

# ARTIGOS

Submetido 12-09-2023. Aprovado 29-04-2024

Avaliado pelo sistema de revisão duplo-anônimo. Editor Associado: Cham Tat-Huei

Avaliadores/as: Chin Chee Hua , University of Technology Sarawak, Sarawak, Malaysia, mas ele não autorizou a divulgação do seu relatório de revisão por pares. O/A segundo/a avaliador/a não autorizou a divulgação de sua identidade e do relatório de revisão por pares.

Versão traduzida | DOI: <http://dx.doi.org/10.1590/S0034-759020240408x>

## ALGORITMO DE APRENDIZAGEM PROFUNDA COMPREENSIVA PARA COMPREENDER O PAPEL DAS REDES SOCIAIS NA PERCEPÇÃO DO CONSUMIDOR EM RELAÇÃO AO CONSUMO SUSTENTÁVEL

*A comprehensive deep learning algorithm to understand the role of social media in consumer perception towards green consumption*

*Un algoritmo de aprendizaje profundo comprensivo para comprender el papel de las redes sociales en la percepción del consumidor hacia el consumo verde*

Ceren Cubukcu Cerasi\*<sup>1</sup> | [cerencubukcu@gtu.edu.tr](mailto:cerencubukcu@gtu.edu.tr) | ORCID: 0000-0002-9253-2826

Yavuz Selim Balcioğlu<sup>1</sup> | [ysbalcioglu@gtu.edu.tr](mailto:ysbalcioglu@gtu.edu.tr) | ORCID: 0000-0001-7138-2972

Farid Huseynov<sup>1</sup> | [fhuseynov@gtu.edu.tr](mailto:fhuseynov@gtu.edu.tr) | ORCID: 0000-0002-9936-0596

Asli Kilic<sup>1</sup> | [aslikilic@gtu.edu.tr](mailto:aslikilic@gtu.edu.tr) | ORCID: 0000-0002-6621-965X

\*Autor correspondente

<sup>1</sup>Gebze Technical University, Management Information Systems Department, Gebze, Kocaeli, Turquia

### RESUMO

Esta pesquisa propõe um algoritmo abrangente de aprendizado profundo para compreender o papel das redes sociais na percepção do consumidor em relação ao consumo sustentável. Após o surgimento da Covid-19, a sociedade tem apresentado um foco maior sobre a relação entre as pessoas e a natureza. Alcançar os objetivos de desenvolvimento sustentável requer a promoção do consumo verde, o que exige compreender e influenciar as atitudes públicas em relação à sustentabilidade. Enquanto estudos anteriores exploraram o consumo sustentável usando modelos comportamentais e pesquisas, muitas vezes negligenciaram a perspectiva das redes sociais. Utilizando técnicas de aprendizado profundo, este estudo visa analisar dados das redes sociais, incluindo conteúdo de texto e vídeo, para obter insights sobre o comportamento e preferências do consumidor. O estudo envolve a coleta de dados do X (antigo Twitter) e do YouTube, o desenvolvimento de algoritmos de aprendizado profundo para classificação de texto e a criação de um sistema de visualização e relatório. Mais especificamente, este estudo visa analisar o impacto do compartilhamento de informações nas redes sociais nas intenções de compra sustentável da sociedade e propor arquiteturas avançadas para mineração de texto, especificamente o método LDA. Este estudo destaca os insights obtidos da análise do discurso das redes sociais sobre o consumo sustentável. Tendências, atitudes emocionais e engajamento foram examinados usando mineração de texto e análise de sentimento. O estudo revela diferenças específicas da plataforma no sentimento e identifica palavras-chave e frases influentes. A análise também revela respostas emocionais e fatores-chave associados ao discurso sobre consumo sustentável. Os resultados podem apoiar na construção de futuras estratégias para promover o consumo sustentável. O estudo conclui enfatizando a importância de pesquisas adicionais para explorar as discrepâncias entre as plataformas e aproveitar as implicações dessas descobertas para estratégias de consumo sustentável. **Palavras-chave:** aprendizado profundo, redes sociais, sustentabilidade, consumo verde, objetivos de desenvolvimento sustentável.

### ABSTRACT

This research proposes a comprehensive deep-learning algorithm to understand the role of social media in consumer perception of green consumption. After the COVID-19 pandemic, society has shown increased focus on the relationship between people and nature. Achieving sustainable development goals requires promoting green consumption, which necessitates understanding and influencing public attitudes toward sustainability. While previous studies have explored green consumption using behavioral models and surveys, they often overlook the perspective of social media. This study uses deep learning techniques to analyze social media data, including text and video content, to gain insights into consumer behavior and preferences. The study entails collecting data from X (former Twitter) and YouTube, developing deep learning algorithms for text classification, and creating a visualization and reporting system. More specifically, this study aims to analyze the impact of social media information sharing on society's green purchasing intentions and proposes advanced architectures for text mining specifically the LDA method. This study highlights the valuable insights from analyzing social media discourse on green consumption. Trends, emotional attitudes, and engagement were examined using text mining and sentiment analysis. The study reveals platform-specific differences in sentiment and identifies influential keywords and phrases. The analysis also uncovers emotional responses and key factors associated with the discourse on green consumption. The findings can inform future strategies for promoting sustainable consumption. The study concludes by emphasizing the importance of further research to explore the discrepancies between platforms and harness the implications of these findings for sustainable consumption strategies.

**Keywords:** deep learning, social media, sustainability, green consumption, sustainable development goals.

### RESUMEN

Este artículo de investigación propone un algoritmo de aprendizaje profundo integral para comprender el papel de las redes sociales en la percepción del consumidor hacia el consumo verde. Después del brote de la COVID-19, la sociedad ha mostrado un mayor enfoque en la relación entre las personas y la naturaleza. Lograr los objetivos de desarrollo sostenible requiere promover el consumo verde, lo que implica comprender e influir en las actitudes públicas hacia la sostenibilidad. Si bien investigaciones previas han explorado el consumo verde utilizando modelos de comportamiento y encuestas, a menudo pasan por alto la perspectiva de las redes sociales. Aprovechando técnicas de aprendizaje profundo, este estudio tiene como objetivo analizar datos de las redes sociales, incluido contenido de texto y video, para obtener información sobre el comportamiento y las preferencias del consumidor. El estudio implica la recopilación de datos de X (anteriormente Twitter) y YouTube, el desarrollo de algoritmos de aprendizaje profundo para la clasificación de texto y la creación de un sistema de visualización e informes. Más específicamente, este estudio tiene como objetivo analizar el impacto de compartir información en las redes sociales en las intenciones de compra verde de la sociedad y proponer arquitecturas avanzadas para la minería de texto, específicamente el método LDA. Este estudio destaca las valiosas ideas obtenidas al analizar el discurso en las redes sociales sobre el consumo verde. Se examinaron tendencias, actitudes emocionales y participación mediante minería de texto y análisis de sentimiento. El estudio revela diferencias específicas de la plataforma en el sentimiento e identifica palabras clave y frases influentes. El análisis también descubre respuestas emocionales y factores clave asociados con el discurso sobre el consumo verde. Los hallazgos pueden informar estrategias futuras para promover el consumo sostenible. El estudio concluye enfatizando la importancia de investigaciones adicionales para explorar las discrepancias entre plataformas y aprovechar las implicaciones de estos hallazgos para las estrategias de consumo sostenible.

**Palabras clave:** aprendizaje profundo, redes sociales, sostenibilidad, consumo verde, objetivos de desarrollo sostenible

## INTRODUÇÃO

O mundo passa por grandes desafios tais como as mudanças climáticas e a crise de recursos, que têm exigido dos governos uma procura ativa por caminhos de desenvolvimento regenerativo nos quais os humanos e a natureza possam coexistir harmoniosamente (Sharafi, 2021). Especialmente após a pandemia de Covid-19, a sociedade passou a concentrar-se cada vez mais nas pessoas e na natureza (Jian et al., 2020; Sun et al., 2021). O público começou a compreender que, embora a civilização industrial traga conveniência à vida pública, também traz consigo muitos problemas ambientais, especialmente a elevada quantidade de emissões de carbono que impedirão a realização dos Objetivos de Desenvolvimento Sustentável (ODS) propostos pela ONU em 2015 (Yang Y. et al., 2022; Yang et al., 2023).

Para alcançar os ODS não será suficiente apenas os esforços do governo e das empresas. Será necessário também que a sociedade escolha um estilo de vida mais verde, ou seja, um consumo verde em sua vida cotidiana (Akhtar et al., 2021). A chave para promover tal consumo reside na mudança de atitudes públicas. Compreender a tendência da sociedade para o consumo verde contribui para que se encontre o ponto-chave da mudança de atitude dos consumidores e para o aumento da percepção de sua eficácia, influenciando o comportamento.

Os estudos anteriores, em sua maioria, utilizaram modelos e teorias comportamentais para explorar as intenções e comportamentos de compra do consumidor (Costa et al., 2021; Mamun et al., 2018; Zaremohzzabieh et al., 2021). Embora muitos estudiosos tenham examinado a lacuna atitude-comportamento no consumo verde, a maior parte das pesquisas baseia-se em *surveys* ou formulários de entrevistas, e não consideram a perspectiva das redes sociais para avaliar esse tipo de consumo. Com o desenvolvimento do *deep learning* (aprendizagem profunda), tornou-se possível analisar o comportamento do consumidor utilizando técnicas de mineração de texto e LDA (*Latent Dirichlet Allocation* - Alocação Latente de Dirichlet). A análise de dados de texto e comentários de vídeos das redes sociais por algoritmos de aprendizagem profunda pode fornecer informações para governos e executivos das empresas, de forma que compreendam as mudanças na psicologia do consumidor ao longo do tempo. Além disso, essa análise permite o ajuste de estratégias e o desenvolvimento de planos de marketing ecológico com base nas necessidades reais.

Este estudo propõe um algoritmo multimodal de aprendizagem profunda usando mídias sociais para medir a perspectiva das pessoas sobre o consumo verde. Existem poucos estudos na literatura que analisam dados de mídias sociais usando técnicas de mineração de texto e aprendizagem profunda. Os principais propósitos deste estudo são: primeiro, compreender as percepções e atitudes das pessoas sobre o consumo verde através da utilização de dados das redes sociais. Através da análise de dados de texto e comentários em vídeos obtidos de plataformas como o X (antigo Twitter) e YouTube, tenta-se determinar o que o público pensa e como se comporta em relação ao consumo ambientalmente amigável; em segundo, gerar resultados de visualização e relatórios, apresentando os dados obtidos de forma compreensível e eficaz. Isso permite que governos e empresas, ao acessar o presente estudo, moldem as suas estratégias

de consumo verde e planos de marketing com base nas necessidades reais. Ao analisar esses resultados, os tomadores de decisão obtêm informações valiosas sobre as preferências e tendências dos consumidores, permitindo escolhas informadas que promovam práticas sustentáveis. Além disso, a comunicação clara desses dados estimula a colaboração e a compreensão entre os stakeholders envolvidos em iniciativas ambientalmente conscientes, contribuindo para o sucesso global das iniciativas verdes. Em terceiro, examinar como o compartilhamento de informação nas redes sociais afeta as intenções de compra ecológica da sociedade e sua importância em termos de compreensão do tipo de mudanças que ocorreram nos hábitos de consumo e nas atitudes ambientais do público. Finalmente, a pesquisa contribui para o desenvolvimento de arquiteturas novas e eficazes de análise de texto e aplicações LDA. Tais técnicas são ferramentas importantes para melhor compreensão do comportamento e das intenções do consumidor em relação ao consumo verde.

