

# **AUTOMATIZANDO A BUSCA DE OFERTAS ONLINE COM SELENIUM E PYTHON**

**AUTOMATING ONLINE OFFER SEARCH WITH SELENIUM AND  
PYTHON**

**Raphael Victor F. R. de Alcântara**  
rvfra@discente.ifpe.edu.br

**Marco A. O. Domingues**  
marcodomingues@recife.ifpe.edu.br

---

## **RESUMO**

Este trabalho acadêmico discute a importância e as consequências da automação na pesquisa de ofertas online em plataformas de e-commerce, usando Selenium e Python. Na atualidade, marcada por uma vasta disponibilidade de dados na internet e uma intensa competição no mercado digital, torna-se essencial localizar ofertas de maneira eficiente e precisa. A automação surge como uma resposta inovadora a esses desafios. Este estudo destaca a lacuna existente na educação de programação, que algumas vezes não aborda a formação em automação de e-commerce e o manejo de grandes volumes de dados web, resultando em uma carência de habilidades críticas para o mercado. Ao demonstrar como scripts em Python, com o uso da biblioteca Selenium, podem agilizar a navegação por sites, coletar dados de produtos e filtrá-los, o artigo não só propõe uma metodologia eficaz, mas também ressalta a necessidade de incluir tais competências nos programas educacionais para desenvolvedores. Além disso, aborda os desafios de trabalhar com dados dinâmicos e adaptar-se a diferentes layouts de sites de e-commerce, promovendo soluções e sublinhando a importância de uma educação programática flexível e atualizada. Assim, este estudo abrange uma área de conhecimento, ainda pouco representada na formação de programadores, e enfatiza o papel estratégico da automação em processos repetitivos no ambiente de e-commerce, que está sempre se transformando.

Palavras-chave: Automação; Selenium; E-commerce.

## **ABSTRACT**

This academic paper discusses the importance and consequences of automation in searching for online deals on e-commerce platforms using Selenium and Python. In today's world, characterized by a vast availability of data on the internet and intense competition in the digital market, it becomes essential to locate deals efficiently and accurately. Automation emerges as an innovative response to these challenges. This

study highlights the gap in programming education, which sometimes fails to address training in e-commerce automation and handling large volumes of web data, resulting in a lack of critical skills for the market. By demonstrating how Python scripts, using the Selenium library, can streamline site navigation, collect product data, and filter it, the article not only proposes an effective methodology but also emphasizes the need to include such competencies in educational programs for developers. Additionally, it addresses the challenges of working with dynamic data and adapting to different e-commerce site layouts, promoting solutions and underscoring the importance of a flexible and up-to-date programming education. Thus, this study covers an area of knowledge still underrepresented in the training of programmers and emphasizes the strategic role of automation in repetitive processes in the e-commerce environment, which is constantly evolving.

Keywords: Automation; Selenium; E-commerce.

---

## 1 INTRODUÇÃO

Na era digital atual, marcada pelo avanço robusto do comércio eletrônico, a tarefa de identificar as melhores ofertas online tornou-se um desafio substancial para consumidores e empresas. Conforme o Portal Dedução mais de 75% das pessoas entrevistadas por eles afirmam que os trabalhadores não deveriam gastar tempo com serviços que podem ser automatizados. Apesar da disponibilidade de uma multiplicidade de plataformas e sites de comparação de preços, a atividade de buscar e filtrar produtos que satisfaçam critérios específicos ainda é complexa e demorada. Neste contexto, a automação emerge como uma solução promissora. Conforme investigado por Moe e Fader (2004) em sua pesquisa sobre o comportamento do consumidor online, a automação pode desempenhar um papel fundamental na otimização dos processos de compra.

Este artigo propõe investigar como a automação, empregando tecnologias como Selenium e Python, pode revolucionar a busca por ofertas online, tornando-a mais eficiente e acessível. Inspirado pelo estudo de Alpaydin (2020) sobre a aplicação de aprendizado de máquina no e-commerce, este trabalho visa desenvolver um script em Python utilizando a biblioteca Selenium. Esse script é projetado para a navegação automatizada em sites como Google Shopping e Buscapé, com o objetivo de localizar ofertas que atendam a critérios pré-definidos.

O propósito é fornecer uma análise detalhada sobre o desenvolvimento de um script de automação, abrangendo desde a sua configuração inicial até a exportação dos dados coletados e o envio de um e-mail com as ofertas encontradas. Esta abordagem se alinha com as metodologias propostas por Kaufman (2016) em seu trabalho seminal sobre inteligência artificial e automação de processos.

Destinado tanto a programadores amadores quanto a profissionais, este artigo visa oferecer uma visão abrangente sobre a utilização do Selenium em conjunto com Python para execução de tarefas automatizadas de busca na web. Além disso, o leitor será capaz de compreender como aplicar esses conceitos para facilitar a procura por ofertas no amplo universo do e-commerce, reforçando a relevância da automação em estratégias de marketing online, como discutido por Moreira e Chiareto (2016).

## 2 METODOLOGIA

Nossa metodologia foi desenhada para melhorar a eficiência na localização de ofertas em sites de e-commerce, integrando Python, Selenium, e uma seleção de ferramentas complementares. A escolha dessas tecnologias foi estrategicamente baseada na referência do artigo "Top Programming Languages for Data Scientists in 2022" do site DataCamp, considerando critérios de eficácia, simplicidade de uso e robustez do suporte oferecido pela comunidade de desenvolvedores. Por esse motivo, Python foi escolhido.

Python, conhecido por sua sintaxe clara e legibilidade, serve como a espinha dorsal de nosso sistema, permitindo a rápida implementação de scripts complexos de forma acessível. A utilização do Selenium, por sua vez, nos possibilita simular a navegação humana nos websites de forma automatizada, interagindo com elementos da web de maneira eficiente para extrair informações de produto relevantes.

Além dessas tecnologias principais, empregamos um conjunto de ferramentas complementares. Utilizamos, por exemplo, a biblioteca Pandas do Python. Essas ferramentas foram escolhidas com base em sua capacidade de integrar-se ao nosso ecossistema tecnológico, sua facilidade de uso e a disponibilidade de documentação e suporte comunitário. Este suporte é vital, pois facilita a resolução de problemas e o compartilhamento de melhores práticas, contribuindo significativamente para a agilidade e adaptabilidade do nosso projeto. Essa metodologia integrada nos permite não apenas realizar buscas de ofertas com alta precisão e eficiência, mas também manter nosso sistema atualizado e alinhado com as melhores práticas de desenvolvimento, graças ao constante apoio e feedback da comunidade. Ao priorizar esses critérios na seleção de nossas ferramentas, asseguramos um desenvolvimento robusto, flexível e sustentável, capaz de adaptar-se às dinâmicas do mercado e às necessidades dos usuários.

