

Regular Expressions [REGEX] Reference Sheet

Rebecca Ramnauth | github.com/rramnauth2220/referencesheets

Character

- [*set*] – in the set of
- [*^set*] – not in the set of
- [*a-z*] – in the *a* – *z* range
- [*^a-z*] – not in the range
- .* – any character except *\n* (new line)
- \char* – escape special character

Control Characters

- `\t` – horizontal tab
- `\v` – vertical tab
- `\b` – backspace
- `\e` – escape
- `\r` – carriage return
- `\f` – form feed
- `\n` – new line
- `\a` – bell (alarm)
- `\c char` – ASCII control character

Anchors

- ^ – start of string/ line in multi-line pattern
- \A – start of string
- \$ – end of string
- \Z – end of string
- \b – word boundary
- \B – not word boundary
- < – start of word
- > – end of word

Character Classes

- `\c` – control character
- `\s` – white space
- `\S` – not white space
- `\d` – digit
- `\D` – not digit
- `\w` – word
- `\W` – not word
- `\x` – hexadecimal digit
- `\o` – octal digit
- `\p{ctgry}` – in the Unicode category/block
- `\P{ctgry}` – not in Unicode category/block

POSIX

- `[:upper:]` – upper case letters
- `[:lower:]` – lower case letters
- `[:alpha:]` – all letters
- `[:alnum:]` – all letters
- `[:digit:]` – digits
- `[:xdigit:]` – hexadecimal digits
- `[:punct:]` – punctuation

- [**:blank:**] – space and tab
- [**:space:**] – blank characters
- [**:cntrl:**] – control characters
- [**:graph:**] – printed characters
- [**:print:**] – printed characters and spaces
- [**:word:**] – digits, letters, and underscore

Assertions

- ?= – lookahead assertion
- ?! – negative lookahead
- ?<= – lookbehind assertion
- ?!= or ?<! – negative lookbehind
- ?> – once-only subexpression
- ?() – condition [if then]
- ?()! – condition [if then else]
- ?# – comment

Substitution

- \$n** – substring matched by group number **n**
- \${name}** – substring matched by group name
- \$\$** – literal **\$** character
- &** – copy of whole match
- `** – text before the match
- '** – text after the match
- +** – last captured group
- _** – entire input string

Inline Options

- i** – case-insensitive
- m** – multiline mode
- n** – explicit (named)
- s** – single-line mode
- x** – ignore white space

Quantifiers

GREEDY	LAZY	MATCHES
*	*?	0 or more times
+	+?	1 or more times
?	??	0 or 1 time
{n}	{n}?	Exactly n times
{n, }	{n, }?	At least n times
{n,m}	{n,m}?	n to m times

Alternation

a b	either a or b
(? (exp) yes no)	yes if exp is matched no if exp isn't matched
(? (name) yes no)	yes if name is matched no if name isn't matched