

Case Study: Create Regular Expressions

- Extract the complete timestamps from each line.

```
import re

s = """
Date: Sat, 06 Jan 2016 07:50:17 GMT
Hello This is testing
Date: Mon, 26 Jan 2016 07:50:17 GMT
Have a nice time...
Date: Wed, 16 Jan 2016 07:50:17 GMT
"""

p=re.compile("\w{3}, \d{2} \w{3} \d{4} \d{2}:\d{2}:\d{2} \w{3}")

for i in re.findall(p,s):
    print('%s' % i)
```

- Extract the complete e-mail address from each line.

```
import re

s = """
should we use regex???
contact me at 32laddf@abc.com
hello pz contact
at rjtdon1@xyz.com
or you can contact me at
contact@yourrisk.com
"""

p = re.compile(r'[\w\.-] + @ [\w\.-] +')

for i in re.findall(p,s):
    print('%s' % i)
```

- Extract only the months from the timestamps.

```
regex = re.compile('(P<month>[a-zA-Z]+)\s+
                    (P<year>\d{4})')
m = regex.search('Hello How are you May 2013')
m.group('month')
```

```
import datetime
m = datetime.datetime.now()
m.strftime("%B")
m.strftime("%b")
```

- Extract only the years from the timestamps.

```
regex = re.compile('(P<month>[a-zA-Z]+)\s+
                    (P<year>\d{4})')
m = regex.search('Hello How are you May 2013')
m.group('year')
```

```
import datetime
m = datetime.datetime.now()
m.strftime("%Y")
m.strftime("%y")
```

- Extract login, domain names from the given string

```
plogin = r'^([^\@]+)\@([^\@]+)$'
pdomain = '@[\w.]+'
email = 'rpbc@rajkot.com'
mlogin = re.search(plogin,email)
mlogin.group(1)#Display only login name
mdomain = re.search(pdomain,email)
mdomain.group()

#Extract only domains like .com/.in
Str1 = """
    my first email is user1@gmail.com
    second email is enter code
    hereuser2@yahoo.in and third email
    is user3@outlook.com
    """
print(re.findall('@+\S+[\.in|.com|]',Str1))
```

- Processing Telephone Numbers

```
p = '(\d{3}[-\.\s]??\d{3}[-\.\s]??\d{4}
    |(\d{3})\s*\d{3}[-\.\s]??\d{4}|
    \d{3}[-\.\s]??\d{4})'
s = '(123)-456-7890'
re.match(p,s).group()
```

000-000-0000
000 000 0000
000.000.0000

(000) 000-0000
(000) 000 0000
(000) 000.0000
(000) 000-0000
(000) 000 0000
(000) 000.0000

000-0000
000 0000
000.0000

0000000
0000000000

The above regex for telephone number
matches all these possible values