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?

* backslashes dont have to be escaped

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

* Match object

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

* 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?

* gets 1st set of parentheses i. e 3 digits, gets 2nd set of parentheses 3 digits - 4 digits respectively Ex: (‘111’, ‘212-3456’)

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?

* can be escaped with backslash i.e \. , \(, and\)

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

* If the regex had no groups, it returns a list of all match strings. If there are groups, return a list of tuples of string.

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

* either, or between groups

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

* Regex had basic and special characters, basic characters have no special meaning (they are general keyboard letters, numbers); special cha

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

* \* indicates 0 or more occurrences of the search term or match
* + indicates 1 or more occurrences of the search term

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

* {4} indicates regex match string exactly 4 repeats of search obj Ex: (Ma){4} = MaMaMaMa
* {4,5} indicates regex match string exactly 4 or 5 repeats of search obj (Ex: (Ma){4, 5} = MaMaMaMa | MaMaMaMaMa

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

* \d, \w, and \s match a single digit between 0 to 9, word or space character respectively

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

* \D, \W, and \S matches a single character that is not a digit, word or space character respectively

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

* .\*? and .\*? look same, question is not clear to me.

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

* [a-z0-9]

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

* re.IGNORECASE as second argument in 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 . matches any character except new line characters.
* re.DOTALL, dot match new line 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?

* it allows to add white spaces and comments to re.compile().

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][a-z]\*\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|pets|throws)\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.'