

# A simple WebScraper with BeautifulSoup

May 7, 2023

# A simple Webscrapper with BeautifulSoup

This code scrapes a website for book information and stores it in a pandas DataFrame for further analysis. It uses the requests and BeautifulSoup libraries to retrieve and parse HTML content, and extracts book titles, prices, and availabilities from the webpage using BeautifulSoup. The resulting DataFrame contains information about all of the books on the webpage.

```
[2]: # Import the necessary libraries

import requests # used for making HTTP requests
from bs4 import BeautifulSoup # used for web scraping
import pandas as pd # used for data manipulation and analysis
```

This part of the code imports the necessary libraries to scrape data from a website, parse the HTML content, and manipulate the data using pandas.

```
[3]: # Set the URL of the website we want to scrape

url = "https://books.toscrape.com/"
```

```
[4]: # Send an HTTP GET request to the specified URL and get the response

response = requests.get(url)
```

```
[21]: response
```

```
[21]: <Response [200]>
```

```
[5]: # Extract the content of the HTTP response

content = response.content
```

```
[23]: content
```

```
[23]: b'<!DOCTYPE html>\n<!--[if lt IE 7]>      <html lang="en-us" class="no-js lt-ie9
lt-ie8 lt-ie7"> <![endif]-->\n<!--[if IE 7]>      <html lang="en-us"
class="no-js lt-ie9 lt-ie8"> <![endif]-->\n<!--[if IE 8]>      <html
lang="en-us" class="no-js lt-ie9"> <![endif]-->\n<!--[if gt IE 8]><!--> <html
lang="en-us" class="no-js"> <!--<![endif]-->\n      <head>\n          <title>\n
```

```

All products | Books to Scrape - Sandbox\n</title>\n\n      <meta http-
equiv="content-type" content="text/html; charset=UTF-8" /\n      <meta
name="created" content="24th Jun 2016 09:29" /\n      <meta
name="description" content="" /\n      <meta name="viewport"
content="width=device-width" /\n      <meta name="robots"
content="NOARCHIVE,NOCACHE" /\n\n      <!-- Le HTML5 shim, for IE6-8 support
of HTML elements -->\n      <!--[if lt IE 9]>\n      <script
src="//html5shim.googlecode.com/svn/trunk/html5.js"></script>\n
<![endif]-->\n\n      \n      <link rel="shortcut icon"
href="static/oscar/favicon.ico" /\n      \n\n      \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\n      \n      \n\n      \n\n
</head>\n\n      <body id="default" class="default">\n      \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      \n      </div>\n
</div>\n      </header>\n\n      \n      \n<div class="container-fluid page">\n
<div class="page_inner">\n      \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      \n      \n      \n      <div class="side_categories">\n      <ul
class="nav nav-list">\n      \n      <li>\n
<a href="catalogue/category/books_1/index.html">\n      \n      \n      \n
Books\n      \n      </a>\n\n      <li>\n
<a href="catalogue/category/books/travel_2/index.html">\n
\n      \n      \n      \n      Travel\n      \n      \n
</a>\n\n      </li>\n      \n      <li>\n      <a
href="catalogue/category/books/mystery_3/index.html">\n
\n      \n      \n      \n      Mystery\n      \n      \n
</a>\n\n      </li>\n      \n      <li>\n      <a
href="catalogue/category/books/historical-fiction_4/index.html">\n
\n      \n      \n      \n      Historical Fiction\n      \n      \n
</a>\n\n      </li>\n      \n      <li>\n      <a
href="catalogue/category/books/sequential-art_5/index.html">\n
\n      \n      \n      \n      Sequential Art\n      \n      \n
</a>\n\n      </li>\n      \n      <li>\n      <a

```

```

href="catalogue/category/books/classics_6/index.html">\n
\n                                Classics\n                                \n
</a>\n\n                        </li>\n                                \n
\n                                <li>\n                                <a
href="catalogue/category/books/philosophy_7/index.html">\n
\n                                Philosophy\n                                \n
</a>\n\n                        </li>\n                                \n
\n                                <li>\n                                <a
href="catalogue/category/books/romance_8/index.html">\n
\n                                Romance\n                                \n
</a>\n\n                        </li>\n                                \n
\n                                <li>\n                                <a
href="catalogue/category/books/womens-fiction_9/index.html">\n
\n                                Womens Fiction\n                                \n
</a>\n\n                        </li>\n                                \n
\n                                <li>\n                                <a
href="catalogue/category/books/fiction_10/index.html">\n
\n                                Fiction\n                                \n
</a>\n\n                        </li>\n                                \n
\n                                <li>\n                                <a
href="catalogue/category/books/childrens_11/index.html">\n
\n                                Childrens\n                                \n
</a>\n\n                        </li>\n                                \n
\n                                <li>\n                                <a
href="catalogue/category/books/religion_12/index.html">\n
\n                                Religion\n                                \n
</a>\n\n                        </li>\n                                \n
\n                                <li>\n                                <a
href="catalogue/category/books/nonfiction_13/index.html">\n
\n                                Nonfiction\n                                \n
</a>\n\n                        </li>\n                                \n
\n                                <li>\n                                <a
href="catalogue/category/books/music_14/index.html">\n
\n                                Music\n                                \n
</a>\n\n                        </li>\n                                \n
\n                                <li>\n                                <a
href="catalogue/category/books/default_15/index.html">\n
\n                                Default\n                                \n
</a>\n\n                        </li>\n                                \n
\n                                <li>\n                                <a
href="catalogue/category/books/science-fiction_16/index.html">\n
\n                                Science Fiction\n
\n                                </a>\n\n                                </li>\n
\n                                \n                                <li>\n                                <a
href="catalogue/category/books/sports-and-games_17/index.html">\n
\n                                Sports and Games\n
\n                                </a>\n\n                                </li>\n

```

```

\n                \n                <li>\n                <a
href="catalogue/category/books/add-a-comment_18/index.html">\n
\n                Add a comment\n                \n
</a>\n\n                </li>\n                \n
\n                <li>\n                <a
href="catalogue/category/books/fantasy_19/index.html">\n
\n                Fantasy\n                \n
</a>\n\n                </li>\n                \n
\n                <li>\n                <a
href="catalogue/category/books/new-adult_20/index.html">\n
\n                New Adult\n                \n
</a>\n\n                </li>\n                \n
\n                <li>\n                <a
href="catalogue/category/books/young-adult_21/index.html">\n
\n                Young Adult\n                \n
</a>\n\n                </li>\n                \n
\n                <li>\n                <a
href="catalogue/category/books/science_22/index.html">\n
\n                Science\n                \n
</a>\n\n                </li>\n                \n
\n                <li>\n                <a
href="catalogue/category/books/poetry_23/index.html">\n
\n                Poetry\n                \n
</a>\n\n                </li>\n                \n
\n                <li>\n                <a
href="catalogue/category/books/paranormal_24/index.html">\n
\n                Paranormal\n                \n
</a>\n\n                </li>\n                \n
\n                <li>\n                <a
href="catalogue/category/books/art_25/index.html">\n
\n                Art\n                \n
</a>\n\n                </li>\n                \n
\n                <li>\n                <a
href="catalogue/category/books/psychology_26/index.html">\n
\n                Psychology\n                \n
</a>\n\n                </li>\n                \n
\n                <li>\n                <a
href="catalogue/category/books/autobiography_27/index.html">\n
\n                Autobiography\n                \n
</a>\n\n                </li>\n                \n
\n                <li>\n                <a
href="catalogue/category/books/parenting_28/index.html">\n
\n                Parenting\n                \n
</a>\n\n                </li>\n                \n
\n                <li>\n                <a
href="catalogue/category/books/adult-fiction_29/index.html">\n
\n                Adult Fiction\n                \n

```

```

</a>\n\n                                </li>\n                                \n
\n                                <li>\n                                <a
href="catalogue/category/books/humor_30/index.html">\n
\n                                Humor\n                                \n
</a>\n\n                                </li>\n                                \n
\n                                <li>\n                                <a
href="catalogue/category/books/horror_31/index.html">\n
\n                                Horror\n                                \n
</a>\n\n                                </li>\n                                \n
\n                                <li>\n                                <a
href="catalogue/category/books/history_32/index.html">\n
\n                                History\n                                \n
</a>\n\n                                </li>\n                                \n
\n                                <li>\n                                <a
href="catalogue/category/books/food-and-drink_33/index.html">\n
\n                                Food and Drink\n                                \n
</a>\n\n                                </li>\n                                \n
\n                                <li>\n                                <a
href="catalogue/category/books/christian-fiction_34/index.html">\n
\n                                Christian Fiction\n
\n                                </a>\n\n                                </li>\n
\n                                \n                                <li>\n                                <a
href="catalogue/category/books/business_35/index.html">\n
\n                                Business\n                                \n
</a>\n\n                                </li>\n                                \n
\n                                <li>\n                                <a
href="catalogue/category/books/biography_36/index.html">\n
\n                                Biography\n                                \n
</a>\n\n                                </li>\n                                \n
\n                                <li>\n                                <a
href="catalogue/category/books/thriller_37/index.html">\n
\n                                Thriller\n                                \n
</a>\n\n                                </li>\n                                \n
\n                                <li>\n                                <a
href="catalogue/category/books/contemporary_38/index.html">\n
\n                                Contemporary\n                                \n
</a>\n\n                                </li>\n                                \n
\n                                <li>\n                                <a
href="catalogue/category/books/spirituality_39/index.html">\n
\n                                Spirituality\n                                \n
</a>\n\n                                </li>\n                                \n
\n                                <li>\n                                <a
href="catalogue/category/books/academic_40/index.html">\n
\n                                Academic\n                                \n
</a>\n\n                                </li>\n                                \n
\n                                <li>\n                                <a
href="catalogue/category/books/self-help_41/index.html">\n

```

```

\n                                Self Help\n\n                                \n
</a>\n\n                                </li>\n\n                                \n
\n                                <li>\n\n                                <a
href="catalogue/category/books/historical_42/index.html">\n
\n                                Historical\n\n                                \n
</a>\n\n                                </li>\n\n                                \n
\n                                <li>\n\n                                <a
href="catalogue/category/books/christian_43/index.html">\n
\n                                Christian\n\n                                \n
</a>\n\n                                </li>\n\n                                \n
\n                                <li>\n\n                                <a
href="catalogue/category/books/suspense_44/index.html">\n
\n                                Suspense\n\n                                \n
</a>\n\n                                </li>\n\n                                \n
\n                                <li>\n\n                                <a
href="catalogue/category/books/short-stories_45/index.html">\n
\n                                Short Stories\n\n                                \n
</a>\n\n                                </li>\n\n                                \n
\n                                <li>\n\n                                <a
href="catalogue/category/books/novels_46/index.html">\n
\n                                Novels\n\n                                \n
</a>\n\n                                </li>\n\n                                \n
\n                                <li>\n\n                                <a
href="catalogue/category/books/health_47/index.html">\n
\n                                Health\n\n                                \n
</a>\n\n                                </li>\n\n                                \n
\n                                <li>\n\n                                <a
href="catalogue/category/books/politics_48/index.html">\n
\n                                Politics\n\n                                \n
</a>\n\n                                </li>\n\n                                \n
\n                                <li>\n\n                                <a
href="catalogue/category/books/cultural_49/index.html">\n
\n                                Cultural\n\n                                \n
</a>\n\n                                </li>\n\n                                \n
\n                                <li>\n\n                                <a
href="catalogue/category/books/erotica_50/index.html">\n
\n                                Erotica\n\n                                \n
</a>\n\n                                </li>\n\n                                \n
\n                                <li>\n\n                                <a
href="catalogue/category/books/crime_51/index.html">\n
\n                                Crime\n\n                                \n
</a>\n\n                                </li>\n\n                                \n
</ul></li>\n\n                                \n                                \n                                </ul>\n
</div>\n\n\n\n                                </aside>\n\n\n                                <div class="col-sm-8
col-md-9">\n\n                                \n                                <div class="page-header action">\n
<h1>All products</h1>\n\n                                </div>\n\n                                \n\n
\n\n\n\n<div id="messages">\n\n\n</div>\n\n\n\n                                <div

```



```

text="Adding...">Add to basket</button>\n      </form>\n\n\n      \n
</div>\n      \n      </article>\n\n</li>\n      \n
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">\n\n\n\n\n\n\n      <article
class="product_pod">\n      \n      <div class="image_container">\n
\n      \n      <a
href="catalogue/soumission_998/index.html"></a>\n      \n      \n
</div>\n      \n\n      \n      \n      <p class="star-
rating One">\n      <i class="icon-star"></i>\n
<i class="icon-star"></i>\n      <i class="icon-star"></i>\n
<i class="icon-star"></i>\n      <i class="icon-star"></i>\n
</p>\n      \n      \n\n      \n      <h3><a
href="catalogue/soumission_998/index.html"
title="Soumission">Soumission</a></h3>\n      \n\n      \n      <div
class="product_price">\n      \n\n\n\n\n\n\n      \n      <p
class="price_color">\xc2\xa350.10</p>\n      \n\n<p class="instock
availability">\n      <i class="icon-ok"></i>\n      \n      In stock\n
\n</p>\n\n      \n      \n\n\n\n\n\n\n      \n      <form>\n
<button type="submit" class="btn btn-primary btn-block" data-loading-
text="Adding...">Add to basket</button>\n      </form>\n\n\n\n      \n
</div>\n      \n      </article>\n\n</li>\n      \n
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">\n\n\n\n\n\n\n      <article
class="product_pod">\n      \n      <div class="image_container">\n
\n      \n      <a href="catalogue/sharp-
objects_997/index.html"></a>\n      \n      \n
</div>\n      \n\n      \n      \n      <p class="star-
rating Four">\n      <i class="icon-star"></i>\n
<i class="icon-star"></i>\n      <i class="icon-star"></i>\n
<i class="icon-star"></i>\n      <i class="icon-star"></i>\n
</p>\n      \n      \n      \n\n      \n      <h3><a
href="catalogue/sharp-objects_997/index.html" title="Sharp Objects">Sharp
Objects</a></h3>\n      \n\n      \n      <div
class="product_price">\n      \n\n\n\n\n\n\n      \n      <p
class="price_color">\xc2\xa347.82</p>\n      \n\n<p class="instock
availability">\n      <i class="icon-ok"></i>\n      \n      In stock\n
\n</p>\n\n      \n      \n\n\n\n\n\n\n      \n      <form>\n
<button type="submit" class="btn btn-primary btn-block" data-loading-
text="Adding...">Add to basket</button>\n      </form>\n\n\n\n      \n
</div>\n      \n      </article>\n\n</li>\n      \n
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">\n\n\n\n\n\n\n      <article
class="product_pod">\n      \n      <div class="image_container">\n
\n      \n      <a href="catalogue/sapiens-a-brief-
history-of-humankind_996/index.html"><img
src="media/cache/be/a5/bea5697f2534a2f86a3ef27b5a8c12a6.jpg" alt="Sapiens: A

```



```

Brief History of Humankind" class="thumbnail"></a>\n
\n
</div>\n
\n\n
\n
\n
<p
class="star-rating Five">\n
\n
<i class="icon-star"></i>\n
<i class="icon-star"></i>\n
<i class="icon-star"></i>\n
<i class="icon-star"></i>\n
</p>\n
\n
\n\n
\n
\n
<h3><a
href="catalogue/sapiens-a-brief-history-of-humankind_996/index.html"
title="Sapiens: A Brief History of Humankind">Sapiens: A Brief History
...</a></h3>\n
\n\n
\n
\n
<div class="product_price">\n
\n\n\n\n\n\n\n
\n
<p class="price_color">\xc2\xa354.23</p>\n
\n\n\n<p
class="instock availability">\n
<i class="icon-ok"></i>\n
\n
In
stock\n
\n</p>\n\n\n
\n
\n\n\n\n\n\n\n\n
\n
<form>\n
<button type="submit" class="btn btn-primary btn-block" data-
loading-text="Adding...">Add to basket</button>\n
</form>\n\n\n\n
\n
</div>\n
\n
</article>\n\n</li>\n
\n
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">\n\n\n\n\n\n\n\n
<article
class="product_pod">\n
\n
<div class="image_container">\n
\n
\n
<a href="catalogue/the-requiem-
red_995/index.html"></a>\n
\n
\n
</div>\n
\n\n
\n
\n
<p class="star-
rating One">\n
\n
<i class="icon-star"></i>\n
<i class="icon-star"></i>\n
<i class="icon-star"></i>\n
<i class="icon-star"></i>\n
</p>\n
\n
\n\n
\n
\n
<h3><a
href="catalogue/the-requiem-red_995/index.html" title="The Requiem Red">The
Requiem Red</a></h3>\n
\n\n
\n
\n
<div
class="product_price">\n
\n\n\n\n\n\n\n\n
\n
<p
class="price_color">\xc2\xa322.65</p>\n
\n\n\n<p class="instock
availability">\n
<i class="icon-ok"></i>\n
\n
In stock\n
\n</p>\n\n\n
\n
\n\n\n\n\n\n\n\n
\n
<form>\n
<button type="submit" class="btn btn-primary btn-block" data-loading-
text="Adding...">Add to basket</button>\n
</form>\n\n\n\n
\n
</div>\n
\n
</article>\n\n</li>\n
\n
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">\n\n\n\n\n\n\n\n
<article
class="product_pod">\n
\n
<div class="image_container">\n
\n
\n
<a href="catalogue/the-dirty-little-
secrets-of-getting-your-dream-job_994/index.html"></a>\n
\n
\n
</div>\n
\n\n
\n
\n
<p class="star-rating Four">\n
\n
<i class="icon-star"></i>\n
<i class="icon-star"></i>\n
<i class="icon-star"></i>\n
<i class="icon-star"></i>\n
</p>\n
\n
\n\n
\n
\n
<h3><a
href="catalogue/the-dirty-little-secrets-of-getting-your-dream-

```

```

job_994/index.html" title="The Dirty Little Secrets of Getting Your Dream
Job">The Dirty Little Secrets ...</a></h3>\n          \n\n          \n
<div class="product_price">\n          \n\n\n\n\n\n\n          \n          <p
class="price_color">\xc2\xa333.34</p>\n          \n\n<p class="instock
availability">\n          <i class="icon-ok"></i>\n          \n          In stock\n
\n</p>\n\n          \n          \n\n\n\n\n\n\n          \n          <form>\n
<button type="submit" class="btn btn-primary btn-block" data-loading-
text="Adding...">Add to basket</button>\n          </form>\n\n\n          \n
</div>\n          \n          </article>\n\n</li>\n          \n
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">\n\n\n\n\n\n\n          <article
class="product_pod">\n          \n          <div class="image_container">\n
\n          \n          <a href="catalogue/the-coming-woman-
a-novel-based-on-the-life-of-the-infamous-feminist-victoria-
woodhull_993/index.html"></a>\n          \n          \n
</div>\n          \n\n          \n          \n          <p class="star-
rating Three">\n          <i class="icon-star"></i>\n
<i class="icon-star"></i>\n          <i class="icon-star"></i>\n
<i class="icon-star"></i>\n          <i class="icon-star"></i>\n
</p>\n          \n          \n\n          \n          <h3><a
href="catalogue/the-coming-woman-a-novel-based-on-the-life-of-the-infamous-
feminist-victoria-woodhull_993/index.html" title="The Coming Woman: A Novel
Based on the Life of the Infamous Feminist, Victoria Woodhull">The Coming Woman:
A ...</a></h3>\n          \n\n          \n          <div class="product_price">\n
\n\n\n\n\n\n\n          \n          <p class="price_color">\xc2\xa317.93</p>\n          \n\n<p
class="instock availability">\n          <i class="icon-ok"></i>\n          \n          In
stock\n          \n</p>\n\n          \n          \n\n\n\n\n\n\n          \n
<form>\n          <button type="submit" class="btn btn-primary btn-block" data-
loading-text="Adding...">Add to basket</button>\n          </form>\n\n\n\n
\n          </div>\n          \n          </article>\n\n</li>\n          \n
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">\n\n\n\n\n\n\n          <article
class="product_pod">\n          \n          <div class="image_container">\n
\n          \n          <a href="catalogue/the-boys-in-the-
boat-nine-americans-and-their-epic-quest-for-gold-at-the-1936-berlin-
olympics_992/index.html"></a>\n          \n          \n
</div>\n          \n\n          \n          \n          <p class="star-
rating Four">\n          <i class="icon-star"></i>\n
<i class="icon-star"></i>\n          <i class="icon-star"></i>\n
<i class="icon-star"></i>\n          <i class="icon-star"></i>\n
</p>\n          \n          \n\n          \n          <h3><a
href="catalogue/the-boys-in-the-boat-nine-americans-and-their-epic-quest-for-
gold-at-the-1936-berlin-olympics_992/index.html" title="The Boys in the Boat:

