Python Training.....

-- Jeetendra Bhattad

Regular Expressions

re – module:methods.

- Regular Expressions
  - Powerful String Manipulation tool
- Functions
  - match : Match regular expression pattern at the beginning of a string
  - search : Searches string for presence of pattern
  - sub : Substitue occurences of a pattern found in a string
  - subn : Same as sub, but also returns the number of replacement made
  - split : Split a string by the occurrences of a pattern
  - findall : Find all occurences of a pattern in a string
  - finditer : Return an iterator yeilding a match object for each match
  - compile : Compile a pattern into RegexObject
  - purge : Clear the regular expression cache
  - escape : Backslash all non-alphanumerics in the string

Regular Expressions
Special Characters

- Special Characters
  - "." matches any character except newline
  - "^" matches the start of string
  - "$" matches at the end or just before the newline at the end of string
  - "*" matches 0 or more repetitions of preceding RE
  - "+" matches 1 or more repetitions of preceding RE
  - "?" matches 0 or 1 occurrence of the preceding RE
  - *?, +?, ?? non greedy versions
  - {m,n} matches from to m to n repetitions of the preceding RE
  - {m,n}? non-greedy version
  - [] indicates set of character, ^ as the 1st character indicates complement
  - "|" A|B matches either A or B

Regular Expressions

Special Sequences

- Special Sequences
  - \A matches only at the start of string
  - \Z matches only at the end of string
  - \b matches empty string but only at the start or end of the word
  - \B matches empty string but not at the start or end of the word
  - \d matches any decimal digit, same as [0-9]
  - \D matches any non digit character, same as [^0-9]
  - \s matches any whitespace character, same as [\t\n\r\f\v]
  - \S matches any non-whitespace character, same as [^ \t\n\r\f\v]
  - \w matches any alphanumeric character, same as [a-zA-Z0-9_]
  - \W matches any non alphanumeric character, same as [^a-zA-Z0-9_]
  - \\ matches a literal \

Assignments

- Write a program to accept pattern and string from user & check if given pattern is present at the begining of the string.
- Write a program to accept pattern and string from user & check if given pattern is present at the end of the string.
- Write a program to accept pattern and string from user & count the occurences of the given pattern in the string.
- Write a program accept string, replaceTo string and replaceBy string. Replace all occurences of replaceTo string by replaceBy string using RE.
- Write a program to accept an alphanumeric string and remove all the characters other than digits.
- Write a program to delete python style single line comments.
- Write a program to validate email-id.

```
"\w+@\w+\.+(com|org|net|edu)"
x = re.search( "(\w+)@(\w+)\.+(com|org|net|edu)", "bhattad.jeetendra@gmail.com")
x.groups()
```

- Write a program to accept file name from user & print all single line comments from the given file.
- Write a program to split words & count the same.

Assignments

Look for each pattern in the text and print the results

either a or b

a followed by one or more a or b

a followed by one or more a or b, not greedy

sequences without -, ., or space

sequences of lower case letters

sequences of upper case letters

sequences of lower or upper case letters

one upper case letter followed by lower case letters

one upper case letter followed by one lower case

a followed by any one character

b followed by any one character

a followed by anything, ending in b

Assignments

sequence of digits

sequence of non-digits

sequence of whitespace

sequence of non-whitespace

alphanumeric characters

non-alphanumeric

alphanumeric word at start of string

alphanumeric word at end of string, with optional punctuation

word containing 's'

's' at start of word

's' at end of word

's', not start or end of word

'b' followed by literal 'ab'

'a' followed by 0-n 'a' and 0-n 'b'

'a' followed by 0-n 'ab'

'a' followed by 1-n 'ab'