Assignment\_7

1) re.compile() function returns Regex objects.

2) Raw strings are used because backslashes should not have to be escaped.

3) The search() method returns Match objects.

4) The actual strings that match the pattern is obtained by using group() method.

5) Group0 is the entire match, group1 covers the first set of parentheses, and group2 covers the second set of parentheses.

6) Periods and parentheses can be escaped with a backslash : \., \(, and \).

7) If the regex has no groups when using findall(), a list of strings is returned. If the regex has groups then tuples of strings is returned.

8) The | character signifies matching either , or between two groups.

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

Ans→ A character set is an explicit list of the characters that may qualify for a match in a search.

10) The + matches one or more. The \* matches zero or more.

11) The{4} matches exactly four instances of the preceding group. The {4,5} matches between four and five instances.

12) The \d, \w, and \s shorthand character classes match a single digit, word, and space character, respectively.

13) The \D, \W, and \S shorthand character classes match a single character that is not a digit, word, or space character, respectively.

14) The.\*? performs a non-greedy match, and the .\* performs a greedy match.

15) [a-z0-9] or [0-9a-z] is the syntax for matching both numbers and lowercase letters with a character class.

16) Passing re.I or re.IGNORECASE as the second argument to re.compile() will make the matching case insensitive.\

17) The . character normally matches any character except the newline character. If re.DOTALL is passed as the second argument to re.compile(), then the dot will also match newline characters.

18) It will return 'X drummers, X pipers, five rings, X hens' because sub is done based on digit.

19) The re.VERBOSE argument allows us to add whitespace and comments to the string when passed as a second argument to re.compile().

20) re.compile(r'^\d{1,3}(,\d{3})\*$') will match .

21) re.compile(r”[A-Z][a-z]\*\sWatanabe”)

“””

import re

str1 = "Haruto Watanabe,Alice Watanabe,haruto Watanabe,Mr. Watanabe,Haruto watanabe"

string\_pattern = r"[A-Z][a-z]\*\sWatanabe"

regex\_pattern = re.compile(string\_pattern)

result = regex\_pattern.findall(str1)

print(result)

“””

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