```

Nine Americans and Their Epic Quest for Gold at the 1936 Berlin Olympics">The Boys in the ...</a></h3>\n \n \n \n <div class="product\_price">\n \n \n \n \n \n \n \n \n <p class="price\_color">\xc2\xa322.60</p>\n \n \n <p class="instock availability">\n <i class="icon-ok"></i>\n \n In stock\n \n</p>\n \n \n \n \n \n \n \n \n <form>\n <button type="submit" class="btn btn-primary btn-block" data-loading-text="Adding...">Add to basket</button>\n </form>\n \n \n \n </div>\n \n </article>\n \n</li>\n \n <li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">\n \n \n \n \n \n \n \n <article class="product\_pod">\n \n <div class="image\_container">\n \n <a href="catalogue/the-black-maria\_991/index.html"></a>\n \n </div>\n \n \n \n <p class="star-rating One">\n <i class="icon-star"></i>\n <i class="icon-star"></i>\n <i class="icon-star"></i>\n <i class="icon-star"></i>\n </p>\n \n \n \n <h3><a href="catalogue/the-black-maria\_991/index.html" title="The Black Maria">The Black Maria</a></h3>\n \n \n <div class="product\_price">\n \n \n \n \n \n \n \n \n <p class="price\_color">\xc2\xa352.15</p>\n \n \n <p class="instock availability">\n <i class="icon-ok"></i>\n \n In stock\n \n</p>\n \n \n \n \n \n \n \n \n <form>\n <button type="submit" class="btn btn-primary btn-block" data-loading-text="Adding...">Add to basket</button>\n </form>\n \n \n \n </div>\n \n </article>\n \n</li>\n \n <li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">\n \n \n \n \n \n \n \n <article class="product\_pod">\n \n <div class="image\_container">\n \n <a href="catalogue/starving-hearts-triangular-trade-trilogy-1\_990/index.html"></a>\n \n </div>\n \n \n \n <p class="star-rating Two">\n <i class="icon-star"></i>\n <i class="icon-star"></i>\n <i class="icon-star"></i>\n <i class="icon-star"></i>\n </p>\n \n \n \n <h3><a href="catalogue/starving-hearts-triangular-trade-trilogy-1\_990/index.html" title="Starving Hearts (Triangular Trade Trilogy, #1)">Starving Hearts (Triangular Trade ...</a></h3>\n \n \n <div class="product\_price">\n \n \n \n \n \n \n \n \n <p class="price\_color">\xc2\xa313.99</p>\n \n \n <p class="instock availability">\n <i class="icon-ok"></i>\n \n In stock\n \n</p>\n \n \n \n \n \n \n \n \n <form>\n

```

<button type="submit" class="btn btn-primary btn-block" data-loading-
text="Adding...">Add to basket</button>\n      </form>\n\n\n      \n
</div>\n      \n      </article>\n\n\n</li>\n      \n
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">\n\n\n\n\n\n\n      <article
class="product_pod">\n      \n      <div class="image_container">\n
\n      \n      <a href="catalogue/shakespeares-
sonnets_989/index.html"></a>\n      \n
\n      </div>\n      \n\n      \n      \n      <p
class="star-rating Four">\n      \n      <i class="icon-star"></i>\n
<i class="icon-star"></i>\n      <i class="icon-star"></i>\n
<i class="icon-star"></i>\n      <i class="icon-star"></i>\n
</p>\n      \n      \n\n      \n      <h3><a
href="catalogue/shakespeares-sonnets_989/index.html" title="Shakespeare&#39;s
Sonnets">Shakespeare&#39;s Sonnets</a></h3>\n      \n\n      \n
<div class="product_price">\n      \n\n\n\n\n\n\n      \n      <p
class="price_color">\xc2\xa320.66</p>\n      \n\n<p class="instock
availability">\n      <i class="icon-ok"></i>\n      \n      In stock\n
\n</p>\n\n      \n      \n\n\n\n\n\n\n      \n      <form>\n
<button type="submit" class="btn btn-primary btn-block" data-loading-
text="Adding...">Add to basket</button>\n      </form>\n\n\n\n      \n
</div>\n      \n      </article>\n\n\n</li>\n      \n
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">\n\n\n\n\n\n\n      <article
class="product_pod">\n      \n      <div class="image_container">\n
\n      \n      <a href="catalogue/set-me-
free_988/index.html"></a>\n      \n      \n
</div>\n      \n\n      \n      \n      <p class="star-
rating Five">\n      \n      <i class="icon-star"></i>\n
<i class="icon-star"></i>\n      <i class="icon-star"></i>\n
<i class="icon-star"></i>\n      <i class="icon-star"></i>\n
</p>\n      \n      \n\n      \n      <h3><a
href="catalogue/set-me-free_988/index.html" title="Set Me Free">Set Me
Free</a></h3>\n      \n\n      \n      <div class="product_price">\n
\n\n\n\n\n\n\n      \n      <p class="price_color">\xc2\xa317.46</p>\n      \n\n<p
class="instock availability">\n      <i class="icon-ok"></i>\n      \n      In
stock\n      \n</p>\n\n      \n      \n\n\n\n\n\n\n      \n
<form>\n      <button type="submit" class="btn btn-primary btn-block" data-
loading-text="Adding...">Add to basket</button>\n      </form>\n\n\n\n
\n      </div>\n      \n      </article>\n\n\n</li>\n      \n
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">\n\n\n\n\n\n\n      <article
class="product_pod">\n      \n      <div class="image_container">\n
\n      \n      <a href="catalogue/scott-pilgrims-
precious-little-life-scott-pilgrim-1_987/index.html"></i>\n                <i class="icon-star"></i>\n
</p>\n                \n                \n                \n                <h3><a
href="catalogue/our-band-could-be-your-life-scenes-from-the-american-indie-
underground-1981-1991_985/index.html" title="Our Band Could Be Your Life: Scenes
from the American Indie Underground, 1981-1991">Our Band Could Be ...</a></h3>\n
\n                \n                <div class="product_price">\n
\n\n\n\n\n\n\n                \n                <p class="price_color">\xc2\xa357.25</p>\n                \n\n\n<p
class="instock availability">\n                <i class="icon-ok"></i>\n                \n                In
stock\n                \n</p>\n\n                \n                \n\n\n\n\n\n\n                \n
<form>\n                <button type="submit" class="btn btn-primary btn-block" data-
loading-text="Adding...">Add to basket</button>\n                </form>\n\n\n\n
\n                </div>\n                \n                </article>\n\n\n</li>\n                \n
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">\n\n\n\n\n\n\n\n                <article
class="product_pod">\n                \n                <div class="image_container">\n
\n                \n                <a
href="catalogue/olio_984/index.html"></a>\n                \n                \n
</div>\n                \n\n                \n                \n                <p class="star-
rating One">\n                <i class="icon-star"></i>\n
<i class="icon-star"></i>\n                <i class="icon-star"></i>\n
<i class="icon-star"></i>\n                <i class="icon-star"></i>\n
</p>\n                \n                \n\n                \n                <h3><a
href="catalogue/olio_984/index.html" title="Olio">Olio</a></h3>\n                \n\n
\n                <div class="product_price">\n                \n\n\n\n\n\n\n\n                \n
<p class="price_color">\xc2\xa323.88</p>\n                \n\n\n<p class="instock
availability">\n                <i class="icon-ok"></i>\n                \n                In stock\n
\n</p>\n\n                \n                \n\n\n\n\n\n\n\n                \n                <form>\n
<button type="submit" class="btn btn-primary btn-block" data-loading-
text="Adding...">Add to basket</button>\n                </form>\n\n\n\n                \n
</div>\n                \n                </article>\n\n\n</li>\n                \n
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">\n\n\n\n\n\n\n\n                <article
class="product_pod">\n                \n                <div class="image_container">\n
\n                \n                <a href="catalogue/mesaerion-the-
best-science-fiction-stories-1800-1849_983/index.html"></a>\n
\n                \n                </div>\n                \n\n                \n                \n
<p class="star-rating One">\n                <i class="icon-star"></i>\n
<i class="icon-star"></i>\n                <i class="icon-star"></i>\n
<i class="icon-star"></i>\n                <i class="icon-star"></i>\n
</p>\n                \n                \n\n                \n                <h3><a
href="catalogue/mesaerion-the-best-science-fiction-
stories-1800-1849_983/index.html" title="Mesaerion: The Best Science Fiction
Stories 1800-1849">Mesaerion: The Best Science ...</a></h3>\n                \n\n
\n                <div class="product_price">\n                \n\n\n\n\n\n\n\n                \n
<p class="price_color">\xc2\xa337.59</p>\n                \n\n\n<p class="instock

