Regex Reference Guide

Metacharacters

Metacharacters are characters with a special meaning:

Character	Description	Example
[]	A set of characters	"[a-m]"
\	Signals a special sequence or escapes	"\d"
.	Any character except newline	"he..o"
^	Starts with	"^hello"
$	Ends with	"planet$"
*	Zero or more occurrences	"he.*o"
+	One or more occurrences	"he.+o"
?	Zero or one occurrence	"he.?o"
{}	Exact occurrences	"he.{2}o"
|	Either or	"falls|stays"
()	Capture and group	-

Special Sequences

Special sequences start with a \ followed by a character:

Character	Description	Example
\A	Match at the start of the string	"\AThe"
\b	Match at word boundary	r"\bain", r"ain\b"
\B	Match NOT at word boundary	r"\Bain", r"ain\B"
\d	Match any digit (0-9)	"\d"

\D   Match any non-digit   "\D"

\s   Match any whitespace character   "\s"

\S   Match any non-whitespace character   "\S"

\w   Match word characters (a-z, A-Z, 0-9, _)   "\w"

\W   Match non-word characters   "\W"

\Z   Match at the end of the string   "Spain\Z"

Sets

A set is a collection of characters enclosed in []:

Set   Description

[arn]   Match any of the characters (a, r, n)

[a-n]   Match lowercase characters between a and n

[^arn]   Match any character except a, r, n

[0123]   Match any of the digits 0, 1, 2, 3

[0-9]   Match any digit between 0 and 9

[0-5][0-9]   Match numbers 00 to 59

[a-zA-Z]   Match letters a-z (case insensitive)

[+]   Match literal '+' (special chars lose meaning in sets)