

# Atenção digital e rompimentos de mercado evidências com influenciadores no Brasil

Bruno Alves de Oliveira<sup>†, a</sup> 

Francisco Gildemir Ferreira da Silva<sup>‡, a</sup> 

Victor Aguiar Evangelista de Farias<sup>\*, a</sup> 

Julio Fernando Seara Sequeira da Mota Lobão<sup>\*\*\*, b</sup> 

<sup>a</sup>Universidade Federal do Ceará

<sup>b</sup>Universidade do Porto

**Resumo** Este artigo investiga se a atividade de influenciadores financeiros no X (antigo *Twitter*) está associada a rompimentos de curto prazo nas zonas de suporte e resistência do Ibovespa. A base de postagens foi coletada via *Apify* a partir de contas identificadas nos relatórios *FInfluence* da ANBIMA. Os *tweets* foram classificados tematicamente e quanto ao sentimento com o auxílio do modelo *GPT-4o-mini*, com validação manual em amostras balanceadas. Mantiveram-se apenas as publicações classificadas como Economia e Mercado Financeiro. A análise utiliza bandas calibradas pela volatilidade e medidas diárias agregadas. A estratégia empírica combina estudo de eventos e modelos logísticos. Os resultados indicam que o engajamento agregado apresenta associação mais consistente com os rompimentos do que o sentimento. Assim, no nosso artigo contribui para o debate sobre a eficiência informacional dos mercados e, em termos práticos, apresenta implicações relevantes para diferentes agentes do mercado, incluindo investidores, reguladores e formuladores de políticas públicas.

**Palavras-chave:** X (*Twitter*); atenção do investidor; engajamento; sentimento; Ibovespa  
**Códigos JEL:** G14, G41, C55.

## 1. Introdução

As redes sociais passaram a ocupar papel crescente na circulação de informações econômicas e financeiras, ampliando a velocidade com que opiniões,

**Como citar:** Oliveira, B. A., Silva, F. G. F., Farias, V. A. E. & Lobão, J. F. S. S. M. (2025). Atenção digital e rompimentos de mercado evidências com influenciadores no Brasil. *Revista Brasileira de Finanças*, 24, e202601. <https://doi.org/10.xxxx/xxxx>  
Submetido em January 1, 2026. Publicado *on-line* em March 31, 2026.

Editor responsável: Mr. Editor.

<sup>†</sup>[brunoalvesoliveira96@alu.ufc.br](mailto:brunoalvesoliveira96@alu.ufc.br)

<sup>‡</sup>[gildemir@ufc.br](mailto:gildemir@ufc.br)

<sup>\*</sup>[victor.farias@ufc.br](mailto:victor.farias@ufc.br)

<sup>\*\*</sup>[jlobao@fep.up.pt](mailto:jlobao@fep.up.pt)

interpretações e narrativas alcançam investidores e demais participantes do mercado. Nesse contexto, o X (antigo *Twitter*) destaca-se como ambiente particularmente relevante por combinar difusão instantânea de conteúdo, elevada visibilidade pública e forte capacidade de interação entre usuários. A literatura já documenta que informações extraídas dessa plataforma podem conter conteúdo preditivo para retornos, volatilidade e volume de negociação, sobretudo quando associadas a medidas de sentimento e atenção do investidor (Fan e Gordon, 2014; Yuexin Mao e Liu, 2012).

Do ponto de vista teórico, esse fenômeno dialoga com os limites da Hipótese de Mercado Eficiente (HME). Embora a formulação clássica de Fama (1970) sustente que preços incorporam rapidamente a informação disponível, a literatura de finanças comportamentais mostra que os agentes enfrentam restrições cognitivas, informacionais e operacionais para processar todo o fluxo informacional disponível. A crítica a essa noção de racionalidade plena remonta a Simon (1955), que propõe substituir a racionalidade global do agente econômico por um comportamento compatível com limites efetivos de informação e capacidade de processamento. Em complemento, Kahneman e Riepe (1998) mostram que decisões financeiras são frequentemente afetadas por vieses de julgamento e de preferência, ao passo que Barberis e Thaler (2003) sintetizam essa agenda ao destacar que a literatura de finanças comportamentais se estrutura em torno de dois pilares: limites à arbitragem e desvios sistemáticos de racionalidade. Em consequência, investidores tendem a reagir de forma mais intensa a conteúdos que capturam sua atenção, especialmente quando tais conteúdos são amplificados por agentes com grande exposição pública. Nesse ambiente, influenciadores digitais podem atuar como geradores de relevância informacional, ao reduzir custos de busca, direcionar o foco dos investidores e, potencialmente, afetar a dinâmica de preços. Essa interpretação também encontra respaldo em evidências sobre o comportamento de participantes profissionais do mercado. Taylor e Allen (1992) mostram que a análise técnica recebe peso relevante entre operadores do mercado cambial, sobretudo em horizontes mais curtos, enquanto Menkhoff (2010) documentam que gestores profissionais também atribuem importância expressiva a esse tipo de análise, especialmente em janelas de curto prazo e em contextos nos quais fatores psicológicos são percebidos como relevantes para a formação de preços.

A literatura recente sobre mídias sociais e mercados financeiros sugere que esse efeito pode operar por diferentes canais. Um deles é o sentimento agregado das mensagens, que expressa a direção qualitativa do conteúdo disseminado. Outro, menos explorado no contexto brasileiro, é a intensidade de atenção associada à circulação dessas mensagens, mensurada por interações

observáveis como curtidas, respostas, citações e republicações. Em vez de refletir apenas opinião, essas métricas capturam o grau de repercussão e difusão do conteúdo, funcionando como medida da atenção coletiva despertada por publicações financeiras em plataformas digitais.

Este artigo investiga se a atividade de grandes influenciadores financeiros brasileiros no X está associada a rompimentos de curto prazo nas zonas de suporte e resistência do Ibovespa. A escolha por modelar a probabilidade de rompimento, em vez da volatilidade condicional, decorre do entendimento de que choques informacionais oriundos de postagens específicas tendem a se manifestar como descontinuidades discretas — ultrapassagens de limiares de preço — mais do que como alterações persistentes na variância dos retornos. Para operacionalizar esse problema, utiliza-se a relação de influenciadores identificados nos relatórios *FInfluence*, publicados pela ANBIMA (ANBIMA, 2020-2026), a partir dos quais se constrói uma base de postagens públicas coletadas via *Apify*.

Metodologicamente, o estudo combina três etapas. Primeiro, os *tweets* são classificados tematicamente e quanto ao sentimento com o auxílio de um modelo de linguagem de grande porte (GPT-4o-mini), sendo mantidas apenas as publicações relacionadas a Economia e Mercado Financeiro. Segundo, essas postagens são agregadas em frequência diária para a construção de medidas de engajamento e sentimento. Terceiro, tais variáveis são relacionadas a bandas de suporte e resistência calibradas pela volatilidade histórica do Ibovespa, por meio de estudos de evento e modelos logísticos voltados à identificação da probabilidade de rompimentos.

A contribuição do artigo é tripla. Em primeiro lugar, o trabalho estende a literatura de atenção limitada ao contexto de influenciadores financeiros nativos digitais em um mercado emergente. Em segundo, propõe uma estratégia empírica que desloca o foco da previsão direta de retornos para a modelagem de rompimentos técnicos, mais aderente à natureza discreta de choques informacionais de curto prazo. Em terceiro, oferece evidência sobre o papel do engajamento agregado como medida da atenção no mercado acionário brasileiro, distinguindo sua contribuição da informação contida no sentimento das mensagens.

Os resultados indicam que o engajamento agregado apresenta associação mais consistente com os rompimentos das bandas de suporte e resistência do que as medidas de sentimento. Em nível descritivo, a reação da atenção digital se concentra em janelas curtas ao redor dos eventos; em nível econométrico, a associação entre atenção e rompimentos mostra-se mais estável em janelas mais amplas de observação. Em conjunto, os achados sugerem que a atividade

de influenciadores financeiros nas redes sociais atua menos como determinante autônomo da direção dos preços e mais como mecanismo de relevância e amplificação informacional em torno de episódios de maior tensão no mercado.

O restante do artigo está organizado da seguinte forma. A Seção 2 apresenta a revisão de literatura. A Seção 3 descreve a estratégia empírica, a construção da base e das variáveis. A Seção 4 discute os resultados. Por fim, a Seção 5 conclui.

## 2. Revisão de literatura

### 2.1 Hipótese de Mercado Eficiente e limites comportamentais

A obra seminal de Fama (1970) estruturou a Hipótese de Mercado Eficiente (HME) e distinguiu três formas de eficiência informacional com base na teoria da utilidade esperada. Na forma fraca, os preços incorporam integralmente a informação histórica; na semiforte, ajustam-se de maneira rápida e não viesada às informações públicas; e, na forma forte, nem mesmo informações privilegiadas permitiriam retornos anormais sistemáticos. Contudo, a literatura de finanças comportamentais relativiza o pressuposto de racionalidade plena. A teoria do prospecto de Tversky e Kahneman (1979) demonstra que agentes exibem aversão ao risco no domínio dos ganhos e propensão ao risco nas perdas, o que implica decisões influenciadas por vieses cognitivos e emocionais. Nesse contexto, o sentimento agregado pode constituir fator explicativo adicional para retornos e volatilidade, mesmo sob ampla disponibilidade informacional.