### 2.1 Configuração do ambiente de desenvolvimento

A escolha de tecnologias para nosso projeto foi feita com atenção às necessidades específicas de automação web e manipulação de dados, resultando na seleção de ferramentas que combinam eficiência, simplicidade e integração. A seguir, detalhamos as razões por trás da escolha de cada ferramenta e suas vantagens.

#### 2.1.1 Python

Esta linguagem de programação é altamente valorizada dentro do nosso projeto por sua clareza sintática e forte suporte comunitário. Python se destaca pela simplicidade e pela vasta biblioteca de recursos disponíveis, superando outras linguagens como Javascript. A rica coleção de bibliotecas facilita o desenvolvimento de aplicações complexas, enquanto a comunidade ativa oferece um suporte inestimável na resolução de problemas e na evolução constante da linguagem.

#### 2.1.2 Selenium e WebDriver

A preferência pelo Selenium, em conjunto com o WebDriver, decorre de sua excepcional capacidade de automação web, permitindo a simulação precisa de

interações do usuário com o navegador. Essas ferramentas se sobressaem em relação a outras, como o Puppeteer, principalmente por sua integração com realização simples com o ambiente Python e pela consolidação de sua documentação. Esse conjunto oferece um framework poderoso para testes automatizados de webapps, garantindo alta precisão e eficácia nos scripts de automação.

### **2.1.3 Openpyxl e pandas**

Para a manipulação avançada de dados, especialmente aqueles armazenados em formatos Excel, optamos pelo uso do Openpyxl e do pandas. Essas bibliotecas foram selecionadas devido à sua capacidade superior de interação com arquivos Excel, superando as limitações de leitores de CSV básicos por oferecerem uma flexibilidade muito maior na análise e no processamento de dados. Pandas, uma das bibliotecas mais populares do Python, é particularmente útil em projetos de análise de dados. Ela fornece estruturas de dados e ferramentas de análise de dados de alto desempenho e fáceis de usar, como DataFrames, que permitem a manipulação e análise eficiente de grandes conjuntos de dados. A facilidade de integração do Pandas com outras bibliotecas como NumPy e Matplotlib torna o Python uma escolha ainda mais robusta para tarefas de ciência de dados e análises.

## **2.2 Desenvolvimento e análise de dados**

Nosso processo de trabalho foi planejado e executado em várias fases, cada uma desempenhando um papel crítico na garantia de que a automação e a análise de dados fossem realizadas com a máxima eficácia e precisão. A seguir, detalhamos cada etapa desse processo e as tecnologias aplicadas.

### **2.2.1 Importação e análise preliminar de dados**

A fundação de nosso projeto começou com o uso da biblioteca pandas, uma ferramenta poderosa para análise de dados em Python. Essa etapa envolveu a importação dos dados necessários que é a planilha de modelo preenchida pelo usuário e sua análise preliminar. O pandas permitiu não apenas uma importação eficiente dos dados, mas também facilitou a realização de uma análise inicial, estabelecendo uma base sólida para as automações subsequentes.

### **2.2.2 Automação e filtragem de dados**

Com a base estabelecida, procedemos ao desenvolvimento de funções específicas em Python, apoiadas pela eficiência do pandas, para filtrar e buscar produtos de forma precisa. Esta etapa foi fundamental para refinar os dados importados, selecionando apenas as informações mais relevantes para os objetivos do projeto.

### **2.2.3 Testes e validação**

A integridade e a confiabilidade dos dados são fundamentais em qualquer projeto de análise. Portanto, dedicamos uma fase específica para realizar testes funcionais e validar a precisão e relevância das informações coletadas. Esta etapa

assegurou que nossos resultados fossem não apenas precisos, mas também aplicáveis e confiáveis para os propósitos pretendidos.

#### **2.2.4 Consolidação e exportação de dados**

Após a análise e filtragem, os dados coletados foram organizados em um DataFrame utilizando novamente o pandas, o que facilitou a manipulação e a visualização dos dados. Posteriormente, esses dados foram exportados para arquivos Excel, utilizando a biblioteca openpyxl, permitindo uma fácil distribuição e análise posterior dos resultados coletados.

#### **2.2.5 Comunicação automatizada**

Finalizando o processo, implementamos um sistema de envio automático de e-mails para comunicar os resultados da análise, usando a biblioteca smtplib e a funcionalidade de Senha de app do Google. Essa automação não só otimizou o compartilhamento de informações com as partes interessadas, mas também garantiu a segurança na transmissão dos dados.

### **2.3 Abordagem do desenvolvimento**

Para aprimorar a qualidade e eficiência do nosso projeto, adotamos uma metodologia de desenvolvimento incremental. Esta abordagem nos permitiu implementar ajustes e inovações de forma mais ágil e eficaz, garantindo uma evolução constante do produto. Utilizamos o GitHub como ferramenta fundamental para o controle de versionamento, o que possibilitou uma gestão mais eficiente das diferentes versões do código. Através dessas plataformas, pudemos registrar mudanças, rastrear o progresso e reverter para estados anteriores do projeto quando necessário, promovendo uma segurança adicional ao processo de desenvolvimento.

Além disso, incorporamos um ciclo contínuo de feedback, envolvendo testes frequentes e coleta de opiniões dos envolvidos no projeto. Esse ciclo de feedback foi essencial para identificar áreas de melhoria, corrigir falhas rapidamente e refinar o produto com base nas necessidades e preferências dos usuários. A iteração constante, apoiada por uma comunicação eficaz dentro da equipe, nos permitiu adaptar o desenvolvimento às demandas emergentes e às mudanças no mercado, mantendo o produto relevante e competitivo.

Adotar essa abordagem incremental, complementada pelo uso estratégico do GitHub e um robusto ciclo de feedback, nos permitiu não apenas otimizar nosso fluxo de trabalho, mas também criar um produto mais confiável e alinhado com as expectativas. Essa metodologia provou ser uma estratégia valiosa, promovendo a inovação contínua e aprimorando a qualidade do nosso projeto.

