

```

1 Sure, here are some of the methods of BeautifulSoup in web scraping and their explanations:
2
3 soup.find(): This method is used to find the first occurrence of a specified HTML tag in the BeautifulSoup object. You can search for tags based on their name, attributes, and text content.
4
5 soup.find_all(): This method is used to find all occurrences of a specified HTML tag in the BeautifulSoup object. You can search for tags based on their name, attributes, and text content.
6
7 soup.select(): This method is used to select HTML elements based on CSS selectors. You can use CSS selectors to select tags based on their attributes, class names, and parent-child relationships.
8
9 soup.get_text(): This method is used to extract the text content of an HTML document. It returns all the text in the document, including the text inside tags.
10
11 soup.prettify(): This method is used to format the HTML document in a human-readable way. It adds line breaks and indentation to the document to make it easier to read.
12
13 soup.contents: This method returns a list of all the direct child elements of an HTML tag. It can be used to iterate over the child elements of a tag and extract their contents.
14
15 soup.parent: This method returns the parent element of an HTML tag. It can be used to navigate up the HTML document and extract information from parent elements.
16
17 soup.next_sibling: This method returns the next sibling element of an HTML tag. It can be used to navigate to the next sibling element and extract information from it.
18
19 soup.previous_sibling: This method returns the previous sibling element of an HTML tag. It can be used to navigate to the previous sibling element and extract information from it.
20
21 These are just some of the many methods available in BeautifulSoup. Using these methods, you can extract information from HTML documents and scrape data from websites.
22

```

```

1 #find() method: This method returns the first matching tag, given the tag name and attributes.
2 from bs4 import BeautifulSoup
3
4 html_doc = """
5 <html>
6 <head>
7 <title>My Website</title>
8 </head>
9 <body>
10 <div id="content">
11 <h1>Welcome to my website!</h1>
12 <p>This is a paragraph of text.</p>
13 <a href="http://www.google.com">Google</a>
14 </div>
15 </body>
16 </html>
17 """
18 soup = BeautifulSoup(html_doc, 'html.parser')
19 # Find the first <div> tag with id="content"
20 div_tag = soup.find('div', {'id': 'content'})
21 print(div_tag)
22

```

```

<div id="content">
<h1>Welcome to my website!</h1>
<p>This is a paragraph of text.</p>
<a href="http://www.google.com">Google</a>
</div>

```

```

1 #find_all() method: This method returns a list of all matching tags, given the tag name and attributes.
2 from bs4 import BeautifulSoup

```

```

3
4 html_doc = """
5 <html>
6 <head>
7 <title>My Website</title>
8 </head>
9 <body>
10 <div class="article">
11 <h2>Article Title</h2>
12 <p>This is the first paragraph.</p>
13 <p>This is the second paragraph.</p>
14 </div>
15 <div class="article">
16 <h2>Another Article Title</h2>
17 <p>This is another paragraph.</p>
18 </div>
19 </body>
20 </html>
21 """
22
23 soup = BeautifulSoup(html_doc, 'html.parser')
24
25 # Find all <div> tags with class="article"
26 #div_tags = soup.find_all('div', {'class': 'article'})
27 div_tags = soup.find_all('h2')
28 for div_tag in div_tags:
29     print(div_tag)
30

```

```

<h2>Article Title</h2>
<h2>Another Article Title</h2>

```

```

1 #select() method: This method allows you to use CSS selectors to find matching tags.
2 from bs4 import BeautifulSoup
3
4 html_doc = """
5 <html>
6 <head>
7 <title>My Website</title>
8 </head>
9 <body>
10 <div class="article">
11 <h2>Article Title</h2>
12 <p>This is the first paragraph.</p>
13 <p>This is the second paragraph.</p>
14 </div>
15 <div class="article">
16 <h2>Another Article Title</h2>
17 <p>This is another paragraph.</p>
18 </div>
19 </body>
20 </html>
21 """
22
23 soup = BeautifulSoup(html_doc, 'html.parser')
24

```

```
25 # Find all <div> tags with class="article"
26 #div_tags = soup.select('div.article')
27 div_tags = soup.select('h2')
28 for div_tag in div_tags:
29     print(div_tag)
30
```

```
<h2>Article Title</h2>
<h2>Another Article Title</h2>
```

```
1 #Accessing tag attributes: You can use the dictionary-like syntax to access the attributes of a tag.from bs4 import BeautifulSoup
2
3 html_doc = """
4 <html>
5 <head>
6 <title>My Website</title>
7 </head>
8 <body>
9 <a href="http://www.google.com">Google</a>
10 </body>
11 </html>
12 """
13
14 soup = BeautifulSoup(html_doc, 'html.parser')
15
16 # Find the <a> tag and print its href attribute
17 a_tag = soup.find('a')
18 print(a_tag['href'])
19
```