### 2.2 Sentimento agregado, conteúdo informacional e previsibilidade

Evidências empíricas indicam que métricas emocionais extraídas do *Twitter* possuem conteúdo informacional relevante. Zhang et al. (2011) mensuram a frequência diária de palavras associadas a estados emocionais — como *hope*, *fear* e *worry* — e encontram correlação negativa significativa com os principais índices acionários e positiva com o VIX (índice do medo), o que sugere que picos emocionais antecedem deterioração de mercado. De forma mais robusta, Bollen et al. (2011) identificam que dimensões específicas de humor, particularmente *calm*, apresentam poder preditivo estatisticamente significativo sobre o índice *Dow Jones Industrial Average* (DJIA) com defasagens de três a quatro dias.

Complementarmente, Sul et al. (2017) demonstram que estratégias baseadas em sentimento extraído do *Twitter* geram retornos anormais economicamente relevantes, mesmo após controle por fatores tradicionais de risco, o que indica que o mercado não incorpora instantaneamente a informação

difusa proveniente das mídias sociais. No que tange ao conteúdo informacional, [Sprenger e Welpe \(2011\)](#) distinguem empiricamente *news* de *noise* e evidenciam que o impacto sobre retornos, volume e volatilidade depende da natureza econômica da mensagem. Em linha com essa literatura, [Baker e Wurgler \(2006\)](#) e [Lee et al. \(2002\)](#) mostram que o sentimento do investidor afeta retornos esperados, volatilidade condicional e prêmio de risco, especialmente em ativos mais sensíveis a expectativas.

### 2.3 Influenciadores políticos e reação de mercado

Uma vertente recente desloca o foco do sentimento difuso para agentes com elevada centralidade e capacidade de mobilização. [Machus et al. \(2022\)](#) analisam *tweets* de Donald Trump direcionados a empresas específicas e identificam retornos anormais e aumento significativo de volume no momento da publicação, seguidos de reversão parcial, o que sugere reação de curto prazo possivelmente associada à sobre-reação.

Em horizonte mais amplo, [Agrawal e Agarwal \(2025\)](#) documentam a existência de um *pre-post-Twitter-drift* associado a *tweets* de Elon Musk e Donald Trump, com persistência de retornos anormais por até 250 dias. O fenômeno não é observado para anúncios corporativos divulgados na mesma plataforma, o que indica que a anomalia decorre da influência do emissor e não do canal tecnológico, tensionando diretamente a forma semiforte da HME.

### 2.4 Influenciadores corporativos e ativos específicos

No âmbito corporativo, [Šević et al. \(2023\)](#) encontram correlação de intensidade média a alta entre o conteúdo e o sentimento de *tweets* de Elon Musk e o desempenho das ações da Tesla no período 2019–2022, o que sugere associação estatisticamente relevante entre comunicação digital e variações de preço no curto prazo. De maneira convergente, [Ante \(2022\)](#) demonstram que *tweets* de Musk sobre criptomoedas produzem retornos anormais e aumentos significativos de volume, com efeitos particularmente pronunciados em ativos de maior componente especulativo, o que evidencia a capacidade de influenciadores de coordenar expectativas e desencadear ajustes rápidos de preço.

A literatura converge para três proposições centrais: (i) o sentimento agregado mensurado em mídias sociais possui conteúdo informacional relevante para retornos e volatilidade; (ii) a natureza da mensagem condiciona a magnitude e a persistência da reação; e (iii) agentes com elevada exposição pública amplificam esses efeitos e podem gerar tanto respostas transitórias com reversão quanto distorções persistentes. Esse arcabouço teórico-empírico fun-

damenta a investigação do papel de influenciadores digitais na dinâmica de volatilidade e formação de preços no mercado acionário brasileiro.

## 2.5 Atenção do investidor, *herding* e mercados emergentes

Uma dimensão central deste estudo reside na ideia de que a informação disseminada por influenciadores financeiros não afeta o mercado apenas por seu conteúdo, mas também por sua capacidade de capturar atenção e coordenar o foco dos investidores. Essa perspectiva dialoga com a agenda de finanças comportamentais e eficiência de mercado desenvolvida por Lobão, particularmente em trabalhos que examinam *herding*, reação agregada dos investidores e limites da incorporação plena de informação em mercados acionários.

No contexto brasileiro, Lobão et al. (2024) mostram que o comportamento dos investidores pode exibir dinâmica coletiva relevante no mercado acionário, ao propor que decisões individuais não são formadas de maneira totalmente isolada, mas também em resposta ao ambiente informacional e ao comportamento observado dos demais agentes. Em paralelo, Costa et al. (2023) evidenciam que diferentes estados de *herding* estão associados ao comportamento dos retornos, o que reforça a interpretação de que episódios de atenção e coordenação informacional podem ter implicações para a dinâmica de preços.

Em escala regional, Lobão e Almeida (2024) documentam evidências de comportamento de manada nos principais mercados acionários da América Latina, o que amplia a relevância dessa agenda para economias emergentes e sugere que desvios em relação à hipótese de eficiência semiforte podem ser particularmente relevantes em ambientes marcados por maior fricção informacional e menor profundidade de mercado. Essa linha de pesquisa é especialmente importante para o presente artigo, pois o mercado brasileiro permanece menos explorado do que os mercados desenvolvidos na literatura sobre mídias sociais, atenção do investidor e reação de preços.

Mais recentemente, a produção de Lobão passou a incorporar de forma explícita o papel das redes sociais e da influência pública na eficiência de mercado. Em Cunha et al. (2025), argumenta-se que plataformas digitais e emissores com elevada visibilidade pública podem alterar a difusão de informação e afetar o processo de formação de preços, o que fornece enquadramento particularmente próximo ao deste estudo. A contribuição do presente artigo consiste, portanto, em avançar essa agenda ao conectar influenciadores financeiros nativos do ambiente digital, atenção medida por engajamento agregado e rompimentos técnicos no Ibovespa.

### 3. Metodologia e dados

Este artigo analisa a relação entre picos de atenção gerados por influenciadores financeiros — medidos por engajamento e análise de sentimento — e a dinâmica estrutural de suporte e resistência no mercado acionário brasileiro [Guerra \(2022\)](#). A estrutura de mercado é capturada por zonas de suporte e resistência calibradas a partir da volatilidade histórica dos retornos, o que permite avaliar se choques informacionais oriundos das redes sociais estão associados a rompimentos dessas zonas.

#### 3.1 Zonas de Suporte e Resistência

Com base em [Guerra \(2022\)](#), é possível estimar zonas de suporte e resistência de um determinado ano a partir das informações observadas no ano imediatamente anterior. O presente estudo incorpora a extensão semanal proposta no mesmo trabalho, o que permite acompanhar com maior granularidade a dinâmica das bandas ao longo do tempo.

A construção dessas zonas parte da ideia de que a volatilidade histórica resume a amplitude usual de oscilação dos preços em torno de um valor de referência. Sob essa perspectiva, a banda superior representa uma região em que o preço se afasta positivamente desse valor em magnitude compatível com a volatilidade estimada, ao passo que a banda inferior representa o afastamento negativo correspondente. Assim, a volatilidade não é utilizada apenas como medida de risco, mas também como parâmetro para delimitar regiões de tensão no processo de formação de preços.

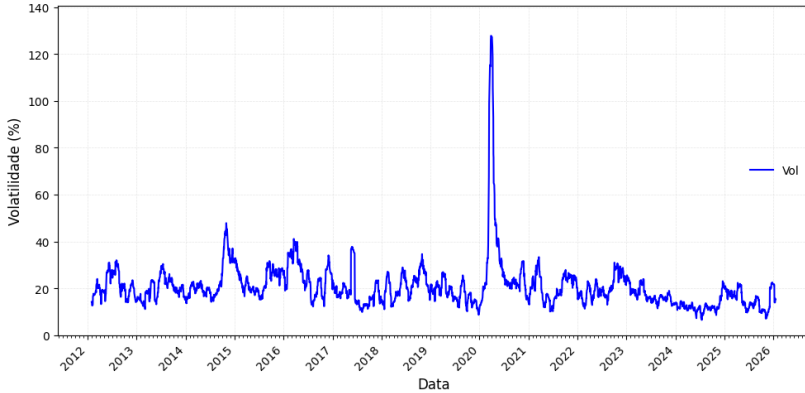
Nesse contexto, a volatilidade para um dado horizonte temporal é definida da seguinte forma (Figura 1):

$$Vol_t = \sigma \sqrt{T} \quad (1)$$

onde  $Vol_t$  representa a volatilidade no horizonte considerado,  $\sigma$  corresponde ao desvio-padrão dos retornos e  $T$  denota o número de períodos incluídos no intervalo de análise.

A Equação (1) decorre da regra da raiz do tempo, usualmente empregada para escalar a volatilidade entre diferentes horizontes temporais sob a hipótese de independência serial dos retornos. A anualização da volatilidade cumpre, portanto, função metodológica específica: fornecer medida padronizada e comparável da dispersão dos retornos, que pode ser convertida para outros horizontes, como o semanal, sem perda de consistência interna na construção das bandas.

