Title : Python Basic Assignment_7
Name : Bhavikkumar Modi
E-mail : bmodi700@gmail.com
Website : https://keenbm.github.io/

1. What is the name of the feature responsible for generating Regex objects?

Ans:

re.compile()

2. Why do raw strings often appear in Regex objects?

Ans:

Raw string notation keeps regular expression meaningful and confusion free. Without it every backslash(' \ ') in regular expression would have to be prefixed with another one to escape it.

3.What is the return value of the search() method?

Ans:

It returns a match object if there is a match else returns none.

4. From a Match item, how do you get the actual strings that match the pattern?

Ans:

group() method returns the actual string that match the pattern.

5. In the regex which created from the r'(\d\d\d)-(\d\d\d-\d\d\d\d)', what does group zero cover?Group 2? Group 1?

Ans:

- group 0 covers (\d\d\d)-(\d\d\d-\d\d\d\d)
- group 1 covers (\d\d\d)
- group 2 covers (\d\d\d-\d\d\d\d)

6. In standard expression syntax, parentheses and intervals have distinct meanings. How can you tell a regex that you want it to fit real parentheses and periods?

Ans:

`# \., \(, \)`

7.The findall() method returns a string list or a list of string tuples. What causes it to return one of the two options?

Ans:

If there is a one group in the expression, findall() returns list of strings. If there is several groups in the expression, findall() returns list of tuple of strings.

8.In standard expressions, what does the | character mean?

Ans:

It means either or

9.In regular expressions, what does the character stand for?

Ans:

Zero or one occerences

10. In regular expressions, what is the difference between the + and * characters?

Ans:

- ' + ' refers one or more occurences
- ' * ' refers zero or more occurences.

11. What is the difference between {4} and {4,5} in regular expression?

Ans:

- {4} means preceding character is to be repeated 4 times.
- {4,5} means preceding character is repeated minimum 4 times and maximum 5 times.

12. What do you mean by the \d, \w, and \s shorthand character classes signify in regular expressions?

Ans:

- \d returns a match where string contains digit(number from 0-9)
- \w returns a match where string contains any word character(character from a to Z, digit from 0 to 9 and underscore character)
- \s returns a match where string contains a whitespace character.

14. What is the difference between .? and .?

Ans:

- '.' accepts as many sequence as available(Greedy approach)
- '.?' accepts the first matched sequence and then stops(Non-Greedy approach).

15. What is the syntax for matching both numbers and lowercase letters with a character class?

Ans:

[a-z0-9]

16. What is the procedure for making a normal expression in regax case insensitive?

Ans:

pass re.IGNORECASE to re.compile()

17. What does the . character normally match? What does it match if re.DOTALL is passed as 2nd argument in re.compile()?

Ans:

'.' character normally pass any character except newline character. With the help of DOTALL, '.' character can match any character including newline.

18. If numReg = re.compile(r'\d+'), what will numReg.sub('X', '11 drummers, 10 pipers, five rings, 4 hen') return?

Ans:

'X drummers, X pipers, five rings, X hen'

19. What does passing re.VERBOSE as the 2nd argument to re.compile() allow to do?

Ans:

It allows to write regular expressin that look nicer and are more readable by allowing visually separate logical section of the pattern and add comments.

In [ ]: