

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

`re.compile()` is the feature responsible generating regex objects.

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

Raw strings are used so that backslashes do not have to be escaped.

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

The `group()` method returns strings of the matched text.

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

The `group()` method is used to get actual strings that matched 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?

Group 0 is the entire match, group 1 covers the first set of parentheses, and group 2 covers the second set of parentheses.

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?

We will use escape character.

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

If the regex has no groups, a list of strings is returned. If the regex has groups, a list of tuples of strings is returned.

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

This is used in case of either scenario. The `|` character signifies matching "either, or" between two groups.

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

The `?` character can either mean "match zero or one of the preceding group" or be used to signify non greedy matching.

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

The `+` matches one or more. The `*` matches zero or more.

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

`{4}` - this means matching next 4 patterns.

`{4,5}` - this matches next 4 to 5 pattern.

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

`\w`: the “word character” class represents the regex range `[A-Za-z0-9_]`, and it matches a single uppercase character, lowercase character, digit or underscore

`\d`: the “digit character” class represents the regex range `[0-9]`, and it matches a single digit character

`\s`: the “whitespace character” class represents the regex range `[\t\r\n\f\v]`, matching a single space, tab, carriage return, line break, form feed, or vertical tab

13. What do means by `\D`, `\W`, and `\S` shorthand character classes signify in regular expressions?

`\D` – matches any character except a digit e.g., a letter.

`\S` – matches any character except a whitespace e.g., a letter

`\W` – matches any character except a word character e.g., non-Latin letter or space.

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

`. *` matches any character (except for line terminators)

`*?` matches the previous token between zero and unlimited times, as few times as possible

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

`\w`

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

Using `re.IGNORECASE` will make insensitive

For example - `re.match('test', 'TeSt', re.IGNORECASE)`

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

`.` Matches anything

If `re.DOTALL` is passed as the second argument to `re.compile()`, then the dot will also match newline characters.

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

`[11,10,4]`

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

The `re.VERBOSE` argument allows you to add whitespace and comments to the string passed to `re.`

20. How would you write a regex that match a number with comma for every three digits? It must match the given following:

'42'

'1,234'

'6,368,745'

but not the following:

'12,34,567' (which has only two digits between the commas)

'1234' (which lacks commas)

`(^\d{1,3})(,\d{3})*$`

21. How would you write a regex that matches the full name of someone whose last name is Watanabe? You can assume that the first name that comes before it will always be one word that begins with a capital letter. The regex must match the following:

'Haruto Watanabe'

'Alice Watanabe'

'RoboCop Watanabe'

but not the following:

'haruto Watanabe' (where the first name is not capitalized)

'Mr. Watanabe' (where the preceding word has a nonletter character)

'Watanabe' (which has no first name)

'Haruto watanabe' (where Watanabe is not capitalized)

`[A-Z]\w* Watanabe`

22. How would you write a regex that matches a sentence where the first word is either Alice, Bob, or Carol; the second word is either eats, pets, or throws; the third word is apples, cats, or baseballs; and the sentence ends with a period? This regex should be case-insensitive. It must match the following:

'Alice eats apples.'

'Bob pets cats.'

'Carol throws baseballs.'

'Alice throws Apples.'

'BOB EATS CATS.'

but not the following:

'RoboCop eats apples.'

'ALICE THROWS FOOTBALLS.'

'Carol eats 7 cats.'

```
re.compile(r'(Alice|Bob|Carol)\s(eats|pets|throws)\s(apples|cats|baseballs).', re.I|re.DOTALL)
```