

UNIVERSIDADE ESTADUAL DE MARINGÁ
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS
DEPARTAMENTO DE CIÊNCIAS CONTÁBEIS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS CONTÁBEIS
ÁREA DE CONCENTRAÇÃO: CONTROLADORIA

BEATRIZ AMARAL TUFURETI

**ALÉM DO *FEED*: RELAÇÃO ENTRE TWITTER E COMPORTAMENTO DAS
AÇÕES DE EMPRESAS BRASILEIRAS**

Maringá

2024

BEATRIZ AMARAL TUFURETI

**ALÉM DO *FEED*: RELAÇÃO ENTRE TWITTER E COMPORTAMENTO DAS
AÇÕES DE EMPRESAS BRASILEIRAS**

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Contábeis, área de concentração Controladoria, da Universidade Estadual de Maringá como requisito parcial para obtenção do título de Mestre em Ciências Contábeis.

Orientadora: Prof(a). Dr(a). Joyce Menezes da
Fonseca Tonin

Coorientador: Prof. Dr. Julyerme Matheus
Tonin

Maringá

2024

Dados Internacionais de Catalogação-na-Publicação (CIP)
(Biblioteca Central - UEM, Maringá - PR, Brasil)

T915a

Tufureti, Beatriz Amaral

Além do *Feed* : relação entre Twitter e comportamento das ações de empresas brasileiras / Beatriz Amaral Tufureti. -- Maringá, PR, 2024.
111 f. : il. color., figs., tabs.

Orientadora: Profa. Dra. Joyce Menezes da Fonseca Tonin.

Coorientador: Prof. Dr. Julyerme Matheus Tonin.

Dissertação (mestrado) - Universidade Estadual de Maringá, Centro de Ciências Sociais Aplicadas, Departamento de Ciências Contábeis, Programa de Pós-Graduação em Ciências Contábeis, 2024.

1. Twitter. 2. Sentimento. 3. Mercado acionário - Brasil. 4. Eficiência informacional. I. Tonin, Joyce Menezes da Fonseca, orient. II. Tonin, Julyerme Matheus, coorient. III. Universidade Estadual de Maringá. Centro de Ciências Sociais Aplicadas. Departamento de Ciências Contábeis. Programa de Pós-Graduação em Ciências Contábeis. IV. Título.

CDD 23.ed. 657

DECLARAÇÃO

Aos dezoito dias do mês de dezembro do ano de dois mil e vinte e quatro, às 14h00min., no Bloco B-12, sala 010, realizou-se, no *Campus* Sede da Universidade Estadual de Maringá, a Defesa Pública da Dissertação de Mestrado, sob o título: “**Além do Feed: Relação Entre Twitter e Comportamento das Ações de Empresas Brasileiras**”, de autoria de **Beatriz Amaral Tufureti**, aluna do Programa de Pós-Graduação em Ciências Contábeis – Mestrado – Área de Concentração: Controladoria - Linha de pesquisa em Contabilidade para Usuários Externos. A Banca Examinadora foi composta pelos professores:

Membros da Banca	Função	IES
Prof. ^a Dr. ^a Joyce Menezes da Fonseca Tonin	Presidente	PCO/UEM
Prof. Dr. Julyerme Matheus Tonin	Coorientador	PCE/UEM
Prof. ^a Dr. ^a Kelly Cristina Mucio Marques	Examinadora	PCO/UEM
Prof. Dr. Wilson Nakamura	Examinador	UPM

Concluídos os trabalhos de apresentação e arguição, a aluna foi **APROVADA** pela Banca Examinadora. A validação da aprovação na Defesa Pública está condicionada à aluna apresentar a versão definitiva da Dissertação, no prazo de 60 (sessenta) dias, de acordo com Art. 72 da Resolução nº 095/2018-CI/CSA, para a expedição do Diploma de Mestre. Para constar, a presente Ata foi lavrada e assinada pelo Coordenador do Programa e pelos membros da Banca Examinadora.

Maringá Pr., 18 de dezembro de 2024.


PROF. DR. ROMILDO DE OLIVEIRA MORAES
Coordenador do Programa de Pós-Graduação em
Ciências Contábeis

DEDICATÓRIA

Dedico este trabalho à minha mãe, Vera, ao meu pai, Almir, ao meu irmão, Geovane, e à minha cunhada, Rosângela, por todo o amor, apoio e inspiração. E ao meu amor, Leonardo, por ser minha fortaleza e por compartilhar comigo os sonhos e desafios desta jornada chamada vida.

AGRADECIMENTOS

A realização desta dissertação marca o encerramento de uma etapa desafiadora e enriquecedora, que só foi possível graças à colaboração e ao apoio de muitas pessoas.

Agradeço a Deus por me guiar e fortalecer ao longo desta jornada. À minha mãe, Vera, e ao meu pai, Almir, pelo apoio constante e confiança no meu potencial. Ao meu irmão, Geovane, cuja motivação foi fundamental para minha decisão de ingressar no mestrado.

Ao meu companheiro de vida, Leonardo, meu parceiro em todos os momentos, obrigada por estar ao meu lado em cada desafio, vibrar com as conquistas e oferecer apoio nos dias difíceis.

À Profa. Dra. Joyce, minha orientadora, e ao Prof. Dr. Julyerme, meu coorientador, pela orientação e dedicação durante todo o processo. Agradeço também à banca examinadora, composta pela Profa. Dra. Kelly e pelo Prof. Dr. Wilson, pelas valiosas contribuições e sugestões. Um agradecimento especial à Profa. Dra. Deisy Cristina, que foi minha supervisora no estágio de docência.

Aos amigos que o mestrado me trouxe, como Samuel Belinato, Leonardo Rossini, Bruna, Marya, Suzy e tantos outros colegas, minha sincera gratidão por tornarem essa jornada mais leve com seu apoio, parceria e amizade. Aos amigos de fora do ambiente acadêmico, Rayani, Bruna e João Paulo, meu obrigado por sua amizade e palavras de incentivo.

Aos colaboradores do Programa de Pós-Graduação e do Departamento, especialmente Leoni, Ana e Sr. Amauridio, pelo carinho e suporte essencial ao bom andamento das atividades. Agradeço ainda aos professores do programa, que contribuíram com conhecimento e inspiração ao longo dessa trajetória.

Por fim, agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo suporte financeiro, que viabilizou a realização deste mestrado.

A todos vocês, meu sincero agradecimento por contribuírem, direta ou indiretamente, para que esta conquista se tornasse realidade.

*A flecha de sol
pinta estrelas na vidraça.
Despede-se o dia*

(“Flecha de sol” por Helena Kolody)

Tufureti, B. A. (2024). *Além do feed: A relação entre o Twitter e o comportamento das ações de empresas brasileiras*. Dissertação de Mestrado em Ciências Contábeis, Universidade Estadual de Maringá, Maringá, PR, Brasil.

RESUMO

As redes sociais têm transformado a disseminação de informações no mercado acionário, permitindo interações em tempo real entre investidores e empresas, o que influencia diretamente as decisões de investimento. Este estudo buscou preencher a lacuna existente na literatura, decorrente da escassez de pesquisas sobre o impacto das informações provenientes das redes sociais nos mercados acionários de países emergentes. Para isso, investigou-se a relação entre o volume e o sentimento presente nos *tweets* com o retorno, o volume negociado e a volatilidade das ações de 24 empresas brasileiras com ADRs listados na NYSE, integrantes dos índices IBOV, IBRA e IBRX100, no período de 2022 a 2023. Trata-se de uma pesquisa descritiva e quantitativa-dedutiva, caracterizada como netnográfica e bibliográfica, com dados coletados online. Foram analisados 529.180 *tweets*, extraídos com por meio de linguagem *python*, com sentimentos atribuídos por técnicas de Processamento de Linguagem Natural (PNL). A análise incluiu a identificação de picos de atividade, retornos anormais, volume negociado e volatilidade das ações, além da aplicação de correlação de Pearson e testes de causalidade de Granger para examinar as relações entre sentimento, volume de *tweets* e métricas de desempenho acionário. Os resultados revelaram correlações significativas entre o volume de *tweets* e o volume de negociações, enquanto o impacto do sentimento diário foi mais limitado e menos consistente, especialmente em relação ao retorno diário das ações. O mercado mostrou-se eficiente em sua forma semiforte, com o sentimento positivo exercendo maior impacto em termos de amplitude e duração do que o negativo, principalmente durante picos de atividade no Twitter e anúncios de resultados. Conclui-se que o Twitter influencia aspectos específicos do mercado acionário brasileiro de maneira não uniforme, evidenciando uma relação multifacetada entre seu conteúdo e o desempenho das ações.

Palavras-chave: Twitter. Sentimento. Mercado Acionário. Eficiência Informacional. Brasil.

Tufureti, B. A. (2024). *Beyond the Feed: The relationship between Twitter and the stock performance of Brazilian companies*. Master's Dissertation in Accounting Sciences, State University of Maringá, Maringá, PR, Brazil.

ABSTRACT

Social media has transformed the dissemination of information in the stock market, enabling real-time interactions between investors and companies, directly influencing investment decisions. This study aimed to fill a gap in the literature regarding the scarcity of research on the impact of social media information on the stock markets of emerging countries. To this end, it investigated the relationship between tweet volume and sentiment with the return, trading volume, and volatility of stocks from 24 Brazilian companies with ADRs listed on the NYSE, included in the IBOV, IBRA, and IBRX100 indices, during the period from 2022 to 2023. The research is descriptive and quantitatively deductive, characterized as netnographic and bibliographic, with data collected online. A total of 529,180 tweets were analyzed, extracted using Python-based tools, with sentiments attributed through Natural Language Processing (NLP) techniques. The analysis included identifying activity peaks, abnormal returns, trading volume, and stock volatility, as well as applying Pearson correlation and Granger causality tests to examine the relationships between sentiment, tweet volume, and stock performance metrics. The results revealed significant correlations between tweet volume and trading volume, while the impact of daily sentiment was more limited and less consistent, particularly concerning daily stock returns. The market was found to be efficient in its semi-strong form, with positive sentiment having a greater impact in terms of magnitude and duration than negative sentiment, especially during Twitter activity peaks and earnings announcements. It is concluded that Twitter influences specific aspects of the Brazilian stock market in a non-uniform manner, highlighting a multifaceted relationship between its content and stock performance.

Keywords: Twitter. Sentiment. Stock Market. Informational Efficiency. Brazil.

LISTA DE ABREVIATURAS E SIGLAS

ADR	<i>American Depositary Receipt</i>
API	<i>Application Programming Interface</i>
AR	Retorno anormal
B3	Brasil, Bolsa, Balcão
CAR	Retorno anormal acumulado
CVM	Comissão de Valores Mobiliários
DJIA	<i>Dow Jones Industrial Average</i>
EA	<i>Earnings Announcements</i> (Anúncios de Resultados)
GNL	<i>Google Cloud Natural Language</i>
IBOV	Índice Bovespa
IBRA	Índice Brasil Amplo
IBRX100	Índice Brasil 100
NASDAQ	<i>National Association of Securities Dealers Automated Quotations</i>
NYSE	<i>New York Stock Exchange</i>
PNL	Processamento de Linguagem Natural
Reg FD	<i>Regulation Fair Disclosure</i>
SEC	<i>Securities and Exchange Commission</i>
VSCode	Visual Studio Code

LISTA DE FIGURAS

Figura 1 - Mercado eficiente, suas condições e categorias	23
Figura 2 - Hierarquia dos níveis de eficiência.....	23
Figura 3 - Escala de sentimento e magnitude	38
Figura 4 - Fluxograma da proposta metodológica.....	40
Figura 5 - Janela de estimação e de evento	41
Figura 6 - Volume de Tweets por Polaridade por dia	53
Figura 7 - Volume de Tweets por dia.....	54
Figura 8 - Distribuição do sentimento para os 304 picos detectados no Twitter	57
Figura 9 - Série diária do volume e picos tweets e anúncios de resultados da Gerdau	58
Figura 10 - Retornos Anormais Médios Acumulados para série temporal acionária.....	60
Figura 11 - Retornos Anormais Médios Acumulados para pico de atividade no Twitter	61
Figura 12 - Retornos Anormais Médios Acumulados para picos e anúncio de resultados	62
Figura 13 - Retornos Anormais Médios Acumulados para picos não-EA	63

LISTA DE TABELAS

Tabela 1 – Principais anomalias no mercado	25
Tabela 2 – Amostra de Empresas brasileiras de capital aberto listadas nos índices amplos da B3 e com ADRs listados na NYSE	35
Tabela 3 - Exemplos de tweets	39
Tabela 4 - Níveis de significância	43
Tabela 5 - Distribuição setor econômico, subsetor e segmento Bovespa.....	48
Tabela 6 – Volume de tweets e de dias com tweets por empresa.....	49
Tabela 7 - Estatística descritiva (01/01/2022 a 31/12/2023)	51
Tabela 8 – Matriz correlação entre Retorno e Volatilidades	52
Tabela 9 – Anúncios de Resultado	55
Tabela 10 – Estatísticas descritivas em dias de anúncio de resultado	56
Tabela 11 - Valores de AR e CAR médios para cada tipo de evento	59
Tabela 12 - Resultados dos testes de Correlação de Pearson	65
Tabela 13 - Resultados dos testes de Causalidade de Granger	67

LISTA DE EQUAÇÕES

Equação 1 – Retorno diário da ação.....	41
Equação 2 – Retorno diário do índice de mercado.....	41
Equação 3 – Retorno esperado.....	41
Equação 4 – Retorno Anormal (AR).....	42
Equação 5 – Retorno Anormal Acumulado (CAR).....	42
Equação 6 – Teste t para Retorno Anormal (AR).....	42
Equação 7 – Teste t para Retorno Anormal Acumulado (CAR).....	42
Equação 8 – Sentimento Diário.....	43
Equação 9 – Volatilidade de Parkinson.....	44
Equação 10 – Volatilidade de Garman e Klass.....	44
Equação 11 – Volatilidade de Rogers-Satchell.....	44
Equação 12 – Volatilidade de Yang-Zhang.....	44
Equação 13 – Identificação de Picos de Atividade no Twitter.....	45
Equação 14 – Correlação de Pearson.....	46
Equação 15.1 – Causalidade de Granger.....	46
Equação 15.2 – Causalidade de Granger.....	47

SUMÁRIO

LISTA DE FIGURAS.....	10
LISTA DE TABELAS.....	12
LISTA DE EQUAÇÕES.....	13
LISTA DE ABREVIATURAS E SIGLAS.....	10
SUMÁRIO.....	14
1 INTRODUÇÃO.....	16
1.1 CONTEXTUALIZAÇÃO E PROBLEMA DE PESQUISA.....	16
1.2 OBJETIVO GERAL.....	17
1.2.1 Objetivos Específicos.....	18
1.3 JUSTIFICATIVA E CONTRIBUIÇÕES.....	18
1.4 DELIMITAÇÃO E ORGANIZAÇÃO DA PESQUISA.....	20
2 FUNDAMENTAÇÃO TEÓRICA.....	21
2.1 EFICIÊNCIA INFORMACIONAL DE MERCADO.....	21
2.2 REDES SOCIAIS E MERCADO ACIONÁRIO.....	27
3 PROCEDIMENTOS METODOLÓGICOS.....	33
3.1 DEFINIÇÃO DO TIPO DE PESQUISA.....	33
3.2 AMOSTRA E COLETA DE DADOS.....	34
3.3 ATRIBUIÇÃO DE SENTIMENTO.....	37
3.4 CONSTRUTOS/VARIÁVEIS E MÉTODOS DE ANÁLISE.....	39
3.4.1 Picos de Atividade e Eventos.....	45
3.4.2 Correlação de Pearson e Causalidade de Granger.....	46
4 ANÁLISE E DISCUSSÃO DOS RESULTADOS.....	47
4.1 ESTATÍSTICA DESCRITIVA.....	47
4.2 IDENTIFICAÇÃO DOS PICOS DE ATIVIDADE NO TWITTER.....	56
4.3 ESTUDO DE EVENTOS.....	58
4.4 CORRELAÇÃO DE PEARSON E CAUSALIDADE DE GRANGER.....	64
5 CONCLUSÃO.....	68
REFERÊNCIAS.....	70
APÊNDICES.....	82
Apêndice A - 304 picos de atividade do Twitter identificados.....	82
Apêndice B – 30 picos de atividade no Twitter que coincidem com os anúncios de resultados.....	85
Apêndice C – 233 picos de atividades no Twitter não relacionados aos EA.....	86
Apêndice D – Série diária do volume de tweets, picos de tweets e anúncios de resultados.....	89

Apêndice E –Figuras: Retornos Anormais Médios	92
Apêndice F – Sentimento Diário por Setor.....	93
Apêndice G – Rotina STATA 16 – Estudo de Eventos.....	94
Apêndice H – Rotina STATA 16 – Identificação dos Picos	97
Apêndice I – Rotina STATA 16 – Retornos Anormais Médios	100
Apêndice J – Rotina STATA 16 – Correlação e Granger	107
ANEXOS.....	109
Anexo 1 – Ferramenta simplificada de coleta dos tweets via Twscrape.....	109
Anexo 2 – Ferramenta de atribuição do sentimento via Google Natural Language API ...	110

1 INTRODUÇÃO

1.1 CONTEXTUALIZAÇÃO E PROBLEMA DE PESQUISA

O uso crescente das redes sociais alterou como as informações do mercado de ações são obtidas e compartilhadas, facilitando a compreensão do seu funcionamento (Müller et al., 2023; Shen et al., 2018). A ascensão das redes sociais, aliada aos avanços computacionais e técnicas de análise de linguagem, como a análise de sentimento, têm despertado interesse da comunidade acadêmica e financeira, que vem reconhecendo as redes sociais como uma importante fonte de coleta dados e informação, incluindo para decisões de investimento (Asur & Huberman, 2010; Müller et al., 2023; Teti et al., 2019).

Diante dessas novas fontes de informação, diversas pesquisas têm analisado a relação das redes sociais e o mercado bursátil (Bollen et al., 2011; Ganesh & Iyer, 2023; Ge et al., 2018; Karabulut, 2012; Kolchyna et al., 2015; Ranco et al., 2015; Rao & Srivastava, 2012; Scaramozzino et al., 2021; Souza et al., 2019; Teti et al., 2019; Tommasetti et al., 2021; Wei et al., 2016). Nesta temática, uma das áreas de estudo se concentra na análise da relação do sentimento expresso *online* com, por exemplo, métricas de desempenho dos mercados. Neste campo de estudo, há evidências de que retornos, volume de negociação e volatilidade dos ativos financeiros podem ser afetados pela reação dos investidores às informações, indicando que o humor e o sentimento público via redes sociais podem influenciar os valores do mercado de ações tanto quanto as notícias e documentos oficiais (Bollen et al., 2011; Klibanoff et al., 1998).

A análise de sentimento é um método apropriado para captar o sentimento *online* do mercado. A partir do Processamento de Linguagem Natural (PLN), essa abordagem tem como objetivo extrair indicadores de sentimento e opiniões a partir de textos. Pode ser aplicado em mercados financeiros para identificar e prever movimentos de preços (tendência, ciclos, volatilidade etc.), analisar produtos, gerenciar crises etc. (Fan & Gordon, 2014; Mukherjee & Bhattacharyya, 2013). Assim, em contraste com meios convencionais, como pesquisas de opinião ou variáveis de mercado, que também podem apresentar limitações, observa-se uma expansão do uso de fontes não tradicionais para capturar o sentimento, com destaque para o Twitter que possui forte capacidade analítica em alta frequência (Aggarwal, 2022; Souza, 2020).

O Twitter vem se tornando popular entre *traders*, investidores, analistas financeiros, área empresarial de relação com investidores, *broadcasting*, agências de notícias, entre outros. A comunidade financeira utiliza essa ferramenta para compartilhar opiniões, análises de

mercado, conteúdo de influenciadores e educadores financeiros. Todos estes elementos, tornam a plataforma útil para obtenção de informações e identificação de tendências no mercado acionário (Souza, 2020; Wei et al., 2016). A descentralização da informação, com preservação da autonomia (Teti et al., 2019; Zenha, 2018), juntamente com a facilidade e velocidade no compartilhamento de informações (Souza, 2020), tornam o Twitter atraente para uma diversidade de usuários, bem como uma ferramenta capaz de aprimorar a eficiência do mercado.

O conceito de eficiência de mercado se consubstancia na Hipótese do Mercado Eficiente (HME), que conceitua um mercado eficiente como aquele em que todas as informações disponíveis são refletidas nos preços dos ativos, ajustando-os rapidamente (Fama, 1970, 1991). Isso indica que as informações em tempo real propagadas pelo Twitter podem influenciar a velocidade com que se repercutem nos preços dos ativos no mercado acionário. Portanto, observa-se um crescimento no número de estudos que utilizam redes sociais em busca de informações e notícias com potencial de impactar o mercado de ações (Smith & O'Hare, 2022).

Apesar do crescente interesse acadêmico na relação entre redes sociais e comportamento do mercado acionário, Agarwal et al. (2019) revelaram uma escassez de estudos que explorem como as informações disponíveis na *internet* influenciam os mercados de ações de países emergentes, como o Brasil. Nas pesquisas que exploram a análise de sentimento textual para prever o retorno de ações no mercado de capitais, há uma predominância do mercado estadunidense, enquanto países emergentes ocupam uma pequena parcela (Carvalho et al., 2023). Isso ressalta a necessidade de se compreender melhor o efeito e a relação das informações *online* com os comportamentos dos mercados não desenvolvidos.

Portanto, ao considerar a lacuna na literatura e a relevância que o Twitter tem como fonte de informação, surge o questionamento: **Qual a relação entre o conteúdo do Twitter sobre companhias de capital aberto no Brasil e o desempenho das ações destas companhias?**

1.2 OBJETIVO GERAL

O objetivo desta pesquisa consiste em identificar qual a relação do volume de *tweets* e do sentimento presente neles com o retorno, volume negociado e volatilidade das ações de 24 empresas brasileiras que têm *American Depositary Receipts* (ADR) listados na Bolsa de Valores de Nova Iorque (NYSE) e que fazem parte dos índices amplos do mercado brasileiro (IBOV, IBRA e IBRX100).

1.2.1 Objetivos Específicos

Para alcançar o propósito geral, é necessário atingir os seguintes objetivos específicos:

- a) Implementar ferramentas para extração de *tweets* e para mensuração do sentimento;
- b) Mensurar retorno e volatilidade das ações;
- c) Analisar a relação do sentimento e volume dos *tweets* com o retorno das ações, volume de ações negociada e volatilidade das ações.

1.3 JUSTIFICATIVA E CONTRIBUIÇÕES

A transição do pregão viva-voz para o pregão eletrônico representou uma transformação importante, tornando as bolsas mais transparentes, acessíveis e eficientes (Dutta et al., 2017). Essa evolução incentivou a concorrência entre corretores, reduzindo o controle sobre informações de preços e promovendo a integração global dos mercados (Bodie et al., 2015; Pinheiro, 2019). O surgimento dos *Depository Receipts* (DRs) contribuiu com o processo de globalização ao atrair investidores internacionais e para digitalização do mercado acionário (Assaf Neto, 2021; Bodie et al., 2015; Pinheiro, 2019).

Conforme o mercado financeiro se adapta à era digital, a transmissão de informações aos investidores torna-se mais ágil, acessível e abrangente, principalmente devido ao uso das redes sociais (Umar et al., 2021). Isso ocorre porque as plataformas se converteram em canais de disseminação de opiniões e ações dos seus usuários, com poder de impactar ações, títulos, moedas, *commodities*, criptomoedas etc. (Pedersen, 2022). Assim, a digitalização do mercado acionário não apenas melhorou a eficiência operacional, mas também reconfigurou a forma como a informação é consumida e compartilhada, transformando as redes sociais em importantes catalisadores.

Um exemplo da influência catalisadora das redes sociais é o caso da GameStop, ocorrido em janeiro de 2021. Por meio de plataformas como Twitter, YouTube e, sobretudo, o Reddit, investidores individuais expressaram opiniões que resultaram em uma valorização sem precedentes das ações, no volume de negociações e na volatilidade deste ativo (Pedersen, 2022; Umar et al., 2021). Em contraste com a análise tradicional baseada em fundamentos econômico-financeiros, este episódio evidenciou o efeito do sentimento dos investidores no mercado de ações (Nyakurukwa & Seetharam, 2023).

Esse evento evidencia a necessidade de aprofundamento das pesquisas sobre um tema proeminente, o sentimento nas redes sociais no contexto do mercado de ações (Nyakurukwa & Seetharam, 2023; Umar et al., 2021). Nesse contexto, o presente estudo busca realizar uma

análise abrangente sobre a relação do sentimento do investidor com o mercado acionário. Como *proxy* do sentimento do investidor, utiliza-se a informação divulgada presente em duas variáveis obtidas por meio do Twitter: sentimento e volume dos *tweets*. Por sua vez, como métrica de desempenho no mercado acionário, utiliza-se o retorno, volume de negociação e volatilidade.

A utilização do Twitter como fonte de dados é justificada por ser o principal *microblog* usado no mundo. Conforme dados de agosto de 2023, do verificador de tráfego *Cloudflare Radar* (2023), o Twitter é a rede social mais utilizada para compartilhar informações do setor financeiro. O recorte temporal consiste no período de 01/01/2022 a 31/12/2023 e é fundamentado pela estabilização da pandemia da COVID-19 e pela eleição presidencial no Brasil de 2022. A concentração em empresas brasileiras é motivada pelo destaque que o Brasil obteve como o primeiro país emergente e o terceiro global no uso do Twitter em agosto de 2023, ficando apenas atrás dos Estados Unidos e do Japão (*Cloudflare Radar*, 2023).

A abordagem amostral, em termos de recorte espacial de análise, centrada em empresas brasileiras com ADR na NYSE, justifica-se pela visibilidade internacional dessas empresas. A listagem de ativos na NYSE, a maior bolsa de valores do mundo, garante regulamentações mais rigorosas e com isso mais credibilidade para as empresas listadas (Assaf Neto, 2021; Bodie et al., 2015). Além disso, a escolha dos índices IBOV, IBRA e IBrX 100 se justifica por serem os principais índices amplos das ações negociadas na Brasil, Bolsa, Balcão, a entidade que administra o mercado em que ações e outros valores mobiliários são negociados no Brasil, conhecida também pela sigla e identidade visual [B]³. Cabe destacar que estes índices não possuem restrições quanto a características como setor e nível de governança.

Além disso, as conclusões deste estudo podem oferecer contribuições para diversos agentes envolvidos no mercado acionário. Para os investidores, o estudo pode ser capaz de ajudá-los em ajustes de estratégias de investimentos, potencializando suas tomadas de decisões, uma vez que o sentimento capturado *online* pode ser capaz de impactar o mercado acionário (Ismail et al., 2022; Souza et al., 2019). No âmbito empresarial, as empresas podem utilizar este estudo para aprimorar estratégias de comunicação nas redes sociais, pois fortalecer a presença nessas plataformas pode resultar em benefícios como visibilidade, reconhecimento de marca, diferenciação competitiva, captação de novos investidores e clientes, entre outros benefícios que podem influenciar positivamente a percepção dos investidores e melhorar o desempenho no mercado de ações (Blankespoor et al., 2014; Jena & Majhi, 2023; Ramassa & Fabio, 2016).

Para reguladores, formuladores de políticas e autoridades monetárias, a pesquisa tem o potencial de contribuir com uma base argumentativa para a análise de ações e regulamentações

adicionais relacionadas à divulgação de informação nas redes sociais, permitindo a implementação de medidas que reflitam a dinâmica das interações *online* no contexto do mercado acionário (Almeida et al., 2022). Ademais, uma das contribuições no âmbito teórico reside na viabilidade de replicação metodológica e na comparabilidade de resultados. Entretanto, é importante observar a constante evolução das diretrizes da plataforma ao longo do tempo, podendo requerer ajustes ou alterações na forma de captação dos *tweets*.

Destaca-se a falta de exploração dos países emergentes em estudos sobre a interação entre informações *online*, especialmente das redes sociais, com o mercado acionário (Agarwal et al., 2019; Carvalho et al., 2023). Nesse contexto, a singularidade dessas economias em desenvolvimento justifica estudos adicionais, motivando este trabalho a contribuir teórica e empiricamente sobre a relação de informações oriundas do Twitter com o mercado acionário emergente brasileiro (Ismail et al., 2022; Souza et al., 2019).

Adicionalmente, ressalta-se que, embora as redes sociais tenham o potencial de aprimorar a eficiência da informação e a qualidade das decisões de investimento, elas também carregam o risco de disseminar informações incorretas e contribuir para comportamentos irracionais ou manipulativos (Hirshleifer et al., 2021). Apesar de não ser o foco desta pesquisa, destaca-se a necessidade de conscientização sobre os potenciais impactos das informações disseminadas no mercado financeiro, especialmente diante da propagação de postagens falsas ou tendenciosas. Assim, o estudo abre espaço para debates sobre ética na comunicação *online*, estimulando reflexões sobre práticas responsáveis nas redes sociais e incentivando a busca por soluções que promovam uma interação mais transparente entre usuários, empresas e o mercado.

1.4 DELIMITAÇÃO E ORGANIZAÇÃO DA PESQUISA

A pesquisa se delimita ao escopo de *tweets* públicos publicados no período de 01 de janeiro de 2022 a 31 de dezembro de 2023, abrangendo 24 meses. A amostra analisada consiste em 24 empresas brasileiras que possuem ADR listados na Bolsa de Valores de Nova Iorque (NYSE) e que também estão incluídas nos índices amplos brasileiros IBOV, IBRA e IBRX100. Foram coletados *tweets* em português usando termos relacionados às 24 entidades, por meio de um *script* em linguagem de programação *Python* (Anexo 1).

Quanto à organização desta pesquisa, está estruturado em cinco capítulos. O Capítulo 1, que é o presente capítulo, compreende a introdução, abordando a contextualização e o problema de pesquisa, delineando os objetivos, além de discutir as justificativas, contribuições e delimitações do estudo. O Capítulo 2 aborda a revisão de literatura, explorando a eficiência

informacional e a relação entre mídias sociais e o mercado acionário. O Capítulo 3 detalha os procedimentos metodológicos, desde a categorização da pesquisa até a descrição da amostra, incluindo o procedimento de coleta de dados e a apresentação dos métodos de análise utilizados. O Capítulo 4 apresenta a estatística descritiva e os resultados da identificação dos picos de atividade no Twitter, estudos de eventos e testes de correlação e causalidade. Por fim, o Capítulo 5 reúne as principais conclusões, destacando as contribuições e limitações da pesquisa, além de propor sugestões para estudos futuros. O trabalho também conta com Apêndices e Anexos que complementam os dados e as análises realizadas.

2 FUNDAMENTAÇÃO TEÓRICA

Na seção intitulada "Fundamentação Teórica", abordam-se duas áreas fundamentais. O tópico 2.1 explora a Eficiência Informacional bem como a Hipótese do Mercado Eficiente (HME), destacando suas características, a racionalidade do indivíduo e anomalias do mercado. No tópico 2.2, a atenção volta-se para as interconexões entre mídias e redes sociais e o mercado acionário, explorando como a dinâmica *online* pode influenciar o comportamento no mercado. Incorporará no subtópico 2.2 os estudos anteriores, com ênfase nas implicações do Twitter, explorando como a plataforma e os sentimentos extraídos dela têm sido explorados, fornecendo subsídios para a análise dos resultados.

2.1 EFICIÊNCIA INFORMACIONAL DE MERCADO

A Hipótese do Mercado Eficiente (HME), formalizada por Fama (1970), consolidou o conceito de eficiência de mercado como um pilar fundamental da teoria financeira contemporânea, contribuindo para a compreensão da formação dos preços dos ativos em mercados competitivos (Alajbeg et al., 2012; de Camargos & Barbosa, 2015; Mendes & Lopes Lucena, 2022; Tıtan, 2015). Contudo, a preocupação com a eficiência informacional antecede a HME. De forma preliminar, Gibson (1889) argumentou que, em um mercado aberto, o valor das ações reflete o julgamento da melhor inteligência disponível assim que as informações se tornam públicas. Essa observação já considerava a racionalidade do investidor e incorporava, ainda que implicitamente, o conceito de mercado eficiente, ou seja, que os preços refletem todas as informações disponíveis (Fama, 1970).

Esse entendimento foi moldado por abordagens como "*fair game*" (jogo justo) e "*martingale*". Samuelson (1965) relacionou a eficiência dos mercados ao modelo de

martingale que, fundamentado no princípio de um jogo justo e com origens nas teorias sobre jogos de azar, descreve um processo estocástico em que a melhor previsão para o valor futuro é o valor atual, com ganho esperado de zero, caracterizando tanto a *martingale* quanto jogo justo como expressões do equilíbrio nos mercados financeiros (LeRoy, 1989; Lo, 2000; Samuelson, 1965). Além disso, Bachelier (1900) contribuiu ao aplicar o conceito de movimento browniano, um movimento aleatório observado em partículas em um fluido, para modelar as flutuações dos preços dos ativos. Esse feito influenciou o desenvolvimento do modelo de “*random walk*” que se tornou um dos pilares da HME (Camargos & Barbosa, 2010; Naseer & Bin Tariq, 2016; Sewell, 2011).