<http://www.google.com>

```
1 #Accessing tag text: You can use the .text property to access the text content of a tag.
2 from bs4 import BeautifulSoup
3
4 html_doc = """
5 <html>
6 <head>
7 <title>My Website</title>
8 </head>
9 <body>
10 <div id="content">
11 <h1>Welcome to my website!</h1>
12 <p>This is a paragraph of text.</p>
13 <a href="http://www.google.com">Google</a>
14 </div>
15 </body>
16 </html>
17 """
18
19 soup = BeautifulSoup(html_doc, 'html.parser')
20
21 # Find the <h1> tag and print its text content
22 h1_tag = soup.find('h1')
```

24

Welcome to my website!

```
1 from bs4 import BeautifulSoup
2 import requests
3 #get data from html page to a file text
4 f=requests.get("https://www.timesjobs.com/candidate/job-search.html?searchType=personalizedSearch&from=submit&txtKeywords=Python&txtLocation=Mumbai&cboWorkExp1=0").text
5
6 soup=BeautifulSoup(f,'lxml')
7 print(soup)
```

```
<!DOCTYPE html>
<html><head>
<link href="https://fonts.googleapis.com/css?family=Poppins:400,500,600,700" rel="stylesheet"/>
<link href="https://fonts.googleapis.com/icon?family=Material+Icons" rel="stylesheet"/>
<link href="https://static.timesjobs.com/newtj_css/css/tj_css/usability/materialize.css?v=7.1.7" media="all" rel="stylesheet" type="text/css"/>
<link href="https://static.timesjobs.com/newtj_css/css/tj_css/usability/global-usability.css?v=7.1.7" media="all" rel="stylesheet" type="text/css"/>
<link href="https://static.timesjobs.com/newtj_css/css/tj_css/usability/srp-usability.css?v=7.1.7" media="all" rel="stylesheet" type="text/css"/>
<script src="https://static.timesjobs.com/newtj_js/scripts/tj_scripts/usability/jquery-3.3.1.min.js" type="text/javascript"></script>
<script src="https://static.timesjobs.com/newtj_js/scripts/jquery.tokeninput.js" type="text/javascript"></script>
<!-- <script type="text/javascript" src="https://code.angularjs.org/1.7.0/angular.min.js"></script> -->
<script src="https://static.timesjobs.com/newtj_js/scripts/tj_scripts/usability/angular1.7.0.min.js" type="text/javascript"></script>
<script src="https://static.timesjobs.com/newtj_js/scripts/tj_scripts/usability/materialize.min.js" type="text/javascript"></script>
<script src="https://static.timesjobs.com/newtj_js/scripts/groupjs/dwrGroup.js?v=7.1.7" type="text/javascript"></script>
<script src="https://static.timesjobs.com/newtj_js/scripts/tj_scripts/mustache.js" type="text/javascript"></script>
<link href="https://www.timesjobs.com/candidate/job-search.html" rel="canonical"/>
<link href="https://www.timesjobs.com/candidate/job-search.html?searchType=personalizedSearch&from=submit&txtKeywords=Python&txtLocation=Mumbai&cboworkExp1=0&sequence=2&...">
<title>Python Jobs in Mumbai with Entry Level experience - TimesJobs</title>
<meta content="Jobs in Python in Mumbai with Entry Level experience. Search and apply jobs online at Timesjobs.com" name="description"/>
<!-- Google Tag Manager -->
<script>(function(w,d,s,l,i){w[l]=w[l]||[];w[l].push({'gtm.start':
new Date().getTime(),event:'gtm.js'});var f=d.getElementsByTagName(s)[0],
j=d.createElement(s),dl=l!='dataLayer'? '&l='+l:'';j.async=true;j.src=
'https://www.googletagmanager.com/gtm.js?id='+i+dl;f.parentNode.insertBefore(j,f);
})(window,document,'script','dataLayer','GTM-59SJDG4');</script>
<!-- End Google Tag Manager -->
<script src="https://www.gstatic.com/charts/loader.js" type="text/javascript"></script>
<script>
//google.charts.load('current', {'packages':['bar']});
google.charts.load('current', {'packages': ['corechart', 'bar']});
</script>
<script>!function(e){var n="https://s.go-mpulse.net/boomerang/";if("False"=="True")e.BOOMR_config=e.BOOMR_config||{},e.BOOMR_config.PageParams=e.BOOMR_config.PageParams||{},e.BOOMR_config.PageP
<body>
<div id="site">
<!-- Header Starts -->
<!-- Google Tag Manager (noscript) -->
<noscript>
<iframe height="0" src="https://www.googletagmanager.com/ns.html?id=GTM-59SJDG4" style="display:none;visibility:hidden" width="0"></iframe>
</noscript>
<!-- End Google Tag Manager (noscript) -->
<meta content="text/html; charset=utf-8" http-equiv="Content-Type"/>
<script src="https://wchat.freshchat.com/js/widget.js" type="text/javascript"></script>
<script>
window.fcWidget.init({
token: "d7130618-a917-400b-804c-bdf44d8d178b",
host: "https://wchat.freshchat.com",
config: {
headerProperty: {
```

```

hideChatButton: true
}
    }
});

</script>
<!--[if IE]>
<script src="scripts/html5shiv.js"></script>
<![endif]-->
<script src="https://static.timesjobs.com/newtj_js/scripts/ajax-upload.js" type="text/javascript"></script>

```

1

```

1 from bs4 import BeautifulSoup
2 import requests
3 #get data from html page to a file text
4 f=requests.get("https://www.timesjobs.com/candidate/job-search.html?searchType=personalizedSearch&from=submit&txtKeywords=python&txtLocation=mumbai&cboworkExp1=0").text
5
6 soup=BeautifulSoup(f,'lxml')
7 tags=soup.find_all('span')
8 for i in tags:
9     print(i.text)
10 #print(soup)

```

close

Live Chat (10am to 6pm IST)

1

Job Alerts(10)

3 hrs ago

Morgan Stanley Pvt Ltd

(12)

Crimson Interactive Pvt Ltd

(3)

Hiotron India Pvt Ltd

(3)

Verisk Analytics India Private..

(3)

Boyen Haddin Consultants LLP

(2)

Dg7

(2)

Interactive Brokers

(2)

Larsen & Toubro Infotech L..

(2)

Pagalguy

(2)

Publicis Groupe

(2)

Rangam Infotech Pvt. Ltd.

(2)

Tee Security

```
1 tag=soup.find_all('ul')
2 for i in tag:
3     print(i.text)
```

TimesJobs
JobBuzz

Gulf Jobs
Career Insight

call0120 6358222

Live Chat (10am to 6pm IST)

notifications_none
1

To see all recommended jobs, profile views, recruiter's contacted information, updates and more...

Login
or
Register

Login to View Notifications!

Forgot Password?

Login-ID/Password cannot be left blank. Enter the full email ID (like: ram@gmail.com), if you are using registered email-id for login.

Job Alerts(10)

Maxgen Technologies Pvt Ltd

```
1 tag=soup.find_all('ul',class_="list-job-dtl clearfix")
2 for i in tag:
3     print(i.text)
```

Job Description:

Maxgen technologies pvt ltd offers live project internships in Mumbai. We are offering internships in core python , django and machine learning.benefit to join maxgen ?Colleg... More Details

KeySkills:

python , Css

directions_walk

Walk-in Details 10-Apr-2023 - 08-Jun-2023 09:00 AM - 04:00 PM Maxgen Technologies Pvt Ltd.

Ambiance 1505,Sector 19 D,
Opposite R.T.O,
Backside Of Fortune Hotel,
Vashi , Navi Mumbai

Job Description:

Maxgen technologies pvt ltd offers live project internships in Mumbai. We are offering internships in python , django and machine learning.Reasons behind choosing us?1 - Real... More Details

KeySkills:

python , new

directions_walk

Walk-in Details 13-Mar-2023 - 12-May-2023 09:30 AM - 05:00 PM Maxgen Technologies Pvt Ltd.

Ambiance 1505,Sector 19 D,
Opposite R.T.O,
Backside Of Fortune Hotel,
Vashi , Navi Mumbai

Job Description:

Maxgen technologies pvt ltd offers live project internships in Mumbai. We are offering internships in python , django and machine learning.benefit to join maxgen ?College pro... More Details

KeySkills:

python , hire

directions_walk

Walk-in Details 29-Mar-2023 - 27-May-2023 10:30 AM - 04:30 PM Maxgen Technologies Pvt Ltd.
Ambiance 1505,Sector 19 D,
Opposite R.T.O,
Backside Of Fortune Hotel,
Vashi , Navi Mumbai

