

Homework 1

Ryan Reynolds

ryreynol

Due: 9/13/18

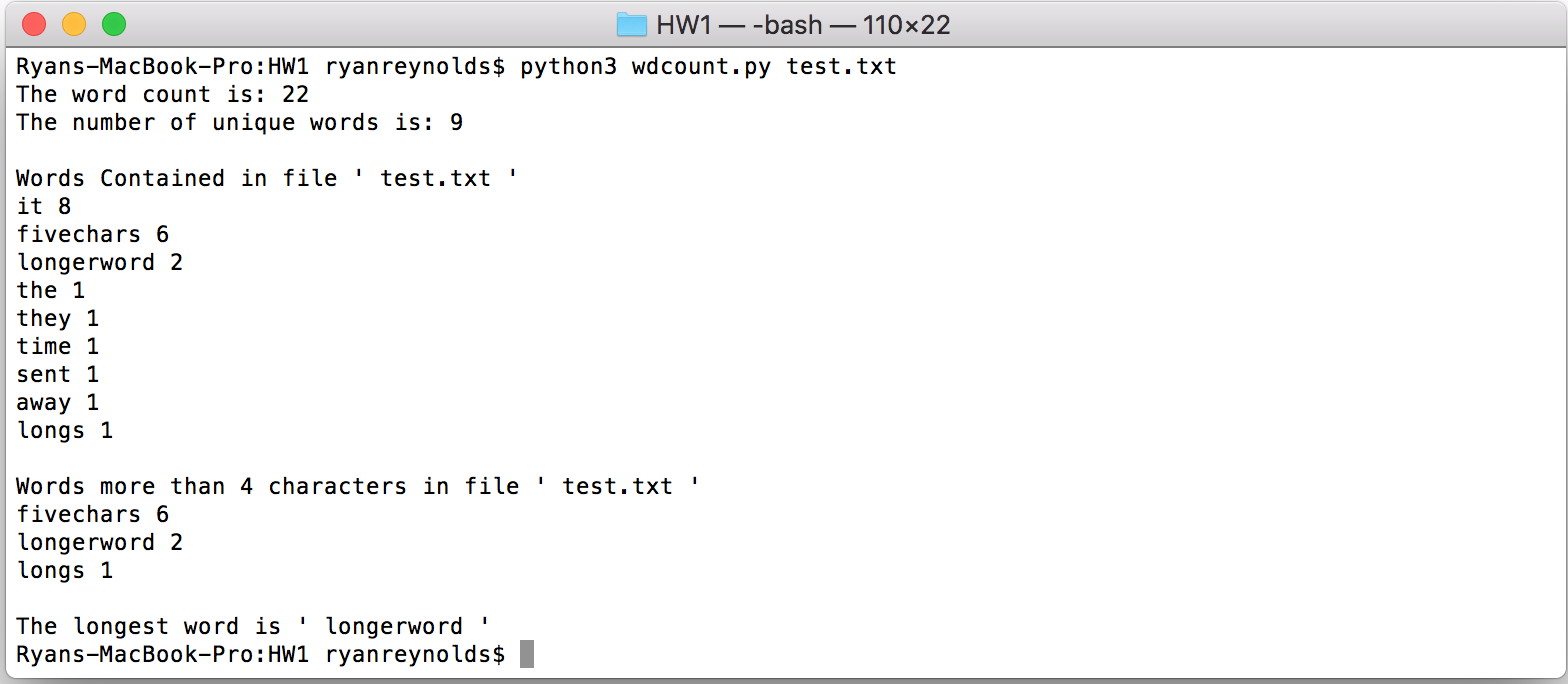
Code Description and Debugging Report:

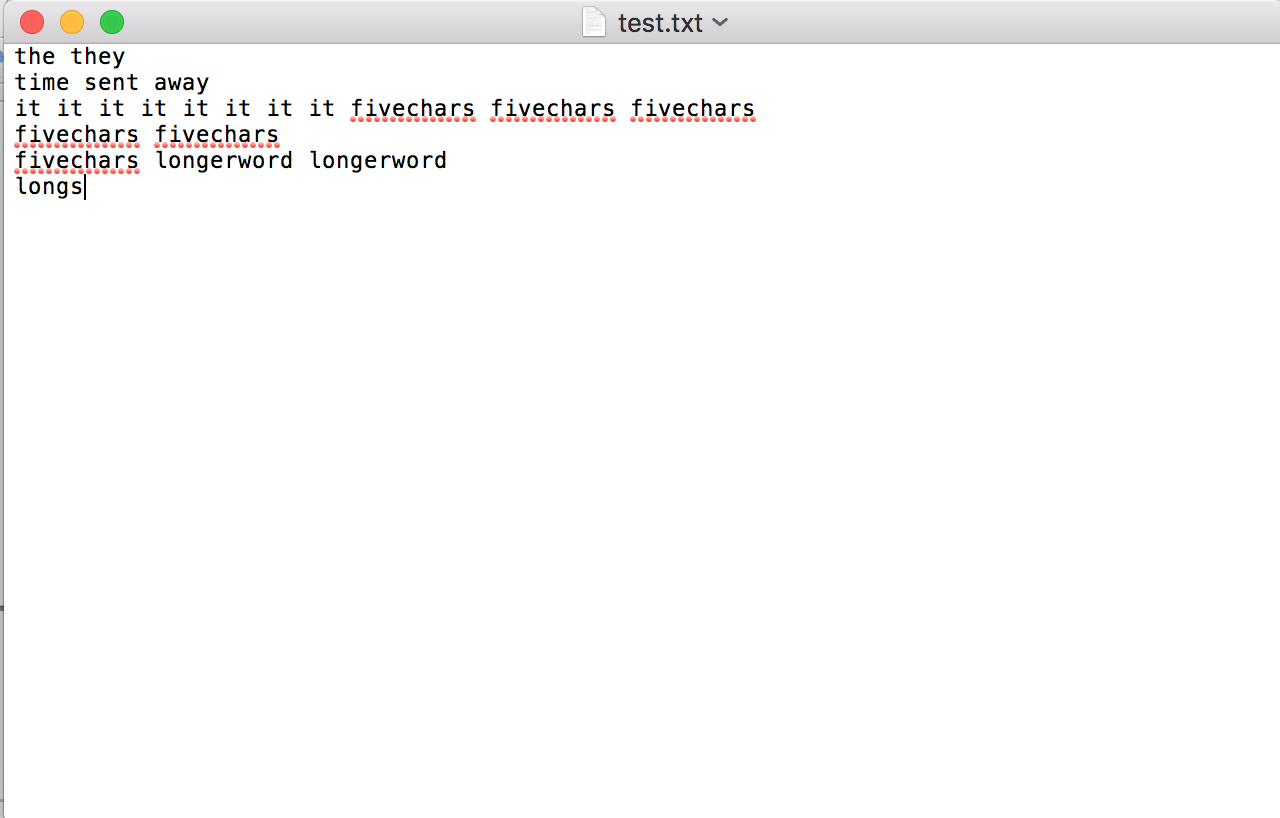
The attached code reads in a text file word by word from the command line. The program outputs the total number of words and the number of unique words. Then the program lists each unique word and its number of occurrences in ascending order. After, the program lists each word with more than four characters and its number of occurrences in the file. Finally, the program lists the longest word in the file.

Two dictionaries-wdict and longwdict- were implemented to create the above functionality. wdict stores all of the unique words in the file and counts their occurrences using a for loop. longwdict stores all of the words longer than 4 characters. The dictionary is populated within for loop that sorts and prints the wdict dictionary. Each word is converted to lowercase to prevent false unique words due to capitalization. If the word being sorted is longer than 4 characters the word is stored as a key in longwdict and its occurrences as the value. The total word count is calculated by summing all of the values stored in wdict. The number of unique words is calculated by taking the length of the wdict dictionary. The lognest word was determined by checking the length of each word as it is read into the file. A slightly more efficient way would be to apply this within the sorting loop instead of the file.read().split() loop.

The program was tested by creating a sample test file, the output and test file are provided below. The only logic bugs that occurred were formatting errors with the print statements. Specifically, spacing in between sorted words. All of the debugging were syntax related errors. Specifically, forgetting the “:” at the ends of if and for loop statements.

Output:

**

Input File:

Source Code:

