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

The re.compile() function is responsible for generating Regex objects.

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

Raw strings (prefixed with r) are often used in Regex objects to ensure that backslashes (\) are treated as literal characters. This prevents Python from interpreting them as escape characters.

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

The search() method returns a Match object, which contains information about the first occurrence of the pattern in the searched string.

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

To get the actual strings that match the pattern from a Match object, you can use the .group() method with the desired group number. For example, match.group(0) returns the entire matched string, and match.group(1) returns the content of the first captured group.

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?

 In the regex r'(\d\d\d)-(\d\d\d-\d\d\d\d)', group zero covers the entire matched string, group 2 covers the part after the hyphen (e.g., the phone number), and group 1 covers the part before the hyphen (e.g., the area code).

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?

To match real parentheses and periods in a regex, you can use the backslash (\) as an escape character. For example, \( matches a literal opening parenthesis, and \. matches a literal period.

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

The findall() method returns a list of matching substrings if the regex pattern contains capturing groups. If the pattern has no capturing groups, it returns a list of plain strings.

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

In regular expressions, the | character represents an OR operator, allowing you to match multiple alternatives. For example, apple|banana matches either "apple" or "banana."

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

In regular expressions, the . character (dot) matches any character except a newline character. It's a wildcard character.

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

In regular expressions:

* The + character matches one or more occurrences of the preceding element.
* The \* character matches zero or more occurrences of the preceding element.

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

In regular expressions, {4} specifies exactly four repetitions of the preceding element, while {4,5} specifies a range of four to five repetitions.

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

In regular expressions, shorthand character classes signify:

* \d: Matches any digit (0-9).
* \w: Matches any word character (alphanumeric or underscore).
* \s: Matches any whitespace character (space, tab, newline, etc.).

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

In regular expressions, the shorthand character classes signify:

* \D: Matches any non-digit character.
* \W: Matches any non-word character.
* \S: Matches any non-whitespace character.

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

The .\*? is a non-greedy match that matches as few characters as possible, while .\* is a greedy match that matches as many characters as possible.

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

To match both numbers and lowercase letters with a character class, you can use [0-9a-z]. This matches any character that is a digit (0-9) or a lowercase letter (a-z).

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

To make a regular expression case insensitive, you can pass the re.IGNORECASE or re.I flag as the second argument to re.compile().

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

The . character in regular expressions normally matches any character except a newline character. If re.DOTALL (or re.S) is passed as the second argument to re.compile(), it also matches a newline character.

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

numRegex.sub('X', '11 drummers, 10 pipers, five rings, 4 hen') will 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?

Passing re.VERBOSE as the second argument to re.compile() allows you to add comments and whitespace to the regular expression pattern for better readability.

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)

To match a number with commas for every three digits, you can use the following regex: 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'

'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)

To match the full name of someone with the last name "Watanabe," you can use the regex: 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.'

To match sentences with specific words, you can use the following regex: r'(Alice|Bob|Carol)\s(eats|pets|throws)\s(apples|cats|baseballs)\.'. This regex is case-insensitive.