```

availability">\n <i class="icon-ok"></i>\n \n In stock\n  
 \n</p>\n\n \n \n\n\n\n\n\n\n \n <form>\n  
 <button type="submit" class="btn btn-primary btn-block" data-loading-  
 text="Adding...">Add to basket</button>\n </form>\n\n\n \n  
 </div>\n \n </article>\n\n</li>\n \n  
 <li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">\n\n\n\n\n\n\n\n <article  
 class="product\_pod">\n \n <div class="image\_container">\n  
 \n \n <a href="catalogue/libertarianism-  
 for-beginners\_982/index.html"></a>\n \n \n  
 </div>\n \n\n \n \n <p class="star-  
 rating Two">\n <i class="icon-star"></i>\n  
 <i class="icon-star"></i>\n <i class="icon-star"></i>\n  
 <i class="icon-star"></i>\n <i class="icon-star"></i>\n  
 </p>\n \n \n\n \n <h3><a  
 href="catalogue/libertarianism-for-beginners\_982/index.html"  
 title="Libertarianism for Beginners">Libertarianism for Beginners</a></h3>\n  
 \n\n \n <div class="product\_price">\n  
 \n\n\n\n\n\n\n \n <p class="price\_color">\xc2\xa351.33</p>\n \n\n<p  
 class="instock availability">\n <i class="icon-ok"></i>\n \n In  
 stock\n \n</p>\n\n \n \n\n\n\n\n\n\n \n  
 <form>\n <button type="submit" class="btn btn-primary btn-block" data-  
 loading-text="Adding...">Add to basket</button>\n </form>\n\n\n\n  
 \n </div>\n \n </article>\n\n</li>\n \n  
 <li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">\n\n\n\n\n\n\n\n <article  
 class="product\_pod">\n \n <div class="image\_container">\n  
 \n \n <a href="catalogue/its-only-the-  
 himalayas\_981/index.html"></a>\n \n \n  
 </div>\n \n\n \n \n <p class="star-  
 rating Two">\n <i class="icon-star"></i>\n  
 <i class="icon-star"></i>\n <i class="icon-star"></i>\n  
 <i class="icon-star"></i>\n <i class="icon-star"></i>\n  
 </p>\n \n \n\n \n <h3><a  
 href="catalogue/its-only-the-himalayas\_981/index.html" title="It&#39;s Only the  
 Himalayas">It&#39;s Only the Himalayas</a></h3>\n \n\n \n  
 <div class="product\_price">\n \n\n\n\n\n\n\n\n \n <p  
 class="price\_color">\xc2\xa345.17</p>\n \n\n<p class="instock  
 availability">\n <i class="icon-ok"></i>\n \n In stock\n  
 \n</p>\n\n \n \n\n\n\n\n\n\n\n \n <form>\n  
 <button type="submit" class="btn btn-primary btn-block" data-loading-  
 text="Adding...">Add to basket</button>\n </form>\n\n\n\n \n  
 </div>\n \n </article>\n\n</li>\n \n  
 </ol>\n \n\n\n\n\n <div>\n <ul class="pager">\n  
 \n <li class="current">\n \n Page 1 of 50\n

```
</li>\n
href="catalogue/page-2.html">next</a></li>\n
</div>\n\n\n
</div><!-- /row -->\n
-->\n</div><!-- /container-fluid -->\n\n\n\n
fluid">\n
<!-- jQuery -->\n
query/1.9.1/jquery.min.js"></script>\n
document.write('<script
src="static/oscar/js/jquery/jquery-1.9.1.min.js"><\/script>')</script>\n
\n
\n
\n
\n
<script type="text/javascript"
src="static/oscar/js/bootstrap3/bootstrap.min.js"></script>\n
-->\n
<script src="static/oscar/js/oscar/ui.js" type="text/javascript"
charset="utf-8"></script>\n
<script src="static/oscar/js/bootstrap-
datetimepicker/bootstrap-datetimepicker.js" type="text/javascript"
charset="utf-8"></script>\n
<script src="static/oscar/js/bootstrap-
datetimepicker/locales/bootstrap-datetimepicker.all.js" type="text/javascript"
charset="utf-8"></script>\n
\n
\n
\n
<script type="text/javascript">\n
\n
\n
oscar.init();\n
oscar.search.init();\n
</script>\n
\n
<!-- Version: N/A -->\n
</body>\n</html>\n'
```

```
[6]: # Use BeautifulSoup to parse the HTML content of the website
```

```
soup = BeautifulSoup(content, 'html.parser')
```

```
[22]: soup
```

```
[22]: <!DOCTYPE html>
```

```
<!--[if lt IE 7]>          <html lang="en-us" class="no-js lt-ie9 lt-ie8 lt-ie7">
<![endif]-->
<!--[if IE 7]>             <html lang="en-us" class="no-js lt-ie9 lt-ie8">
<![endif]-->
<!--[if IE 8]>             <html lang="en-us" class="no-js lt-ie9"> <![endif]-->
<!--[if gt IE 8]><!--> <html class="no-js" lang="en-us"> <!--<![endif]-->
<head>
<title>
    All products | Books to Scrape - Sandbox
</title>
<meta content="text/html; charset=utf-8" http-equiv="content-type"/>
<meta content="24th Jun 2016 09:29" name="created"/>
<meta content="" name="description"/>
<meta content="width=device-width" name="viewport"/>
<meta content="NOARCHIVE,NOCACHE" name="robots"/>
```



```

<!-- Le HTML5 shim, for IE6-8 support of HTML elements -->
<!--[if lt IE 9]>
    <script src="//html5shim.googlecode.com/svn/trunk/html5.js"></script>
    <![endif]-->
<link href="static/oscar/favicon.ico" rel="shortcut icon"/>
<link href="static/oscar/css/styles.css" rel="stylesheet" type="text/css"/>
<link href="static/oscar/js/bootstrap-datetimepicker/bootstrap-
datetimepicker.css" rel="stylesheet"/>
<link href="static/oscar/css/datetimepicker.css" rel="stylesheet"
type="text/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> We love
being scraped!</small>
</div>
</div>
</div>
</div>
</header>
<div class="container-fluid page">
<div class="page_inner">
<ul class="breadcrumb">
<li>
<a href="index.html">Home</a>
</li>
<li class="active">All products</li>
</ul>
<div class="row">
<aside class="sidebar col-sm-4 col-md-3">
<div id="promotions_left">
</div>
<div class="side_categories">
<ul class="nav nav-list">
<li>
<a href="catalogue/category/books_1/index.html">

Books

</a>

</li>
</ul>
<li>
<a href="catalogue/category/books/travel_2/index.html">

Travel

```

```

        </a>
</li>
<li>
<a href="catalogue/category/books/mystery_3/index.html">

        Mystery

    </a>
</li>
<li>
<a href="catalogue/category/books/historical-fiction_4/index.html">

        Historical Fiction

    </a>
</li>
<li>
<a href="catalogue/category/books/sequential-art_5/index.html">

        Sequential Art

    </a>
</li>
<li>
<a href="catalogue/category/books/classics_6/index.html">

        Classics

    </a>
</li>
<li>
<a href="catalogue/category/books/philosophy_7/index.html">

        Philosophy

    </a>
</li>
<li>
<a href="catalogue/category/books/romance_8/index.html">

        Romance

    </a>
</li>
<li>
<a href="catalogue/category/books/womens-fiction_9/index.html">

```

Womens Fiction

</a>

</li>

<li>

<a href="catalogue/category/books/fiction\_10/index.html">

Fiction

</a>

</li>

<li>

<a href="catalogue/category/books/childrens\_11/index.html">

Childrens

</a>

</li>

<li>

<a href="catalogue/category/books/religion\_12/index.html">

Religion

</a>

</li>

<li>

<a href="catalogue/category/books/nonfiction\_13/index.html">

Nonfiction

</a>

</li>

<li>

<a href="catalogue/category/books/music\_14/index.html">

Music

</a>

</li>

<li>

<a href="catalogue/category/books/default\_15/index.html">

Default

</a>

</li>

<li>

[<a href="catalogue/category/books/science-fiction\\_16/index.html">](catalogue/category/books/science-fiction_16/index.html)

Science Fiction

[</a>](catalogue/category/books/science-fiction_16/index.html)

</li>

<li>

[<a href="catalogue/category/books/sports-and-games\\_17/index.html">](catalogue/category/books/sports-and-games_17/index.html)

Sports and Games

[</a>](catalogue/category/books/sports-and-games_17/index.html)

</li>

<li>

[<a href="catalogue/category/books/add-a-comment\\_18/index.html">](catalogue/category/books/add-a-comment_18/index.html)

Add a comment

[</a>](catalogue/category/books/add-a-comment_18/index.html)

</li>

<li>

[<a href="catalogue/category/books/fantasy\\_19/index.html">](catalogue/category/books/fantasy_19/index.html)

Fantasy

[</a>](catalogue/category/books/fantasy_19/index.html)

</li>

<li>

[<a href="catalogue/category/books/new-adult\\_20/index.html">](catalogue/category/books/new-adult_20/index.html)

New Adult

[</a>](catalogue/category/books/new-adult_20/index.html)

</li>

<li>

[<a href="catalogue/category/books/young-adult\\_21/index.html">](catalogue/category/books/young-adult_21/index.html)

Young Adult

[</a>](catalogue/category/books/young-adult_21/index.html)

</li>

<li>

[<a href="catalogue/category/books/science\\_22/index.html">](catalogue/category/books/science_22/index.html)

