

BeautifulSoup的CSS选择器

jupyter

CSS选择器

选择器	例子	例子描述	CSS
.class	.intro	选择 class="intro" 的所有元素。	1
#id	#firstname	选择 id="firstname" 的所有元素。	1
*	*	选择所有元素。	2
element	p	选择所有 <p> 元素。	1
element,element	div,p	选择所有 <div> 元素和所有 <p> 元素。	1
element element	div p	选择 <div> 元素内部的所有 <p> 元素。	1
element>element	div>p	选择父元素为 <div> 元素的所有 <p> 元素。	2
element+element	div+p	选择紧接在 <div> 元素之后的所有 <p> 元素。	2
[attribute]	[target]	选择带有 target 属性所有元素。	2
[attribute=value]	[target=_blank]	选择 target="_blank" 的所有元素。	2
[attribute~=value]	[title~=flower]	选择 title 属性包含单词 "flower" 的所有元素。	2
[attribute =value]	[lang =en]	选择 lang 属性值以 "en" 开头的元素。	2
:first-child	p:first-child	选择属于父元素的第一个子元素的每个 <p> 元素。	2
element1~element2	p~ul	选择前面有 <p> 元素的每个 元素。	3
[attribute^=value]	a[src^="https"]	选择其 src 属性值以 "https" 开头的每个 <a> 元素。	3
[attribute&=value]	a[src&=".pdf"]	选择其 src 属性以 ".pdf" 结尾的所有 <a> 元素。	3
[attribute*=value]	a[src*="abc"]	选择其 src 属性中包含 "abc" 子串的每个 <a> 元素。	3
:first-of-type	p:first-of-type	选择属于其父元素的首个 <p> 元素的每个 <p> 元素。	3
:last-of-type	p:last-of-type	选择属于其父元素的最后 <p> 元素的每个 <p> 元素。	3
:only-of-type	p:only-of-type	选择属于其父元素唯一的 <p> 元素的每个 <p> 元素。	3
:only-child	p:only-child	选择属于其父元素的唯一子元素的每个 <p> 元素。	3
:nth-child(n)	p:nth-child(2)	选择属于其父元素的第二个子元素的每个 <p> 元素。	3
:nth-last-child(n)	p:nth-last-child(2)	同上，从最后一个子元素开始计数。	3
:nth-of-type(n)	p:nth-of-type(2)	选择属于其父元素第二个 <p> 元素的每个 <p> 元素。	3
:nth-last-of-type(n)	p:nth-last-of-type(2)	同上，但是从最后一个子元素开始计数。	3
:last-child	p:last-child	选择属于其父元素最后一个子元素每个 <p> 元素。	3
:root	:root	选择文档的根元素。	3
:empty	p:empty	选择没有子元素的每个 <p> 元素（包括文本节点）。	3
:not(selector)	:not(p)	选择非 <p> 元素的每个元素。	3

```
In [1]: from bs4 import BeautifulSoup
```

```
In [25]: html='''
<div class="panel">
  <div class="panel-heading">
    <h4>Hello</h4>
  </div>
  <div class="panel-body">
    <ul class="list" id="list-1">
      <li class="element" data="favorite">Foo</li>
      <li class="element">Bar</li>
      <li class="element">Jay</li>
    </ul>
    <ul class="list list-small" id="list-2">
      <li class="element">Foo</li>
      <li class="element">Bar</li>
    </ul>
    <ul data="poor"></ul>
  </div>
</div>
'''

bs = BeautifulSoup(html, features='html5lib')
bs.select('ul[data]')
```

```
Out[25]: [<ul data="poor"></ul>]
```

```
In [ ]:
```

```
In [ ]:
```