

Extração de datasets em marketplaces de anúncios: experiências com Facebook, Olx e MercadoLivre

Dataset extraction from advertising marketplaces: experiences with Facebook, Olx, and MercadoLivre

Eduardo Diniz Amaral

Universidade Estadual de Montes Claros / Universidade Federal de Santa Catarina, Brasil.

E-mail: eduardodiniz@gmail.com

Gustavo Medeiros de Araújo

Universidade Federal de Santa Catarina, Brasil.

E-mail: gustavo.araujo@ufsc.br

RESUMO

Diante da quantidade cada vez maior de informação produzida em marketplaces digitais, bem como a crescente adesão e estes tipos de serviços por parte dos usuários de internet no Brasil e no mundo, bem como da falta de trabalhos relacionados a este tema, esta pesquisa tem como objetivo realizar experiências com extração de datasets de anúncios. Após análise dos principais espaços de e-commerce no Brasil, os marketplaces escolhidos foram: Mercado Livre, Facebook e OLX A linguagem de programação escolhida foi Python e as ferramentas para extração scrapy, bibliotecas request, beautifulsoup e Selenium Webdriver. Após análise da estrutura web das páginas dos resultados de anúncios, foram elaborados scripts de extração das principais variáveis do anúncio, dentro de uma categoria em comum entre os marketplaces. Os resultados apontam que é possível extrair datasets de anúncios nestas plataformas em diversos formatos por meio de scrapers e que tais informações tem potencial para exploração em diversos segmentos da ciência de dados.

Palavras-chave: Datasets, Python, Ecommerce, Marketplaces, Anúncios.

Como citar: Amaral, E. D.; & Araújo, G. M. (2022). Extração de datasets em marketplaces de anúncios: experiências com Facebook, Olx e MercadoLivre. En T. M. R. Dias (Ed.), Informação, Dados e Tecnologia. *Advanced Notes in Information Science, volume 2* (pp. 63-73). Tallinn, Estônia: ColNes Publishing. DOI: 10.47909/anis.978-9916-9760-3-6.103.

Copyright: © 2022, The author(s). This is an open access work distributed under the terms of the CC BY-NC 4.0 license which permits copying and redistributing the material in any medium or format, adapting, transforming and building upon the material as long as the license terms are followed.

ABSTRACT

Considering the increasing amount of information produced in digital marketplaces, the growing adhesion to these types of services by internet users in Brazil and the world, and the lack of work related to this topic, this research aims to experiment with dataset extraction ads. After analyzing the main e-commerce spaces in Brazil, the chosen marketplaces were: Mercado Livre, Facebook, and OLX. Python's programming language uses the following libraries: scrapy, BeautifulSoup, and Selenium Webdriver. After analyzing the web structure of the ad results pages, scripts were created to extract the main variables of the ad within a common category among the marketplaces. The results show that scrapers can remove datasets from advertisements on these platforms in different formats. Such information has potential for exploration in various segments of data science.

Keywords: Datasets, Python, E-commerce, Marketplaces. Advertisings

1. INTRODUÇÃO

OS DADOS estatísticos apresentados pelo CETIC.BR (2020) mostram uma adesão e participação cada vez maior do Brasil no e-commerce. Dados apontam que o e-commerce brasileiro deve crescer mais de 28% em 2021, impulsionado pelo aumento no número de consumidores, consolidação de lojas virtuais locais, fortalecimento de marketplaces e aumento na eficiência logística. (ECOMMERCEBRASIL, 2020). As perspectivas valem para os mais diversos espaços e categorias de comércio eletrônico.

Os volumes de anúncios e transações são desafiadores. Para se ter ideia e exemplo, de acordo com OLX (2021), o site de anúncios OLX contabiliza cerca de 500.000 anúncios por dia, 50 vendas por minuto, 100 milhões de buscas por mês. Informa ainda que cerca de 50% dos internautas brasileiros já passaram pela OLX. São expressivos volumes de dados que compõem anúncios - gênero textual usado para promover uma marca, produto, serviço ou ideia, envolvendo títulos, descrições, fotos, vídeos, categorias, localizações e contatos dos anunciantes.

