

Task 1: Web Scraping from an API

The task is to extract information of a number of books by category, from the following site:
<https://books.toscrape.com/>

There are some 50 categories of books on this site. You need to extract the following information for the first 30 (or less if so is the case) for any 10 categories of your choice:

1. Book Name
2. Rating
3. Price
4. Availability
5. Image

Of course, it's a mammoth task to do this manually! So you are required to create a script in Python to automate the task for you. As a starting point, you may (and probably will) refer to the following libraries and their usage to help you achieve the task.

BeautifulSoup, requests, json, csv, regex

How to submit the task?

You are required to submit the extracted information (apart from the images) for each of the books in the following form:

1. An .xlsx file or .csv file for each category of books with columns for name, rating, price and availability
2. A combined .xlsx or .csv file to store all books in the single file. There should be one extra column now for each book - that's the category.

For the images, download and store each book's image in a folder having the name of the category of the book. All these folders should be stored inside one combined folder named as "book_images".

Prescribed Time for submission : 3-4 Days