

Regex Cheat Sheet

Regex Syntax

Characters

Character Matches

a	a character
.	Any character (except newline)
\.	. character
\\	\ character
*	* character

Character Classes

	Matches	Description
[abcd]	Any one of the letters a through d	Set of characters
[^abcd]	Any character but a, b, c, or d	Complement of a set of characters
[a-d]	Any one of the letters a through d	Range of characters
[a-dz]	Any of a, b, c, d, or z	Range of characters

Special Sequences

Type	Expression	Equivalent To	Description
Word Character	\w	[a-zA-Z0-9_]	Alphanumeric or underscore
Non-word Character	\W	[^a-zA-Z0-9_]	Anything but a word character
Digit Character	\d	[0-9]	Numeric
Non-digit Character	\D	[^0-9]	Non-numeric
Whitespace Character	\s	[\t\n\r\f\v]	Whitespace
Non-whitespace Character	\S	[^\t\n\r\f\v]	Anything but a whitespace character

Anchors

Anchor Matches

^	Start of the string
\$	End of the string
\b	Boundary between word and non-word characters

Groups

Group Type	Expression
Capturing	(...)
Non-capturing	(?: ...)

Quantifiers/Repetition

Quantifier	Modification
{5}	Match expression exactly 5 times
{2,5}	Match expression 2 to 5 times
{2,}	Match expression 2 or more times
{,5}	Match expression 0 to 5 times
*	Match expression 0 or more times
{,}	Match expression 0 or more times
?	Match expression 0 or 1 times
{0,1}	Match expression 0 or 1 times
+	Match expression 1 or more times
{1,}	Match expression 1 or more times

Non-greedy quantifiers

Quantifier	Modification
{2,5}?	Match 2 to 5 times (less preferred)
{2,}?	Match 2 or more times (less preferred)
{,5}?	Match 0 to 5 times (less preferred)
*?	Match 0 or more times (less preferred)
{,}?	Match 0 or more times (less preferred)
??	Match 0 or 1 times (less preferred)
{0,1}?	Match 0 or 1 times (less preferred)
+?	Match 1 or more times (less preferred)
{1,}?	Match 1 or more times (less preferred)

Alternators

Quantifier	Modification
ABC DEF	Match string ABC or string DEF

Lookaround

Quantifier	Modification
(?=abc)	Zero-width match confirming abc will match upcoming chars
(?!abc)	Zero-width match confirming abc will not match upcoming chars

Python

functions

Function	Purpose	Usage
<code>re.search</code>	Return a match object if pattern found in string	<code>re.search(r'[pat]tern', 'string')</code>
<code>re.finditer</code>	Return an iterable of match objects (one for each match)	<code>re.finditer(r'[pat]tern', 'string')</code>
<code>re.findall</code>	Return a list of all matched strings (different when capture groups)	<code>re.findall(r'[pat]tern', 'string')</code>
<code>re.split</code>	Split string by regex delimiter & return string list	<code>re.split(r'[-]', 'st-ring')</code>
<code>re.compile</code>	Compile a regular expression pattern for later use	<code>re.compile(r'[pat]tern')</code>

flags

Flag	Description
<code>re.IGNORECASE</code>	Match uppercase and lowercase characters interchangeably
<code>re.VERBOSE</code>	Ignore whitespace characters and allow # comments