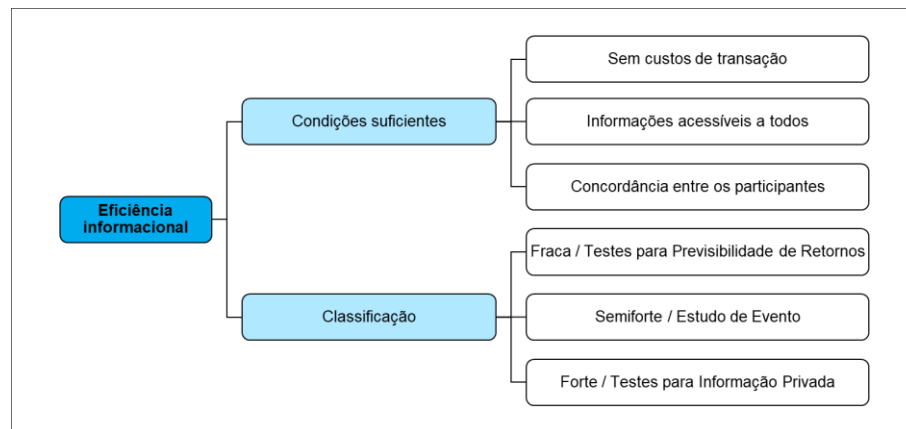
O modelo de passeio aleatório caracteriza as mudanças nos preços dos ativos como sucessivas, independentes e identicamente distribuídas, ou seja, uma sucessão de passos aleatórios (Fama, 1970). Dessa forma, a HME afirma que, com muitos investidores negociando ativamente, os preços se ajustam rapidamente às novas informações, tornando o mercado mais eficiente (Fama, 1970, 1991; Naseer & Bin Tariq, 2016; Souza & Martins, 2022). Este ajuste rápido resulta em mudanças de preços aleatórias, dificultando a previsão de movimentos futuros e tornando impraticável usar tendências passadas para prever preços com precisão (Malkiel, 2003; Naseer & Bin Tariq, 2016).

Nesse cenário, o retorno obtido nos mercados contemporâneos está diretamente relacionado ao risco associado dos ativos, uma vez que os preços atuais refletem o valor presente dos títulos, evitando, assim, ganhos anormais baseados em informações (Naseer & Bin Tariq, 2016; Pinheiro, 2019; Vieira, 2021). Isso resulta em equivalência entre os retornos esperados e requeridos, eliminando a necessidade dos investidores em buscar ativos precificados incorretamente (Gitman, 2010). Como consequência, investidores que desejam retornos mais elevados devem considerar ativos de maior risco, visto que a recompensa está associada à disposição de assumir riscos, não à habilidade de prever com precisão o futuro em comparação com um mercado eficientemente informado (Keynes, 1923; Souza, 2020).

Para que um mercado seja considerado informacionalmente eficiente, Fama (1970) descreve que três condições são suficientes, mas não obrigatórias, são elas: a) a ausência de custos de transação na negociação de títulos, facilitando a fluidez das transações; b) disponibilidade de informação sem custo para todos os participantes, garantindo acesso equitativo às informações relevantes; e, c) a concordância total entre os participantes com as implicações das informações disponíveis para o preço atual e as distribuições de preços futuros

de cada título. Conforme proposto por Fama (1970, 1991), quando efetivamente caracterizado como eficiente, o mercado pode ser classificado em três formas distintas (Figura 1).

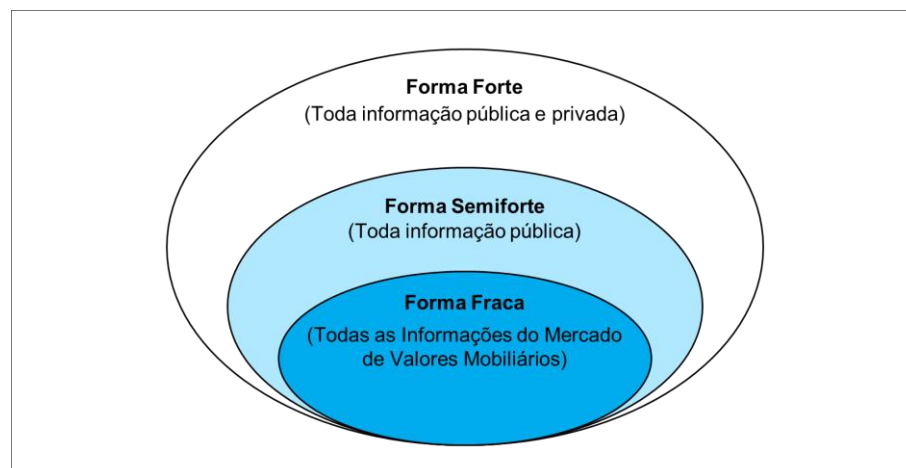
Figura 1 - Mercado eficiente, suas condições e categorias



Fonte: Elaborado pela autora, com base em Fama (1970, 1991).

A eficiência de mercado é categorizada em fraca, semiforte ou forte. A primeira postula que os preços refletem informações passadas, a segunda inclui informações públicas passadas e atuais, e a última abrange todas as informações disponíveis, públicas e privadas, incluindo informações privilegiadas (Fama, 1970; Mustafa & Ahmed, 2020). Nota-se que a classificação é hierárquica, com a eficiência fraca contida na semiforte, que, por sua vez, está contida na eficiência forte (Figura 2). É importante ressaltar que Fama (1991) propôs uma renomeação nas categorias, designando os testes de eficiência fraca como “testes de previsibilidade de retornos”, e os testes de eficiência semiforte e forte como “estudos de eventos” e “testes para informações privadas”, respectivamente.

Figura 2 - Hierarquia dos níveis de eficiência



Fonte: Elaborado pela autora, com base em Naseer e Bin Tariq (2016).

A HME passou a ser amplamente testada e discutida a partir da década de 1970 (Alajbeg et al., 2012; Malkiel, 2003). Nesse ínterim, Jensen (1978) argumentou que a Hipótese de Mercados Eficientes é uma das proposições econômicas com grande apoio de evidências empíricas, sustentada por diversos estudos que demonstraram a reação dos preços das ações às novas informações. Além disso, a HME influenciou tanto estratégias de investimento (Abreu et al., 2023) quanto políticas econômicas, sendo utilizada, por exemplo, para justificar a falta de regulamentação rigorosa nos mercados financeiros que, por sua vez, contribuiu para a crise financeira de 2007-2008 (Ball, 2009).

A crise de 2007-2008, junto com outras evidências de anomalias, revelou que os mercados podem não ser tão estáveis, previsíveis e eficientes como a HME originalmente sugeria, destacando inconsistências no conceito central da hipótese (Brown & Cliff, 2004; Jensen, 1978; Pereira & Urpia, 2011; Shiller, 2003). Dentre as evidências de ineficiência de mercado, cabe destacar o estudo de Bernard e Thomas (1990) que, a partir de uma amostra de 2.649 empresas no período de 1974 a 1986, identificou que os preços das ações não incorporaram completamente todas as informações disponíveis sobre os ganhos futuros, resultando em reações previsíveis nos preços das ações após os anúncios de ganhos futuros.

Ao longo do tempo, as investigações sobre ineficiências informacionais ganharam importância. Diante de comportamentos inesperados no preço dos ativos, os investidores começam a demonstrar interesse em retornos anormais, gerando incentivos para buscar ou usar informações privilegiadas (Camargos & Barbosa, 2010; Magalhães-Timotio & Barbosa, 2020; Malkiel, 2003). Embora a HME seja uma hipótese bem fundamentada, que se alinha com a noção de mercados racionais, têm sido identificado comportamentos nos preços dos ativos que parecem desafiar a ideia de eficiência informacional, ou seja, há indícios de anomalias nesses mercados (Jensen, 1978; Mustafa & Ahmed, 2020; Tıtan, 2015; Woo et al., 2020).

As anomalias podem ser categorizadas em três categorias: i) de calendário, ii) fundamentais e, iii) técnicas. As anomalias de calendário, como o Efeito Janeiro e o Efeito Final de Semana, estão associadas a períodos específicos e desafiam a forma fraca de eficiência de mercado (Latif et al., 2012; Naseer & Bin Tariq, 2016; Rossi, 2015). As fundamentais originam-se da Análise Fundamentalista e estão ligada ao valor intrínseco das ações, como o Efeito *Book-to-Market* e o Efeito Tamanho (Camargos & Barbosa, 2010; Famá et al., 2008). Por fim, as anomalias técnicas estão relacionadas a previsões futuras, baseando-se no comportamento passado dos ativos, como o Efeito Momento (Famá et al., 2008). A Tabela 1 apresenta as principais anomalias identificadas no mercado.

Tabela 1 – Principais anomalias no mercado

Tipo de Anomalia	Anomalia	Descrição
Calendário	Efeito Dia da Semana	Retornos diários variam ao longo da semana, geralmente menores nas segundas-feiras.
	Efeito Janeiro	Retornos em janeiro tendem a ser maiores em comparação com outros meses.
	Efeito Mudança de Mês	Ações apresentam retornos maiores no último e nos quatro primeiros dias do mês.
Fundamentais ou de Valor	Efeito de Sobrereação	Investidores reagem exageradamente a boas ou más notícias, resultando em movimentos extremos.
	Efeito Tamanho	Ações de empresas pequenas têm retornos ajustados ao risco maiores que as grandes.
	Efeito do Índice Preço/Lucro	Retornos ajustados ao risco de ações com baixo P/L são superiores aos previstos.
Técnicas	Efeito Momento	Ativos com bom desempenho passado tendem a ter bom desempenho futuro.
	Uso de Médias Móveis	Estratégia baseada em médias móveis para prever movimentos de preços.
	Quebra de Faixa de Negociação	Identifica oportunidades de negociação quando os preços rompem faixas estabelecidas.

Fonte: Adaptado de Camargos e Barbosa (2010), Famá et al. (2008) e Latif et al. (2011).

Por sua vez, as Finanças Comportamentais buscaram identificar os elementos causadores de ineficiências, analisando como fatores psicológicos influenciam os investidores e os preços dos ativos (Famá et al., 2008; Kimura, 2010; Naseer & Bin Tariq, 2016; Woo et al., 2020). Os investidores nem sempre agem de forma racional, sendo frequentemente influenciados por vieses comportamentais, como o excesso de confiança, aversão à perda e comportamento de manada, que podem desviar os preços de mercado dos seus valores fundamentais (Cornell, 2018; Kahneman & Tversky, 1979; Kimura, 2010; Tversky & Kahneman, 1974; Valcanover et al., 2020). Ou seja, há desvio da premissa de racionalidade, sendo que o comportamento dos investidores muda ao longo do tempo, tornando a eficiência do mercado dependente do contexto e da forma como eventos e informações são interpretados e absorvidos (Cornell, 2018; de Camargos & Barbosa, 2015).

Ademais, a assimetria informacional, ou seja, a distribuição desigual de conhecimento entre os participantes, pode contribuir para a ineficiência do mercado (Akerlof, 1970; Belo & Brasil, 2006; Goel et al., 2020; Komalasari, 2016). Em um mercado eficiente todos teriam igualdade de acesso às informações. Mas na prática, há informações assimétricas e negociações não simultâneas (Belo & Brasil, 2006; Fama, 1970; Goel et al., 2020). Isso pode gerar dois fatos estilizados: i) seleção adversa, quando a qualidade do bem é desconhecida, e ii) risco moral, quando as ações de uma parte não são observadas após o acordo (Belo & Brasil, 2006; Cruz, 2016; Guruganesh et al., 2021). Como consequência, pode ocorrer precificação incorreta ou

atrasos nos ajustes de preços, com o ativo sendo “justamente” precificado apenas quando a informação relevante se tornar pública (Belo & Brasil, 2006; Di Toro, 2022).

Esses vieses podem ser atenuados pelo avanço tecnológico e a digitalização do mercado. Os avanços que reduzem os custos de aquisição e de processamento de informações, ou que aumentam a escala ou tornam os processos mais rápido, tem contribuído com a eficiência informacional (Barbopoulos et al., 2021). O aprendizado de máquina, por exemplo, tem sido utilizado para prever movimentos de preços e identificar padrões, desafiando as suposições tradicionais da HME (Diamond & Perkins, 2022; Seo & Yang, 2024). No entanto, algoritmos e sistemas automatizados podem perder nuances subjetivas e cometer erros, podendo comprometer a qualidade das informações refletidas nos preços (Barbopoulos et al., 2021).

Assim como os algoritmos, a mídia e as redes sociais, embora eficazes na disseminação de informações relevantes, podem tanto aprimorar a qualidade das decisões quanto podem propagar crenças errôneas e estratégias ingênuas (Hirshleifer et al., 2021). Essas plataformas, ao influenciar a atenção pública, podem levar os investidores a reações exageradas ou irracionais, distorcendo os preços e gerando movimentos anormais no mercado (Woo et al., 2020). Em suma, o avanço tecnológico tem, por um lado, desafiado a HME, mas também tem proporcionando aos investidores um grande volume de dados ou informações e novas ferramentas para a tomada de decisões (Arruda et al., 2015; Barber & Odean, 2001; Costa & Famá, 2007; Jensen, 1978; Souza & Martins, 2022).

Com base nas críticas à HME, Lo (2004, 2005) propôs a Hipótese dos Mercados Adaptativos (HMA) como uma alternativa para reconciliar e possibilitar a coexistência entre a Hipótese de Mercados Eficientes e as finanças comportamentais (Dhankar & Shankar, 2016; Lim & Brooks, 2011). A HMA propõe que a eficiência do mercado é dinâmica e varia com o tempo e entre os mercados, influenciada pela capacidade dos participantes de se adaptarem às mudanças nos mercados e pelas dinâmicas complexas, como bolhas, ciclos e crises (Mirzaee Ghazani & Khalili Araghi, 2014). Esta abordagem afirma que os mercados não convergem para um equilíbrio perfeito, mas sim que estão em um processo dinâmico, volátil e com padrões comportamentais recorrentes. Tal perspectiva contrasta com a HME, que pressupõe uma ideia de eficiência estática e universal (Dhankar & Shankar, 2016; Lim & Brooks, 2011).

A HMA sugere que, em mercados habitualmente eficientes, há momentos em que ineficiências podem ser exploradas por investidores adaptativos que conseguem identificar e capitalizar essas oportunidades (Lim & Brooks, 2011). A HMA se baseia nos preceitos de comportamento individual e evolução de mercado, em que os participantes ao agir em seu

próprio interesse, cometem erros, aprendem e se adaptam. Há uma competição e seleção natural que molda a “ecologia de mercado” e a evolução determina sua dinâmica (Lim & Brooks, 2011; Santos et al., 2024; Xiong et al., 2019). Esta visão adaptativa do mercado considera a coexistência de eficiência e ineficiência, que variam ao longo do tempo e entre diferentes mercados, dependendo das circunstâncias e dos participantes (Mirzaee Ghazani & Khalili Araghi, 2014; Rönkkö et al., 2024; Santos et al., 2024).

Em suma, a HME postula que o mercado é eficiente ao refletir todas as informações disponíveis nos preços dos ativos financeiros, o que torna difícil a obtenção consistente de retornos acima da média. Contudo, embora a HME forneça uma estrutura robusta, a HMA sugere que os mercados nem sempre reagem de maneira perfeita às informações disponíveis, por isso há momentos de ineficiência no mercado. Nesse contexto, as, especificamente em sua forma semiforte, uma vez que essa categoria de eficiência considera que os preços incorporam todas as informações públicas disponíveis. Enfim, este estudo adota a perspectiva de eficiência na forma semiforte, que considera informações e notícias públicas disponíveis. Ao avaliar como as informações originárias e propagadas pelas redes sociais, especificamente do Twitter, são incorporadas no preço do mercado, o presente estudo pretende contribuir com a literatura sobre a eficiência do mercado bursátil brasileiro.

2.2 REDES SOCIAIS E MERCADO ACIONÁRIO

O avanço tecnológico resultou na criação de redes sociais, para suprir a crescente demanda por comunicação e compartilhamento de opiniões e informações de forma ágil, ampliando as interações sociais em ambiente virtual (Lira et al., 2022; Teti et al., 2019). Assim, uma rede social *online* é um ambiente ou interface *web* que reúne pessoas com interesses comuns, que interagem, trocam informação, se conectam ou, de forma geral, se relacionam (Zenha, 2018). A interação pode ocorrer de forma síncrona ou assíncrona, por meio de plataformas como o Twitter, Facebook, Instagram etc. Para os investidores individuais, esses ambientes se tornaram uma fonte relevante de informações sobre economia e investimentos, (Comissão de Valores Mobiliários, 2023; Garton et al., 1997; Recuero, 2009; Zenha, 2018).

Ao atuar como fonte de informações e notícias, inclusive como parte integrante do processo de divulgação de informações corporativas, essas plataformas passaram a fazer parte do ambiente empresarial (Cocilova, 2015; Farkas & Keshk, 2019; Teti et al., 2019). Diante dessa capacidade informativa das redes sociais, a *Securities and Exchange Commission* (SEC), órgão equivalente à Comissão de Valores Mobiliários (CVM) no Brasil, reconheceu

oficialmente, em abril de 2013, o uso corporativo das redes sociais por empresas de capital aberto. A SEC esclareceu que a *Regulation Fair Disclosure* (Reg FD), na qual desde 2008 orientava as empresas norte-americanas a efetuarem suas divulgações em *sites* corporativos, passaria a abranger às redes sociais e outros meios de comunicação também, estando sujeitas as mesmas condições da regulamentação (SEC, 2013).

A necessidade de ampliar o alcance da Reg FD, para além dos *websites* corporativos, encontra obstáculos no reconhecimento de que as redes sociais não são apenas um canal de comunicação, mas são também uma oportunidade estratégica para estabelecerem interações relevantes. As organizações que não adotam uma estratégia que inclua as redes sociais reduzem a possibilidade de atrair potenciais investidores (Boylan & Boylan, 2017). A importância das redes sociais para o mercado financeiro se torna mais evidente diante do aumento no número de investidores. No Brasil, o número de investidores individuais passou de aproximadamente 700 mil em maio de 2019 para cerca de 4,9 milhões em setembro de 2023 (B3, 2022, 2023a).

Com o aumento de investidores, em abril de 2023, a CVM analisou o impacto dos “influenciadores digitais” no mercado de valores mobiliários brasileiro. A pesquisa identificou que 73% das pessoas físicas que realizaram o primeiro investimento basearam-se em informações de Youtubers ou Influenciadores e 45% em plataformas *online*. Em relação a busca de conhecimento sobre investimentos, os investidores recorriam a *internet* em geral (73%), a Youtubers ou Influenciadores (60%), as redes sociais, *e-mails* e notificações de bancos e corretoras (38%), e redes sociais como um todo (36%) (Comissão de Valores Mobiliários, 2023). Esses dados da CVM corroboram os achados de Siikanen et al. (2018) que identificaram que investidores menos ativos e organizações sem fins lucrativos tendem a ser mais influenciadas pelas informações do Facebook do que investidores mais ativos, como instituições financeiras.

As plataformas de redes sociais podem agir como catalisadores e influenciar os investidores, impactando assim os mercados acionário (Umar et al., 2021). Um caso ilustrativo desta ação catalisadora é a GameStop (GME). Esta varejista de produtos eletrônicos e *videogames*, enfrentava dificuldades financeiras, agravadas pela pandemia da COVID-19. Em janeiro de 2021, investidores individuais compartilharam suas observações *online* sobre grandes fundos de investimento que apostavam contra a GME. A compra de ações da empresa, em massa, causou rápido aumento no preço que chegou a atingir quase \$500 por ação em 28 de janeiro de 2021 (Chohan, 2021; Pedersen, 2022; Umar et al., 2021). O súbito e expressivo aumento de preços e volume de negociações, e consequente aumento na volatilidade, resultou

em perdas substanciais para os fundos de investimento, que com vendas a descoberto, precisaram proteger suas posições (Chohan, 2021; Pedersen, 2022; Umar et al., 2021).

O movimento iniciou uma comunidade do Reddit denominada “WallStreetBets” e se espalhou para outras redes sociais como Twitter e Youtube, com o apoio de figuras influentes como Elon Musk e Keith Gill (Chohan, 2021; Pedersen, 2022; Umar et al., 2021), chamando a atenção de investidores “amadores”, mídia e órgãos reguladores como a SEC. Assim, uma característica marcante do episódio foi o papel do sentimento do investidor (Chohan, 2021; Umar et al., 2021), uma vez que os sentimentos expressos Reddit tiveram um papel no evento, especialmente durante os momentos de mercado em alta (Long et al., 2022).

Assim como o Reddit, o Twitter teve sua contribuição nesse caso, quando Elon Musk sinalizou apoio ao movimento postando o *link* da comunidade do Reddit juntamente com a mensagem “Gamestonk!!”¹, dando mais visibilidade a ação. O Twitter desempenhou esse papel por ser uma das primeiras e mais populares plataformas de redes sociais utilizadas no tanto por usuários comuns como pelas empresas para divulgar informações corporativas, notícias e atualizações sobre produtos e serviços (Amin et al., 2021; Farkas & Keshk, 2019). Lançado em 2006 por Jack Dorsey, Evan Williams e Biz Stone, é um *microblog* que possibilita enviar e receber, em tempo real, atualizações denominadas *tweets* (Y. Mao et al., 2012; Zenha, 2018).

Por meio dos *tweets* a plataforma se tornou popular, pois permite compartilhar grande quantidade de informações, tais como bem-estar, opiniões sobre marcas, produtos, políticas, eventos sociais, mercado financeiro etc. (Kolchyna et al., 2015). Desde a aquisição por Elon Musk, em 2022, a plataforma passou por transformações, como a mudança do nome “Twitter” e do icônico logo do passarinho azul para a letra “X”, introdução de um limite diário de leitura de *tweets*, entre outras mudanças (*Além do X*, 2023). Apesar das alterações, a comunidade permanece se referindo à plataforma pelo seu nome original, assim, este estudo utilizará a denominação mais comum, preservando a familiaridade estabelecida ao longo dos anos.

Destaca-se que, mesmo com as mudanças na plataforma e na marca, em agosto de 2023, o domínio “*twitter.com*” se encontrava na posição 40^a no mundo e 72^a no Brasil, entre os mais acessados, conforme dados do verificador de tráfego *Cloudfare* (Cloudflare Radar, 2023). Além disso, é relevante ressaltar que, em agosto de 2023, o Brasil ocupou a 3^a posição, atrás de Estados Unidos e Japão, respectivamente, entre os 10 países em que o Twitter foi mais acessado, (Cloudflare Radar, 2023). Tal fato, destaca a popularidade e adoção da plataforma

¹ “Gamestonk” é uma expressão que combina “GameStop” (o nome da empresa de varejo de videogames) e “stonk”, uma gíria da internet usada para descrever um forte aumento no preço de uma ação.

nacionalmente. Assim, esses indicadores sugerem que, apesar das mudanças, o Twitter continua se apresentando como uma plataforma de influência e alcance global, mantendo sua posição como uma das principais redes sociais no cenário mundial.

A popularidade do Twitter é atribuída à ênfase na "troca de informações" em vez da "interação" profunda, como observado em outras plataformas (Zenha, 2018). Essa característica é resultado do foco na comunicação rápida por meio de texto curto, limitando os caracteres a 280 por *tweet* para os usuários não verificados. Isso induz os usuários a serem diretos em suas postagens ou a criar "*threads*", ou seja, uma sequência de mensagens encadeadas que contam narrativas completas "ultrapassando" os limites impostos (Falconi-Pires & Lourenço, 2022). Assim, o aumento do uso das redes sociais, especialmente do Twitter, e as evidências de sua utilidade para investidores em suas decisões financeiras (Bartov et al., 2017) e para estratégias de comunicação empresarial (Boylan & Boylan, 2017; Teti et al., 2019) têm despertado interesse em explorar a relação entre as atividades nessas plataformas e o mercado de ações.

Investigando o uso das redes sociais como uma ferramenta para investir e a relação com os preços das ações, Teti et al. (2019) observaram que quanto maior a presença e interação nas redes sociais, mais influência isso parece ter sobre o preço das ações de empresas da indústria de tecnologia nos Estados Unidos. Corroborando com essa perspectiva, Scaramozzino et al. (2021) observaram uma influência mútua entre sentimentos e preços, especialmente no setor de tecnologia, e identificaram um fluxo causal entre os preços das ações e o sentimento expresso em *blogs* e mídias para as 50 principais empresas do índice S&P.

Buscando verificar a relação entre o conteúdo dos *tweets* e das informações contidas nas demonstrações contábeis brasileiras, o estudo conduzido por Ismail et al. (2022) combinando estudo de evento e análise de conteúdo, evidenciou uma relação fraca entre o conteúdo dos *posts* e os retornos anormais, possivelmente devido à predominância de postagens de empresas de maior porte. Tal predominância também foi observado por Amin et al. (2021), no Reino Unido, no qual identificaram que empresas maiores tendem a divulgar mais informações de Responsabilidade Social Corporativa (CSR) no Twitter em comparação com as menores.

A análise dos sentimentos expressos nas redes sociais, tem se tornado um campo de estudo em expansão, sendo geralmente avaliado em termos de polaridade (positivo, negativo ou neutro). Araújo e Fernandes (2021) realizaram um estudo com cinco empresas, destacando a polaridade do sentimento dos *tweets* de perfis voltados para o mercado acionário. Por sua vez, Almeida et al. (2022) analisaram as ações da Petrobras ao longo de uma década. Ambos os estudos revelaram uma associação entre o sentimento expresso nos *tweets* e as flutuações nos

preços das ações, sugerindo que variações no sentimento podem influenciar os retornos das ações. Os achados destes estudos corroboram com os resultados de pesquisas anteriores.

Neste âmbito, Bollen et al. (2011) que identificaram correlações entre certos estados de humor e mudanças nos valores de fechamento do DJIA e Rao e Srivastava (2012) que encontraram uma correlação de 0,88 entre os retornos e o sentimento polarizado do Twitter. Da mesma forma, Ranco et al. (2015) identificaram uma dependência entre o sentimento do Twitter e os retornos anormais durante os picos de volume dos *tweets*, apesar da baixa correlação de Pearson. Utilizando aprendizado de máquina supervisionado, o estudo demonstrou que a polaridade do sentimento (positivo, negativo e neutro) nas postagens influenciava a direção dos retornos, mesmo que estes últimos fossem baixos, mantendo a dependência estatisticamente significativa dias após os eventos.

Outros estudos, como Karabulut (2012) também observaram efeitos ao longo de dias. O autor, identificou que seu indicador de sentimento, formulado a partir de atualizações de *status* e interações dos usuários do Facebook, pode influenciar, os retornos do mercado de ações dos EUA. No entanto, esse efeito é temporário, sendo totalmente revertido durante as duas semanas subsequentes a de negociação. Em contrapartida, para Rautiainen & Jokinen (2022), o sentimento nas redes sociais no contexto de empresas finlandesas não foi significativo na previsão dos preços das ações. No entanto, o estudo evidenciou que a atividade e popularidade de uma empresa nas mídias sociais são variáveis relevantes para prever os preços das ações.

Além do sentimento, pesquisas têm explorado o volume e picos de postagens nas redes sociais. Mao et al. (2012) descobriram que número diário de *tweets* sobre as ações do S&P 500 está fortemente correlacionado com o preço de fechamento diário, bem como com as mudanças diárias de preço e mudanças de preço absolutas do S&P 500. Nessa mesma linha, Wei et al. (2016) encontraram uma forte associação entre o volume de mensagens postadas no Twitter e o retorno das ações.

O uso estratégico das redes sociais pelas entidades pode impactar positivamente os mercados financeiros e o valor de suas ações (Boylan & Boylan, 2017). Em relação aos estudos de postagens das próprias companhias, Broadstock e Zhang (2019) descobriram que o sentimento expresso no Twitter, tanto pelas empresas quanto pelo sentimento geral do mercado, influencia os retornos intradiários das ações de empresas americanas. Ganesh & Iyer (2023) utilizando *tweets* corporativos de 30 empresas do DJIA, evidenciaram que a frequência dos posts influencia positivamente os retornos das ações e o volume de negociação. Boylan e Boylan (2017) concluíram que as divulgações financeiras no Twitter e Facebook, pelas próprias

entidades, melhoram o valor para os acionistas em 17 instituições financeiras, resultando em um aumento do valor das ações, com base no Retorno Anormal Cumulativo (CARs) dos três dias subsequentes à disseminação das informações financeiras.

Em termos de recorte analítico, cabe destacar o estudo de períodos ou eventos específicos. No caso da pandemia de COVID-19, Jena e Majhi (2023) investigaram como o sentimento do Twitter durante a pandemia pode prever os índices do mercado de ações nos EUA e na Índia. Com base na relação entre sentimento no Twitter e desempenho do mercado, os autores descobriram que os setores de finanças e bens de consumo foram os mais impactados pela pandemia em comparação com tecnologia e farmacêutico. Por sua vez, no caso de períodos eleitorais, Souza et al. (2019) examinaram o impacto do sentimento expresso em *tweets* relacionados ao índice IBOVESPA no mercado acionário brasileiro durante as eleições de 2018 e observaram que o volume diário de *tweets* afetou o volume de negociações, resultando em um aumento nas transações no mercado acionário. Além disso, descobriram uma influência positiva nos retornos do índice a curto prazo, mas revertendo essa tendência quatro dias depois.

As contas do Twitter de pessoas com relevância e influência na sociedade, como figuras políticas e grandes empresários, têm sido objeto de análise. Por exemplo, Alves (2022) concluiu que não se pode descartar a hipótese de que os *tweets* do Elon Musk não influenciam os preços das ações da Tesla, portanto, os mercados seriam eficiente em sua forma semiforte. No contexto político, Ge et al. (2018) identificaram resultados semelhantes ao de Alves (2022) ao concluir que os *tweets* do então Presidente Donald Trump que mencionam empresas cotadas em bolsa influenciam os preços das ações, bem como aumentam o volume de negociação, a volatilidade e a tenção dos investidores. Esse resultado difere do encontrado por Tommasetti et al. (2021) que não encontraram nenhuma relação significativa dos *tweets* do então presidente Bolsonaro sobre bancos com o preço das ações relacionadas ou os volumes negociados.

Outros estudos, como de Arruda et al. (2015), sugeriram que o uso de Facebook, Twitter e YouTube, tanto oficiais (vinculados ao site da empresa) quanto não oficiais (páginas encontradas por busca manual), poderia reduzir a assimetria informacional e influenciar o valor das empresas no Brasil e nos Estados Unidos em 2012. Contudo, os resultados mostraram que, para as empresas brasileiras, apenas o Facebook não oficial teve um impacto significativo. Por outro lado, para as empresas americanas, as redes sociais confirmaram a redução da assimetria informacional, mas não influenciaram o valor das empresas. No geral, concluíram que, apesar de as redes sociais poderem reduzir a assimetria, elas não geram retornos anormais, estando em conformidade com a HME.

Diante dos estudos apresentados, verifica-se uma necessidade crescente de avaliação da dinâmica das interações *online*, especialmente no contexto dos mercados de ações emergentes. Em um ambiente de comunicação fácil e ágil, observa-se que o Twitter tem se destacado na comunidade acadêmica que investiga o mercado financeiro, seja como fonte de informação ou pela sua capacidade de propagação de notícias. Portanto, a escolha do Twitter como fonte de dados para esta pesquisa está alinhada com a relevância e popularidade que a plataforma detém tanto na sociedade quanto no meio acadêmico. Por fim, a escolha de variáveis de análise como o volume de *tweets* e os sentimentos extraídos deles são considerados indicadores capazes de explicar determinados comportamentos no mercado acionário, como o retorno, volume negociado e volatilidade das ações.

3 PROCEDIMENTOS METODOLÓGICOS

3.1 DEFINIÇÃO DO TIPO DE PESQUISA

Quanto ao objetivo da pesquisa, este pode ser considerado como descritivo, uma vez que nesse tipo de pesquisa procuram descrever as características de uma população ou fenômeno específico, bem como a identificação e, por vezes, a análise da natureza das relações entre diferentes variáveis, buscando identificar quais variáveis são relevantes em uma determinada situação e como se relacionam (Gil, 2019; Marconi & Lakatos, 2023).

O problema desta pesquisa adota uma abordagem quantitativa-dedutiva de natureza aplicada. Nessa abordagem, os pesquisadores buscam analisar de forma neutra e objetiva os fenômenos sociais por meio da utilização de número e medidas estatísticas que possibilitam descrever populações e examinar relação entre variáveis, partindo de princípios gerais, teorias objetivas ou premissas amplamente aceitas e progride em direção ao particular, focando na aplicação do conhecimento em contextos específicos, em vez do desenvolvimento de teorias universais (Creswell & Creswell, 2021; Gil, 2019).

Esta pesquisa é considerada netnográfica devido a origem *on-line* dos dados coletados, uma vez que essa modalidade examina o comportamento de indivíduos e grupos sociais na *internet*, bem como as interações e dinâmicas que ocorrem nesse ambiente virtual (Gil, 2019). A pesquisa também é considerada bibliográfica, pois essas pesquisas são elaboradas com base em material existente, ou seja, abrange a bibliografia tornada pública em relação ao tema de estudo, constituído principalmente de livros e artigos científicos (Gil, 2019; Marconi & Lakatos, 2023).