De modo geral, o objetivo deste estudo é medir a percepção do consumo verde por meio de dados estruturados e semiestruturados coletados de postagens do X e comentários de vídeos do YouTube. São discutidas três etapas do processo de detecção do comportamento do consumidor: coleta de dados sobre consumo verde nas plataformas de mídia social X e YouTube, o que reduz a subjetividade trazida anteriormente por pesquisas e entrevistas executadas por meio de amostragem em larga escala; desenvolvimento de aprendizagem profunda em tempo real e algoritmos transformadores para classificação de texto de mídia social; e desenvolvimento do sistema de visualização e relatórios. O resultado apresenta uma análise sobre como o compartilhamento de informações nas mídias sociais afeta as intenções de compra verdes da sociedade e sugere o desenvolvimento de arquiteturas de ponta para análise de texto e aplicações LDA com a capacidade de analisar e detectar intenções e comportamentos de compra do consumidor.

## REVISÃO DA LITERATURA

Os comportamentos de consumo que impõem menos carga ao meio ambiente são chamados de consumo verde (Li, 2020). Na literatura, o comportamento verde do consumidor, a produção verde corporativa e o marketing verde nas mídias sociais são os tópicos de consumo verde mais estudados (Yao et al., 2022). A compreensão do consumidor sobre o consumo verde e as suas perspectivas sobre as questões ambientais são fundamentais para ajudar os consumidores a se conscientizarem do consumo verde e a adotarem comportamentos sustentáveis (Wang, 2021). No entanto, estudos afirmam que (Ahmad & Ariffin, 2018; ElHaffar et al., 2020; Groening et al., 2018; Huang et al., 2022), apesar do elevado nível de conhecimento das questões ambientais, os consumidores apresentam baixa aderência. Portanto, examinar as perspectivas dos consumidores sobre o consumo verde é uma questão importante para compreender melhor o seu comportamento.

De acordo com Sajeewanie et al. (2019), os primeiros e mais fundamentais modelos de comportamento pró-ambiental basearam-se no desenvolvimento linear do conhecimento sobre o meio ambiente. A influência do ambiente na qual o consumidor está inserido é ignorada nos

estudos atuais sobre o seu comportamento verde, e os modelos de regressão quantitativa são favorecidos, uma vez que o problema é tratado como estático, e não dinâmico (Yang M. et al., 2022).

Estudos sobre o consumo verde e seus sub-ramos foram realizados para diferentes setores (Zhao et al., 2020), incluindo política (Jiang & Gao, 2023), logística (Agyabeng-Mensah et al., 2020) e até manufatura (Kluczek, 2017). Utilizando como base tais estudos que consideram fatores como idade, raça, gênero, crença, características pessoais e outras, é possível desenvolver estratégias de segmentação de mercado (Huseynov & Yildirim, 2019) para aumentar o consumo verde. Descobriram ainda que o consumo sustentável varia em decorrência do gênero do consumidor (Sarac, 2022), indicando que os homens de uma determinada faixa etária estão aquém das mulheres no consumo de produtos verdes (Bedard & Reisdorf, 2018).

A maioria dos estudos realizados sobre a perspectiva das pessoas para com o consumo verde obtiveram dados através de questionários (Jain et al., 2020; Jalali & Khalid, 2019; Tang et al., 2020). Biswas (2016), Djafarova e Rushworth (2017), Bedard e Reisdorf (2018), Xie e Madni (2023) usaram pesquisas por questionário para coletar dados e entender a percepção das pessoas sobre o consumo verde, e chegaram a descobertas complementares, observando que a mídia social influencia o comportamento de escolha verde do consumidor de forma favorável (Biswas, 2016). Celebidades do Instagram, blogueiros, proprietários de canais do YouTube e proprietários de perfis “Instafamosos” têm o poder de influenciar o comportamento de compra de meninas (Djafarova & Rushworth, 2017) e da geração *millennials*. A intenção de fazer compras verdes está positivamente correlacionada com as mídias sociais e a influência interpessoal online (Bedard & Reisdorf, 2018). Especialmente para os jovens (Xie & Madni, 2023), as redes sociais são a principal fonte de informação sobre o consumo verde (Ahamad & Ariffin, 2018) e têm a maior influência no comportamento do consumidor (Jain et al., 2020), portanto, torna-se evidente a importância dos estudos sobre o impacto das redes sociais no consumo verde e no comportamento do consumidor. Jalali e Khalid (2019) usaram dados de pesquisas de usuários do Instagram para realizar um estudo sobre os hábitos de compra de produtos verdes, valores e precauções dos influenciadores. Eles conduziram suas pesquisas usando a Teoria dos Usos e Gratificações (TUG) e a Teoria da Ação Racional (TAR) para descobrir como a cognição verde dos consumidores afeta seu comportamento de compra.

Embora a aprendizagem automática esteja sendo utilizada na investigação do consumo verde e seus sub-ramos (Ma & Qiao, 2021; Tanveer et al., 2020), os números ainda não são significativos, uma vez que a maioria dos investigadores utilizou modelos e teorias comportamentais (Costa et al., 2021) e os que utilizaram aprendizado de máquina aplicaram pesquisas e questionários para coleta dos dados (Jain et al., 2020). A mineração de texto é uma ferramenta amplamente utilizada em dados complexos pela produção de resultados rápidos e precisos em análises de dados não estruturados. Então, passou a ser utilizada em diferentes campos de análise, como a de dados de redes sociais (Kunte & Panicker, 2019; Park et al., 2022), de testes de políticas (Jiang & Gao, 2023; Lu & Park, 2022), médica (Balcioglu, 2022; Yazdavar et al., 2020), análise de turismo (Imamah et al., 2020) e muito mais. O uso da mineração de texto previne a subjetividade que

possa surgir nas pesquisas pois permite informações antecipadas. Desse modo, resultados mais confiáveis podem ser obtidos.

A análise de sentimentos é outra área onde a mineração de texto pode ser aplicada. Quando realizada em usuários de redes sociais, essa análise é um dos métodos mais eficazes para examinar as emoções, preferências (Serrano et al., 2021) e comportamentos do usuário. Estudos nesse sentido podem ajudar os tomadores de decisão a desenvolver novas estratégias. Wu et al. (2021) e Huang et al. (2021) analisaram os dados de uma popular plataforma social chinesa chamada Sina Weibo (www.weibo.com) utilizando mineração de texto e análise de sentimentos. Como resultado, observaram que a maioria dos chineses tem pensamentos positivos sobre o consumo verde (Huang et al., 2021), mas as políticas de reciclagem têm um impacto negativo devido a fatores como taxas, penalidades e procedimentos irregulares (Wu et al., 2021). O estudo utilizando mineração de texto de Jiang e Gao (2023) sobre as políticas chinesas confirma a afirmação, demonstrando que o país presta pouca atenção à cadeia de reciclagem.

Brzustewicz e Singh (2021) usaram mineração de texto para categorizar os dados. De acordo com amostras retiradas do X, os temas que os usuários abordam sobre consumo sustentável podem ser categorizados como consumo de alimentos orgânicos, desperdício de alimentos, comida vegana, turismo sustentável, transporte sustentável e consumo de energia sustentável; e a maioria dos usuários tem uma perspectiva positiva sobre o assunto (Brzustewicz & Singh, 2021).

O presente estudo oferece uma nova abordagem do assunto usando conjuntos de dados do X e do YouTube. As publicações e os comentários dos vídeos serão submetidos a análise através de algoritmos de aprendizado de máquina.

## COLETA DE DADOS

Os dados para este estudo foram coletados do X e do YouTube, duas das mais proeminentes plataformas de mídia social. Para o X, foram criados dois conjuntos de dados separados usando sua própria API (*Application Programming Interface* – Interface de Programação de Aplicação). O primeiro conjunto de dados foi construído com base em uma coleção de IDs de postagens adquiridos por meio de pesquisas por palavras-chave relacionadas ao consumo geral. O segundo conjunto de dados consistiu em postagens relativas ao consumo ecológico, coletadas pelo período de um ano. Para recuperar os dados necessários associados a cada ID de postagem, foi empregada a técnica conhecida como mineração de texto. A biblioteca Tweepy Python (biblioteca em linguagem Python de interação com a API do X) foi utilizada para coletar informações sobre cada postagem e condensá-las em um arquivo no formato JSON (*JavaScript Object Notation*). O segundo conjunto de dados foi compilado usando a API X *Streaming*, que fornece acesso em tempo real aos dados de publicações do X. No YouTube, os dados foram extraídos dos comentários de vídeos relacionados ao consumo ambientalmente consciente. Este processo de coleta de dados utilizou a API do YouTube, empregando para a obtenção de dados o cliente API do Google para linguagem Python.

Esta pesquisa foi conduzida como parte de um estudo mais amplo, que analisou postagens ligadas ao consumo ambientalmente consciente. Os dados estudados foram originalmente obtidos em dois estudos independentes que analisaram as discussões no X sobre a ingestão de verduras durante os primeiros meses da pandemia. O primeiro conjunto de dados foi construído com base em uma coleção de IDs de postagens coletadas usando palavras-chave vinculadas ao consumo geral. Empregou-se uma técnica conhecida como mineração de texto, que consiste em compilar todas as informações relevantes que pudessem ser encontradas sobre cada postagem em um arquivo escrito no formato JSON usando a API X Search em cada ID de postagem. Este procedimento coletou informações dos quatro milhões de postagens que ainda estavam acessíveis na plataforma no mês de dezembro de 2020. O segundo conjunto de dados consistiu em cerca de 6 milhões de postagens relacionadas ao consumo ecologicamente correto, que foram coletados entre o mês de dezembro de 2019 a dezembro de 2020, utilizando a API de streaming no X. Ambos os conjuntos de dados passaram então por filtragem, de forma que apenas as datas de sobreposição foram consideradas. Além disso, quaisquer postagens ou comentários repetidos em ambos os conjuntos de dados foram excluídos. O resultado consistiu em 1.227.170 postagens e comentários distintos. As entidades objeto da representação JSON de cada postagem foi analisado para quaisquer URLs (endereços de páginas de internet) que pudessem ser usadas para acessar vídeos do YouTube, e, os comentários dos vídeos encontrados foram coletados e analisados. A Figura 1 apresenta um fluxograma gráfico do processo de seleção e exclusão de dados.

## Explicação do processo de coleta de dados

### Identificação de fontes de dados

#### *X e YouTube*

Essas plataformas foram escolhidas pelo seu destaque como mídias sociais e por seu papel significativo na formação de discurso público, especialmente no que diz respeito às questões ambientais.

### Justificativa para seleção das plataformas

Essas plataformas foram selecionadas pela diversidade em sua base de usuários e pelos diferentes tipos de envolvimento que promovem (postagens, repostagens, comentários).

### Métodos de coleta de dados

#### *X*

Uso da API: Utilizamos a API do X para coletar postagens e repostagens. Isso envolveu o uso de palavras-chave específicas relacionadas ao consumo verde.

Conjuntos de dados: Foram criados dois conjuntos de dados independentes – um de postagens relacionadas ao consumo geral e outro com foco no consumo ambientalmente sustentável.

API de streaming: para coleta de dados em tempo real, empregamos a API de streaming do X.

## YouTube

API para comentários: a API do YouTube foi utilizada para extrair comentários de vídeos relacionados ao consumo ambientalmente consciente.

Processo de Filtragem: Teve como foco vídeos e comentários diretamente relevantes para o tema consumo verde.

## Parâmetros de coleta de dados

### *Período*

Examinamos um período significativo para captar a evolução do discurso, destacando particularmente as mudanças durante os primeiros meses da pandemia de Covid-19.

Coleta baseada em palavras-chave: Para filtragem e coleta de dados, foram utilizadas palavras-chave e hashtags condizentes com o consumo verde.

Considerações demográficas e de idioma: caso aplicável, foi mencionado se o estudo se limitou a idiomas ou dados demográficos específicos.

## Processamento de dados

### *Técnicas de mineração de texto*

Empregamos mineração de texto para analisar os dados coletados, garantindo a extração de informações relevantes.

## Limpeza de dados

Foram aplicadas etapas de pré-processamento para o refino do conjunto de dados removendo duplicatas e conteúdo irrelevante.

## Considerações éticas

### *Dados públicos*

Garantimos que todos os dados coletados estivessem disponíveis publicamente, respeitando as diretrizes éticas e os termos de serviço das plataformas.

## Anonimização e privacidade

Todas as medidas para anonimizar os dados e respeitar a privacidade dos usuários foram tomadas.

