

REGEX CHEAT SHEET FOR 'egrep', 'grep -E' and 'sed -r'		
META-CHARACTER	MEANING	EXAMPLES
^	beginning of line position; also exclusion if first character inside character class	'^mold' matches 'mold' and 'molded' if these words are at the beginning of the line, but not 'there was mold' or 'smoldering'. '[^0-9]' matches anything except digits 0 through 9.
		'mol[^t]' matches 'mole', 'mold' and 'smoldering' but not 'molt' or 'molten'.
\$	end of line position	'mold\$' matches 'mold' and 'there was mold' if these words are at the end of the line, but not 'molded' or 'smoldering'. '^\$' matches empty lines.
[]	character class	'mol[t d]' matches 'mold' and 'molt'. '[0-9]' matches digits 0 through 9. '[0-9^&*]' matches digits 0 through 9, as well as characters '^', '&', and '*'.
+	match one or more of preceding item	'[AT]+' matches characters 'A' or 'T' one or more times. '[1-5]+' matches digits 1 through 5 one or more times.
.	match any one character	'[AT]+...' matches uppercase 'A' or 'T' one or more times, followed three other characters.
*	match zero, one, or more of preceding item	'.*' matches any character zero or more number of times.
\<	beginning of word (i.e. alphanumeric text)	
\>	end of word (i.e. alphanumeric text)	'[a-z]+eed\>' matches 'succeed' and 'weed' but not 'succeeded' or 'weeded'.
\b	word boundary, either beginning or end	'[a-z]+eed\b' matches 'succeed' and 'weed' but not 'succeeded' or 'weeded'.
()	define scope of ' ' or '?' and backreferences	'mol(d t)' matches 'mold' or 'molt'.
 	match either side	'AC(GT AG)' matches 'ACGT' or 'ACAG'
?	match zero or one of preceding item	'data ?set' matches 'dataset' and 'data set' 'data ?set'
{ }	match item {number} of times or match {min,max} number of times	'[a-z]+[^c]c{1}e{2,}d' matches 'exceed' and 'proceed' but not 'succeed'
\	escape: force literal interpretation of metacharacter	'got \\$\\$\\$\\$?' matches 'got \$\$\$\$'
\1\2	first backreference, second backreference	In '(data) (set)' \1 stands for 'data' and \2 for 'set'.
\w	letters, numbers, underscores; same as [A-Za-z0-9_]	
\W	anything except letters, numbers, underscores, blank characters; same as [^A-Za-z0-9_]	
\s	blank characters: space, tab	
\S	non-blank characters, not space or tab	
\t	tab character	
\n	newline (linefeed)	matches line breaks in Unix (Linux, Mac OS)
\n\r	newline (linefeed + carriage return)	matches line breaks in Windows