words from the list had special characters. I used the following methods to clean up the words.



However, there were still some words having contractions like “he’ll” or “we’ll”, so I also deleted these in order to have the single word itself. To do so, I used the built-in function .replace() in order to replace ’ll from the words:

Text

Description automatically generated with medium confidence

The next step was to get rid of the duplicate words. So, I decided to do a for-loop. I used the membership operator “in” to check if the word is present in the no\_duplicate\_list. So, if the word is not in the no\_duplicate\_list, it will append the word to the no\_duplicate\_list, otherwise will not append.

Text

Description automatically generated

For the list of characters, I did a for loop to go through each words from the list of words, and count the length of each word as characters. The following code demonstrates:

Text

Description automatically generated

Part 2:

The second part of the assignment was the easier one. I needed to implement a function to determine the total purchase price (of the portfolio) and implement another function to determine the profit margin of the portfolio.

First of all, I processed the file .csv by using the module csv, so I can read the csv file and get the whole portfolio in a list of Lists.

To do the functions, I had to use a for loop to access the dataList (table); however, the first row of the table is the header and since I only need to loop from the data, I used the following code to loop from a certain row of the table:



Where [1:] means to start the loop from (index 1 = second row) from the table.

The next step is to find the desire column from each row to retrieve the data. However, the data stored in the table have the following format: ‘ $4.55 ’. In order to get rid of the leading and trailing spaces, I did the following code:



After this code, the new data will be ‘$4.55’. To get rid of the symbol ‘$’, I used the built-in method lstrip**(‘character’**), which means that will strip from the left side of the string the desire **‘character’**



Now, I have a number but in a string variable. To convert from a string to a number, I used the float() function.



Finally, I added this number in an accumulative variable that will be the total purchase price



Likewise, I implemented in the same way for the function of profit margin.

The enhancement for this program was that I formatted the float number. I used the format() function from the str class to format the float number into two decimal places.

Overall, the assignment was fun to code and useful to sharp my programming skills.

**Part 1 Test 1**

**Text

Description automatically generated**

**Part 2 Test 1**

**Text

Description automatically generated**