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

Re.complie() method with which we can get the regex object.

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

If we do not use the raw string in the regex objects then we need to use the escape character to keep the operator in the given string as string literal for every operator like ‘\’.

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

It will return the matched object if matched else None will be returned.

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

With group() method we can get the matched string.

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 1 will cover 3 digits.

Group2 will cover the 3 digits - 4 digits

Group 0 will cover 3 digits - 3 digits - 4 digits

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?

Here we can use the escape characters or a raw string to represent the parenthesis and periods while using in a regex.

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

It will return the string list if there are no groups in it.

If there are multiple groups are present then it will return the string of tuples for different groups.

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

It means that the regex will match the first string before | or the 2nd string after the |

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

After the character ? if there are some pattern string then it will represented as one or more of such strings.

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

\* means 0 or more repetitions and + means 1 or more repetitions.

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

{4} means the pattern should contain or match 4 characters.

{4,5} means the pattern should match between 4 to 5 not less than 4.

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

\d means decimal from 0-9

\w means it matches the word which is alphanumeric including underscore.

\s means the matching the characters including white spaces.

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

\D matching the character which is not decimal.

\W matching a character which is not a word.

\S matching a character which is not a whitespace.

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

\*? Means it will match at least one character from pattern

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

[0-9a-z]

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

By passing the parameter re.ignorecase.

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

. will match any character except newline. Whereas DOTALL will match any character including newline.

18. If numRegex = 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?

With this we can remove the white-spaces and can also write the comments and re expression in more suitable way which can be readable.

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{1,2}

^\d{1,2}(,\d{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)

==

re.search([+?A-Z\Sa-z]Watanabe$,pattern\_string)

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)