**Edit Distance**

Kerja kelompok 2-3 orang. Bobot nilai 50% dari nilai akhir. Tanggal kumpul 30 November 23.55.

Upload isi dokumen yang diattach [ini](http://learning.surya.ac.id/pluginfile.php/7828/mod_assign/intro/hound).

1. Konversi kata-kata dan huruf  
a. Selain angka 0-9 dan huruf a-z atau ', hapuslah semua tanda baca  
b. Ubahlah semua huruf ke huruf kecil  
Jadi kalimat-kalimat seperti:  
"The exact date is 1742." Dr. Mortimer drew it from his breast-pocket.

Akan dijadikan kata-kata  
the  
exact  
date  
is  
1742  
dr  
mortimer  
drew  
it  
from  
his  
breast  
pocket

Kalimat seperti:

"A man's or a woman's?"

Akan dijadikan kata-kata

a  
man's  
or  
a  
woman's

2. Hitung frekuensi kata-kata tersebut. Kata dan frekuensi bisa disimpan di array atau map.

3. Ingat tutup file setelah selesai dibaca.

4. Masukkan sebuah kata dan edit distance

Misalnya

Masukkan kata:

hound

Masukkan edit distance (minimum 0)

2

5. Pergunakan Algoritma Edit Distance. Sort hasil berdasarkan edit distance menggunakan algoritma QuickSort atau MergeSort.

6. Ouput data harus benar. Urutkan output berdasarkan edit distance dulu baru urutan abjad  
Kalau user memasukkan edit distance 2 untuk kata hound, maka semua kata dengan edit distance maksimum 2 (termasuk 0, 1, 2) akan diprint out beserta edit distancenya. Kolom terakhir berisi frekuensi kata tersebut. Perhatikan yang di bawah ini cuma contoh, bukan hasil yang persis.

hound 0 76  
bound 1 5  
found 1 47  
hounds 1 4  
pound 1 2  
round 1 46  
sound 1 25  
wound 1 4  
around 2 10  
ground 2 17  
house 2 73  
pounds 2 6

Kalau user memasukkan input yang tidak mempunyai match di edit distance, tulis  
"No Match Found"

7. Ingat untuk menyertakan reference untuk bahan yang didapat dari internet, biarpun algoritma maupun source code. Usahakan program bisa dicompile di Linux atau cygwin.

8. Tuliskan user manual bagaimana cara menggunakan program tersebut.

Bobot nilai (105 jadi ada bonus nilai 5).  
-Membuka, membaca, dan menutup file dengan benar (10%)  
-Error Checking (10%)  
-Map atau Array (5%)  
-Edit Distance (30%)  
-Quick Sort/Merge Sort (20%)  
-Output benar (10%)  
-User manual (20%)

Kalau telat lebih semenit sampai sehari, nilai jadi 50%.