## Justificativa para abordagem de coleta de dados

### *Alinhamento com objetivos de pesquisa*

O objetivo principal deste estudo foi compreender o papel das mídias sociais na formação da percepção do consumidor em relação ao consumo verde. Isto envolveu a análise do discurso público, tendências e sentimentos em plataformas onde estas discussões prevalecem.

## Seleção de plataformas de mídia social (X e YouTube)

X

A plataforma foi selecionada por sua natureza dinâmica e em tempo real, oferecendo insights imediatos sobre reações e discussões públicas. O formato do X incentiva a expressão de opiniões concisas e focadas, tornando-o uma fonte rica para análise de sentimentos e identificação de tendências.

*YouTube*

O YouTube foi selecionado por seu papel como plataforma para discussões mais detalhadas, frequentemente encontradas em comentários dos vídeos relacionados ao consumo verde. Os comentários do YouTube proporcionam profundidade à compreensão das percepções do consumidor, complementando os dados concisos obtidos do X.

## Coleta de dados via APIs

*Uso da API do X*

A API do X permitiu reunir uma vasta quantidade de dados relacionados ao consumo verde, incluindo postagens, repostagens e hashtags, garantindo uma visão abrangente do tema ao longo do tempo.

*Uso da API do YouTube*

O uso da API do YouTube permitiu a extração de comentários de vídeos relevantes, fornecendo insights sobre opiniões e perspectivas mais detalhadas dos consumidores sobre o consumo verde.

## Técnicas de coleta de dados

*Pesquisas de palavras-chave*

Foram aplicados termos específicos relacionados ao consumo verde para filtragem dos dados, garantindo que fossem relevantes ao foco deste estudo.

*Mineração de texto e análise de sentimentos*

Essenciais para a análise de grandes volumes de dados de texto, essas técnicas permitiram identificar temas, sentimentos e tendências principais no discurso em torno do consumo verde.

*Escopo temporal da coleta de dados*

O período de coleta foi estrategicamente escolhido para cobrir fases significativas do discurso público, especialmente durante os primeiros meses da pandemia da Covid-19, um período em que a consciência ambiental teve um notável aumento.

## Volume e diversidade de dados

Para captura de uma ampla gama de pontos de vista e discussões, garantimos a coleta de uma amostra de grande volume e diversidade. Esta abordagem abrangente permitiu análises robustas e resultados mais generalizáveis.

## Abrangência

O estudo é rigoroso na abordagem para capturar uma amostra representativa do discurso sobre o consumo verde.

## MÉTODOS

Esta pesquisa empregou uma abordagem computacional baseada em dados usando Python, uma linguagem de programação versátil e amplamente utilizada em ciência de dados e pesquisas computacionais em ciências sociais. O estudo envolveu duas etapas principais: coleta e análise de dados, onde cada uma delas utilizou bibliotecas e técnicas Python específicas para ser realizada.

## Perspectiva teórica

### Relevância para os objetivos da pesquisa

O tamanho da amostra foi dimensionado para garantir uma representação abrangente do discurso sobre o consumo verde no X e no YouTube. Dada a grande quantidade de dados gerados diariamente nestas plataformas, tornou-se necessária a coleta de uma amostra de maior porte, capturando uma ampla gama de opiniões e sentimentos.

## Alinhamento com o desenho do estudo

Este estudo visa analisar tendências, atitudes emocionais e envolvimento no decorrer de um período significativo (cobrindo os primeiros meses da pandemia de Covid-19). Este período assistiu a um aumento do interesse público pelas questões ambientais, exigindo a captação de uma amostra de tamanho substancial que refletisse com precisão as mudanças no discurso.

## Perspectiva estatística

### Disponibilidade de dados

O volume de dados disponíveis no X e no YouTube permitiu a coleta de uma amostra de grande porte, conferindo robustez à análise.

## Método de amostragem

A utilização de APIs do X e do YouTube possibilitou a coleta de um amplo e variado conjunto de dados, incluindo repostagens, comentários e palavras-chave relacionadas ao consumo verde.

## Diversidade de dados

Para garantir uma visão holística, os dados foram recolhidos no período de um ano, captando várias fases e picos da discussão, o que é fundamental para a compreensão de tendências e padrões de longo prazo.

## Poder estatístico

Uma amostra de tamanho substancial aumenta o poder estatístico do estudo, reduzindo a probabilidade de erros do Tipo II (não detectar a presença de um efeito) e aumentando a confiabilidade dos resultados.

## Generalização

Um tamanho substancial de uma amostra contribui para a generalização dos resultados do estudo, permitindo uma extrapolação mais precisa dos resultados para uma população ampla de usuários de redes sociais interessados no consumo verde.

## Abordando a fundamentação teórica e estatística

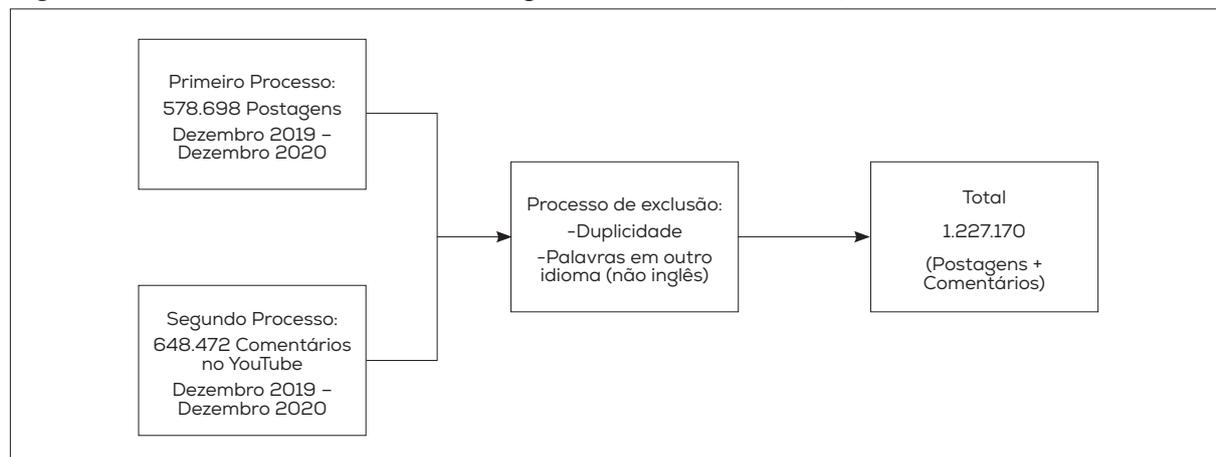
### Conformidade com as normas de amostragem

O tamanho da amostra segue práticas comuns de pesquisas computacionais de ciências sociais, onde grandes conjuntos de dados são frequentemente analisados para discernir padrões na comunicação digital.

## Garantindo a riqueza dos dados

A riqueza e profundidade de dados é garantida pelo tamanho da amostra coletado, permitindo uma análise diferenciada das palavras-chave e tendências de sentimentos.

Figura 1. Processo de combinação, filtragem e depuração do conjunto de dados



## Pré-processamento de dados

Antes de prosseguir com a análise, os dados foram depurados e pré-processados. Os dois conjuntos de dados foram filtrados de forma que apenas as datas de sobreposição fossem incluídas e quaisquer postagens ou comentários duplicados em ambos os conjuntos foram removidos, criando um conjunto final de postagens e comentários únicos. A biblioteca Pandas do Python foi utilizada para tarefas de depuração e pré-processamento de dados.

## Coleta robusta de dados

### Amostra de grande porte

Reunimos extensos conjuntos de dados, incluindo milhões de postagens e comentários do YouTube, garantindo um conjunto substancial e variado de opiniões públicas.

## Considerações sobre prazos

A coleta de dados abrangeu um período significativo, captando a evolução do discurso ao longo do tempo e garantindo que fossem incluídas tendências transitórias e sustentadas.

## Recuperação de dados focada

### Coleta baseada em palavras-chave

Selecionamos cuidadosamente palavras-chave relacionadas ao consumo verde, garantindo que os dados fossem precisamente direcionados e relevantes para os objetivos da pesquisa.

## Utilização de API

Empregamos recursos avançados de API para extrair e compilar dados sistematicamente, garantindo eficiência e abrangência em nosso processo de coleta.

## Processamento de dados aprofundado

### Depuração rigorosa de dados

Implementamos procedimentos completos de depuração e refino de dados, incluindo a remoção de dados duplicados e entradas irrelevantes.

## Mineração de texto e análise de sentimentos

Aplicamos técnicas sofisticadas de mineração de texto e ferramentas de análise de sentimentos para refinar e interpretar grandes quantidades de dados, extraindo temas, sentimentos e padrões importantes.

## Rigor analítico

### Técnicas analíticas diversas

Nossa análise não se limitou a estatísticas descritivas básicas, estendendo-se a técnicas avançadas, como LDA, para modelagem de tópicos e análise de sentimentos, proporcionando uma compreensão mais profunda do discurso público.

## Análise de dados

A análise do conjunto de dados foi realizada utilizando vários métodos em linguagem Python. Para visualizar a frequência de postagens e comentários do YouTube ao longo do tempo, usamos o Matplotlib, uma biblioteca de visualização de dados em Python amplamente utilizada. Foi executada uma análise de frequência de palavras em postagens e comentários para compreender os temas relacionados ao consumo verde mais discutidos. As bibliotecas Python NLTK (*Natural Language Toolkit*) e *Scikit-learn* foram aplicadas para tarefas de processamento de linguagem natural, incluindo tokenização e remoção de palavras de parada. A análise de sentimentos foi realizada tanto nas postagens quanto nos comentários do YouTube a fim de compreender o tom emocional subjacente às discussões. O processo de análise de sentimentos envolveu pré-processamento de texto, tokenização e pontuação de sentimentos usando léxicos. Usamos a biblioteca Python TextBlob, que fornece uma API simples para aplicação em tarefas comuns de processamento de linguagem natural, como marcação de classes gramaticais, extração de

sintagmas nominais e análise de sentimentos. Para uma análise de texto mais avançada, foi utilizado o método LDA, um tipo de modelo probabilístico de tópicos que extrai os tópicos subjacentes das postagens e comentários, implementada utilizando a biblioteca Python Gensim.

Neste estudo, centrado no papel das redes sociais na formação da percepção do consumidor em relação ao consumo verde, utilizamos um conjunto de ferramentas de software e bibliotecas automatizadas. Essas ferramentas foram fundamentais para depurar com eficiência e precisão o grande volume de dados coletados do X e do YouTube.

A linguagem Python serviu como a espinha dorsal do nosso processo de depuração de dados e é reconhecida por sua robustez e versatilidade no tratamento de grandes conjuntos de dados. Ela forneceu uma plataforma flexível para integração de várias bibliotecas de depuração de dados e scripts personalizados.

A biblioteca Pandas é utilizada por seus poderosos recursos de manipulação de dados. Nos permitiu realizar tarefas como a filtragem de dados irrelevantes, remoção de duplicatas e reestruturação dos conjuntos de dados, facilitando a análise. Possibilitou o manuseio facilitado e transformação de grandes conjuntos de dados para um formato adequado.

O conjunto de bibliotecas e programas NLTK foi empregado para pré-processar dados textuais de postagens em mídias sociais. Foi essencial para tarefas como tokenização, remoção de palavras de parada e normalização de texto, que são obrigatórias na depuração e preparo de dados de texto para análise de sentimentos. Expressões regulares (Regex) foram usadas para identificar e remover padrões de texto indesejados, tal como URLs (endereços de páginas da internet), caracteres especiais e símbolos não padrão. Possibilitou o refino dos dados textuais para inclusão apenas de conteúdo relevante.

