

基本記号

記号	意味
?	0回または1回
*	0回以上
+	1回以上
{ }	指定回数
( )	グループ
	OR
^	行頭
[]	範囲
a-z	小文字
A-Z	大文字
0-9	数字

記号	意味
\d	数字
\w	文字
\s	空白文字
\D	数字以外
\W	文字以外
\S	空白文字以外
.	任意の文字
^	行頭
\$	行末

match

一致しない場合はNone

```
import re

re.match(pattern, string)
## pattern 正規表現
## string 文字列
```

```

a = re.match(r'test','testasdtest')
print(a) # 匹配成功
print(a.group()) # test
print(a.span()) # (0, 4)
print(re.match(r'test','atestasdtest')) # None

```

## search

匹配成功返回匹配对象

```

import re

## search与match
a = re.match(r'test','atestasdtest')
b = re.search(r'test','atestasdtest')
print(a) # None
print(b) # 匹配成功

```

## findall

返回所有匹配对象的列表

- `re.S` 标志

```

import re

result = re.findall(r'test','123test123test')
print(result) # ['test', 'test']

## 多行匹配
a = """aaatestaa
aaaa123"""
print(re.findall(r'test',a)) # None
print(re.findall(r'test',a,re.S)) # 匹配成功

```

## sub

替换匹配到的内容

```

import re

## sub(替换匹配到的内容)
result = re.sub('php','python','php匹配成功--php')
print(result) # python匹配成功--python

```

## split

按匹配到的内容分割字符串

```
import re

s = "itcast,java:php-php3;html"
print(re.split(r',',s))          # [], [] [] [] []
['itcast', 'java:php-php3;html']
print(re.split(r',|:|-|;',s))    # [], [], [], [], []
['itcast', 'java', 'php', 'php3', 'html']
print(re.split(r',|:|-|%',s))    # [], [], [], [], []
['itcast', 'java', 'php', 'php3;html']
```

正则表达式

- 正则表达式
- 正则表达式

```
import re

text = 'abc123def456'
## 
match = re.search(r'\d+.*',text)
print(match.group())          # abc123def456
## 
match = re.search(r'\d+.*?',text)
print(match.group())          # abc123

## 
text = "<div>content</div><p>paragraph</p>"
match = re.search(r'<.*>', text)
print(match.group())          # <div>content</div><p>paragraph</p>
## 
match = re.search(r'<.*?>', text)
print(match.group())          # <div>
```