3.2 AMOSTRA E COLETA DE DADOS

Inicialmente, para a composição da população da pesquisa, considerou-se a utilização de empresas que estivessem tanto em bolsa brasileira, a B3, quanto na Bolsa de Valores de Nova Iorque (NYSE) ou na *National Association of Securities Dealers Automated Quotations* (NASDAQ), devido à visibilidade e rigor regulatório das bolsas norte-americanas (Assaf Neto, 2021). No entanto, como não havia amostra relevante de empresas listadas na B3 e na NASDAQ, optou-se por usar exclusivamente a B3 e a NYSE, uma vez que esta última é reconhecida como a maior e mais influente bolsa de valores dos Estados Unidos, abrigando empresas com as maiores capitalizações de mercado e exercendo influência global sobre outros mercados acionários (Assaf Neto, 2021; Bodie et al., 2015; InfoMoney, 2022). Assim, a população da pesquisa ficou composta por empresas brasileiras listadas na B3 e que possuem *American Depositary Receipts* (ADRs) listados na NYSE.

A relação das empresas foi obtida por meio do banco de dados Economática® em 18/07/2023, totalizando 31 empresas. A partir deste recorte inicial, foram filtradas as empresas presentes nos índices amplos da B3: IBOV, IBRA e IBRX 100. Esses índices são resultantes de carteiras teóricas de ações em um determinado intervalo de tempo com objetivos distintos (Pinheiro, 2019). O IBOV, como principal indicador de desempenho das ações na B3, busca refletir o desempenho médio dos negócios à vista ocorridos nos preços, o IBRX 100 visa ser o indicador do desempenho médio das cotações das 100 ações de maior negociabilidade (número de negócio e volume financeiro) e representatividade e o IBRA tem como objetivo ser o indicador do desempenho médio das cotações de todos os ativos negociados no mercado à vista (Assaf Neto, 2021; B3, 2023b).

Após concluir o processo de seleção, a amostra final ficou composta por 24 empresas, abrangendo tanto empresas que atuam no mercado acionário nacional quanto no mercado internacional (Tabela 2), totalizando 529.180 tweets, sendo 291.003 *tweets* em 2022 e 238.177 em 2023.

Tabela 2 – Amostra de Empresas brasileiras de capital aberto listadas nos índices amplos da B3 e com ADRs listados na NYSE

Empresas	Governança ^a	Setor Econômico	Subsetor Bovespa	Segmento Bovespa	Tweets
Ambev	Tradicional	Consumo não cíclico	Bebidas	Cervejas e refrigerantes	19.800
Assai	NM	Consumo não cíclico	Comércio e distribuição	Alimentos	19.019
Azul	N2	Bens industriais	Transporte	Transporte aéreo	1.489
Bradesco	N1	Financeiro	Intermediários financeiros	Bancos	73.378
Braskem	N1	Materiais básicos	Químicos	Petroquímicos	8.551
BRF SA	NM	Consumo não cíclico	Alimentos processados	Carnes e derivados	1.385
Cemig	N1	Utilidade pública	Energia elétrica	Energia elétrica	18.862
Copel	N2	Utilidade pública	Energia elétrica	Energia elétrica	12.379
Cosan	NM	Petróleo gás e biocombustíveis	Petróleo gás e biocombustíveis	Exploração refino e distribuição	749
Eletrobras	N1	Utilidade pública	Energia elétrica	Energia elétrica	7.969
Embraer	NM	Bens industriais	Material de transporte	Material aeronáutico e de defesa	4.816
Gerdau	N1	Materiais básicos	Siderurgia e metalurgia	Siderurgia	2.726
Gol	N2	Bens industriais	Transporte	Transporte aéreo	1.795
Grupo Natura	NM	Consumo não cíclico	Produtos de uso pessoal e de limpeza	Produtos de uso pessoal	1.041
ItauUnibanco	N1	Financeiro	Intermediários financeiros	Bancos	145.529
P. Acucar-Cbd	NM	Consumo não cíclico	Comércio e distribuição	Alimentos	1.160
Petrobras	N2	Petróleo gás e biocombustíveis	Petróleo gás e biocombustíveis	Exploração refino e distribuição	114.800
Sabesp	NM	Utilidade pública	Água e saneamento	Água e saneamento	20.219
Santander BR	Tradicional	Financeiro	Intermediários financeiros	Bancos	62.660
Sid Nacional	Tradicional	Materiais básicos	Siderurgia e metalurgia	Siderurgia	1.543
Suzano	NM	Materiais básicos	Madeira e papel	Papel e celulose	1.283
Telef Brasil	Tradicional	Comunicações	Telecomunicações	Telecomunicações	730
Ultrapar	NM	Petróleo gás e biocombustíveis	Petróleo gás e biocombustíveis	Exploração refino e distribuição	572
Vale	NM	Materiais básicos	Mineração	Minerais metálicos	6.725

Nota. NM = Novo Mercado; N1 = Nível 1; N2 = Nível 2; PT = português; Qtd = Quantidade. Fonte: Elaborado pela autora (2024) conforme dados coletados.

^a Classificação do Nível de Governança, tem como referência o mês de Julho de 2023.

Para coletar os *tweets*, foi elaborado um *script* em linguagem de programação *Python* usando o *Twscrape* (Twitter API scrapper)², que consiste em uma biblioteca *Python* que permite extrair resultados de busca, perfis de usuários, *tweets* e outras informações do Twitter. Para executar o *script*, foi utilizado o ambiente de desenvolvimento *Visual Studio Code* (VSCode), com o código sendo executado em um arquivo *Jupyter Notebook*. O desenvolvedor do *Twscrape* "@vladkens", explicou que, desde 1º de julho de 2023, o Twitter implementou novos limites de busca para cada conta, o que pode ser considerado um desafio adicional do presente estudo, em relação aos esforços e tempo dedicado a coleta de dados.

² Pesquisar pelo usuário @vladkens, desenvolvedor principal da ferramenta Twscrape, ou pelo link: <https://github.com/vladkens/twscrape>.

Essas limitações impostas pelo Twitter foram o principal motivo para o uso do *Twscrape*. Anteriormente, o Twitter disponibilizava uma API³ (*Application Programming Interface*) gratuita para coleta de dados, contudo, atualmente se tornou paga, com valores que podem atingir até 5 mil dólares, devido a nossas estratégias de monetização da plataforma, tornando seu uso inviável para pesquisas acadêmicas ou outras aplicações com baixo orçamento. Além disso, outras mudanças, como a necessidade de estar logado para visualizar os *tweets* e limites de visualização e de busca, tornaram várias ferramentas de coleta de *tweets* de estudos anteriores impraticáveis, como Rao e Srivastava (2012), Ranco et al. (2015), Souza et al. (2019), Souza (2020), Nti et al. (2020), Almeida et al. (2022), Alves (2022), Ismail et al. (2022), Jena e Majhi (2023), Sarrang et al. (2024) e Lu (2024). Outras ferramentas pagas, como *Bloomberg*, utilizado por Teti et al. (2019), também são inviáveis devido aos custos envolvidos.

Diante das restrições, o *Twscrape* destaca-se pelo uso do método *web scraping*⁴ combinado com a função de autenticação e *login* das contas. Até o momento desta pesquisa, não foram identificados trabalhos que tenham utilizado o *Twscrape* em cenários similares ao do problema de pesquisa, o que pode ser atribuído ao tempo necessário para a realização e publicação de pesquisas, uma vez que a biblioteca foi disponibilizada apenas em 6 de maio de 2023, tornando-o uma ferramenta nova e inexplorada na literatura acadêmica. Portanto, a escolha do *Twscrape* é justificada por oferecer uma solução viável e eficiente para a coleta de dados, viabilizando a execução desta pesquisa.

No que se refere aos termos usados na busca, foram selecionados aqueles que identificassem cada empresa da amostra de forma específica. Por exemplo, para coletar os *tweets* referente a empresa Gol foram incluídos no *script* os seguintes termos-chave: “GOL LINHAS AEREAS”, “GOLL4”, “GOLL3”, “GOLL11” e “GOLL12”. É importante observar que o termo “GOL” não foi informado na busca, por se tratar de um homônimo (palavras de grafia e pronúncia igual outra, mas com significado diferente), uma vez que isso resultaria na coleta de *tweets* não apenas relacionados à empresa da amostra, devido a coincidência do nome com outro substantivo. O mesmo critério foi aplicado para a coleta dos *tweets* da empresa Vale, Azul, Suzano e Natura.

³ As APIs são interfaces especiais projetadas para facilitar a comunicação entre aplicativos e servidores, permitindo a interconexão de pessoas, aplicativos e sistemas (Khder, 2021; Ofoeda et al., 2019).

⁴ *Web scraping* é a extração automática de dados de *sites* usando *software*, convertendo dados não estruturados da *web* em dados estruturados para análise em uma planilha ou banco de dados centralizado (Khder, 2021).

Os *tweets* coletados abrangem um período de dois anos civis, compreendido de 01 de janeiro de 2022 até 31 de dezembro de 2023. A seleção iniciando no começo de 2022 é justificada pela estabilização da pandemia da COVID-19, proporcionando uma janela de observação menos influenciada pela crise global de saúde. Além disso, o ano de 2022 coincide com o período de eleição presidencial no Brasil, conferindo à análise uma dimensão temporal relevante dada a influência que eventos e figuras públicas e políticas podem exercer nas interações e sentimentos expressos nas redes sociais, conforme apontado nos estudos de Ge et al. (2018) e Souza et al. (2019).

Após a coleta e pré-tratamento dos *tweets*, que incluiu a remoção de postagens que não estivessem em português, as patrocinadas e de *spam*, obteve-se um total de 529.180 postagens em português, conforme detalhado na Tabela 2. Esses *tweets* abrangem diversos tópicos, incluindo elogios, reclamações, notícias, opiniões e debates sobre produtos, serviços e aspectos financeiros das empresas. Portanto, representam uma variedade de *stakeholders* e não se limitam a um único tipo de conteúdo. Os *tweets* foram coletados de diversos perfis abertos no Twitter, excluindo os perfis das próprias empresas, cujas publicações não foram utilizadas, visto que algumas não estão presentes na plataforma.

Os dados históricos dos mercados e das ações foram coletados na plataforma Economática®, incluindo informações como o preço de abertura, fechamento, máximo e mínimo do dia, corrigidos pela inflação. Ressalta-se que algumas das empresas da amostra possuem ações ordinárias e preferenciais listadas na B3 e estão presentes nos índices amplos IBOV, IBRX e IBRA, sendo elas Bradesco, Cemig e Itaú Unibanco. Devido ao volume de negociação, optou-se por utilizar apenas as ações preferenciais para a análise dessas companhias.

3.3 ATRIBUIÇÃO DE SENTIMENTO

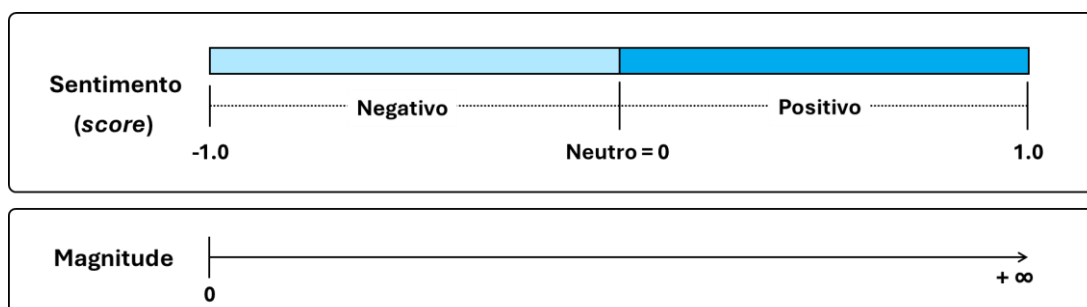
A análise de sentimento envolve o estudo computacional das opiniões, sentimentos e emoções expressas em texto e possui duas abordagens principais: aprendizado supervisionado e aprendizado não supervisionado (Indurkha & Damerau, 2010; Yu et al., 2013). No aprendizado supervisionado, a máquina é treinada com exemplos humanos pré-determinados e, em seguida, classifica a base de dados (Oliveira et al., 2019). Em relação ao aprendizado não supervisionado, conjuntos predefinidos de palavras positivas e negativas são usados para determinar a polaridade do texto, dispensando a classificação prévia da base de dados (Yu et

al., 2013). Diante disso, para atribuir o sentimento, foi adotada a abordagem não supervisionada, utilizando os modelos pré-treinados pela *Google*.

Dito isso, esta pesquisa faz uso de um Processador de Linguagem Natural (PNL) denominado *Google Cloud Natural Language API*⁵ (GNL) para atribuir sentimento a cada *tweet* coletado. Procedimento similar foi adotado nas pesquisas de Bollen et al. (2011), Ge et al. (2018) e Souza e Martins (2022). O GNL examina o texto e identifica a opinião emocional dominante por meio de uma “pontuação” (*score*) que varia de -1.0 a 1.0 para indicar a inclinação emocional geral daquele texto e uma “magnitude” que reflete a força das emoções identificadas, variando de 0.0 a $+\infty$ (Google Cloud, 2023). Ressalta-se que um texto com pontuação neutra pode indicar baixa emoção ou emoções mistas (e.g. quando sentimentos positivos e negativos se anulam) e, para distinguir esses casos, utiliza-se a magnitude, pois textos verdadeiramente neutros têm magnitudes baixas ou zero, enquanto os mistos têm magnitudes maiores (Google Cloud, 2023).

No processo operacional de atribuição de sentimento aos *tweets*, um *script* em linguagem *Python* foi estruturado seguindo a biblioteca e documentação de referência da *API Natural Language Python* (Anexo 2). Para viabilizar o uso da API, foram executadas as orientações de autenticação e a configuração das credenciais. Para executar o roteiro, foi utilizado o ambiente de desenvolvimento VSCode em um arquivo *Jupyter Notebook*. No *script*, foram estabelecidas as faixas de pontuação seguidas por Souza e Martins (2022), assim, considera-se como "Positivo" quando abrange *score* acima de zero, "Neutro" para as com valores de *score* iguais 0, e "Negativo" se aplica para *score* menor que 0 (Figura 3).

Figura 3 - Escala de sentimento e magnitude



Fonte: Elaborado pela autora (2024).

A Tabela 3 apresenta exemplos de *tweets* coletados e como estão classificados.

⁵ Disponível em https://cloud.google.com/natural-language?hl=pt_br

Tabela 3 - Exemplos de *tweets*

<i>Tweet</i>	Score	Magnitude	Sentimento
Aqui vai uma recomendação especial, fiquem de olho na ABEV3, na BRFS3, para elas o 4º Tri é muito importante, elas costumam apresentar resultados espetaculares no fim de ano, tanto por conta das festas natalinas, quanto por conta do verão, a época de ouro para uma cerveja!	0,90	0,90	Positivo
a pessoa responsável por desenvolver o app do Bradesco devia ser imediatamente demitida, como pode um app tão ruim??????	-0,60	0,60	Negativo
Meta é até o final do ano alcançar 1K ABEV3 apenas com os prêmios recebidos	0,0	0,0	Neutro

Fonte: Elaborado pela autora (2024) com base nos dados coletados.

É relevante mencionar que a magnitude, ao ser uma evidência da existência, ainda que mínima, de sentimentos, fundamentou a decisão de classificar como "Neutro" aquele *tweet* cujos valores de *score* e magnitude fossem ambas iguais a 0. Essa escolha baseou-se na compreensão de que, mesmo em casos em que o *score* sugere neutralidade, a presença de magnitude zero indica uma ausência de qualquer inclinação emocional. Ademais, é importante destacar que o objetivo desta pesquisa não é avaliar a eficiência da medição do sentimento, mas sim analisar a relação entre variáveis do Twitter e o mercado acionário.

3.4 CONSTRUTOS/VARIÁVEIS E MÉTODOS DE ANÁLISE

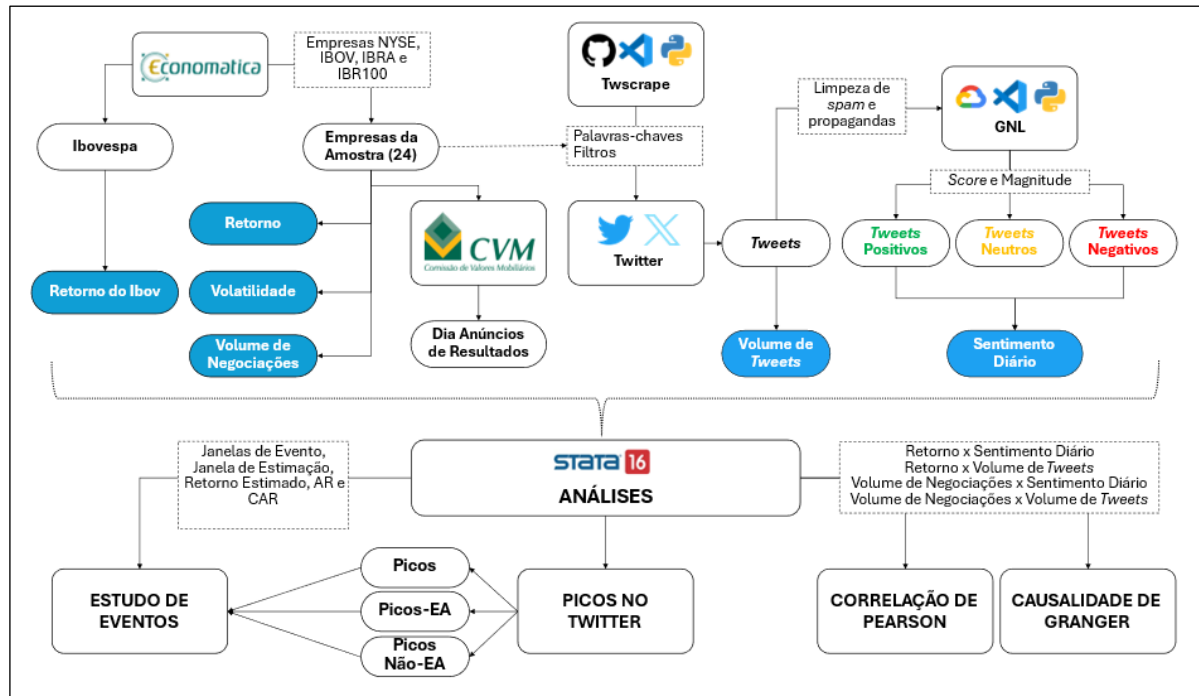
Este tópico aborda os construtos, variáveis e método de análises. Para uma compreensão mais visual e esquematizada do processo metodológico que será descrito neste tópico, veja a Figura 4 que consta com o fluxograma da proposta metodológica.

Um dos métodos aplicados nesta pesquisa consiste no Estudo de Evento. Esse método foi formalizado por Fama (1970, 1991) e tornou-se um dos métodos mais utilizados para se testar a forma semiforte de eficiência do mercado. Assim, a utilidade desse método consiste na verificação do impacto de um evento no valor de uma empresa, ou seja, se o mercado for eficiente o evento será refletido diretamente no preço das ações (MacKinlay, 1997). Para executar esse método, Campbell et al. (1997) define etapas a serem seguidas, são elas: (1) definição do evento de interesse e da janela do evento; (2) critérios de seleção (3) medição dos retornos normais e anormais (4) procedimento de estimativa (5) procedimento de teste (6) resultados empíricos (7) interpretações e conclusões.

As postagens do Twitter foram estabelecidas como os eventos de interesse, sob a percepção de que o mercado acionário não reage de forma anormal a essas postagens, conforme proposto pela HME. Dessa forma, o evento zero corresponde aos dias em que os *tweets* foram feitos, durante o período de 01 de janeiro de 2022 a 31 de dezembro de 2023. A janela do evento

representa o período durante o qual os preços dos ativos das empresas envolvidas no evento são analisados (Campbell et al., 1997; MacKinlay, 1997).

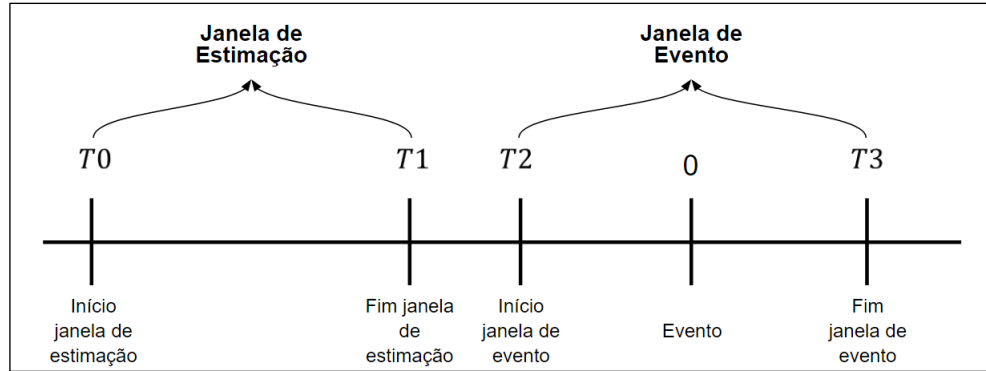
Figura 4 - Fluxograma da proposta metodológica



Fonte: Elaborado pela pesquisa (2024).

Neste contexto, optou-se por utilizar uma janela centrada de 11 dias ao redor do evento, incluindo 5 dias antes e 5 dias após o evento (-5, +5). Estudos anteriores demonstraram que, independentemente do tema da notícia, ela tende a permanecer nos *trending topics* do Twitter por aproximadamente 5 a 6 dias (Koul et al., 2022). A mesma abordagem de janela de -5 a +5 dias foi adotada por Ranco et al. (2015) e Lu (2024) reforçando a adequação deste período para capturar os efeitos do evento. Por fim, a janela de estimação nesta pesquisa se apresenta antes da janela de evento e possui 30 dias. A Figura 5 oferece uma representação visual das janelas temporais. Na figura 5, a linha do tempo destaca a sequência temporal de um evento, na qual o comprimento da janela de estimação é denotado de T0 a T1. O próprio evento ocorre no tempo 0, e a janela do evento é representada de T2 a T3.

Figura 5 - Janela de estimação e de evento



Fonte: Adaptado de Benninga (2014) e MacKinlay (1997).

Seguindo as etapas de execução do estudo de evento, é necessário calcular o retorno diário de cada ação. Assim, o retorno diário das ações foi calculado conforme a Equação 1.

$$Ret_{i,t} = \ln \left(\frac{Close_{i,t}}{Close_{i,t-1}} \right) \quad (1)$$

Em que:

$Ret_{i,t}$ representa o logaritmo do retorno (log-retorno) diário da ação i no tempo t .

$Close_{i,t}$ e $Close_{i,t-1}$ representa valor de fechamento da ação i no tempo t no e tempo $t - 1$, respectivamente.

Além dos retornos das ações individuais, é necessário calcular o retorno diário do índice de mercado, pois ele será utilizado como uma variável de referência para avaliar o desempenho das ações no período de análise. O retorno do índice de mercado, nesta pesquisa representado pelo IBOVESPA (IBOV), foi calculado conforme a Equação 2:

$$Ret_{m,t} = \ln \left(\frac{Close_{m,t}}{Close_{m,t-1}} \right) \quad (2)$$

Em que:

$Ret_{i,t}$ é o retorno diário normal do índice de mercado m no tempo t .

$\ln$ representa o logaritmo neperiano.

$P_{m,t}$ e $P_{m,t-1}$ representam o valor de fechamento do mercado m tempo t no e tempo $t - 1$, respectivamente.

Para calcular os retornos esperados de uma ação, geralmente se utiliza a janela de estimação para criar um modelo baseado na regressão dos retornos da ação em relação aos retornos do índice de mercado (Benninga, 2014). Este modelo, conhecido como modelo de mercado, pode ser visualizado na Equação 3.

$$R_{i,t} = \alpha_i + \beta_i R_{m,t} \quad (3)$$

Em que:

$R_{i,t}$ e $R_{m,t}$ representam o retorno da ação i e o retorno do mercado m no tempo t

α_i e β_i são os parâmetros do modelo, estimados a partir de uma regressão de mínimos quadrados ordinários sobre a janela de estimação.

Conhecido o retorno estimado, é feito o cálculo do retorno anormal (AR) da ação que consiste na diferença entre seu retorno real e seu retorno previsto, com o intuito de medir o impacto de um evento no retorno da ação na janela de evento (Benninga, 2014). A fórmula do retorno anormal pode ser definida como Equação 4:

$$AR_{i,t} = R_{i,t} - E(R_{i,t}) \quad (4)$$

Em que:

$AR_{i,t}$ representa o retorno anormal da ação i no tempo t .

$R_{i,t}$ retorno diário real observado da ação i no tempo t .

$E(R_{i,t}) = \alpha_i + \beta_i R_{m,t}$ representa o retorno esperado calculado na Equação 3.

Com os retornos anormais auferidos, calcula-se o retorno anormal acumulado (CAR), ou seja, a soma de todos os retornos anormais durante a janela de evento. Assim, é possível obter uma medida do impacto total do evento ao longo do período estabelecido. A fórmula pode ser visualizada na Equação 5.

$$CAR_{i,j} = \sum_{t=T2}^{T3} AR_{i,t} \quad (5)$$

Em que:

$CAR_{i,j}$ denota o retorno anormal acumulado da ação i durante a janela de evento j .

$T2$ e $T3$ representa o início e o fim da janela de evento, conforme ilustrado na Figura 5.

$AR_{i,t}$ representa o retorno anormal da ação i no tempo t .

Por meio de teste t de *Student*, o passo subsequente envolve a avaliação da significância dos valores de AR (Equação 6) e CAR (Equação 7). O teste t permite avaliar a significância estatística da diferença entre duas médias de amostras independentes para uma única variável dependente (Hair Jr et al., 2009), sob a hipótese ($H1$) que os retornos anormais são nulos, indicando que o evento (*tweets*) não afeta as ações em análise.

$$tAR_{i,t} = \frac{AR_{i,t}}{RMSE_{i,t}} \quad (6)$$

Em que:

$tAR_{i,t}$ é o valor do teste t para o retorno anormal da ação i no tempo t .

$AR_{i,t}$ representa o retorno anormal da ação i no tempo t .

$RMSE_{i,t}$ consiste raiz do erro quadrático médio obtido a partir da regressão dos retornos estimados para a ação i no tempo t .

$$tCAR_{i,j} = \frac{CAR_{i,j}}{RMSE_{i,t}} \quad (7)$$

Em que:

$tCAR$ é o valor do teste t para o retorno anormal acumulado da ação i na janela de evento j .
 $CAR_{i,j}$ é o retorno anormal acumulado da ação i durante a janela de evento j .
 $RMSE_{i,t}$ consiste na raiz do erro quadrático médio obtido a partir da regressão dos retornos estimados para a ação i no tempo t .

A raiz do erro quadrático médio (RMSE) foi utilizada para calcular a significância dos retornos anormais e acumulados porque ela permite mensurar de forma consistente a variabilidade dos erros de previsão dos retornos. Os níveis de significância adotados e os valores críticos correspondentes do teste t podem ser visualizados na Tabela 4 - Níveis de significância.

Tabela 4 - Níveis de significância

Valor Crítico do Teste t	Nível de significância	Hipótese	Legenda
2,58	1%	Se $ tAR_{i,t} $ ou $ tCAR_{i,j} > 2,58$	Significativo a 1% (***)
1,96	5%	Se $ tAR_{i,t} $ ou $ tCAR_{i,j} > 1,96$	Significativo a 5% (**)
1,64	10%	Se $ tAR_{i,t} $ ou $ tCAR_{i,j} > 1,64$	Significativo a 10% (*)

Fonte: Adaptado de Hair Jr et al. (2009).

Após atribuir um sentimento a cada *tweet* coletado, devido à possibilidade de múltiplos *posts* por dia, foi adaptado o modelo de polaridade proposto por Zhang e Skiena (2010), Ranco et al. (2015) e Luciani et al. (2023) para calcular uma variável que representasse o sentimento diário de cada empresa. Essa variável, denominada Sentimento Diário (Equação 8) é calculada subtraindo a média dos *scores* negativos da média dos *scores* positivas e dividindo pela soma das duas médias de cada empresa i no dia t de negociação. Além do sentimento, esta pesquisa também considera como variável explicativa a quantidade de *tweets* em determinado dia, assim, será utilizado o Volume de *Tweets* relacionado a empresa i no dia t ($VT_{i,t}$), tal como empregado em Rao e Srivastava (2012).

$$SD_{i,t} = \frac{s_{i,t}^+ - s_{i,t}^-}{s_{i,t}^+ + s_{i,t}^-} \quad (8)$$

Em que:

$SD_{i,t}$ representa o sentimento diário dos *tweets* da empresa i no tempo t

$s_{i,t}^+$ representa os valores positivos da empresa i no tempo t , em termos médios

$s_{i,t}^-$ representa os valores negativos da empresa i no tempo t , em termos médios

Diante de evidências que indicam que o volume negociado, além de fonte de informação, tem influência sobre o retorno das ações (Chen et al., 2015; Do et al., 2014; Kao et al., 2020; Sheng et al., 2017), esta pesquisa utiliza o Volume Negociação das ações i negociadas no tempo t ($VN_{i,t}$).

Para captar informações adicionais da volatilidade diária, além do cálculo convencional baseado na informação de fechamento do preço das ações ($Close_{i,t}$), complementarmente serão utilizados cálculos de volatilidade que incluem informações de preços de abertura, preços mais baixos e preços mais altos. Nesse contexto, para o cálculo da volatilidade diária, será utilizada as medidas de volatilidade de Parkinson (1980), Garman e Klass (1980), Rogers e Satchell (1991) e Yang e Zhang (2000), assim, seus estimadores $\sigma P_{i,t}$, $\sigma GK_{i,t}$, $\sigma RS_{i,t}$ e $\sigma YZ_{i,t}$ podem ser observados, respectivamente, nas Equações 9, 10, 11, e 12.

$$\sigma P_{i,t} = \sqrt{\frac{N}{4n \log 2} \sum_{i=1}^n \left[\log \frac{H_{i,t}}{L_{i,t}} \right]^2} \quad (9)$$

Em que:

$\sigma P_{i,t}$ representa Volatilidade de Parkinson do ativo i no tempo t

n representa o número de períodos observados

$H_{i,t}$ representa o preço mais alto do ativo i no tempo t

$L_{i,t}$ representa o preço mais baixo do ativo i no tempo t

$$\sigma GK_{i,t} = \sqrt{\frac{1}{n} \sum \frac{1}{2} \left[\ln \frac{H_{i,t}}{L_{i,t}} \right]^2 - [2 \ln 2 - 1] \left[\ln \frac{C_{i,t}}{O_{i,t}} \right]^2} \quad (10)$$

Em que:

$\sigma GK_{i,t}$ representa a Volatilidade de Garman-Klass do ativo i no tempo t

n representa o número de períodos observados

$H_{i,t}$ representa o preço mais alto do ativo i no tempo t

$L_{i,t}$ representa o preço mais baixo do ativo i no tempo t

$C_{i,t}$ representa o preço de fechamento do ativo i no tempo t

$O_{i,t}$ representa o preço de abertura do ativo i no tempo t

$$\sigma RS_{i,t} = \left(\ln \frac{H_{i,t}}{C_{i,t}} \right) \times \left(\ln \frac{H_{i,t}}{O_{i,t}} \right) + \left(\ln \frac{L_{i,t}}{C_{i,t}} \right) \times \left(\ln \frac{L_{i,t}}{O_{i,t}} \right) \quad (11)$$

Em que:

$\sigma RS_{i,t}$ representa a Volatilidade de Rogers-Satchell do ativo i no tempo t

n representa o número de períodos observados

$H_{i,t}$ representa o preço mais alto do ativo i no tempo t

$L_{i,t}$ representa o preço mais baixo do ativo i no tempo t

$C_{i,t}$ representa o preço de fechamento do ativo i no tempo t

$O_{i,t}$ representa o preço de abertura do ativo i no tempo t

$$\sigma YZ_{i,t} = \left(\ln \frac{O_{i,t}}{C_{i,t}} \right)^2 + k \times \left(\ln \frac{C_{i,t}}{O_{i,t}} \right)^2 + (1 - k) \times \sigma RS_{i,t} \quad (12)$$

Em que:

$\sigma YZ_{i,t}$ representa a Volatilidade de Yang-Zhang do ativo i no tempo t

$C_{i,t}$ representa o preço de fechamento do ativo i no tempo t

$O_{i,t}$ representa o preço de abertura do ativo i no tempo t
representa

k é uma constante calculada por $k = \left(\frac{0,34}{1,34 + \frac{n+1}{n-1}} \right)$, em que n se refere ao número de períodos

$\sigma RS_{i,t}$ representa a Volatilidade de Rogers-Satchell do ativo i no tempo t

Em suma, para avaliar o impacto de um evento no valor da empresa, definiu-se as postagens do Twitter como eventos de interesse e o impacto em termos de retornos anormais acumulados. A abordagem metodológica descrita, confere robustez a investigação, aplicando testes convencionais de análise de volume de *tweets* como atribuição de sentimento diário. Na análise dos retornos também há um resgate histórico de diferentes cálculos de retorno, que incorporam informações adicionais na análise, como preço de abertura, fechamento, mínimo e máximo diário, no cômputo do retorno de preços. Na próxima seção há um cotejamento com picos de atividade, baseados na frequência de postagem e com períodos em que houve anúncio de resultados das empresas.

3.4.1 Picos de Atividade e Eventos

Com as variáveis do Twitter coletadas, foi possível detectar os picos de atividade na plataforma com base na metodologia de Ranco et al. (2015), que utiliza a série temporal do volume diário de *tweets* de cada empresa. Para cada dia (t_0), foi considerada uma janela de 11 dias ($-5; t_0; +5$). Dentro dessa janela, é calculada a mediana do volume de *tweets* (VT_b), que serve como valor de base. Em seguida, a atividade do dia central (VT_{t_0}) é comparada com a mediana da janela (VT_b), utilizando a fórmula da Equação 13:

$$\phi(t_0) = \frac{VT_{t_0} - VT_b}{\max(VT_b, n_{min})} \quad (13)$$

Em que, $n_{min} = 5$ foi usado para evitar problemas com valores muito baixos de atividade. Este valor foi ajustado em relação ao estudo de Ranco et al. (2015), que usou $n_{min} = 10$, devido ao menor volume geral de dados coletados nesta pesquisa, o que permitiu capturar variações menores na atividade de *tweets*, que podem ser significativas em um contexto de menor escala. Posteriormente, o dia foi considerado um pico de atividade quando $\phi(t_0)$ ultrapassou o limite ϕ_t , que foi definido como 2, tal como em Ranco et al. (2015), que identificou que esse limite reduz sobreposições entre os picos e permite detectar um bom número de eventos relevantes.

Após a identificação dos picos, eles foram classificados com base no sentimento predominante dos *tweets* (positivo, neutro ou negativo) e na sua associação com anúncios de resultados financeiros. Foi realizada uma verificação manual, a partir de dados obtidos da

Comissão de Valores Mobiliários (CVM)⁶, para identificar quantos anúncios de resultados (EA) ocorreram no período, e quantos dias de picos coincidem diretamente com os dias dos anúncios de resultado. Posteriormente, foram identificados os picos que não estavam relacionados aos EA. Para isso, foi removido o intervalo $[t - 1; t + 1]$ em torno de cada dia de EA, em que t representa o dia do anúncio, e a detecção de picos foi realizada novamente, resultando na detecção dos picos não-EA.

3.4.2 Correlação de Pearson e Causalidade de Granger

Para explorar a relação entre sentimento diário e volume de *tweets* e com os retornos das ações e volume de negociações, foram aplicados o teste de correlação de Pearson e o teste de causalidade de Granger. A correlação de Pearson foi utilizada para avaliar a dependência linear entre as variáveis do Twitter e as variáveis de mercado. Em conformidade com Belfiore (2015), o coeficiente de correlação de Pearson para duas séries temporais X_i e Y_i , é calculado conforme segue na equação 14:

$$\rho(X, Y) = \frac{\sum_{i=1}^n (X_i - \bar{X}) \times (Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \times \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (14)$$

Em que:

ρ é o coeficiente de correlação de Pearson

X_i e Y_i são os valores individuais das variáveis

$\bar{X}$ e $\bar{Y}$ são as médias das variáveis

n é o número de pares de dados

O teste de causalidade de Granger foi aplicado para verificar se as variáveis originadas do Twitter ajudam a prever os retornos das ações e o volume de negociações. Isso, porque, a causalidade de granger é um método estatístico utilizado para determinar se uma variável temporal x ajuda a prever outra variável temporal y (Stern, 2011). Diz-se que x granger-causa y se os valores passados de x fornecem informações úteis para prever os valores de y além do que é fornecido pelos valores passados de y sozinho (Granger, 1969). A causalidade de Granger requer estimar as regressões:

$$y_t = \beta_{1,0} + \sum_{i=1}^p \beta_{1,i} y_{t-i} + \sum_{j=1}^p \beta_{1,p+j} x_{t-j} + \varepsilon_{1t} \quad (15.1)$$

⁶ Disponível em <https://sistemas.cvm.gov.br/>.

$$x_t = \beta_{2,0} + \sum_{i=1}^p \beta_{2,i} y_{t-i} + \sum_{j=1}^p \beta_{2,p+j} x_{t-j} + \varepsilon_{1t} \quad (15.2)$$

Em que:

y e x são as variáveis

p é número de defasagens

β é o intercepto da equação

t é o tempo atual

i é a defasagem da variável dependente

j é a defasagem da variável independente

ε representa o termo de erro

Foi realizado o teste de estacionariedade usando o teste de Dickey-Fuller aumentado (ADF) para dados em painel, garantindo que as séries temporais fossem estacionárias antes da aplicação dos modelos vetoriais autorregressivos (VAR). Com a confirmação da estacionariedade, os modelos VAR foram executados e o teste de Granger foi utilizado para avaliar a causalidade entre as variáveis, considerando até 4 dias de defasagens, o que permitiu analisar as interações entre Sentimento Diário, Volume de *Tweets*, Retornos das Ações e Volume Negociado em cada empresa.

4 ANÁLISE E DISCUSSÃO DOS RESULTADOS

4.1 ESTATÍSTICA DESCRITIVA

A estatística descritiva tem como objetivo apresentar e avaliar as características da amostra, sem realizar conclusões ou inferências sobre uma população maior, detalhando as características do banco de dados por meio de medidas de frequência, posição ou dispersão (Fávero et al., 2009, 2014). Este tópico, com especial enfoque na frequência, desagrega a análise em relação a distribuição do nível de governança, as classes de ações das empresas, a distribuição dos setores econômicos, subsetores e segmentos correspondentes, proporcionando uma visão detalhada da diversidade da amostra. Além disso, serão fornecidas informações sobre a quantidade de *tweets* extraídos, em português, por empresa e quantidade de dias nos quais foram registrados *tweets*, após a atribuição do Sentimento Diário.

Em relação à classificação do nível de governança das 24 empresas-amostra, destaca-se que o tipo NM emerge como o mais prevalente, com 10 empresas representadas, totalizando 41,67%. Em seguida, o tipo N1 é observado em 6 empresas, perfazendo 25% do total. Assim, as empresas NM e N1 representam 66,67% da amostra. Por outro lado, tanto o tipo tradicional quanto o N2 são igualmente representados, cada um em 4 empresas, o que corresponde a

16,67%. Portanto, evidencia-se uma clara predominância do Novo Mercado em comparação com as demais classificações.

Quanto às classes de ações das empresas da amostra, destaca-se a predominância da classe ON (ordinárias), representando 45,83% do total, com 11 observações. Em seguida, surge a classe PN (preferenciais) com 9 observações, correspondendo a 37,50% do total. As classes menos representadas são PNB (preferenciais tipo B) com 2 observações (8,33%), e PNA (preferenciais tipo A) e UNT (combinação de ações ON e PN), cada uma com apenas uma observação. Assim, evidencia-se a predominância das classes ON e PN, enquanto as demais classes apresentam uma presença consideravelmente menor no conjunto de dados.

Na análise da amostra por setores, subsetor e segmentos econômicos, os de Materiais Básicos e Consumo Não Cíclico são os mais representativos, cada um com cinco empresas (20,83%). O setor de Utilidade Pública vem em seguida com quatro empresas (16,67%). Os setores de Petróleo, Gás e Biocombustíveis, Bens Industriais e Financeiro têm três empresas cada (12,5%), enquanto Comunicações é o menos representado, com uma empresa (4,17%). Nos subsetores, Intermediários Financeiros, Petróleo, Gás e Biocombustíveis e Energia Elétrica se destacam, cada um com três empresas (12,5%). Siderurgia e Metalurgia, Comércio e Distribuição e Transporte possuem duas empresas cada (8,33%). Os demais subsetores— Água e Saneamento, Alimentos Processados, Bebidas, Madeira e Papel, Material de Transporte, Mineração, Produtos de Uso Pessoal e de Limpeza, Químicos e Telecomunicações —são representados por uma empresa cada (4,7%) (Tabela 5).

Tabela 5 - Distribuição setor econômico, subsetor e segmento Bovespa

Setor Econômico Bovespa	Subsetor Bovespa	Segmento Bovespa	Qtd	%
Bens industriais	Material de transporte	Material aeronáutico e de defesa	1	4,2%
	Transporte	Transporte aéreo	2	8,3%
Comunicações	Telecomunicações	Telecomunicações	1	4,2%
Consumo não cíclico	Alimentos processados	Carnes e derivados	1	4,2%
	Bebidas	Cervejas e refrigerantes	1	4,2%
	Comércio e distribuição	Alimentos	2	8,3%
	Produtos de uso pessoal e de limpeza	Produtos de uso pessoal	1	4,2%
Financeiro	Intermediários financeiros	Bancos	3	12,5%
Materiais básicos	Madeira e papel	Papel e celulose	1	4,2%
	Mineração	Minerais metálicos	1	4,2%
	Químicos	Petroquímicos	1	4,2%
	Siderurgia e metalurgia	Siderurgia	2	8,3%
Petróleo gás e biocombustíveis	Petróleo gás e biocombustíveis	Exploração refino e distribuição	3	12,5%
Utilidade pública	Água e saneamento	Água e saneamento	1	4,2%
	Energia elétrica	Energia elétrica	3	12,5%

Nota. Qtd = Quantidade. Fonte: Elaborado pela pesquisa (2024) a partir dos dados coletados.

Em relação ao volume de *tweets*, foram coletadas 529.180 postagens para o período analisado, sendo classificados em negativos (42%), neutros (19%) e positivos (39%). Entre as empresas com maior volume de postagens, destacam-se Itaú Unibanco (145.529), Petrobras (114.800), Bradesco (73.378) e Santander BR (62.660). Em termos de predominância de *tweets* negativos, cabe destacar Bradesco (56% do total) e Santander BR (55%). No entanto, empresas como Sid Nacional, Azul e Suzano mostraram uma predominância de *tweets* positivos, com 87%, 86% e 84%, respectivamente. Em relação ao número de dias com *tweets*, 63% dos dias apresentaram predominância de positivos, enquanto 30% foram negativos e 7% neutros. Ambev e Vale se destacaram, com 97% dos dias com *tweets* positivos cada, seguidas por Petrobras, que apresentou 90% dos dias com *tweets* positivos, apesar do elevado volume de postagens negativos (Tabela 6).

Tabela 6 – Volume de tweets e de dias com tweets por empresa

Painel a) Volume de tweets

Empresa	Negativo		Neutro		Positivo		Total
	Qtd	%	Qtd	%	Qtd	%	
Assai	5.784	30%	3.236	17%	9.999	53%	19.019
Cemig	9.351	50%	3.213	17%	6.298	33%	18.862
Sabesp	8.553	42%	4.523	22%	7.143	35%	20.219
Copel	5.350	43%	2.035	16%	4.994	40%	12.379
Petrobras	32.044	28%	33.175	29%	49.581	43%	114.800
Embraer	766	16%	972	20%	3.078	64%	4.816
ItaúUnibanco	75.576	52%	23.140	16%	46.813	32%	145.529
Ambev	4.162	21%	3.275	17%	12.363	62%	19.800
Bradesco	41.224	56%	11.016	15%	21.138	29%	73.378
Braskem	2.731	32%	2.034	24%	3.786	44%	8.551
Vale	285	4%	1.009	15%	5.431	81%	6.725
Santander BR	34.749	55%	9.592	15%	18.319	29%	62.660
Gerdau	385	14%	500	18%	1.841	68%	2.726
Gol	200	11%	289	16%	1.306	73%	1.795
Eletrobras	1.848	23%	2.377	30%	3.744	47%	7.969
Sid Nacional	26	2%	181	12%	1.336	87%	1.543
Azul	39	3%	165	11%	1.285	86%	1.489
Suzano	34	3%	175	14%	1.074	84%	1.283
BRF	26	2%	152	11%	1.207	87%	1.385
Cosan	41	5%	130	17%	578	77%	749
P. Acucar-Cbd	29	3%	204	18%	927	80%	1.160
Telef Brasil	23	3%	117	16%	590	81%	730
Grupo Natura	49	5%	108	10%	884	85%	1.041
Ultrapar	21	4%	95	17%	456	80%	572
Total	223.296	42%	101.713	19%	204.171	39%	529.180

Painel b) Contagem de dias com tweets por empresa

Empresa	Negativo		Neutro		Positivo		Total
	Qtd	%	Qtd	%	Qtd	%	
Assai	72	10%	35	5%	622	85%	729
Cemig	527	72%	47	6%	155	21%	729
Sabesp	444	61%	63	9%	221	30%	728
Copel	430	59%	65	9%	230	32%	725
Petrobras	64	9%	7	1%	630	90%	701
Embraer	62	9%	63	9%	572	82%	697
ItauUnibanco	618	89%	14	2%	63	9%	695
Ambev	9	1%	10	1%	665	97%	684
Bradesco	613	93%	11	2%	32	5%	656
Braskem	139	21%	106	16%	411	63%	656
Vale	4	1%	14	2%	623	97%	641
Santander BR	531	91%	11	2%	44	8%	586
Gerdaul	59	10%	73	13%	442	77%	574
Gol	67	14%	62	13%	342	73%	471
Eletrobras	59	13%	57	12%	342	75%	458
Sid Nacional	2	0%	30	7%	397	93%	429
Azul	3	1%	24	6%	356	93%	383
Suzano	4	1%	31	8%	330	90%	365
BRF	3	1%	25	7%	334	92%	362
Cosan	13	4%	41	13%	258	83%	312
P. Acucar-Cbd	7	2%	29	9%	272	88%	308
Telef Brasil	8	3%	49	17%	236	81%	293
Grupo Natura	11	4%	18	6%	256	90%	285
Ultrapar	14	6%	33	13%	205	81%	252
Total	3.763	30%	918	7%	8.038	63%	12.719

Nota. Qtd = Quantidade. Fonte: Elaborado pela pesquisa (2024) a partir dos dados coletados.

Em relação as estatísticas descritivas, durante o período analisado, constataram-se que o retorno das ações (Ret) e o índice IBOVESPA (Ret_Ibov), que representa o desempenho agregado do mercado, exibiram baixos valores diários médios positivos de 0,0001 e 0,0005, respectivamente. Quando esses retornos são anualizados, os valores médios aumentam para 0,0289 (Ret_Anual) e 0,1249 (Ret_Ibov_Anual). A média do Ibovespa foi inferior àquela reportada por Souza e Martins (2022) para o período de 2017 a 2018, porém superior ao valor obtido por Silva (2018) no período de 2011 a 2017. Assim como nos estudos anteriores, também nesta pesquisa verificou a expectativa de um retorno positivo, corroborando a evidência de que mercados emergentes podem proporcionar uma melhor diversificação e potencialmente maiores retornos (Atilgan et al., 2015), o que torna a bolsa brasileira particularmente atrativa para investidores tanto nacionais quanto internacionais (Tabela 7).

Tabela 7 - Estatística descritiva (01/01/2022 a 31/12/2023)*Painel a) Sentimento e Volume de Tweets*

Variáveis	Média	Mediana	Desvio Padrão	Assimetria	Curtose	Obs ^a	Mínimo	Máximo
SD	0,409	0,481	0,565	-0,423	1,950	9.515,000	-1,000	1,000
SD_Neg	-0,343	-0,304	0,231	-1,391	4,866	2.528,000	-1,000	-0,004
SD_Neu	0,000	0,000	0,000	-	-	593,000	0,000	0,000
SD_Pos	0,744	1,000	0,320	-0,774	2,057	6.394,000	0,001	1,000
VT	46,511	11,000	237,107	51,267	3.423,620	9.515,000	1,000	17.524,000
VTD_Pos	18,725	6,000	67,806	39,344	2.322,709	9.116,000	1,000	4.543,000
VTD_Neu	12,938	4,000	54,567	32,388	1.471,664	6.534,000	1,000	2.939,000
VTD_Neg	32,301	8,000	167,501	46,204	2.526,422	5.799,000	1,000	10.042,000

Painel b) Retornos e Volatilidades

Variáveis ^b	Média	Mediana	Desvio Padrão	Assimetria	Curtose	Mínimo	Máximo
Ret	0,00011	0,00000	0,02616	0,27349	10,71786	-0,21561	0,32196
Ret_Anual ^c	0,02889	0,00000	6,59337	0,27349	10,71786	-54,33350	81,13477
Ret_Ibov	0,00050	0,00036	0,01208	0,01881	3,57323	-0,03407	0,05393
Ret_Ibov_Anual ^c	0,12496	0,09092	3,04485	0,01881	3,57323	-8,58595	13,59134
PK	0,00057	0,00032	0,00099	18,13594	661,94600	0,00002	0,04643
GK	0,00057	0,00035	0,00094	22,87938	1.008,40900	0,00002	0,04766
RS	0,00057	0,00034	0,00098	23,89216	1.078,75500	0,00000	0,04990
YZ	0,00070	0,00040	0,00158	23,78673	839,13880	0,00002	0,06986

Painel c) Volumes de Negociações

Variáveis ^b	Média	Mediana	Desvio Padrão	Assimetria	Curtose	Mínimo	Máximo
VN	26.916	18.762	24.879,04000	3,23447	20,81849	2.458	355.101
VN_Ibov	2.457.576	2.422.204	529.351,40000	0,84658	5,12254	1.219.872	5.234.183

Notas. Obs = Número de observações. SD = Sentimento Diário; SD_Neg, SD_Neu, SD_Pos = Sentimento Diário Negativo, Neutro e Positivo, respectivamente; VT = Volume Total de Tweets; VT_Pos, VT_Neu, VT_Neg = Volume de Tweets Positivos, Neutros e Negativos; Ret = Retorno Diário da Ação; Ret_Ibov = Retorno Diário do Ibovespa; PK, GK, RS, YZ = Volatilidades de Parkinson, Garman-Klass, Rogers-Satchell e Yang-Zhang; Ret_Anual = Retorno Anualizado das Ações; Ret_Ibov_Anual = Retorno Anualizado do Ibovespa. VN = Volume de Negociações; VN_Ibov = Volume de Negociações do Ibovespa; Fonte: Elaborado pela pesquisa (2024) a partir dos dados coletados.

^a Refere-se às 24 empresas analisadas considerando o número de dias com *tweets* de cada uma pós união com a base de atividade da bolsa de valores brasileira.

^b Cada variável possui 11.952 observações.

^c Retorno diário na forma de retorno anualizado simples, por meio da equação $R_{anualizado} = \frac{1}{N} \sum_{i=1}^N R_t * 252$.

O volume de negócios das ações (VN) apresentou uma média de 26.916 operações diárias e uma mediana de 18.762, evidenciando uma heterogeneidade entre os dias. Essa variação é confirmada pelo elevado desvio padrão (0,02616) e pela amplitude de oscilação, dada pela diferença entre o valor mínimo (2.458) e o máximo (355.101). Por sua vez, o índice Ibovespa (VN_Ibov) registrou um valor médio de operações consideravelmente maior, com uma média de 2.457.576 e uma mediana de 2.422.204. Comparativamente, mercados emergentes apresentam volumes inferiores ao de países desenvolvidos, como o Reino Unido com média de 669.805.966 (Nisar & Yeung, 2018) e a indústria de tecnologia dos EUA com 4.162.000 (Teti et al., 2019), por exemplo.

As volatilidades, medidas por diferentes metodologias — Parkinson (PK), Garman e Klass (GK), Rogers e Satchell (RS) e Yang e Zhang (YZ) —, apresentam médias estáveis, variando entre 0,00057 e 0,00070. Essa consistência nas médias e a alta correlação (Tabela 8)

entre as medidas de volatilidade (todas acima de 0,74, com significância de 1%) indicam que os diferentes métodos produzem resultados semelhantes. As correlações entre o retorno diário (Ret) e as volatilidades, embora com coeficientes não tão elevados, também se apresentaram significativas ao nível de 1%. Esses resultados mostram que a escolha do método para calcular o retorno — seja pelo *log* dos preços, pela amplitude diária (*high - low*) ou pelos preços de abertura e fechamento — não influencia significativamente a análise. Dessa forma, para simplificar as etapas subsequentes, é utilizado o retorno convencional.

Tabela 8 – Matriz correlação entre Retorno e Volatilidades

Variáveis	Ret	PK	GK	RS	YZ
Ret	1,000				
PK	0,089***	1,000			
GK	0,107***	0,946***	1,000		
RS	0,111***	0,863***	0,975***	1,000	
YZ	0,109***	0,744***	0,798***	0,782***	1,000

Nota. *** Significativo ao nível de 1%. Fonte: Dados da pesquisa (2024).

As médias de sentimento diário variaram significativamente entre os setores analisados (Apêndice F). Setores como Materiais básicos (0,73295), Comunicações (0,79204) e Bens industriais (0,69620) apresentaram as maiores médias em termos absolutos, sugerindo uma percepção amplamente favorável. Em contraste, o setor Financeiro (-0,27932) e de Utilidade Pública (-0,01662) tiveram uma média negativa, indicando uma visão desfavorável ao longo do período. No geral, o Sentimento Diário (SD) extraído dos *tweets* apresentou uma média (Tabela 7) de 0,4089 e uma mediana de 0,4815, revelando que, embora existam diferenças setoriais, o sentimento predominante na amostra foi positivo. O desvio padrão de 0,5652 revela uma variabilidade considerável nos sentimentos ao longo do período, embora a dispersão seja limitada em razão da natureza do intervalo fixo de -1 a 1 da variável.

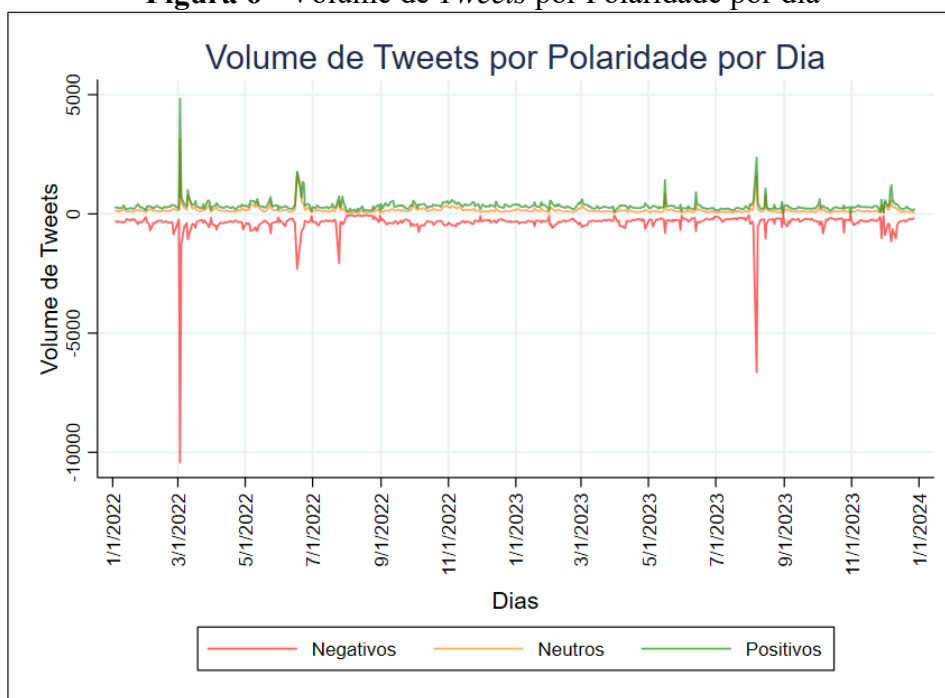
De forma semelhante, Souza e Martins (2022) também identificaram uma média positiva nos sentimentos dos *tweets* no mercado brasileiro sobre o Ibovespa, e Gu e Kurov (2020) encontraram um sentimento médio positivo de 0,063 no Twitter, em relação as empresas do índice Russell 3000. Por outro lado, Silva (2018) e Galdi e Gonçalves (2018) ao analisarem sentimentos a partir de notícias de mídia especializada, identificaram que as notícias no mercado brasileiro eram predominantemente negativas. Similarmente, Almeida et al. (2022) encontraram uma tendência ao pessimismo nos *tweets* sobre a Petrobras, o que se difere dos achados desta pesquisa, que identificou uma média de sentimento diário de 0,579 para o setor de petróleo, gás e biocombustíveis (Apêndice F).

A diferença entre os resultados pode ser atribuída a diversos fatores. Por exemplo, o período analisado pode ter impacto significativo, uma vez que diferentes períodos capturam

distintos contextos econômicos, sociais e políticos que afetaram o comportamento dos investidores e o sentimento do público. Além disso, a origem e o conteúdo dos dados analisados, uma vez que esta pesquisa considera *tweets* de diversas empresas, o que implica um público mais amplo e heterogêneo, muitas vezes emocional e reativo aos eventos, como divulgações de resultados.

Em relação ao Volume de *Tweets* (VT), a média diária registrada foi de aproximadamente 46 por empresa, enquanto a mediana foi de 11, evidenciando a variabilidade nas postagens diárias ao longo do período analisado. O volume mínimo diário foi de 1 *tweet* por empresa, e o máximo chegou a 17.524, revelando oscilações significativas na atividade. A média foi inferior ao patamar de 640 *tweets* sobre o IBOVESPA encontrado por Souza e Martins (2022). Os volumes diários de tweets positivos (VTD_Pos) e negativos (VTD_Neg) apresentaram médias de 18 e 32 tweets, respectivamente. Assim, observa-se que, em média, há mais mensagens negativas do que positivas diariamente, com períodos de expansão no volume de tweets negativos em determinados dias (Figura 6). Entretanto, como os valores de sentimentos positivos são, em geral, superiores aos negativos em termos absolutos, os positivos prevaleceram na média geral, resultando em uma média de sentimento diário positivo.

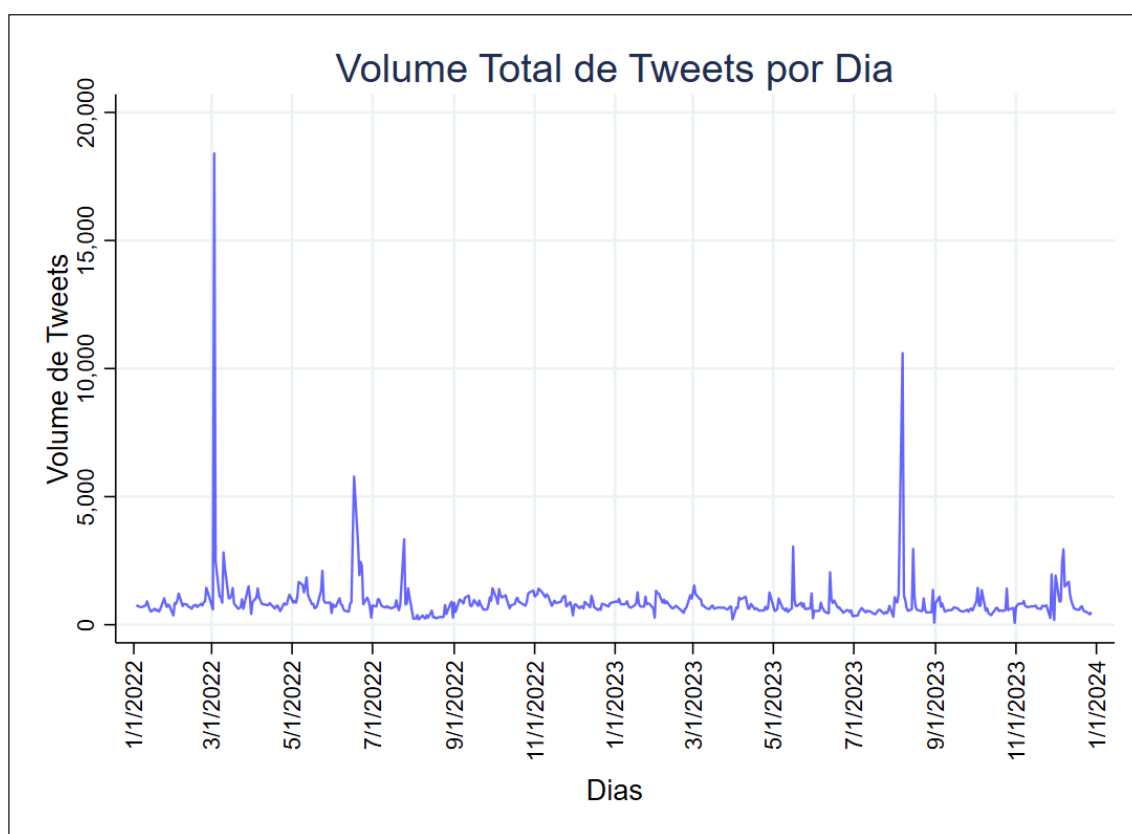
Figura 6 - Volume de *Tweets* por Polaridade por dia



Nota. Os dias no gráfico estão apresentados no formato: Mês/Dia/Ano. O volume de *tweets* negativos foi multiplicado por -1 e plotado para baixo de 0 no eixo y para facilitar a distinção visual das polaridades. Fonte: A partir dos dados coletados.

Na análise do volume total de *tweets* por dia durante o período em estudo, evidencia-se que, embora o volume diário de postagens se mantenha relativamente constante na maior parte do tempo, há picos esporádicos e acentuados que indicam aumentos significativos na atividade (Figura 7). Esses aumentos sugerem momentos em que eventos específicos ou notícias relevantes ocorreram, impulsionando um aumento abrupto no volume de interações. A presença desses picos revela que, embora o fluxo de *tweets* diários pareça estável, eventos têm o potencial de gerar um interesse e uma interação pública substancialmente maiores em relação às empresas.

Figura 7 - Volume de *Tweets* por dia



Nota. Os dias no gráfico estão apresentados no formato: Mês/Dia/Ano. Fonte: A partir dos dados coletados.

Referente aos anúncios de resultados, foram identificadas 192 datas de divulgação, distribuídas uniformemente entre os anos de 2022 e 2023, com 96 anúncios em cada ano (Tabela 9). Em 2022, as divulgações foram referentes ao 4º trimestre de 2021 e aos 1º, 2º e 3º trimestres de 2022, com 24 anúncios em cada período. Da mesma forma, em 2023, ocorreram divulgações referentes ao 4º trimestre de 2022 e aos 1º, 2º e 3º trimestres de 2023, também com 24 anúncios por período.

Tabela 9 – Anúncios de Resultado*Painel a) Anúncios de resultados por setor*

Setores	2022	2023	Total	%
Consumo não cíclico	20	20	40	21%
Materiais básicos	20	20	40	21%
Utilidade pública	16	16	32	17%
Bens Industriais	12	12	24	13%
Financeiro	12	12	24	13%
Petróleo gás e biocombustíveis	12	12	24	13%
Comunicações	4	4	8	4%
Total Geral	96	96	192	100%

Painel b) Anúncios de resultados por dias da semana

Dia da Semana	2022	2023	Total	%
quinta-feira	33	29	62	32%
quarta-feira	34	24	58	30%
segunda-feira	10	26	36	19%
terça-feira	9	15	24	13%
sexta-feira	10	1	11	6%
sábado		1	1	1%
Total Geral	96	96	192	100%

Painel c) Anúncios de resultados por setores e por dias da semana

Setores	Dias da semana						Total
	segunda-feira	terça-feira	quarta-feira	quinta-feira	sexta-feira	sábado	
Consumo não cíclico	12	3	13	12	-	-	40
Materiais básicos	5	4	21	10	-	-	40
Utilidade pública	4	4	5	12	6	1	32
Bens Industriais	8	2	2	11	1	-	24
Financeiro	5	5	4	10	-	-	24
Petróleo, gás e biocombustíveis	2	2	10	6	4	-	24
Comunicações	-	4	3	1	-	-	8
Total	36	24	58	62	11	1	192

Fonte: A partir de dados coletados pela pesquisa (2024).

Devido às diferenças no número de empresas por setor, os setores mais representativos na amostra — Consumo não cíclico e Materiais Básicos — corresponderam cada um por 21% das datas de anúncios de resultados. Observa-se ainda uma concentração dos anúncios às quartas e quintas-feiras, totalizando 62% do total, o que sugere uma preferência por divulgar resultados no meio da semana para maximizar a visibilidade e mitigar possíveis impactos externos ou efeito-calendário. Vale destacar que, após a integração dos dados com a base principal, que considera apenas dias úteis de operação da bolsa, o total de divulgações foi reduzido para 191, devido a uma data que coincidiu com um sábado, quando a bolsa não opera. Essa ocorrência foi associada a um anúncio do setor de Utilidade Pública em 2023.

Conforme a Tabela 10, nos dias de anúncio de resultados, observou-se um aumento nas médias de diversas métricas em comparação ao período total analisado. O Sentimento Diário (SD) apresentou uma média mais alta (0,5211 em comparação com 0,409) e uma mediana

superior (0,6491 em comparação com 0,481), indicando uma percepção mais positiva do mercado nesses dias. O Volume de *Tweets* (VT) também aumentou, com uma média diária de 104 *tweets* em relação aos 46 do período geral, refletindo um maior engajamento na plataforma. Dos 191 dias de anúncio registrados na base de dados, houve publicações no Twitter em 176 deles, o que demonstra um aumento do envolvimento dos *stakeholders* nesses dias.

