



CENTRO UNIVERSITÁRIO RUY BARBOSA WYDEN

ADLEI SANTOS PEREIRA
ANDRÉ FELIPE XAVIER OLIVEIRA SANTOS
ARTHUR MOTTA CUMMING
JORDYSON SANTOS DA CRUZ
MATHEUS VINÍCIUS FERREIRA COSTA

DOCUMENTAÇÃO TÉCNICA - COLETA AUTOMATIZADA DE PRODUTOS DA
CESTA BÁSICA

Salvador

2025

ADLEI SANTOS PEREIRA
ANDRÉ FELIPE XAVIER OLIVEIRA SANTOS
ARTHUR MOTTA CUMMING
JORDYSON SANTOS DA CRUZ
MATHEUS VINÍCIUS FERREIRA COSTA

DOCUMENTAÇÃO TÉCNICA - COLETA AUTOMATIZADA DE PRODUTOS DA
CESTA BÁSICA

Trabalho apresentado aos cursos de Graduação em Ciência da computação e Análise e desenvolvimento de sistemas do Centro Universitário Ruy Barbosa Wyden, como requisito parcial obrigatório da disciplina Tópicos de Big Data em Python.

Orientador(a): Roney Camargo Malaguti

Salvador

2025

SUMÁRIO

1. INTRODUÇÃO	6
2. TECNOLOGIAS UTILIZADAS	6
3. EXECUÇÃO DO CÓDIGO.....	7
4. POSSÍVEIS MELHORIAS FUTURAS	7
5. CONSIDERAÇÕES FINAIS.....	8

1. INTRODUÇÃO

Este projeto foi desenvolvido por um grupo de cinco alunos, com o objetivo de aplicar conhecimentos adquiridos em disciplinas dos cursos de Ciência da computação e Análise e desenvolvimento de sistemas do Centro Universitário Ruy Barbosa Wyden.

O projeto consiste em um dashboard que realiza a comparação de preços em tempo real de produtos da cesta básica entre diversos supermercados utilizando técnicas de web scraping e exibição visual com Streamlit. O sistema coleta dados de cinco sites distintos (Mercantil Atacado, Hiperideal, GBarbosa, Pão de Açúcar e Redemix), analisando os nomes dos produtos e seus respectivos preços. Após a coleta, os dados são tratados e organizados para exibição em um dashboard interativo, permitindo ao usuário identificar onde cada item está mais barato e acessar diretamente a página do supermercado correspondente.

Objetivo geral:

- Automatizar a coleta de preços de produtos essenciais (como arroz, feijão, leite, etc.) nos sites dos supermercados Redemix, GBarbosa, Mercantil, Hiperideal e Pão de Açúcar, com o objetivo de construir uma base de comparação de preços atualizada para uso acadêmico, social ou comercial.

2. TECNOLOGIAS UTILIZADAS

Para construção do projeto, foram utilizadas as seguintes ferramentas e tecnologias:

- Playwright (Python): Framework moderno e robusto para web scraping, ideal para sites com carregamento dinâmico. Usamos o modo headless para garantir que a coleta aconteça em segundo plano, sem abrir janelas visíveis.
- Pandas: Usado para manipulação e tratamento dos dados coletados, transformando-os em tabelas organizadas e prontos para exibição.
- Streamlit: Biblioteca de Python para construção de dashboards e aplicações web de forma rápida e simples. Foi utilizada para criar a interface de visualização dos preços.
- CSV: Utilizado no projeto para registrar as informações coletadas (nome, preço, link e mercado de origem) em softwares como Excel, Google Sheets e ferramentas de análise de dados. O uso do `pandas.to_csv()` garante a exportação dos dados de forma automatizada ao final da coleta.

A estrutura do código separa os scrapers por supermercado, com funções independentes para facilitar manutenção e expansão futura.

3. EXECUÇÃO DO CÓDIGO

Para executar o projeto, siga os seguintes passos:

1. Certifique-se de ter o Python instalado (recomenda-se versão 3.10 ou superior).
2. Execute o `python test.py` no terminal e aguarde finalizar a coleta.
3. Após finalizada a coleta, execute o `app.py` através do comando “`streamlit run app.py`”.
4. O dashboard será aberto automaticamente no navegador, exibindo os preços captados dos mercados em tempo real.

4. POSSÍVEIS MELHORIAS FUTURAS

O projeto desenvolvido apresenta um grande potencial de evolução e pode ser expandido de diversas formas, incorporando funcionalidades adicionais e melhorias técnicas que ampliam sua utilidade prática e integração com outras ferramentas. Abaixo, destacamos algumas dessas possibilidades:

- ❖ Adicionar mais supermercados e categorias de produtos além da cesta básica.
- ❖ Implementar agendamentos automáticos de scraping com salvamento em banco de dados.
- ❖ Aplicar técnicas de machine learning para prever variações de preços.
- ❖ Disponibilizar a aplicação em nuvem, acessível por qualquer dispositivo conectado à internet.
- ❖ Melhorar a experiência do usuário com filtros mais dinâmicos e histórico de preços.

A estrutura modular já preparada facilita a continuidade e o escalonamento do projeto.

5. CONSIDERAÇÕES FINAIS

Este projeto consolidou, na prática, diversas competências fundamentais do curso, como programação orientada a objetos, automação com Playwright, desenvolvimento de dashboards com Streamlit, tratamento de dados com Pandas e organização de código em ambiente colaborativo com forte aplicabilidade no mundo real, representando um passo importante na formação técnica e no preparo para projetos profissionais e acadêmicos.