Apesar de tamanha quantidade de dados, as potencialidades informacionais destes espaços são pouco discutidas na comunidade científica. A literatura aponta que existem trabalhos que buscam capturar e interpretar estes dados, gerando indicadores e novos conhecimentos, em diversas áreas de pesquisa: econômicas, sociais aplicadas, comportamentais e mercadológicas. O trabalho de Wijaya (2020) utilizou dados de e-commerce para predição de taxa de pobreza em uma cidade da Indonésia utilizando para isso machine learning. Pandey (2020) demonstra que é possível utilizar algoritmos do tipo random forest para predição de preço

de venda de automóveis. Mohamed (2020) mostra que é possível, por meio de deep learning, obter predição de preço de itens usados disponibilizados em plataformas de anúncios. O trabalho de Xu (2020) discute técnicas e metodologias para identificar anúncios ilegais na rede social Facebook envolvendo animais selvagens na China. Zaheer (2018) mostra a utilização de algoritmos do tipo random forest em datasets extraídos do marketplace OLX.

Existem também trabalhos que buscam alternativas e ferramentas para a extração de informações disponíveis em websites, portais, marketplaces, lojas e afins.

Neste panorama, a ciência de dados não poderia se posicionar de maneira indiferente. Poderiam dados públicos de plataformas de anúncios serem extraídos, organizados e utilizados pelas ciências de dados?

2. OBJETIVOS

O objetivo deste trabalho é propor técnicas e ferramentas capazes de extrair datasets em marketplaces cujo objeto principal de tráfego são anúncios de produtos. Serão testadas algumas ferramentas de extração de dados disponíveis na web em formato HTML. Os testes serão realizados em três marketplaces de formatos distintos: OLX, MercadoLivre e Facebook.

3. PROCEDIMENTOS METODOLÓGICOS

De acordo com Gerhardt e Silveira (org., 2009), esta pesquisa é de caráter exploratório e bibliográfico, de natureza aplicada., visto que faz uso de artigos, manuais e outros documentos a respeito de plataformas e ferramentas de softwares. Adiante, os métodos são descritos em etapas. O primeiro passo será analisar os principais marketplaces de anúncios no Brasil e selecionar três deles para, em seguida, realizar uma análise estrutural dos dados públicos disponíveis, em termos de metadados e estruturas semânticas.

A seguir, será definida uma categoria de anúncios em comum entre os três marketplaces para mineração de anúncios disponíveis, bem como variáveis e abrangência geográfica dos anúncios. A próxima etapa envolve a análise da estrutura web dos portais e disposição das variáveis escolhidas em cada um dos marketplaces. Em seguida, serão propostas ferramentas necessárias para a extração de dados dos anúncios disponíveis, configuração e execução dos procedimentos de extração. Por último, serão realizadas as devidas formatações aos dados obtidos e discutidos os resultados e métodos.

4. RESULTADOS

4.1. Seleção das plataformas de anúncios

Para o desenvolvimento deste trabalho foram selecionados três marketplaces de anúncios: OLX (disponível na URL <https://www.olx.com.br>), Mercado Livre (disponível na URL <https://www.mercadolivre.com.br>) e Facebook Marketplace (disponível na URL: <http://facebook.com/marketplace>).

A seleção destas plataformas utilizou o critério de visibilidade no cenário nacional (TRENDS, 2021), e expressiva quantidade de volume de dados, bem como especificidades diferentes. O MercadoLivre é um marketplace de anúncios que obrigatoriamente intermedia transações e pagamentos entre compradores e vendedores. Já o OLX permite esta modalidade, mas como uma opção. Ou seja, permite que os anunciantes apenas coloquem dados dos anúncios e informações de contato. Diferente dos anteriores, o Facebook Marketplace permite somente exposição do anúncio, sem possibilidade de intermediação de pagamento. Entretanto, traz forte conexão com os perfis dos usuários que utilizam a rede social.