A *Scikit-learn*, biblioteca de aprendizado de máquina, foi utilizada para implementar algumas das técnicas mais avançadas de depuração e pré-processamento de dados, especialmente na preparação do conjunto de dados para análise de sentimentos e modelagem de tópicos. *Tweepy* (para dados do X), uma biblioteca Python para acesso à API do X, foi fundamental na fase inicial de coleta de dados e também desempenhou um importante papel na depuração preliminar dos dados do X.

cliente API do Google (para dados do YouTube) facilitou a extração eficiente e a depuração inicial dos dados do YouTube, garantindo a coleta de comentários relevantes para análise. O Jupyter Notebooks (organização sem fins lucrativos criada para “desenvolver software de código aberto, padrões abertos e serviços para computação interativa em dezenas de linguagens de programação”) forneceu um ambiente interativo para codificação, depuração e análise preliminar de dados, permitindo um processo documentado e simplificado.

Por fim, usamos o WordCloud, uma biblioteca Python para nuvens de palavras. A nuvem de palavras foi gerada a partir dos termos mais utilizados nas postagens relacionadas ao consumo verde, oferecendo uma representação visual dos temas centrais do discurso. Em resumo, através da integração de várias bibliotecas em linguagem Python e técnicas de análise de dados, este estudo oferece uma exploração aprofundada do discurso público ocorrido no X e no YouTube em torno do consumo verde.

Reconhecemos que a variância do método comum (VMC) é uma preocupação crítica na investigação, particularmente em estudos que utilizam dados autorrelatados ou que dependem

de métodos de coleta de dados de fonte única. A VMC pode levar a correlações inflacionadas ou espúrias entre variáveis, comprometendo potencialmente a validade dos resultados da investigação.

Embora o nosso estudo tenha analisado principalmente dados de plataformas de redes sociais e não se baseou fortemente em medidas autorrelatadas, compreendemos a importância de mitigar a VMC em todas as formas de investigação.

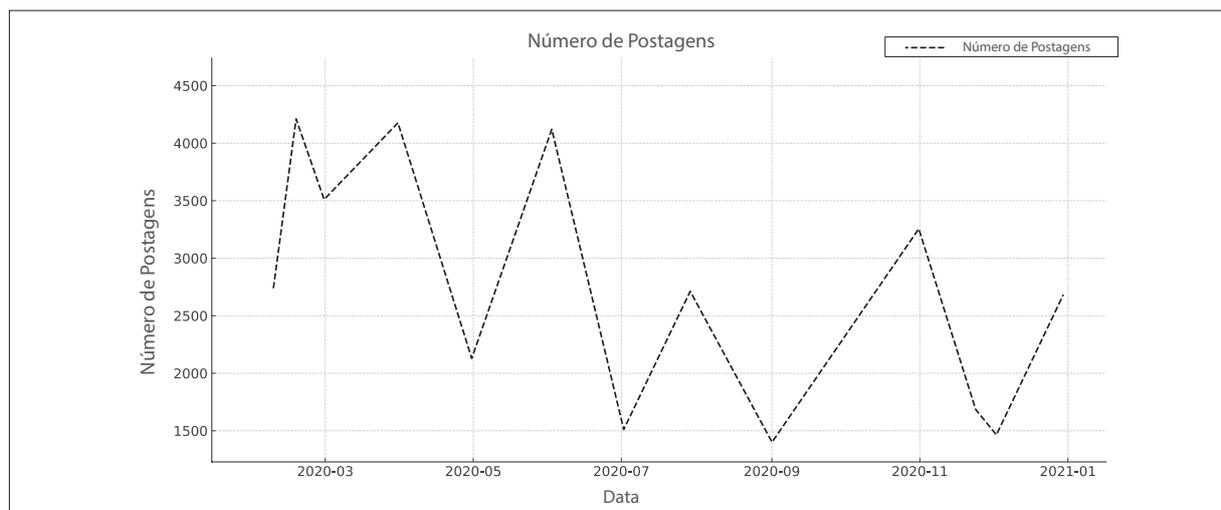
Para estudos que envolvem dados autorreferenciados, defendemos o uso de diversos métodos e fontes de coleta de dados para reduzir o risco da VMC. Isto inclui a triangulação de dados com outras fontes, o emprego de abordagens de métodos mistos e a integração de insights qualitativos. São estratégias metodológicas para abordar a VMC:

- Anonimato e Confidencialidade: garantir aos entrevistados o anonimato e a confidencialidade para reduzir os vieses de conveniência social.
- Separação temporal: coletar variáveis preditoras e de resultado em momentos diferentes para reduzir a VMC.
- Separação metodológica: utilizar diferentes métodos para medir diferentes construtos, reduzindo assim a probabilidade de vieses associados aos métodos.

## RESULTADOS E DISCUSSÃO

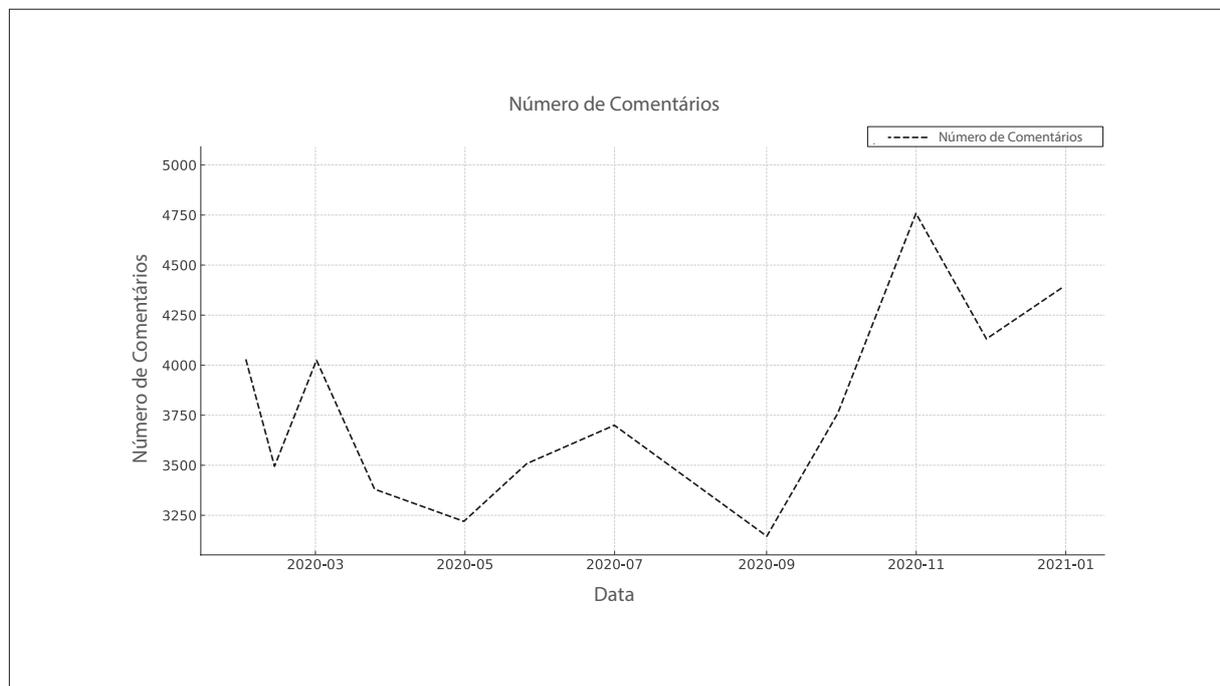
A Figura 2 é uma ilustração da frequência de postagens e repostagens relevantes sobre o consumo verde. Mostra que a linha de tendência subiu de janeiro a fevereiro, quando atingiu o seu primeiro máximo. A partir de janeiro, foram observados os maiores picos de intensidade, seguidos por um segundo pico em abril e um terceiro pico em junho. Em novembro ocorreu o aparecimento do quarto pico menos intenso. Esses dados demonstram a intensidade das atividades de discussão sobre o consumo verde no X já no primeiro período, iniciando em janeiro e atingindo o ponto máximo no final do mesmo mês. Isto indica que a consciência pública aumentou neste período.

Figura 2. Número de Postagens



A frequência de comentários no YouTube relativos ao consumo verde é mostrada na Figura 3, Figura 3, onde a linha de tendência subiu de forma constante de janeiro até novembro, quando atingiu o seu ponto mais alto. Em janeiro já foi observado o primeiro pico de intensidade, seguido por um segundo ocorrido em março, e um terceiro pico ocorrido em julho. Em novembro ocorreu o aparecimento do quarto pico, alcançando o ponto máximo. Esses dados demonstram a conscientização do público, pois representam a intensidade das atividades de discussão sobre o consumo ambientalmente responsável no YouTube começando em janeiro e atingindo seu pico no final do ano.

**Figura 3.** Número de comentários nos vídeos do YouTube



A Figura 4 mostra o número de postagens que incluem a palavra-chave “sustentabilidade”. A trajetória ascendente do termo “sustentabilidade” atingiu seu ponto mais alto em março. Entre abril e setembro, a linha apresentou tendência decrescente. A Figura 5 ilustra vários níveis de utilização do X durante o período. A trajetória ascendente deste gráfico para o uso do termo “consumo” atingiu seu ponto máximo em julho. Entre março e maio, houve uma mudança descendente. Essa mudança pode ter conexão com o anúncio feito pelas autoridades chinesas sobre a descoberta de um novo vírus pertencente à família do coronavírus à medida que o número de casos confirmados da doença crescia. Foi identificado pela primeira vez como 2019-nCoV, mas seu nome foi posteriormente alterado para Covid-19. Isto marcou o início da pandemia de Covid-19, provocando um aumento significativo no número de postagens que usaram o termo surto.

Figura 4. Frequência de citações da palavra-chave: Sustentabilidade em postagens do X

