

```
from bs4 import BeautifulSoup
```

```
import requests
```

```
response=requests.get('https://books.toscrape.com/index.html')
response.text

k rel="shortcut icon" href="static/oscar/favicon.ico" />\n
\n
\n
\n
<link rel="stylesheet" type="text/css" href="static/oscar/css/styles.css" />\n
\n
<link rel="stylesheet" href="static/oscar/js/bootstrap-datetimepicker/bootstrap-datetimepicker.css" />\n
<link rel="stylesheet" type="text/css" href="static/oscar/css/datetimepicker.css" />\n\n\n\n
\n
\n\n
\n
</head>\n\n
<body id="default" class="default">\n
\n
\n
\n
<header class="header container-fluid">\n
<div class="page_inner">\n
<div class="row">\n
<div class="col-sm-8 h1"><a href="index.html">Books to Scrape</a><small> We love being scraped!</small>\n</div>\n\n
</div>\n
</div>\n
</header>\n\n
\n
<div class="container-fluid page">\n
<div class="page_inner">\n
<ul class="breadcrumb">\n
<li>\n
<a href="index.html">Home</a>\n
</li>\n
<li class="active">All products</li>\n
</ul>\n\n
<div class="row">\n\n
<aside class="sidebar col-sm-4 col-md-3">\n
\n
<div id="promotions_left">\n
\n
</div>\n
\n
<div class="side_categories">\n
<ul class="nav nav-list">\n
\n
<li>\n
```

```
Soup=BeautifulSoup(requests.get('https://books.toscrape.com/index.html').text, 'html.parser')
Soup
<link href="static/oscar/css/styles.css" rel="stylesheet" type="text/css" /
>
<link href="static/oscar/js/bootstrap-datetimepicker/bootstrap-datetimepic
ker.css" rel="stylesheet"/>
<link href="static/oscar/css/datetimepicker.css" rel="stylesheet" type="te
xt/css"/>
</head>
<body class="default" id="default">
<header class="header container-fluid">
<div class="page_inner">
<div class="row">
<div class="col-sm-8 h1"><a href="index.html">Books to Scrape</a><small> W
e love being scraped!</small>
</div>
</div>
</div>
</header>
<div class="container-fluid page">
<div class="page_inner">
<ul class="breadcrumb">
```

In [14]:

```
Soup.find_all('article',attrs={'class':'product_pod'})
```

```
</p>
<form>
  <button class="btn btn-primary btn-block" data-loading-text="Adding..." t
ype="submit">Add to basket</button>
</form>
</div>
</article>,
<article class="product_pod">
  <div class="image_container">
    <a href="catalogue/tipping-the-velvet_999/index.html"></a>
  </div>
  <p class="star-rating One">
    <i class="icon-star"></i>
    <i class="icon-star"></i>
    <i class="icon-star"></i>
    <i class="icon-star"></i>
    <i class="icon-star"></i>
  </p>
  </div>
```

In [15]:

```
name=Soup.find_all('article',attrs={'class':'product_pod'})[0].find('h3').find('a').get('ti
name
```

Out[15]:

```
'A Light in the Attic'
```

In [19]:

```
rating_dic={'One':1,'Two':2,'Three':3,'Four':4,'Five':5}
rate=rating_dic[Soup.find_all('article',attrs={'class':'product_pod'})[0].find('p').get('cl
rate
```

Out[19]:

```
3
```

In [24]:

```
p=float(Soup.find_all('article',attrs={'class':'product_pod'})[0].find('div',attrs={'class'
p
```

Out[24]:

```
51.77
```

In [26]:

```
category=Soup.find('title').get_text().split('|')[0].strip()
category
```

Out[26]:

```
'All products'
```

In [27]:

```
name=Soup.find_all('article',attrs={'class':'product_pod'})[1].find('h3').find('a').get('title')
name
```

Out[27]:

'Tipping the Velvet'

In [35]:

```
rating_dic={'One':1,'Two':2,'Three':3,'Four':4,'Five':5}
rate=rating_dic[Soup.find_all('article',attrs={'class':'product_pod'})[1].find('p').get('class')]
rate
```

Out[35]:

1

In [29]:

```
p=float(Soup.find_all('article',attrs={'class':'product_pod'})[1].find('div',attrs={'class':'product_pod'})[1].find('p').get('text'))
p
```

Out[29]:

53.74

In [30]:

```
category=Soup.find('title').get_text().split('|')[1].strip()
category
```

Out[30]:

'Books to Scrape - Sandbox'

In [31]:

```
name=Soup.find_all('article',attrs={'class':'product_pod'})[2].find('h3').find('a').get('title')
name
```

Out[31]:

'Soumission'

In [32]:

```
rating_dic={'One':1,'Two':2,'Three':3,'Four':4,'Five':5}
rate=rating_dic[Soup.find_all('article',attrs={'class':'product_pod'})[2].find('p').get('class')]
rate
```

Out[32]:

1

In [34]:

```
import csv
```

In [40]:

```
with open('books.csv','w',newline="")as f:
    writer=csv.writer(f)
    writer.writerow(['Book Name','Price','Rate','Category'])
    category=Soup.find('title').get_text().split('|')[0].strip()
    for book in Soup.find_all('article',attrs={'class':'product_pod'}):
        book_name=book.find('h3').find('a').get('title')
        rate=rating_dic[book.find('p').get('class')[1]]
        price=float(book.find('div',attrs={'class':'product_price'}).find('p',attrs={'class':
        writer.writerow([book_name,price,rate,category])
```

In [41]:

```
import pandas as pd
```

In [42]:

```
df=pd.read_csv('books.csv')
df
```

Out[42]:

	Book Name	Price	Rate	Category
0	A Light in the Attic	51.77	3	All products
1	Tipping the Velvet	53.74	1	All products
2	Soumission	50.10	1	All products
3	Sharp Objects	47.82	4	All products
4	Sapiens: A Brief History of Humankind	54.23	5	All products
5	The Requiem Red	22.65	1	All products
6	The Dirty Little Secrets of Getting Your Dream...	33.34	4	All products
7	The Coming Woman: A Novel Based on the Life of...	17.93	3	All products
8	The Boys in the Boat: Nine Americans and Their...	22.60	4	All products
9	The Black Maria	52.15	1	All products
10	Starving Hearts (Triangular Trade Trilogy, #1)	13.99	2	All products
11	Shakespeare's Sonnets	20.66	4	All products
12	Set Me Free	17.46	5	All products
13	Scott Pilgrim's Precious Little Life (Scott Pi...	52.29	5	All products
14	Rip it Up and Start Again	35.02	5	All products
15	Our Band Could Be Your Life: Scenes from the A...	57.25	3	All products
16	Olio	23.88	1	All products
17	Mesaerion: The Best Science Fiction Stories 18...	37.59	1	All products
18	Libertarianism for Beginners	51.33	2	All products
19	It's Only the Himalayas	45.17	2	All products

In []: