**ALGORITMA DAN STRUKTUR DATA**

**MODUL 7**

**REGULAR EXPRESSIONS**



**Disusun oleh:**

Muhammad Ferizal Fadhli

L200210119

D

**TEKNIK INFORMATIKA**

**FAKULTAS KOMUNIKASI DAN INFORMATIKA**

**UNIVERSITAS MUHAMMADIYAH SURAKARTA**

**2022/2023**

Soal Soal Mahasiswa

1. Meng-ekstrak kata-kata dengan awalan ‘me’. Bukalah halaman web id.wikipedia.org/wiki/Indonesia. Salin seluruh tubuh artikel itu dan tuangkan ke dalam notepad. Simpan sebahai Indonesia.txt. Tugasmu adalah mengekstrak kata yang mempunyai awalan ‘me’, lalu menyimpannya dalam sbuah tuple. Perhatikan bahwa ‘me’ bisa muncul di awal tengah ataupun akhir kalimat.

Berikut Kode Program :

import re

isi = ()

f = open("Indonesia.txt", "r", *encoding*="latin1")

teks = f.read()

p = r"me\w+"

string = re.findall(p, teks)

p2 = r"\w+me\w+"

string2 = re.findall(p2, teks)

p3 = r"\w+me"

string3 = re.findall(p3, teks)

isi = string + string2 + string3

print(isi)

Berikut SS Ketika Program Dijalankan :



1. Meng-ekstrak kata-kata dengan awalan ‘di’. Sama seperti di atas, tapi yang kamu ekstrak adalah kata kata dengan awalan ‘di’. Perhatikan yang dimaksud disini adalah awalan ‘di’ (menuntukkan kata kerja pasif), bukan kata depan ‘di’ (menunjukkan keterangan tempat), yang kita bahas pada nomer dibawah.

Berikut Kode Program :

import re

isi = ()

f = open("Indonesia.txt", "r", *encoding*="latin1")

teks = f.read()

p = r"di\w+"

string = re.findall(p, teks)

isi = string

print(isi)

Berikut SS Ketika Program Dijalankan :



1. Meng-ekstrak kata depan ‘di’ dan tempat yang ditunjukkan. Di sini tugasmu adalah mengkestrak masih pada berkas Indonesia.txt – kata depan ‘di’ beserta satu kata yang mengikutinya.

Berikut Kode Program :

import re

isi = ()

f = open("Indonesia.txt", "r", *encoding*="latin1")

teks = f.read()

p = r"di \w+"

string = re.findall(p, teks)

isi = string

print(isi)

Berikut SS Ketika Program Dijalankan :



1. Mengekstrak data dari berkas html. Tugamu di sini adalah membuat sebuah list of tuples yang memuat nama-nama negara innovation index-nya. Kunjungilah halaman en.wikipedia.org/wiki/Knowladge\_Economic\_index. Perhatikan halaman itu memuat sebuah besar hasil penelitian.

Berikut Kode Program :

import re

f= open("KEI.html", "r", *encoding*="latin1")

teks = f.read()

p = r'([\w+]+)</a></td>\n<td>([0-9.]+)</td>' *#mencari pola*

string = re.findall(p, teks) *#mencari string*

list = []

for i in string:

    a = (i[0], float(i[1]))

    list.append(a)

print(list)

Berikut SS Ketika Program Dijalankan :