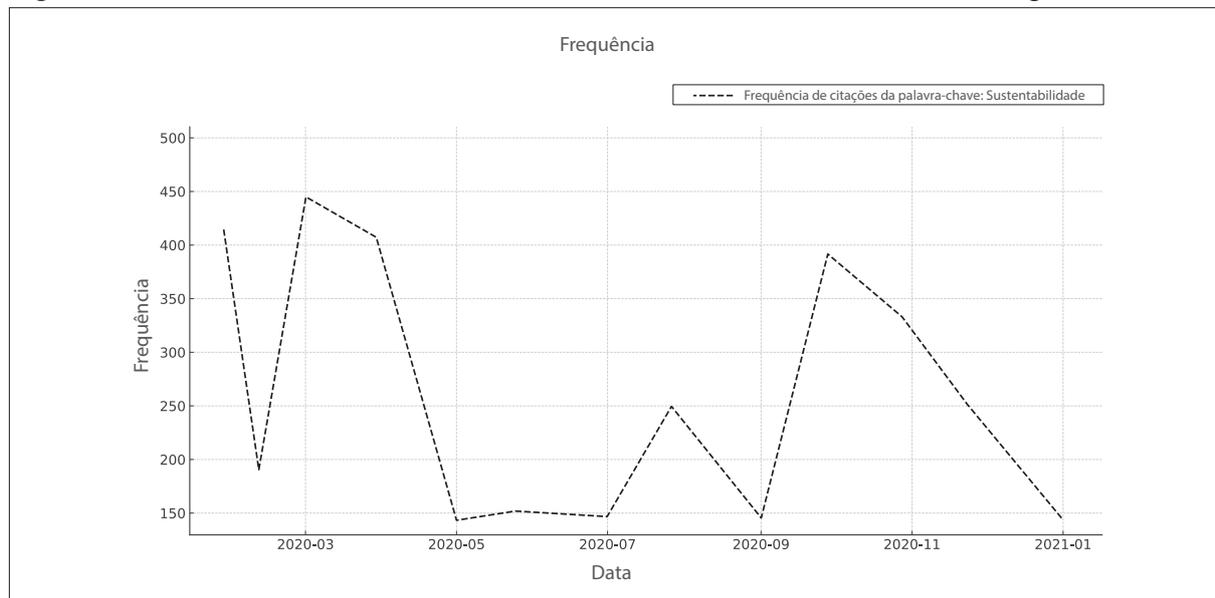


Figura 5. Frequência de citações da palavra-chave: Consumo em postagens do X

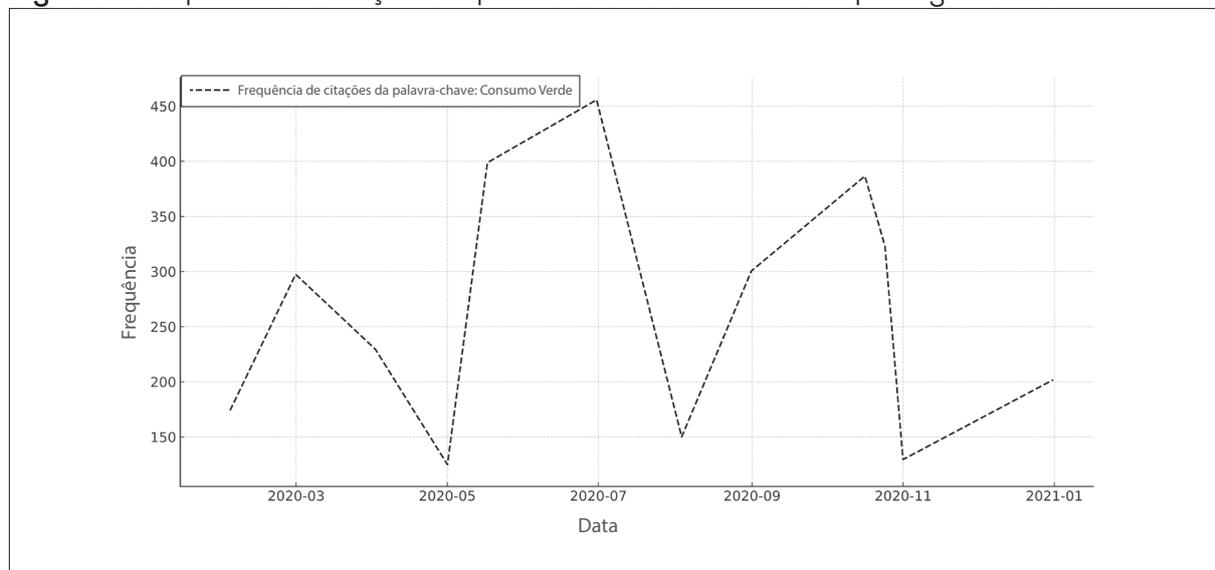
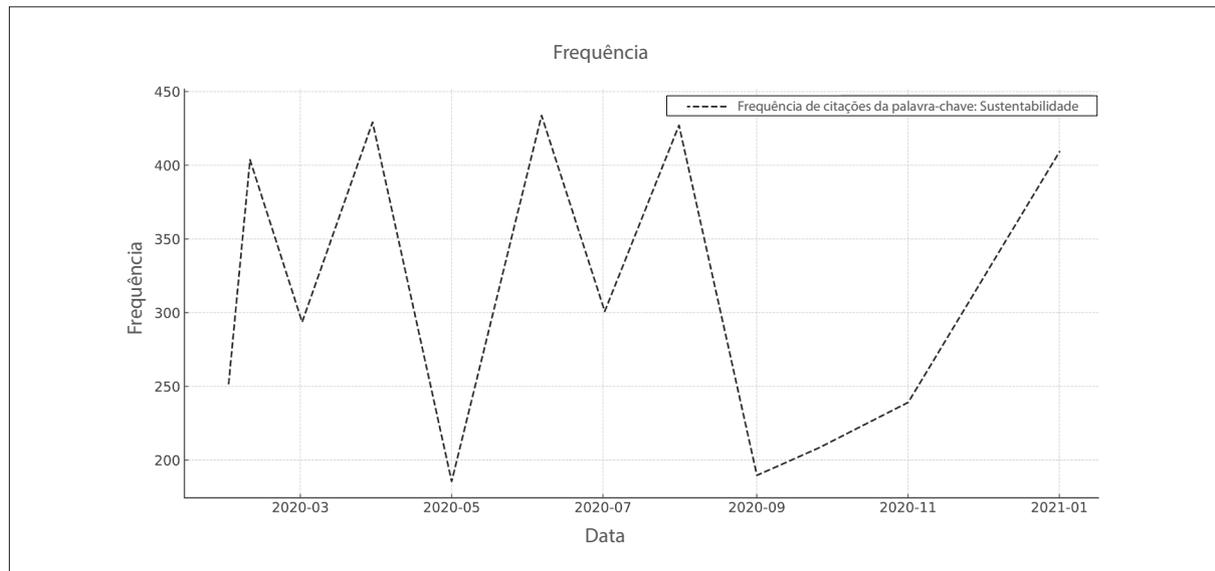
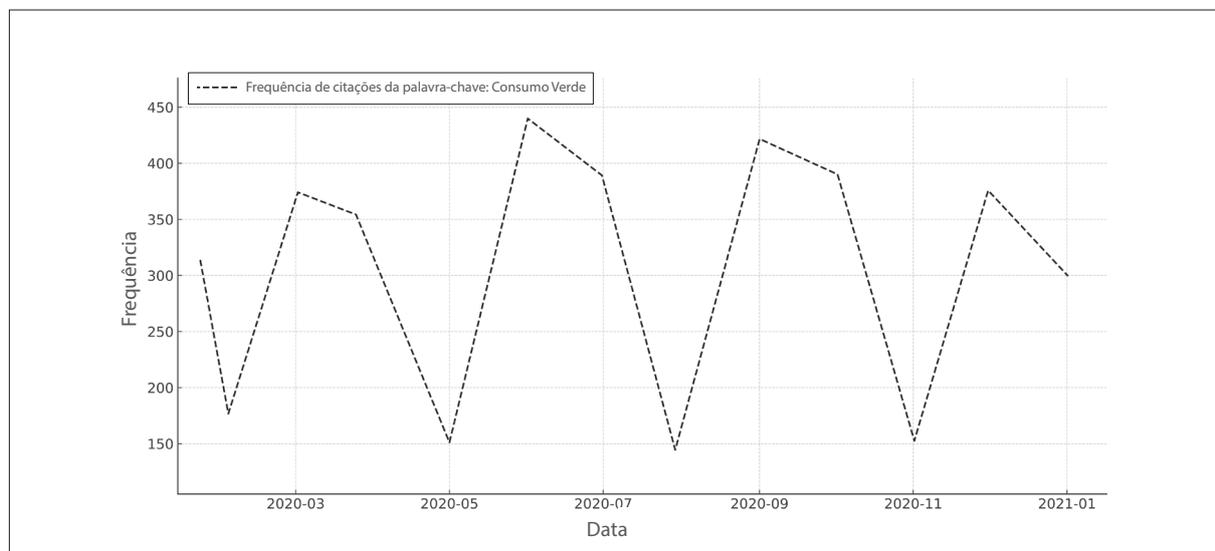


Figura 6 aponta a frequência de citações da palavra-chave “sustentabilidade” na seção de comentários do YouTube. A trajetória ascendente do termo “sustentabilidade” atingiu seus pontos máximos em abril e junho de 2020. Entre agosto e setembro, a linha de tendência tornou-se descendente. Na Figura 7 é apresentada a frequência de citações da palavra-chave “consumo verde” nos comentários dos vídeos do YouTube. A trajetória da linha atingiu seu ponto mais alto no mês de junho e uma tendência decrescente se apresentou para os meses de março, maio, agosto e setembro.

**Figura 6.** Frequência de citações da palavra-chave: Sustentabilidade em comentários de vídeos do YouTube

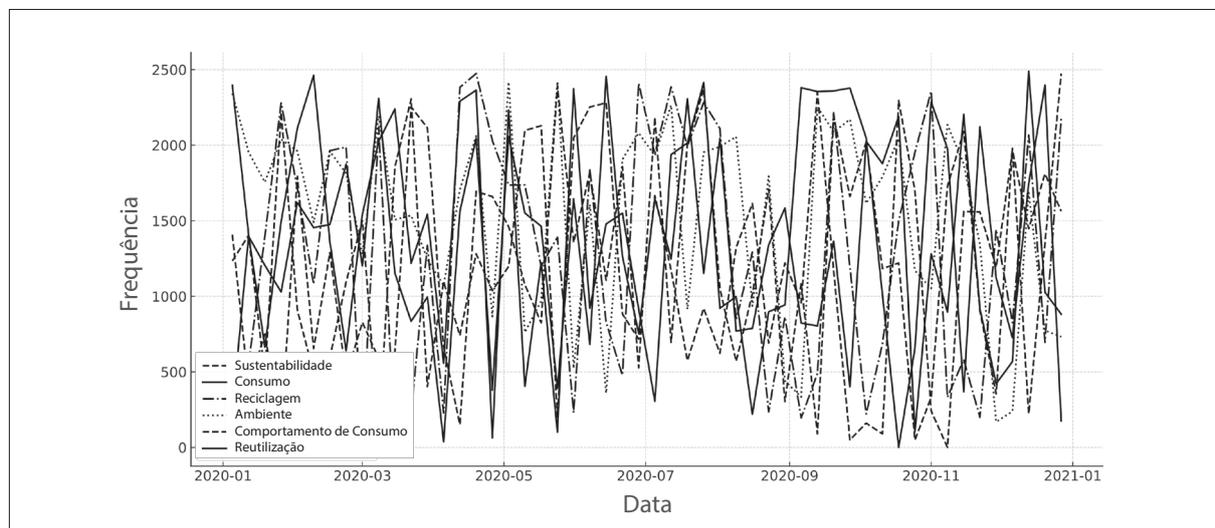


**Figura 7.** Frequência de citações da palavra-chave: Consumo Verde



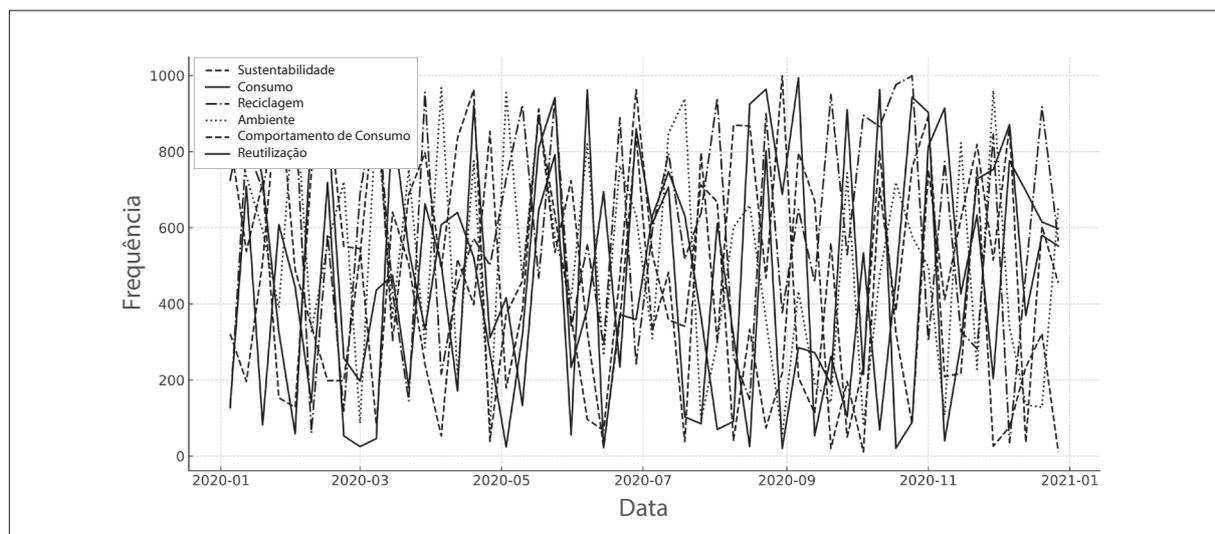
A Figura 8 exibe linhas de tendência que indicam a frequência de citações das seis principais palavras-chave relativas ao consumo verde no X. Essa frequência no aparecimento destas palavras pode representar as perspectivas e preocupações dos usuários do X sobre a poluição. Sustentabilidade e Consumo são os dois aspectos mais importantes do consumo ambientalmente consciente. Como foram utilizados apenas os seis componentes mais importantes na construção do gráfico, outros fatores foram omitidos da Figura 8. Verificamos que as maiores quedas na frequência de citação entre os seis fatores ocorreram em março, porém, neste mesmo mês, dois fatores apresentaram maior aumento: Ambiente e Comportamento de Consumo.