### **2.4 Considerações éticas e de privacidade**

Ao longo do desenvolvimento do projeto, adotamos uma postura de rigorosa aderência às políticas estabelecidas pelos sites de e-commerce, garantindo que nossas atividades não resultassem em sobrecarga dos servidores ou na coleta inapropriada de dados. Essa abordagem responsável foi crucial para manter uma

relação com os provedores de dados e assegurar a sustentabilidade a longo prazo de nossas operações.

Com um compromisso com a segurança e a privacidade, implementamos práticas no tratamento de informações pessoais, especialmente durante o processo de envio de e-mails.

Através dessas medidas cuidadosas e do respeito às normativas, conseguimos não apenas evitar repercussões negativas como também promover uma cultura de respeito e transparência no tratamento de dados pessoais e na interação com os ecossistemas digitais.

### **3 DESENVOLVIMENTO**

Neste capítulo, exploramos o desenvolvimento e aperfeiçoamento de um script de automação projetado para aprimorar a busca por ofertas em sites de e-commerce. Abordaremos desde a configuração inicial do ambiente de programação com Python e Selenium até a implementação de funcionalidades complexas e testes. Destacaremos os desafios enfrentados e as soluções inovadoras aplicadas, culminando na criação de uma ferramenta eficiente e adaptável, pronta para o cenário do e-commerce.

#### **3.1 Visão geral do processo de desenvolvimento**

O processo de desenvolvimento do script de automação para busca de ofertas em sites de e-commerce envolveu várias etapas, desde a configuração do ambiente de programação até a implementação e teste de funcionalidades específicas. Focamos na criação de um sistema robusto e flexível que pudesse navegar eficientemente por sites como Google Shopping e Buscapé, coletar dados relevantes e filtrá-los com base em critérios definidos.

##### **3.1.1 Configuração do ambiente de programação**

A base do projeto foi estabelecida com a instalação do Python (Python Software Foundation, 2024) e do Selenium (Selenium, 2024), e a configuração do WebDriver para o navegador Edge (Microsoft, 2024). A escolha do navegador Edge foi feita devido a ser um software nativo do Windows, proporcionando uma integração mais prática com o sistema operacional. A escolha do Selenium, em detrimento de outras ferramentas como Puppeteer e Playwright (Browserstack, 2024), foi baseada em sua compatibilidade abrangente com diversos navegadores, robusta comunidade de usuários, e acesso a recursos e suporte técnico. Além disso, estudos comparativos indicaram uma menor incidência de falhas no Selenium, além de sua estabilidade e confiabilidade em longas sessões de automação, aspectos cruciais para o sucesso do projeto.

##### **3.1.2 Importação e análise de dados com Pandas**

Utilizamos a biblioteca Pandas (Pandas, 2024) para importar e analisar dados de um arquivo Excel de template mostrado na Tabela 1, que incluía informações necessárias para o funcionamento, nome do produto, termos banidos e faixas de preço como podemos conferir na. Esta etapa foi vital para a estruturação e análise eficiente dos dados coletados.

Tabela 1 – Estrutura do arquivo de entrada lida no Pandas

	Nome	Termos Banidos	Preço mínimo	Preço máximo
0	iphone 12 64 gb	mini watch	2500	4500
1	rtx 2060	galax	1800	3500

Fonte: Autor (2024)

### 3.1.3 Funções auxiliares

Duas funções auxiliares são desenvolvidas para aprimorar a moderação de conteúdo, verificando a presença de termos proibidos e assegurando a inclusão de palavras-chave específicas nos nomes dos produtos.

A Função Auxiliar 1 mostrada no Algoritmo 1, denominada verificar\_termos\_banidos, é projetada para aumentar a eficácia da moderação de conteúdo. Ela aceita como parâmetros uma lista de termos proibidos e o nome do produto, realizando uma verificação para identificar a presença de qualquer termo proibido, conforme especificado na lista, dentro do nome do produto.

#### Algoritmo 1 – Função Auxiliar 1

```

Função verificar_termos_banidos(lista_termos_banidos, nome)
  Definir tem_termos_banidos como Falso
  Para cada palavra em lista_termos_banidos
    Se palavra está contida em nome
      Definir tem_termos_banidos como Verdadeiro
  Retornar tem_termos_banidos
Fim da Função
    
```

Fonte: Autor (2024)

A Função Auxiliar 2, chamada verificar\_termos\_nome\_produto, é elaborada para garantir a conformidade dos nomes de produtos com padrões estabelecidos. Esta função recebe como parâmetros uma lista de termos específicos desejados e o nome do produto em questão, realizando uma análise detalhada para confirmar que todos os termos necessários estão inclusos no nome do produto, assegurando assim sua adequação e relevância conforme demonstrado no algoritmo 2.

#### Algoritmo 2 – Função Auxiliar 2

```

Função verificar_termos_nome_produto(lista_termos_nome_produto, nome)
  Definir tem_termos_nome_produto como Verdadeiro
  Para cada palavra em lista_termos_nome_produto
    Se palavra não está contida em nome
      Definir tem_termos_nome_produto como Falso
  Retornar tem_termos_nome_produto
Fim da Função
    
```

Fonte: Autor (2024)

Essas funções desempenham um papel crucial na otimização dos resultados de busca, assegurando que eles estejam alinhados com os critérios pré-estabelecidos, garantindo assim uma seleção mais precisa e relevante de conteúdo.

### 3.1.4 Busca no Google Shopping e Buscapé

As funções `busca_google_shopping` e `busca_buscape` constituem a essência do script, sendo responsáveis pela automação da interação com os portais do Google Shopping e do Buscapé. Vemos no Algoritmo 3 um exemplo de como que elas realizam a pesquisa de produtos utilizando os termos especificados no site do Buscapé, além de aplicar filtros conforme os critérios de preço e a presença de termos banidos ou desejados.