Science

[</a>](catalogue/category/books/science_22/index.html)

```
</li>
<li>
<a href="catalogue/category/books/poetry_23/index.html">

        Poetry

    </a>
</li>
<li>
<a href="catalogue/category/books/paranormal_24/index.html">

        Paranormal

    </a>
</li>
<li>
<a href="catalogue/category/books/art_25/index.html">

        Art

    </a>
</li>
<li>
<a href="catalogue/category/books/psychology_26/index.html">

        Psychology

    </a>
</li>
<li>
<a href="catalogue/category/books/autobiography_27/index.html">

        Autobiography

    </a>
</li>
<li>
<a href="catalogue/category/books/parenting_28/index.html">

        Parenting

    </a>
</li>
<li>
<a href="catalogue/category/books/adult-fiction_29/index.html">

        Adult Fiction
```

```

        </a>
</li>
<li>
<a href="catalogue/category/books/humor_30/index.html">

        Humor

    </a>
</li>
<li>
<a href="catalogue/category/books/horror_31/index.html">

        Horror

    </a>
</li>
<li>
<a href="catalogue/category/books/history_32/index.html">

        History

    </a>
</li>
<li>
<a href="catalogue/category/books/food-and-drink_33/index.html">

        Food and Drink

    </a>
</li>
<li>
<a href="catalogue/category/books/christian-fiction_34/index.html">

        Christian Fiction

    </a>
</li>
<li>
<a href="catalogue/category/books/business_35/index.html">

        Business

    </a>
</li>
<li>
<a href="catalogue/category/books/biography_36/index.html">

```

Biography

</a>

</li>

<li>

<a href="catalogue/category/books/thriller\_37/index.html">

Thriller

</a>

</li>

<li>

<a href="catalogue/category/books/contemporary\_38/index.html">

Contemporary

</a>

</li>

<li>

<a href="catalogue/category/books/spirituality\_39/index.html">

Spirituality

</a>

</li>

<li>

<a href="catalogue/category/books/academic\_40/index.html">

Academic

</a>

</li>

<li>

<a href="catalogue/category/books/self-help\_41/index.html">

Self Help

</a>

</li>

<li>

<a href="catalogue/category/books/historical\_42/index.html">

Historical

</a>

</li>

```
<li>
<a href="catalogue/category/books/christian_43/index.html">

    Christian

</a>
</li>
<li>
<a href="catalogue/category/books/suspense_44/index.html">

    Suspense

</a>
</li>
<li>
<a href="catalogue/category/books/short-stories_45/index.html">

    Short Stories

</a>
</li>
<li>
<a href="catalogue/category/books/novels_46/index.html">

    Novels

</a>
</li>
<li>
<a href="catalogue/category/books/health_47/index.html">

    Health

</a>
</li>
<li>
<a href="catalogue/category/books/politics_48/index.html">

    Politics

</a>
</li>
<li>
<a href="catalogue/category/books/cultural_49/index.html">

    Cultural
```



```

                </a>
</li>
<li>
<a href="catalogue/category/books/erotica_50/index.html">

                Erotica

                </a>
</li>
<li>
<a href="catalogue/category/books/crime_51/index.html">

                Crime

                </a>
</li>
</ul></li>
</ul>
</div>
</aside>
<div class="col-sm-8 col-md-9">
<div class="page-header action">
<h1>All products</h1>
</div>
<div id="messages">
</div>
<div id="promotions">
</div>
<form class="form-horizontal" method="get">
<div style="display:none">
</div>
<strong>1000</strong> results - showing <strong>1</strong> to
<strong>20</strong>.

        </form>
<section>
<div class="alert alert-warning" role="alert"><strong>Warning!</strong> This is
a demo website for web scraping purposes. Prices and ratings here were randomly
assigned and have no real meaning.</div>
<div>
<ol class="row">
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">
<article class="product_pod">
<div class="image_container">

```

```

<a href="catalogue/a-light-in-the-attic_1000/index.html"></a>
</div>
<p class="star-rating Three">
<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>
<i class="icon-star"></i>
</p>
<h3><a href="catalogue/a-light-in-the-attic_1000/index.html" title="A Light in
the Attic">A Light in the ...</a></h3>
<div class="product_price">
<p class="price_color">£51.77</p>
<p class="instock availability">
<i class="icon-ok"></i>

```

In stock

```

</p>
<form>
<button class="btn btn-primary btn-block" data-loading-text="Adding..."
type="submit">Add to basket</button>
</form>
</div>
</article>
</li>
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">
<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>
<i class="icon-star"></i>
</p>
<h3><a href="catalogue/tipping-the-velvet_999/index.html" title="Tipping the
Velvet">Tipping the Velvet</a></h3>
<div class="product_price">
<p class="price_color">£53.74</p>
<p class="instock availability">

```

<i class="icon-ok"></i>

In stock

</p>

<form>

<button class="btn btn-primary btn-block" data-loading-text="Adding..." type="submit">Add to basket</button>

</form>

</div>

</article>

</li>

<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">

<article class="product\_pod">

<div class="image\_container">

<a href="catalogue/soumission\_998/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>

<h3><a href="catalogue/soumission\_998/index.html"

title="Soumission">Soumission</a></h3>

<div class="product\_price">

<p class="price\_color">£50.10</p>

<p class="instock availability">

<i class="icon-ok"></i>

In stock

</p>

<form>

<button class="btn btn-primary btn-block" data-loading-text="Adding..." type="submit">Add to basket</button>

</form>

</div>

</article>

</li>

<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">

<article class="product\_pod">

<div class="image\_container">

<a href="catalogue/sharp-objects\_997/index.html"></a>
</div>
<p class="star-rating Four">
<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>
<i class="icon-star"></i>
</p>
<h3><a href="catalogue/sharp-objects_997/index.html" title="Sharp Objects">Sharp
Objects</a></h3>
<div class="product_price">
<p class="price_color">£47.82</p>
<p class="instock availability">
<i class="icon-ok"></i>

        In stock

</p>
<form>
<button class="btn btn-primary btn-block" data-loading-text="Adding..."
type="submit">Add to basket</button>
</form>
</div>
</article>
</li>
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">
<article class="product_pod">
<div class="image_container">
<a href="catalogue/sapiens-a-brief-history-of-humankind_996/index.html"></a>
</div>
<p class="star-rating Five">
<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>
<i class="icon-star"></i>
</p>
<h3><a href="catalogue/sapiens-a-brief-history-of-humankind_996/index.html"
title="Sapiens: A Brief History of Humankind">Sapiens: A Brief History
...</a></h3>
<div class="product_price">
<p class="price_color">£54.23</p>
<p class="instock availability">

```

<i class="icon-ok"></i>

In stock

</p>

<form>

<button class="btn btn-primary btn-block" data-loading-text="Adding..." type="submit">Add to basket</button>

</form>

</div>

</article>

</li>

<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">

<article class="product\_pod">

<div class="image\_container">

<a href="catalogue/the-requiem-red\_995/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>

<h3><a href="catalogue/the-requiem-red\_995/index.html" title="The Requiem Red">The Requiem Red</a></h3>

<div class="product\_price">

<p class="price\_color">£22.65</p>

<p class="instock availability">

<i class="icon-ok"></i>

In stock

</p>

<form>

<button class="btn btn-primary btn-block" data-loading-text="Adding..." type="submit">Add to basket</button>

</form>

</div>

</article>

</li>

<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">

<article class="product\_pod">

<div class="image\_container">

<a href="catalogue/the-dirty-little-secrets-of-getting-your-dream-

job\_994/index.html"></a>

</div>

<p class="star-rating Four">

<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>

<h3><a href="catalogue/the-dirty-little-secrets-of-getting-your-dream-job\_994/index.html" title="The Dirty Little Secrets of Getting Your Dream Job">The Dirty Little Secrets ...</a></h3>

<div class="product\_price">

<p class="price\_color">£33.34</p>

<p class="instock availability">

<i class="icon-ok"></i>

In stock

</p>

<form>

<button class="btn btn-primary btn-block" data-loading-text="Adding..." type="submit">Add to basket</button>

</form>

</div>

</article>

</li>

<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">

<article class="product\_pod">

<div class="image\_container">

<a href="catalogue/the-coming-woman-a-novel-based-on-the-life-of-the-infamous-feminist-victoria-woodhull\_993/index.html"></a>

</div>

<p class="star-rating Three">

<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>

<h3><a href="catalogue/the-coming-woman-a-novel-based-on-the-life-of-the-infamous-feminist-victoria-woodhull\_993/index.html" title="The Coming Woman: A Novel Based on the Life of the Infamous Feminist, Victoria Woodhull">The Coming

Woman: A ...</a></h3>

<div class="product\_price">  
<p class="price\_color">£17.93</p>  
<p class="instock availability">  
<i class="icon-ok"></i>

In stock

</p>  
<form>  
<button class="btn btn-primary btn-block" data-loading-text="Adding..."  
type="submit">Add to basket</button>  
</form>  
</div>  
</article>  
</li>  
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">  
<article class="product\_pod">  
<div class="image\_container">  
<a href="catalogue/the-boys-in-the-boat-nine-americans-and-their-epic-quest-for-  
gold-at-the-1936-berlin-olympics\_992/index.html"></a>  
</div>  
<p class="star-rating Four">  
<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>  
<i class="icon-star"></i>  
</p>  
<h3><a href="catalogue/the-boys-in-the-boat-nine-americans-and-their-epic-quest-  
for-gold-at-the-1936-berlin-olympics\_992/index.html" title="The Boys in the  
Boat: Nine Americans and Their Epic Quest for Gold at the 1936 Berlin  
Olympics">The Boys in the ...</a></h3>  
<div class="product\_price">  
<p class="price\_color">£22.60</p>  
<p class="instock availability">  
<i class="icon-ok"></i>

In stock

</p>  
<form>  
<button class="btn btn-primary btn-block" data-loading-text="Adding..."  
type="submit">Add to basket</button>

