

8. ADVANCED FEATURES**Character Slashes****Introduction**

We have already seen the use of the slash to match special characters:

Code	Description	Code	Description
\.	Matches the period.	\?	Matches the question mark.
\\	Matches the vertical bar.	*	Matches the wildcard star.
\(Matches the open round bracket.	\+	Matches the plus symbol.
\)	Matches the close round bracket.	\/	Matches the front slash.
\[Matches the open square bracket.	\{	Matches the open curly braces.
\]	Matches the close square bracket.	\}	Matches the close curly braces.

There are other slash codes to match with character types and white spaces:

Code	Description
\d	Matches any of the digit characters (0-9)
\D	Matches any character other than the digit ones (a-z, A-Z, @, \$, etc.)
\w	Matches any of the word characters (a-z, A-Z, and “_”)
\W	Matches any character other than the word ones (0-9, @, \$, etc.)
\s	Matches any whitespace character (space, tab, newline, etc.).
\S	Matches any character other than the whitespace ones (a-z, A-Z, 0-9, etc.)
\t	Matches the tab character
\r	Matches the return character
\n	Matches the new line character

Examples

- “X” matches pattern “\w”
- “XXX” matches pattern “\w\w\w”
- “XXX” matches pattern “\w{3}”
- “3” matches pattern “\d”
- “3X3” matches pattern “\d\w\d”
- “12.345” matches pattern “\d+\.\d+”
- “[34]” matches pattern “\[\d\d\]”
- “Array[5]” matches pattern “\w+\[\d+\]”
- “(abc)” matches pattern “\(\w\w\w\)”
- “3D ” matches pattern “\d\w\s”

Raw Notation

We mentioned that programming languages prefer to have two backslashes instead of one, so in programming the question mark character “\?” becomes “\\?”. The backslash character itself is matched to “\\” which when programming becomes “\\\\”, which is just too much. If we put the letter “r” at the start of a computer program RegEx, it goes into “raw notation” and we don’t need for the extra slashes.

- “\\?” is the same as r“\?” and “\\\\” is the same as r“\\”