

Homework #4: 統計單字數量和字母出現次數

繳交日期與繳交規定：

請將作業 **turnin** 到 **PD.hw4**

Due Date: 2016/10/16(日)23:59:59

請在 **due date** 前將作業 **turnin**, 超過即算遲交; 遲交請 **turnin** 到 **PD.hw4.delay**

作業需在 **due day** 後兩星期之內 **demo** 完成, 也就是 **10/30(日)** 之前要完成 **demo**

若對繳交規定有任何疑問請來信至 pdta@gais.cs.ccu.edu.tw

測試環境

系上工作站

作業說明

【A】統計單字

請撰寫一個程式，從標準輸入讀取測資，一次讀一個字元直到檔案結尾，必須統計英文單字數量，並印出統計結果及每個單字(一行一個單字)。

【B】統計字母

請撰寫一個程式，從標準輸入讀取測資，一次讀一個字元直到檔案結尾，必須統計字母總數、每個字母(a~z) 的出現次數和出現頻率，並計算頻率的平均值、找出出現頻率最大及最小的字母，將結果依範例格式輸出。

作業提示

1. 從標準輸入讀取測資的方法可以使用 `cat hw4_test.txt | ./a.out`，便可將 `hw4_test.txt` 讀進執行檔。
2. 統計英文單字數量，不包含以數字開頭的單字，但包含字串中間或結尾有數字的單字，**ex.** 不包含 **99apple**，但包含 **app100le** 及 **apple200**。
3. 統計字母時，不分大小寫。
4. 出現頻率 = 出現次數 / 字母總數。
5. 若有多個字母都是 **Max** 或 **min**，則印出第一個最大或最小的即可，**ex.** **c, n, q** 的出現頻率皆為 **0.226**，都是最大值，則印出 **c** 即可。
6. 建議使用 **array** 來實作。

程式要求

1. 出現頻率須輸出並四捨五入到小數點後第三位。

作業配分

【A】

1. 統計英文單字數量並印出結果(30%)
2. 印出每個單字(一行一個單字)(20%)

【B】

1. 統計字母總數、每個字母的出現次數及頻率(30%)
2. 計算頻率的平均值(5%)
3. 找出出現頻率最大及最小的字母並印出(包含字母、次數及頻率)(15%)

輸出結果範例

【範例測資】

There are 3Apples.

There is an orange1.

ORange1 is sweet.

My phone number is 0912-345-678.

【A】

【輸出結果】

Word count: 13

--Words--

There

are

There

...

orange1.

ORange1

...

is

【B】

【輸出結果】

Total : 57

--Alphabet--

a : 5 , 0.088

b : 1 , 0.018

...

z : 0 , 0.000

Avg = 0.038

Max: e : 12, 0.211

Min: c : 0, 0.000

Bonus

【A】 (10%)

找出長度最長的單字，並印出其長度。