

Desafio Tecnico - Web Scraping com BeautifulSoup

Desafio Oficial - Scraping com BeautifulSoup

Objetivo:

Criar um script Python que raspe todos os livros com avaliacao de 5 estrelas do site:

<https://books.toscrape.com>

Tecnologias que voce deve utilizar:

- requests - para fazer as requisicoes HTTP
- BeautifulSoup - para fazer o parsing e extracao dos dados do HTML
- pandas - para salvar os dados em .csv
- Nao usar Selenium (o site e estatico, nao precisa de navegador)

Requisitos funcionais:

Para cada livro com 5 estrelas, colete:

- Titulo
- Preco
- Disponibilidade
- Link completo da pagina do livro

Detalhes tecnicos:

- Os livros aparecem em varias paginas (50 no total). Use for ou while para navegar.
- Os livros com 5 estrelas estao marcados com: `<p class="star-rating Five"></p>`
- O link do livro e relativo (ex: "catalogue/a-book-title_123/index.html")
Concatene com <https://books.toscrape.com/> para montar o link completo.

Saida esperada:

Salve os dados em um arquivo chamado:

livros_5_estrelas.csv

Com as colunas:

Titulo, Preco, Disponibilidade, Link

Exemplo de uma linha esperada no CSV:

A Light in the ..., 51.77, In stock,
https://books.toscrape.com/catalogue/a-light-in-the-attic_1000/index.html

Restricoes:

- Nao use Selenium (esse site e 100% HTML estatico).
- Nao use XPath apenas BeautifulSoup com find, find_all ou select.

Dica de estrategia:

1. Use um for com range de 1 ate 50 (sao 50 paginas).
2. Use requests.get() com <https://books.toscrape.com/catalogue/page-1.html>, [page-2.html](https://books.toscrape.com/catalogue/page-2.html), etc.
3. Dentro de cada pagina, filtre os <article class="product_pod"> com estrela 5.
4. Pegue o titulo com a['title'], o preco, a disponibilidade e o href da tag a.