Tabela 10 – Estatísticas descritivas em dias de anúncio de resultado

Estatísticas	SD	VT	Ret	Ret_Ibov	VN	VN_Ibov
Média	0,5211	104,9545	0,0031	0,0007	35.288,0400	2.658.342,0000
Mediana	0,6491	18,0000	0,0024	0,0007	24.295,0000	2.502.851,0000
Desvio Padrão	0,4898	732,5500	0,0422	0,0135	30.531,1500	693.633,2000
Assimetria	-0,8018	12,8443	1,9043	-0,4350	2,2325	1,4101
Curtose	2,8706	168,5207	21,7167	2,9614	8,8536	5,9657
Obs	176,0000	176,0000	191,0000	191,0000	191,0000	191,0000
Mínimo	-1,0000	1,0000	-0,1964	-0,0341	5.278,0000	1.695.700,0000
Máximo	1,0000	9.691,0000	0,3220	0,0274	183.039,0000	5.234.183,0000

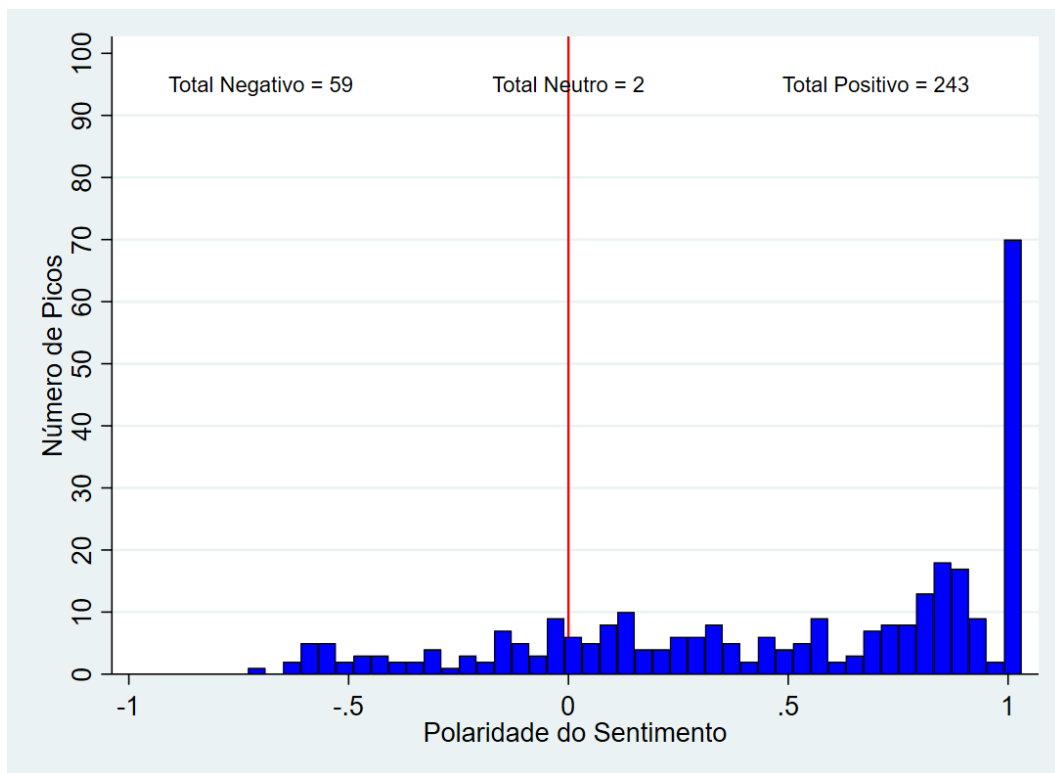
Notas. Obs = Número de observações. SD = Sentimento Diário; VT = Volume Total de Tweets; Ret = Retorno Diário da Ação; Ret_Ibov = Retorno Diário do Ibovespa; VN = Volume de Negociações; VN_Ibov = Volume de Negociações do Ibovespa. Fonte: Elaborado pela pesquisa (2024) a partir dos dados coletados.

Além disso, os retornos das ações e do Ibovespa apresentaram aumentos (0,0031 e 0,0007, respectivamente), sugerindo uma reação favorável do mercado. O volume de negociações, tanto das ações (35.288) quanto do Ibovespa (2.658.342), também foi superior, evidenciando maior atividade dos investidores durante os dias de anúncio. Esses fatores combinados demonstram que os dias de anúncio de resultados são marcados por maior participação do mercado e uma percepção mais otimista.

4.2 IDENTIFICAÇÃO DOS PICOS DE ATIVIDADE NO TWITTER

A análise da identificação de picos de atividade no Twitter, baseada na metodologia de Ranco et al. (2015), resultou na detecção de 304 picos no período de análise (Apêndice A). Esses picos foram classificados de acordo com o sentimento predominante das postagens: 59 picos apresentaram sentimento negativo, 2 picos foram neutros e 243 picos exibiram sentimento positivo. Essa distribuição evidencia que a maioria dos picos de atividade no Twitter está associada a sentimentos predominantemente positivos, conforme ilustrado na Figura 8, que mostra a distribuição da polaridade do sentimento para os picos detectados.

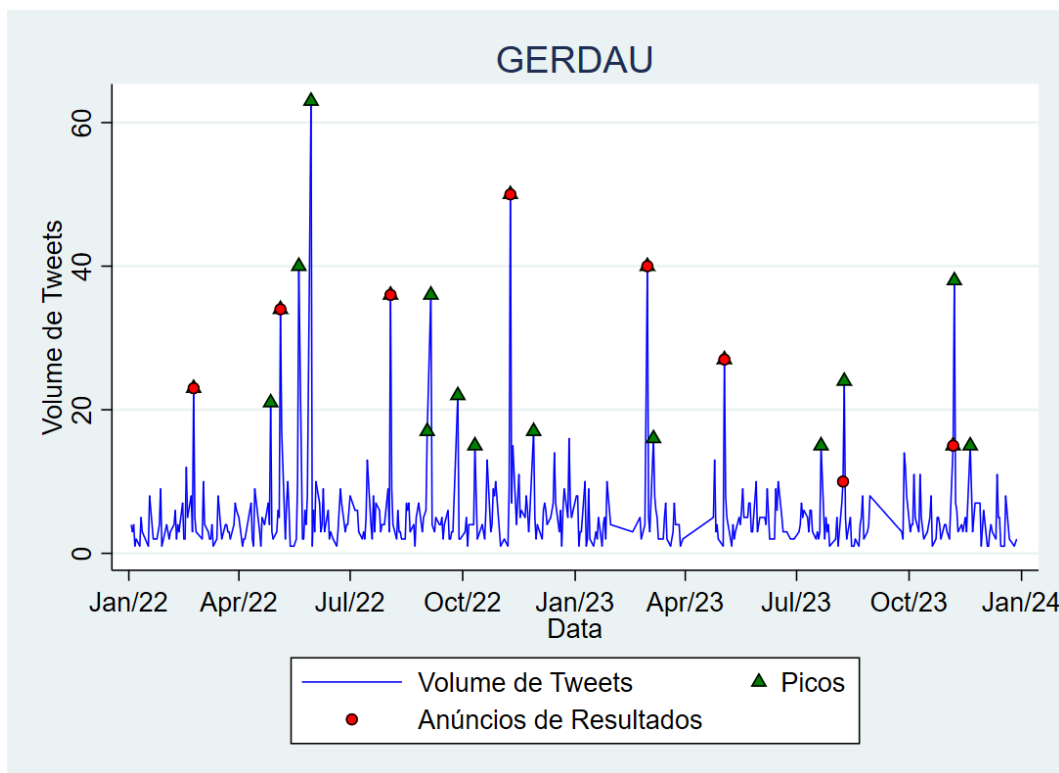
Figura 8 - Distribuição do sentimento para os 304 picos detectados no Twitter



Fonte: A pesquisa (2024).

Durante o período analisado, foram identificados 192 anúncios de resultados financeiros (EA), dos quais 30 picos coincidiram exatamente com as datas desses anúncios (Apêndice B). Isso indica que o *recall*, definido como a proporção de eventos de anúncios de resultados corretamente detectados como EA, foi de 15,6% no procedimento de identificação de picos utilizado. Em comparação, Ranco et al. (2015) reportou uma taxa de coincidência de 78%, enquanto Sprenger et al. (2014) reportaram um *recall* de 33%. O percentual de coincidência mais baixo observado neste estudo pode ser atribuído a diversas razões, incluindo o menor tamanho da base de dados, a abrangência dos tipos de *tweets* coletados, o contexto específico do mercado brasileiro e uma série temporal curta. A Figura 9 ilustra a série diária do volume de *tweets*, destacando os picos detectados e os anúncios de resultados (EA) da empresa Gerdau (Apêndice D).

Figura 9 - Série diária do volume e picos *tweets* e anúncios de resultados da Gerda



Fonte: A pesquisa (2024).

A análise desses picos mostra que, embora alguns coincidam com os dias de divulgação dos resultados financeiros, uma parcela significativa ocorre fora desses períodos. Mais especificamente, foram identificados 233 picos que não estavam associados aos anúncios de resultados financeiros (Apêndice C), considerando a exclusão do intervalo de $[t-1, t+1]$ em torno dos dias de EA. Esse resultado reforça que os eventos de maior atividade no Twitter não podem ser explicados apenas pelos anúncios de resultados, indicando que outros fatores também influenciam os picos de interações na plataforma.

4.3 ESTUDO DE EVENTOS

A Tabela 11 apresenta os retornos anormais médios (AR) de 24 empresas, juntamente com os retornos anormais acumulados médios (CAR) para cada tipo de evento e para as três categorias de polaridade: positiva, negativa e neutra. Os ARs sem significância estatística (até o nível de 10%) foram classificados como neutros, enquanto os valores significativos foram categorizados como positivos ou negativos e, em seguida, acumulados. Foram considerados quatro cenários de eventos na análise: (1) todas as observações tratadas como eventos; (2) picos de atividade; (3) picos coincidentes com anúncios de resultados (EA); e (4) picos não relacionados aos EA. As Figuras 9, 10 e 11 e 12 apresentam, respectivamente, os CARs para

esses cenários, ilustrando o comportamento dos retornos anormais ao longo da janela de -10 a +10. Os Gráficos dos AR podem ser visualizados no Apêndice E.

Tabela 11 - Valores de AR e CAR médios para cada tipo de evento

Painel a) retornos anormais médios na janela de -10 a +10

Dias	Retornos Anormais (AR)											
	Todas as Observações			Picos			EA-Pico			Não-EA Picos		
	Neg	Neu	Pos	Neg	Neu	Pos	Neg	Neu	Pos	Neg	Neu	Pos
-10	0,00032	0,00006	-0,00078	0,0043	0,0011	-0,0035	0,0001	-0,0006	-0,0137	0,0042	0,0019	-0,0049
-9	0,00018	0,00000	-0,00018	0,0003	0,0000	-0,0004	0,0018	0,0003	0,0001	0,0003	0,0000	-0,0007
-8	-0,00032	0,00007	-0,00077	-0,0009	0,0015	0,0024	-0,0023	0,0051	-0,0008	-0,0040	0,0010	0,0030
-7	0,00001	0,00010	-0,00136	0,0025	-0,0006	0,0039	0,0111	-0,0061	0,0026	0,0017	-0,0008	0,0003
-6	0,00069	-0,00003	-0,00056	0,0066	-0,0003	0,0023	0,0025	-0,0065	0,0250	0,0054	0,0011	-0,0006
-5	0,00048	-0,00003	-0,00049	-0,0004	-0,0016	0,0021	-0,0138	-0,0024	-0,0011	0,0035	-0,0018	0,0037
-4	0,00012	-0,00005	0,00003	0,0012	-0,0014	0,0038	-0,0085	-0,0020	0,0051	0,0028	-0,0021	0,0037
-3	-0,00001	-0,00004	0,00038	-0,0015	0,0043	0,0011	0,0134	0,0057	0,0081	-0,0038	0,0051	-0,0005
-2	0,00146	-0,00003	-0,00126	-0,0039	-0,0009	-0,0022	-0,0102	0,0077	-0,0028	-0,0048	-0,0020	-0,0048
-1	-0,00087	-0,00006	0,00156	-0,0039	-0,0006	0,0094	-0,0016	0,0121	0,0045	-0,0072	-0,0026	0,0057
0	-0,04111	-0,00041	0,04454	-0,0637	-0,0015	0,0754	-0,0784	0,0005	0,1007	-0,0547	-0,0018	0,0683
1	-0,00226	0,00007	0,00081	0,0045	-0,0012	0,0008	0,0064	0,0008	0,0092	0,0008	-0,0019	0,0015
2	0,00264	-0,00013	-0,00086	0,0067	0,0005	-0,0012	0,0031	0,0153	0,0070	0,0037	-0,0010	-0,0053
3	-0,00078	0,00000	0,00066	0,0030	0,0004	-0,0011	0,0071	0,0085	0,0117	-0,0019	0,0006	-0,0021
4	0,00130	-0,00004	-0,00054	-0,0002	0,0006	-0,0014	0,0135	0,0026	-0,0010	0,0005	0,0008	-0,0003
5	0,00184	-0,00011	0,00013	0,0000	-0,0009	0,0016	0,0028	0,0052	0,0040	0,0029	-0,0018	0,0011
6	-0,00008	-0,00001	0,00027	0,0012	0,0014	-0,0007	0,0079	-0,0074	-0,0004	0,0014	0,0017	-0,0015
7	0,00040	0,00005	-0,00073	0,0027	-0,0019	-0,0035	-0,0091	-0,0007	-0,0102	0,0035	-0,0031	-0,0045
8	0,00071	0,00007	-0,00118	-0,0015	0,0012	-0,0010	-0,0019	0,0013	-0,0040	-0,0012	0,0017	0,0001
9	0,00145	-0,00001	-0,00064	-0,0009	0,0006	-0,0011	-0,0151	-0,0035	-0,0072	0,0025	0,0012	0,0007
10	0,00093	0,00003	-0,00031	-0,0007	0,0013	-0,0013	-0,0046	0,0066	-0,0092	0,0016	0,0010	-0,0021

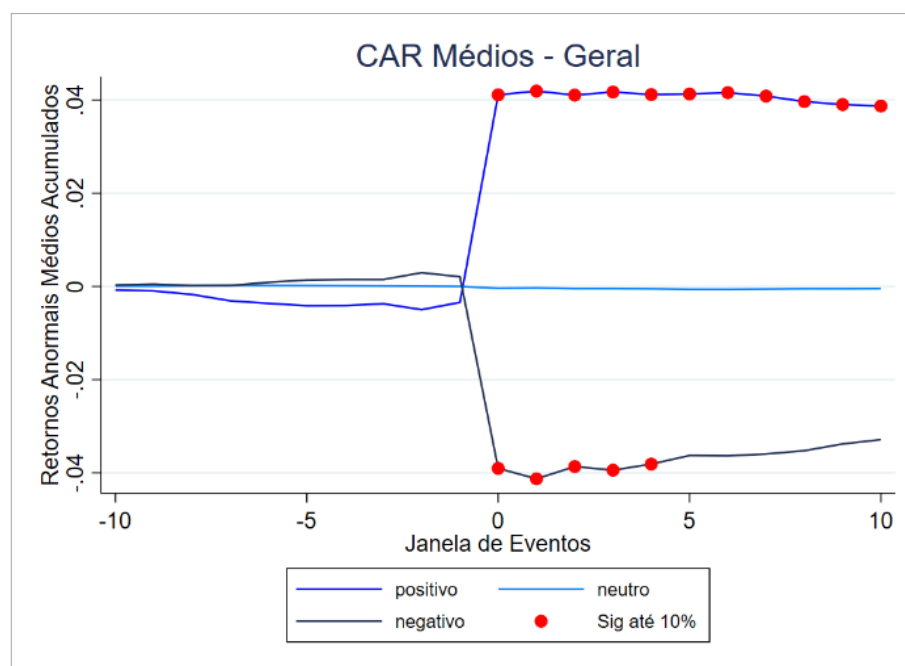
Painel b) retornos anormais médios acumulados na janela de -10 a +10

Dias	Retornos Anormais Acumulados (CAR)											
	Todas as Observações			Picos			EA-Pico			Não-EA Picos		
	Neg	Neu	Pos	Neg	Neu	Pos	Neg	Neu	Pos	Neg	Neu	Pos
-10	0,00032	0,00006	-0,00078	0,0043	0,0011	-0,0035	0,0001	-0,0006	-0,0137	0,0042	0,0019	-0,0049
-9	0,00050	0,00006	-0,00096	0,0045	0,0011	-0,0039	0,0019	-0,0002	-0,0136	0,0045	0,0019	-0,0056
-8	0,00018	0,00013	-0,00173	0,0036	0,0026	-0,0015	-0,0004	0,0049	-0,0145	0,0005	0,0029	-0,0026
-7	0,00020	0,00023	-0,00309	0,0060	0,0019	0,0024	0,0106	-0,0012	-0,0118	0,0022	0,0021	-0,0023
-6	0,00089	0,00021	-0,00365	0,0126	0,0016	0,0047	0,0131	-0,0076	0,0131	0,0076	0,0031	-0,0029
-5	0,00137	0,00018	-0,00413	0,0123	0,0000	0,0068	-0,0007	-0,0101	0,0120	0,0111	0,0013	0,0009
-4	0,00149	0,00013	-0,00410	0,0135	-0,0014	0,0106	-0,0091	-0,0121	0,0171	0,0139	-0,0008	0,0046
-3	0,00148	0,00009	-0,00373	0,0120	0,0029	0,0117	0,0043	-0,0064	0,0252	0,0100	0,0043	0,0041
-2	0,00295	0,00006	-0,00499	0,0081	0,0020	0,0095	-0,0060	0,0013	0,0225	0,0052	0,0023	-0,0007
-1	0,00208	0,00000	-0,00343	0,0041	0,0014	0,0189	-0,0076	0,0135	0,0269	-0,0020	-0,0002	0,0050
0	-0,03903	-0,00040	0,04111	-0,0595	-0,0001	0,0943	-0,0860	0,0140	0,1276	-0,0567	-0,0020	0,0733
1	-0,04128	-0,00034	0,04192	-0,0550	-0,0013	0,0952	-0,0796	0,0148	0,1368	-0,0560	-0,0039	0,0748
2	-0,03865	-0,00047	0,04106	-0,0484	-0,0008	0,0939	-0,0764	0,0301	0,1438	-0,0523	-0,0049	0,0696
3	-0,03943	-0,00046	0,04172	-0,0454	-0,0004	0,0929	-0,0694	0,0386	0,1555	-0,0542	-0,0042	0,0675
4	-0,03812	-0,00050	0,04117	-0,0456	0,0002	0,0914	-0,0558	0,0412	0,1545	-0,0536	-0,0035	0,0672
5	-0,03628	-0,00061	0,04130	-0,0456	-0,0007	0,0930	-0,0530	0,0465	0,1585	-0,0507	-0,0052	0,0683
6	-0,03636	-0,00061	0,04157	-0,0444	0,0007	0,0923	-0,0451	0,0390	0,1581	-0,0493	-0,0035	0,0667
7	-0,03596	-0,00056	0,04084	-0,0417	-0,0013	0,0887	-0,0542	0,0383	0,1479	-0,0458	-0,0066	0,0623
8	-0,03525	-0,00049	0,03966	-0,0432	-0,0001	0,0878	-0,0561	0,0396	0,1439	-0,0469	-0,0049	0,0624
9	-0,03380	-0,00050	0,03903	-0,0441	0,0005	0,0866	-0,0712	0,0361	0,1367	-0,0445	-0,0038	0,0631
10	-0,03287	-0,00047	0,03872	-0,0448	0,0018	0,0853	-0,0758	0,0428	0,1276	-0,0429	-0,0027	0,0610

Nota. Neg = Negativo, Neu = Neutro, Pos = Positivo. Os números destacados indicam significância até o nível de 10%. Fonte: Elaborado pela pesquisa (2024).

A Figura 10 ilustra, por meio de linhas, os Retornos Anormais Médios Acumulados (CAR) ao longo da série temporal, para o conjunto de empresas analisadas, em torno dos dias de picos e anúncios de resultados. Antes do evento ($t=0$), os CAR permanecem estáveis e próximos de zero, sem uma tendência definida. Em $t=0$, ocorre um choque exógeno, sendo representado por um ponto de inflexão. A Figura 10 indica o momento em que as informações do evento são incorporadas aos preços das ações, tanto para o caso de aumento nos CAR positivos quanto para o caso de queda nos negativos.

Figura 10 - Retornos Anormais Médios Acumulados para série temporal acionária



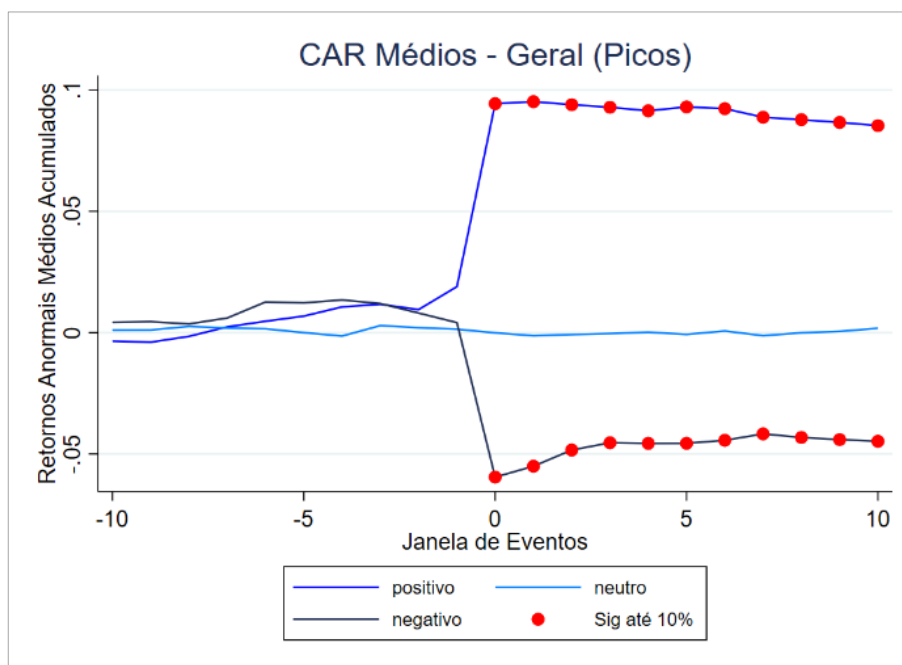
Fonte: Elaborado pela pesquisa (2024).

A Figura 10 sugere um impacto mais duradouro do evento nos retornos positivos (até $t=10$) em comparação aos negativos (até $t=4$). A estabilização observada nos dias subsequentes ao evento aponta que o mercado absorve rapidamente as informações, incorporando-as aos preços em curto prazo. Após o impacto inicial, os retornos apresentam menor variabilidade, indicando uma possível eficiência no processo de assimilação das informações pelo mercado, em linha com a hipótese de mercado eficiente em sua forma semiforte.

A Figura 11 apresenta os CAR para os dias de pico de atividade no Twitter, incluindo os coincidentes com os anúncios de resultados. Observando os dias de picos de atividade de *tweets*, percebe-se que esses eventos exercem um impacto significativo nos retornos anormais acumulados. De forma análoga, o CAR médio dos eventos aumenta após picos de sentimento positivo e diminui após picos de sentimento negativo, entretanto, o impacto dos eventos

positivos é quase duas vezes maior, em termos absolutos, do que os negativos. Isso reflete uma reação mais forte e imediata do mercado a informações favoráveis. Além disso, a amplitude do CAR médio nos dias de pico é maior em comparação à análise da série temporal como um todo (Figura 10) reforçando a importância desses eventos para o comportamento do mercado. Esse comportamento pode ser explicado, em parte, pela predominância de picos no Twitter associados a sentimentos positivos (Figura 8).

Figura 11 - Retornos Anormais Médios Acumulados para pico de atividade no Twitter

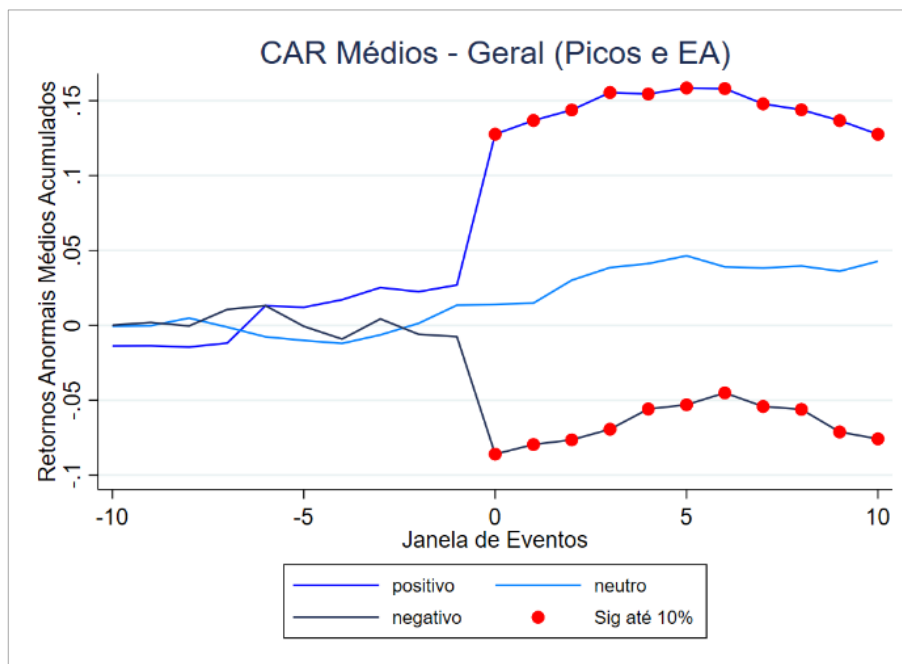


Fonte: Elaborado pela pesquisa (2024).

Os valores de CAR mostram-se significativos ao nível de 10% por até dez dias após os eventos de sentimento positivo e negativo. Esses resultados permitem rejeitar a hipótese nula de que os eventos não impactam os retornos das ações. A estabilização observada no gráfico após $t=0$ sugere que o mercado absorveu rapidamente as informações associadas aos picos detectados no Twitter, incorporando-as aos preços das ações em um curto intervalo de tempo.

Os Retornos Anormais Médios Acumulados nos dias que coincidem com picos de atividade no Twitter e anúncios de resultados (Figura 12) exibem um padrão semelhante ao observado nos picos gerais, mas com amplitudes significativamente maiores para eventos positivos e negativos. Antes do evento ($t=0$), os CAR permanecem relativamente estáveis. No entanto, em $t=0$, ocorre um impacto, marcado por um aumento acentuado nos retornos positivos e uma queda pronunciada nos retornos negativos, acompanhados por pequenas variações que mantêm a significância estatística ao longo do período analisado.

Figura 12 - Retornos Anormais Médios Acumulados para picos e anúncio de resultados

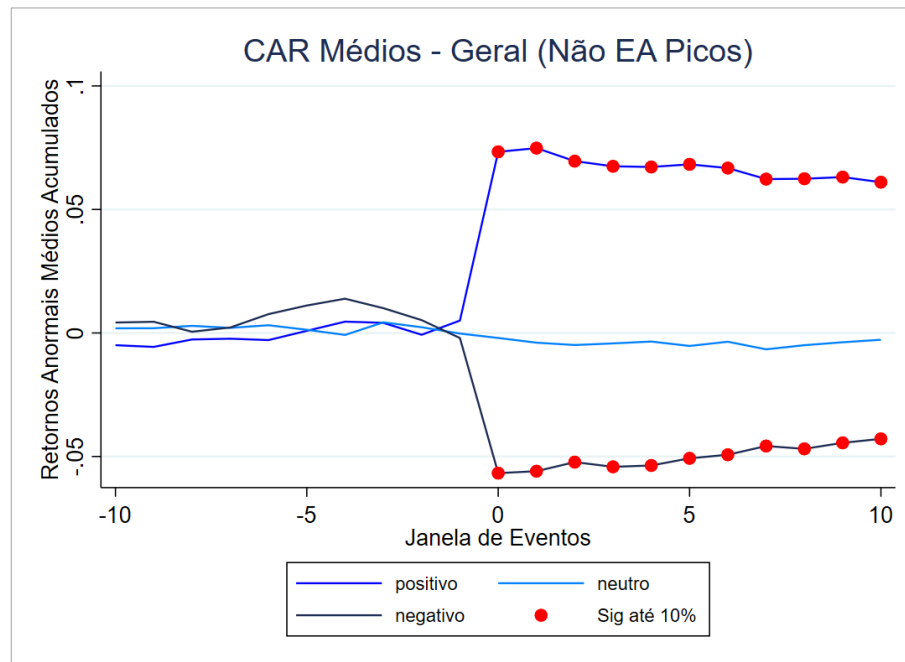


Fonte: Elaborado pela pesquisa (2024).

Adicionalmente, a amplitude dos CAR médios em dias de coincidência entre picos de atividade no Twitter e a divulgação de resultados supera tanto os picos de atividade isolados (Figura 11) quanto a análise da série temporal completa (Figura 10). Esse comportamento indica que a interação na plataforma em dias de divulgação de resultados intensifica a disseminação de informações, aumentando a atenção dos investidores e amplificando as reações nos preços das ações.

A Figura 13 apresenta os CAR médio para os eventos não relacionados a anúncios de resultados (não-EA). Mesmo após remover os anúncios de resultados, cujo impacto nos retornos de preço já é conhecido, ainda é possível rejeitar a hipótese nula. Nesse caso, o CAR médio dos eventos não relacionados a EA aumenta anormalmente após os picos positivos detectados e diminui após os picos negativos. Nota-se um impacto maior nos retornos positivos do que nos negativos, tendo em vista a amplitude que os CAR chegaram. Os CAR são menos do que quando coincidem com os EA, mas, em ambos os casos o impacto é estatisticamente significativo até o final da janela.

Figura 13 - Retornos Anormais Médios Acumulados para picos não-EA



Fonte: Elaborado pela pesquisa (2024).

Com base nos cenários analisados, os resultados indicam que os *tweets* contêm informações que influenciam os retornos das ações, alinhando-se à literatura sobre a capacidade informativa da divulgação de resultado (Ball & Brown, 1968; Campbell et al., 1997) e dos *tweets* (Bartov et al., 2017; Bollen et al., 2011; Broadstock & Zhang, 2019; Jena & Majhi, 2023; Y. Mao et al., 2012; Ranco et al., 2015; Rao & Srivastava, 2012; Wei et al., 2016). Nota-se que picos de sentimentos positivos elevam os CAR, enquanto picos negativos resultam em quedas, em ambos os casos, observa-se significância estatística. Esse efeito é mais forte em eventos de picos coincidentes com divulgações de resultados, em que há a maior amplitude nos CAR positivos, indicando uma reação mais intensa do mercado brasileiro a informações favoráveis. A tendência de estabilização dos CAR após $t=0$ sugere que o mercado brasileiro processa rapidamente as informações, alinhando-se à hipótese de mercado eficiente em sua forma semiforte.

Os resultados divergem de evidências de que o mercado, em especial o brasileiro, enfatiza mais as notícias negativas (Galdi & Gonçalves, 2018; Silva, 2018; Souza & Martins, 2022). Almeida et al. (2022) reforçam essa visão, mostrando que variações negativas de sentimento têm maior impacto no comportamento do mercado e intensificam movimentos pessimistas. Além disso, Ranco et al. (2015) identificaram que CARs negativos são mais acentuados que os positivos, enquanto Farkas e Keshk (2019) observaram que reações a notícias negativas são mais fortes em relação a positivas, diferente do observado nesta pesquisa. Por

fim, Souza e Martins (2022) sugerem que o volume total de *tweets* não tem associação significativa com retornos, contrastando com os resultados desta pesquisa, que indicam que, em dias de pico de atividade, os retornos anormais acumulados são mais acentuados.

4.4 CORRELAÇÃO DE PEARSON E CAUSALIDADE DE GRANGER

Este tópico apresenta os resultados da análise de correlação de Pearson e do teste de causalidade de Granger. Foram testadas as relações entre: a) o retorno diário das ações (Ret) e o sentimento diário dos *tweets* (SD), b) o retorno das ações e o volume diário de *tweets* (VT), c) o volume de negociações (VN) e o sentimento diário dos *tweets*, e d) o volume de negociações e o volume diário de *tweets*. A correlação de Pearson foi aplicada para medir a associação linear entre essas variáveis, enquanto o teste de Granger verificou se SD e VT são capazes de prever os retornos e os volumes de negociações, sugerindo uma possível relação temporal. A Tabela 12 apresenta os resultados das correlações calculadas.

A correlação entre o sentimento diário (SD) e o retorno das ações (Ret_i) foi geralmente positiva, indicando que o aumento no sentimento captado no Twitter está relacionado a retornos mais altos. No entanto, essa relação foi, em grande parte, fraca ou insignificante, com algumas exceções como Bradesco, Cemig, Natura, Petrobras e Siderúrgica Nacional, que apresentaram correlações significativas, mas de baixa magnitude. Quanto ao volume de *tweets* (VT) e o retorno diário, foram correlações predominantemente positiva, sugerindo que maior atividade no Twitter está associada a retornos mais altos. Algumas empresas, como Ambev, Cosan e Petrobras, apresentaram correlações negativas, sugerindo quedas nos retornos em resposta ao aumento de *tweets*.

A relação entre o sentimento diário e o volume de negócios é significativa para a maioria das empresas, indicando que variações de otimismo ou pessimismo podem impactar a atividade de compra e venda de ações. Por exemplo, empresas como Bradesco (0,1659) e Copel (0,1295) apresentam correlações positivas significativas, indicando que o sentimento pode influenciar o volume de negociações. No entanto, empresas como Petrobras (-0,2584) e Ultrapar (-0,1354) mostram correlações negativas, sugerindo que sentimentos mais negativos podem estar associados a volumes de negócios mais altos, possivelmente devido a maior volatilidade.

A relação entre o volume de *tweets* e o volume de negociações é amplamente positiva, com a maioria das empresas apresentando correlações estatisticamente significativas. Esses resultados sugerem que o Twitter pode atuar como uma fonte relevante de informações para o mercado, indicando que a atividade em redes sociais está associada a um aumento na

negociação de ações. Em outras palavras, à medida que o volume de *tweets* sobre uma empresa cresce, o volume de negócios também tende a aumentar. Empresas como Bradesco (0,7484), Cosan (0,7097), Natura (0,7014) e Azul (0,5679) se destacam com correlações muito elevadas e significativas, evidenciando uma forte ligação entre a discussão no Twitter e a movimentação nas negociações, possivelmente impulsionado pela disseminação de informações que estimulam a atividade no mercado.

Tabela 12 - Resultados dos testes de Correlação de Pearson

Empresa	Correlação de Pearson			
	$\rho(SD_i, Ret_i)$	$\rho(VT_i, Ret_i)$	$\rho(SD_i, VN_i)$	$\rho(VT_i, VN_i)$
Ambev	0,0438	-0,1159**	0,0195	0,2272***
Assaí	0,0288	0,0970**	-0,0388	0,3532***
Azul	-0,0047	0,3128***	-0,0356	0,5679***
Bradesco	-0,1039**	-0,0209	0,1659***	0,7484*
Braskem	0,0509	-0,0495	0,0895*	0,2248***
BRF	0,0757	0,0047	0,0747	0,3881***
Cemig	0,0903**	-0,0875*	0,0871*	0,1203***
Copel	0,1053**	0,1026**	0,1295***	0,2314***
Cosan	-0,0537	-0,2088***	-0,0733	0,7097***
Eletrobras	0,0023	-0,0989*	-0,1140**	0,2288***
Embraer	-0,0234	0,0134	0,1210*	0,1800***
Gerdaul	0,0416	0,0378	0,1008**	0,2663***
Gol	0,0703	0,1573***	0,0741	0,4438***
Itaú	-0,0182	-0,0151	0,1212***	0,0744
Natura	0,1313**	0,1018*	-0,0327	0,7014***
Pão de Açúcar	0,0396	0,0008	0,1436**	0,4615***
Petrobras	0,0847*	-0,1204***	-0,2584***	0,3941***
Sabesp	0,1054**	0,0043	0,2721***	0,0610
Santander	0,0075	0,0699	0,0277	0,0535
Sid Nacional	-0,1229**	0,1260**	0,0004	0,4144***
Suzano	0,0772	0,1481***	0,1275**	0,4394***
Telefônica Brasil	0,0670	0,1485**	0,0230	0,2469***
Ultrapar	-0,0553	0,1783***	-0,1354**	0,4027***
Vale	0,0566	0,0426	-0,0850*	0,4599***

Nota. SD = Sentimento Diário; VT = Volume de Tweets; Ret = Retorno da ação; VN = Volume de Negociações; *** Significativo ao nível de 1% ($p < 0,01$); ** Significativo ao nível de 5% ($p < 0,05$); e, * Significativo ao nível de 10% ($p < 0,10$).

De modo geral, as correlações de Pearson apresentaram coeficientes pequenos, mas foram estatisticamente significativas para algumas empresas. A Petrobras Cemig e Copel se destacou, exibindo correlações significativas em todas as relações testadas, sugerindo uma conexão consistente entre sentimento, volume de tweets e comportamento de mercado. A baixa magnitude dos coeficientes se alinha com Mao et al. (2011) e Ranco et al. (2015), que sugerem que dias com menor atividade de *tweets* podem reduzir a força das correlações.

As correlações entre o sentimento diário e o retorno das ações foram, em sua maioria, fracas, indicando que o sentimento captado no Twitter não é um preditor robusto para os retornos. Esses resultados corroboram os achados de Rao e Srivastava (2012), Ranco et al.

(2015), Mao et al. (2012) e Wei et al. (2016) de que o impacto do Twitter sobre os retornos é geralmente limitado e depende de eventos específicos que geram maior atenção. Também é corroborado por Rautiainen e Jokinen (2022) que concluíram que o sentimento geral não afeta significativamente os preços das ações, mas pode ter impacto em casos específicos.

Em contraste, esses resultados divergem do estudo de Bollen et al. (2011), que identificou um efeito mais consistente do humor público sobre os retornos, do DJIA. Além disso, estudos como os de Bartov et al. (2017) e Scaramozzino et al. (2021) destacam que, em certos contextos, há uma relação mais pronunciada entre sentimentos e retornos, incluindo evidências de fluxos causais significativos entre sentimentos e valores de mercado.

Por outro lado, as correlações significativas entre o volume de *tweets* e o volume de negociações sugerem que a atenção gerada nas redes sociais pode influenciar a movimentação do mercado, ainda que essa influência varie conforme o contexto de cada empresa e o nível de atividade *online*. Esses achados também são corroborados por Cocilova (2015), que observou correlação positiva entre o volume diário de *tweets* e negociações. Isso está em linha com estudos como Ganesh e Iyer (2023), Ranco et al. (2015), e Rao e Srivastava (2012) que concluíram que a atenção gerada no Twitter é um preditor relevante de atividade de mercado.

Os resultados dos testes de causalidade de Granger, apresentados na Tabela 13, mostram as direções de causalidade entre as variáveis do Twitter e as de mercado com até quatro dias de defasagem. A defasagem de um dia (D1) indica efeitos imediatos, com impacto no dia seguinte, sugerindo reação rápida dos investidores às discussões no Twitter. Defasagens maiores (D2, D3 e D4) indicam que o impacto demora mais para ocorrer, sugerindo que os investidores levam mais tempo para processar as informações ou que o efeito dos eventos no Twitter demora para atingir o mercado.

Em relação ao teste de causalidade de Granger, foram identificadas 96 relações estatisticamente significativas, representando 25% das 384 combinações avaliadas até o nível de significância de 10%. Dentre essas, 47% indicam que as variáveis de mercado granger-causam as variáveis do Twitter, 36% apontam o oposto, e 17% sugerem uma relação bicausal. Poucas empresas apresentaram relações significativas entre o retorno diário e as variáveis do Twitter, o que está em linha com os achados de Ranco et al. (2015), Bollen et al. (2011) e Ganesh e Iyer (2023), que destacam o impacto limitado das variáveis do Twitter nos retornos das ações em testes de causalidade de Granger.

Tabela 13 - Resultados dos testes de Causalidade de Granger

Empresa	Causalidade de Granger															
	Ret & SD				Ret & VT				VN & SD				VN & VT			
	D1	D2	D3	D4	D1	D2	D3	D4	D1	D2	D3	D4	D1	D2	D3	D4
Ambev													→		→	
Assaí									→			→	→	→	→	→
Azul			←	←			→	←					←		→	
Bradesco										←			←	←	←	←
Braskem	↔			←	→				←							
BRF		→		←		→								←		
Cemig												→	→		↔	←
Copel						→	→						→	→	→	←
Cosan							→							↔	→	
Eletrobras									→	→			←	→	↔	
Embraer					←			←								
Gerdau						↔	←									
Gol			→												→	↔
Itaú			→										→			
Natura	↔				←	→										←
Pão de Açúcar			←			→			→							↔
Petrobras			←					←	↔	↔	↔	→	←	→	→	→
Sabesp																
Santander							→		←	↔	↔	↔		→	↔	
Sid Nacional		→			←								→	←	←	
Suzano		→					←									
Telefônica Brasil							←	←	→	←						←
Ultrapar	←	↔												→	→	
Vale											→					
Total	3	4	5	3	3	6	7	4	7	5	3	4	8	12	13	9

Nota. SD = Sentimento Diário; VT = Volume de *Tweets*; Ret = Retorno da ação; VN = Volume de Negociações; Setas indicam causalidade de Granger com até 10% de significância. Para a esquerda, a variável do Twitter granger-causa a de mercado; para a direita, a de mercado granger-causa a do Twitter; setas bidirecionais indicam causalidade mútua. D1, D2, D3 e D4 representam defasagens de 1 a 4 dias, respectivamente.

As relações entre o volume de negociações e o sentimento diário também não foram amplamente significativas, diferentemente de estudos como Mao et al. (2011) e Rao e Srivastava (2012), que encontraram uma forte relação. No entanto, é importante destacar que empresas como Petrobras e Santander apresentaram relações bidirecionais entre Sentimento e Volume de Negociações em quase todas as defasagens testadas, indicando que, para essas empresas, o sentimento diário influencia o volume de negociações e vice-versa.

O volume de *tweets* mostrou pouco impacto nos retornos, mas evidenciou uma influência mais consistente sobre o volume de negociações. Foi identificado que há causalidade, no sentido de Granger, em aproximadamente 17% das empresas na defasagem D1 e 21% na defasagem D4. Esses resultados sugerem que o volume de *tweets* é um preditor mais eficaz para mudanças no volume de negociações do que nos retornos das ações, corroborando estudos como Ganesh e Iyer (2023) e Ranco et al. (2015). Assim, os resultados destacam que a atenção gerada pelo Twitter tem maior impacto na movimentação do mercado (via volume de negociações) do

que diretamente nos preços das ações, e que esse impacto pode variar substancialmente entre as empresas analisadas.

5 CONCLUSÃO

Compreende-se que as redes sociais aprimoram a comunicação entre organizações e *stakeholders*, possibilitando interação em tempo real, redução de custos e vantagens competitivas em relação a outros meios de divulgação. Assim, esta pesquisa teve como objetivo principal investigar a relação entre o volume e o sentimento presente nos *tweets* com o retorno, o volume negociado e a volatilidade das ações de 24 empresas brasileiras que possuem ADRs listados na NYSE e integram os principais índices IBOV, IBRA e IBRX100, no período de 2022 a 2023, por meio de métodos como estudo de eventos, correlação de Pearson e análise de Causalidade de Granger.

Os resultados desta pesquisa indicam que o sentimento diário e o volume de *tweets* apresentam correlações com o volume de negociações para a maioria das empresas analisadas. Isso reforça a ideia de que o Twitter atua como um catalisador, dada a sua capacidade de disseminar informações rapidamente para uma ampla base de investidores. No entanto, a correlação entre o sentimento do Twitter e os retornos das ações mostrou-se menos representativas, sugerindo que a plataforma influencia mais as decisões de investimento do que mudanças nos preços. O aumento no volume de *tweets* pode sinalizar eventos relevantes, incentivando investidores a investigar, proteger suas posições ou explorar oportunidades especulativas. Essa atividade pode gerar comportamento de manada e atrair novos investidores, ampliando o volume de negociações, mesmo sem fundamentos sólidos.

Outro ponto identificado na pesquisa é que, após um choque inicial causado por eventos específicos como os picos de atividade no Twitter, os retornos anormais acumulados se estabilizam rapidamente, ou seja, entram em um platô, sugerindo que o mercado brasileiro incorpora novas informações de maneira ágil. Esse comportamento está alinhado com a hipótese de mercado eficiente em sua forma semiforte. Embora existam ruídos provenientes de *tweets*, o mercado é capaz de corrigir essas disparidades de forma quase imediata, mostrando eficiência na absorção de informações. Esses resultados divergem do caso GameStop, na qual comportamentos organizados em massa permitiram ganhos anormais elevados, indicando uma ineficiência do mercado.

As evidências mostram que, em dias de picos no Twitter, especialmente durante anúncios de resultados o mercado brasileiro mostra maior engajamento e uma percepção mais

otimista, sugerindo maior sensibilidade a informações positivas em termos de amplitude e duração do impacto. Esse padrão, alinhado com estudos como Boylan e Boylan (2017), Ranco et al. (2015) e Teti et al. (2019), destaca o papel das redes sociais em amplificar impactos positivos, especialmente em empresas com alta presença digital.

Os resultados reforçam a relevância do Twitter como um canal de amplificação de informações, mas também evidenciam limitações em sua capacidade de influenciar os retornos das ações. Com base nessas observações, investidores podem ajustar suas estratégias ao considerar o volume de atividade nas redes sociais como um indicador de atenção do mercado. Por sua vez, aprimorar suas práticas de comunicação para maximizar benefícios e fortalecer sua imagem perante o mercado, enquanto reguladores podem usar essas evidências para aprimorar diretrizes sobre o uso corporativo das redes sociais, promovendo maior transparência e equilíbrio no ambiente financeiro.

O estudo atingiu seu objetivo e respondeu à pergunta de pesquisa ao demonstrar que o Twitter influencia determinados aspectos do mercado acionário brasileiro, embora de forma não uniforme, evidenciando uma relação multifacetada entre seu conteúdo e o desempenho das ações. As hipóteses foram parcialmente confirmadas: o volume de *tweets* apresentou uma relação significativa com o volume de negociações, enquanto o impacto do sentimento mostrou-se limitado. Além disso, a hipótese de eficiência de mercado foi corroborada, indicando que os preços das ações refletem, em grande parte, as informações disponíveis.

No campo teórico, este trabalho contribuiu para preencher a lacuna na literatura ao investigar a interação entre redes sociais e o mercado acionário em um contexto de mercado emergente, como o brasileiro, ainda pouco explorado. Ademais, ao utilizar dados do Twitter para avaliar o sentimento dos investidores, a pesquisa adota uma abordagem alternativa às metodologias tradicionais, como pesquisas de opinião e variáveis de mercado, com potencial para replicações e comparações em outros contextos e mercados.

Algumas limitações devem ser destacadas. Os resultados são restritos ao contexto brasileiro devido ao recorte amostral e temporal. As restrições do Twitter na coleta de dados podem ter reduzido o volume disponível, especialmente para empresas com baixo engajamento. A exclusão de dias não úteis também limitou o conjunto de *tweets* analisados, o que pode ter limitado a identificação de interações relevantes para o mercado. Além disso, o uso de figuras de linguagem, como sarcasmo e gírias, e outros elementos informais como abreviações e emojis, pode ter comprometido a precisão na classificação de sentimentos, embora este não fosse um foco do estudo.

Pesquisas futuras podem desenvolver dicionários personalizados para redes sociais utilizando inteligência artificial e aprendizado de máquina, incluir empresas com maior engajamento e expandir a análise para plataformas como Instagram, LinkedIn, Facebook, TikTok e YouTube. Além disso, investigar o papel das redes sociais no gerenciamento de resultados corporativos pode contribuir para o aprimoramento das práticas financeiras.

REFERÊNCIAS

- Abreu, D. P. A. de, Camargos, M. A. de, & Bressan, A. A. (2023). The (In)Efficiency of Emerging and Developed Markets: An Analysis from Fractal Theory. *BAR - Brazilian Administration Review*, 20(1).
<https://www.redalyc.org/journal/841/84175119006/html/>
- Agarwal, S., Kumar, S., & Goel, U. (2019). Stock market response to information diffusion through internet sources: A literature review. *International Journal of Information Management*, 45, 118–131. <https://doi.org/10.1016/j.ijinfomgt.2018.11.002>
- Aggarwal, D. (2022). Defining and measuring market sentiments: A review of the literature. *Qualitative Research in Financial Markets*, 14(2), 270–288.
<https://doi.org/10.1108/QRFM-03-2018-0033>
- Akerlof, G. A. (1970). The Market for “Lemons”: Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, 84(3), 488.
<https://doi.org/10.2307/1879431>
- Alajbeg, D., Bubas, Z., & Sonje, V. (2012). The efficient market hypothesis: Problems with interpretations of empirical tests. *Financial Theory and Practice*, 36(1), 53–72.
<https://doi.org/10.3326/fintp.36.1.3>
- Além do X: Veja 10 mudanças no Twitter sob o comando de Elon Musk.* (2023, julho 25). G1.
<https://g1.globo.com/tecnologia/noticia/2023/07/25/alem-do-x-veja-10-mudancas-no-twitter-sob-o-comando-de-elon-musk.ghtml>
- Almeida, M. D., Maia, V. M., & Tommasetti, R. (2022). A Influência da Mídia Social no Valor da Petrobras. *Revista Mineira de Contabilidade*, 23(3), 63–75.
<https://doi.org/10.51320/rmc.v23i3.1371>
- Alves, J. B. P. (2022). *A influência dos tweets de Elon Musk no retorno das ações da Tesla: Um estudo de eventos* [Dissertação de Mestrado, Instituto Superior de Contabilidade e Administração do Porto]. <https://recipp.ipp.pt/handle/10400.22/22350>
- Amin, M. H., Mohamed, E. K. A., & Elragal, A. (2021). CSR disclosure on Twitter: Evidence from the UK. *International Journal of Accounting Information Systems*, 40.
<https://doi.org/10.1016/j.accinf.2021.100500>
- Araújo, V. L., & Fernandes, B. J. T. (2021). Análise de Sentimentos de Textos do Twitter sobre Mercado de Ações Brasileiro. *Revista de Engenharia e Pesquisa Aplicada*, 6(5), 18–26. <http://revistas.poli.br/index.php/rep/article/view/1750>

- Arruda, M. P. de, Girão, L. F. de A. P., & Lucena, W. G. L. (2015). Assimetria Informacional e o Preço das Ações: Análise da Utilização das Redes Sociais nos Mercados de Capitais Brasileiro e Norte-americano. *Revista Contabilidade & Finanças*, 26, 317–330. <https://doi.org/10.1590/1808-057x201501540>
- Assaf Neto, A. (2021). *Mercado Financeiro* (15º ed). Grupo GEN. <https://app.minhabiblioteca.com.br/books/9788597028171>
- Asur, S., & Huberman, B. A. (2010). Predicting the Future with Social Media. 2010 *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, 1, 492–499. <https://doi.org/10.1109/WI-IAT.2010.63>
- Atilgan, Y., Demirtas, K. O., & Simsek, K. D. (2015). Studies of Equity Returns in Emerging Markets: A Literature Review. *Emerging Markets Finance & Trade*, 51(4), 757–773. <https://www.jstor.org/stable/26752599>
- B3, B., Bolsa, Balcão. (2022). *Pessoas Físicas—Uma análise da evolução dos investidores na B3 – Setembro de 2022*. B3. <https://api.mziq.com/mzfilemanager/v2/d/5fd7b7d8-54a1-472d-8426-eb896ad8a3c4/e592e4a4-1ad2-61b5-9212-b958a9840577?origin=1>
- B3, B., Bolsa, Balcão. (2023a). *Pessoas Físicas—Uma análise da evolução dos investidores na B3 – Setembro de 2023*. B3. https://www.b3.com.br/data/files/01/61/F9/67/9DE6C8103152D4C8AC094EA8/Book%20Pessoa%20Fisica%20-%203TRI%202023%20_final_.pdf
- B3, B., Bolsa, Balcão. (2023b, novembro 16). *Índices amplos | B3*. https://www.b3.com.br/pt_br/market-data-e-indices/indices/indices-amplos/
- Bachelier, L. (1900). Théorie de la spéculation. *Annales scientifiques de l'École normale supérieure*, 17, 21–86. <https://doi.org/10.24033/asens.476>
- Ball, R. (2009). The Global Financial Crisis and the Efficient Market Hypothesis: What Have We Learned? *Journal of Applied Corporate Finance*, 21(4), 8–16. <https://doi.org/10.1111/j.1745-6622.2009.00246.x>
- Ball, R., & Brown, P. (1968). An Empirical Evaluation of Accounting Income Numbers. *Journal of Accounting Research*, 6(2), 159–178. <https://doi.org/10.2307/2490232>
- Barber, B. M., & Odean, T. (2001). The Internet and the Investor. *Journal of Economic Perspectives*, 15(1), 41–54. <https://doi.org/10.1257/jep.15.1.41>
- Barbopoulos, L. G., Dai, R., Putniņš, T., & Saunders, A. (2021). *Market Efficiency in the Age of Machine Learning*. <https://www.research.ed.ac.uk/en/publications/market-efficiency-in-the-age-of-machine-learning>
- Bartov, E., Faurel, L., & Mohanram, P. S. (2017). Can Twitter Help Predict Firm-Level Earnings and Stock Returns? *Rotman School of Management Working*. <https://doi.org/10.2139/ssrn.2782236>
- Belfiore, P. (2015). *Estatística—Aplicada à Administração, Contabilidade e Economia com Excel e SPSS*. Grupo GEN. [https://app.minhabiblioteca.com.br/reader/books/9788595155596/epubcfi/6/50\[%3Bvnd.vst.idref%3Dbody024\]/4/34/1:12\[%C3%ADtu%2Clo%20\]](https://app.minhabiblioteca.com.br/reader/books/9788595155596/epubcfi/6/50[%3Bvnd.vst.idref%3Dbody024]/4/34/1:12[%C3%ADtu%2Clo%20])

- Belo, N. M., & Brasil, H. G. (2006). Assimetria informacional e eficiência semiforte do mercado. *Revista de Administração de Empresas*, 46, 48–57. <https://doi.org/10.1590/S0034-75902006000500004>
- Benninga, S. (2014). *Financial modeling* (4^o ed). The MIT Press.
- Bernard, V. L., & Thomas, J. K. (1990). Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of Accounting and Economics*, 13(4), 305–340. [https://doi.org/10.1016/0165-4101\(90\)90008-R](https://doi.org/10.1016/0165-4101(90)90008-R)
- Blankespoor, E., Miller, G. S., & White, H. D. (2014). The Role of Dissemination in Market Liquidity: Evidence from Firms' Use of Twitter. *The Accounting Review*, 89(1), 79–112. <https://www.jstor.org/stable/24468513>
- Bodie, Z., Kane, A., & Marcus, A. J. (2015). *Investimentos*. AMGH. <https://app.minhabiblioteca.com.br/reader/books/9788580554205/pageid/0>
- Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1), 1–8. <https://doi.org/10.1016/j.jocs.2010.12.007>
- Boylan, D. H., & Boylan, C. L. (2017). Technology in Accounting: Social Media as Effective Platform for Financial Disclosures. *The International Journal of Digital Accounting Research*, 17, 93–109. https://doi.org/10.4192/1577-8517-v17_4
- Broadstock, D. C., & Zhang, D. (2019). Social-media and intraday stock returns: The pricing power of sentiment. *Finance Research Letters*, 30, 116–123. <https://doi.org/10.1016/j.frl.2019.03.030>
- Brown, G. W., & Cliff, M. T. (2004). Investor sentiment and the near-term stock market. *Journal of Empirical Finance*, 11(1), 1–27. <https://doi.org/10.1016/j.jempfin.2002.12.001>
- Camargos, M. de, & Barbosa, F. V. (2010). Teoria e evidência da eficiência informacional do mercado de capitais brasileiro. *Caderno de Pesquisas em Administração*, 10(1), 43–55.
- Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*. <https://press.princeton.edu/books/hardcover/9780691043012/the-econometrics-of-financial-markets>
- Carvalho, R. S. de, Tonin, J. M. da F., & Sanches, S. L. R. (2023). Previsibilidade do Retorno das Ações pela Análise do Sentimento Textual: Uma revisão. *Revista de Educação e Pesquisa em Contabilidade (REPeC)*, 17(2). <https://doi.org/10.17524/repec.v17i2.3111>
- Chen, G., Qiu, T., Jiang, X.-F., Zhong, L.-X., & Wu, X.-R. (2015). How trading volume responds to return in financial dynamics? *Physica A: Statistical Mechanics and its Applications*, 424, 73–81. <https://doi.org/10.1016/j.physa.2015.01.001>
- Chohan, U. (2021). Counter-Hegemonic Finance: The Gamestop Short Squeeze. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3775127>
- Cloudflare Radar. (2023, agosto 29). Domain Rankings. <https://radar.cloudflare.com/>
- Cocilova, A. (2015). *Twitter data analysis for financial markets* [Dissertação de graduação]. <https://amslaurea.unibo.it/8126/>

- Comissão de Valores Mobiliários. (2023). *Análise de Impacto Regulatório (AIR): Influenciadores Digitais e o Mercado de Capitais Brasileiro*. CVM. <https://www.gov.br/cvm/pt-br/centrais-de-conteudo/publicacoes/estudos/20230418-air-influenciadores.pdf/view>
- Cornell, B. (2018). What Is the Alternative Hypothesis to Market Efficiency? *The Journal of Portfolio Management*, 44(7), 3–6. <https://doi.org/10.3905/jpm.2018.44.7.003>
- Costa, T. A., & Famá, R. (2007). *A Turbulência das Finanças Modernas: Será o fim do comportamento aleatório e o surgimento do comportamento determinístico do Caos*. 10. <https://sistema.semead.com.br/10semead/sistema/resultado/trabalhosPDF/239.pdf>
- Creswell, J. W., & Creswell, J. D. (2021). *Projeto de Pesquisa: Métodos Qualitativo, Quantitativo e Misto* (5º ed). Penso. [https://app.minhabiblioteca.com.br/reader/books/9786581334192/epubcfi/6/28\[%3Bvnd.vst.idref%3Dchapter1.xhtml\]!/4/2/2/18/5:88\[ndi%2Cv%C3%ADd\]](https://app.minhabiblioteca.com.br/reader/books/9786581334192/epubcfi/6/28[%3Bvnd.vst.idref%3Dchapter1.xhtml]!/4/2/2/18/5:88[ndi%2Cv%C3%ADd])
- Cruz, A. F. da. (2016). Assimetria informacional no contexto das finanças corporativas: Determinantes e efeitos no contexto organizacional. *Revista de Economia do Centro-Oeste*, 2(1), Artigo 1. <https://doi.org/10.5216/reoeste.v2i1.42530>
- de Camargos, M. A., & Barbosa, F. V. (2015). *Eficiência informacional do mercado de capitais brasileiro em anúncios de fusões e aquisições*. 25(3).
- Dhankar, R. S., & Shankar, D. (2016). Relevance and evolution of adaptive markets hypothesis: A review. *Journal of Indian Business Research*, 8(3), 166–179. <https://doi.org/10.1108/JIBR-12-2015-0125>
- Di Toro, M. (2022). The Effect of the Asymmetric Information on the Asset's Price and the Strong Efficient Market Hypothesis. *International Journal of Research and Review*, 9(10), 463–470. <https://doi.org/10.52403/ijrr.20221052>
- Do, H. X., Brooks, R., Treepongkaruna, S., & Wu, E. (2014). How does trading volume affect financial return distributions? *International Review of Financial Analysis*, 35, 190–206. <https://doi.org/10.1016/j.irfa.2014.09.003>
- Dutta, S., Essaddam, N., Kumar, V., & Saadi, S. (2017). How does electronic trading affect efficiency of stock market and conditional volatility? Evidence from Toronto Stock Exchange. *Research in International Business and Finance*, 39, 867–877. <https://doi.org/10.1016/j.ribaf.2015.11.001>
- Falconi-Pires, L. M., & Lourenço, J. (2022). Twitter ontem e hoje: Observações metodológicas críticas. *Revista Heterotópica*, 4(Especial), 36–52. <https://doi.org/10.14393/HTP-v4nEspecial-2022-67202>
- Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2), 383–417. <https://doi.org/10.2307/2325486>
- Fama, E. F. (1991). Efficient Capital Markets: II. *The Journal of Finance*, 46(5), 1575–1617. <https://doi.org/10.2307/2328565>
- Famá, R., Cioffi, P. L. de M., & Coelho, P. A. R. (2008). Contexto das finanças comportamentais: Anomalias e eficiência do mercado de capitais brasileiro. *REGE Revista de Gestão*, 15(2), 65–78. <https://doi.org/10.5700/issn.2177-8736.rege.2008.36638>

- Fan, W., & Gordon, M. D. (2014). The power of social media analytics. *Communications of the ACM*, 57(6), 74–81. <https://doi.org/10.1145/2602574>
- Farkas, M., & Keshk, W. (2019). How Facebook influences non-professional investors' affective reactions and judgments: The effect of disclosure platform and news valance. *Journal of Financial Reporting and Accounting*, 17(1), 80–103. <https://doi.org/10.1108/JFRA-10-2017-0092>
- Fávero, L. P., Belfiore, P. P., Silva, F. L. da, & Chan, B. L. (2009). *Análise de dados: Modelagem multivariada para tomada de decisões*. Elsevier.
- Fávero, L. P., Belfiore, P., Takamatsu, R. T., & Suzart, J. (2014). *Métodos quantitativos com stata: Procedimentos, rotinas e análise de resultados*. Elsevier.
- Galdi, F. C., & Gonçalves, A. M. (2018). PESSIMISMO E INCERTEZA DAS NOTÍCIAS E O COMPORTAMENTO DOS INVESTIDORES NO BRASIL. *Revista de Administração de Empresas*, 58, 130–148. <https://doi.org/10.1590/S0034-759020180203>
- Ganesh, A., & Iyer, S. (2023). Impact of Firm-Initiated Tweets on Stock Return and Trading Volume. *Journal of Behavioral Finance*, 24(2), 171–182. <https://doi.org/10.1080/15427560.2021.1949717>
- Garman, M. B., & Klass, M. J. (1980). On the Estimation of Security Price Volatilities from Historical Data. *The Journal of Business*, 53(1), 67–78. <https://www.jstor.org/stable/2352358>
- Garton, L., Haythornthwaite, C., & Wellman, B. (1997). Studying Online Social Networks. *Journal of Computer-Mediated Communication*, 3(1). <https://doi.org/10.1111/j.1083-6101.1997.tb00062.x>
- Ge, Q., Kurov, A., & Wolfe, M. (2018). Do Investors Care about Presidential Company-Specific Tweets? <https://doi.org/10.2139/ssrn.2986350>
- Gil, A. C. (2019). *Métodos e Técnicas de Pesquisa Social* (7º ed). Atlas. <https://app.minhabiblioteca.com.br/books/9788597020991>
- Gitman, Lawrence J. (2010). *Princípios de Administração Financeira* (12º ed). Pearson Prentice Hall.
- Goel, A., Tripathi, V., & Agarwal, M. (2020). Information asymmetry and stock returns. *Journal of Advances in Management Research*, 18(1), 85–112. <https://doi.org/10.1108/JAMR-05-2020-0084>
- Google Cloud. (2023, setembro 21). *Princípios básicos da Natural Language API*. Cloud Natural Language. <https://cloud.google.com/natural-language/docs/basics?hl=pt-br>
- Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, 37(3), 424–438. <https://doi.org/10.2307/1912791>
- Gu, C., & Kurov, A. (2020). Informational role of social media: Evidence from Twitter sentiment. *Journal of Banking & Finance*, 121. <https://doi.org/10.1016/j.jbankfin.2020.105969>

- Guruganesh, G., Schneider, J., & Wang, J. R. (2021). Contracts under Moral Hazard and Adverse Selection. *Proceedings of the 22nd ACM Conference on Economics and Computation*, 563–582. <https://doi.org/10.1145/3465456.3467637>
- Hair Jr, J. F., Black, W. C., Babin, B. J., Anderson, R. E., Tatham, R. L., Gouvêa, M. A., & Sant’Anna, A. S. (2009). *Análise Multivariada de Dados* (6º ed). Bookman.
- Hirshleifer, D. A., Peng, L., & Wang, Q. (2021). News Diffusion in Social Networks and Stock Market Reactions. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3824022>
- Indurkha, N., & Damerau, F. J. (Orgs.). (2010). *Handbook of Natural Language Processing* (2nd edition). Chapman and Hall/CRC.
- InfoMoney. (2022, novembro 8). NYSE: O que é e como funciona a Bolsa de Nova York. *InfoMoney*. <https://www.infomoney.com.br/guias/o-que-e-nyse/>
- Ismail, A. M., Maia, V. M., Tommasetti, R., & Macedo, M. A. da S. (2022). Análise de Tweets ao Redor da Divulgação de Resultados. *Revista Evidenciação Contábil & Finanças*, 10(3), 118–136. <https://doi.org/10.22478/ufpb.2318-1001.2022v10n3.62911>
- Jena, P. R., & Majhi, R. (2023). Are Twitter sentiments during COVID-19 pandemic a critical determinant to predict stock market movements? A machine learning approach. *Scientific African*, 19. <https://doi.org/10.1016/j.sciaf.2022.e01480>
- Jensen, M. C. (1978). Some anomalous evidence regarding market efficiency. *Journal of Financial Economics*, 6(2–3), 95–101. [https://doi.org/10.1016/0304-405X\(78\)90025-9](https://doi.org/10.1016/0304-405X(78)90025-9)
- Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263–291. <https://doi.org/10.2307/1914185>
- Kao, Y.-S., Chuang, H.-L., & Ku, Y.-C. (2020). The empirical linkages among market returns, return volatility, and trading volume: Evidence from the S&P 500 VIX Futures. *The North American Journal of Economics and Finance*, 54. <https://doi.org/10.1016/j.najef.2018.10.019>
- Karabulut, Y. (2012). Can Facebook Predict Stock Market Activity? *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2017099>
- Keynes, J. M. (1923). Some aspects of commodity markets. *Manchester Guardian Commercial: European Reconstruction Series*, 13, 784–786.
- Khder, M. (2021). Web Scraping or Web Crawling: State of Art, Techniques, Approaches and Application. *International Journal of Advances in Soft Computing and its Applications*, 13(3), 145–168. <https://doi.org/10.15849/IJASCA.211128.11>
- Kimura, H. (2010). Aspectos comportamentais associados às reações do mercado de capitais. *RAE eletrônica*, 2. <https://doi.org/10.1590/S1676-56482003000100007>
- Klibanoff, P., Lamont, O., & Wizman, T. A. (1998). Investor Reaction to Salient News in Closed-End Country Funds. *The Journal of Finance*, 53(2), 673–699. <https://doi.org/10.1111/0022-1082.265570>
- Kolchyna, O., Souza, T. T. P., Treleaven, P., & Aste, T. (2015). Twitter Sentiment Analysis: Lexicon Method, Machine Learning Method and Their Combination. *arXiv.Org*. <https://arxiv.org/abs/1507.00955v3>

