

## webscrapcd.py

```
1 import requests
2 from bs4 import BeautifulSoup
3 import csv
4
5 url = "http://books.toscrape.com/"
6 headers = {
7     "User-Agent": "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like
8     Gecko) Chrome/113.0.0.0 Safari/537.36"
9 }
10 response = requests.get(url, headers=headers)
11
12 if response.status_code == 200:
13     soup = BeautifulSoup(response.content, "html.parser")
14     products = soup.find_all("article", class_="product_pod")
15
16     with open("books_data.csv", mode="w", newline="", encoding="utf-8") as file:
17         writer = csv.writer(file)
18         writer.writerow(["Title", "Price", "Image URL"])
19
20         for product in products:
21             title_tag = product.find("h3").find("a")
22             title = title_tag["title"] if title_tag else "No title"
23
24             price_tag = product.find("p", class_="price_color")
25             price = price_tag.text.strip() if price_tag else "No price"
26
27             image_tag = product.find("img")
28             image_url = url + image_tag["src"] if image_tag else "No image"
29
30             writer.writerow([title, price, image_url])
31
32     print("Data has been written to books_data.csv.")
33 else:
34     print(f"Failed to retrieve page: {response.status_code}")
```