1. What is the name of the feature responsible for generating Regex objects? re.compile()

2. Why do raw strings often appear in Regex objects? Because we don’t have to scape backslash in this case

3. What is the return value of the search() method? Get two parameters and return the match object if they are matched return the first occurrence of the match object and if they are not match return None

4. From a Match item, how do you get the actual strings that match the pattern? We can use group() method

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 entire match, group 2 cover second set of parenthesis and group 1 cover first set of parentheses

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

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

mo.group(0) , mo.group(1) , mo.group(2)

**Result is : ('415-555-4242', '415', '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?

We can use backslash to address them **\( \) \,**

period = re.compile(r'(\(\d\,\d\))')

mo = period.search('My period is (4,3).')

mo.group(0)

**Result is : '(4,3)'**

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 one capturing groups is present in the pattern, return a list of groups but if the pattern has more than one group it will return list of tuple.

8. In standard expressions, what does the | character mean? The | character works same as OR and signifies matching "either, or" between two groups.

9. In regular expressions, what does the character stand for? \w (word character) matches any single letter, number or underscore (same as [a-zA-Z0-9\_])

10.In regular expressions, what is the difference between the + and \* characters? + matches One or more and \* matches Zero or more

11. What is the difference between {4} and {4,5} in regular expression? {4} The item matches exactly 4 times , {4,5} the expression matches at least 4 and not more than 5 times

12. What do you mean by the \d, \w, and \s shorthand character classes signify in regular expressions? It matches (a single digit, single space, single character, and single space) character which is digit or word or space

period = re.compile(r'\d, \w, and \s’)

mo = period.search('1, a, and ')

mo.group(0)

**result is: '1, a, and '**

13. What do means by \D, \W, and \S shorthand character classes signify in regular expressions? It matches a single character which are not digit no character and no space respectively

14. What is the difference between .\*? and .\*? nothing

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?

re. IGNORECASE ([A-Z])

17. What does the . character normally match? What does it match if re.DOTALL is passed as 2nd argument in re.compile()? . means any single character, If we send re.DOTALL 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? By this parameter we can 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: re.compile(r'\d{1,3}(,\d{3})\*')

'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)

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:

re.compile(r'[A-Z](?:\w)+\sWatanabe')

'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)

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:

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

'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.'