| **正規表示式** | **說明及範例** | **比對不成立之字串** |
| --- | --- | --- |
| /a/ | 含字母 “a” 的字串，例如 “ab”, “bac”, “cba” | “xyz” |
| /a./ | 含字母 “a” 以及其後任一個字元的字串，例如 “ab”, “bac”（若要比對.，請使用 \.） | “a”, “ba” |
| /^xy/ | 以 “xy” 開始的字串，例如 “xyz”, “xyab”（若要比對 ^，請使用 \^） | “axy”, “bxy” |
| /xy$/ | 以 “xy” 結尾的字串，例如 “axy”, “abxy”以 “xy” 結尾的字串，例如 “axy”, “abxy” （若要比對 $，請使用 \$） | “xya”, “xyb” |
| [13579] | 包含 “1” 或 “3” 或 “5” 或 “7” 或 “9” 的字串，例如：”a3b”, “1xy” | “y2k” |
| [0-9] | 含數字之字串 | 不含數字之字串 |
| [a-z0-9] | 含數字或小寫字母之字串 | 不含數字及小寫字母之字串 |
| [a-zA-Z0-9] | 含數字或字母之字串 | 不含數字及字母之字串 |
| b[aeiou]t | “bat”, “bet”, “bit”, “bot”, “but” | “bxt”, “bzt” |
| [^0-9] | 不含數字之字串（若要比對 ^，請使用 \^） | 含數字之字串 |
| [^aeiouAEIOU] | 不含母音之字串（若要比對 ^，請使用 \^） | 含母音之字串 |
| [^\^] | 不含 “^” 之字串，例如 “xyz”, “abc” | “xy^”, “a^bc” |

.

| 正規表示式的特定字元 | 說明 | 等效的正規表示式 |
| --- | --- | --- |
| \d | 數字 | [0-9] |
| \D | 非數字 | [^0-9] |
| \w | 數字、字母、底線 | [a-zA-Z0-9\_] |
| \W | 非 \w | [^a-zA-Z0-9\_] |
| \s | 空白字元 | [ \r\t\n\f] |
| \S | 非空白字元 | [^ \r\t\n\f] |

.

| 正規表示式 | 說明 |
| --- | --- |
| /a?/ | 零或一個 a（若要比對? 字元，請使用 \?） |
| /a+/ | 一或多個 a（若要比對+ 字元，請使用 \+） |
| /a\*/ | 零或多個 a（若要比對\* 字元，請使用 \\*） |
| /a{4}/ | 四個 a |
| /a{5,10}/ | 五至十個 a |
| /a{5,}/ | 至少五個 a |
| /a{,3}/ | 至多三個 a |
| /a.{5}b/ | a 和 b中間夾五個（非換行）字元 |

.

