

Atividade: Web Scraping com Scrapy

Nesta atividade, você irá praticar a extração de dados de um site usando o framework Scrapy e depois realizará uma análise simples com Pandas em Python.

Objetivos

1. Criar um spider em Scrapy para coletar dados de livros.
2. Exportar os dados coletados em CSV ou JSON.
3. Analisar os dados coletados utilizando Pandas.

Parte 1 – Extraindo os dados

Usaremos o site de testes: <https://books.toscrape.com/>

- 1. Criar um projeto Scrapy:
- 2. Código do spider (spiders/catalogo.py):
- 3. Executar o spider e salvar os dados:

Parte 2 – Analisando os dados

Após extrair os dados, abra um Jupyter Notebook ou Google Colab e carregue o CSV gerado.

Questões para responder

1. Qual é o preço médio dos livros coletados?
2. Qual é o livro mais caro da base?
3. Quantos livros estão em estoque e quantos estão fora de estoque?
4. Faça um gráfico de barras ou histograma mostrando a distribuição dos preços.

Entregáveis

- Arquivo livros.csv gerado pelo Scrapy.
- Notebook ou Colab com a análise dos dados.
- Respostas das perguntas.