**Figura 1**  
**Volatilidade anual do Ibovespa**



A partir dessa definição, as zonas anualizadas de suporte e resistência são obtidas pelas seguintes expressões:

$$Supply_t^{(y)} = (1 + Vol_t^{(y)}) \cdot P_t^{ref} \quad (2)$$

$$Demand_t^{(y)} = (1 - Vol_t^{(y)}) \cdot P_t^{ref} \quad (3)$$

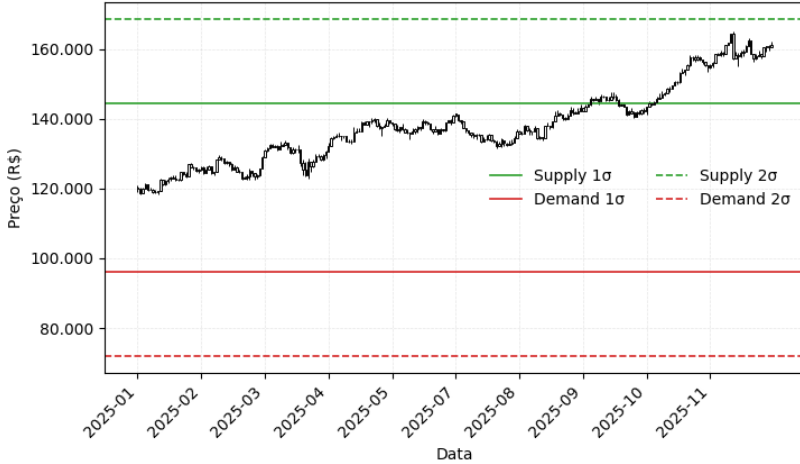
em que  $P_t^{ref}$  representa o preço de referência e  $Vol_t^{(y)}$  corresponde à volatilidade anualizada. Essas expressões definem limites superiores e inferiores para o preço com base em um valor de referência e na volatilidade estimada. Em termos econômicos, as bandas representam desvios percentuais esperados em torno desse preço, de modo que sua ultrapassagem pode ser interpretada como rompimento de uma faixa estatisticamente relevante de oscilação.

Para os fins deste artigo, adotaram-se os parâmetros e transformações sugeridos por [Guerra \(2022\)](#). Inicialmente, estimou-se a volatilidade diária com base em uma janela móvel de 20 pregões. Em seguida, essa medida foi anualizada pela multiplicação pela raiz quadrada de 252 dias de negociação, isto é:

$$Vol_t^{(y)} = \sigma_t^{(d)} \sqrt{252} \quad (4)$$

em que  $\sigma_t^{(d)}$  representa o desvio-padrão diário dos retornos. A Figura 2 ilustra as bandas anuais obtidas a partir desse procedimento.

**Figura 2**  
**Zonas de suporte e resistência com base na volatilidade anual (2025)**  
**(referência calculada a partir de 2024)**



A partir da volatilidade anual, derivou-se a volatilidade semanal pela mesma regra de escalonamento temporal, mediante divisão da medida anual pela raiz quadrada de 52 semanas:

$$Vol_t^{(w)} = \frac{Vol_t^{(y)}}{\sqrt{52}} \quad (5)$$

Com isso, as bandas em frequência semanal são definidas por:

$$Supply_t^{(w)} = \left(1 + Vol_t^{(w)}\right) \cdot P_t^{ref} \quad (6)$$

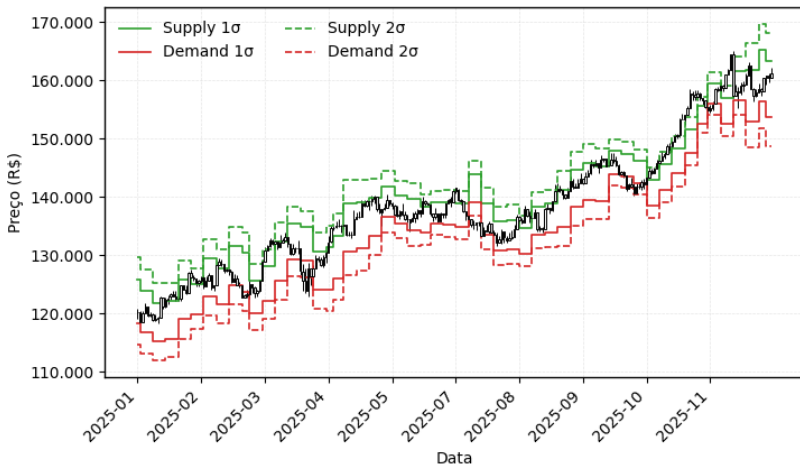
$$Demand_t^{(w)} = \left(1 - Vol_t^{(w)}\right) \cdot P_t^{ref} \quad (7)$$

Esse procedimento preserva a consistência entre anualização e conversão para frequência semanal, uma vez que todas as medidas derivam de uma

mesma base diária e são convertidas entre horizontes temporais pela regra da raiz do tempo. A adoção da frequência semanal é particularmente útil neste artigo, pois permite captar com maior sensibilidade episódios de aproximação e rompimento das bandas ao longo do ano, sem depender exclusivamente de uma referência anual fixa.

Por fim, as grandezas foram expressas em termos percentuais, o que facilita sua interpretação econômica e sua utilização empírica na identificação de rompimentos das zonas de suporte e resistência. A Figura 3 apresenta as bandas semanais construídas a partir desse procedimento.

**Figura 3**  
**Zonas de suporte e resistência com base na volatilidade semanal (2025)**



### 3.2 Extração da base

A coleta de dados foi realizada por meio da plataforma *Apify*, a partir de contas identificadas nos relatórios *FInfluence* da ANBIMA (ANBIMA, 2020-2026). Para cada perfil selecionado, foram extraídas as postagens públicas publicadas no X (antigo *Twitter*) ao longo do período de análise. A Tabela 1 apresenta a distribuição do número de postagens por conta e sua participação relativa no total de observações coletadas.

A base consolidada inclui, entre outras, as seguintes variáveis principais (Tabela A1): *userName* (nome de usuário), *createdAt* (data e horário da publicação), *text* (conteúdo textual da postagem), *fullText* (texto completo), *retweetCount* (número de compartilhamentos), *replyCount* (número de respostas),

*likeCount* (número de curtidas), *quoteCount* (número de citações), *viewCount* (número de visualizações), *isReply* (indicador de resposta a outra publicação) e *isQuote* (indicador de citação de outra publicação).

A distribuição das postagens revela concentração da atividade em conjunto relativamente restrito de contas, com destaque para *ZattarRafael*, *criptofacil* e *crisinveste*. Esse resultado, contudo, não implica que as conclusões do artigo sejam determinadas exclusivamente por esses perfis. A amostra analisada não corresponde ao universo completo de influenciadores financeiros atuantes na plataforma, mas a um subconjunto selecionado a partir das contas identificadas nos relatórios da ANBIMA. Além disso, maior frequência de publicação não implica, necessariamente, maior relevância informacional, uma vez que a estratégia empírica adotada neste estudo também incorpora métricas de engajamento e filtros temáticos, o que reduz a influência mecânica do volume total de postagens sobre os resultados.

**Tabela 1**  
**Principais contas do X por número de postagens e participação**

Usuário	Quantidade	Participação
ZattarRafael	8.385	21,5%
criptofacil	7.242	18,6%
crisinveste	5.697	14,6%
portaldobitcoin	3.842	9,9%
mendlowicz	3.552	9,1%
capitalizo	3.510	9,0%
Felipe_Hermes	3.252	8,4%
fernandoulrich	734	1,9%
nathfinancas	700	1,8%
faveladoinvest	566	1,5%
samydana	562	1,4%
jamargonyc	335	0,9%
PeriniBruno	275	0,7%
ThiagoNigro	112	0,3%
MePoupenaweb	106	0,3%
Riquezaemdias	35	0,1%
dinheirocomvoce	11	0,03%
HermannGreb	1	0,003%

Número de postagens e participação relativa das contas mais ativas no X (antigo *Twitter*) na base referente aos anos de 2024 e 2025.

### 3.3 Classificação temática e de sentimento

As etapas de classificação temática e de sentimento foram conduzidas com o auxílio de um modelo de linguagem de grande porte (*GPT-4o-mini*). A adoção desse procedimento está em linha com a literatura recente que destaca o potencial de modelos de linguagem na análise textual, especialmente em tarefas que exigem compreensão contextual e interpretação semântica mais sofisticada. Em particular, Long et al. (2025) mostram que a utilização de *Large Language Models* (LLMs) na rotulagem de textos financeiros pode contribuir para a construção de bases anotadas de maior qualidade e para o aprimoramento de procedimentos subsequentes de classificação.

Como verificação adicional, realizou-se uma etapa de validação manual da classificação automatizada. Essa estratégia está em linha com trabalhos recentes que combinam rotulagem automatizada por modelos de linguagem e avaliação humana para examinar a consistência de classificações em contextos financeiros. Ainda, Long et al. (2025) utilizam anotações humanas e rótulos gerados por *GPT-4o-mini* na construção e avaliação de bases para análise de sentimento em mercados financeiros, o que reforça a relevância da comparação entre classificações automatizadas e referências humanas.

No presente estudo, selecionou-se uma amostra balanceada de 500 *tweets* para a validação da classificação de sentimento e outra amostra balanceada de 500 *tweets* para a validação da classificação temática. Cada publicação da amostra recebeu classificação manual e, em seguida, foi comparada ao rótulo originalmente atribuído pelo modelo. A qualidade da classificação automatizada foi avaliada por meio de métricas usuais de desempenho em problemas de classificação multiclasse: *acurácia*, *precision*, *recall*, *F1-score* e o coeficiente Kappa de Cohen.

#### 3.3.1 Classificação temática

Na primeira etapa, adotou-se uma abordagem de *zero-shot classification* para atribuir a cada publicação uma entre cinco categorias mutuamente exclusivas (Tabela 5 e Tabela A3): Economia, Mercado Financeiro, Meme, Piadas ou Outro. A classificação foi realizada com base exclusivamente no conteúdo textual da postagem.