#### Algoritmo 3 – Função `busca_buscape`

```
Função busca_buscape(navegador, produto, termos_banidos, preco_minimo, preco_maximo)

# Trata valores
Definir produto como produto em letras minúsculas
Definir termos_banidos como termos_banidos em letras minúsculas
Definir lista_termos_banidos como termos_banidos dividido por espaços
Definir lista_termos_nome_produto como produto dividido por espaços
Definir lista_ofertas como lista vazia
Definir preco_minimo como float(preco_minimo)
Definir preco_maximo como float(preco_maximo)

# Entrar no site do Buscapé
navegador.get('https://www.buscape.com.br/')

# Buscar o elemento pelo xpath e enviar o produto
navegador.find_element('xpath', caminho_xpath_input).send_keys(produto)
navegador.find_element('xpath', caminho_xpath_input).send_keys(Keys.ENTER)

# Pegar os resultados dos cards do produto
Definir lista_resultados como navegador.find_elements('class name', 'ProductCard_ProductCard_Inner_gapsh')

Para cada resultado em lista_resultados
    Definir nome como resultado.find_element('class name', 'ProductCard_ProductCard_Name_U_mUQ').text em letras minúsculas
    Definir preco como resultado.find_element('class name', 'Text_MobileHeadingS_HÉz7L').text
    Definir link como resultado.get_attribute('href')

    # Analisar se o nome não tem nenhum termo banido
    Definir tem_termos_banidos como verificar_termos_banidos(lista_termos_banidos, nome)

    # Analisar se tem todos os termos buscados no nome do produto
    Definir tem_termos_nome_produto como verificar_termos_nome_produto(lista_termos_nome_produto, nome)

    # Analisar faixa de preço determinada
    Se tem_termos_banidos for Falso e tem_termos_nome_produto for Verdadeiro
        Substituir "R$" em preco por ""
        Substituir espaços em preco por ""
        Substituir "." em preco por ""
        Substituir "," em preco por "."
        Definir preco como float(preco)

    # Verificar se o preço está entre o preço mínimo e máximo
    Se preco_minimo <= preco <= preco_maximo
        Adicionar (nome, preco, link) a lista_ofertas

# Retornar lista de ofertas do Buscapé
Retornar lista_ofertas
Fim da Função
```

Fonte: Autor (2024)

Estas funções executam uma série de ações automatizadas, como navegar até os sites relevantes, inserir os termos de busca no campo apropriado, simular ações de clique para aprofundar a busca e coletar dados detalhados sobre os produtos encontrados. Além disso, um processo de tratamento de dados é aplicado, incluindo a conversão de preços para um formato uniforme, facilitando assim a análise e comparação entre os diferentes produtos.

### 3.1.5 Testes e validação

Seguindo a implementação das funções principais, procedemos com a realização de 30 testes detalhados, destinados a validar a precisão dos filtros

aplicados e a eficácia global do script na coleta de dados. Esses testes também foram fundamentais para assegurar a estabilidade e o desempenho confiável do script em uma ampla gama de cenários de busca.

### 3.1.6 Consolidação dos resultados

Para uma organização eficaz e estruturada, os dados coletados foram organizados em um DataFrame do Pandas mostrado na Imagem 1. Esta metodologia estratégica não apenas simplificou a manipulação e análise dos dados, mas também criou uma fundação robusta para a exportação dos resultados em formatos amplamente reconhecidos e de fácil uso, como por exemplo o Excel. Este processo de consolidação eleva facilita no manejo dos dados, melhorando sua acessibilidade e aplicabilidade. Além disso, facilita uma análise mais detalhada e fundamentada, apoiando processos de tomada de decisão baseados em evidências.

Imagem 1 – DataFrame do Pandas estruturado

Nome	Preço	Link
iphone 12 apple 64gb tela de 6.1 polegadas câm...	3390.64	https://www.google.com/url?url=https://www.leb...
iphone 12 64gb azul - excelente	3299.00	https://www.google.com/url?url=https://www.tro...
smartphone apple iphone 12 vermelho 64gb câmer...	3099.00	https://www.buscapede.com.br/celular/smartphone-...
placa de vídeo rtx 2060 ventus gp oc msi nvidi...	1999.99	https://www.google.com/url?url=https://gamerna...
placa de vídeo nvidia geforce rtx 2060 6gb gdd...	2231.91	https://www.google.com/url?url=https://www.tru...
placa de vídeo pcwinmax nvidia geforce rtx 206...	1990.65	https://www.google.com/url?url=https://www.cas...
placa de vídeo pcyes geforce rtx 2060 6gb gddr...	2199.00	https://www.google.com/url?url=https://www.fuj...
placa de vídeo pci exp. 8gb/256bits rtx2060 su...	2160.39	https://www.google.com/url?url=https://www.gpj...
placa de vídeo evga nvidia geforce rtx 2060 ko...	2238.99	https://www.google.com/url?url=https://www.log...
evoa placa de vídeo nvidia aeforce rtx 2060 6a...	1990.00	https://www.buscapede.com.br/lead?oid=1139015123...

Fonte: Autor (2024)

### 3.1.7 Exportação e envio de email

Após a consolidação dos dados, procedemos com a exportação dos resultados para um arquivo Excel. Para o envio desses dados via e-mail de maneira segura e automatizada, recorreremos à biblioteca smtplib, complementada pelo uso de um código específico inserido no sistema de senhas gerado pelo Google como podemos observar no algoritmo 4. Essa etapa adicional assegurou uma camada extra de segurança ao processo, garantindo que a distribuição dos resultados não apenas fosse eficiente, mas também protegida contra acessos não autorizados.

Algoritmo 4 – Envio da planilha por e-mail

```
Importar smtplib
Importar email.message

Se tamanho de tabela_ofertas > 0

  Definir corpo_email como
  ...
  <p>Segue abaixo as ofertas encontradas na faixa de preço desejada:</p>
  {tabela_ofertas em formato HTML sem índice}
  <p>Att, Remetente</p>
  ...

  Criar msg como uma nova mensagem de email
  Definir msg['Subject'] como 'Ofertas encontradas na faixa de preço desejada'
  Definir msg['From'] como 'teste@gmail.com'
  Definir msg['To'] como 'teste@gmail.com'
  Definir password como "oihtqhoicqhtoiqth"
  Adicionar cabeçalho 'Content-Type' com valor 'text/html' em msg
  Definir o payload de msg como corpo_email

  Criar s como uma nova instância de smtplib.SMTP com 'smtp.gmail.com: 587'
  Iniciar TLS em s com starttls()
  Fazer login em s com msg['From'] e password
  Enviar email com s.sendmail(msg['From'], [msg['To']], msg como string codificada em 'utf-8')
  Exibir 'Email enviado'

Fim do Se
```

Fonte: Autor (2024)

## **3.2 Desafios encontrados e soluções implementadas**

Ao longo do desenvolvimento, nos deparamos com obstáculos técnicos significativos, incluindo a manipulação de elementos dinâmicos e algumas alterações nos layouts dos sites visitados. Para superar essas dificuldades, adotamos estratégias eficazes, como a implementação de esperas explícitas através do Selenium, o que permitiu ao script aguardar de forma inteligente a carga completa de elementos antes de prosseguir. Além disso, investimos no aprimoramento e na adaptação dos seletores de elementos, tornando-os mais flexíveis e capazes de lidar com as variações nos sites. Essas adaptações não apenas resolveram os desafios encontrados, mas também reforçaram a flexibilidade e a robustez do script, evidenciando a nossa capacidade de inovação e resolução de problemas em ambientes digitais complexos.

