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

re. compile()

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

python on raw string is created by prefixing a string literal with 'r' or 'R'. Python raw string treats backslash (\) as a literal character. This is useful when we want to have a string that contains backslash and don't want it to be treated as an escape character.

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

The search() method returns Matched Objects.

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

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

e.g->

import re

phoneNumRegex = re.compile(r'\d\d\d-\d\d\d-\d\d\d\d')

mo = phoneNumRegex.search('My number is 415-555-4242.')

print('Phone number found: ' + mo.group())

🡪 Phone number found: 415-555-4242.

1. 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?

the r'(\d\d\d)-(\d\d\d-\d\d\d\d)

Group 0 represents the entire match, Group 1 represents the first 3 digit in first parenthesis Group 2 represents the second parenthesis.

1. 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?

Period and parenthesis can be escaped by escape character

R’\.,\(,and\)’

1. 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 group then it return string list while if regex has group then it return

list of string tuples.

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

It represents ‘either or’ between groups.

1. In regular expressions, what does the character question mark stand for?

The question mark character either mean “match zero or one of the preceding group” or be used to signify nongreedy matching

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

+ signifies one or more character

\* signifies zero or more character.

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

{4} signifies exactly 4 instances of preceeding group.

While {4,5} matches between 4 and 5 instances

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

\d signifies a single digit. Regex range [0-9]

\w signifies word . Regex range [A-Za-z0-9\_]

\s signifies space Regex range [\t\r\n\f\v]

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

\D signifies a single character that is not a digit [^0-9]

\W signies non word character [^A-Za-z0-9\_]

\S signifies a non whitespace character [^\t\r\n\f\v]

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

.\* forms a greedy match

.\*question mark form non greedy match.

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

[0-9a-z] or [a-z0-9]

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

Passing re.I or re.IGNORECAE as the second argument to re.compile will make the matching character insensitive.

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

The . character normally matches any character except the newline character. 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 numRegex.sub('X', '11 drummers, 10 pipers, five rings, 4 hen') return?

X drummers , X pipers, five rings, X hen

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

It is use to add whitespace and comment o the string passing in re.compile().

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)

Re.compile(r’(^\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'

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)

Re.compile(r’[A-Z][a-z]\*\sWatanabe’)

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.IGNORECASE)