Para essa etapa, o *system prompt* foi formulado de modo a delimitar explicitamente as classes admissíveis e o formato esperado de saída. Em termos substantivos, o modelo foi instruído a operar como classificador de *tweets* em português, de modo a distinguir publicações associadas a Economia, Mercado Financeiro, Meme, Piadas e Outro. Além disso, exigiu-se que, para cada observação, o retorno incluísse o identificador do *tweet*, um conjunto

reduzido de termos relevantes para a classificação e a classe temática final. A estrutura resumida do *prompt* utilizado é apresentada na Tabela 2.

**Tabela 2**  
**Estrutura do *prompt* utilizado na classificação temática**

<b>Conteúdo do <i>system prompt</i></b>
Você é um classificador de <i>tweets</i> em português.
Classes possíveis:
Economia (inflação, juros, PIB, macroeconomia)
Mercado Financeiro (ações, bolsa, investimentos, BBAS3)
Meme (humor viral, internet)
Piadas (trocadilhos, anedotas)
Outro (não encaixa nas categorias)
Retorne para cada <i>tweet</i> : <i>id</i> , <i>wordsToClassification</i> (3–8 termos do texto) e <i>classificationTweet</i> .
Estrutura resumida do <i>prompt</i> empregado para orientar o modelo na etapa de classificação temática das postagens.

Para cada *tweet*, o modelo retornou: (i) um identificador único (*id*); (ii) um conjunto de termos considerados relevantes para a classificação (*wordsToClassification*); e (iii) a classificação temática final (*classificationTweet*).

A Tabela 3 apresenta os resultados agregados da validação. Observa-se acurácia de 80,2%, *Macro F1* de 81,1% e Kappa de Cohen de 0,753, o que indica elevado grau de concordância entre a classificação automatizada e a classificação manual. Esses resultados sugerem que a rotulagem temática produzida pelo modelo apresenta desempenho compatível com seu uso na filtragem da base.

A Tabela 4 apresenta a decomposição do desempenho por classe. Os melhores resultados ocorreram nas classes Economia, Piadas e Meme, todas com *F1-score* superior a 83%. A classe Mercado Financeiro apresentou o menor *recall* entre as categorias, o que indica maior dificuldade do modelo em recuperar todas as publicações que, na anotação manual, pertenciam a essa classe. Esse padrão é consistente com a proximidade semântica entre Economia e Mercado Financeiro, bem como com a possibilidade de parte dos conteúdos de Mercado Financeiro se confundir com a categoria residual Outro.

Após essa etapa, mantiveram-se apenas as publicações classificadas como Economia ou Mercado Financeiro. Esse filtro resultou em 17.999 *tweets* elegíveis. Em seguida, os dados foram agregados por dia útil, originando base final com 518 observações.

**Tabela 3**  
**Métricas agregadas da validação manual da classificação temática**

Métrica	Valor
Número de observações válidas	500
Acurácia	80,2%
Macro Precision	80,3%
Macro Recall	86,6%
Macro F1	81,1%
Weighted Precision	84,5%
Weighted Recall	80,2%
Weighted F1	79,3%
Kappa de Cohen	0,753

Resultados agregados da comparação entre a classificação automatizada e a anotação manual da amostra balanceada de validação.

**Tabela 4**  
**Desempenho da classificação temática por classe**

Classe	Precision	Recall	F1-score	Support
Economia	86,0%	88,7%	87,3%	97
Mercado Financeiro	99,0%	55,1%	70,8%	178
Meme	72,0%	100,0%	83,7%	72
Piadas	80,0%	98,8%	88,4%	81
Outro	64,4%	90,3%	75,1%	72

Métricas de desempenho por classe na validação manual da classificação temática.

### 3.3.2 Classificação de sentimento

Na segunda etapa, os *tweets* elegíveis foram submetidos a um procedimento de classificação de sentimento, com instruções específicas para interpretação de contexto financeiro, linguagem figurada, ironia e uso de *emojis*.

Para essa etapa, formulou-se um *system prompt* específico, apresentado na Tabela 6, com o objetivo de orientar o modelo quanto às classes admissíveis, aos critérios de interpretação semântica e ao formato esperado de saída. Assim como na etapa anterior, o modelo foi instruído a classificar cada publicação em uma entre três categorias mutuamente exclusivas — Positivo, Negativo ou Neutro. Além disso, exigiu-se que o retorno incluísse um escore contínuo de confiança no intervalo entre 0 e 1.

Para cada publicação, o modelo retornou (Tabela 9 e Tabela A2): (i) uma classificação categórica de sentimento; e (ii) um escore contínuo de confiança no intervalo entre 0 e 1 (*scoreOpenAI*).

A Tabela 7 apresenta os resultados agregados da validação. Observa-se

**Tabela 5**  
**Distribuição das classificações temáticas**

<i>classificationTweet</i>	Quantidade	Participação
Outro	19.907	48,6%
Mercado Financeiro	13.089	32,0%
Economia	5.503	13,4%
Meme	2.146	5,2%
Piadas	276	0,7%

**Tabela 6**  
**Estrutura do *prompt* utilizado na classificação de sentimento**

Conteúdo do <i>system prompt</i>
Você é um classificador de sentimento de <i>tweets</i> em português, com foco em contexto financeiro.
Regras:
Classifique o sentimento do texto como: Positivo, Negativo ou Neutro.
Considere ironia/sarcasmo, emojis e contexto financeiro.
<i>Score</i> deve ser uma confiança entre 0.0 e 1.0.
Retorne somente JSON no formato solicitado.
Estrutura resumida do <i>prompt</i> empregado para orientar o modelo na etapa de classificação de sentimento das postagens.

acurácia de 74,8%, *Macro F1* de 75,1% e Kappa de Cohen de 0,622, o que sugere concordância substancial entre a classificação automatizada e a classificação manual. Esses resultados indicam que, embora a rotulagem automática não esteja isenta de erro, seu desempenho é compatível com o uso da classificação como base para a construção das medidas agregadas de sentimento empregadas nas etapas subsequentes da análise.

A Tabela 8 apresenta a decomposição do desempenho por classe. O melhor resultado ocorreu na identificação de mensagens negativas, com *F1-score* de 79,5%. As classes Positivo e Neutro também apresentaram desempenho satisfatório, embora a classe Neutro tenha revelado maior sobreposição com as demais categorias. Esse padrão é consistente com a dificuldade inerente de distinguir mensagens efetivamente neutras de publicações com conteúdo ambíguo, irônico ou com polaridade fraca.

Em conjunto, os resultados da validação oferecem suporte empírico adicional ao uso da classificação automatizada nesta pesquisa. A partir dessa etapa, foi possível agregar as medidas de sentimento em frequência diária e utilizá-las como variáveis explicativas nos modelos econométricos.

**Tabela 7**  
**Métricas agregadas da validação manual da classificação de sentimento**

Métrica	Valor
Número de observações válidas	500
Acurácia	74,8%
<i>Macro Precision</i>	74,8%
<i>Macro Recall</i>	77,6%
<i>Macro F1</i>	75,1%
<i>Weighted Precision</i>	76,7%
<i>Weighted Recall</i>	74,8%
<i>Weighted F1</i>	74,5%
Kappa de Cohen	0,622

Resultados agregados da comparação entre a classificação automatizada e a anotação manual da amostra balanceada de validação.

**Tabela 8**  
**Desempenho da classificação de sentimento por classe**

Classe	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	<i>Support</i>
Negativo	70,7%	90,8%	79,5%	130
Neutro	85,5%	62,8%	72,4%	226
Positivo	68,3%	79,2%	73,3%	144

Métricas de desempenho por classe na validação manual da classificação de sentimento.

### 3.4 Construção das variáveis de atenção e sentimento

A principal variável explicativa deste estudo consiste em uma proxy diária de atenção associada à atividade dos influenciadores financeiros no X. A construção dessa medida parte da hipótese de que a relevância informacional de uma publicação pode ser aproximada pelo volume de interação que ela desperta entre os usuários da plataforma. Nesse contexto, o engajamento observado em cada *tweet* é interpretado como uma proxy da capacidade de difusão, alcance e mobilização de atenção do conteúdo publicado.

A literatura recente tem utilizado métricas observáveis de interação em redes sociais como indicadores da intensidade de reação do público ao conteúdo financeiro disseminado nessas plataformas. Em particular, [Bouteska et al. \(2023\)](#) empregam estatísticas como curtidas, respostas e republicações para mensurar o *investor engagement* no *Twitter*, interpretando tais interações como sinais da repercussão e da relevância informacional das postagens. Seguindo essa lógica, este estudo define o engajamento de cada *tweet* como a soma das principais interações públicas disponibilizadas pela plataforma.

Formalmente, para cada publicação *i*, o engajamento é definido como:

**Tabela 9**  
**Distribuição das classificações de sentimento**

<i>sentimentOpenAI</i>	Quantidade	Participação
Positivo	8.461	46,97%
Negativo	6.460	35,86%
Neutro	3.092	17,17%

$$eng_i = like_i + rt_i + quote_i + reply_i + bookmark_i \quad (8)$$

em que  $eng_i$  representa o engajamento total do *tweet*  $i$ , enquanto  $like_i$ ,  $rt_i$ ,  $quote_i$ ,  $reply_i$  e  $bookmark_i$  correspondem, respectivamente, ao número de curtidas, republicações, citações, respostas e marcações de favorito. A Equação (8) sintetiza, em uma única estatística, diferentes formas de reação do público ao conteúdo publicado. Sob essa definição, valores mais elevados de engajamento indicam maior difusão, alcance e relevância informacional da postagem.

