# **Lesson-18:** Regular Expressions

**IMPORTANT NOTE:**

\w (word character) matches any single letter, number or underscore (same as [a-zA-Z0-9_]). The uppercase counterpart \W (non-word-character) matches any single character that doesn't match by \w (same as [^a-zA-Z0-9_]).

*: The preceding item will be matched zero or more times, i.e., 0+
+: The preceding item will be matched one or more times, i.e., 1+

#\d - Represents any digit (0 to 9) characters
#\D - Represents any non-digit characters
#\s - Represents white space characters
#\S - Represents non-white space characters
#\w - Represents any alphanumeric characters
#\W - Represents any non-alphanumeric characters
#\b - Represents a space around words
#\A - Matches only at the start of the string
#\Z - Matches only at the end of the string
# * - 0 or more repetitions of the preceding regex
# + - 1 or more repetitions of the preceding regex

**Program 1:**

```
import re

str = "I love you sub and India"
v = re.compile(r"s\w\w")
result = v.search(str)
if result:
        print(result.group( ))
```

**Output**:

```
sub
```

**Program 2:**

```
import re

str = "I love you sub and india"
res = re.search( r"s\w\w",str)
if result:
      print(result.group( ))
```

**Output**:

```
sub
```

**Program 3:**

```
import re

str = "python is a great programming language and python is a snake"
res = re.findall( r"p\w\w\w\w\w",str)
if res:
    print(res)
else:
    print("Not Found")
```

**Output:**

```
['python', 'progra', 'python']
```

**Program 4:**

```
import re

strone = "C is a programming language and python is a snake"
strtwo = "python is a great programming language and python is a snake"

#match the starting word
resone = re.match( r"p\w\w\w\w\w",strone)
restwo = re.match( r"p\w\w\w\w\w",strtwo)
if resone:
    print("***")
    print(resone.group( ))

if restwo:
    print("---")
    print(restwo.group( ))
```

**Output:**

```
---
python
```

**Program 5:**

```
import re

s = "Subhash+-  Programming*****Classes"
result = re.split(r"\W+", s)
print(result)
```

**Output:**

```
['Subhash', 'Programming',
'Classes']
```

**Program 6:**

**Output:**

```
import re

s = "Subhash Loves Cooking"
result = re.sub(r"Cooking", "Programming", s)
print(result)
```

```
Subhash Loves Programming
```

**Program 7**

```
import re

s = "subhash loves 3 hashing in programming and praying as hash at 5pm or 6"
d = "Subhash celebrates his birthday on 07-06-1985 and marriage anniversary on 28-08-2014"
a = "Subhash: 80808080 Charan: 9090909090 Archita: 1010101010"
b = "Hello World"

res = re.findall(r"h\w\w\w", s )
print(res)

res = re.findall(r"h[\w]*", s)
print(res)

res = re.findall(r"\bh[\w]*\b",s)
print(res)

res = re.findall(r"\d\w", s)
print(res)

res = re.findall(r"\d[\w]*",s)
print(res)

res = re.findall(r"\d[\w]+",s)
print(res)

res = re.findall(r"\b\w{7}\b",s)
print(res)

res = re.findall(r"\b\w{5,}\b", s)
print(res)

res = re.findall(r"\b\w{2,3}\b",s)
print(res)

res = re.findall(r"\b\d\b", s)
```

```
print(res)

res = re.findall(r"[\w]\Z", s)
print(res)

res = re.findall(r"\A[\w]*", s)
print(res)

res = re.findall(r"\Ap[\w]*", s)
print(res)

res = re.findall(r"\As[\w]*",s)
print(res)

res = re.findall(r"\w{1}", s)
print(res)

res = re.findall(r"\b\w{1}\b", s)
print(res)

res = re.findall(r"\D+", s)
print(res)

res = re.findall(r"\d{1,2}-\d{1,2}-\d{1,4}", d)
print(res)

res = re.findall(r"[A-Z][a-z]*", a)
print(res)

res = re.findall(r"\d[0-9]*", a)
print(res)

res = re.search(r"^hello", b, re.IGNORECASE)
if(res):
    print("string starts with hello")
else:
    print("string does not start with hello")

res = re.search(r"world$", b, re.IGNORECASE)
if(res):
    print("string ends with world")
else:
    print("string does not end with world")
```

**Output:**

```
['hash', 'hash', 'hash']
['hash', 'hashing', 'hash']
['hashing', 'hash']
['5p']
['3', '5pm', '6']
['5pm']
['subhash', 'hashing', 'praying']
['subhash', 'loves', 'hashing', 'programming', 'praying']
['in', 'and', 'as', 'at', '5pm', 'or']
['3', '6']
['6']
['subhash']
[]
['subhash']
['s', 'u', 'b', 'h', 'a', 's', 'h', 'l', 'o', 'v', 'e', 's', '3', 'h', 'a', 's', 'h', 'i', 'n', 'g', 'i', 'n', 'p', 'r', 'o', 'g', 'r', 'a', 'm', 'm', 'i', 'n', 'g', 'a', 'n', 'd', 'p', 'r', 'a', 'y', 'i', 'n', 'g', 'a', 's', 'h', 'a', 's', 'h', 'a', 't', '5', 'p', 'm', 'o', 'r', '6']
['3', '6']
['subhash loves ', ' hashing in programming and praying as hash at ', 'pm or ']
['07-06-1985', '28-08-2014']
['Subhash', 'Charan', 'Archita']
['80808080', '9090909090', '1010101010']
string starts with hello
string ends with world
```