

# children\_tags\_solution

March 19, 2023

## 1 TODO: Get The Children from the <title> Tag

In the cell below, print the contents and the number of children of the <title> tag in the sample2.html file. Start by opening the sample2.html file and passing the open filehandle to the BeautifulSoup constructor using the lxml parser. Save the BeautifulSoup object returned by the constructor in a variable called page\_content. Then access the <title> tag and save the tag object in variable called page\_title. Then use the .contents attribute to print the contents and the number of children of the <title> tag.

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

        # Open the HTML file and create a BeautifulSoup Object
        with open('./sample2.html') as f:
            page_content = BeautifulSoup(f, 'lxml')

        # Access the title tag
        page_title = page_content.head.title

        # Print the children of the title tag
        print(page_title.contents)

        # Print the number of children of the title tag
        print('\nThe <title> contains {} children'.format(len(page_title.contents)))

['AI For Trading']
```

The <title> contains 1 children

## 2 TODO: Loop Through The Children The <title> Tag

In the cell below, print the children of the <title> tag in the sample2.html file. Start by opening the sample2.html file and passing the open filehandle to the BeautifulSoup constructor using the lxml parser. Save the BeautifulSoup object returned by the constructor in a variable called page\_content. Then create a loop that prints the children of the <title> tag using the .children attribute.

```
In [2]: # Import BeautifulSoup
        from bs4 import BeautifulSoup

        # Open the HTML file and create a BeautifulSoup Object
        with open('./sample2.html') as f:
            page_content = BeautifulSoup(f, 'lxml')

        # Print the children of the head tag
        for child in page_content.head.title.children:
            print(child)
```

AI For Trading

### 3 TODO: Search For The <head> Tag

In the cell below, search for the <head> tag only in the direct children of the <html> tag in the sample2.html file. Start by opening the sample2.html file and passing the open filehandle to the BeautifulSoup constructor using the lxml parser. Save the BeautifulSoup object returned by the constructor in a variable called page\_content. Then search the html tag's direct children for the <head> tag using the recursive=False argument. Print the result using the .pretty() attribute.

```
In [3]: # Import BeautifulSoup
        from bs4 import BeautifulSoup

        # Open the HTML file and create a BeautifulSoup Object
        with open('./sample2.html') as f:
            page_content = BeautifulSoup(f, 'lxml')

        # Search the html tag's direct children for the head tag
        for tag in page_content.html.find_all('head', recursive = False):
            print(tag.pretty())
```

```
<head>
  <title>
    AI For Trading
  </title>
  <meta charset="utf-8"/>
  <link href="./teststyle.css" rel="stylesheet"/>
  <style>
    .h2style {background-color: tomato;color: white;padding: 10px;}
  </style>
</head>
```