Em seguida, as observações foram agregadas em frequência diária, considera-se apenas os *tweets* previamente classificados nas categorias *Economia* e *Mercado Financeiro*. Para cada dia  $t$ , foram construídas quatro estatísticas: o número de publicações elegíveis, o engajamento total diário, a média diária do escore de sentimento e o desvio-padrão diário do escore de sentimento.

O número diário de publicações elegíveis foi definido como a contagem de *tweets* classificados nas categorias *Economia* e *Mercado Financeiro* em cada dia  $t$ . O engajamento agregado diário, por sua vez, foi obtido pela soma do engajamento individual de todas as publicações pertencentes a esse mesmo dia. A média diária do sentimento foi calculada como a média simples dos escores de sentimento atribuídos aos *tweets* elegíveis, enquanto a dispersão intradiária foi mensurada pelo desvio-padrão desses escores. Nos dias em que apenas uma publicação esteve disponível, o desvio-padrão diário do sentimento foi definido como zero, com o objetivo de preservar a consistência da série temporal e evitar valores ausentes nas etapas subsequentes da modelagem.

Como a distribuição do engajamento apresenta elevada assimetria à direita, marcada por episódios pontuais de grande repercussão, o engajamento total diário foi transformado por meio da função logarítmica  $\log(1 + x)$ , originando um índice diário de atenção:

$$A_t = \log(1 + E_t) \quad (9)$$

em que  $A_t$  representa o índice diário de atenção e  $E_t$  corresponde ao engajamento agregado no dia  $t$ . Essa transformação reduz a influência desproporcional de valores extremos sem eliminar a informação contida em picos de atenção, tornando a variável mais adequada à modelagem econométrica.

Adicionalmente, com o objetivo de captar o componente persistente da atenção e suavizar oscilações de curtíssimo prazo, foi calculada uma média móvel exponencial com janela equivalente a sete períodos para o índice de atenção. A série suavizada foi então defasada em um período, de forma a mitigar problemas de simultaneidade entre atenção e comportamento contemporâneo do mercado:

$$\tilde{A}_t = \text{EMA}_7(A_t)_{t-1} \quad (10)$$

De forma análoga, construiu-se uma medida suavizada para o sentimento médio diário:

$$\tilde{S}_t = \text{EMA}_7(\bar{S}_t)_{t-1} \quad (11)$$

em que  $\tilde{A}_t$  representa a medida suavizada e defasada de atenção, e  $\tilde{S}_t$  representa a medida suavizada e defasada de sentimento médio. No contexto deste artigo, o índice de atenção derivado do engajamento agregado constitui a principal proxy de relevância informacional produzida pelos influenciadores financeiros, ao passo que as medidas de sentimento capturam a direção média e a dispersão qualitativa do conteúdo disseminado. Em conjunto, essas variáveis permitem representar tanto a intensidade quanto o teor informacional do fluxo de publicações nas redes sociais.

## 4. Resultados e discussão

### 4.1 Evidência descritiva: engajamento, sentimento e rompimentos

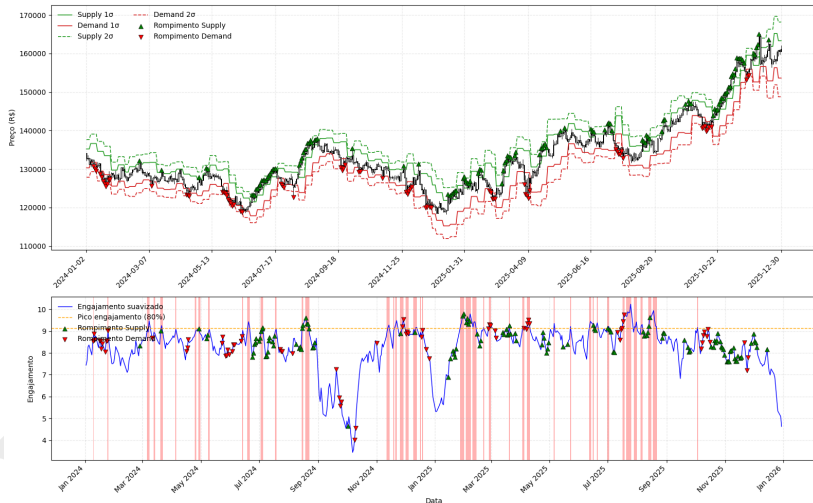
Cabe destacar que as Figuras 4 e 5 têm caráter descritivo e ilustrativo. Assim, seu objetivo é evidenciar padrões visuais da evolução conjunta entre as variáveis de interesse, sem substituir a inferência estatística apresentada na seção [Resultados](#).

A Figura 4 apresenta a evolução conjunta do Ibovespa, das bandas de suporte e resistência calibradas pela volatilidade semanal e da série de engajamento suavizado entre 2024 e 2025. Visualmente, observa-se que episódios de atenção elevada tendem a se concentrar em momentos de maior tensão do

preço em relação às zonas de suporte e resistência, sobretudo quando o índice se aproxima ou ultrapassa a banda superior. Em outras palavras, picos de engajamento não se distribuem de forma uniforme ao longo da amostra, mas se concentram em fases de maior pressão de compra ou de maior sensibilidade do mercado a choques informacionais.

Esse padrão sugere que a atenção capturada nas redes sociais pode estar associada a episódios de descoberta de preço e continuidade de movimento, ainda que a simples inspeção visual não permita distinguir coincidência temporal de associação estatística. Também se nota que nem todo pico de engajamento coincide com rompimentos, o que reforça a necessidade de complementar a evidência gráfica com testes formais e especificações econométricas.

**Figura 4**  
**Engajamento suavizado, zonas de suporte e resistência e rompimentos do Ibovespa**

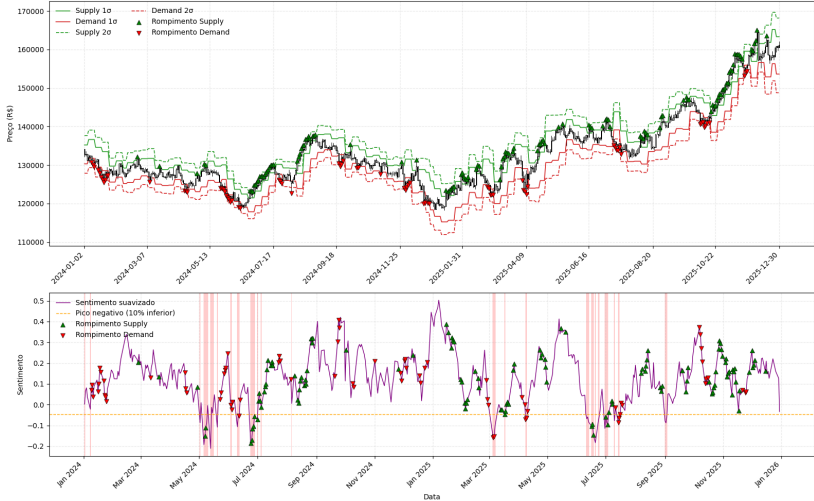


A figura combina, no painel superior, o comportamento do Ibovespa e as bandas de suporte e resistência estimadas a partir da volatilidade, e, no painel inferior, a medida suavizada de engajamento diário. As faixas verticais indicam dias classificados como pico de atenção, definidos pelo percentil anual de engajamento.

A Figura 5 complementa essa evidência ao relacionar os rompimentos com a série de sentimento suavizado. Em comparação com o engajamento, o sentimento apresenta padrão visual menos sistemático. Embora episódios de sentimento mais adverso apareçam em janelas próximas a alguns rompimentos, a associação parece mais heterogênea ao longo do tempo e menos concentrada em torno dos eventos. Esse contraste já sugere, em nível descritivo, que

a intensidade da atenção pode constituir dimensão mais robusta do que a polaridade média das mensagens para explicar a dinâmica dos rompimentos.

**Figura 5**  
**Sentimento suavizado, zonas de suporte e resistência e rompimentos do Ibovespa**



No painel inferior, a série de sentimento suavizado é comparada aos dias de rompimento das bandas de volatilidade. As faixas verticais indicam episódios de sentimento extremo no limite inferior da distribuição anual.

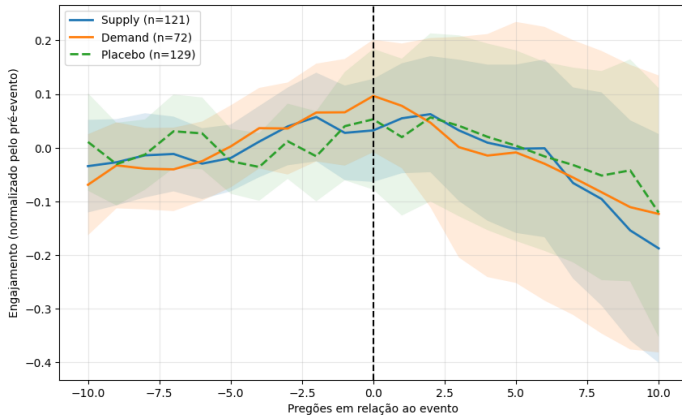
## 4.2 Estudo de eventos em torno dos rompimentos

Para investigar se os rompimentos são acompanhados por alterações sistêmicas na atenção digital, construiu-se um estudo de eventos em torno dos dias de rompimento das bandas de suporte e resistência. A Figura 6 mostra o comportamento médio do engajamento normalizado pelo período pré-evento, com comparação entre janelas associadas a rompimentos de oferta, rompimentos de demanda e datas placebo.

