Cheatography

PCRE Cheat Sheet
by [tempusername12346455](via) via cheatography.com/170864/cs/35809/

Assertions

(?=...)	Positive Lookahead	question(?=s)	questions	question
(?!...)	Negative Lookahead	answer(?!s)	answer	answers
(?<=...)	Positive Lookbehind	(?<=appl)e	apple	application
(?<!...)	Negative Lookbehind	(?<!goo)d	mood	good
(?(condition)...)	if-then pattern	(<)?p(?(1)>)	<p>, p	<p
(?(condition)...|...)	if-then-else pattern	^(?(?=q)que|ans)	question, answer	quote

Anchors

^	Start of string or line
$	End of string, or end of line in multi-line pattern
\A	Start of string, ignores m flag
\z	End of string, ignores m flag
\Z	End of string or char before last new line, ignores m flag
\b	Word boundary; position between a word character (\w), and a nonword character (\W)
\B	Not word boundary
\G	End of the previous match or the start of the string for the first match

Flags or Modifiers

g	Global match
i	Ignore case
m	Multiple lines (^ and $ match start and end of a line)
s	Single line (. matches newline)
x	Specifies that whitespace characters in the pattern are ignored except when escaped or inside a character class
J	Duplicate group names allowed
U	Ungreedy quantifiers
A	Anchored, the pattern is forced to match only at the beginning (^)
E	Anchored, the pattern is forced to match only at the end ($)

Quantifiers

a|b	a or b
+	1 or more
*	0 or more
?	0 or 1
??	0 or 1, as few times as possible (lazy)
+?	1 or more, as few times as possible (lazy)
*?	0 or more, as few times as possible (lazy)
{n}	n times exactly
{n,m}	From n to m times
{n,}	At least n time

Lazy or Ungreedy means match the pattern the least times as possible.
Greedy means match the pattern the most times as possible.
Quantifiers (+, *) are greedy by default

Groups

(...)	Capturing
(?P<obj>...)	Capturing group named "obj"
(?P=obj)	Match the named group "obj"
\g{Z}	Match the named or numbered group Z
\Y	Match the Y'th captured group
(?>...)	Atomic group: This group does not allow any backtracking to occur

Groups (cont)

(?|...)	Duplicate group numbers
(?#...)	Comment between parenthesis
(?:...)	Matches but does not capture the group
(?&obj)	Recurse into the named group "-obj"
(?Y)	Recurse into numbered group Y
(?R)	Recurse into entire pattern
\g<-Z>	Recurse into the named or numbered group Z

Character Classes and Ranges

.	Any character except new line (\n)
[abc-f]	One character of: a, b, c, d, e, f
[^abc-f]	One character except: a, b, c, d, e, f
\d	One digit
\D	One non-digit
\s	One whitespace
\S	One non-whitespace
\w	One word [a-zA-Z0-9_]
\W	One non-word [^a-zA-Z0-9_]

Special Character Classes

\\	General Escape
\n	New line or line feed
\r	Carriage return (used to reset a device's position to the beginning of a line of text)
\t	Tab
\0	Null character
\xHH	Hexademical digit HH
\oOOO	Octal digit OOO
[\b]	Backspace (yes it's inside brackets)
\f	Form feed (next page)
\v	Vertical tab
\Q...\E	Remove special meaning from a sequence of characters

Posix Classes

[:xdigit:]	Hexadecimal digits
[:upper:]	Uppercase letters
[:lower:]	Lowercase letters
[:alpha:]	Letters
[:alnum:]	Letters and digits
[:ascii:]	Ascii codes 0 - 127
[:cntrl:]	Control characters
[:print:]	Visible characters
[:punct:]	Visible punctuation characters