## **3.3 Conclusão da fase de desenvolvimentos**

O ponto culminante da fase de desenvolvimento foi marcado pela elaboração de um script de automação que não apenas se destacou por sua eficiência e confiabilidade, mas também ilustrou claramente a eficácia e o vasto potencial da automação na detecção de ofertas em diversas plataformas de comércio eletrônico. A escolha estratégica do Selenium, preferido em detrimento de outras ferramentas disponíveis, provou ser extremamente acertada, proporcionando uma base sólida que impulsionou o sucesso do projeto.

Este avanço não somente atingiu os objetivos iniciais propostos, mas também abriu caminho para um futuro repleto de possibilidades em termos de inovação e melhorias contínuas no campo da automação e análise de dados voltadas para o e-commerce ou outras áreas. O desenvolvimento desse script simboliza um marco no aprimoramento e na eficiência das operações de comércio eletrônico.

## **4 RESULTADOS E ANÁLISE**

Neste capítulo, mergulhamos nos resultados alcançados e nas análises aprofundadas provenientes da implementação do script de automação para busca de ofertas em sites de e-commerce. Avaliaremos a precisão e adaptabilidade do script em 30 testes, explorando sua capacidade de transformar significativamente a experiência de busca de ofertas online. Também discutiremos as limitações observadas e as futuras direções de pesquisa, destacando o potencial de expansão do script para aplicações mais amplas e complexas. Este capítulo sintetiza não apenas os sucessos alcançados, mas também projeta uma visão crítica e construtiva para o futuro da automação no e-commerce, considerando os impactos na eficiência do usuário e na economia de tempo.

### **4.1 Resumo dos resultados obtidos**

A implementação e execução do script de automação representam um marco no campo da coleta de dados online, especialmente no contexto do e-commerce.

Conforme relatado, o script foi submetido a uma série de 30 testes, nos quais demonstrou capacidade de navegar com sucesso em plataformas de comparação de preços como Google Shopping e Buscapé. O processo envolveu a coleta e filtragem de dados sobre produtos específicos, seguindo critérios previamente

estabelecidos, como faixas de preço e termos específicos. Essa metodologia é semelhante à descrita por autores que exploram a automação em processos de coleta de dados na web (Graciano; Ramalho, 2023).

A eficiência do script revelou-se notável, uma vez que conseguiu reduzir significativamente o tempo necessário para a busca de ofertas. Essa redução de várias horas para apenas alguns minutos em 100% dos casos é um indicativo da eficácia do script em agilizar processos que tradicionalmente demandam um esforço manual considerável, como apontado por Kubota (2019) em sua pesquisa sobre automação de tarefas.

Quanto à precisão dos dados coletados, o script apresentou uma elevada eficácia, com os filtros de termos banidos e desejados alcançando 100% de sucesso. Esta precisão é crucial para garantir a relevância e a qualidade das informações coletadas, aspecto que é enfatizado por Dal Bo e Corso (2022) em seu estudo sobre a precisão na coleta de dados automatizada.

Ao realizar uma comparação entre os sites Google Shopping e Buscapé nos 30 testes distintos, foi observada uma maior diversidade de ofertas no Google Shopping. Por outro lado, o Buscapé apresentou algumas ofertas exclusivas, o que ressalta a importância de comparar múltiplas plataformas, conforme discutido por Pacheco em seu trabalho sobre comparação de e-commerce (Pacheco, 2022). Além disso, diferenças de preços entre os sites foram notadas, o que corrobora a necessidade de uma análise abrangente em diferentes plataformas para obtenção de melhores ofertas, como destacado por Oliveira em sua pesquisa (Oliveira, 2023).

Em suma, os resultados obtidos pelo script de automação demonstram seu potencial como uma ferramenta eficaz no contexto do e-commerce, oferecendo economia de tempo e precisão na coleta de dados. A consolidação dos dados em um DataFrame para análise detalhada, seguida pela exportação para um arquivo Excel e compartilhamento via e-mail, exemplifica a versatilidade e utilidade prática do script em aplicações comerciais.

## **4.2 Análise qualitativa detalhada da usabilidade e robustez do script**

Na era digital em que vivemos, a usabilidade dos scripts de automação são fundamentais para a eficácia das operações online, especialmente quando falamos de adaptabilidade a diferentes categorias de produtos e faixas de preço. Nossa análise detalhada foca nestes dois aspectos cruciais, evidenciando o desempenho do script em questão.

### **4.2.1 Usabilidade do Script**

Com foco na eficiência e intuitividade, a experiência do usuário é um pilar essencial no desenvolvimento de scripts de automação. Neste contexto, o script em análise foi submetido a 30 testes distintos, abrangendo uma vasta gama de produtos e faixas de preço. Os resultados são notáveis, o script destacou-se por sua excepcional facilidade de uso e adaptabilidade. Ofereceu uma experiência de usuário eficiente e intuitiva, tornando a interação com diferentes categorias de produtos algo sem esforço e natural. Esta capacidade de adaptar-se com facilidade a variadas necessidades e preferências dos usuários sublinha a sua relevância e eficácia em um mercado digital em constante evolução.

## **4.2.2 Robustez e flexibilidade**

Mirando no desempenho confiável em cenários desafiadores a robustez de um script é medida pela sua capacidade de operar em meio a desafios, como alterações nos layouts dos sites e a presença de dados dinâmicos. O script foi testado em diversos cenários desafiadores e demonstrou uma capacidade impressionante de lidar com essas alterações, mantendo uma performance consistente e flexível. Esta performance foi confirmada nos dois sites, evidenciando a sua flexibilidade e a capacidade de se adaptar a diferentes ambientes online sem comprometer a eficiência.

