Vanessa Ulloa

Co Sci 436

Final Project

Analysis

# PROBLEM

Write a Java program which accepts the name of a text file, reads in that file, and prints out the following information about the file:

INPUT

1. Text file

OUTPUT

1. Frequency of all Letters (A to Z) in the file
2. Frequency of word use in the file
3. Average number of letters in a word
4. Average number of words in a sentence

# ANALYSIS

1. The class with the main () method will be called FinalProject. This will use Scanner to read in a file name into a String variable, or file name can be assigned to the String variable for testing.
2. A 2nd class will be created that will Read the file and parse the lines into a String Array, this is done to ensure that all data is capture from the text file and not just 1 line into 1 string.
3. A Try/Catch is used to create an instance of the 2nd class and pass the String array onto a Count Results method
4. The Count results method will do all the operations needed to get the results for this problem.
   1. The string array that was created earlier will have all of its elements (lines) concatenated onto 1 string so that we can work with it better.
   2. A new String array is created by splitting the 1 string based on a space deliminator
   3. The count all the sentences we will look for punctuation (.,!,?) and increment an integer variable to keep track of that.
   4. To count all the characters, we loop through the String array and count the length of each element which counts each word
   5. To count all the instances of an Alphabet character we will loop through the String array with all the words and then loop through each word. The outer loop loops through the array and the inner loop then loops through each character in the word. Each character will be compared to the Alphabet in the 3rd (2nd inner loop) and the corresponding value in an integer array will be incremented to keep count.
   6. The Alphabet Array and the integer array used to keep count will be sorted using an insertion sort. If the value is less than the next one, then rearrange the values until we get values in decreasing fashion
   7. The String array with all the words will be copied, these arrays will be compared against each other and a variable will be used to keep track if it appears more than once. This is used to remove duplicates from the 2nd string array in order to have an array to use for comparison against all the words. A final string array is created with all the words listed once and without duplication for count and comparison—similar to the alphabet array and count.
5. Final Method is created to display the results.
   1. Average number of letters is determined by dividing character count by the length of the String Array holding all the words
   2. Average number of words is determined by dividing the length of the String array holding all the words and the number of sentences. 8