Assignment 7
Solutions

1)

We can generate regex objects by passing a string to re.compile() method

2)

raw strings appear often in regex objects because they are used to avoid multiple backlashes in regex

3)

The return value of the search() method is a Match object

4)

We use the group() method to get the actual string from the Match object

5)

Here, group 0 denotes entire matched string, group 1 denotes the string in the first set and group 2 denotes the second set

6)

We can tell a regex fit real parentheses and periods by using a \ escape character

7)

The findall() method returns both a string list and a list of string tuples. It returns a string list when used with regex that do not have groups and it returns a list of string tuples when used with regex that do have groups

8)

The | character is used to match one of many expressions. It works like the “Either Or” concept

9)

10)

The + character is used to match one or more while the * character is used to match zero or more

11)

The difference between {4} and {4,5} in regular expression is that {4} will match exactly 4 instances while {4,5} will match 4 to 5 instances

12)

\d signifies numerical digit from 0 to 9

\w signifies any word character (numerical digit, alphabet or underscore)

\s signifies space, tab or newline character

13)

\D signifies character that is non-numeric digit

\W signifies character that is not a numerical digit, alphabet or underscore

\S signifies character that is not space, tab or newline character

14)

15)

The syntax to match both numbers and lowercase letters is

```
regob=re.compile(r'[0-9a-z]')
```

16)

We can pass re.I as the second argument to re.compile() method to make regex case insensitive

17)

The . character normally matches characters other than a newline character. The dot will also match newline characters if re.DOTALL is passed as 2nd argument in re.compile().

18)

It should be:
numReg.sub('X', '11 drummers, 10 pipers, five rings, 4 hen')

it will return: 'X drummers, X pipers, five rings, X hen'

19)

We can add comments and whitespaces to the strings by using re.VERBOSE as the second argument in re.compile()

20)

```
regob = re.compile(r'^\d+(,\d{3})*$')
```

21)

```
regob = re.compile(r'[A-Z][a-z]*\sWatanabe')
```

22)

```
regob =
re.compile(r'(Alice|Bob|Carol)\s(eats|pets|throws)\s(apples|c
ats|baseballs)\.',re.I)
```