

# REGULAR EXPRESSIONS CHEATSHEET

## Basic RegEx

Expression	Meaning
<b>word</b>	<ol style="list-style-type: none"><li>Matches to 'word'</li><li>This is the most form of regex for pattern matching</li></ol>
<b>[wW]</b>	<ol style="list-style-type: none"><li>[ _ , _ ] represents OR operator</li><li>This matches to 'w' or 'W'</li><li>[0123456789] matches to any single digit</li><li>[A-Z a-z], [0-9] are shorthands</li></ol>
<b>[^a]</b>	<ol style="list-style-type: none"><li>if '^' is first symbol after '[' then this is for <b>negation</b></li><li>This matches anything except 'a'</li><li>[^A-Z]: not uppercase character</li><li>[^\.]: not a period</li><li>[^Ss]: not s or S</li></ol>
<b>countr(y ies)</b>	<ol style="list-style-type: none"><li>The ' ' operator is for disjunction</li><li>The '()' operator is for grouping</li><li>This matches either "country" or "countries"</li></ol>

## Counting and Wildcards

Expression	Meaning
<b>_?</b>	<ol style="list-style-type: none"><li>'_?' matches to 0 or 1 instances of _</li><li>colou?r: matches to 'color' or 'colour'</li></ol>
<b>_*</b>	<ol style="list-style-type: none"><li>'_*' matches to 0 or more instances of _</li><li>'a*': matches to '', 'a', 'aa', 'aaa' ...</li></ol>
<b>_+</b>	<ol style="list-style-type: none"><li>'_+' matches to 1 or more instances of _</li><li>'a+': matches to 'a', 'aa', 'aaa' ...</li></ol>
<b>.</b>	<ol style="list-style-type: none"><li>The '.' is a wildcard that matches with anything</li><li>'.*': Matches to any string including empty string</li></ol>
<b>{n}</b>	<ol style="list-style-type: none"><li>Represents n occurrences</li><li>([a-z]{2}): Represents words with 2 repetitions</li></ol>
<b>{n,m}</b>	<ol style="list-style-type: none"><li>Represents n to m occurrences</li></ol>
<b>{n, }</b>	<ol style="list-style-type: none"><li>Represents n or more occurrences</li></ol>

## Anchors and other operators

Expression	Meaning
<code>^_</code>	<ol style="list-style-type: none"><li>1. Represents start of line</li><li>2. <code>^[0-9]+</code>: matches to strings that start with integer</li></ol>
<code>_\$</code>	<ol style="list-style-type: none"><li>1. Represents end of line</li><li>2. <code>[a-z A-Z]+\$</code>: matches to strings that end with a word</li></ol>
<code>\b.*\b</code>	<ol style="list-style-type: none"><li>1. Represents boundary</li></ol>
<code>\d</code>	<ol style="list-style-type: none"><li>1. Represents digit</li><li>2. <code>\D</code> represents non digit</li></ol>
<code>\w</code>	<ol style="list-style-type: none"><li>1. Represents alphanumeric or underscore</li><li>2. <code>\W</code> represents non alphanumeric</li></ol>
<code>\s</code>	<ol style="list-style-type: none"><li>1. Represents whitespace or newline</li><li>2. <code>\W</code> represents non whitespace</li></ol>
<b>Representing characters like ., ? etc</b>	<ol style="list-style-type: none"><li>1. If we want to represent characters used in symbols we use <code>\</code></li><li>2. <code>\?</code> represents <code>?</code>; <code>\.</code> represents <code>.</code> and so on</li></ol>