# Assignment 16 Solutions

### Q1. What is the benefit of regular expressions?

**Ans:** Regular Expressions, also known as `regex` or `regexp` , regular expression is a special sequence of characters that helps us to match or find other strings or sets of strings, using a specialized syntax held in a pattern. Regular expressions are widely used in UNIX world. The Python module re provides full support for Perl-like regular expressions in Python. The re module raises the exception re.error if an error occurs while compiling or using a regular expression.

Python's regular expression library is called "re" and is based on Perl's regular expression engine, which is pretty much an industry standard now.

### Q2. Describe the difference between the effects of "(ab)c+" and "a(bc)+" Which of these, if any, is the unqualified pattern "abc+"?

**Ans:** Both `(ab)c+` and `a(bc)+` are valid patterns. The difference between both these patterns is that in `(ab)c+` `ab` is a group whereas in `a(bc)+` `bc` is a group. Capturing groups are a way to treat multiple characters as a single unit. They are created by placing the characters to be grouped inside a set of parentheses.

### Q3. How much do you need to use the following sentence while using regular expressions?

```
import re
```

**Ans:** `import re` statement always has to be imported before using regular expressions it's a basic import statement to use regular expressions in python

### Q4. Which characters have special significance in square brackets when expressing a range, and under what circumstances?

**Ans:** The Characters `.` , `*` , `?` , `^` ,or, `()` , have a special signiface when used with square brackets. They need not be be explicitly escaped by `\` as in case of pattern texts in a raw string. Range in Regular Expressions Ranges of characters can be indicated by giving two characters and separating them by a '-', for example [a-z] will match any lowercase ASCII letter, [0-5][0-9] will match all the two-digits numbers from 00 to 59.

### Q5. How does compiling a regular-expression object benefit you?

**Ans:** We can combine a regular expression pattern into pattern objects, which can be used for pattern matching. It also helps to search a pattern again without rewriting it

### Q6. What are some examples of how to use the match object returned by re.match and re.search?

**Ans:** `re.search()` is returning match object and implies that first match found at index 69.

`re.match()` is returning none because match exists in the second line of the string and `re.match()` only works if the match is found at the beginning of the string.

`re.IGNORECASE` is used to ignore the case sensitivity in the strings.

Both `re.search()` and `re.match()` returns only the first occurrence of a substring in the string and ignore others.

In [4]:

```
# import re module
import re
Substring ='string'
String1 ='''We are learning regex with geeksforgeeks
        regex is very useful for string matching.
        It is fast too.'''
String2 ='''string We are learning regex with geeksforgeeks
        regex is very useful for string matching.
        It is fast too.'''
# Use of re.search() Method
print(re.search(Substring, String1, re.IGNORECASE))
# Use of re.match() Method
print(re.match(Substring, String1, re.IGNORECASE))
# Use of re.search() Method
print(re.search(Substring, String2, re.IGNORECASE))
# Use of re.match() Method
print(re.match(Substring, String2, re.IGNORECASE))
```

```
<re.Match object; span=(68, 74), match='string'>
None
<re.Match object; span=(0, 6), match='string'>
<re.Match object; span=(0, 6), match='string'>
```

### Q7. What is the difference between using a vertical bar (|) as an alteration and using square brackets as a character set?

**Ans:** When **|** us used then patterns searches for **or** option. i.e **<pattern_1>|<pattern_2>** means it searches as **<pattern_1>or<<pattern_2>** in the searched string. The first occurance of matched string will be returned as the Match Object. Using Character set in square Brackets searches for all the character set in the square bracket and if match is found, it returns it.

In [5]:

```
#For example
#[p - t] = [ pqrst ]
#For example, (x|y|z)ab match by any string-like, x, y, z, a, b.
```

### Q8. In regular-expression search patterns, why is it necessary to use the raw-string indicator (r)? In replacement strings?

**Ans:** Raw Strings are used in the regular-expression search patterns, so that backslash do not have to be escaped. When one wants to match a literal backslash, it must be escaped in the regular expression. With raw string notation, this means r"\". Without raw string notation, one must use 4 times back slashes