

# Regularni izrazi

Podrška za rad sa regularnim izrazima data je u *Python* modulu *re*.

1. Napisati Python skript koji provjerava da li se na početku stringa nalaze 2 velika slova.

```
#!/usr/bin/python3

import re

s = input('Unesite string: ')

rezultat = re.match(r'[A-Z]{2}', s)

if rezultat is1 not None:
    print('Na pocetku stringa se nalaze 2 velika slova')
else:
    print('Na pocetku stringa se NE nalaze 2 velika slova')
```

2. Napisati Python skript koji provjerava da li se na početku stringa nalazi string 'pocetak' pri čemu veličina slova nije bitna.

```
#!/usr/bin/python3

import re

s = input('Unesite string: ')

rezultat = re.match(r'pocetak', s, re.I)

if rezultat is not None:
    print('Nalazi se')
```

---

<sup>1</sup> Razlika između `==` i operatora `is`:

1. operator `==` poredi da li su jednake vrijednosti objekata
2. operator `is` poredi da li se radi o istom objektu u memoriji

```
else:  
    print('Ne nalazi se')
```

3. Napisati Python skript koji provjerava da li se u stringu uzastopno nalaze 2 velika slova i ispisati ih.

```
#!/usr/bin/python3  
  
import re  
  
s = input('Unesite string: ')  
  
rezultat = re.search(r'[A-Z]{2}', s)  
  
if rezultat is not None:  
    print('U stringu se nalaze 2 velika slova')  
    print('Nadjena slova su: ' + rezultat.group())  
else:  
    print('U stringu se NE nalaze 2 velika slova')
```

4. Napisati Python skript koji ispisuje svako pojavljivanje slova a jedan ili više puta u zadatom stringu.

```
#!/usr/bin/python3  
  
import re  
  
s = input('Unesite string: ')  
  
rezultat = re.findall(r'a+', s)  
  
n = len(rezultat)  
  
for i in range(n):  
    print(f'{i}. poklapanje je: {rezultat[i]}')
```

5. Napisati Python skript koji ispisuje svaki početak i kraj pojavljivanja slova a jedan ili više puta u zadanom stringu.

```
#!/usr/bin/python3

import re

s = input('Unesite string: ')

rezultat = re.finditer(r'a+', s)

for poklapanje in rezultat:
    print(f'Poklopljeno: {poklapanje.group()}, \
pocetak - {poklapanje.start()}, kraj - {poklapanje.end()}')
```

6. Prepoznati string ako je riječ. Riječ može da sadrži slova, brojeve i \_.

```
#!/usr/bin/python3

import re

s = input('Unesite string: ')

rezultat = re.fullmatch(r'\w+', s)
if rezultat is None:
    print('Nije rijec')
else:
    print('Jeste rijec')
```

7. Zamijeniti svako pojavljivanje a jedan ili više puta (nije bitna veličina slova) slovom X.

```
#!/usr/bin/python3

import re

s = input('Unesite string: ')

novi_s = re.sub(r'a+', 'X', s, flags = re.I)
print(novi_s)
```

8. Prebrojati koliko ima riječi u tekstu. Riječ može da sadrži slova, brojeve, \_ i nakon riječi eventualno može da se nađe neki interpunkcijski znak ili bjelina.

```
#!/usr/bin/python3

import re

tekst = input('Unesite tekst: ')

rijeci = tekst.split()

brojac = 0

for rijec in rijeci:
    rezultat = re.fullmatch(r'\w+|\w+[,!?!?]', rijec)
    if rezultat is not None:
        brojac += 1

print(f'Broj rijeci je {brojac}')
```

9. U zadatom tekstu ispitati da li svaka linija počinje velikim slovom.

```
#!/usr/bin/python3

import re

tekst = """Ovo je neki tekst.
Svaka njegova linija pocinje velikim slovom.
To bi trebalo provjeriti."""

linije = tekst.split('\n')

automat = re.compile(r'[A-Z]')

ind = True

for linija in linije:
    rezultat = automat.match(linija)
```

```

        if rezultat is None:
            ind = False
            break

if ind:
    print('Svaka linija pocinje velikim slovom')
else:
    print('Ne pocinje svaka linija velikim slovom')

```

10. Napisati Python skript koji kao argument komandne linije dobija putanju do html fajla. U fajlu je data tabela koja sadrži informacije o studentima. Na standardni izlaz ispisati ime, prezime, indeks i broj poena koje je student osvojio ukupno na ispitu pri čemu studenti treba da budu sortirani prema ukupnom broju poena opadajuće.

```

#!/usr/bin/python3

import sys
import re

if len(sys.argv) != 2:
    sys.exit('Neispravan broj argumenata komandne linije')

naziv_fajla = sys.argv[1]

if not naziv_fajla.lower().endswith('.html'):
    sys.exit('Nije html fajl')

try:
    f = open(naziv_fajla, 'r')
    sadrzaj = f.read()
    f.close()
except:
    sys.exit('Neuspesno otvaranje fajla')

reg_izraz = r'<tr[^>]*>\s*'
reg_izraz += r'<td[^>]*>\s*([A-Z][a-z]+)\s+([A-Z][a-z]+)\s*</td>\s*'
reg_izraz += r'<td[^>]*>\s*(m[mnvrli]\d{5})\s*</td>\s*'
reg_izraz += r'<td[^>]*>\s*([0-9][1-4][0-9]50)\s*</td>\s*'

```

```
reg_izraz += r'<td[^>]*>\s*([0-9]([1-4][0-9])?50)\s*</td>\s*'
reg_izraz += r'</tr>'
```

```
automat = re.compile(reg_izraz)
```

```
studenti = {}
```

```
'''
```

```
# 1. nacin:
```

```
for poklapanje in automat.finditer(sadrzaj):
```

```
    ime = poklapanje.group(1)
```

```
    prezime = poklapanje.group(2)
```

```
    alas = poklapanje.group(3)
```

```
    prakticni = poklapanje.group(4)
```

```
    teorija = poklapanje.group(5)
```

```
    ime_i_prezime = ime + ' ' + prezime
```

```
    ukupni_poeni = int(prakticni) + int(teorija)
```

```
    indeks = str(int(alas[4:])) + '/20' + alas[2:4]
```

```
    studenti[indeks] = (ime_i_prezime, ukupni_poeni)
```

```
'''
```

```
# 2. nacin:
```

```
for poklapanje in automat.findall(sadrzaj):
```

```
    ime = poklapanje[0]
```

```
    prezime = poklapanje[1]
```

```
    alas = poklapanje[2]
```

```
    prakticni = poklapanje[3]
```

```
    teorija = poklapanje[4]
```

```
    ime_i_prezime = ime + ' ' + prezime
```

```
    ukupni_poeni = int(prakticni) + int(teorija)
```

```
    indeks = str(int(alas[4:])) + '/20' + alas[2:4]
```

```
    studenti[indeks] = (ime_i_prezime, ukupni_poeni)
```

```
kljucevi = studenti.keys()
```

```
sortirani_kljucevi = sorted(kljucevi, key = lambda x : studenti[x][1],
reverse = True)
```

```
for kljuc in sortirani_kljucevi:  
    print(f'{studenti[kljuc][0]}, broj indeksa {kljuc}, broj poena  
{studenti[kljuc][1]}')
```