```

</form>
</div>
</article>
</li>
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">
<article class="product_pod">
<div class="image_container">
<a href="catalogue/the-black-maria_991/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>
<i class="icon-star"></i>
</p>
<h3><a href="catalogue/the-black-maria_991/index.html" title="The Black
Maria">The Black Maria</a></h3>
<div class="product_price">
<p class="price_color">£52.15</p>
<p class="instock availability">
<i class="icon-ok"></i>

```

In stock

```

</p>
<form>
<button class="btn btn-primary btn-block" data-loading-text="Adding..."
type="submit">Add to basket</button>
</form>
</div>
</article>
</li>
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">
<article class="product_pod">
<div class="image_container">
<a href="catalogue/starving-hearts-triangular-trade-
trilogy-1_990/index.html"></a>
</div>
<p class="star-rating Two">
<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>
<h3><a href="catalogue/starving-hearts-triangular-trade-
trilogy-1_990/index.html" title="Starving Hearts (Triangular Trade Trilogy,
#1)">Starving Hearts (Triangular Trade ...</a></h3>
<div class="product_price">
<p class="price_color">£13.99</p>
<p class="instock availability">
<i class="icon-ok"></i>

```

In stock

```

</p>
<form>
<button class="btn btn-primary btn-block" data-loading-text="Adding..."
type="submit">Add to basket</button>
</form>
</div>
</article>
</li>
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">
<article class="product_pod">
<div class="image_container">
<a href="catalogue/shakespeares-sonnets_989/index.html"></a>
</div>
<p class="star-rating Four">
<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>
<h3><a href="catalogue/shakespeares-sonnets_989/index.html" title="Shakespeare's
Sonnets">Shakespeare's Sonnets</a></h3>
<div class="product_price">
<p class="price_color">£20.66</p>
<p class="instock availability">
<i class="icon-ok"></i>

```

In stock

```

</p>
<form>
<button class="btn btn-primary btn-block" data-loading-text="Adding..."

```

```

type="submit">Add to basket</button>
</form>
</div>
</article>
</li>
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">
<article class="product_pod">
<div class="image_container">
<a href="catalogue/set-me-free_988/index.html"></a>
</div>
<p class="star-rating Five">
<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>
<i class="icon-star"></i>
</p>
<h3><a href="catalogue/set-me-free_988/index.html" title="Set Me Free">Set Me
Free</a></h3>
<div class="product_price">
<p class="price_color">£17.46</p>
<p class="instock availability">
<i class="icon-ok"></i>

```

In stock

```

</p>
<form>
<button class="btn btn-primary btn-block" data-loading-text="Adding..."
type="submit">Add to basket</button>
</form>
</div>
</article>
</li>
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">
<article class="product_pod">
<div class="image_container">
<a href="catalogue/scott-pilgrims-precious-little-life-scott-
pilgrim-1_987/index.html"></a>
</div>
<p class="star-rating Five">
<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>
<h3><a href="catalogue/scott-pilgrims-precious-little-life-scott-
pilgrim-1_987/index.html" title="Scott Pilgrim's Precious Little Life (Scott
Pilgrim #1)">Scott Pilgrim's Precious Little ...</a></h3>
<div class="product_price">
<p class="price_color">£52.29</p>
<p class="instock availability">
<i class="icon-ok"></i>

```

In stock

```

</p>
<form>
<button class="btn btn-primary btn-block" data-loading-text="Adding..."
type="submit">Add to basket</button>
</form>
</div>
</article>
</li>
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">
<article class="product_pod">
<div class="image_container">
<a href="catalogue/rip-it-up-and-start-again_986/index.html"></a>
</div>
<p class="star-rating Five">
<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>
<i class="icon-star"></i>
</p>
<h3><a href="catalogue/rip-it-up-and-start-again_986/index.html" title="Rip it
Up and Start Again">Rip it Up and ...</a></h3>
<div class="product_price">
<p class="price_color">£35.02</p>
<p class="instock availability">
<i class="icon-ok"></i>

```

In stock

```

</p>
<form>

```

```

<button class="btn btn-primary btn-block" data-loading-text="Adding..."
type="submit">Add to basket</button>
</form>
</div>
</article>
</li>
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">
<article class="product_pod">
<div class="image_container">
<a href="catalogue/our-band-could-be-your-life-scenes-from-the-american-indie-
underground-1981-1991_985/index.html"></a>
</div>
<p class="star-rating Three">
<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>
<h3><a href="catalogue/our-band-could-be-your-life-scenes-from-the-american-
indie-underground-1981-1991_985/index.html" title="Our Band Could Be Your Life:
Scenes from the American Indie Underground, 1981-1991">Our Band Could Be
...</a></h3>
<div class="product_price">
<p class="price_color">£57.25</p>
<p class="instock availability">
<i class="icon-ok"></i>

```

In stock

```

</p>
<form>
<button class="btn btn-primary btn-block" data-loading-text="Adding..."
type="submit">Add to basket</button>
</form>
</div>
</article>
</li>
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">
<article class="product_pod">
<div class="image_container">
<a href="catalogue/olio_984/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>
<h3><a href="catalogue/olio_984/index.html" title="Olio">Olio</a></h3>
<div class="product_price">
<p class="price_color">£23.88</p>
<p class="instock availability">
<i class="icon-ok"></i>

```

In stock

```

</p>
<form>
<button class="btn btn-primary btn-block" data-loading-text="Adding..."
type="submit">Add to basket</button>
</form>
</div>
</article>
</li>
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">
<article class="product_pod">
<div class="image_container">
<a href="catalogue/mesaerion-the-best-science-fiction-
stories-1800-1849_983/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>
<h3><a href="catalogue/mesaerion-the-best-science-fiction-
stories-1800-1849_983/index.html" title="Mesaerion: The Best Science Fiction
Stories 1800-1849">Mesaerion: The Best Science ...</a></h3>
<div class="product_price">
<p class="price_color">£37.59</p>
<p class="instock availability">
<i class="icon-ok"></i>

```

In stock