- Komalasari, P. (2016). INFORMATION ASYMMETRY AND HERDING BEHAVIOR. *Jurnal Akuntansi dan Keuangan Indonesia*, 13(1).
<https://doi.org/10.21002/jaki.2016.04>
- Koul, Y., Mamgain, K., & Gupta, A. (2022). Lifetime of tweets: A statistical analysis. *Social Network Analysis and Mining*, 12(1), 101. <https://doi.org/10.1007/s13278-022-00926-4>
- Latif, M., Arshad, S., Fatima, M., & Farooq, S. (2012). Market Efficiency, Market Anomalies, Causes, Evidences, and Some Behavioral Aspects of Market Anomalies. *Research Journal of Finance and Accounting*, 2(9–10), 1.
<https://www.iiste.org/Journals/index.php/RJFA/article/view/1287>
- LeRoy, S. F. (1989). Efficient Capital Markets and Martingales. *Journal of Economic Literature*, 27(4), 1583–1621. <https://www.jstor.org/stable/2727024>
- Lim, K.-P., & Brooks, R. (2011). The Evolution of Stock Market Efficiency Over Time: A Survey of the Empirical Literature. *Journal of Economic Surveys*, 25(1), 69–108.
<https://doi.org/10.1111/j.1467-6419.2009.00611.x>
- Lira, A. C. da S., Pachú, C. O., & Araújo, L. M. A. de. (2022). Consequências negativas da utilização e da comunicação em redes sociais: Uma revisão narrativa. *Cadernos da Escola de Comunicação*, 18(1), 30–44.
<https://portaldeperiodicos.unibrasil.com.br/index.php/cadernoscomunicacao/article/view/6357>
- Lo, A. W. (2000). Finance: A Selective Survey. *Journal of the American Statistical Association*, 95(450), 629–635. <https://doi.org/10.2307/2669406>
- Lo, A. W. (2004). *The Adaptive Markets Hypothesis: Market Efficiency from an Evolutionary Perspective*. <https://papers.ssrn.com/abstract=602222>
- Lo, A. W. (2005). *Reconciling Efficient Markets with Behavioral Finance: The Adaptive Markets Hypothesis*. <https://papers.ssrn.com/abstract=1702447>
- Long, S., Lucey, B. M., Yarovaya, L., & Xie, Y. (2022). “I Just Like the Stock”: The Role of Reddit Sentiment in the GameStop Share Rally (SSRN Scholarly Paper 3822315).
<https://doi.org/10.2139/ssrn.3822315>
- Lu, D. W. (2024). *Twitter Sentiments and Stock Prices: An Event Study on the Role of Influencers* [Dissertation, Georgia State University].
<https://doi.org/10.57709/36943015>
- Luciani, A., Liu, C., & Peterson, R. (2023). Media Sentiment Momentum. Em *Handbook of Alternative Data in Finance* (Vol. 1). Chapman and Hall/CRC.
- MacKinlay, A. C. (1997). Event Studies in Economics and Finance. *Journal of Economic Literature*, 35(1), 13–39. <https://www.jstor.org/stable/2729691>
- Magalhães-Timotio, J. G., & Barbosa, F. V. (2020). Anomalias no Mercado de Capitais: Um Estudo Bibliométrico. *Revista de Administração de Roraima - RARR*, 10.
<https://doi.org/10.18227/2237-8057rarr.v10i0.5823>
- Malkiel, B. G. (2003). The Efficient Market Hypothesis and Its Critics. *Journal of Economic Perspectives*, 17(1), 59–82. <https://doi.org/10.1257/089533003321164958>

- Mao, H., Counts, S., & Bollen, J. (2011). *Predicting Financial Markets: Comparing Survey, News, Twitter and Search Engine Data* (arXiv:1112.1051). arXiv. <http://arxiv.org/abs/1112.1051>
- Mao, Y., Wei, W., Wang, B., & Liu, B. (2012). Correlating S&P 500 stocks with Twitter data. *Proceedings of the First ACM International Workshop on Hot Topics on Interdisciplinary Social Networks Research*, 69–72. <https://doi.org/10.1145/2392622.2392634>
- Marconi, M. de A., & Lakatos, E. M. (2023). *Fundamentos de Metodologia Científica* (9º ed). Atlas. [https://app.minhabiblioteca.com.br/reader/books/9788597026580/epubcfi/6/38\[%3Bvnd.vst.idref%3Dhtml17\]!/4/68/4](https://app.minhabiblioteca.com.br/reader/books/9788597026580/epubcfi/6/38[%3Bvnd.vst.idref%3Dhtml17]!/4/68/4)
- Mendes, M. S., & Lopes Lucena, W. G. (2022). Youtube e o Retornos das Ações: Um Estudo de Eventos no Canal “O Primo Rico”. Em *RC&C: Revista de Contabilidade e Controladoria* (2022º-12º–24º ed, Vol. 14, Número 3). <https://revistas.ufpr.br/rcc/article/view/81700>
- Mirzaee Ghazani, M., & Khalili Araghi, M. (2014). Evaluation of the adaptive market hypothesis as an evolutionary perspective on market efficiency: Evidence from the Tehran stock exchange. *Research in International Business and Finance*, 32, 50–59. <https://doi.org/10.1016/j.ribaf.2014.03.002>
- Mukherjee, S., & Bhattacharyya, P. (2013). *Sentiment Analysis: A Literature Survey*. arXiv. <https://doi.org/10.48550/arXiv.1304.4520>
- Müller, K., Pan, Y., & Schwarz, C. (2023). Social Media and Stock Market Participation. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4557783>
- Mustafa, K. H., & Ahmed, A. M. (2020). A Critical Review of the Market Efficiency Concept. *Shirkah: Journal of Economics and Business*, 5(2), 271. <https://doi.org/10.22515/shirkah.v5i2.326>
- Naseer, M., & Bin Tariq, D. Y. (2016). *The Efficient Market Hypothesis: A Critical Review of the Literature* [SSRN Scholarly Paper]. <https://papers.ssrn.com/abstract=2714844>
- Nisar, T. M., & Yeung, M. (2018). Twitter as a tool for forecasting stock market movements: A short-window event study. *The Journal of Finance and Data Science*, 4(2), 101–119. <https://doi.org/10.1016/j.jfds.2017.11.002>
- Nti, I. K., Adekoya, A. F., & Weyori, B. A. (2020). Predicting Stock Market Price Movement Using Sentiment Analysis: Evidence From Ghana. *Applied Computer Systems*, 25(1), 33–42. <https://doi.org/10.2478/acss-2020-0004>
- Nyakurukwa, K., & Seetharam, Y. (2023). The evolution of studies on social media sentiment in the stock market: Insights from bibliometric analysis. *Scientific African*, 20. <https://doi.org/10.1016/j.sciaf.2023.e01596>
- Ofoeda, J., Boateng, R., & Effah, J. (2019). Application Programming Interface (API) Research: A Review of the Past to Inform the Future. *International Journal of Enterprise Information Systems (IJEIS)*, 15(3), 76–95. <https://doi.org/10.4018/IJEIS.2019070105>

- Oliveira, D. J. S., Bermejo, P. H. de S., Pereira, J. R., & Barbosa, D. A. (2019). A aplicação da técnica de análise de sentimento em mídias sociais como instrumento para as práticas da gestão social em nível governamental. *Revista de Administração Pública*, 53, 235–251. <https://doi.org/10.1590/0034-7612174204>
- Parkinson, M. (1980). The Extreme Value Method for Estimating the Variance of the Rate of Return. *The Journal of Business*, 53(1), 61–65. <https://www.jstor.org/stable/2352357>
- Pedersen, L. H. (2022). Game on: Social networks and markets. *Journal of Financial Economics*, 146(3), 1097–1119. <https://doi.org/10.1016/j.jfineco.2022.05.002>
- Pereira, E. J. de A. L., & Urpia, A. G. B. da C. (2011). Hipótese dos mercados eficientes vis-à-vis incerteza, convenção e especulação: Por uma mudança de paradigma nos mercados financeiros. *Pesquisa & Debate Revista do Programa de Pós-Graduação em Economia Política*, 22(1(39)). <https://revistas.pucsp.br/index.php/rpe/article/view/7393>
- Pinheiro, J. L. (2019). *Mercado de Capitais* (9º ed). Grupo GEN. <https://app.minhabiblioteca.com.br/books/9788597021752>
- Ramassa, P., & Fabio, C. D. (2016). *Social media for investor relations: A literature review and future directions*. 16, 117–135. https://doi.org/10.4192/1577-8517-v16_5
- Ranco, G., Aleksovski, D., Caldarelli, G., Grčar, M., & Mozetič, I. (2015). The Effects of Twitter Sentiment on Stock Price Returns. *PLOS ONE*, 10(9). <https://doi.org/10.1371/journal.pone.0138441>
- Rao, T., & Srivastava, S. (2012). Analyzing Stock Market Movements Using Twitter Sentiment Analysis. *Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2012)*, 119–123. <https://dl.acm.org/doi/pdf/10.5555/2456719.2456923>
- Rautiainen, A., & Jokinen, J. (2022). The value-relevance of social media activity of Finnish listed companies. *International Journal of Accounting & Information Management*, 30(2), 301–323. <https://doi.org/10.1108/IJAIM-04-2021-0076>
- Recuero, R. (2009). *Redes Sociais na Internet*. Sulina.
- Rogers, L. C. G., & Satchell, S. E. (1991). Estimating Variance From High, Low and Closing Prices. *The Annals of Applied Probability*, 1(4), 504–512. <https://www.jstor.org/stable/2959703>
- Rönkkö, M., Holmi, J., Niskanen, M., & Mättö, M. (2024). The adaptive markets hypothesis: Insights into small stock market efficiency. *Applied Economics*, 56(25), 3048–3062. <https://doi.org/10.1080/00036846.2024.2326039>
- Rossi, M. (2015). The efficient market hypothesis and calendar anomalies: A literature review. *International Journal of Managerial and Financial Accounting*, 7(3/4), 285. <https://doi.org/10.1504/IJMFA.2015.074905>
- Samuelson, P. (1965). Proof That Properly Anticipated Prices Fluctuate Randomly. *Industrial Management Review*, 6(2). <https://capital-gain.ru/wp-content/uploads/Proof-that-properly-anticipated-prices-fluctuate-randomly.pdf>

- Santos, M. A. dos, Fávero, L. P. L., Brugni, T. V., & Serra, R. G. (2024). Adaptive markets hypothesis and economic-institutional environment: A cross-country analysis. *REGE Revista de Gestão*, 31(2), Artigo 2. <https://doi.org/10.1108/REGE-06-2022-0095>
- Sarrang, J. B., Zuhdi, A., & Ariwibowo, A. B. (2024). Sentiment Analysis from Twitter Regarding E-SIM Services by Smartfren Using the Naïve Bayes Classifier Method. *Intelmatix*, 4(1), Artigo 1. <https://doi.org/10.25105/itm.v4i1.17620>
- Scaramozzino, R., Cerchiello, P., & Aste, T. (2021). Information Theoretic Causality Detection between Financial and Sentiment Data. *Entropy*, 23(5), 621. <https://doi.org/10.3390/e23050621>
- SEC. (2013). *SEC Says Social Media OK for Company Announcements if Investors Are Alerted* (Press Release 2013–51). Securities & Exchange Commission (SEC). <https://www.sec.gov/news/press-release/2013-2013-51htm>
- Sewell, M. (2011). *History of the efficient market hypothesis* [Research Note No. RN/11/ 04, UCL Department of Computer Science.]. http://www.cs.ucl.ac.uk/fileadmin/UCL-CS/research/Research_Notes/RN_11_04.pdf
- Shen, D., Liu, L., & Zhang, Y. (2018). Quantifying the cross-sectional relationship between online sentiment and the skewness of stock returns. *Physica A: Statistical Mechanics and Its Applications*, 490, 928–934. <https://doi.org/10.1016/j.physa.2017.08.036>
- Sheng, X., Brzeszczyński, J., & Ibrahim, B. M. (2017). International stock return co-movements and trading activity. *Finance Research Letters*, 23, 12–18. <https://doi.org/10.1016/j.frl.2017.06.006>
- Shiller, R. J. (2003). From Efficient Markets Theory to Behavioral Finance. *Journal of Economic Perspectives*, 17(1), 83–104. <https://doi.org/10.1257/089533003321164967>
- Siikanen, M., Baltakys, K., Kanninen, J., Vatrapi, R., Mukkamala, R., & Hussain, A. (2018). Facebook drives behavior of passive households in stock markets. *Finance Research Letters*, 27, 208–213. <https://doi.org/10.1016/j.frl.2018.03.020>
- Silva, M. D. D. O. P. D. (2018). *O efeito do sentimento das notícias sobre o comportamento dos preços no mercado acionário brasileiro* [Doctorate, Universidade de Brasília]. <https://doi.org/10.26512/2018.02.T.32178>
- Smith, S., & O'Hare, A. (2022). Comparing traditional news and social media with stock price movements; which comes first, the news or the price change? *Journal of Big Data*, 9(1), 47. <https://doi.org/10.1186/s40537-022-00591-6>
- Souza, D. M. S. de. (2020). *Efeito do sentimento do investidor manifesto via Twitter sobre os retornos e o volume negociado no mercado acionário brasileiro* [Dissertação, Universidade Federal da Paraíba]. <https://repositorio.ufpb.br>
- Souza, D. M. S. de, Lucena, W. G. L., & Queiroz, D. B. de. (2019). *O Efeito do Sentimento do Investidor Expresso via Twitter sobre o Comportamento do Mercado Acionário Brasileiro Durante o Período Eleitoral*. XIX USP Internacional Conference in Accounting, São Paulo. <https://congress USP.fipecafi.org/anais/19UspInternational/ArtigosDownload/1621.pdf>

- Souza, D. M. S. de, & Martins, O. S. (2022). Brazilian stock market performance and investor sentiment on Twitter. *Revista de Gestão*, 31(1), 18–33. <https://doi.org/10.1108/REG-07-2021-0145>
- Sprenger, T. O., Sandner, P. G., Tumasjan, A., & Welpe, I. M. (2014). News or Noise? Using Twitter to Identify and Understand Company-specific News Flow. *Journal of Business Finance & Accounting*, 41(7–8), 791–830. <https://doi.org/10.1111/jbfa.12086>
- Stern, D. I. (2011). *From Correlation to Granger Causality* (SSRN Scholarly Paper 1959624). Social Science Research Network. <https://doi.org/10.2139/ssrn.1959624>
- Teti, E., Dallochio, M., & Aniasi, A. (2019). The relationship between twitter and stock prices. Evidence from the US technology industry. *Technological Forecasting and Social Change*, 149. <https://doi.org/10.1016/j.techfore.2019.119747>
- Țițan, A. G. (2015). The Efficient Market Hypothesis: Review of Specialized Literature and Empirical Research. *Procedia Economics and Finance*, 32, 442–449. [https://doi.org/10.1016/S2212-5671\(15\)01416-1](https://doi.org/10.1016/S2212-5671(15)01416-1)
- Tommasetti, R., Maia, V. M., & Macedo, M. Á. da S. (2021). Twitter as ‘bully pulpit’: Brazilian banking sector case. *Applied Economics Letters*, 28(1), 61–64. <https://doi.org/10.1080/13504851.2020.1730753>
- Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *Science*, 185(4157), 1124–1131. <https://doi.org/10.1126/science.185.4157.1124>
- Umar, Z., Gubareva, M., Yousaf, I., & Ali, S. (2021). A tale of company fundamentals vs sentiment driven pricing: The case of GameStop. *Journal of Behavioral and Experimental Finance*, 30, 100501. <https://doi.org/10.1016/j.jbef.2021.100501>
- Valcanover, V. M., Sonza, I. B., & Da Silva, W. V. (2020). Behavioral Finance Experiments: A Recent Systematic Literature Review. *SAGE Open*, 10(4), 215824402096967. <https://doi.org/10.1177/2158244020969672>
- Vieira, M. A. H. (2021). *Hipótese eficiência de mercado: O caso da Gamestop* [Monografia, Universidade de Brasília]. <https://bdm.unb.br/handle/10483/29383>
- Wei, W., Mao, Y., & Wang, B. (2016). Twitter volume spikes and stock options pricing. *Computer Communications*, 73, 271–281. <https://doi.org/10.1016/j.comcom.2015.06.018>
- Woo, K.-Y., Mai, C., McAleer, M., & Wong, W.-K. (2020). Review on Efficiency and Anomalies in Stock Markets. *Economies*, 8(1), Artigo 1. <https://doi.org/10.3390/economies8010020>
- Xiong, X., Meng, Y., Li, X., & Shen, D. (2019). An empirical analysis of the Adaptive Market Hypothesis with calendar effects: Evidence from China. *Finance Research Letters*, 31. <https://doi.org/10.1016/j.frl.2018.11.020>
- Yang, D., & Zhang, Q. (2000). Drift-Independent Volatility Estimation Based on High, Low, Open, and Close Prices. *The Journal of Business*, 73(3), 477–492. <https://doi.org/10.1086/209650>

- Yu, Y., Duan, W., & Cao, Q. (2013). The impact of social and conventional media on firm equity value: A sentiment analysis approach. *Decision Support Systems*, 55(4), 919–926. <https://doi.org/10.1016/j.dss.2012.12.028>
- Zenha, L. (2018). Redes sociais online: O que são as redes sociais e como se organizam? *Caderno de Educação*, 49, 19–42. <https://revista.uemg.br/index.php/cadernodeeducacao/article/view/2809>
- Zhang, W., & Skiena, S. (2010). Trading Strategies to Exploit Blog and News Sentiment. *Proceedings of the International AAAI Conference on Web and Social Media*, 4(1), Artigo 1. <https://doi.org/10.1609/icwsm.v4i1.14075>

APÊNDICES

Apêndice A - 304 picos de atividade do Twitter identificados

Empresa	Data	SD	Polaridade	VT	Empresa	Data	SD	Polaridade	VT
Ambev	18/07/2022	0,500	Positivo	97	Gerdau	21/07/2023	0,833	Positivo	15
Ambev	01/02/2023	0,356	Positivo	574	Gerdau	09/08/2023	0,800	Positivo	24
Ambev	02/02/2023	0,336	Positivo	429	Gerdau	06/11/2023	0,818	Positivo	15
Ambev	03/02/2023	0,212	Positivo	191	Gerdau	07/11/2023	0,833	Positivo	38
Azul	10/11/2022	0,867	Positivo	20	Gerdau	20/11/2023	1,000	Positivo	15
Azul	14/11/2022	1,000	Positivo	16	Gol	28/04/2022	0,867	Positivo	17
Azul	02/02/2023	1,000	Positivo	13	Gol	06/07/2022	0,833	Positivo	13
Azul	13/02/2023	1,000	Positivo	17	Gol	27/07/2022	0,692	Positivo	13
Azul	06/03/2023	0,878	Positivo	55	Gol	28/07/2022	0,714	Positivo	18
Azul	07/03/2023	1,000	Positivo	21	Gol	27/10/2022	1,000	Positivo	15
Azul	15/05/2023	0,750	Positivo	25	Gol	06/12/2022	1,000	Positivo	20
Azul	10/08/2023	0,556	Positivo	21	Gol	04/01/2023	0,818	Positivo	16
Azul	20/09/2023	1,000	Positivo	23	Gol	07/02/2023	1,000	Positivo	25
Azul	02/10/2023	0,818	Positivo	13	Gol	06/03/2023	1,000	Positivo	35
Bradesco	25/02/2022	-0,628	Negativo	755	Gol	08/03/2023	0,733	Positivo	37
Bradesco	06/06/2022	-0,463	Negativo	455	Gol	11/04/2023	1,000	Positivo	16
Bradesco	28/09/2022	-0,418	Negativo	492	Gol	26/04/2023	0,909	Positivo	28
Bradesco	10/04/2023	-0,573	Negativo	421	Gol	06/06/2023	0,889	Positivo	18
Bradesco	28/04/2023	-0,564	Negativo	617	Gol	11/07/2023	0,846	Positivo	17
Bradesco	30/05/2023	-0,538	Negativo	727	Gol	14/07/2023	0,455	Positivo	16
Bradesco	07/08/2023	-0,456	Negativo	380	Gol	15/08/2023	0,818	Positivo	15
Bradesco	06/10/2023	-0,543	Negativo	792	Gol	20/09/2023	1,000	Positivo	17
Bradesco	25/10/2023	-0,604	Negativo	885	Gol	06/11/2023	1,000	Positivo	21
Bradesco	03/11/2023	-0,584	Negativo	384	Gol	04/12/2023	0,833	Positivo	14
Bradesco	28/11/2023	-0,530	Negativo	1.099	Itau	03/03/2022	-0,377	Negativo	17.524
Bradesco	11/12/2023	-0,610	Negativo	1.025	Itau	04/03/2022	-0,350	Negativo	1.801
Braskem	17/01/2022	0,714	Positivo	17	Itau	17/03/2022	-0,033	Negativo	668
Braskem	28/01/2022	0,765	Positivo	26	Itau	22/06/2022	0,586	Positivo	1.269
Braskem	30/03/2022	-0,091	Negativo	15	Itau	23/06/2022	0,517	Positivo	1.331
Braskem	07/04/2022	0,714	Positivo	22	Itau	07/08/2023	-0,505	Negativo	9.691
Braskem	11/04/2022	1,000	Positivo	18	Itau	08/08/2023	-0,393	Negativo	553
Braskem	11/10/2022	0,947	Positivo	45	Itau	07/12/2023	0,130	Positivo	1.395
Braskem	13/10/2022	1,000	Positivo	20	Itau	08/12/2023	0,001	Positivo	866
Braskem	28/11/2022	1,000	Positivo	14	Natura	09/03/2022	0,895	Positivo	20
Braskem	17/01/2023	-0,111	Negativo	46	Natura	15/06/2022	1,000	Positivo	16
Braskem	18/01/2023	0,006	Positivo	604	Natura	12/08/2022	0,733	Positivo	19
Braskem	19/01/2023	-0,138	Negativo	70	Natura	24/08/2022	1,000	Positivo	13
Braskem	27/01/2023	-0,100	Negativo	69	Natura	18/10/2022	1,000	Positivo	37
Braskem	16/03/2023	0,125	Positivo	24	Natura	30/01/2023	1,000	Positivo	19
Braskem	05/05/2023	0,773	Positivo	54	Natura	14/03/2023	0,182	Positivo	27
Braskem	08/05/2023	0,697	Positivo	41	Natura	04/04/2023	0,939	Positivo	39
Braskem	09/05/2023	0,778	Positivo	44	Natura	05/04/2023	0,900	Positivo	20
Braskem	10/05/2023	0,840	Positivo	29	Natura	09/05/2023	0,818	Positivo	40
Braskem	15/05/2023	0,688	Positivo	41	Pao De Acucar	02/06/2022	1,000	Positivo	15
Braskem	22/06/2023	0,143	Positivo	57	Pao De Acucar	05/09/2022	1,000	Positivo	17
Braskem	05/07/2023	0,304	Positivo	30	Pao De Acucar	04/11/2022	0,867	Positivo	18
Braskem	12/07/2023	0,667	Positivo	20	Pao De Acucar	10/01/2023	1,000	Positivo	30
Braskem	21/07/2023	0,565	Positivo	32	Pao De Acucar	29/06/2023	1,000	Positivo	19
Braskem	09/08/2023	0,286	Positivo	22	Pao De Acucar	19/07/2023	1,000	Positivo	17
Braskem	30/10/2023	1,000	Positivo	14	Pao De Acucar	27/07/2023	-0,143	Negativo	15
Braskem	09/11/2023	0,677	Positivo	48	Pao De Acucar	20/09/2023	1,000	Positivo	17
Braskem	01/12/2023	-0,207	Negativo	1.220	Pao De Acucar	16/10/2023	0,917	Positivo	27
Brf	02/02/2022	1,000	Positivo	15	Pao De Acucar	11/12/2023	1,000	Positivo	30
Brf	27/05/2022	1,000	Positivo	19	Pao De Acucar	13/12/2023	1,000	Positivo	17
Brf	23/06/2022	1,000	Positivo	14	Petrobras	11/01/2022	0,312	Positivo	335
Brf	11/08/2022	0,875	Positivo	18	Petrobras	10/03/2022	-0,024	Negativo	2.248
Brf	25/10/2022	1,000	Positivo	27	Petrobras	11/03/2022	-0,128	Negativo	1.714
Brf	29/12/2022	1,000	Positivo	22	Petrobras	28/03/2022	0,252	Positivo	803
Brf	24/03/2023	1,000	Positivo	20	Petrobras	29/03/2022	0,099	Positivo	898
Brf	16/05/2023	0,857	Positivo	19	Petrobras	24/05/2022	-0,048	Negativo	1.406
Brf	18/05/2023	0,909	Positivo	23	Petrobras	17/06/2022	-0,134	Negativo	5.263

Brf	01/06/2023	1,000	Positivo	13	Petrobras	20/06/2022	0,039	Positivo	2.626
Brf	15/08/2023	0,692	Positivo	15	Petrobras	28/07/2022	0,598	Positivo	630
Brf	16/10/2023	1,000	Positivo	15	Petrobras	29/07/2022	0,466	Positivo	387
Cemig	16/11/2022	-0,139	Negativo	177	Petrobras	15/08/2022	0,566	Positivo	435
Cemig	14/04/2023	0,349	Positivo	55	Petrobras	25/08/2022	0,071	Positivo	617
Cemig	15/08/2023	-0,304	Negativo	92	Petrobras	30/09/2022	0,074	Positivo	1.259
Cemig	14/09/2023	-0,244	Negativo	58	Petrobras	31/10/2022	0,117	Positivo	1.047
Copel	08/04/2022	-0,135	Negativo	44	Petrobras	14/12/2022	0,362	Positivo	516
Copel	07/06/2022	-0,429	Negativo	35	Petrobras	02/01/2023	0,319	Positivo	443
Copel	21/11/2022	0,451	Positivo	132	Petrobras	24/01/2023	0,497	Positivo	393
Copel	22/11/2022	0,243	Positivo	138	Petrobras	28/02/2023	0,343	Positivo	498
Copel	23/11/2022	0,414	Positivo	342	Petrobras	02/03/2023	0,147	Positivo	763
Copel	24/11/2022	0,123	Positivo	191	Petrobras	05/04/2023	0,330	Positivo	250
Copel	09/12/2022	-0,494	Negativo	97	Petrobras	16/05/2023	0,340	Positivo	2.689
Copel	23/05/2023	0,522	Positivo	78	Petrobras	30/06/2023	0,240	Positivo	222
Copel	01/06/2023	0,524	Positivo	62	Petrobras	04/08/2023	0,317	Positivo	504
Copel	12/07/2023	-0,038	Negativo	58	Petrobras	15/08/2023	0,224	Positivo	457
Copel	13/07/2023	-0,154	Negativo	59	Sabesp	01/02/2022	-0,045	Negativo	61
Copel	09/08/2023	0,146	Positivo	129	Sabesp	03/03/2022	-0,476	Negativo	106
Copel	15/08/2023	0,290	Positivo	139	Sabesp	03/05/2022	-0,181	Negativo	135
Copel	04/09/2023	0,000	Neutro	48	Sabesp	27/09/2022	0,455	Positivo	57
Copel	05/09/2023	0,083	Positivo	31	Sabesp	10/10/2022	0,132	Positivo	155
Copel	25/09/2023	-0,556	Negativo	36	Sabesp	27/10/2022	-0,010	Negativo	281
Copel	04/10/2023	-0,091	Negativo	60	Sabesp	03/01/2023	-0,321	Negativo	66
Copel	07/12/2023	-0,059	Negativo	20	Sabesp	04/01/2023	0,036	Positivo	70
Copel	18/12/2023	-0,317	Negativo	45	Sabesp	13/02/2023	0,077	Positivo	48
Cosan	07/10/2022	0,760	Positivo	31	Sabesp	14/02/2023	0,366	Positivo	50
Cosan	10/10/2022	0,643	Positivo	36	Sabesp	10/04/2023	0,200	Positivo	79
Cosan	02/06/2023	1,000	Positivo	13	Sabesp	01/08/2023	0,108	Positivo	197
Eletrobras	10/10/2022	0,297	Positivo	46	Sabesp	18/09/2023	-0,447	Negativo	160
Eletrobras	27/10/2022	0,545	Positivo	36	Sabesp	02/10/2023	0,356	Positivo	246
Eletrobras	28/10/2022	0,071	Positivo	40	Sabesp	03/10/2023	0,117	Positivo	787
Eletrobras	23/11/2022	0,882	Positivo	27	Sabesp	04/10/2023	0,162	Positivo	99
Eletrobras	14/12/2022	0,458	Positivo	151	Sabesp	17/10/2023	0,259	Positivo	89
Eletrobras	22/12/2022	0,864	Positivo	56	Sabesp	18/10/2023	0,070	Positivo	117
Eletrobras	30/03/2023	0,976	Positivo	105	Sabesp	16/11/2023	0,079	Positivo	200
Eletrobras	14/04/2023	0,697	Positivo	67	Sabesp	28/11/2023	0,185	Positivo	385
Eletrobras	15/06/2023	0,471	Positivo	52	Sabesp	06/12/2023	0,051	Positivo	1.286
Eletrobras	29/06/2023	0,524	Positivo	29	Sabesp	07/12/2023	-0,085	Negativo	1.123
Eletrobras	21/07/2023	0,825	Positivo	81	Sabesp	08/12/2023	-0,231	Negativo	242
Eletrobras	15/08/2023	0,017	Positivo	1.774	Santander	24/01/2022	-0,027	Negativo	495
Eletrobras	16/08/2023	-0,134	Negativo	393	Santander	04/02/2022	-0,729	Negativo	434
Eletrobras	17/08/2023	-0,013	Negativo	116	Santander	29/04/2022	-0,576	Negativo	546
Eletrobras	07/11/2023	0,000	Neutro	21	Santander	25/07/2022	-0,570	Negativo	2.766
Eletrobras	08/11/2023	0,579	Positivo	28	Santander	13/06/2023	0,127	Positivo	1.595
Embraer	12/01/2022	0,846	Positivo	16	Santander	16/06/2023	0,067	Positivo	441
Embraer	09/02/2022	0,167	Positivo	18	Santander	23/08/2023	-0,302	Negativo	345
Embraer	03/03/2022	0,700	Positivo	23	Santander	30/08/2023	-0,011	Negativo	924
Embraer	10/03/2022	0,789	Positivo	25	Santander	01/09/2023	-0,228	Negativo	340
Embraer	28/04/2022	0,909	Positivo	29	Santander	04/09/2023	-0,336	Negativo	603
Embraer	24/05/2022	0,500	Positivo	21	Santander	12/12/2023	-0,078	Negativo	318
Embraer	19/07/2022	1,000	Positivo	35	Santander	13/12/2023	0,030	Positivo	337
Embraer	04/08/2022	0,765	Positivo	19	Sid Nacional	19/05/2022	1,000	Positivo	19
Embraer	22/08/2022	0,294	Positivo	19	Sid Nacional	09/09/2022	1,000	Positivo	15
Embraer	24/08/2022	-0,286	Negativo	20	Sid Nacional	21/09/2022	1,000	Positivo	19
Embraer	15/09/2022	0,905	Positivo	28	Sid Nacional	01/11/2022	1,000	Positivo	19
Embraer	20/09/2022	1,000	Positivo	28	Sid Nacional	11/11/2022	1,000	Positivo	18
Embraer	17/10/2022	0,882	Positivo	25	Sid Nacional	22/11/2022	1,000	Positivo	32
Embraer	07/11/2022	0,500	Positivo	33	Sid Nacional	25/11/2022	1,000	Positivo	19
Embraer	14/11/2022	0,800	Positivo	37	Sid Nacional	26/12/2022	1,000	Positivo	20
Embraer	16/01/2023	0,818	Positivo	22	Sid Nacional	01/03/2023	0,846	Positivo	13
Embraer	17/02/2023	0,889	Positivo	20	Sid Nacional	09/03/2023	1,000	Positivo	28
Embraer	10/03/2023	0,778	Positivo	20	Sid Nacional	14/04/2023	1,000	Positivo	13
Embraer	24/04/2023	0,571	Positivo	31	Sid Nacional	04/05/2023	0,800	Positivo	24
Embraer	26/04/2023	0,818	Positivo	28	Sid Nacional	14/11/2023	1,000	Positivo	18
Embraer	04/05/2023	0,583	Positivo	28	Sid Nacional	20/11/2023	0,833	Positivo	13
Embraer	09/05/2023	0,579	Positivo	21	Suzano	10/02/2022	1,000	Positivo	16