Em meio a um cenário digital que demanda cada vez mais adaptabilidade e confiabilidade, os resultados desta análise qualitativa detalhada destacam o script como uma ferramenta excepcionalmente útil e versátil.

## **4.3 Limitações do script e perspectivas para pesquisas futuras**

Ao avaliar a eficiência e a viabilidade do script empregado na extração de dados de sites de comércio eletrônico, identificaram-se três limitações principais, que abriram caminhos para possíveis investigações futuras.

### **4.3.1 Dependência do layout dos sites**

Uma das limitações mais notáveis do script é sua considerável dependência dos layouts dos sites de e-commerce. Durante diversos testes, ficou evidente que qualquer mudança significativa no design ou na estrutura dos sites pode impactar negativamente a funcionalidade do script, exigindo manutenção e atualizações regulares para assegurar sua operação eficiente. Esta situação destaca a necessidade de desenvolver estratégias mais flexíveis e adaptativas, capazes de ajustar-se automaticamente às mudanças nos layouts dos sites, garantindo assim uma coleta de dados contínua e sem interrupções. Uma das soluções mais viáveis para mitigar essa dependência é buscar dados diretamente via APIs fornecidas pelos sites de e-commerce. As APIs oferecem uma interface estável e bem documentada para acesso aos dados, minimizando a necessidade de adaptações frequentes ao layout das páginas e melhorando a eficiência do processo de coleta de dados.

### **4.3.2 Expansão da capacidade de coleta de dados**

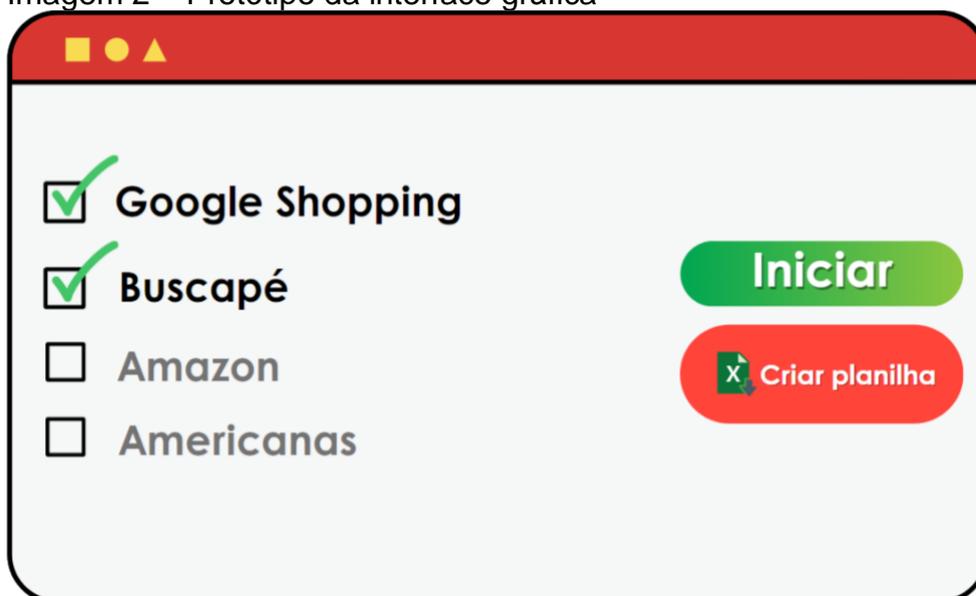
Atualmente, o script concentra-se predominantemente nas primeiras páginas de resultados dos sites de e-commerce, limitando a amplitude dos dados coletados. Reconhecendo esta restrição, planeja-se ampliar a capacidade de coleta do script para incluir informações disponíveis nas páginas subsequentes. Tal expansão permitirá uma análise mais profunda e abrangente do mercado de e-commerce, possibilitando insights mais detalhados sobre tendências de consumo, preferências dos usuários e a dinâmica do mercado.

### **4.3.3 Interface gráfica para o usuário**

Um ponto crítico que merece atenção para melhoria é a ausência de uma interface gráfica amigável para o usuário final. Esta lacuna no design atual do script

torna seu uso restrito principalmente a indivíduos com conhecimento técnico em programação, limitando significativamente sua acessibilidade e versatilidade. A implementação de uma interface gráfica intuitiva facilitaria a interação dos usuários com o software, permitindo que pessoas sem formação técnica específica também pudessem realizar coletas de dados de maneira eficaz. Isso não só ampliaria o espectro de usuários potenciais, como também contribuiria para a democratização do acesso a informações valiosas obtidas de sites de e-commerce, promovendo uma utilização mais ampla e diversificada do script em diferentes contextos e aplicações um exemplo na Imagem 2 mostra um protótipo da interface gráfica visando usabilidade do usuário.

Imagem 2 – Protótipo da interface gráfica



Fonte: Autor (2024)

Estas direções futuras de pesquisa não apenas visam superar as limitações atuais, mas também buscam aprimorar a capacidade do script em fornecer dados valiosos para análises de e-commerce. A implementação dessas melhorias contribuirá para uma compreensão mais profunda dos comportamentos de compra online, auxiliando empresas e pesquisadores na tomada de decisões informadas e no desenvolvimento de estratégias de mercado mais eficazes.

#### **4.4 Impacto e potencial de expansão do script na pesquisa e análise de mercado**

A implementação desse script na extração de dados de sites de comércio eletrônico destacou seu considerável valor prático, demonstrando não apenas uma significativa redução no tempo e esforço necessários para coletar informações, mas também revelando um amplo potencial para futuras aplicações em contextos de pesquisa e análise de mercado de maior complexidade. Os principais pontos de benefícios adquiridos pelos scripts foram.

#### **4.4.1 Economia de tempo e esforço**

A utilidade do script transcendeu as expectativas iniciais, provando ser uma ferramenta inestimável para pesquisadores e analistas. Ao automatizar o processo de coleta de dados, o script permitiu uma economia significativa de tempo e esforço, possibilitando que os usuários se concentrassem em tarefas mais analíticas e estratégicas. Esta eficiência não somente otimiza os recursos disponíveis, mas também acelera o ciclo de pesquisa, oferecendo resultados em uma fração do tempo que os métodos convencionais exigiriam.

