

Individual Assignment 1

You are asked to collect data from a test website: books.toscrape.com. Please analyze the HTML tags carefully.

In the dataset, we want to collect the information for all 1000 books. For each book, the following information is required with the example value shown in the (parenthesis):

1. Title ("A light in the Attic")
2. Price (51.77)
3. Star ("Three")
4. URL (http://books.toscrape.com/catalogue/a-light-in-the-attic_1000/index.html)

Follow each book's link and go to its own page, collect:

5. Availability ("In stock (22 available)")
6. Description ("It's hard to imagine a world without A Light in the Attic. ...")

Please submit two files:

1. A word file includes python code with comments # or your spider .py file, and a screenshot on your terminal showing that your code has run through successfully.
2. A CSV file saves your data, each book should be a record, and each type of information is a field.