


```

1
2 log=''192.168.198.92 - - [22/Jan/2002:23:08:37 -0400] "GET
3 / HTTP/1.1" 200 6394 www.yahoo.com
4 "-" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1...)" "-"
5 192.168.198.92 - - [22/Mar/2002:23:08:38 -0400] "GET
6 /images/logo.gif HTTP/1.1" 200 807 www.yahoo.com
7 "http://www.some.com/" "Mozilla/4.0 (compatible; MSIE 6...)" "-"
8 192.168.72.177 - - [22/Jun/2002:23:32:14 -0400] "GET
9 /news/sports.html HTTP/1.1" 200 3500 www.yahoo.com
10 "http://www.some.com/" "Mozilla/4.0 (compatible; MSIE ...)" "-"
11 192.168.72.177 - - [22/Jul/2002:23:32:14 -0400] "GET
12 /favicon.ico HTTP/1.1" 404 1997 www.yahoo.com
13 "-" "Mozilla/5.0 (Windows; U; Windows NT 5.1; rv:1.7.3)..." "-"
14 192.168.72.177 - - [22/Aug/2002:23:32:15 -0400] "GET
15 /style.css HTTP/1.1" 200 4138 www.yahoo.com
16 "http://www.yahoo.com/index.html" "Mozilla/5.0 (Windows..." "-"
17 192.168.72.177 - - [22/Dec/2002:23:32:16 -0400] "GET
18 /js/ads.js HTTP/1.1" 200 10229 www.yahoo.com
19 "http://www.search.com/index.html" "Mozilla/5.0 (Windows..." "-"
20 192.168.72.177 - - [22/Dec/2002:23:32:19 -0400] "GET
21 /search.php HTTP/1.1" 400 1997 www.yahoo.com
22 "-" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; ...)" "-"'''
23 p="\d+\. \d+\. \d+\. \d+"
24 pdate=p+"\d+/\w{3}/\d+"
25 pwebsite="\w{4}\:+\//\w{3}\.+\w*\.\w*"
26 list=re.findall(pwebsite,log)
27 print(list)

```

 ['<http://www.some.com>', '<http://www.some.com>', '<http://www.yahoo.com>', '<http://www.search.com>']

1

```

1 import re
2 log=''192.168.198.92 - - [22/Dec/2002:23:08:37 -0400] "GET
3 / HTTP/1.1" 200 6394 www.yahoo.com
4 "-" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1...)" "-"
5 192.168.198.92 - - [22/Dec/2002:23:08:38 -0400] "GET
6 /images/logo.gif HTTP/1.1" 200 807 www.yahoo.com
7 "http://www.some.com/" "Mozilla/4.0 (compatible; MSIE 6...)" "-"
8 192.168.72.177 - - [22/Dec/2002:23:32:14 -0400] "GET
9 /news/sports.html HTTP/1.1" 200 3500 www.yahoo.com
10 "http://www.some.com/" "Mozilla/4.0 (compatible; MSIE ...)" "-"
11 192.168.72.177 - - [22/Dec/2002:23:32:14 -0400] "GET
12 /favicon.ico HTTP/1.1" 404 1997 www.yahoo.com
13 "-" "Mozilla/5.0 (Windows; U; Windows NT 5.1; rv:1.7.3)..." "-"
14 192.168.72.177 - - [22/Dec/2002:23:32:15 -0400] "GET
15 /style.css HTTP/1.1" 200 4138 www.yahoo.com
16 "http://www.yahoo.com/index.html" "Mozilla/5.0 (Windows..." "-"
17 192.168.72.177 - - [22/Dec/2002:23:32:16 -0400] "GET
18 /js/ads.js HTTP/1.1" 200 10229 www.yahoo.com
19 "http://www.search.com/index.html" "Mozilla/5.0 (Windows..." "-"
20 192.168.72.177 - - [22/Dec/2002:23:32:19 -0400] "GET
21 /search.php HTTP/1.1" 400 1997 www.yahoo.com
22 "-" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; ...)" "-"'''