| 字元 | 說明 | 簡單範例 |
| --- | --- | --- |
| \ | 避開特殊字元 | /A\\*/ 可用於比對 “A\*”，其中 \* 是一個特殊字元，為避開其特殊意義，所以必須加上 “\” |
| ^ | 比對輸入列的啟始位置 | /^A/ 可比對 “Abcd” 中的 “A”，但不可比對 “aAb” |
| $ | 比對輸入列的結束位置 | /A$/ 可比對 “bcdA” 中的 “A”，但不可比對 “aAb” |
| \* | 比對前一個字元零次或更多次 | /bo\*/ 可比對 “Good boook” 中的 “booo”，亦可比對 “Good bk” 中的 “b” |
| + | 比對前一個字元一次或更多次，等效於 {1,} | /a+/ 可比對 “caaandy” 中的 “aaa”，但不可比對 “cndy” |
| ? | 比對前一個字元零次或一次 | /e?l/ 可比對 “angel” 中的 “el”，也可以比對 “angle” 中的 “l” |
| . | 比對任何一個字元（但換行符號不算） | /.n/ 可比對 “nay, an apple is on the tree” 中的 “an” 和 “on”，但不可比對 “nay” |
| (x) | 比對 x 並將符合的部分存入一個變數 | /(a\*) and (b\*)/ 可比對 “aaa and bb” 中的 “aaa” 和 “bb”，並將這兩個比對得到的字串設定至變數 RegExp.$1 和 RegExp.$2。 |
| xy | 比對 x 或 y | /a\*b\*/g 可比對 “aaa and bb” 中的 “aaa” 和 “bb” |
| {n} | 比對前一個字元 n 次，n 為一個正整數 | /a{3}/ 可比對 “lllaaalaa” 其中的 “aaa”，但不可比對 “aa” |
| {n,} | 比對前一個字元至少 n 次，n 為一個正整數 | /a{3,}/ 可比對 “aa aaa aaaa” 其中的 “aaa” 及 “aaaa”，但不可比對 “aa” |
| {n,m} | 比對前一個字元至少 n 次，至多 m 次，m、n 均為正整數 | /a{3,4}/ 可比對 “aa aaa aaaa aaaaa” 其中的 “aaa” 及 “aaaa”，但不可比對 “aa” 及 “aaaaa” |
| [xyz] | 比對中括弧內的任一個字元 | /[ecm]/ 可比對 “welcome” 中的 “e” 或 “c” 或 “m” |
| [^xyz] | 比對不在中括弧內出現的任一個字元 | /[^ecm]/ 可比對 “welcome” 中的 “w”、”l”、”o”，可見出其與 [xyz] 功能相反。（同時請注意 /^/ 與 [^] 之間功能的不同。） |
| [\b] | 比對退位字元（Backspace character） | 可以比對一個 backspace ，也請注意 [\b] 與 \b 之間的差別 |
| \b | 比對英文字的邊界，例如空格 | 例如 /\bn\w/ 可以比對 “noonday” 中的 ‘no’ ; /\wy\b/ 可比對 “possibly yesterday.” 中的 ‘ly’ |
| \B | 比對非「英文字的邊界」 | 例如, /\w\Bn/ 可以比對 “noonday” 中的 ‘on’ , 另外 /y\B\w/ 可以比對 “possibly yesterday.” 中的 ‘ye’ |
| \cX | 比對控制字元（Control character），其中 X 是一個控制字元 | /\cM/ 可以比對 一個字串中的 control-M |
| \d | 比對任一個數字，等效於 [0-9] | /[\d]/ 可比對 由 “0” 至 “9” 的任一數字 但其餘如字母等就不可比對 |
| \D | 比對任一個非數字，等效於 [^0-9] | /[\D]/ 可比對 “w” “a”… 但不可比對如 “7” “1” 等數字 |
| \f | 比對 form-feed | 若是在文字中有發生 “換頁” 的行為 則可以比對成功 |
| \n | 比對換行符號 | 若是在文字中有發生 “換行” 的行為 則可以比對成功 |
| \r | 比對 carriage return |  |
| \s | 比對任一個空白字元（White space character），等效於 [ \f\n\r\t\v] | /\s\w\*/ 可比對 “A b” 中的 “b” |
| \S | 比對任一個非空白字元，等效於 [^ \f\n\r\t\v] | /\S/\w\* 可比對 “A b” 中的 “A” |
| \t | 比對定位字元（Tab） |  |
| \v | 比對垂直定位字元（Vertical tab） |  |
| \w | 比對數字字母字元（Alphanumerical characters）或底線字母（”\_”），等效於 [A-Za-z0-9\_] | /\w/ 可比對 “.A \_!9″ 中的 “A”、”\_”、”9″。 |
| \W | 比對非「數字字母字元或底線字母」，等效於 [^A-Za-z0-9\_] | /\W/ 可比對 “.A \_!9″ 中的 “.”、” “、”!”，可見其功能與 /\w/ 恰好相反。 |
| \o*octal* | 比對八進位，其中*octal是八進位數目* | /\oocetal123/ 可比對 與 八進位的ASCII中 “123” 所相對應的字元值。 |
| \x*hex* | 比對十六進位，其中*hex是十六進位數目* | /\xhex38/ 可比對 與 16進位的ASCII中 “38” 所相對應的字元。 |

Unless an 'r' or 'R' prefix is present, escape sequences in strings are interpreted according to rules similar to those used by Standard C. The recognized escape sequences are:

| **Escape Sequence** | **Meaning** | **Notes** |
| --- | --- | --- |
| \newline | Backslash and newline ignored |  |
| \\ | Backslash (\) |  |
| \' | Single quote (') |  |
| \" | Double quote (") |  |
| \a | ASCII Bell (BEL) |  |
| \b | ASCII Backspace (BS) |  |
| \f | ASCII Formfeed (FF) |  |
| \n | ASCII Linefeed (LF) |  |
| \r | ASCII Carriage Return (CR) |  |
| \t | ASCII Horizontal Tab (TAB) |  |
| \v | ASCII Vertical Tab (VT) |  |
| \ooo | Character with octal value *ooo* | (1,3) |
| \xhh | Character with hex value *hh* | (2,3) |

Escape sequences only recognized in string literals are:

| **Escape Sequence** | **Meaning** | **Notes** |
| --- | --- | --- |
| \N{name} | Character named *name* in the Unicode database | (4) |
| \uxxxx | Character with 16-bit hex value *xxxx* | (5) |
| \Uxxxxxxxx | Character with 32-bit hex value *xxxxxxxx* | (6) |

Notes:

1.As in Standard C, up to three octal digits are accepted.

2.Unlike in Standard C, exactly two hex digits are required.

3.In a bytes literal, hexadecimal and octal escapes denote the byte with the given value. In a string literal, these escapes denote a Unicode character with the given value.

4.Changed in version 3.3: Support for name aliases [1] has been added.

5.Individual code units which form parts of a surrogate pair can be encoded using this escape sequence. Exactly four hex digits are required.

6.Any Unicode character can be encoded this way. Exactly eight hex digits are required.

<https://docs.python.org/3/reference/lexical_analysis.html#line-structure>

<http://atedev.wordpress.com/2007/11/23/%E6%AD%A3%E8%A6%8F%E8%A1%A8%E7%A4%BA%E5%BC%8F-regular-expression/>