🡪Python正则表达式分组

groups () 以tuple形式返回group (1), group (2) …内容

group () 和group (0) 一样，返回整个字符串

group (1), group (2) 返回该序号得分组得捕获内容

🡪

Python正则表达式中

. 如果被[.]包裹则.不表示匹配除换行以外得所有字符，只匹配.符号

🡪re模块方法

|  |  |
| --- | --- |
| 方法名 | 描述 |
| re.match | re.match 尝试从字符串的起始位置匹配一个模式，如果不是起始位置匹配成功的话，match()就返回none。 |
| re.search | re.search 扫描整个字符串并返回第一个成功的匹配。  re.match与re.search的区别:  re.match只匹配字符串的开始，如果字符串开始不符合正则表达式，则匹配失败，函数返回None；而re.search匹配整个字符串，直到找到一个匹配。 |
| re.sub | 替换字符串中的匹配项。 |

🡪正则表达式修饰符-可选标志

|  |  |
| --- | --- |
| 修饰符 | 描述 |
| re.I | 使匹配对大小写不敏感 |
| re.L | 做本地化识别匹配 |
| re.M | 多行匹配，影响`和$。  不指定re.M情况下，^匹配整体字符串的开头，$匹配整体字符串的结尾。  指定re.M情况下，^匹配整体字符串的开头，以及每一行的开头。$匹配整体字符串的结尾，以及每一行的结尾。 |
| re.S/re.DOTALL | 当指定时候，使.匹配包括换行在内的所有字符。  当不指定时候，使.匹配除了换行符以外的所有字符。 |
| re.U | 根据Unicode字符集解析字符。这个标志影响\w, \W \b, \B |
| re.X | 允许在正则表达式中加入注释 |

🡪正则表达式特殊字符

|  |  |
| --- | --- |
| 特殊字符 | 效果 |
| \w | 匹配字母数字及下划线 |
| \W | 匹配非字母数字及下划线以外的所有字符 |
| \s | 匹配任意空白字符，等价于 [ \t\n\r\f]。 |
| \S | 匹配任意非空字符 |
| \d | 匹配任意数字，等价于[0-9] |
| \D | 匹配任意非数字 |
| \b | 匹配一个单词边界，也就是指单词和空格间的位置。例如， 'er\b' 可以匹配"never" 中的 'er'，但不能匹配 "verb" 中的 'er'。 |
| \B | 匹配非单词边界。'er\B' 能匹配 "verb" 中的 'er'，但不能匹配 "never" 中的 'er'。 |
| ^ | 匹配字符串的开头。如果^在字符集[]中，则表示非。 |
| $ | 匹配字符串的末尾。$在字符集[]中使用只代表普通字符$. |
| . | 匹配任意字符，除了换行符，当re.DOTALL标记被指定时，则可以匹配包括换行符的任意字符。 |
| [...] | 用来表示一组字符集,单独列出：[amk] 匹配 'a'，'m'或'k' |
| [^...] | 不在[]中的字符：[^abc] 匹配除了a,b,c之外的字符。 |

🡪

Python字符串replace方法，只能使用固定字符串不能指定正则表达式。

Javascript字符串的replace方法，可以使用固定字符串，也可以指定正则表达式。

Python正则表达式模块re.sub以及re.subn可以使用正则表达进行替换。