ALGORITHM AND DATA STRUCTURE PRACTICUM MODULE 7 REGULAR EXPRESSIONS



CREATED BY :
KURNIAWAN BAGASKARA
L200214253

INFORMATICS STUDY PROGRAM FACULTY OF COMMUNICATION AND INFORMATION SCIENCE MUHAMMADIYAH SURAKARTA UNIVERSITY

```
import re
s = 'sebuah contoh kata:teh!!'
cocok = re.findall(r'kata:\w\w\w', s)
if cocok:
  print('menemukan', cocok)
else:
  print('tidak menemukan')
# 'menemukan [kata:teh]'
#LATIHAN 7.
import re
s = 'sebuah contoh kata:batagor!!'
cocok = re.findall(r'kata:\w\w\w', s)
if cocok:
  print('menemukan', cocok)
else:
  print('tidak menemukan')
## 'menemukan [kata:bat]'
import re
s = 'sebuah contoh kata:es teh!!'
cocok = re.findall(r'kata:\w\w\w', s)
if cocok:
  print('menemukan', cocok)
else:
  print('tidak menemukan')
#tidak menemukan karena tidak ada 3 huruf sebelum spasi
```

EXERCISE:

```
import re
cocok1 = re.findall(r'eee', 'teeeh') #mencari huruf eee
cocok2 = re.findall(r'ehs', 'teeeh') #mencari huruf ehs
cocok3 = re.findall(r'..h', 'teeeh') #mencari 2 huruf sebelum huruf h
cocok4 = re.findall(r'\d\d\d', 't123h di 2019 bulan 02') #mencari angka dengan per kumpulan
angka sebanyak 3 angka
cocok5 = re.findall(r\w\w\w', '@@a*bc#def*tghh!!') #mencari huruf dengan per kumpulan
huruf sebanyak 3 huruf
##untuk yang \d\d\d maupun \w\w\w merupakan banyak kumpulan huruf yang dapat dicari
jadi apabila kumpulan huruf lebih
##dari 3 huruf maka akan dicari dan disimpan hanya 3 huruf
##apabila kurang dari 3 huruf maka tidak akan disimpan
print(cocok1, cocok2, cocok3, cocok4, cocok5)
print ('LATIHAN 7.3.1')
import re
cocoka = re.findall(r'e+', 'ghdteeeh')
cocokb = re.findall(r'e+', 'teeheeee')
polanya = r' d s* d s* d'
cocok1 = re.findall(polanya, 'xx1 2 3xx')
cocok2 = re.findall(polanya, 'xx12 3xx')
cocok3 = re.findall(polanya, 'xx123xx')
cocok4 = re.findall(r'^k\w+', 'mejakursi')
```

```
cocok5 = re.findall(r'k[k\w\s]+', 'mejakursi tamu saya')
print(cocoka, cocokb, cocok1, cocok2, cocok3, cocok4, cocok5)
print ('LATIHAN 7.3')
import re
s = 'Alamatku adalah dita-b@google.com mas'
cocok6 = re.findall(r'\w+@\w+', s)
print(cocok6[0]) #output : b@google
print ('LATIHAN 7.3.2')
import re
cocok = re.findall(r'[\w.-]+@[\w.-]+', s)
print(cocok[0]) #output : dita-b@google.com
import re
cocok = re.findall(r'[a-zA-Z0-9_.+-]+@[a-zA-Z0-9-]+\.[a-zA-Z0-9-.]+', s)
print(cocok[0]) #output : dita-b@google.com
print('LATIHAN 7.4')
import re
s = 'Alamatku adalah dita-b@google.com mas'
cocok = re.findall(r'([\w.-]+)@([\w.-]+)',s)
print(cocok[0][0]) #output : dita-b
print(cocok[0][1]) #output : google.com
```

import re

s = 'Alamatku sri@google.com serta joko@abc.com ok bro. atau don@email.com bisajadi babejo@gmail.com atau lalalong@mail.com'

```
pola = r'[\w\.-] + @[\w\.-] +'
e = re.findall(pola, s) #mengembalikan list beranggotakan string alamat
print(e) # output : ['sri@google.com', 'joko@abc.com', 'don@email.com']
pola2 = r'(\lceil w \cdot - \rceil +) @(\lceil w \cdot - \rceil +)'
e2 = re.findall(pola2, s)
print(e2) # output : [('sri', 'google.com'), ('joko', 'abc.com'), ('don', 'email.com')]
for tup in e:
  print('user', tup[0], 'dengan host: ', tup[1])
#analisa:program hanya akan mengambil dan memproses data yang didalamnya tergabung
dan ditengahnya ada tanda {@}
##output:
##user s dengan host: r
##user j dengan host: o
##user d dengan host: o
# LATIHAN 7.5
import re
s = 'Alamatku sri@google.com serta joko@abc.com ok bro. atau don@email.com bisajadi
babejo@gmail.com atau lalalong@mail.com'
```

e = re.findall(pola, s) #mengembalikan list beranggotakan string alamat

 $pola = r' \lceil \langle w \rangle - \rceil + @ \lceil \langle w \rangle - \rceil + '$