Os resultados indicam que a reação do engajamento se concentra em janelas curtas ao redor do evento, sobretudo entre três pregões antes e três pregões após o rompimento. Em torno de  $t = 0$ , há elevação moderada do engajamento médio tanto para rompimentos de oferta quanto para rompimentos de demanda, com maior intensidade relativa neste último caso. Ainda assim, parte desse movimento também aparece na trajetória placebo, o que recomenda cautela na interpretação dos efeitos médios em nível.

Por essa razão, a Figura 7 apresenta o excesso de engajamento em relação

**Figura 6**  
**Engajamento médio em torno dos rompimentos e datas placebo**



A figura reporta o engajamento médio normalizado pelo pré-evento em janelas de  $-10, +10$  pregões em torno dos rompimentos das bandas de suporte e resistência e de datas placebo pareadas. As áreas sombreadas representam intervalos de confiança.

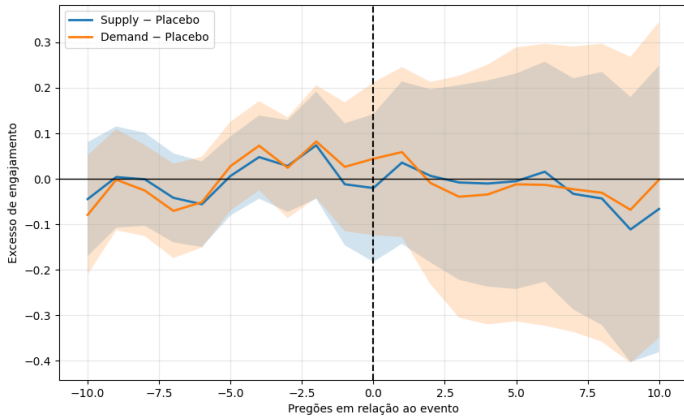
ao placebo. Nessa métrica, o diferencial entre eventos reais e placebo permanece relativamente contido e concentrado na vizinhança imediata do evento, novamente com maior nitidez no intervalo aproximado de  $[-3, +3]$  pregões. Isso sugere que a atenção digital funciona sobretudo como mecanismo de amplificação de curtíssimo prazo, mais associado ao entorno imediato do rompimento do que a um processo prolongado de antecipação.

Em conjunto, as duas figuras apontam para um padrão coerente com a hipótese de que as redes sociais operam como mecanismo de relevância informacional em torno de dias de maior tensão do preço, mas com efeito médio agregado predominantemente transitório.

### 4.3 Resultados

A análise econométrica foi conduzida por meio de modelos logísticos, tendo como variável dependente a ocorrência de rompimento das bandas de suporte ou resistência em janelas centradas ao redor de cada observação. Essa escolha é compatível com a natureza binária do problema empírico: o interesse do artigo recai sobre a probabilidade de ocorrência de um evento discreto, e não sobre a magnitude contínua da variação de preços. Nesse contexto, a especificação logística permite interpretar diretamente a associação entre variáveis informacionais e a chance de rompimento. Modelos *probit* forneceriam leitura semelhante, ao passo que modelos de duração não foram

**Figura 7**  
**Excesso de engajamento em torno dos rompimentos, em relação ao placebo**



A figura mostra a diferença entre o engajamento médio em torno dos rompimentos e o engajamento médio observado nas datas placebo. Valores positivos indicam atenção acima do padrão contrafactual.

adotados porque o foco não está no tempo até o evento, mas em sua ocorrência dentro de janelas discretas de curto prazo.

A especificação básica estimada é dada por:

$$P(\text{Breakout}_t = 1) = \Lambda \left( \beta_0 + \beta_1 \tilde{A}_{t-1} + \beta_2 \tilde{S}_{t-1}^+ + \beta_3 \tilde{S}_{t-1}^- + \beta_4 Z_t \right) \quad (12)$$

em que  $\Lambda(\cdot)$  denota a função logística,  $\text{Breakout}_t$  representa a ocorrência de rompimento na janela centrada em torno de  $t$ ,  $\tilde{A}_{t-1}$  corresponde à medida defasada de atenção,  $\tilde{S}_{t-1}^+$  e  $\tilde{S}_{t-1}^-$  representam indicadores associados a episódios de sentimento positivo e negativo extremos, e  $Z_t$  reúne os controles adicionais incluídos nas estimações, como tendência temporal, volatilidade e efeito de ano.

As janelas de 3, 5 e 7 pregões foram avaliadas como exercício de robustez. A escolha não foi arbitrária: o estudo de eventos apontou concentração da reação média da atenção digital em torno de  $[-3, +3]$ , enquanto janelas mais amplas permitiram verificar se a associação persistia sob definições menos estritas do evento. De modo semelhante, os percentis adotados para engajamento e sentimento foram tratados como definições alternativas de extremidade da distribuição, com o objetivo de avaliar a estabilidade dos resultados.

A análise distingue dois níveis de evidência. O primeiro corresponde à

inferência principal do artigo, baseada nas especificações com resultados mais estáveis e economicamente coerentes. O segundo corresponde a um exercício mais restritivo de robustez, no qual as especificações também precisam preservar desempenho preditivo fora da amostra.

No plano inferencial, a evidência mais forte permanece associada às especificações com janela centrada de  $[-7, +7]$  pregões, apresentadas na Tabela 10. Para os rompimentos da banda superior (*supply breakout*), a combinação com percentil 75 para engajamento, percentil 75 para sentimento positivo e percentil 10 para sentimento negativo produz coeficiente positivo para a variável de atenção. Para os rompimentos da banda inferior (*demand breakout*), a combinação com percentil 85 para engajamento, percentil 75 para sentimento positivo e percentil 5 para sentimento negativo gera coeficiente negativo para o engajamento. Em termos econômicos, esses resultados são consistentes com a hipótese de que picos de atenção se associam mais frequentemente a episódios de pressão compradora do que a movimentos de deterioração do mercado.

Em conjunto, essa primeira evidência sugere que o conteúdo informacional proveniente das redes sociais opera mais por meio da intensidade da atenção do que do sinal emocional isolado das mensagens. O efeito do sentimento permanece presente, mas mostra menor estabilidade entre especificações.

Quando o critério de seleção passa a exigir, simultaneamente, robustez inferencial e desempenho preditivo fora da amostra, as especificações selecionadas migram para janelas centradas de  $[-3, +3]$  pregões, conforme mostra a Tabela 11. Esse resultado é compatível com o estudo de eventos, que já havia sugerido concentração da reação média da atenção em horizontes curtos.

A não coincidência integral entre as especificações principais e as especificações da interseção não caracteriza inconsistência metodológica. Os critérios de seleção não são os mesmos: uma configuração pode ser mais informativa do ponto de vista inferencial sem ser a melhor em previsão, e o inverso também pode ocorrer. Por essa razão, a análise separa evidência principal e robustez complementar.

Nas especificações finais da interseção, os coeficientes de atenção e sentimento deixam de apresentar a mesma força observada na evidência principal. Para os rompimentos de oferta, a combinação vencedora utiliza janela centrada de  $[-3, +3]$  pregões, percentil 95 para engajamento, percentil 95 para sentimento positivo e percentil 25 para sentimento negativo. Para os rompimentos de demanda, a combinação selecionada utiliza janela centrada de  $[-3, +3]$  pregões, percentil 75 para engajamento, percentil 95 para sentimento positivo e percentil 10 para sentimento negativo. Em ambos os casos, os sinais permanecem economicamente interpretáveis, mas a evidência estatística é mais

**Tabela 10**  
**Especificações logísticas centrais para rompimentos de suporte e resistência**

<b>Rompimento</b>	<b>Janela</b>	<b>Eng.</b>	<b>Sent.+</b>	<b>Sent.-</b>	<b>Coef. Eng.</b>	<b>Coef. Sent.+</b>	<b>Coef. Sent.-</b>	<b>Pseudo-<math>R^2</math></b>	<b>AIC</b>	<b>BIC</b>
<i>Supply</i> (1 $\sigma$ )	7	75	75	10	1,005*** (0,002)	0,587*** (0,045)	1,005** (0,039)	0,167	435,419	464,751
<i>Demand</i> (1 $\sigma$ )	7	85	75	5	-0,759*** (0,007)	-0,863*** (0,000)	-1,273*** (0,006)	0,115	607,251	636,583

As colunas Eng., Sent.+ e Sent.- indicam os percentis anuais utilizados para definir picos de engajamento, sentimento positivo e sentimento negativo. \* Nível de significância de 10%. \*\* Nível de significância de 5%. \*\*\* Nível de significância de 1%.

fraca.

Em síntese, os resultados apontam para uma leitura em dois níveis. A evidência principal é mais favorável às especificações com janela centrada de  $[-7, +7]$  pregões. A robustez complementar desloca a seleção para janelas centradas de  $[-3, +3]$  pregões, em linha com o estudo de eventos, mas sem preservar a mesma força dos resultados principais. Assim, os achados são consistentes com a hipótese de que a informação oriunda das redes sociais é mais relevante em horizontes curtos, embora sua interpretação dependa do critério de seleção adotado.

Do ponto de vista econômico, os resultados são compatíveis com um ambiente em que influenciadores financeiros ampliam a visibilidade de temas e narrativas já em circulação. Nessa interpretação, o engajamento agregado não atua como determinante autônomo da direção dos preços, mas como mecanismo de relevância e propagação informacional em episódios de maior tensão de mercado.

#### 4.4 Discussão

Em conjunto, os resultados sugerem que a atividade dos influenciadores financeiros no X contém informação útil para compreender episódios de rompimento das zonas de suporte e resistência do mercado acionário brasileiro. A principal contribuição empírica do artigo reside em mostrar que medidas agregadas de atenção digital, construídas a partir do engajamento observado nas postagens, apresentam associação sistemática com a dinâmica de rompimentos do Ibovespa.