Job Description:

```
1 from bs4 import BeautifulSoup
2 import requests
3 #get data from html page to a file text
4 f=requests.get("https://www.timesjobs.com/candidate/job-search.html?searchType=personalizedSearch&from=submit&txtKeywords=python&txtLocation=mumbai&cboworkExp1=0").text
5
6 soup=BeautifulSoup(f,'lxml')
7 print(soup)
8 job=soup.find('h3',class_="joblist-comp-name")
9 #company=job.find('h3')
10 print(company.text.strip())
```

```
<!DOCTYPE html>
<html><head>
<link href="https://fonts.googleapis.com/css?family=Poppins:400,500,600,700" rel="stylesheet"/>
<link href="https://fonts.googleapis.com/icon?family=Material+Icons" rel="stylesheet"/>
<link href="https://static.timesjobs.com/newtj_css/css/tj_css/usability/materialize.css?v=7.1.7" media="all" rel="stylesheet" type="text/css"/>
<link href="https://static.timesjobs.com/newtj_css/css/tj_css/usability/global-usability.css?v=7.1.7" media="all" rel="stylesheet" type="text/css"/>
<link href="https://static.timesjobs.com/newtj_css/css/tj_css/usability/srp-usability.css?v=7.1.7" media="all" rel="stylesheet" type="text/css"/>
<script src="https://static.timesjobs.com/newtj_js/scripts/tj_scripts/usability/jquery-3.3.1.min.js" type="text/javascript"></script>
<script src="https://static.timesjobs.com/newtj_js/scripts/jquery.tokeninput.js" type="text/javascript"></script>
<!-- <script type="text/javascript" src="https://code.angularjs.org/1.7.0/angular.min.js"></script> -->
<script src="https://static.timesjobs.com/newtj_js/scripts/tj_scripts/usability/angular1.7.0.min.js" type="text/javascript"></script>
<script src="https://static.timesjobs.com/newtj_js/scripts/tj_scripts/usability/materialize.min.js" type="text/javascript"></script>
<script src="https://static.timesjobs.com/newtj_js/scripts/groupJs/dwrGroup.js?v=7.1.7" type="text/javascript"></script>
<script src="https://static.timesjobs.com/newtj_js/scripts/tj_scripts/mustache.js" type="text/javascript"></script>
<link href="https://www.timesjobs.com/candidate/job-search.html" rel="canonical"/>
<link href="https://www.timesjobs.com/candidate/job-search.html?searchType=personalizedSearch&from=submit&txtKeywords=python&txtLocation=mumbai&cboworkExp1=0&sequence=2&sequence=2" rel="canonical"/>
<title>python Jobs in mumbai with Entry Level experience - TimesJobs</title>
<meta content="Jobs in python in mumbai with Entry Level experience. Search and apply jobs online at Timesjobs.com" name="description"/>
<!-- Google Tag Manager -->
<script>(function(w,d,s,l,i){w[l]=w[l]||[];w[l].push({'gtm.start':
new Date().getTime(),event:'gtm.js'});var f=d.getElementsByTagName(s)[0],
```



```

j=d.createElement(s),dl=l!='dataLayer'?'&l='+l:'';j.async=true;j.src=
'https://www.googletagmanager.com/gtm.js?id='+i+dl;f.parentNode.insertBefore(j,f);
})(window,document,'script','dataLayer','GTM-59SJDG4');
<!-- End Google Tag Manager -->
<script src="https://www.gstatic.com/charts/loader.js" type="text/javascript"></script>
<script>
    //google.charts.load('current', {'packages':['bar']});
    google.charts.load('current', {packages: ['corechart', 'bar']});
</script>
<script>!function(e){var n="https://s.go-mpulse.net/boomerang/";if("False"=="True")e.BOOMR_config=e.BOOMR_config||{},e.BOOMR_config.PageParams=e.BOOMR_config.PageParams||{},e.BOOMR_config.PageP
<body>
<div id="site">
<!-- Header Starts -->
<!-- Google Tag Manager (noscript) -->
<noscript>
<iframe height="0" src="https://www.googletagmanager.com/ns.html?id=GTM-59SJDG4" style="display:none;visibility:hidden" width="0"></iframe>
</noscript>
<!-- End Google Tag Manager (noscript) -->
<meta content="text/html; charset=utf-8" http-equiv="Content-Type"/>
<script src="https://wchat.freshchat.com/js/widget.js" type="text/javascript"></script>
<script>
window.fcWidget.init({
token: "d7130618-a917-400b-804c-bdf44d8d178b",
host: "https://wchat.freshchat.com",
config: {
headerProperty: {
hideChatButton: true
}
}
});

</script>
<!--[if IE]>
<script src="scripts/html5shiv.js"></script>
<![endif]>
<script src="https://static.timesinhs.com/newti is/scripts/ajax-unload.js" type="text/javascript"></script>

```