 $pola2 = r'([\w\.-]+)@([\w\.-]+)'$

e2 = re.findall(pola2, s)

print(e)

```
print(e2)
for tup in e:
  print('user', tup[0], 'dengan host: ', tup[1])
##output:
##['sri@google.com', 'joko@abc.com', 'don@email.com', 'babejo@gmail.com',
'lalalong@mail.com']
##[('sri', 'google.com'), ('joko', 'abc.com'), ('don', 'email.com'), ('babejo', 'gmail.com'),
('lalalong', 'mail.com')]
##user s dengan host: r
##user j dengan host: o
##user d dengan host: o
##user b dengan host: a
##user l dengan host: a
        mukan
| ['eeh'] ['123', '201'] ['def', 'tgh']
3.1
|ee', 'eeee'] ['1 2 3'] ['123'] [] ['kursi tamu saya']
Task:
1.import re
a = open('Indonesia.txt','r')
baca = a.read()
a.close
cocok= re.findall(r"menu\w+",baca)
print(cocok)
print("(Double Klik)")
print ("Kurniawan Bagaskara")
print ("L200214253")
```

```
Shell:

>>> XRun nol.py
['menuduh', 'menunjuk', 'menunjukkan', 'menurut', 'menurut', 'menurut', 'menurut']
[Obuble Klik)
Kurniawan Bagaakara
L200214253

>>> XRun nol.py
['disn.' (disn.' (dis
```

```
2.import re

a = open('Indonesia.txt','r')

baca = a.read()

a.close

cocok= re.findall(r"di\w+",baca)

print(cocok)

print(" ")

print ("Kurniawan Bagaskara")

print ("L200214253")
```

3.import re

a = open('Indonesia.txt','r')

```
baca = a.read()
a.close
cocok= re.findall(r"di\s\w+",baca)
print(cocok)
print(" ")
print ("Kurniawan Bagaskara")
print ("L200214253")
     lose
ok= re.findall(r"di\s\w+",baca)
nt(cocok)
4.import re
f = open('KEI.html','r',encoding='latin1')
teks = f.read()
f.close()
\#pola5 = r'(w+) </a >  (d.\d+)'
##tupless = re.findall(pola5,teks)
##print(tupless)
```

print ("Kurniawan Bagaskara")

print ("L200214253")

```
Shell:

>>> numd = re.findall(pola,teks)

*Traceback (most recent call last):

*Traceback (most recent call last):

*Indicate the state of the state
NameFror: name 'pola' is not defined

>>> pola4 = r'(\w+\/a>c/Kd)'

>>> nua4 = re.findall(pola4,teks)

>>> pola4 = re.findall(pola4,teks)

>>> pola5 = re.findall(pola4,teks)

>>> print (num4)

['Denmark', 'Sweden', 'Finland', 'Netherlands', 'Norway', 'Canada', 'Switzerland', 'Kingdom', 'States', 'Australia', 'Ireland', 'Austria', 'Iceland', 'Germany', 'Zealand', 'Selgium', 'Taiwan', 'Latvia', 'Cpyrus', 'Fortugal', 'Greece', 'Foland', 'Slovakia', 'Barbados', 'Croatia', 'Chile', 'Bulgaria', 'Emirat es', 'Komania', 'Uruyay', 'Qatar', 'Domninca', 'Rica', 'Malaysia', 'Federation', 'Bahranin', 'Kumatit', 'Urarian', 'Torenina', '
```

```
'Indonesia', 'Honduras', 'India', 'Guatemala', 'Vietnam', 'Swaziland', 'Republic', 'Nicaragua', 'Kenya', 'Tajikistan', 'Senegal', 'Zimbabwe', 'Ghana' 'Uganda', 'Madagascar', 'Mauritania', 'Tanzania', 'Pakistan', 'Lesotho', 'Benin', 'Nigeria', 'Yemen', 'Mali', 'Mozambique', 'Angola', 'Cameroon', 'Faso 'Nepal', 'Malawi', 'Laoo', 'Hangladesh', 'Myanmar', 'Rwanda', 'Ethiopia', 'Djibouti', 'Eritrea', 'Leone', 'Herzegovina', 'Montenegro', 'Haiti', 'Ivoire' 'Sudan', 'Zambia']

>>> pola4 = r'(\wh\/3>\/td>\\rd>\(\dot\)'
>>> tuples = re.findall(pola5,teks)
Traceback (most recent call last);
```