Esse resultado é particularmente relevante porque desloca o foco da literatura de sentimento agregado para uma noção mais ampla de ambiente informacional, na qual a intensidade de circulação e reação ao conteúdo importa tanto quanto sua polaridade. No contexto analisado, a atenção digital parece funcionar como variável de amplificação: em nível descritivo, sua reação se concentra em janelas curtas, especialmente entre três pregões antes e três pregões após o rompimento; na modelagem econométrica, sua associação com os rompimentos mostra maior estabilidade quando observada em janelas de sete pregões.

Ao mesmo tempo, a evidência não sugere poder explicativo irrestrito. Os estudos de evento mostram que o excesso de engajamento em torno dos rompimentos é relativamente modesto e concentrado no entorno imediato do evento, enquanto os resultados envolvendo sentimento apresentam menor estabilidade entre especificações. Esses achados recomendam interpretar a atenção digital como fator complementar, e não como substituto de fundamentos, liquidez ou

**Tabela 11**  
**Especificações selecionadas no exercício de robustez complementar**

<b>Rompimento</b>	<b>Janela</b>	<b>Eng.</b>	<b>Sent.+</b>	<b>Sent.-</b>	<b>Coef. Eng.</b>	<b>Coef. Sent.+</b>	<b>Coef. Sent.-</b>	<b>Pseudo-R<sup>2</sup></b>	<b>AIC</b>	<b>BIC</b>
<i>Supply</i> (1 $\sigma$ )	3	95	95	25	0,575 (0,193)	-0,057 (0,895)	0,048 (0,834)	0,069	641,088	670,421
<i>Demand</i> (1 $\sigma$ )	3	75	95	10	0,001 (0,998)	-0,241 (0,619)	-0,102 (0,764)	0,052	605,724	635,057

As colunas Eng., Sent.+ e Sent.- indicam os percentis anuais utilizados para definir picos de engajamento, sentimento positivo e sentimento negativo. \* Nível de significância de 10%. \*\* Nível de significância de 5%. \*\*\* Nível de significância de 1%.

condições macroeconômicas.

Assim, a principal leitura substantiva deste estudo é que choques de atenção gerados nas redes sociais estão associados à dinâmica de rompimentos técnicos do mercado brasileiro, mas seu efeito parece operar sobretudo como mecanismo de reforço e difusão de informação em janelas de curto prazo.

## 5. Conclusão

Este artigo investigou se a atividade de influenciadores financeiros no X (antigo *Twitter*) está associada a rompimentos de curto prazo nas zonas de suporte e resistência do Ibovespa. Para isso, combinou-se a extração de postagens públicas de influenciadores identificados nos relatórios *FInfluence* da ANBIMA com um procedimento de classificação temática e de sentimento assistido por modelo de linguagem de grande porte, seguido da construção de medidas diárias de engajamento e sentimento agregados.

Os resultados indicam que a dimensão mais consistente do ambiente informacional capturado nas redes sociais é o engajamento agregado das publicações. Em nível descritivo, os estudos de evento sugerem que a reação da atenção digital se concentra em torno de janelas curtas, particularmente entre três pregões antes e três pregões após o rompimento. Em nível econométrico, contudo, a associação entre atenção e rompimentos mostrou-se mais estável em janelas mais amplas, de sete pregões. De modo geral, picos de atenção estiveram associados a maior probabilidade de rompimentos da banda superior e a menor probabilidade de rompimentos da banda inferior.

Esses achados contribuem para a literatura ao mostrar que, no contexto do mercado acionário brasileiro, a atenção gerada por influenciadores digitais parece atuar menos como determinante autônomo da direção dos preços e mais como mecanismo de relevância e amplificação informacional. Assim, o artigo desloca o foco da análise exclusiva de sentimento para uma abordagem mais ampla do ambiente informacional, em que a intensidade de circulação e reação ao conteúdo também desempenha papel relevante.

Em termos teóricos, os resultados são difíceis de acomodar não apenas em versões estritas da hipótese de eficiência de mercado, mas também na ideia de passeio aleatório associada à forma fraca de Fama. Se choques de atenção oriundos de redes sociais se associam de maneira sistemática a episódios de rompimento em horizontes curtos, então parte da dinâmica dos preços parece refletir mecanismos de difusão, saliência e coordenação informacional que escapam à noção de que novas variações de preços seguem trajetória essencialmente imprevisível a partir da informação observável. Nessa perspectiva, a evidência apresentada reforça argumentos da literatura comportamental

segundo os quais a atenção limitada, os filtros cognitivos e a influência de agentes com elevada exposição pública podem afetar o processo de formação de preços mesmo em mercados organizados.

O estudo apresenta, contudo, algumas limitações. Em primeiro lugar, a amostra está restrita a influenciadores selecionados a partir dos relatórios da ANBIMA, o que não esgota o universo de agentes relevantes na formação de atenção digital. Em segundo lugar, a mensuração de engajamento e sentimento depende das informações públicas disponibilizadas pela plataforma e de procedimentos automatizados de classificação textual. Em terceiro lugar, a estratégia empírica identifica associações estatísticas entre atenção e rompimentos, mas não permite afirmar causalidade estrita.

Ainda assim, os resultados oferecem implicações práticas relevantes. Para formuladores de política pública e reguladores, a evidência sugere que o ambiente digital passou a integrar de forma mais direta o processo de circulação de informação financeira, o que reforça a importância de iniciativas voltadas à educação financeira e ao aprimoramento da capacidade dos investidores de interpretar informações de mercado, sobretudo no que se refere à leitura de conteúdos disseminados por influenciadores em redes sociais. Também se torna pertinente acompanhar com maior atenção os mecanismos de difusão, amplificação e repercussão de mensagens financeiras em plataformas digitais, uma vez que esses canais podem alterar a intensidade com que determinados temas alcançam o público investidor.

Para investidores e participantes de mercado, os achados indicam que métricas de engajamento podem conter informação útil sobre episódios de maior sensibilidade do mercado a choques informacionais de curto prazo. Isso não implica que postagens de influenciadores devam ser interpretadas como sinais autônomos e suficientes para a tomada de decisão, mas sugere que medidas de atenção digital podem ser incorporadas como variáveis complementares em estratégias de monitoramento, gestão de risco e identificação de contextos de maior probabilidade de ruptura técnica. Em termos práticos, os resultados também abrem espaço para investigar se esse tipo de informação pode ser explorado na concepção de estratégias economicamente úteis ou potencialmente lucrativas, desde que esse uso seja avaliado com cautela, custos de transação, testes fora da amostra e critérios robustos de validação empírica.

Pesquisas futuras podem expandir a análise para outros índices, ativos individuais e plataformas digitais, bem como explorar heterogeneidade entre perfis de influenciadores, tipos de mensagem e regimes de mercado. Também parece promissor investigar interações entre atenção digital, liquidez e volatilidade intradiária, aprofundar a compreensão dos canais pelos quais redes

sociais afetam a dinâmica de preços em mercados emergentes e avaliar em que medida essas informações podem ser convertidas em sinais economicamente úteis para estratégias quantitativas e sistemas de alerta de mercado.

**Agradecimentos** O autor agradece à sua esposa, Stefany Silva Rolim de Oliveira, e ao seu filho, José Rafael Rolim de Oliveira, pelo apoio e pela compreensão ao longo da elaboração deste trabalho. Agradece, ainda, ao Prof. Dr. Francisco Gildemir Ferreira da Silva, pela confiança depositada desde a graduação. O autor também registra seu agradecimento aos Profs. Drs. Victor Aguiar Evangelista de Farias e Julio Fernando Seara Sequeira da Mota Lobão pela oportunidade acadêmica e pelo incentivo à realização desta pesquisa.

**Conflito de interesse** Os autores declaram que não há conflito de interesse.

**Inteligência Artificial** Esta pesquisa utilizou ferramentas de inteligência artificial para na redação do manuscrito e parte da geração de figuras. Todo o conteúdo gerado com o apoio dessas ferramentas foi criticamente revisado e validado pelos autores, de modo a assegurar sua precisão e conformidade com a integridade científica do estudo. O uso de inteligência artificial seguiu diretrizes éticas, transparência e aderência aos padrões acadêmicos. Os autores declara que o uso dessas ferramentas não comprometeu a originalidade nem a integridade do trabalho.

**Disponibilidade de dados** Todos os dados utilizados neste estudo estão disponíveis mediante solicitação razoável aos autores correspondente.

## Referências

Agarwal, P. e Agarwal, R. (2025). *A longer-term evaluation of information releases by influential market agents and the semi-strong market efficiency*, *Journal of Behavioral Finance*, 24, 385–402.

ANBIMA (2020-2026). *Finfluence: influenciadores de finanças no brasil*, *Technical Report 1-9*, Associação Brasileira das Entidades dos Mercados Financeiro e de Capitais.

**URL:** [https://www.anbima.com.br/pt\\_br/especial/influenciadores-de-investimentos.htm](https://www.anbima.com.br/pt_br/especial/influenciadores-de-investimentos.htm)

Ante, L. (2022). *How elon musk's twitter activity moves cryptocurrency markets*, *Technological Forecasting and Social Change*, 186.