**Figura 8.** Linhas de tendência que indicam a frequência de citações de palavras relativas aos seis principais fatores de consumo verde no X



A frequência de palavras para os principais fatores relativos ao consumo verde que ocorreram nos comentários de vídeos do YouTube são mostradas como linhas de tendência na Figura 9, apresentada abaixo. Esta frequência de palavras pode ser uma representação dos pontos de vista e preocupações das pessoas que usam o YouTube sobre poluição. O consumo ecologicamente sensível deve ter como principal ênfase a preservação dos recursos naturais e a redução do consumo global. A Figura 9 não inclui quaisquer outras considerações, uma vez que o gráfico foi construído utilizando apenas os seis elementos considerados mais significativos. Examinando a Figura 9, observamos que a maior diferença ocorre no mês de junho, quando um fator dentre todos os seis apresentou queda máxima, enquanto todos os restantes apresentaram uma tendência ascendente.

**Figura 9.** Linhas de tendência que indicam a frequência de citações de palavras relativas aos seis principais fatores de consumo verde no YouTube



A nuvem de palavras apresentada foi criada a partir dos termos mais utilizados neste estudo e ajudaram os pesquisadores a compreender melhor as postagens dos usuários do X relacionados ao consumo ambientalmente consciente. Como pode ser visto na Figura 10, os termos que apareceram com mais frequência estavam ligados ao conceito de ser sustentável. Os termos “conservação”, “verde”, “clima”, “limpo”, “resíduos” e “energia” são as palavras secundárias. Os termos “sustentabilidade” e “consumo verde” foram utilizados para que possamos investigar uma variedade de pontos de vista sobre as diversas formas de consumo ecologicamente corretas, ao mesmo tempo em que consideramos a frequência de determinadas palavras-chave em uma determinada pesquisa.

**Figura 10.** Nuvem de palavras apresentando as palavras-chave (em inglês) relacionadas ao consumo verde citadas frequentemente em postagens



A análise de sentimentos aprimorou os resultados ao permitir distinguir claramente entre assuntos negativos e positivos sobre o consumo ambientalmente consciente. De acordo com os resultados do estudo das emoções das postagens, apenas 39,62% deles incluíam pensamentos otimistas, enquanto 60,38% continham sentimentos negativos. Isto indica que os usuários do X têm uma atitude desfavorável em relação às compras ambientalmente conscientes. A Figura 11 ilustra que ao longo do mês de janeiro, o sentimento negativo aumentou mais do que o sentimento positivo. Após o mês de fevereiro, houve um incremento da emoção positiva maior do que do sentimento negativo. De fevereiro a abril, a emoção positiva pareceu ter uma influência mais significativa do que a negativa.

De acordo com os resultados da análise das emoções dos comentários do YouTube, 57,19% deles incluíam bons pensamentos, enquanto 42,81% continham sentimentos negativos. Isto sugere que os espectadores do YouTube têm uma atitude favorável em relação às compras ambientalmente conscientes. De acordo com a Figura 12, os níveis mais elevados de sentimento negativo foram registrados no final do mês de fevereiro até o início de março, depois em abril, junho e setembro, enquanto os níveis mais elevados de sentimento bom foram registrados em janeiro e final de fevereiro, ao longo de março, em abril, julho, outubro, novembro e em dezembro.

Figura 11. Análise de sentimentos de postagens negativas e positivas relacionados ao consumo

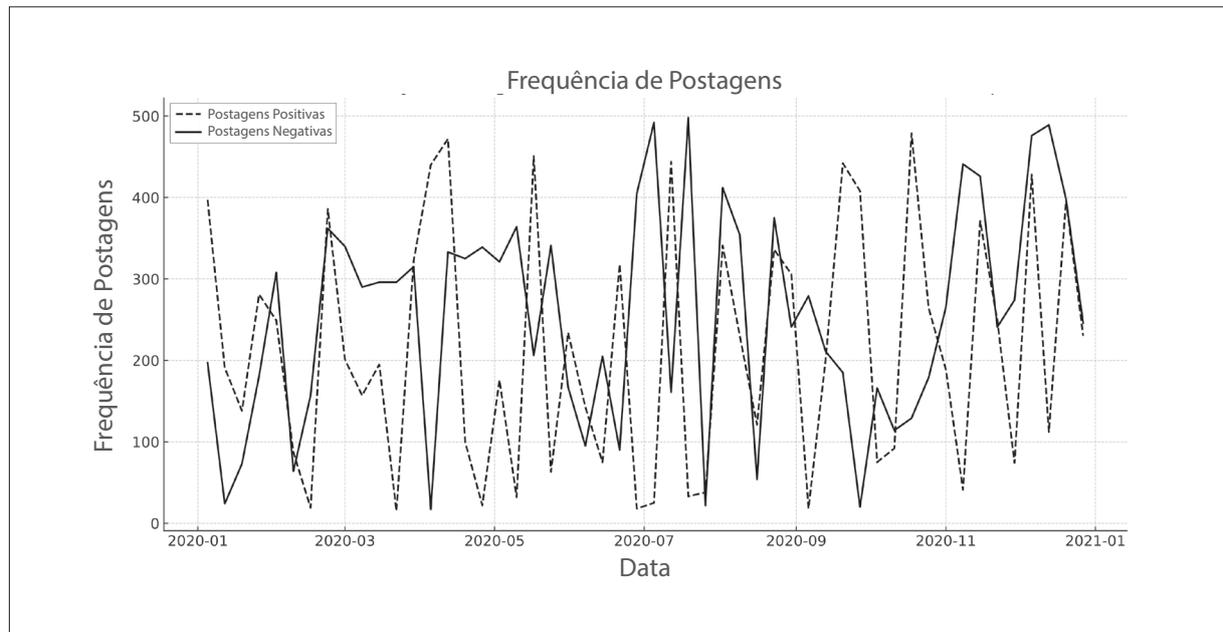
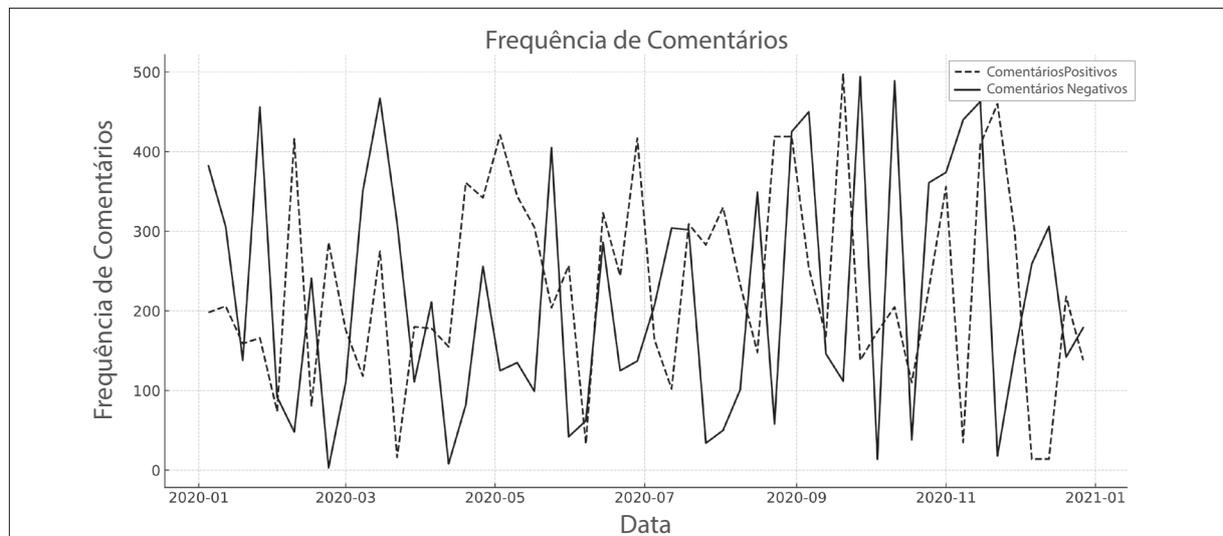
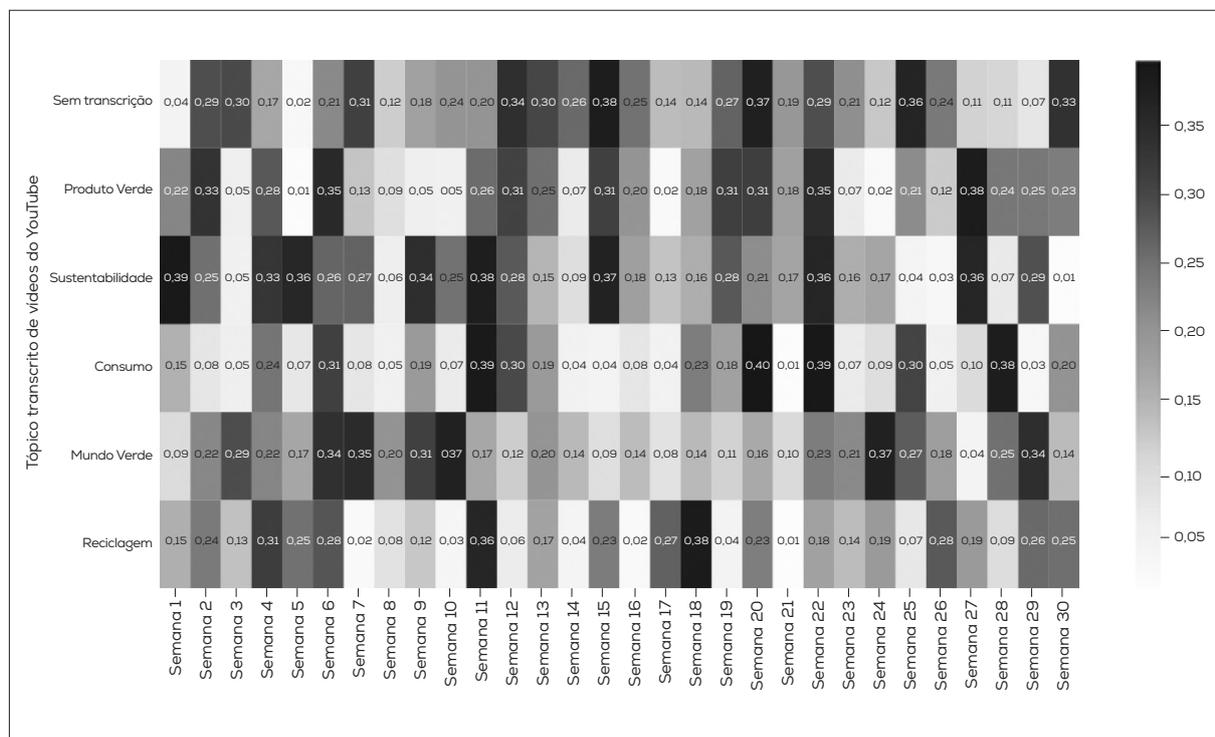


Figura 12. Análise de sentimentos de comentários negativos e positivos relacionados ao consumo



A Figura 13 ilustra, através de um mapa de calor, a dinâmica temporal dos diversos temas dividida em semanas. Ela revela que houve mudanças perceptíveis na frequência relativa de alguns temas ao longo do tempo. Em particular, os vídeos que apoiavam o argumento de consumo verde dominaram a primeira semana, mas foram rapidamente substituídos como quadro dominante após a segunda semana, tornando-se menos frequentes nas semanas seguintes, mas não desaparecendo completamente e figurando como quinto na classificação geral.

