Regular Expressions: python re module

Use raw strings to avoid double escapes: r'python\.org' instead of 'python\\.org'

Char	Description		
.	any character except a newline		
^ or \A	the start of the string		
$ or \Z	the end of the string		
*	**0 or more** repetitions of the preceding RE. ab* will match 'a', 'ab', or 'a' followed by any number of 'b's. Greedy, meaning as many repetitions as are possible		
+	**1 or more** repetitions of the preceding RE. ab+ will match 'a' followed by any non-zero number of 'b's; it will not match just 'a'. Greedy.		
?	**0 or 1** repetitions of the preceding RE. ab? will match either 'a' or 'ab'.		
*?, +?, ??	The '*', '+', and '?' qualifiers are all greedy; they match as much text as possible. If <.*> is matched against '<H1>title</H1>', it will match the entire string, Adding '?' is non-greedy. <.*?> matches only '<H1>'.		
{m}	exactly m copies of the previous RE		
{m,n}	from m to n repetitions of the preceding RE. If omitted, m: 0, n: infinity. Greedy		
{m,n}?	from m to n repetitions of the preceding RE. non-greedy.'aaaaaa', a{3,5} matches 5 'a' characters, while a{3,5}? matches 3 characters.		
\	Escapes special characters (characters like '*', '?', and so forth)		
[ ]	1) Characters listed individually, e.g. [amk] matches 'a', 'm', or 'k'. 2) Ranges of characters, [a-z] matches any lowercase letter, [0-5][0-9] will match all the two-digits numbers from 00 to 59, [A-Za-z] matches any alpha character		
[^]	Not what's inside		
A|B	A or B, use multiple times, such as A|B|C|D, first match and out		
(...)	Match and save the regular expression inside the parentheses as a group. Group contents can be obtained using group() method of MatchObject object.		
(?P<name> pattern)	Symbolic group where **name** is the name of the group and **pattern** is some pattern to match. The pattern (?P<quote>['"]).*?(?P=quote) matches a single or double quoted string. Groups can be referenced by number: (?P=quote)could have been \1		
\w	**[a-zA-Z0-9_]**	**\W**	**[^a-zA-Z0-9_]**
\b	**First or last character**	**\B**	**mid character**
\d	**[0-9]**	**\D**	**[^0-9]**
\s	**Any whitespace [\t\n\r\f\v]**	**\S**	**[^ \t\n\r\f\v]**