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

re module used to generate Regex object.

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

Row string is used to keep regular expression meaningful and confusion free. It is a simple string that match regular expression its avoid to used backslash in re.compile() method

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

Return None or match object if there is any match

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

Using 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 : (\d\d\d)-(\d\d\d-\d\d\d\d)

Group 1 : (\d\d\d)

Group 2: (\d\d\d-\d\d\d\d)

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?

Parentheses is used to make group of match and interval is used and we made regex with parentheses with interval like that r.()

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

When called on a regex with no groups, such as \d\d\d-\d\d\d-\d\d\d\d, the method findall() returns a list of string matches, such as ['(\d\d\d)-(\d\d\d\d-\d\d\d)', '(\d\d\d)-(\d\d\d\d-\d\d\d)']

When called on a regex that has groups, such as (\d\d\d)-(\d\d\d)-(\d\d\d\d), the method findall() returns a list of tuples of strings(one string for each group), such as [('415', '555', '9999'),('212', '555','0000')]

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. For example, the regular expressions r'Batman|Tina Fey' will match either 'Batman' or 'Tina Fey'

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

zero or One occurrence of group

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

\* Match Zero or more

+ Match one or more

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

{4} match this patter 4 times

{4,5} match this pattern 4 or 5 times

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

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

a \d in a regex stands for a digit character - that is, any single numberal 0 to 9.

\w means any letter, numeric digit, or the underscore character.

\s means any space, tab, or newline character.

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

Matches all or one group

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?

Pass re.IGNORECASE or re.I as a second argument to re.compile(). For example, ro = re.compile(r'robocop', re.I)

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

. match any character except new line character and with re.DOTALL we can match new line character also.

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

X drummer

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})\*$')

res=numCommas.search('123,334,567').group()

print(res)

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)

Fn=Re.compile([r’A-Z]\w [A-Z]\w’)

Fn.findAll('Haruto Watanabe, Alice Watanabe, RoboCop Watanabe, haruto Watanabe, Mr. 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.'

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

senRegex.findall('''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.''')