Figura 13. Mapa de calor temporal dos tópicos transcritos do YouTube usados semanalmente

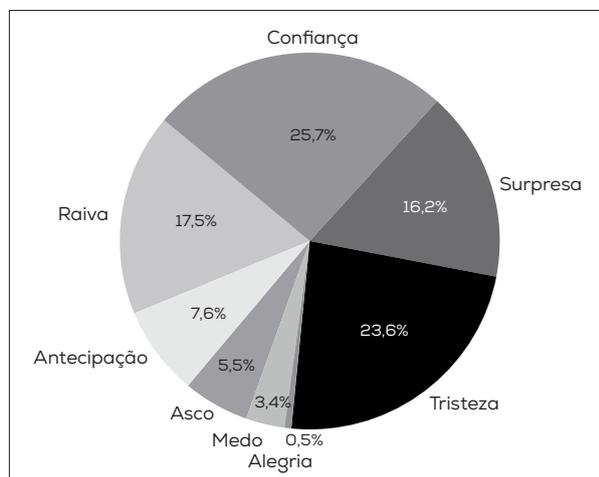


Nota: uma célula mais escura indica uma proporção maior de vídeos abordando o tópico naquela semana.

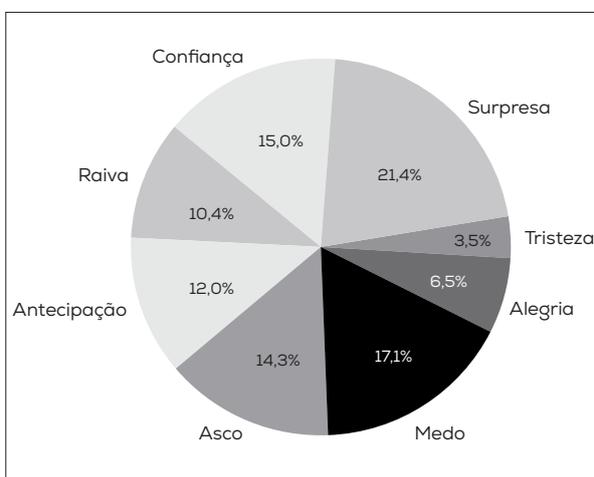
Depois de analisar as postagens usando o método LDA para determinar o seu quociente emocional, os pesquisadores descobriram que mais da metade das postagens em todo o mundo podem ser compreendidas entre três sentimentos: raiva, confiança ou tristeza. Como pode ser visto na Figura 14, uma parcela significativa do total de postagens analisados continha referências ao sentimento de confiança. O sentimento subsequente, tristeza, sugere que os indivíduos ansiavam pela reabilitação ou por receber respostas dos profissionais. Na mesma linha, o sentimento de raiva esteve associado a cerca de 17,5% de todas as postagens, o que apoia ainda mais as atitudes pessimistas da maioria da população. Algumas delas exibiram sentimentos negativos como nojo e medo, com proporções correspondentes de 5,5% e 3,4% do total de postagens. A categoria de alegria, que é um sentimento positivo, e aplicou-se apenas a uma percentagem muito pequena (0,5%) das citações.

O método LDA também foi utilizado para realizar um estudo do quociente emocional dos comentários feitos no YouTube, e os resultados revelaram que mais da metade dos comentários enviados para aquela plataforma de todo o mundo podem ser compreendidos entre três sentimentos: surpresa, medo ou confiança. Conforme mostra a Figura 15, os comentários que expressaram o sentimento de surpresa compuseram a maior parte do gráfico, seguidos pelo sentimento de medo. Na mesma linha, o sentimento de confiança esteve ligado a cerca de 15% dos comentários, o que reforça as opiniões favoráveis das pessoas. Alguns dos comentários apresentaram sentimentos negativos como desprezo, raiva e desespero, com proporções correspondentes de 14,3, 10,4% e 3,5% do total de comentários.

**Figura 14.** Roda de sentimentos mostrando os quocientes emocionais dos postagens examinados



**Figura 15.** Roda de sentimentos mostrando os quocientes emocionais dos comentários de YouTube examinados.



## Impacto nos resultados do estudo

- O conjunto de dados depurado estava livre de ruídos e informações irrelevantes, permitindo uma análise focada e aprofundada.
- A qualidade do conjunto de dados pós-depuração permitiu uma identificação mais clara e precisa de tendências, sentimentos e padrões no discurso sobre o consumo verde.
- Ao analisar um conjunto de dados refinado, foram obtidos insights mais profundos sobre como os usuários das redes sociais percebem e discutem o consumo verde, especialmente sobre sustentabilidade e impacto ambiental.
- A confiabilidade das conclusões do nosso estudo foi significativamente reforçada pelo rigoroso processo de depuração de dados.
- Os dados depurados reduziram o risco de conclusões erradas e garantiram que as interpretações refletissem com precisão o discurso público.
- Nossas descobertas, exploradas a partir de um conjunto de dados meticulosamente depurados, oferecem insights confiáveis que podem informar estratégias para promover o consumo sustentável e compreender o comportamento do consumidor na era digital.

## CONCLUSÃO

Os dados coletados por este estudo oferecem informações valiosas sobre o discurso em torno do consumo ambientalmente consciente nas plataformas de redes sociais X e YouTube. Utilizando métodos inovadores, como mineração de texto e análise de sentimentos, conseguimos discernir tendências, atitudes emocionais e envolvimento com o tema em vários momentos, especialmente durante os primeiros meses da pandemia da Covid-19. A análise de tendências

mostrou a crescente consciência e interesse do público no consumo verde. Os dados do X e do YouTube demonstraram vários picos na intensidade das discussões, mostrando um aumento no envolvimento do público com o tema. No entanto, o sentimento associado a estas discussões variou entre as plataformas. Nossa análise de sentimentos produziu resultados contrastantes: os usuários do X expressaram predominantemente sentimentos negativos em relação ao consumo verde (60,38%), enquanto os comentários do YouTube foram mais positivos (57,19%). Isto sugere que a percepção do consumo verde difere entre plataformas e justifica mais pesquisas para explorar estas discrepâncias.

## Explicação detalhada dos métodos analíticos

### Visão geral da análise de dados

Nosso estudo envolveu uma abordagem de análise multifacetada, utilizando métodos quantitativos e qualitativos para extrair insights significativos dos dados de mídias sociais.

### Análise de sentimentos

Detalhamos como a análise de sentimentos foi conduzida, incluindo os algoritmos ou modelos específicos utilizados (por exemplo, TextBlob ou modelos de análise de sentimentos personalizados). Explicamos o processo de categorização dos sentimentos em positivos, negativos e neutros e como essas categorias foram analisadas quantitativamente.

### Modelagem de tópico

Descrevemos o uso do método LDA (*Latent Dirichlet Allocation*) para revelar tópicos predominantes no conjunto de dados. Explicamos como os tópicos foram identificados e categorizados, e os critérios utilizados para determinar a relevância e importância de cada tópico.

### Análise de tendências

Descrevemos os métodos utilizados para identificar e analisar tendências no conjunto de dados ao longo do tempo. Esclarecemos como os altos e baixos nos dados, demonstrados em gráficos, foram correlacionados com eventos externos e períodos de tempo.

## Técnicas de visualização de dados

Fornecemos detalhes sobre os tipos de visualizações utilizadas (por exemplo, gráficos de linhas, nuvem de palavras) e como elas foram empregadas para representar eficazmente as descobertas dos dados. Discutimos a lógica por trás da escolha de cada tipo de visualização em relação aos dados apresentados.

A análise de palavras-chave e frases relacionadas ao consumo verde trouxe tendências notáveis. A frequência de citações revelou os altos e baixos do interesse público em termos como “sustentabilidade” e “consumo verde”. Essas mudanças coincidiram frequentemente com acontecimentos globais e ciclos de notícias, indicando que o discurso público em torno do consumo verde é suscetível a influências externas. O conteúdo emocional das discussões, revelado através do método LDA, matizou ainda mais a nossa compreensão do discurso público. As respostas emocionais dos usuários do X foram caracterizadas principalmente por sentimentos de confiança, raiva e tristeza. Por outro lado, os comentários do YouTube foram marcados principalmente por surpresa, medo e confiança. Isto sublinha a complexa gama de respostas emocionais ao tema do consumo ambientalmente responsável. Nosso estudo também indicou os principais fatores associados ao discurso sobre o consumo verde. Entre os seis fatores considerados, o ambiente e o comportamento de consumo pareciam ser primordiais. Essas conclusões podem ajudar na definição de estratégias futuras para promover o consumo sustentável. Finalmente, este estudo empregou com sucesso uma nuvem de palavras para visualizar a linguagem comum nas postagens examinadas. A nuvem de palavras para esta pesquisa ressaltou termos como “conservação”, “verde”, “clima”, “limpo”, “resíduos” e “energia”, sugerindo uma ampla gama de preocupações associadas ao discurso sobre o consumo verde.

Nossa pesquisa sugere várias implicações. Em primeiro lugar, destaca o importante papel das plataformas de redes sociais na promoção de discussões em torno do consumo verde. Esta informação pode servir como um farol para empresas, decisores políticos e ativistas sociais, destacando a necessidade de envolverem-se e responder aos argumentos. As conclusões sublinham o potencial de alavancar as redes sociais como uma ferramenta poderosa para promoção de práticas sustentáveis e influência positiva no comportamento do consumidor. À medida que enfrentamos os desafios da sustentabilidade ambiental, o envolvimento ativo de todos stakeholders na definição destes diálogos online torna-se cada vez mais crucial para promover um compromisso coletivo com escolhas ambientalmente amigáveis. Os dados também destacam a importância da sustentabilidade e das preocupações ambientais no comportamento de consumo, destacando a crescente procura por produtos verdes e práticas sustentáveis. Além disso, assinala uma mudança de paradigma na dinâmica do mercado, instando as empresas a reavaliarem as suas estratégias e a incorporarem iniciativas ambientalmente amigáveis nos seus produtos e serviços. À luz desta mudança, as empresas que adotam proativamente práticas sustentáveis não só se alinham com as preferências dos consumidores, mas também se posicionam como líderes em práticas empresariais responsáveis e éticas, promovendo a fidelidade à marca a longo prazo e impacto social positivo. Em segundo lugar, a análise dos padrões emocionais revela respostas aos argumentos. Isto proporciona uma compreensão mais profunda do sentimento público em relação ao consumo verde, o que pode orientar a criação de estratégias de marketing e políticas públicas mais eficazes e direcionadas. Além disso, compreender os aspectos emocionais nas discussões sobre o ambiente permite que as empresas e os tomadores de decisão política ajustem as suas mensagens e projetos para se conectarem com gatilhos emocionais específicos. Isto aumenta as chances de mais pessoas adotarem práticas e políticas verdes. Tal consciência emocional na tomada de decisões tem o potencial de construir uma relação mais empática e conectada

entre os stakeholders, apoiando, em última análise, o sucesso global das iniciativas verdes. Em terceiro, as implicações reveladas neste estudo mostram que as discussões e sensibilidades sobre o consumo verde nas redes sociais têm um efeito decisivo nas tendências dos consumidores à aquisição de produtos ecológicos. Observou-se que conteúdos positivos e informativos sobre produtos sustentáveis, especialmente compartilhados em plataformas de redes sociais, aumentam a probabilidade por sua escolha. Portanto, a análise das redes sociais contribui significativamente para os esforços de sensibilização do público sobre a sustentabilidade, fornecendo informações valiosas sobre o comportamento e as preferências de compra dos consumidores. À luz destas conclusões, as empresas e os grupos de defesa de direitos podem aproveitar estrategicamente as plataformas de redes sociais para disseminar conteúdos positivos e informativos, influenciando assim as escolhas dos consumidores por opções mais sustentáveis. Além disso, os decisores políticos podem utilizar as informações para conceber campanhas direcionadas que promovam comportamentos e produtos ecológicos, propiciando uma cultura de responsabilidade ambiental.

Resumindo, a nossa investigação lançou luz sobre a intrincada dinâmica do discurso público sobre o consumo verde em duas plataformas populares de redes sociais. Elucida como as plataformas moldam o discurso e, por sua vez, as atitudes do público em relação ao consumo ambientalmente consciente. À medida que o interesse nesta área cresce, esperamos ver mais pesquisas explorando as implicações destas descobertas e o seu potencial para informar estratégias que promovam o consumo sustentável.

