

DASAR-DASAR PEMROGRAMAN 1 (CSGE601020)

Semester Gasal 2019/2020

Tugas Pemrograman 2

Deadline: Selasa, 15 Oktober 2019 jam 23:55

Petunjuk Umum:

- Submit program Anda di Scele sebelum deadline yang ditentukan dengan format nama file: *TP2_<nama>_<NPM>_<kode_asdos>.py*
- Tugas terlambat tidak diterima
- Plagiarisme, nilai E

Petunjuk Khusus

WORD COUNTER

Anda diminta membuat program untuk menghitung distribusi kemunculan kata pada sebuah himpunan pesan teks.

Input Program

Program akan meminta user untuk memasukkan beberapa pesan teks. Proses input akan berhenti ketika user memasukkan pesan kosong. Contoh interaksi dengan user:

```
=====
Masukkan pesan: (untuk berhenti masukkan string kosong)
=====
Pesan: Halo, Bunga. Apa kabar kamu? Apakah baik-baik saja??
Pesan: Baik! Apakah kamu sudah mengerjakan TP2 (DDP-1) !?
Pesan: Belum! Dan kamu?
Pesan: Aku juga belum, hehe... Masih 'sibuk' !
Pesan: Ayo belajar bersama, katanya kita harus mempelajari library Matplotlib untuk
menggambar grafik.
Pesan: Oh, ya? Menarik sekali! Kamu sudah membaca soal tampaknya ;- )
Pesan:
```

Apa yang dikerjakan oleh Program Anda:

1. Setiap pesan (berupa string) yang dimasukkan oleh user harus dimasukkan ke dalam sebuah *list of String*.
2. Kemudian setiap string harus **ditokenisasi** (dipecah menjadi token-token) untuk mendapatkan daftar kata. Perhatikan bahwa Anda perlu:
 - Membuang tanda baca, kecuali tanda baca pada *reduplicated word* seperti anak-anak, bersusah-payahContoh:
 - Siapa? → siapa, Ya!!! → ya, (rahasia) → rahasia
 - Anak-anak! → anak-anak
- Membuang kata-kata yang termasuk ke dalam kata penghubung dan preposisi. Daftar kata yang termasuk kategori ini sudah disediakan dalam

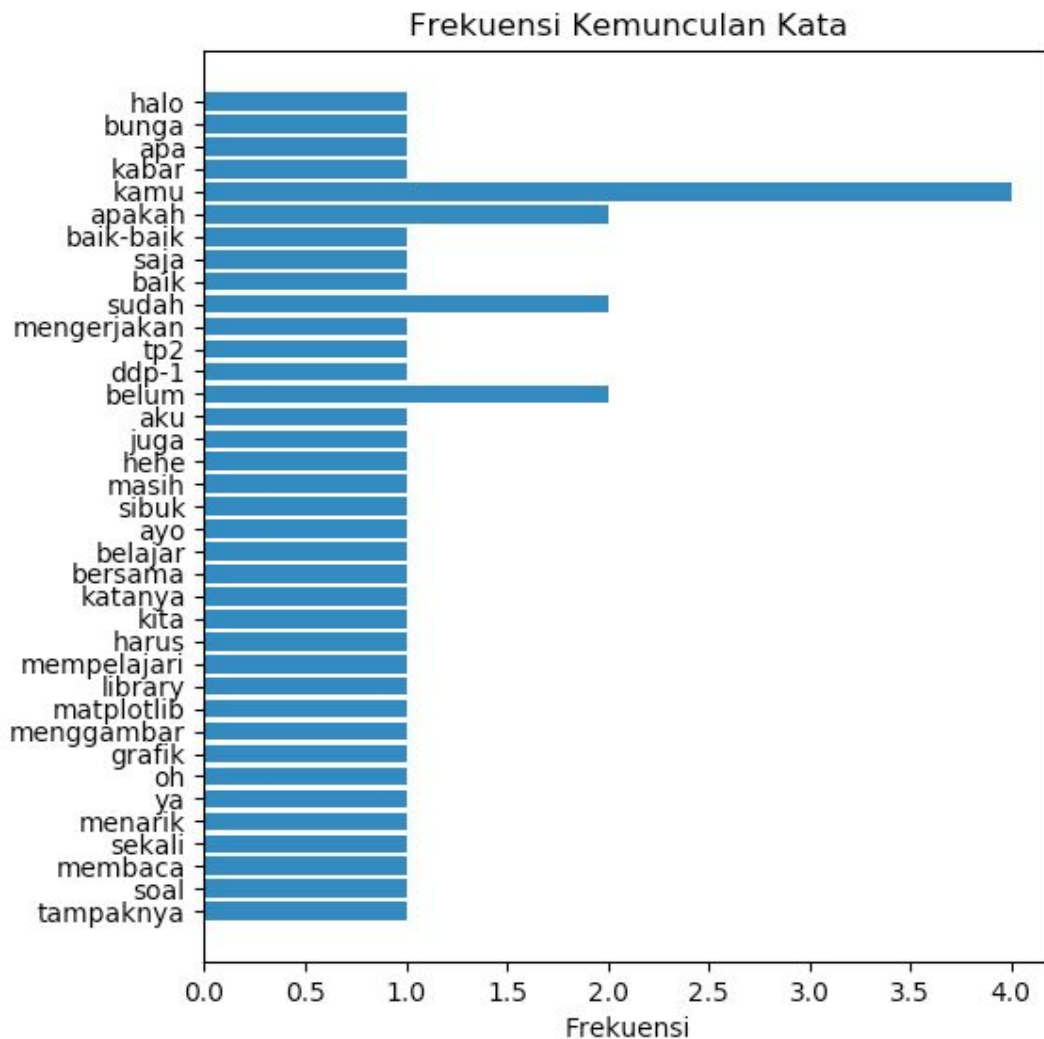
sebuah file bernama **TP2-stopword.txt**. Buatlah fungsi untuk menyalin isi file ini ke dalam sebuah list.