Estas diferenças entre as plataformas fazem com que a análise proposta neste trabalho se torne mais abrangente do que se analisássemos plataformas com perfis idênticos ou similares.

4.2. Seleção de categoria e variáveis

Para padronizar o conteúdo do dataset e suas variáveis, a análise dos anúncios resultantes das buscas, iremos selecionar uma categoria comum e presente nos três marketplaces escolhidos: a categoria VEÍCULOS. As variáveis selecionadas foram: título do anúncio (normalmente contendo marca, modelo e ano do veículo), valor do veículo, quilometragem, ano, localidade do anúncio, link e outras informações (combustível, câmbio, opcionais etc).

A região geográfica selecionada para os testes foi BELO HORIZONTE – MG, por conter número expressivo de anúncios nesta categoria e por ser uma cidade conhecida pelo autor do trabalho.

4.3. Análise de estruturas e variáveis

Utilizando a ferramenta de desenvolvedor WEB INSPECTOR (disponíveis nos navegadores mais recentes) foi analisada o código-fonte das páginas de anúncios no Mercado Livre, Facebook e OLX. O procedimento a ser feito é ativar o mapeamento conforme apontamento do cursor do mouse sobre o anúncios para, em seguida, identificar os elementos HTML que encapsulam as informações referentes aos anúncios. Nos

três marketplaces analisados, observou-se que os itens de anúncios são disponibilizados utilizando estruturas HTML `<div>` `` e `` de maneira hierárquica, e identificados por classes, conforme ilustrado na figura 1, em um item de busca do Mercado Livre. Em 1), é possível perceber que cada item é um objeto `` (lista) de classe `“ui-search-layout__item”`. Em 2) temos o elemento `` aberto. Em 3) estão todas as informações do item exibido na página. Em 4) o item completo em destaque, contendo preço, título, descrição e outras informações.

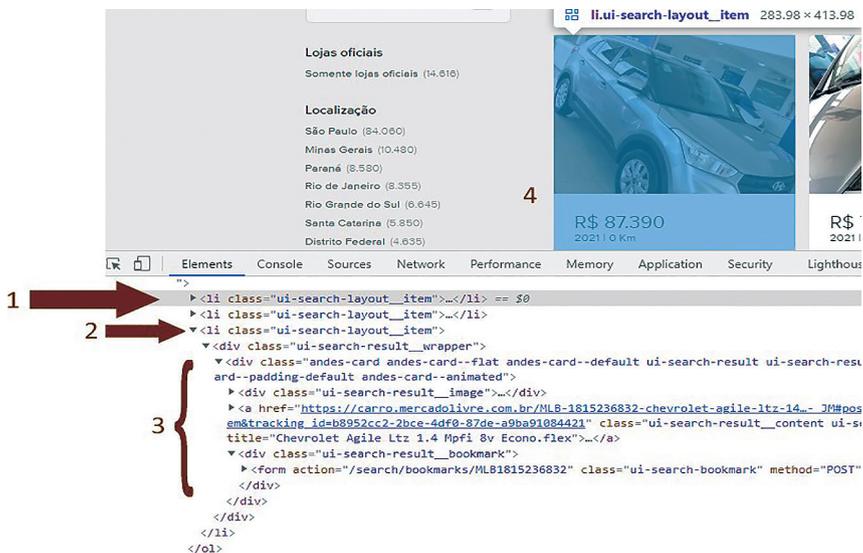


Figura 1. Estrutura HTML do item de busca no Mercado Livre.

Fonte: Dados da pesquisa, 2021.

O Facebook Marketplace e OLX seguem lógica similar. No caso do Facebook, chama a atenção a nomenclatura utilizada na classe (parâmetro *class*) dos elementos, que traz uma chave contendo 13 grupos alfanuméricos de 8 caracteres cada, sendo talvez um mecanismo para dificultar a raspagem automática de dados. Dessa maneira, entende-se que esta análise estrutural deve ser feita sempre antes de se iniciarem os procedimentos de extração.

4.4. Ferramentas de extração

É importante registrar que, das plataformas escolhidas, nenhuma delas disponibilizam APIs ou qualquer tipo de serviço para leitura automática de anúncios disponíveis online. Assim, a solução tecnológica escolhida

foi a utilização de ferramentas capazes de acessar, via URL, dados disponíveis em uma página e, por meio de parâmetros, ler e copiar, de maneira cirúrgica, as informações desejadas. Dentre as diversas técnicas e metodologias para extração de dados, destacamos a definição dada por Thivaharan (*et al.*, 2020), que define web scraper como ferramentas para realizar a chamada ‘raspagem web’ – uma forma de mineração que viabiliza a extração massiva de dados disponíveis em sites web, convertendo-os em informações estruturadas e que podem ser utilizados para análise e ou retroalimentação em outros sistemas.

Para esta tarefa, sugere-se a adoção da linguagem de programação Python, amplamente utilizada em técnicas de mineração de dados e machine learning, bem como algumas bibliotecas e módulos: Scrapy, BeautifulSoup e Selenium Webdriver. O Scrapy é um framework open source desenvolvido em Python para extrair dados de websites (Scrapy, 2021). O BeautifulSoup é uma biblioteca Python capaz de extrair e organizar informações de conteúdos HTML e XML (Crummy, 2020). O Selenium Webdriver é uma biblioteca que pode ser utilizada em Python e oferece como funcionalidade a execução e automação de tarefas em navegadores web e, por meio disso, analisar, extrair e coletar dados.

4.5. Scripts de Extração

Os scripts de execução foram escritos na linguagem PYTHON utilizando o editor VISUAL STUDIO CODE, em um computador com WINDOWS 10 64 bits, processador Intel core i7, 8 GB de RAM, e internet banda larga de 50 Mbps. Considerando os parâmetros definidos na seção 4.3 deste trabalho, as URLs bases para cada um dos marketplaces foram as seguintes:

Facebook Marketplace: <http://facebook.com/marketplace/belohorizonte/carros/>

Mercado Livre: <https://lista.mercadolivre.com.br/veiculos/carros-camionetes-em-belo-horizonte-minas-gerais/>

OLX: <https://mg.olx.com.br/belo-horizonte-e-regiao/autos-e-pecas/carros-vans-e-utilitarios?o=1>.

O script extrator seguiu as seguintes etapas: 1) Definir a URL inicial dos marketplaces; 2) Realizar um request da referida URL; 3) Obter o resultado da URL em HTML ou XML; 4) Tratamento de dados obtidos por meio da biblioteca BeautifulSoup; 5) Identificação das tags HTML e classes dos campos que se deseja extrair; 6) Extrair e armazenar os dados em uma lista em formato .JSON; 7) Voltar ao passo 1 com a URL da próxima página. Um trecho do script, escrito em Python, é exibido na figura 2.

```
page = requests.get(url = url, headers = PARAMS)
soup = BeautifulSoup(page.content, 'lxml')

results = soup.findAll("section", {"class": "ui-search-results ui-search-results--without-disclai

try:
    itens = results[0].findAll("div", class_="ui-search-result_wrapper")
except:
    print("Erro na busca da classe na URL: "+url)

for a in itens:
    try:
        nomeVeiculo = a.findAll("h2", class_="ui-search-item_title ui-search-item_group_eleme
precoVeiculo = a.findAll("span", class_="price-tag-fraction")[0].contents[0]
anoVeiculo = a.findAll("li", class_="ui-search-card-attributes_attribute")[0].contents[
kmVeiculo = a.findAll("li", class_="ui-search-card-attributes__attribute")[1].contents[0
kmVeiculo = kmVeiculo.replace(" Km", "").strip()
urlVeiculo = a.findAll("a", class_="ui-search-result_content ui-search-link")[0]["href"
regiaoVeiculo = a.findAll("span", class_="ui-search-item_group_element ui-search-item_
regiaoVeiculoCidade = regiaoBuscar[regiao]
```

Figura 2. Trecho do script que utiliza requests e BeautifulSoup para coletar dados das páginas web do MercadoLivre e OLX. **Fonte:** Dados da pesquisa, 2021.

O passo 7 descrito anteriormente não se aplica ao Facebook Marketplace, já que os anúncios não possuem paginação, mas sim sistema de rolagem com carregamento dinâmico. Foi necessária a utilização da biblioteca Python SELENIUM WEBDRIVER para execução da rolagem (scroll) automática e captura do HTML exibido no navegador. As demais etapas (identificação e tratamento de tags) seguem da mesma forma. Para isso, as funções para extração dos dados de veículos foram as seguintes, com os devidos parâmetros: buscarFacebook(rolagem=120); buscarDadosMercadoLivre(pages = 10); e buscarDadosOLX(pages = 10)

5. RESULTADOS E DISCUSSÃO

Para fins de testes, foram estabelecidos os seguintes parâmetros: a função de busca do Facebook retorna todos os veículos após 120 rolagens da página. A função de busca do Mercado Livre busca veículos até a 10ª página. A função de dados da OLX também busca todos os veículos contidos entre as páginas 1 a 10. Com tais parâmetros, a extração retornou dos 3 marketplaces 3.110 veículos com os dados especificados na seção 4.3, com tempo de execução de aproximadamente 480 segundos. Embora o script gaste cerca de 1.5 segundos para extração de cada página, o tempo total foi elevado por causa da necessidade de visitar a URL de cada anúncio da OLX para coletar todas as variáveis estipuladas. Após a execução das extrações, as listas foram concatenadas em uma lista principal, conforme ilustrado na figura 3:

```

json = {
    "fonte": "OLX",
    "dia_postagem" : diaPostagem,
    "hora_postagem" : horaPostagem,
    "nomeVeiculo" : nomeVeiculo,
    "kmVeiculo" : kmVeiculo,
    "precoVeiculo": precoVeiculo,
    "regiaoVeiculo":regiaoVeiculo,
    "regiaoVeiculoCidade":regiaoVeiculoCidade,
    "urlVeiculo":urlVeiculo,
    "anoVeiculo":anoVeiculo,
    "transmissaoVeiculo": "n",
    "combustivelVeiculo": "n",
}
listaJson.append(json)

```

Figura 3. Formação da lista com dados recuperados em formato .JSON. Após a captura, a lista atual (variável ‘json’) é acrescida à lista principal (variável ‘listaJson’) por meio da função ‘append’. **Fonte:** Dados da pesquisa, 2021.

Por fim, a variável acumulada com os veículos capturados foi convertida e exportada em formato .XLS por meio da função DATAFRAME da biblioteca PANDAS. A figura 4 mostra trecho da planilha contendo os dados capturados:

	A	B	C	D	E	F	G
1		fonte	a_postagem	ra_postagem	nomeVeiculo	kmVeiculo	precoV
2	0	Facebook	n	n	Ford Fusion	200000	
3	1	Facebook	n	n	NÃO CONSULTAMOS SER	n	
4	2	Facebook	n	n	Volkswagen Golf	175000	
5	3	Facebook	n	n	Peugeot 207	19000	
6	4	Facebook	n	n	FIAT Palio	150000	
7	5	Facebook	n	n	Honda Civic	192000	
8	6	Facebook	n	n	Volkswagen Gol	160000	
9	7	Facebook	n	n	Renault Logan	150000	
10	8	Facebook	n	n	Volkswagen Gol	162000	
11	9	Facebook	n	n	FIAT Palio	165000	
12	10	Facebook	n	n	Chevrolet Classic	123000	
13	11	Facebook	n	n	Volkswagen Jetta	117000	
14	12	Facebook	n	n	Hyundai Azera	121000	
15	13	Facebook	n	n	FIAT Strada	140000	
16	14	Facebook	n	n	Honda Civic	148000	
17	15	Facebook	n	n	Citroën C3	11000	
18	16	Facebook	n	n	FIAT Idea	13000	
19	17	Facebook	n	n	Citroën C4	139000	
20	18	Facebook	n	n	FIAT Uno	20000	
21	19	Facebook	n	n	FIAT Stilo	30 mil milhas	