## REFERÊNCIAS

- Agyabeng-Mensah, Y., Afum, E., & Ahenkorah, E. (2020). Exploring financial performance and green logistics management practices: Examining the mediating influences of market, environmental and social performances. *Journal of Cleaner Production*, 258, 120613. <https://doi.org/10.1016/j.jclepro.2020.120613>
- Ahamad, N. R., & Ariffin, M. (2018). Assessment of knowledge, attitude, and practice towards sustainable consumption among university students in Selangor, Malaysia. *Sustainable Production and Consumption*, 16, 88-98. <https://doi.org/10.1016/j.spc.2018.06.006>
- Akhtar, R., Sultana, S., Masud, M. M., Jafrin, N., & Al-Mamun, A. (2021). Consumers' environmental ethics, willingness, and green consumerism between lower and higher income groups. *Resources, Conservation and Recycling*, 168, 105274. <https://doi.org/10.1016/j.resconrec.2020.105274>
- Balcioğlu, Y. S. (2022). Detection of depression and anxiety symptoms via Twitter after Covid-19 with machine learning., 2. In *Başkent International Conference On Multidisciplinary Studies* (pp. 261-265).
- Bedard, S., & Reisdorf, C. A. (2018). Millennials' green consumption behaviour: Exploring the role of social media. *Corporate Social Responsibility and Environmental Management*, 25(1), 1388-1396. <https://doi.org/10.1002/csr.1654>
- Biswas, A. (2016). Impact of social media usage factors on green consumption behavior based on technology acceptance model. *Journal of Advanced Management Science*, 4(2), 92-97. <https://doi.org/10.12720/joams.4.2.92-97>

- Brzustewicz, P., & Singh, A. (2021). Sustainable consumption in consumer behavior in the time of covid-19: Topic modeling on twitter data using LDA. *Energies*, 14(18), 5787. <https://doi.org/10.3390/en14185787>
- Costa, C. S. R., da Costa, M. F., Maciel, R. G., Aguiar, E. C., & Wanderley, L. O. (2021). Consumer antecedents towards green product purchase intentions. *Journal of Cleaner Production*, 313, 127964. <https://doi.org/10.1016/j.jclepro.2021.127964>
- Djafarova, E., & Rushworth, C. (2017). Exploring the credibility of online celebrities' Instagram profiles in influencing the purchase decisions of young female users. *Computers in Human Behavior*, 68, 1-7. <https://doi.org/10.1016/j.chb.2016.11.009>
- ElHaffar, G., Durif, F., & Dubé, L. (2020). Towards closing the attitude-intention-behavior gap in green consumption: A narrative review of the literature and an overview of future research directions. *Journal of Cleaner Production*, 275, 122556. <https://doi.org/10.1016/j.jclepro.2020.122556>
- Groening, C., Sarkis, J., & Zhu, Q. (2018). Green marketing consumer-level theory review: A compendium of applied theories and further research directions. *Journal of Cleaner Production*, 172, 1848-1866. <https://doi.org/10.1016/j.jclepro.2017.12.002>
- Huang, H., Long, R., Chen, H., Sun, K., & Li, Q. (2022). Exploring public attention about green consumption on Sina Weibo: Using text mining and deep learning. *Sustainable Production and Consumption*, 30, 674-685. <https://doi.org/10.1016/j.spc.2021.12.017>
- Huseynov, F., & Yildirim, S. O. (2019). Online consumer typologies and their shopping behaviors in B2C e-commerce platforms. *Sage Open*, 9(2), 1-19. <https://doi.org/10.1177/2158244019854639>
- Imamah, I., Husni, H., Rachman, E. M., Suzanti, I. O., & Mufarroha, F. A. (2020). Text mining and support vector machine for sentiment analysis of tourist reviews in Bangkalan Regency. *Journal of Physics*, 1477, 022023. <https://doi.org/10.1088/1742-6596/1477/2/022023>
- Jain, V. K., Gupta, A., Tyagi, V., & Verma, H. (2020). Social media and green consumption behavior of millennials. *Journal of Content, Community & Communication*, 11, 221-230. <https://doi.org/10.31620/JCCC.06.20/16>
- Jalali, S. S., & Khalid, H. (2019). Understanding Instagram influencers values in green consumption behaviour: A review paper. *Open International Journal of Informatics*, 7(Special Issue 1), 47-58. <https://oiji.utm.my/index.php/oiji/article/view/115>
- Jian, Y., Yu, I. Y., Yang, M. X., & Zeng, K. J. (2020). The impacts of fear and uncertainty of Covid-19 on environmental concerns, brand trust, and behavioral intentions toward green hotels. *Sustainability*, 12(20), 8688. <https://doi.org/10.3390/su12208688>
- Jiang, Z., & Gao, X. (2023). Text mining and quantitative evaluation of China's green consumption policies based on green consumption objects. *Environment, Development and Sustainability*, 26(3), 6601-6622. <https://doi.org/10.1007/s10668-023-02976-w>
- Kluczek, A. (2017). Quick green scan: A methodology for improving green performance in terms of manufacturing processes. *Sustainability*, 9(1), 88. <https://doi.org/10.3390/su9010088>
- Kunte, A. V., & Panicker, S. (2019). *Using textual data for personality prediction: A machine learning approach*. 4th International Conference on Information Systems and Computer Networks (ISCON).
- Li, M. (2020). Review of consumers' green consumption behavior. *American Journal of Industrial and Business Management*, 10, 585-599. <https://doi.org/10.4236/ajibm.2020.103039>

- Lu, Y., & Park, S. D. (2022). Time series analysis of policy discourse on green consumption in China: Text mining and network analysis. *Sustainability*, 14(22), 14704. <https://doi.org/10.3390/su142214704>
- Ma, Y., & Qiao, E. (2021). *Research on Accurate prediction of operating energy consumption of green buildings based on improved machine learning*. IEEE International Conference on Industrial Application of Artificial Intelligence (IAAI).
- Al Mamun, A., Mohamad, M. R., Yaacob, M. R. B., & Mohiuddin, M. (2018). Intention and behavior towards green consumption among low-income households. *Journal of Environmental Management*, 227, 73-86. <https://doi.org/10.1016/j.jenvman.2018.08.061>
- Park, J. Y., Mistur, E., Kim, D., Mo, Y., & Hoefler, R. (2022). Toward human-centric urban infrastructure: Text mining for social media data to identify the public perception of Covid-19 policy in transportation hubs. *Sustainable Cities and Society*, 76, 103524. <https://doi.org/10.1016/j.scs.2021.103524>
- Sajeewanie, L. A. C., Ab Yajid, M. S., Khatibi, A., Azam, F., & Tham, J. (2019). Integrated model for green purchasing intention and green adoption: Future research direction. *Journal of Sociological Research*, 10(2), 23-66. <https://doi.org/10.5296/jsr.v10i2.14996>
- Saraç, Ö. (2022). Kültür Turistlerinin Sürdürülebilir Tüketim Davranışlarının Cinsiyete Göre Farklılıkları Safranbolu Üzerinde Bir Araştırma. *Journal of Humanities and Tourism Research*, 12(2), 265-283.
- Serrano, L., Ariza-Montes, A., Nader, M., Sianes, A., & Law, R. (2021). Exploring preferences and sustainable attitudes of Airbnb green users in the review comments and ratings: A text mining approach. In *Sustainable Consumer Behaviour and the Environment* (pp. 114-132). Routledge.
- Sharifi, A. (2021). Co-benefits and synergies between urban climate change mitigation and adaptation measures: A literature review. *Science of the Total Environment*, 750, 141642. <https://doi.org/10.1016/j.scitotenv.2020.141642>
- Sun, X., Su, W., Guo, X., & Tian, Z. (2021). The impact of awe induced by Covid-19 pandemic on green consumption behavior in China. *International Journal of Environmental Research and Public Health*, 18(2), 543. <https://doi.org/10.3390/ijerph18020543>
- Tang, H., Xu, Y., Lin, A., Heidari, A. A., Wang, M., Chen, H., ... & Li, C. (2020). Predicting green consumption behaviors of students using efficient firefly grey wolf-assisted k-nearest neighbor classifiers. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2020.2973763>
- Tanveer, M., Richhariya, B., Khan, R. U., Rashid, A. H., Khanna, P., Prasad, M., & Lin, C. T. (2020). Machine learning techniques for the diagnosis of Alzheimer's disease: A review. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 16(1s), 1-35. <https://doi.org/10.1145/3344998>
- Wang, Y. (2021). *Research on the influence mechanism of green cognition level on consumers' green consumption behavior: An empirical study based on SPSS*. International Conference on Management Science and Software Engineering (ICMSSE). <https://doi.org/10.1109/ICMSSE53595.2021.00044>
- Wu, Z., Zhang, Y., Chen, Q., & Wang, H. (2021). Attitude of Chinese public towards municipal solid waste sorting policy: A text mining study. *Science of the Total Environment*, 756, 142674. <https://doi.org/10.1016/j.scitotenv.2020.142674>
- Xie, S., & Madni, G. (2023). Impact of social media on young generation's green consumption behavior through subjective norms and perceived green value. *Sustainability*, 15, 3739. <https://doi.org/10.3390/su15043739>

- Yang, M., Chen, H., Long, R., & Yang, J. (2022). The impact of different regulation policies on promoting green consumption behavior based on social network modeling. *Sustainable Production and Consumption*, 32, 468-478. <https://doi.org/10.1016/j.spc.2022.05.007>
- Yang, W., Feng, L., Wang, Z., & Fan, X. (2023). Carbon emissions and national sustainable development goals coupling coordination degree study from a global perspective: Characteristics, heterogeneity, and spatial effects. *Sustainability*, 15(11), 9070. <https://doi.org/10.3390/su15119070>
- Yang, Y., Li, Y., & Guo, Y. (2022). Impact of the differences in carbon footprint driving factors on carbon emission reduction of urban agglomerations given SDGs: A case study of the Guanzhong in China. *Sustainable Cities and Society*, 85, 104024. <https://doi.org/10.1016/j.scs.2022.104024>
- Yao, J., Guo, X., Wang, L., & Jiang, H. (2022). Understanding green consumption: A literature review based on factor analysis and bibliometric method. *Sustainability*, 14, 8324. <https://doi.org/10.3390/su14148324>
- Yazdavar, A. H., Mahdavinejad, M. S., Bajaj, G., Romine, W., Sheth, A., Monadjemi, A. H., ... & Hitzler, P. (2020). Multimodal mental health analysis in social media. *Plos ONE*, 15(4), 1-27. <https://doi.org/10.1371/journal.pone.0226248>
- Zaremohzzabieh, Z., Ismail, N., Ahrari, S., & Samah, A. A. (2021). The effects of consumer attitude on green purchase intention: A meta-analytic path analysis. *Journal of Business Research*, 132, 732-743. <https://doi.org/10.1016/j.jbusres.2020.10.053>
- Zhao, G., Geng, Y., Sun, H., Tian, X., Chen, W., & Wu, D. (2020). Mapping the knowledge of green consumption: A meta-analysis. *Environmental Science and Pollution Research*, 27, 44937-44950. <https://doi.org/10.1007/s11356-020-11029-y>

## NOTA

Este estudo é um dos resultados do projeto numerado 2022-A-113-07, financiado pelo Fundo de Pesquisa da Gebze Technical University.

## CONFLITOS DE INTERESSE

Os autores não têm conflitos de interesse a declarar.

## CONTRIBUIÇÃO DOS AUTORES

Ceren Cubukcu Cerasi: Conceitualização; aquisição de financiamento; Investigação; Administração de projetos; Supervisão; Validação; Visualização; Redação – revisão e edição.

Yavuz Selim Balcioğlu: Análise formal; Investigação; Metodologia; Validação; Visualização; Redação – rascunho original.

Farid Huseynov: Conceituação; Investigação; Metodologia; Recursos; Programas; Redação – revisão e edição.

Asli Kilic: Curadoria de dados; Análise formal; Metodologia; Validação; Visualização; Redação – rascunho original.