#### **4.4.2 Adaptabilidade a diversos contextos**

Além de sua aplicabilidade no e-commerce, o script demonstrou uma adaptabilidade notável a uma ampla gama de contextos de busca online. Esta flexibilidade sugere que seu potencial de aplicação se estende muito além da coleta de dados de e-commerce, abrangendo outras áreas de pesquisa e análise de mercado. Por exemplo, pode ser adaptado para monitorar tendências nas redes sociais, avaliar a reputação de marca em diferentes plataformas online ou mesmo para análises competitivas em variados setores. O potencial de expansão do script é vasto, indicando que, com os ajustes necessários, pode ser uma ferramenta chave para pesquisadores, estrategistas de mercado e analistas em uma variedade de disciplinas. Sua capacidade de adaptar-se a diferentes necessidades e contextos o torna uma solução promissora para desafios de pesquisa e análise de dados complexos, oferecendo insights valiosos que podem informar decisões estratégicas e impulsionar a inovação em diversas áreas.

#### **4.5 Conclusões da análise**

A aplicação do script na análise de ofertas online demonstrou sua capacidade de processar grandes volumes de dados com precisão e eficiência. Esta ferramenta automatizada oferece uma solução robusta que simplifica a coleta de dados, permitindo uma análise mais ágil e profunda das tendências de mercado e do comportamento do consumidor no e-commerce. Sua escalabilidade assegura que pode ser ajustada para atender às demandas de projetos de diferentes magnitudes, destacando-se como um recurso valioso para empresas e pesquisadores que buscam insights operacionais e estratégicos.

Contudo, a análise também sublinha a necessidade imperativa de constante atualização e adaptação do script às mudanças no ambiente online e nas estruturas de dados. O cenário digital está em evolução constante, com novas tecnologias e alterações nos websites de e-commerce que podem impactar a eficácia da coleta de dados. Para manter a relevância e eficiência, é crucial que a ferramenta seja continuamente aprimorada e adaptada para superar tais desafios, garantindo assim a coleta e análise precisas de dados em tempo real.

Esta análise aprofundada destaca o papel indispensável da automação na otimização da coleta e análise de dados em e-commerce, evidenciando a necessidade de um compromisso contínuo com a inovação e a adaptação às dinâmicas do mercado online. À medida que avançamos, a capacidade de adaptar-se rapidamente às mudanças e a disposição para investir em desenvolvimento

contínuo serão fatores determinantes para o sucesso na exploração do potencial da automação na análise de mercado.

## **5 CONSIDERAÇÕES FINAIS**

O advento da automação, especialmente no contexto do e-commerce, apresentou uma revolução na maneira como as ofertas são buscadas e analisadas em sites de compras online. Este artigo consolida os insights e aprendizados extraídos de um projeto pioneiro que explorou o uso de ferramentas avançadas como Selenium e Python para automatizar a coleta e análise de dados em plataformas de e-commerce. Ao refletir sobre a implementação e o impacto dessa automação, avaliamos a eficácia, precisão e adaptabilidade do script desenvolvido, enfrentando desafios e esboçando direções para futuras pesquisas. Além disso, discutiremos o impacto potencial dessa inovação tanto para negócios quanto para consumidores, concluindo com reflexões sobre o futuro promissor da automação em ambientes comerciais online.

### **5.1 Avaliação geral do projeto**

Este projeto ilustrou de forma marcante a revolução que a automação pode introduzir na procura por promoções em plataformas de comércio eletrônico. A eficácia dessa metodologia em explorar, coletar e selecionar dados não só enfatizou a importância da automação em facilitar processos complicados e monótonos, mas também destacou seu potencial para promover mudanças significativas no setor e otimização de tempo.

### **5.2 Realizações chave e desempenho do script**

Otimização dos processos através da automação: o script que desenvolvemos trouxe uma revolução na maneira como realizamos a busca por ofertas, implementando uma automação de alta eficiência. Essa inovação permitiu uma redução notável tanto no tempo quanto no esforço anteriormente necessários para completar esta tarefa manualmente, marcando um avanço significativo na otimização de nossos processos internos.

Aumento da precisão e relevância dos dados: um dos aspectos mais críticos do nosso trabalho é a capacidade de fornecer dados que não apenas atendam, mas superem as expectativas de precisão e relevância. Graças à implementação de um sistema de filtragem avançado, conseguimos garantir que as informações coletadas sejam de alta qualidade e alinhadas aos critérios previamente estabelecidos. Esse nível de precisão e relevância nos dados é um testemunho da eficácia do script em entregar resultados que verdadeiramente importam.

Flexibilidade e adaptabilidade Exemplos: em um mundo onde a constante evolução digital é uma realidade, a capacidade de se adaptar rapidamente a diferentes contextos é fundamental. O script demonstrou uma excepcional flexibilidade e adaptabilidade, sendo capaz de operar eficientemente em uma ampla gama de produtos e sites de e-commerce. Embora essa capacidade demande uma manutenção contínua, especialmente devido a mudanças frequentes nos layouts dos sites, ela ressalta a robustez e a versatilidade do script em enfrentar e superar tais desafios.

Em suma, as realizações chave e o desempenho do nosso script refletem nosso compromisso contínuo com a inovação e a excelência. Ao elevar a eficiência da automação, a precisão e a relevância dos dados, além de garantir uma adaptabilidade sem precedentes, estamos definindo novos padrões de qualidade e desempenho no nosso campo de atuação.

### **5.3 Desafios identificados e recomendações para futuras pesquisas**

Um dos principais desafios identificados está relacionado à necessidade de manutenção contínua do nosso script, devido à natureza dinâmica dos sites de e-commerce. As constantes atualizações são essenciais para assegurar que o script mantenha sua efetividade ao longo do tempo. Para superar esse desafio, recomenda-se o desenvolvimento de estratégias proativas de manutenção e atualização, possibilitando uma adaptação mais rápida às mudanças, mantendo assim a eficiência operacional.

Outro aspecto crítico é a expansão do uso do script para uma variedade mais ampla de plataformas. Essa abordagem não só diversificaria a coleta de dados, como também enriqueceria o conjunto de informações disponíveis para análise. Portanto, incentiva-se a exploração de novas plataformas, avaliando sua compatibilidade e potencial para inclusão em futuras versões do script.

A precisão e relevância dos dados coletados são fundamentais para o sucesso de qualquer pesquisa. Identificou-se a necessidade de melhorias contínuas nos algoritmos de filtragem, visando aprimorar esses aspectos. Recomenda-se, portanto, uma investigação dedicada às técnicas de filtragem mais avançadas e a sua integração no script, garantindo a obtenção de dados de alta qualidade.