```

</p>
<form>
<button class="btn btn-primary btn-block" data-loading-text="Adding..."
type="submit">Add to basket</button>
</form>
</div>
</article>
</li>
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">
<article class="product_pod">
<div class="image_container">
<a href="catalogue/libertarianism-for-beginners_982/index.html"></a>
</div>
<p class="star-rating Two">
<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>
<h3><a href="catalogue/libertarianism-for-beginners_982/index.html"
title="Libertarianism for Beginners">Libertarianism for Beginners</a></h3>
<div class="product_price">
<p class="price_color">£51.33</p>
<p class="instock availability">
<i class="icon-ok"></i>

```

In stock

```

</p>
<form>
<button class="btn btn-primary btn-block" data-loading-text="Adding..."
type="submit">Add to basket</button>
</form>
</div>
</article>
</li>
<li class="col-xs-6 col-sm-4 col-md-3 col-lg-3">
<article class="product_pod">
<div class="image_container">
<a href="catalogue/its-only-the-himalayas_981/index.html"></a>
</div>
<p class="star-rating Two">

```

```

<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>
<h3><a href="catalogue/its-only-the-himalayas_981/index.html" title="It's Only
the Himalayas">It's Only the Himalayas</a></h3>
<div class="product_price">
<p class="price_color">£45.17</p>
<p class="instock availability">
<i class="icon-ok"></i>

```

In stock

```

</p>
<form>
<button class="btn btn-primary btn-block" data-loading-text="Adding..."
type="submit">Add to basket</button>
</form>
</div>
</article>
</li>
</ol>
<div>
<ul class="pager">
<li class="current">

```

Page 1 of 50

```

</li>
<li class="next"><a href="catalogue/page-2.html">next</a></li>
</ul>
</div>
</div>
</div>
</section>
</div>
</div><!-- /row -->
</div><!-- /page_inner -->
</div><!-- /container-fluid -->
<footer class="footer container-fluid">
</footer>
<!-- jQuery -->
<script
src="http://ajax.googleapis.com/ajax/libs/jquery/1.9.1/jquery.min.js"></script>
<script>window.jQuery || document.write('<script
src="static/oscar/js/jquery/jquery-1.9.1.min.js"></script>')</script>

```

```

<!-- Twitter Bootstrap -->
<script src="static/oscar/js/bootstrap3/bootstrap.min.js"
type="text/javascript"></script>
<!-- Oscar -->
<script charset="utf-8" src="static/oscar/js/oscar/ui.js"
type="text/javascript"></script>
<script charset="utf-8" src="static/oscar/js/bootstrap-datetimepicker/bootstrap-
datetimepicker.js" type="text/javascript"></script>
<script charset="utf-8" src="static/oscar/js/bootstrap-
datetimepicker/locales/bootstrap-datetimepicker.all.js"
type="text/javascript"></script>
<script type="text/javascript">
    $(function() {

        oscar.init();

        oscar.search.init();

    });
</script>
<!-- Version: N/A -->
</body>
</html>

```

These lines of code retrieve the HTML content from the specified URL, parse it using BeautifulSoup, and store it as a variable called soup. This will allow us to extract specific pieces of data from the webpage.

```

[7]: # Find the first article element with the class 'product_pod' using
↳ BeautifulSoup

```

```

product_pod = soup.find('article', {'class' : 'product_pod'})

```

```

[24]: product_pod

```

```

[24]: <article class="product_pod">
<div class="image_container">
<a href="catalogue/a-light-in-the-attic_1000/index.html"></a>
</div>
<p class="star-rating Three">
<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>
<h3><a href="catalogue/a-light-in-the-attic_1000/index.html" title="A Light in
the Attic">A Light in the ...</a></h3>
<div class="product_price">
<p class="price_color">£51.77</p>
<p class="instock availability">
<i class="icon-ok"></i>

```

In stock

```

</p>
<form>
<button class="btn btn-primary btn-block" data-loading-text="Adding..."
type="submit">Add to basket</button>
</form>
</div>
</article>

```

```
[8]: # Extract the title of the book from the product_pod element
```

```
title = product_pod.h3.a.get('title')
```

```
[25]: title
```

```
[25]: 'A Light in the Attic'
```

```
[9]: # Find the price of the book from the first 'p' element with class_
     ↪ 'price_color' using BeautifulSoup
```

```
price = soup.find('p', {'class' : 'price_color'}).text.strip('&')
price = float(price)
```

```
[26]: price
```

```
[26]: 51.77
```

```
[10]: # Find the availability of the book from the first 'p' element with class_
     ↪ 'instock availability' using BeautifulSoup
```

```
stock = soup.find('p', {'class' : 'instock availability'}).text.strip()
```

```
[27]: stock
```

```
[27]: 'In stock'
```

These lines of code extract specific pieces of information from the parsed HTML content using BeautifulSoup. The product\_pod variable contains information about a single product on the

webpage, and the remaining variables extract the book title, price, and availability from that product information.

```
[11]: # Find all 'article' elements with class 'product_pod' on the page using BeautifulSoup
product_pod_list = soup.find_all('article', {'class' : 'product_pod'})

[12]: # Extract a list of book titles from the 'title' attribute of the 'a' tag
      # inside each 'product_pod' element

title_list = [tag.h3.a.get('title') for tag in product_pod_list]

# Extract a list of book prices (as floats) from the 'price_color' class inside
# each 'product_pod' element

price_list = [float(tag.find('p', {'class' : 'price_color'}).text.strip('&'))
              for tag in product_pod_list]

# Extract a list of book availabilities from the 'instock availability' class
# inside each 'product_pod' element

availability_list = [tag.find('p', {'class' : 'instock availability'}).text.
                     strip() for tag in product_pod_list]

[28]: title_list

[28]: ['A Light in the Attic',
      'Tipping the Velvet',
      'Soumission',
      'Sharp Objects',
      'Sapiens: A Brief History of Humankind',
      'The Requiem Red',
      'The Dirty Little Secrets of Getting Your Dream Job',
      'The Coming Woman: A Novel Based on the Life of the Infamous Feminist, Victoria
Woodhull',
      'The Boys in the Boat: Nine Americans and Their Epic Quest for Gold at the 1936
Berlin Olympics',
      'The Black Maria',
      'Starving Hearts (Triangular Trade Trilogy, #1)',
      'Shakespeare's Sonnets',
      'Set Me Free',
      'Scott Pilgrim's Precious Little Life (Scott Pilgrim #1)',
      'Rip it Up and Start Again',
      'Our Band Could Be Your Life: Scenes from the American Indie Underground,
1981-1991',
      'Olio',
```

'Mesaerion: The Best Science Fiction Stories 1800-1849',  
'Libertarianism for Beginners',  
"It's Only the Himalayas"]

```
[29]: price_list
```

[29]: [51.77, 53.74, 50.1, 47.82, 54.23, 22.65, 33.34, 17.93, 22.6, 52.15, 13.99, 20.66, 17.46, 52.29, 35.02, 57.25, 23.88, 37.59, 51.33, 45.17]

```
[30]: availability_list
```

[illegible]

```
'In stock',
'In stock']
```

These lines of code extract information about all of the products on the webpage by finding all instances of the `product_pod` class using BeautifulSoup. The resulting lists of book titles, prices, and availabilities can be used to create a DataFrame and perform further analysis.

```
[16]: # Create a dictionary with the book titles, prices, and availabilities
      ↪ extracted earlier
```

```
dict_books = {
    'Title of Book' : title_list,
    'Price' : price_list,
    'Availability' : availability_list
}

# Convert the dictionary to a pandas DataFrame

df_books = pd.DataFrame(dict_books)
```

```
[17]: df_books
```

```
[17]:
```

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

These lines of code create a dictionary from the extracted book data, with keys corresponding to the column names in the resulting DataFrame. The dictionary is then converted to a pandas DataFrame for further analysis and manipulation.

[ ]: