Regular Expression

In [ ]:
```
Regular Expression: if we want to represent a group of strings , according to a particular
                    pattern then we should use Regular Expression.
```

Examples

In [ ]:
```
Example of RE --> fixed pattern is there for our mobile numbers
                  (0-9 ,10numbers starting_no --> 6,7,8,9)

Example of RE --> Fixed Pattern for PAN Card

Example of RE -->fixed pattern for Vechile number

Example of RE --> Fixed Pattern is for Password --> strong password,weak,medium

Example of RE --> Fixed Pattern for IP Address

Example of RE --> Fixed Pattern for IFSC Code

Example of RE --> Fix Pattern for websites(www.google.com  www.amityuniversity.edu.in)

Example of RE --> Fixed pattern is given to our mail id
```

Mobile Number Pattern

In [ ]:
```
Mobile Number Pattern:
    Mobile Numbers --> 10 --> Starting digit --> 7,8,9,6
    In India There is one fixed pattern for the mobile Number:
        1.only 10 digits are  allowed.
        2.No two Mobile numbers are same
        3.Starting digit of a mobile Number is either 6,7,8,9
```

Email ID Pattern:

In [ ]:
```
Email ID Pattern:
    Email Account --> username@domain_name
    domain_name --> yahoo, hotmail, gmail

    Can we tell that after @ we will get the domain name
```

Uses of Regular Expression

In [ ]:
```
Where we need to use Regular Expression:
    1.Password
    2.Validation of Data
    3.Email
    4.Translator --> TOC , COMPILER DESIGN
    5.Digital Circuit
    6.Protocols like TCP/IP
```

How we can Implement regular Expression in Python

In [ ]:
```
--> In python Programming for implementing regular expression we are having a module
   which is known as re and This module is used for regular expression.

--> re is the module that we need to import for regular expression.
```

Few Important Function related to Regular Expression(Regex)

Finditer() , start(),end() and group() Function

In [ ]:
```
finditer() --> that will take two arguments substring(pattern) and string(input).
           --> return the iterator object which yields match object for every match.

start() --> return the start index of the match

end () --> return end+1 index of the match

group()--> return the matched string
```

Example

In [3]:
```python
import re
matcher = re.finditer("abc","abcabcbabc")
for i in matcher:
    print(i.start())
    print(i.end())
    print(i.group())
```
```
0
3
abc
3
6
abc
7
10
abc
```

Python Program to Return the Domain of a Mail

In [14]:
```python
#Domain of a Mail
import re
Email_id = input("Enter Your Mail Id : ")
matcher = re.finditer("@",Email_id)
for i in matcher:
    y=i.start()
if Email_id[y+1:]=="hotmail.com":
    print("It is a hotmail Id")

elif Email_id[y+1:]=="yahoo.com":
    print("It is a Yahoo Id")

elif Email_id[y+1:]=="outlook.com":
    print("It is a outlook Id")

elif Email_id[y+1:]=="gmail.com":
    print("It is a Gmail Id")

else:
    print("Enter Valid mail Id")
```
```
Enter Your Mail Id : ashu@gmail.com
It is a Gmail Id
```

Character Classes in RE

In [ ]:
```
--> We can customize all character classes based on our Requirements.
```

In [ ]:
```
Character Classes:
    1.[abc] --> either a or b or c
    2.[^abc] --> except a ,b,c
    3.[a-z] --> any lower case alphabet
    4.[A-Z] --> any upper case alphabet
    5.[A-Za-z] --> any alphabet either upper case or lower case
    6.[0-9] --> Any digit from 0 to 9
    7.[A-Za-z0-9] --> any alphanumeric symbols
    8.[^A-Za-z0-9] --> any special symbol
```

Example of Each Character Classes

[abc] --> either a or b or c

In [4]:
```python
import re
matcher = re.finditer("[abc]","a7b@9xbs")
for i in matcher:
    print(str(i.start())+"    "+str(i.end())+"    "+str(i.group()) )
                 #0                #1                    a
                 #2                #3                    b
                 #6                #7                    b
```
```
0    1    a
2    3    b
6    7    b
```

[^abc] --> except a ,b,c

In [7]:
```python
import re
matcher = re.finditer("[^abc]","a7b@9xbs")
for i in matcher:
    print(str(i.start())+"    "+str(i.end())+"    "+str(i.group()) )
            #1                     #2                   7
            #3                     #4                   @
            #4                     #5                   9
            #5                     #6                   x
            #7                     #8                   s
```
```
1    2    7
3    4    @
4    5    9
5    6    x
7    8    s
```

[a-z] --> any lower case alphabet

In [6]:
```python
import re
matcher = re.finditer("[a-z]","a7b@9xbs")
for i in matcher:
    print(str(i.start())+"    "+str(i.end())+"    "+str(i.group()) )
           #0                     #1                a
           #2                     #3                b
           #5                     #6                x
           #6                     #7                b
           #7                     #8                s
```
```
0    1    a
2    3    b
5    6    x
6    7    b
7    8    s
```

[A-Z] --> any upper case alphabet

In [8]:
```python
import re
matcher = re.finditer("[A-Z]","A77#234Bsca")
for i in matcher:
    print(str(i.start())+"    "+str(i.end())+"    "+str(i.group()) )
         #0                     #1                   #A
         #7                     #8                   #B
```
```
0    1    A
7    8    B
```

[A-Za-z] --> any alphabet either upper case or lower case

In [9]:
```python
import re
matcher = re.finditer("[A-Za-z]","A77#ab34B")
for i in matcher:
    print(str(i.start())+"    "+str(i.end())+"    "+str(i.group()) )
```
```
0    1    A
4    5    a
5    6    b
8    9    B
```

[0-9] --> Any digit from 0 to 9

In [22]:
```python
import re
matcher = re.finditer("[0-9]","A77#234Bsca")
for i in matcher:
    print(str(i.start())+"    "+str(i.end())+"    "+str(i.group()) )
        #1                        #2                   7
        #2                        #3                   7
        #4                        #5                   2
        #5                        #6                   3
        #6                        #7                   4
```
```
1    2    7
2    3    7
4    5    2
5    6    3
6    7    4
```

[A-Za-z0-9] --> any alphanumeric symbols

In [10]:
```python
import re
matcher = re.finditer("[A-Za-z0-9]","A77#234Bs")
for i in matcher:
    print(str(i.start())+"    "+str(i.end())+"    "+str(i.group()) )
        #0                        #1                   A
        #1                        #2                   7
        #2                        #3                   7
        #4                        #5                   2
        #5                        #6                   3
        #6                        #7                   4
        #7                        #8                   B
        #8                        #9                   s
```
```
0    1    A
1    2    7
2    3    7
4    5    2
5    6    3
6    7    4
7    8    B
8    9    s
```

[^A-Za-z0-9] --> any special symbol

In [25]:
```python
import re
matcher = re.finditer("[^A-Za-z0-9]","A77#2 34Bs")
for i in matcher:
    print(str(i.start())+"    "+str(i.end())+"    "+str(i.group()) )
      #3                          #4                      3
      #5                          #6

#Note --> Space is also considered as a Special Characters.
```
```
3    4    #
5    6
```

Customization and Practice Problem Based on Character Classes

In [11]:
```python
import re
matcher = re.finditer("[^A-Za-z0-9][A-Za-z0-9]","A77#23%a4Bs")
for i in matcher:
    print(str(i.start())+"    "+str(i.end())+"    "+str(i.group()) )
```
```
3    5    #2
6    8    %a
```

In [12]:
```python
#[klm]
import re
matcher = re.finditer("[klm]","A77#2k%alm")
for i in matcher:
    print(str(i.start())+"    "+str(i.end())+"    "+str(i.group()) )
```
```
5    6    k
8    9    l
9    10    m
```

In [13]:
```python
#[^klm]
import re
matcher = re.finditer("[^klm]","abcklm")
for i in matcher:
    print(str(i.start())+"    "+str(i.end())+"    "+str(i.group()) )
        #0                        #1                   a
        #1                        #2                   b
        #2                        #3                   c
```
```
0    1    a
1    2    b
2    3    c
```

In [14]:
```python
#[a-k]
import re
matcher = re.finditer("[a-k]","abcklm")
for i in matcher:
    print(str(i.start())+"    "+str(i.end())+"    "+str(i.group()) )
        #0                        #1                   a
        #1                        #2                   b
        #2                        #3                   c
        #3                        #4                   k
```
```
0    1    a
1    2    b
2    3    c
3    4    k
```

In [15]:
```python
#[A-K]
import re
matcher = re.finditer("[A-K]","ABCKLM")
for i in matcher:
    print(str(i.start())+"    "+str(i.end())+"    "+str(i.group()) )
        #0                        #1                   a
        #1                        #2                   b
        #2                        #3                   c
        #3                        #4                   k
```
```
0    1    A
1    2    B
2    3    C
3    4    K
```

Important Functions Related to RE Module

re.Match() Function

In [ ]:
```
re.Match --> check the string at the begining of the string
         --> re.match() searches only from the beginning of the string and return match object
             if found. But if a match of substring is found somewhere in the middle of the string,
             it returns none.
```

Example

In [39]:
```python
import re
s=input("Enter a String :")
m=re.match(s,"abcabdefg")
if m!=None:
    print("Match found")
else:
    print("Not Found")
```
```
Enter a String :cab
Not Found
```

re.FullMatch Function()

In [ ]:
```
re.fullmatch --> This function is used to check whole string.
             --> re.fullmatch() returns a match object if and only if the entre string matches the
                 pattern. Otherwise, it will return None.
```

Example

In [45]:
```python
import re
s=input("Enter a String :")
m=re.fullmatch(s,"abcabdefg")
if m!=None:
    print("Match found")
else:
    print("Not Found")
```
```
Enter a String :abcabdefg
Match found
```

re.search () Function

In [ ]:
```
re.search()--> search the given pattern
           --> re.search() method looks for occurrences of the regex pattern inside the entire
               target string and returns the corresponding Match Object instance where the
               match found
```

Example

In [48]:
```python
import re
s=input("Enter a String :")
m=re.search(s,"abcabdefg")
if m!=None:
    print("Match found")
    print(m.start())
else:
    print("Not Found")
```
```
Enter a String :cab
Match found
2
```