3. Setelah itu hitung kemunculan setiap kata. Perhatikan bahwa kata “Kamu” dan “kamu” dianggap sebagai kata yang sama.
4. Tampilkan hasil perhitungan distribusi kemunculan kata dalam 2 cara:
 - Dalam bentuk tabel
 - Dalam bentuk diagram batang. *Anda boleh menampilkan diagram batang yang standard dengan batang yang vertikal, atau menampilkan batang secara horizontal seperti dicontohkan pada dokumen tugas ini.*

Output Program:

Distribusi frekuensi kata:

No	Kata	Frekuensi
1	halo	1
2	bunga	1
3	apa	1
4	kabar	1
5	kamu	4
6	apakah	2
7	baik-baik	1
8	saja	1
9	baik	1
10	sudah	2
11	mengerjakan	1
12	tp2	1
13	ddp-1	1
14	belum	2
15	aku	1
16	juga	1
17	hehe	1
18	masih	1
19	sibuk	1
20	ayo	1
21	belajar	1
22	bersama	1
23	katanya	1
24	kita	1
25	harus	1
26	mempelajari	1
27	library	1
28	matplotlib	1
29	menggambar	1
30	grafik	1
31	oh	1
32	ya	1
33	menarik	1
34	sekali	1
35	membaca	1
36	soal	1
37	tampaknya	1



Petunjuk Teknis:

- Tuliskan program secara modular dengan membuat fungsi-fungsi yang mengerjakan hal yang spesifik.
- Anda **belum** diperbolehkan menggunakan struktur data **Dictionary** untuk menghitung frekuensi kemunculan kata
- Gunakan library **matplotlib** untuk menggambar grafik.

Test Case

Untuk menguji program Anda, masukkan string-string berikut ini:

Halo, Bunga. Apa kabar kamu? Apakah baik-baik saja??
 Baik! Apakah kamu sudah mengerjakan TP2 (DDP-1) !?
 Belum! Dan kamu?
 Aku juga belum, hehe... Masih 'sibuk' !
 Ayo belajar bersama, katanya kita harus mempelajari library Matplotlib untuk menggambar grafik.
 Oh, ya? Menarik sekali! Kamu sudah membaca soal tampaknya ;-)

Komponen Penilaian “Correctness”:

1. Bisa menyimpan input string ke dalam list dengan benar (10%)
2. Bisa menyalin isi file TP2-stopword.txt ke dalam list (10%)
3. Bisa melakukan tokenisasi dan mem-filter tanda baca dan stopwords dengan benar (40%)
4. Bisa menghitung frekuensi kata dengan benar (20%)
5. Bisa menampilkan frekuensi kata dalam bentuk tabel dengan rapi (10%)
6. Bisa menampilkan frekuensi kata dalam bentuk grafik dengan rapi (10%)