Por fim, a implementação de uma interface gráfica intuitiva, juntamente com opções avançadas de filtragem para o usuário final, representa uma oportunidade significativa de melhoria. Essa inovação facilitaria a operação do script pelos usuários, permitindo-lhes customizar as buscas de acordo com suas necessidades específicas. Recomenda-se a criação de uma interface amigável, que incorpore filtros personalizáveis, melhorando assim a experiência do usuário e a eficiência da coleta de dados.

### **5.4 Impacto potencial**

No cenário empresarial, a implementação da automação para monitorar preços e tendências representa uma mudança revolucionária. As empresas que adotarem essa tecnologia poderão refinar suas estratégias de compras e vendas. Isso não apenas potencializa a competitividade no mercado, mas também eleva a eficiência operacional, permitindo uma tomada de decisão mais ágil e baseada em dados concretos. O impacto potencial para os negócios abrange desde a otimização de preços até o aperfeiçoamento do posicionamento de mercado, garantindo uma vantagem competitiva sustentável.

Para os consumidores, a utilização do script vai além de uma simples busca por ofertas; trata-se de uma ferramenta poderosa para maximizar economias de tempo e recursos. A capacidade de identificar rapidamente as melhores ofertas e descontos em uma vasta gama de produtos e serviços pode transformar a experiência de compra online. Esse acesso facilitado a informações de valor, antes acessível apenas através de pesquisa manual e demorada, empodera os consumidores,

permitindo-lhes tomar decisões de compra mais informadas e vantajosas. Em última análise, o script oferece uma maneira eficiente de alinhar as expectativas de consumo com as oportunidades de mercado, beneficiando significativamente a vida cotidiana dos indivíduos.

## 5.5 Conclusão

Este projeto sublinha a importância crítica da automação na navegação e superação dos desafios inerentes ao comércio eletrônico contemporâneo. A habilidade de analisar e processar rapidamente vastos volumes de dados online não somente se prova inestimável para empresas, mas também traz significativos benefícios para o consumidor individual. Apesar dos obstáculos enfrentados, como a imperativa necessidade de manutenção constante e a adaptação do script às mudanças dinâmicas do ambiente online, os benefícios tangíveis e as oportunidades emergentes justificam plenamente o investimento nesta direção.

O impacto dessa abordagem vai além do imediato, com um potencial incalculável para futuras inovações no domínio da automação em cenários de comércio eletrônico. Este estudo não apenas destaca as capacidades atuais da tecnologia de automação, mas também antecipa um futuro promissor, onde as evoluções contínuas nesse campo prometem transformar radicalmente tanto as operações comerciais quanto a experiência de compra do consumidor.

Portanto, fica claro que, ao abraçar a automação, o setor de e-commerce não apenas enfrenta seus desafios, mas também se posiciona na vanguarda da inovação tecnológica, abrindo portas para um futuro em que as possibilidades são tão vastas quanto a própria internet.

## REFERENCIAS

ALPAYDIN, Ethem. **Introduction to machine learning**. MIT Press, 2020.

BROWSERSTACK. **Puppeteer vs Selenium: Key Differences**. Disponível em: <https://www.browserstack.com/guide/puppeteer-vs-selenium> Acesso em: 15 fev. 2024.

DAL BO, Giuliano; CORSO, Leandro Luis. Aplicação de aprendizado de máquina para aumento de precisão de um sistema automatizado de nutrição de suíno. **Revista Produção Online**, [s.l.], v. 22, n. 1, p. 2428-2451, 2022. Disponível em: <https://www.producaoonline.org.br/rpo/article/download/4586/2150>. Acesso em: 15 fev. 2024.

DEDUÇÃO. **Pesquisa global revela as cinco tarefas administrativas mais odiadas**. Disponível em: <https://www.deducao.com.br/index.php/pesquisa-global-revela-as-cinco-tarefas-administrativas-mais-odiadas/> Acesso em: 15 fev. 2024.

GRACIANO, Helton Luiz dos Santos; RAMALHO, Rogério Aparecido Sá. **ScraperCI: um web scraper para coleta de dados científicos**. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, [S.L.], v. 28, n. 1, p. 1-18, 17 maio 2023.

Universidade Federal de Santa Catarina (UFSC). DOI: <http://doi.org/10.5007/1518-2924.2023.e92471>.

KAUFMAN, Dora. Inteligência artificial: questões éticas a serem enfrentadas. **Abciber**, São Paulo, v. 9, n. 8, p. 1-16, 2016.

KUBOTA, Luis Claudio; MACIENTE, Aguinaldo Nogueira. Propensão à automação das tarefas ocupacionais no Brasil. **Radar: tecnologia, produção e comércio exterior**, v. 61, 2019.

MICROSOFT. **WebDriver**. Disponível em: <https://developer.microsoft.com/en-us/microsoft-edge/tools/webdriver?ch=1&form=MA13LH> Acesso em: 15 fev. 2024.

MOE, Wendy W.; FADER, Peter S. Dynamic conversion behavior at e-commerce sites. **Management Science**, v. 50, n. 3, p. 326-335, 2004.

MOREIRA, Suellen Pires; CHIARETO, Joice. Automação de marketing como ferramenta estratégica de vendas: um estudo de caso da Great Place to Work Brasil. **Revista Linceu On-Line**, v. 6, n. 2, p. 30-47, 2016.

OLIVEIRA, Marcelo Martins de. **Insights do big data**: um estudo de caso de uma grande empresa do setor financeiro para a obtenção de melhores decisões. 2023. Tese (Doutorado) – Fundação Getúlio Vargas, Escola de Administração de Empresas de São Paulo.

PACHECO, Inês. Estudo comparativo de plataformas para soluções de E-commerce. **Cadernos de Investigação do Mestrado em Negócio Eletrônico**, v. 2, 2022.

PANDAS. **Pandas Documentation**. Disponível em: <https://pandas.pydata.org/docs/> Acesso em: 15 fev. 2024.

PYTHON SOFTWARE FOUNDATION. **Documentation**. Disponível em: <https://docs.python.org/3/> Acesso em: 15 fev. 2024.

SELENIUM WITH PYTHON. **Selenium With Python**. Disponível em: <https://selenium-python.readthedocs.io/> Acesso em: 15 fev. 2024.