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

🡪 compile() returns a Regex pattern object (or simply, a Regex object).

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

🡪 Since regular expressions frequently use backslashes in them, it is convenient to pass raw strings to the re.compile() function instead of typing extra backslashes.

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

🡪 search() method returns -1 if no match is found.

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

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

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?

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(mo.group(0))

output🡪 415-555-4242

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(mo.group(1))

output 🡪 415

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(mo.group(2))

output🡪 555-4242

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?

🡪 Periods and parentheses can be escaped with a backslash: \., \(, and \).

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

🡪 findall() finds \*all\* the matches and returns them as a list of strings, with each string representing one match.

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

🡪 The | character is called a pipe. You can use it anywhere you want to match one of many expressions

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

🡪 The term Regex stands for Regular expression. The regex or regexp or regular expression is a sequence of different characters which describe the particular search pattern. It is also referred/called as a Rational expression

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

🡪 \* means zero or more occurrence of a in the string while + means that one or more occurrence of a in the string.

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

🡪 {4} is saying "match this pattern four times." {4,5} means "match this pattern four to five times."

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

🡪 \d- for matching strings. \w for word character and \s stands for “whitespace character”

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

🡪 The \d, \w, and \s shorthand character classes match a single character that is not a digit, word, or space character, respectively.

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

🡪 \*?) matches any character ( . ) any number of times ( \* ), as few times as possible to make the regex match ( ? ). You'll get a match on any string, but you'll only capture a blank string because of the question mark

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

🡪 ') You can also include ranges of letters or numbers by using a hyphen. For example, the character class [a-zA-Z0-9] will match all lowercase letters, uppercase letters, and numbers.

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

🡪 Use the regexp function with the same syntax as regexpi to perform case-sensitive matching.

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

🡪 “.” -- Matches any single character except a 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 hens

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

🡪 To ignore whitespace and comments inside the regular expression string.

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)

🡪 numCommas = re.compile(r'(^\d{1,3})(,\d{3})\*$') numCommas.search('12,34,567').group()

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)

🡪 fullName = re.compile(r'[A-Z]\w [A-Z]\w') mo = fullName.findall('Haruto Watanabe, 'Haruto Watanabe, Alice Watanabe', Watanabe', 'Haruto Watanabe, Robocop Watanabe') mo.group()

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.'

* Test result: [('Alice', 'eats', 'apples'), ('Bob', 'pets', 'cats'), ('Carol', 'throws', 'baseballs'), ('Alice', 'throws', 'Apples'), ('BOB', 'EATS', 'CATS')]