Figura 4. Trecho da planilha resultante após conversão .json para .xls, contendo dados de 3.110 veículos. **Fonte:** Autoria própria, 2021.

Os scripts completos podem ser acessados na plataforma GITHUB por meio da URL: <https://github.com/edudinizamaral/webscrapedinter>.

6. CONSIDERAÇÕES FINAIS

Foi possível concluir que a combinação da biblioteca requests e BeautifulSoup para tratamento dos dados podem trazer ótimos resultados para a formação de datasets. Em análise preliminar e pelas experiências obtidas nesta pesquisa, e em sintonia com trabalhos atuais existentes na literatura, que os dados disponíveis em marketplaces de anúncios são acessíveis e podem ser recuperados com certa facilidade para compor datasets para as mais diversas aplicações e finalidades em *data science*.

Acredita-se que a metodologia aplicada aqui, semelhante àquelas descritas nos trabalhos relacionados, pode ser adotada também para outros marketplaces e espaços digitais, observando-se a análise das estruturas de dados em HTML, a identificação dos dados que se deseja e a correta programação dos scripts. Observou-se que, se necessário, é possível ampliar o espectro de captura para outras categorias dos marketplaces de anúncios, bem como para outros tipos de dados envolvendo download de imagens, captura de CEP e veiculação com APIs de geolocalização, por exemplo. A extração de anúncios por meio das bibliotecas e funções aqui utilizadas demonstrou ser bastante eficiente, podendo ser utilizadas de maneira a estender os horizontes para a formação de datasets para aplicações diversas.

Fiesler (2020) levanta a questão da legalidade dos dados relacionadas a scrapers, spiders e afins. Por isso, os web scrapers aqui utilizado focaram em dados ‘públicos’ disponíveis para qualquer usuário, em marketplaces de anúncios em URLs abertas, acessíveis ao usuário comum da plataforma. Por isso, preferimos não utilizar spiders que pudessem trazer riscos, acessos não autorizados ou sobrecargas das plataformas estudadas.

É importante salientar, por fim, que a falta de padronização e desorganização em determinados dados inseridos manualmente pelos usuários destas plataformas é também desafiadora: foram encontrados anúncios com telefone em campo título, propaganda em campo de contato etc. Tais inconsistências que aparentemente podem ser tratadas por scripts de manipulação, normalização e filtragem dos dados extraídos.

CONFLITOS DE INTERESSES

Os autores declaram que não há conflitos de interesses em relação ao artigo intitulado “Extração de datasets em marketplaces de anúncios: experiências com Facebook, OLX e Mercado Livre”, submetido a este instrumento científico para apreciação.

Declaram ainda que não possuem quaisquer vínculos com os marketplaces pesquisados. Ainda, que as informações de anúncios foram extraídas das áreas de acesso público dos referidos sites, e que foram utilizados exclusivamente para fins acadêmicos e científicos.

DECLARAÇÃO DE CONTRIBUIÇÃO

Conceitualização, metodologia, software, validação, análise formal, investigação, recursos, curadoria de dados, escrita (original): Eduardo Diniz Amaral

Conceitualização, metodologia, validação, análise formal, escrita (revisão), visualização, supervisão, administração do projeto: Gustavo Medeiros de Araújo.

DECLARAÇÃO DE CONSENTIMENTO DE DADOS

Os scripts obtidos neste trabalho e utilizados para extração de dados, bem como o dataset gerado a partir da extração, foram depositados no repositório GITHUB, disponível por meio do seguinte endereço eletrônico: <https://github.com/edudinizamaral/webscrapedinter>. Ressaltamos que tais dados e codificações não devem ser utilizados para fins comerciais em hipótese alguma. Fica proibido o uso e comercialização dos scripts desenvolvidos neste trabalho para extração de dados em marketplaces. ©

REFERÊNCIAS

- CASTILLO, B. A. V. (2020). *Desarrollo de sistema de análisis de empleabilidad en portales web de empleos*. Escuela Politécnica Nacional: ECUADOR. Disponível em <https://bibdigital.epn.edu.ec/handle/15000/21177>
- CETIC.BR. (2019). *Pesquisa sobre o uso das Tecnologias de Informação e Comunicação nos domicílios brasileiros - TIC Domicílios 2019*. Disponível em: <https://cetic.br/pt/publicacoes/indice/pesquisas/>
- CRUMMY. (2020). *Beautiful Soup Documentation for Python*. Disponível em: <https://www.crummy.com/software/BeautifulSoup/bs4/doc/>
- ECOMMERCEBRASIL. (2021). *E-commerce brasileiro cresce 73,88% em 2020, revela índice MCC-ENET. 2021*. Disponível em <https://www.e-commercebrasil.com.br/noticias/e-commerce-brasileiro-cresce-dezembro/>
- FIESLER, C., BEARD, N., & KEEGAN, B. C. (2020). No Robots, Spiders, or Scrapers: Legal and Ethical Regulation of Data Collection Methods in Social Media Terms of Service. Proceedings of the International AAAI

- Conference on Web and Social Media*, 14(1), 187-196. Retrieved from <https://ojs.aaai.org/index.php/ICWSM/article/view/7290>
- GERHARDT, T. E., & SILVEIRA, D. T. (2009). Métodos de pesquisa. Plageder.
- FATHALLA, A., SALAH, A., LI, K., LI, K., & FRANCESCO, P. (2020). Deep end-to-end learning for price prediction of second-hand items. *Knowledge and Information Systems*, 62(12), 4541-4568. <https://doi.org/10.1007/s10115-020-01495-8>
- OLX. (2021). *Institucional: Quem somos*. 2021. Disponível em: <https://portalolx.olx.com.br/quem-somos/>
- PANDEY, A. *Car's Selling Price Prediction using Random Forest Machine Learning Algorithm*. Março de 2020. 5th International Conference on Next Generation Computing Technologies (NGCT-2019). <http://dx.doi.org/10.2139/ssrn.3702236>
- SCRAPY. (2021). *An open source and collaborative framework for extracting the data you need from websites*. Disponível em: <https://scrapy.org/>
- THIVAHARAN, S., SRIVATSUN, G., & SARATHAMBEKAI, S. (2020, September). A survey on python libraries used for social media content scraping. In *2020 International Conference on Smart Electronics and Communication (ICOSEC)* (pp. 361-366). <https://doi.org/10.1109/ICOSEC49089.2020.9215357>
- TRENDS. (2021). Google Trends. *Pesquisas relacionadas a marketplaces OLX, Facebook e Mercado Livre*.
- WIJAYA, D. R., PARAMITA, N. L. P. S. P., ULUWIYAH, A., RHEZA, M., ZAHARA, A., & PUSPITA, D. R. (2020). Estimating city-level poverty rate based on e-commerce data with machine learning. *Electronic Commerce Research*, 1-27. <http://dx.doi.org/10.1007/s10660-020-09424-1>
- XU, Q., CAI, M., & MACKAY, T. K. (2020). The illegal wildlife digital market: an analysis of Chinese wildlife marketing and sale on Facebook. *Environmental Conservation*, 47(3), 206-212. <http://dx.doi.org/10.1017/S0376892920000235>
- ZAHEER, M. S. *Random Forest Regression on OLX's Dataset*. 2018. Medium.com. Disponível em: <https://medium.com/@msz991/random-forest-regression-on-olxs-dataset-5d108f027257>

