

Puppeteer + Adidas Web Scraper Explicação Técnica Detalhada

puppeteer: Biblioteca em JavaScript usada para automatizar interações com o navegador. Permite abrir páginas, clicar, preencher formulários e capturar dados.

fs: Módulo nativo do Node.js usado para ler e escrever arquivos no sistema (ex: salvar os dados coletados em JSON).

`{ headless: false }`: Parâmetro que indica se o navegador será exibido.

- `true`: roda em segundo plano (sem abrir o navegador visivelmente)
- `false`: mostra o navegador em tempo real (útil para testes)

`{ waitUntil: 'domcontentloaded' }`: Define que o código só continua após o carregamento do DOM da página.

`waitForSelector('selector')`: Aguarda até que o elemento identificado por um seletor (classe, id, atributo etc.) esteja disponível no DOM.

- Exemplo: ``#glass-gdpr-default-consent-accept-button`` botão de aceitar cookies

`new Promise(resolve => setTimeout(resolve, 10000))`: Usado para pausar a execução por 10 segundos. Isso ajuda a garantir que o conteúdo carregue antes de interagir com ele.

`$$eval()`: Executa uma função para extrair dados de vários elementos no DOM (NodeList), por exemplo, todos os cards de produto.

- Exemplo: ``.product-card_product-card-content___bjeq`` extrai dados de cada card

`querySelector()`: Seleciona um único elemento com base no seletor CSS.

Seletores CSS usados no projeto:

- .product-card_product-card-content__bjeq: Card de produto individual
- .product-card-description_name__xHvJ2: Nome do produto
- ._priceComponent_1dbqy_14: Preço do produto
- .product-card-description_info__z_CcT: Categoria ou tipo
- img: Imagem principal do produto
- #glass-gdpr-default-consent-accept-button: Botao de aceitar cookies (GDPR)
- button[data-auto-id="gl-modal-close"]: Botao de fechar modal AdiClub
- a[data-testid="pagination-next-button"]: Botao para avançar para a proxima pagina

page.url(): Retorna a URL atual da pagina. Usado para verificar se mudou apos clicar em "proximo".

page.waitForFunction(() => window.location.href !== oldUrl): Aguarda a URL mudar util apos clicar no botao "proximo".

fs.writeFileSync('arquivo.json', JSON.stringify(dados)): Salva os dados coletados no formato JSON no sistema de arquivos.

while (true) { ... }: Laco que permite continuar clicando no botao "proximo" e coletando dados de todas as paginas, ate que nao exista mais botao de continuar.

page.evaluate(): Executa uma funcao dentro do navegador para manipular elementos (ex: dar scroll, clicar via DOM, remover modal).

Este script percorre toda a lista de calçados da Adidas Brasil, extrai nome, preço, categoria e imagem de cada produto, pagina por pagina, e salva tudo em um JSON.