```

23

```
1 import re
2 data="mail:amartoHELP@yahoo.com +91-022-9821601163 mail:amartechNO@yahoo.com +91-011-9899887766 mail:champa.chameli@gmail.com +91-033-8899776655"
3 pip="^(mail:)"
4 listed=re.findall(pip,data)
5 print(listed)
```

▼ validation done through RE

```
1 import re
2 p="[a-z]"
3 p2="\d"
4 i=input("Enter data to check:")
5 if re.search(p2,i):
6     print("yes")
7 else:
8     print("no")
9
```

Enter data to check:23456789
yes

```
1 import re
2 p="[a-zA-Z0-9]"
3 i=input("Enter data to check:")
4 if re.search(p,i):
5     print("yes")
6 else:
7     print("no")
8
```

Enter data to check:@#%\$%^&*()_
no

```
1 import re
2 p="(com|colin|man)"
3 i=input("Enter data to check:")
4 if re.search(p,i):
5     print("yes")
6 else:
7     print("no")
```

Enter data to check:AMAR.COM
no

```
1 import re
2 p="[a-zA-Z0-9]+\w."
3 i=input("Enter data to check:")
4 if re.search(p,i):
5     print("yes")
```

```
6 else:
7     print("no")
8
```

Enter data to check:asdf123%^&*(
no

```
1 import re
2 p="[0-9]+\.[0-9]+\.[0-9]+\.[0-9]"#IP 0-9 0 TO 9 DIGIT
3 i=input("Enter data to check:")
4 if re.search(p,i):
5     print("yes")
6 else:
7     print("no")
8
```

Enter data to check:123.12.34.45
yes

```
1 #email
2 p="[a-zA-Z0-9]+\@[a-zA-Z]+\.[a-zA-Z]"#email
3 i=input("Enter data to check:")
4 if re.search(p,i):
5     print("yes")
6 else:
7     print("no")
```

Enter data to check:AMAR@PANCHAL.COM
yes

```
1 p="\d{10}"#mobile 0-10 DIGITS
2 i=input("Enter data to check:")
3 if re.search(p,i):
4     print("yes")
5 else:
6     print("no")
```

Enter data to check:98745
no

```
1 p="^\d{2}"#start with 2 digits
2 i=input("Enter data to check:")
3 if re.search(p,i):
4     print("yes")
5 else:
6     print("no")
```

Enter data to check:12AMAR
yes

```
1 p="\d{2}$"#end with 2 digits
2 i=input("Enter data to check:")
3 if re.search(p,i):
4     print("yes")
```

```
5 else:
6     print("no")
```

```
Enter data to check:AMAR007
yes
```

```
1 p="\+(\d\d)-(\d\d\d)-(\d\d\d\d\d\d\d\d\d)"
2 newp=r"Area:\1\2\3"#r raw string
3 i=input("Enter data to check:")
4 newi=re.sub(p,newp,i)
5 print("Final Result:",newi)
```

```
Enter data to check:+91-022-2357890123
Final Result: Area:910222357890123
```

▼ Extract

```
1 p="[\d]+"#word seprated
2 i=input("Enter data to check:")
3 data=re.findall(p,i)
4 print("Final Result:",data)
```

```
Enter data to check:AMAR HIS NUMBER IS 9821601163 CALL HIM AT ONCE
Final Result: ['9821601163']
```

```
1
```

```
1 p="\+(\d{2})-(\d{3})-(\d{10})"
2 newp=r"Country:\1 Area:\2 Number\3"#r raw string
3 i=input("Enter data to check:")
4 newi=re.sub(p,newp,i)
5 print("Final Result:",newi)
```

```
Enter data to check:+91-022-9876543210
Final Result: Country:91 Area:022 Number9876543210
```