- Baker, M. e Wurgler, J. (2006). [Investor sentiment and the cross-section of stock returns](#), *The Journal of Finance*, 61, 1645—1680.
- Barberis, N. e Thaler, R. (2003). [A survey of behavioral finance](#), *Handbook of the Economics of Finance*, 9(3), 1053–1128.
- Bollen, J., Mao, H. e Zeng, X. (2011). [Twitter mood predicts the stock market](#), *Journal of Computational Science*, 2, 1–8.
- Bouteska, A., Hajek, P., Abedin, M. e Dong, Y. (2023). [Effect of twitter investor engagement on cryptocurrencies during the covid-19 pandemic](#), *Research in International Business and Finance*, 64, 101850.
- Costa, F., Fortuna, N. e Lobão, J. (2023). [Herding states and stock market returns](#), *Research in International Business and Finance*, 68, 102163.
- Cunha, A. M., Lobão, J. e Alves, J. (2025). [Social media, market regulation and ceo influence: Lessons for market efficiency](#), *Corvinus Journal of Sociology and Social Policy*, 16(1), 3–27.
- Fama, E. (1970). Efficient capital markets: a review of theory and empirical work, *The Journal of Finance*, 25, 383–417.  
**URL:** <https://doi.org/10.1111/j.1540-6261.1970.tb00518.x>
- Fan, W. e Gordon, M. D. (2014). The power of social media analytics, *Communications of the ACM*, 57, 2.  
**URL:** <https://dl.acm.org/doi/epdf/10.1145/2602574>
- Guerra, L. (2022). Supply and demand levels forecasting based on returns volatility, Disponível em: [<https://om-qs.com/quantlab/blog/1578/>].
- Kahneman, D. e Riepe, M. W. (1998). [Aspects of investor psychology](#), *The Journal of Portfolio Management*, 10(1), 52–65.
- Lee, W. Y., Jiang, C. X. e Indro, D. C. (2002). [Stock market volatility, excess returns, and the role of investor sentiment](#), *Journal of Banking & Finance*, 26, 2277–2299.
- Lobão, J. e Almeida, B. (2024). [Exploring market-wide herding behavior in the major stock markets of latin america](#), *Academia Revista Latinoamericana de Administración*, 37(4), 601–616.
- Lobão, J., Pacheco, L. e Naia, M. B. (2024). [Navigating the herd: The dynamics of investor behavior in the brazilian stock market](#), *Quantitative Finance and Economics*, 8(3), 635–657.

- Long, S. C., Xie, Y., Zhou, Z., Lucey, B. e Urquhart, A. (2025). From whales to waves: Social media sentiment, volatility, and whales in cryptocurrency markets, *The British Accounting Review*, p. 101682.
- Machus, T., Mestel, R. e Theissen, E. (2022). Heroes, just for one day: the impact of donald trump’s tweets on stock prices, *Journal of Behavioral and Experimental Finance*, 33, 100594.
- Menkhoﬀ, L. (2010). The use of technical analysis by fund managers: International evidence, *Journal of Banking & Finance*, 34(11), 2573–2586.
- Simon, H. A. (1955). A behavioral model of rational choice, *The Quarterly Journal of Economics*, 7(8), 99–118.
- Sprenger, T. O. e Welppe, I. M. (2011). News or noise? the stock market reaction to different types of company-specific news events, *Working Paper 1734632*, SSRN.  
URL: <https://ssrn.com/abstract=1734632>
- Sul, H. K., Dennis, A. R. e Yuan, L. (2017). Trading on twitter: Using social media sentiment to predict stock returns, *Decision Sciences*, 48, 454–488.
- Taylor, M. P. e Allen, H. (1992). The use of technical analysis in the foreign exchange market, *Journal of International Money and Finance*, 11(3), 304–314.
- Tversky, A. e Kahneman, D. (1979). Prospect theory: An analysis of decision under risk, *The Econometric Society*, 47, 263–291.
- Yuexin Mao, Wei Wei, B. W. e Liu, B. (2012). Correlating sp 500 stocks with twitter data, *Processing of the First ACM International Workshop on Hor Topics on Interdisciplinary Social Network Research*, pp. 69–72.
- Zhang, X., Fuehres, H. e Gloor, P. A. (2011). Predicting stock market indicators through twitter “i hope it is not as bad as i fear”, *Procedia - Social and Behavioral Sciences*, pp. 55–62.
- Šević, J. S., Stakić, N. e Stakić, A. J. (2023). Impact of twitter on stock market performance: Evidence from elon musk and tesla, *1st International Conference on Innovation in Information Technology and Business (ICIITB 2022)*, pp. 47–52.

## A. Tabelas e Figuras

Nesta seção, apresentam-se tabelas e figuras complementares que detalham aspectos operacionais e ilustrativos da base de dados e dos procedimentos de classificação adotados no artigo. O objetivo é oferecer maior transparência quanto às etapas de coleta, rotulagem temática e classificação de sentimento das postagens, além de documentar exemplos representativos das observações utilizadas na análise empírica. Esse material suplementar busca facilitar a replicabilidade do estudo e aprofundar a compreensão dos dados que fundamentam os resultados apresentados no texto principal.

DRAFT

**Tabela A1**  
**Exemplos de publicações extraídas e respectivas métricas de engajamento**

userName	createdAt	text	retweet Count	reply Count	like Count	quote Count	view Count
ZaitarRafael	Thu Dec 28 19:08:14 +0000 2023	#ZFlix Já morou de favor. Hoje ele é multimilionário 41 anos Casado, 2 filhos. Família humilde. Ele tem 2 faculdades. Comprou a empresa que começou trabalhando em uma posição inicial. Começou a investir em 2018 com apenas 1,5K no Tesouro Direto. Patrimônio: prédio comercial; 4M; apartamento: 800k; casa: 2,5M; carros: 400k; investimento: 1,5M. O que você diria para ele?!	2	28	186	2	49,775
ZaitarRafael	Thu Dec 28 21:52:24 +0000 2023	Quem investe tentando adivinhar o curto prazo quase sempre se ferra. Independente do que seja. <a href="https://t.co/662vBEUXPo">https://t.co/662vBEUXPo</a>	0	8	30	0	6,373
ZaitarRafael	Thu Dec 28 23:30:53 +0000 2023	Qual foi o seu melhor investimento de 2023?	3	280	112	7	56,468
ZaitarRafael	Thu Dec 28 20:33:00 +0000 2023	Um professor meu de Matemática Financeira dizia que a probabilidade de vc ganhar na mega-sena jogando é a mesma de não jogar. É tão difícil ganhar que, estatisticamente, a chance é a mesma de vc achar um bilhete no chão da rua com os números certos. O que vocês acham?	40	492	1,926	52	254,028
ZaitarRafael	Wed Dec 27 13:47:34 +0000 2023	Dólar no menor patamar dos últimos anos e ainda dá pra ganhar até 31 dólares na conta de presente pra começar bem 2024! Já tem seguidor aproveitando a promoção de aniversário da Nomad. É por tempo limitado! Temos uma parceria e TODO MUNDO que se tornar cliente e utilizar o cupom CARTEIRAGANOS pode comprar dólar com 100% de desconto no JOF e na taxa operacional na primeira operação de câmbio até US 1000. Esse desconto será aplicado na forma de cashback. Abra sua conta no link <a href="https://t.co/maftrUXMco">https://t.co/maftrUXMco</a> e não esqueça do cupom CARTEIRA3ANOS.	1	9	56	0	23,285

**Tabela A2**  
**Exemplos de tweets classificados quanto ao sentimento**

sentimentOpenAI	fullText	userName	createdAt
Negativo	Saindo de 60 graus do Rio pra -5 graus.	ZattarRafael	Thu Mar 21 18:53:55 +0000 2024
Neutro	O mercado é f&#amp;#223; e parceiro. Volatilidade. Rapaz... suporte e resist&#223;ncia... o cara precisa se destacar mto pra dar bom. <a href="https://t.co/RAXLnh4SmN">https://t.co/RAXLnh4SmN</a>	ZattarRafael	Fri Mar 14 00:58:29 +0000 2025
Positivo	O cara ganha entre 50 e 150 pau por m&#223;s e tinha R\$ 2 milh&#225;es aplicados no CDB do ITA&#227;. Eu fiz um v&#227;deo mostrando como eu alocaria essa grana. Feij&#223;o e arroz bem feito ganha da maioria dos portf&#225;lios cheios de "FRU FRU". Coloquei no v&#227;deo abaixo minha vis&#225;o.	ZattarRafael	Tue Apr 16 16:43:50 +0000 2024

**Tabela A3**  
**Exemplos de tweets classificados tematicamente**

classificationTweet	#Zlhx	fullText	userName	createdAt
Economia		Pegou todo dinheiro que tinha e foi tentar a sorte na cidade grande. 23 anos Trabalha com Tecnologia da Informação (TI) Renda: 4,6k + beneficios	ZattarRafael	Fri Feb 16 22:35:51 +0000 2024
Meme		De família humilde de uma comunidade no interior do Rio. Começou a trabalhar aos 14 anos. Sempre foi de juntar <a href="https://t.co/LuAIMstlJ7">https://t.co/LuAIMstlJ7</a>	crinveste	Wed May 28 20:46:29 +0000 2025
Mercado Financeiro		Kkkkkkkkkk <a href="https://t.co/V15US2CfrR">https://t.co/V15US2CfrR</a> Com quantos anos vc descobriu que dos 65% dos americanos que investem em bolsa americana, grande parte não sabe que investe em bolsa americana?	ZattarRafael	Tue Apr 30 01:12:42 +0000 2024
Outro		Quem está errado? <a href="https://t.co/gp2pSybTo">https://t.co/gp2pSybTo</a>	ZattarRafael	Sat Mar 09 12:29:41 +0000 2024