Embraer	25/05/2023	0,727	Positivo	22	Suzano	05/05/2022	1,000	Positivo	21
Embraer	20/07/2023	0,857	Positivo	18	Suzano	28/07/2022	1,000	Positivo	13
Embraer	04/08/2023	1,000	Positivo	32	Suzano	02/12/2022	0,913	Positivo	24
Embraer	14/08/2023	1,000	Positivo	21	Suzano	28/04/2023	0,905	Positivo	23
Embraer	23/08/2023	0,204	Positivo	148	Suzano	23/05/2023	1,000	Positivo	26
Embraer	24/08/2023	0,267	Positivo	40	Suzano	03/08/2023	0,846	Positivo	13
Embraer	03/10/2023	0,368	Positivo	35	Telef Brasil	27/07/2022	0,846	Positivo	13
Embraer	04/10/2023	0,391	Positivo	27	Telef Brasil	26/10/2022	1,000	Positivo	17
Embraer	29/11/2023	0,579	Positivo	28	Telef Brasil	12/12/2022	1,000	Positivo	17
Gerdau	23/02/2022	1,000	Positivo	23	Telef Brasil	26/07/2023	1,000	Positivo	16
Gerdau	27/04/2022	0,889	Positivo	21	Telef Brasil	18/09/2023	1,000	Positivo	13
Gerdau	05/05/2022	0,926	Positivo	34	Telef Brasil	01/11/2023	1,000	Positivo	15
Gerdau	20/05/2022	0,935	Positivo	40	Ultrapar	20/04/2023	1,000	Positivo	12
Gerdau	30/05/2022	0,240	Positivo	63	Ultrapar	04/05/2023	1,000	Positivo	18
Gerdau	03/08/2022	0,938	Positivo	36	Ultrapar	10/08/2023	1,000	Positivo	14
Gerdau	02/09/2022	0,875	Positivo	17	Vale	06/04/2022	1,000	Positivo	19
Gerdau	05/09/2022	0,613	Positivo	36	Vale	28/07/2022	1,000	Positivo	35
Gerdau	27/09/2022	1,000	Positivo	22	Vale	29/07/2022	0,795	Positivo	46
Gerdau	11/10/2022	0,636	Positivo	15	Vale	18/10/2022	0,886	Positivo	44
Gerdau	09/11/2022	0,955	Positivo	50	Vale	17/02/2023	0,931	Positivo	35
Gerdau	28/11/2022	0,733	Positivo	17	Vale	28/07/2023	0,846	Positivo	63
Gerdau	01/03/2023	0,946	Positivo	40	Vale	01/09/2023	0,806	Positivo	33
Gerdau	06/03/2023	0,273	Positivo	16	Vale	14/09/2023	0,760	Positivo	26
Gerdau	03/05/2023	1,000	Positivo	27	Vale	27/10/2023	0,871	Positivo	34

Apêndice B – 30 picos de atividade no Twitter que coincidem com os anúncios de resultados

Empresa	Data	SD	Polaridade	VT	Empresa	Data	SD	Polaridade	VT
Azul	10/11/2022	0,867	Positivo	20	Gerdau	09/11/2022	0,955	Positivo	50
Azul	06/03/2023	0,878	Positivo	55	Gerdau	01/03/2023	0,946	Positivo	40
Azul	15/05/2023	0,750	Positivo	25	Gerdau	03/05/2023	1,000	Positivo	27
Azul	10/08/2023	0,556	Positivo	21	Gerdau	06/11/2023	0,818	Positivo	15
Braskem	08/05/2023	0,697	Positivo	41	Gol	28/04/2022	0,867	Positivo	17
Eletrobras	08/11/2023	0,579	Positivo	28	Gol	28/07/2022	0,714	Positivo	18
Embraer	28/04/2022	0,909	Positivo	29	Gol	27/10/2022	1,000	Positivo	15
Embraer	04/08/2022	0,765	Positivo	19	Gol	26/04/2023	0,909	Positivo	28
Embraer	14/11/2022	0,800	Positivo	37	Gol	06/11/2023	1,000	Positivo	21
Embraer	10/03/2023	0,778	Positivo	20	Itaú	07/08/2023	-0,505	Negativo	9.691
Embraer	04/05/2023	0,583	Positivo	28	Natura	09/03/2022	0,895	Positivo	20
Embraer	14/08/2023	1,000	Positivo	21	Petrobras	28/07/2022	0,598	Positivo	630
Gerdau	23/02/2022	1,000	Positivo	23	Telef Brasil	27/07/2022	0,846	Positivo	13
Gerdau	05/05/2022	0,926	Positivo	34	Telef Brasil	26/10/2022	1,000	Positivo	17
Gerdau	03/08/2022	0,938	Positivo	36	Vale	28/07/2022	1,000	Positivo	35

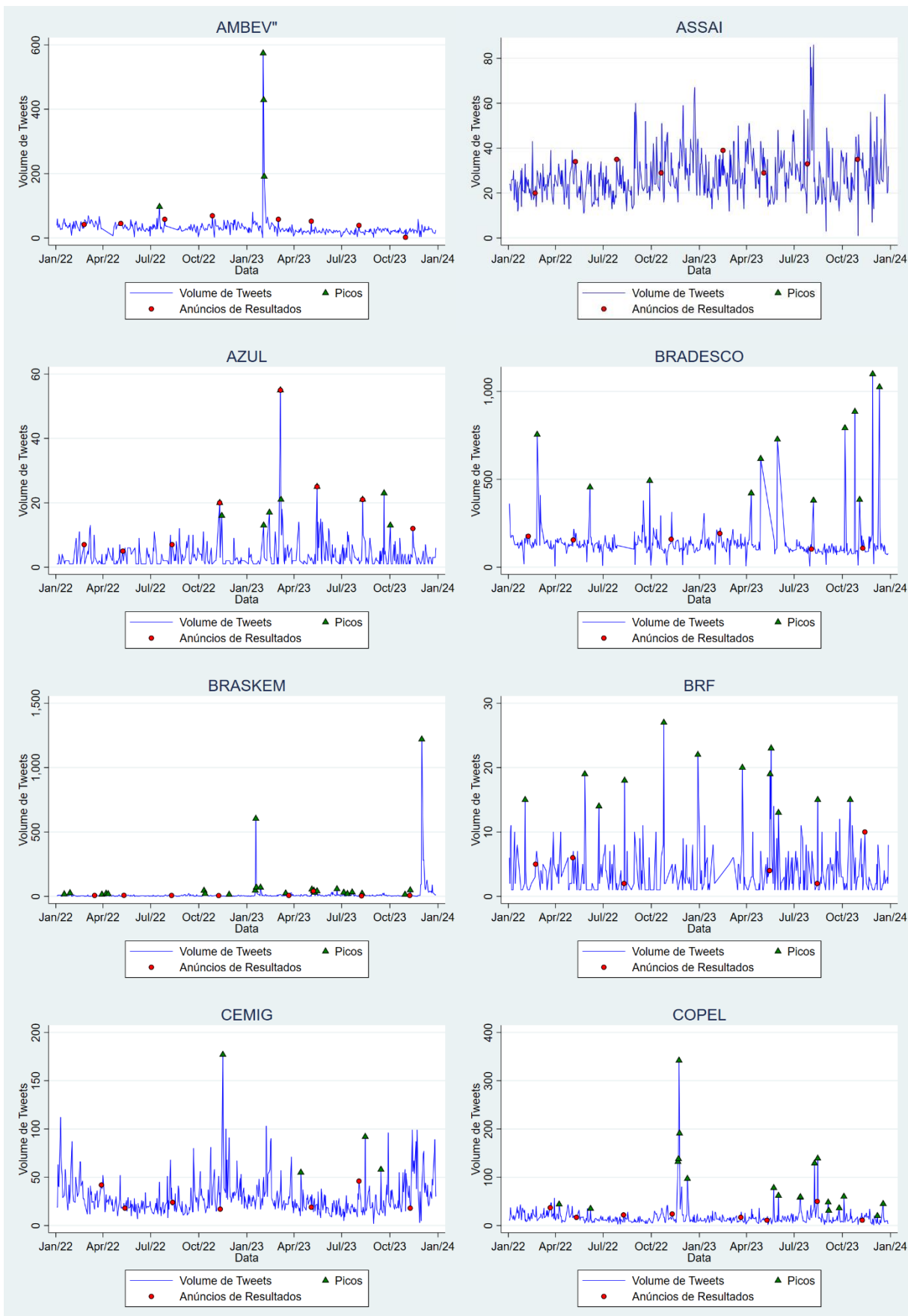
Apêndice C – 233 picos de atividades no Twitter não relacionados aos EA

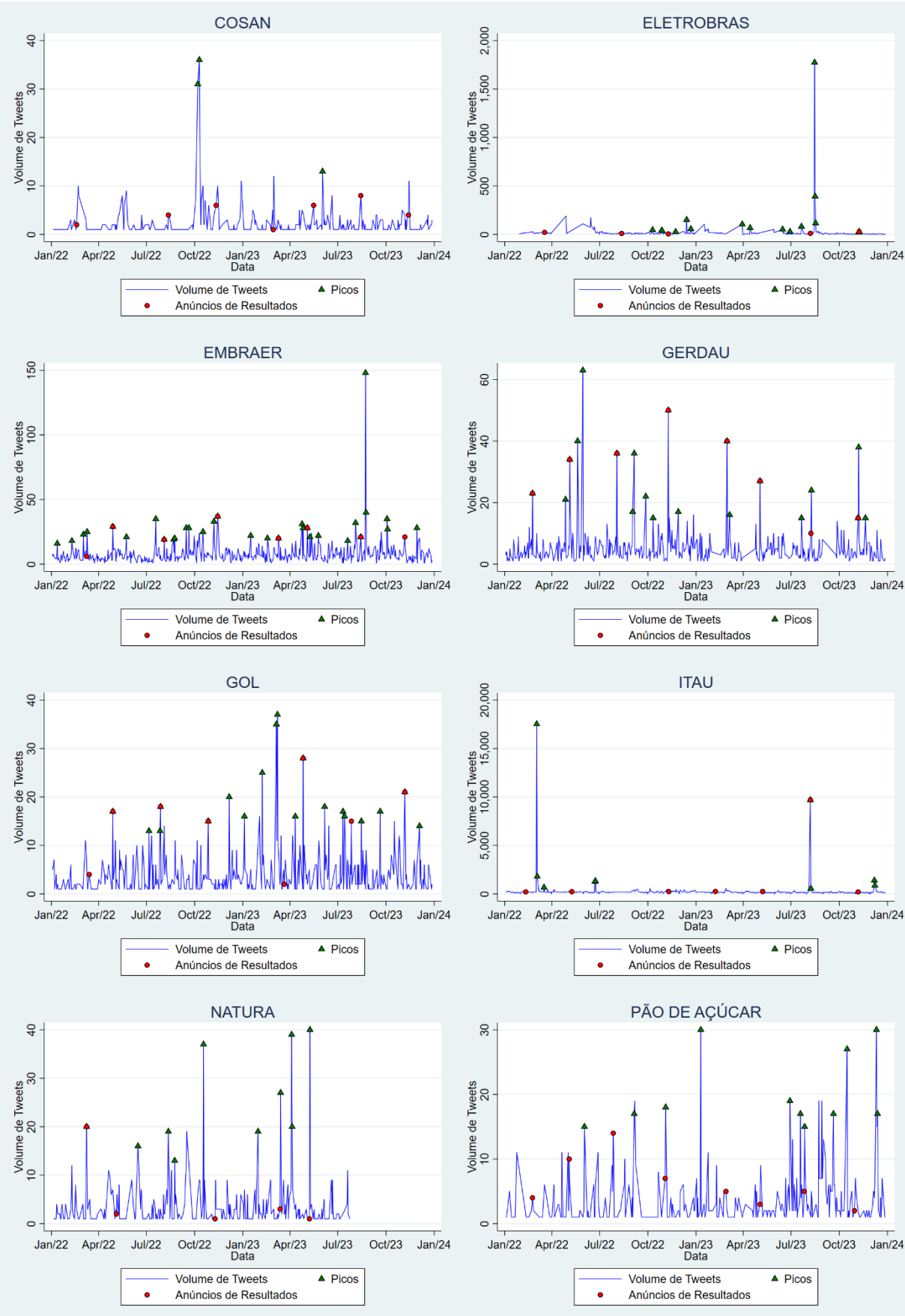
Empresa	Data	SD	Polaridade	VT	Empresa	Data	SD	Polaridade	VT
Ambev	18/07/2022	0,5000	Positivo	97	Gerdau	02/09/2022	0,8750	Positivo	17
Ambev	01/02/2023	0,3560	Positivo	574	Gerdau	05/09/2022	0,6129	Positivo	36
Ambev	02/02/2023	0,3356	Positivo	429	Gerdau	27/09/2022	1,0000	Positivo	22
Ambev	03/02/2023	0,2121	Positivo	191	Gerdau	11/10/2022	0,6364	Positivo	15
Azul	14/11/2022	1,0000	Positivo	16	Gerdau	28/11/2022	0,7333	Positivo	17
Azul	02/02/2023	1,0000	Positivo	13	Gerdau	06/03/2023	0,2727	Positivo	16
Azul	13/02/2023	1,0000	Positivo	17	Gerdau	21/07/2023	0,8333	Positivo	15
Azul	20/09/2023	1,0000	Positivo	23	Gerdau	20/11/2023	1,0000	Positivo	15
Azul	02/10/2023	0,8182	Positivo	13	Gol	06/07/2022	0,8333	Positivo	13
Bradesco	25/02/2022	-0,6277	Negativo	755	Gol	06/12/2022	1,0000	Positivo	20
Bradesco	06/06/2022	-0,4629	Negativo	455	Gol	04/01/2023	0,8182	Positivo	16
Bradesco	28/09/2022	-0,4177	Negativo	492	Gol	07/02/2023	1,0000	Positivo	25
Bradesco	10/04/2023	-0,5733	Negativo	421	Gol	06/03/2023	1,0000	Positivo	35
Bradesco	28/04/2023	-0,5638	Negativo	617	Gol	08/03/2023	0,7333	Positivo	37
Bradesco	30/05/2023	-0,5385	Negativo	727	Gol	11/04/2023	1,0000	Positivo	16
Bradesco	07/08/2023	-0,4563	Negativo	380	Gol	06/06/2023	0,8889	Positivo	18
Bradesco	06/10/2023	-0,5432	Negativo	792	Gol	11/07/2023	0,8462	Positivo	17
Bradesco	25/10/2023	-0,6037	Negativo	885	Gol	14/07/2023	0,4545	Positivo	16
Bradesco	03/11/2023	-0,5838	Negativo	384	Gol	15/08/2023	0,8182	Positivo	15
Bradesco	28/11/2023	-0,5300	Negativo	1.099	Gol	20/09/2023	1,0000	Positivo	17
Bradesco	11/12/2023	-0,6098	Negativo	1.025	Gol	04/12/2023	0,8333	Positivo	14
Braskem	17/01/2022	0,7143	Positivo	17	Itaú	03/03/2022	-0,3770	Negativo	17.524
Braskem	28/01/2022	0,7647	Positivo	26	Itaú	04/03/2022	-0,3501	Negativo	1.801
Braskem	30/03/2022	-0,0909	Negativo	15	Itaú	17/03/2022	-0,0330	Negativo	668
Braskem	07/04/2022	0,7143	Positivo	22	Itaú	22/06/2022	0,5865	Positivo	1.269
Braskem	11/04/2022	1,0000	Positivo	18	Itaú	23/06/2022	0,5174	Positivo	1.331
Braskem	11/10/2022	0,9474	Positivo	45	Itaú	07/12/2023	0,1299	Positivo	1.395
Braskem	13/10/2022	1,0000	Positivo	20	Itaú	08/12/2023	0,0014	Positivo	866
Braskem	28/11/2022	1,0000	Positivo	14	Natura	15/06/2022	1,0000	Positivo	16
Braskem	17/01/2023	-0,1111	Negativo	46	Natura	24/08/2022	1,0000	Positivo	13
Braskem	18/01/2023	0,0062	Positivo	604	Natura	18/10/2022	1,0000	Positivo	37
Braskem	19/01/2023	-0,1379	Negativo	70	Natura	30/01/2023	1,0000	Positivo	19
Braskem	27/01/2023	-0,1000	Negativo	69	Natura	04/04/2023	0,9394	Positivo	39
Braskem	16/03/2023	0,1250	Positivo	24	Natura	05/04/2023	0,9000	Positivo	20
Braskem	10/05/2023	0,8400	Positivo	29	Pão de Açúcar	02/06/2022	1,0000	Positivo	15
Braskem	15/05/2023	0,6875	Positivo	41	Pão de Açúcar	05/09/2022	1,0000	Positivo	17
Braskem	22/06/2023	0,1429	Positivo	57	Pão de Açúcar	10/01/2023	1,0000	Positivo	30
Braskem	05/07/2023	0,3043	Positivo	30	Pão de Açúcar	29/06/2023	1,0000	Positivo	19
Braskem	12/07/2023	0,6667	Positivo	20	Pão de Açúcar	19/07/2023	1,0000	Positivo	17
Braskem	21/07/2023	0,5652	Positivo	32	Pão de Açúcar	20/09/2023	1,0000	Positivo	17
Braskem	30/10/2023	1,0000	Positivo	14	Pão de Açúcar	16/10/2023	0,9167	Positivo	27
Braskem	01/12/2023	-0,2065	Negativo	1.22	Pão de Açúcar	11/12/2023	1,0000	Positivo	30
BRF	02/02/2022	1,0000	Positivo	15	Pão de Açúcar	13/12/2023	1,0000	Positivo	17
BRF	27/05/2022	1,0000	Positivo	19	Petrobras	11/01/2022	0,3117	Positivo	335
BRF	23/06/2022	1,0000	Positivo	14	Petrobras	10/03/2022	-0,0240	Negativo	2.248
BRF	25/10/2022	1,0000	Positivo	27	Petrobras	11/03/2022	-0,1284	Negativo	1.714
BRF	29/12/2022	1,0000	Positivo	22	Petrobras	28/03/2022	0,2522	Positivo	803
BRF	24/03/2023	1,0000	Positivo	20	Petrobras	29/03/2022	0,0989	Positivo	898
BRF	18/05/2023	0,9091	Positivo	23	Petrobras	24/05/2022	-0,0483	Negativo	1.406
BRF	01/06/2023	1,0000	Positivo	13	Petrobras	17/06/2022	-0,1344	Negativo	5.263
BRF	16/10/2023	1,0000	Positivo	15	Petrobras	20/06/2022	0,0390	Positivo	2.626
Cemig	16/11/2022	-0,1391	Negativo	177	Petrobras	15/08/2022	0,5660	Positivo	435
Cemig	14/04/2023	0,3488	Positivo	55	Petrobras	25/08/2022	0,0709	Positivo	617
Cemig	15/08/2023	-0,3043	Negativo	92	Petrobras	30/09/2022	0,0744	Positivo	1.259

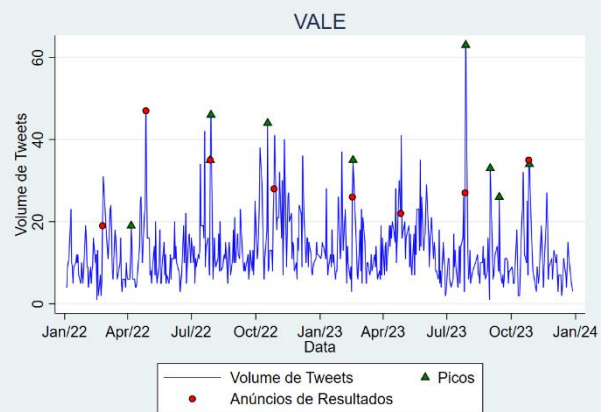
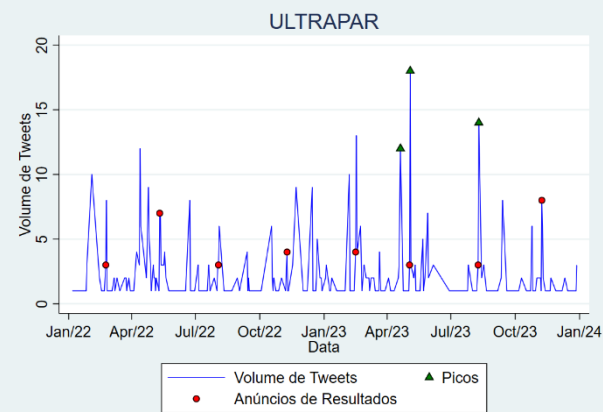
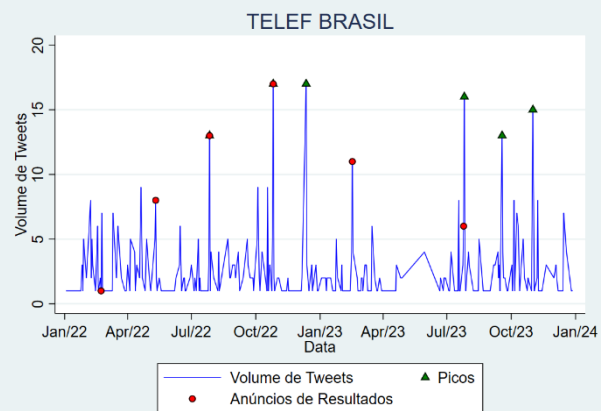
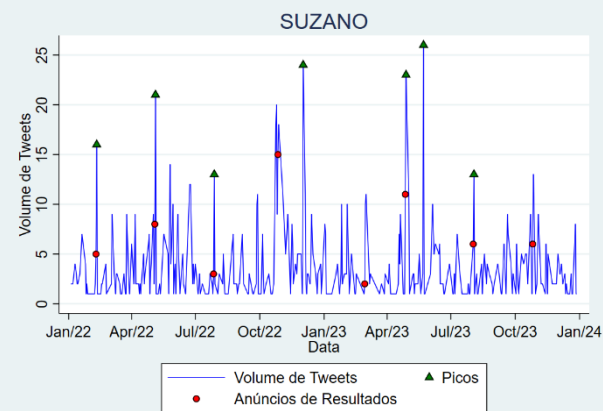
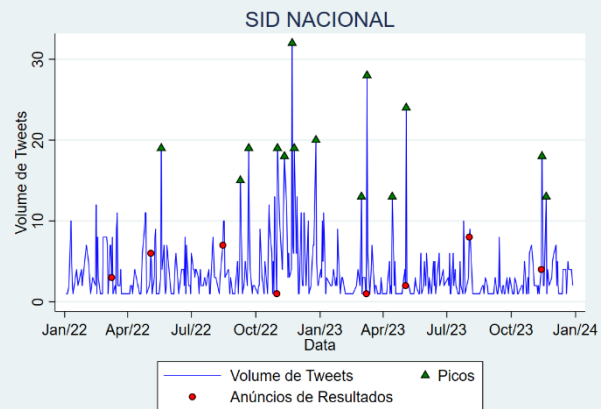
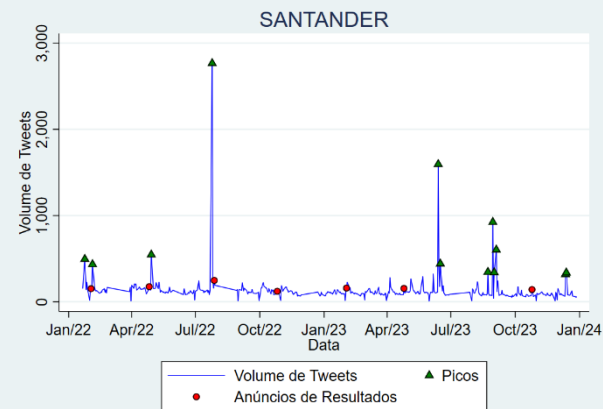
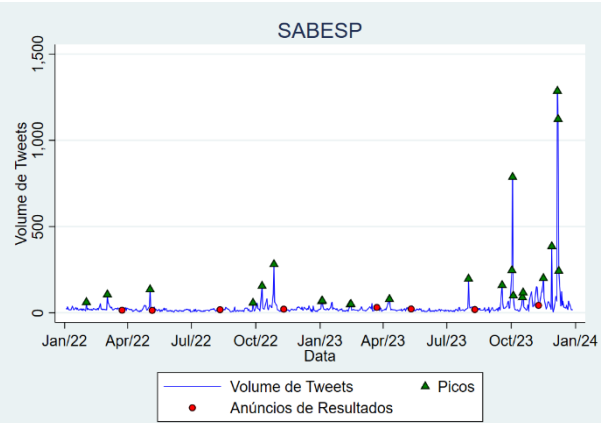
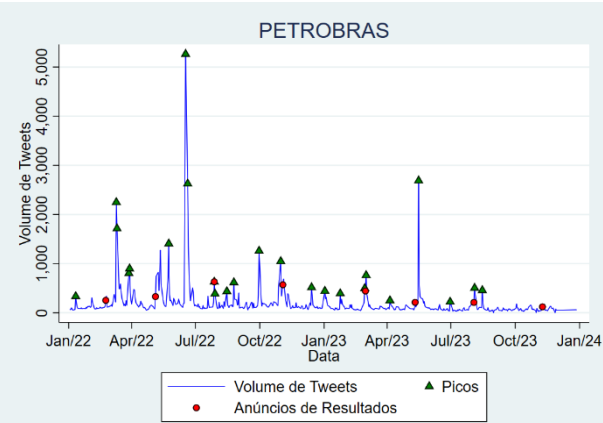
Cemig	14/09/2023	-0,2444	Negativo	58	Petrobras	31/10/2022	0,1167	Positivo	1.047
Copel	08/04/2022	-0,1351	Negativo	44	Petrobras	14/12/2022	0,3619	Positivo	516
Copel	07/06/2022	-0,4286	Negativo	35	Petrobras	02/01/2023	0,3188	Positivo	443
Copel	21/11/2022	0,4510	Positivo	132	Petrobras	24/01/2023	0,4967	Positivo	393
Copel	22/11/2022	0,2427	Positivo	138	Petrobras	05/04/2023	0,3297	Positivo	250
Copel	23/11/2022	0,4138	Positivo	342	Petrobras	16/05/2023	0,3403	Positivo	2.689
Copel	24/11/2022	0,1226	Positivo	191	Petrobras	30/06/2023	0,2398	Positivo	222
Copel	09/12/2022	-0,4940	Negativo	97	Petrobras	15/08/2023	0,2240	Positivo	457
Copel	23/05/2023	0,5224	Positivo	78	Sabesp	01/02/2022	-0,0455	Negativo	61
Copel	01/06/2023	0,5238	Positivo	62	Sabesp	03/03/2022	-0,4762	Negativo	106
Copel	12/07/2023	-0,0385	Negativo	58	Sabesp	03/05/2022	-0,1810	Negativo	135
Copel	13/07/2023	-0,1538	Negativo	59	Sabesp	27/09/2022	0,4545	Positivo	57
Copel	09/08/2023	0,1456	Positivo	129	Sabesp	10/10/2022	0,1321	Positivo	155
Copel	04/09/2023	0,0000	Neutro	48	Sabesp	27/10/2022	-0,0096	Negativo	281
Copel	05/09/2023	0,0833	Positivo	31	Sabesp	03/01/2023	-0,3208	Negativo	66
Copel	25/09/2023	-0,5556	Negativo	36	Sabesp	04/01/2023	0,0357	Positivo	70
Copel	04/10/2023	-0,0909	Negativo	60	Sabesp	13/02/2023	0,0769	Positivo	48
Copel	07/12/2023	-0,0588	Negativo	20	Sabesp	14/02/2023	0,3659	Positivo	50
Copel	18/12/2023	-0,3171	Negativo	45	Sabesp	10/04/2023	0,2000	Positivo	79
Cosan	07/10/2022	0,7600	Positivo	31	Sabesp	01/08/2023	0,1079	Positivo	197
Cosan	10/10/2022	0,6429	Positivo	36	Sabesp	18/09/2023	-0,4472	Negativo	160
Cosan	02/06/2023	1,0000	Positivo	13	Sabesp	02/10/2023	0,3563	Positivo	246
Eletrobras	10/10/2022	0,2973	Positivo	46	Sabesp	03/10/2023	0,1169	Positivo	787
Eletrobras	27/10/2022	0,5455	Positivo	36	Sabesp	04/10/2023	0,1622	Positivo	99
Eletrobras	28/10/2022	0,0714	Positivo	40	Sabesp	17/10/2023	0,2593	Positivo	89
Eletrobras	23/11/2022	0,8824	Positivo	27	Sabesp	18/10/2023	0,0698	Positivo	117
Eletrobras	14/12/2022	0,4576	Positivo	151	Sabesp	16/11/2023	0,0791	Positivo	200
Eletrobras	22/12/2022	0,8636	Positivo	56	Sabesp	28/11/2023	0,1847	Positivo	385
Eletrobras	30/03/2023	0,9756	Positivo	105	Sabesp	06/12/2023	0,0505	Positivo	1.286
Eletrobras	14/04/2023	0,6970	Positivo	67	Sabesp	07/12/2023	-0,0849	Negativo	1.123
Eletrobras	15/06/2023	0,4706	Positivo	52	Sabesp	08/12/2023	-0,2308	Negativo	242
Eletrobras	29/06/2023	0,5238	Positivo	29	Santander	24/01/2022	-0,0269	Negativo	495
Eletrobras	21/07/2023	0,8246	Positivo	81	Santander	04/02/2022	-0,7292	Negativo	434
Eletrobras	15/08/2023	0,0175	Positivo	1.774	Santander	29/04/2022	-0,5759	Negativo	546
Eletrobras	16/08/2023	-0,1343	Negativo	393	Santander	25/07/2022	-0,5698	Negativo	2.766
Eletrobras	17/08/2023	-0,0127	Negativo	116	Santander	13/06/2023	0,1267	Positivo	1.595
Embraer	12/01/2022	0,8462	Positivo	16	Santander	16/06/2023	0,0674	Positivo	441
Embraer	09/02/2022	0,1667	Positivo	18	Santander	23/08/2023	-0,3022	Negativo	345
Embraer	03/03/2022	0,7000	Positivo	23	Santander	30/08/2023	-0,0115	Negativo	924
Embraer	24/05/2022	0,5000	Positivo	21	Santander	01/09/2023	-0,2285	Negativo	340
Embraer	19/07/2022	1,0000	Positivo	35	Santander	04/09/2023	-0,3360	Negativo	603
Embraer	22/08/2022	0,2941	Positivo	19	Santander	12/12/2023	-0,0775	Negativo	318
Embraer	24/08/2022	-0,2857	Negativo	20	Santander	13/12/2023	0,0296	Positivo	337
Embraer	15/09/2022	0,9048	Positivo	28	Sid Nacional	19/05/2022	1,0000	Positivo	19
Embraer	20/09/2022	1,0000	Positivo	28	Sid Nacional	09/09/2022	1,0000	Positivo	15
Embraer	17/10/2022	0,8824	Positivo	25	Sid Nacional	21/09/2022	1,0000	Positivo	19
Embraer	07/11/2022	0,5000	Positivo	33	Sid Nacional	11/11/2022	1,0000	Positivo	18
Embraer	16/01/2023	0,8182	Positivo	22	Sid Nacional	22/11/2022	1,0000	Positivo	32
Embraer	17/02/2023	0,8889	Positivo	20	Sid Nacional	25/11/2022	1,0000	Positivo	19
Embraer	24/04/2023	0,5714	Positivo	31	Sid Nacional	26/12/2022	1,0000	Positivo	20
Embraer	26/04/2023	0,8182	Positivo	28	Sid Nacional	01/03/2023	0,8462	Positivo	13
Embraer	09/05/2023	0,5789	Positivo	21	Sid Nacional	14/04/2023	1,0000	Positivo	13
Embraer	25/05/2023	0,7273	Positivo	22	Sid Nacional	20/11/2023	0,8333	Positivo	13
Embraer	20/07/2023	0,8571	Positivo	18	Suzano	02/12/2022	0,9130	Positivo	24
Embraer	04/08/2023	1,0000	Positivo	32	Suzano	23/05/2023	1,0000	Positivo	26
Embraer	23/08/2023	0,2043	Positivo	148	Telef Brasil	12/12/2022	1,0000	Positivo	17
Embraer	24/08/2023	0,2667	Positivo	40	Telef Brasil	18/09/2023	1,0000	Positivo	13

Embraer	03/10/2023	0,3684	Positivo	35	Ultrapar	20/04/2023	1,0000	Positivo	12
Embraer	04/10/2023	0,3913	Positivo	27	Vale	06/04/2022	1,0000	Positivo	19
Embraer	29/11/2023	0,5789	Positivo	28	Vale	18/10/2022	0,8857	Positivo	44
Gerdau	27/04/2022	0,8889	Positivo	21	Vale	01/09/2023	0,8065	Positivo	33
Gerdau	20/05/2022	0,9355	Positivo	40	Vale	14/09/2023	0,7600	Positivo	26
Gerdau	30/05/2022	0,2400	Positivo	63					

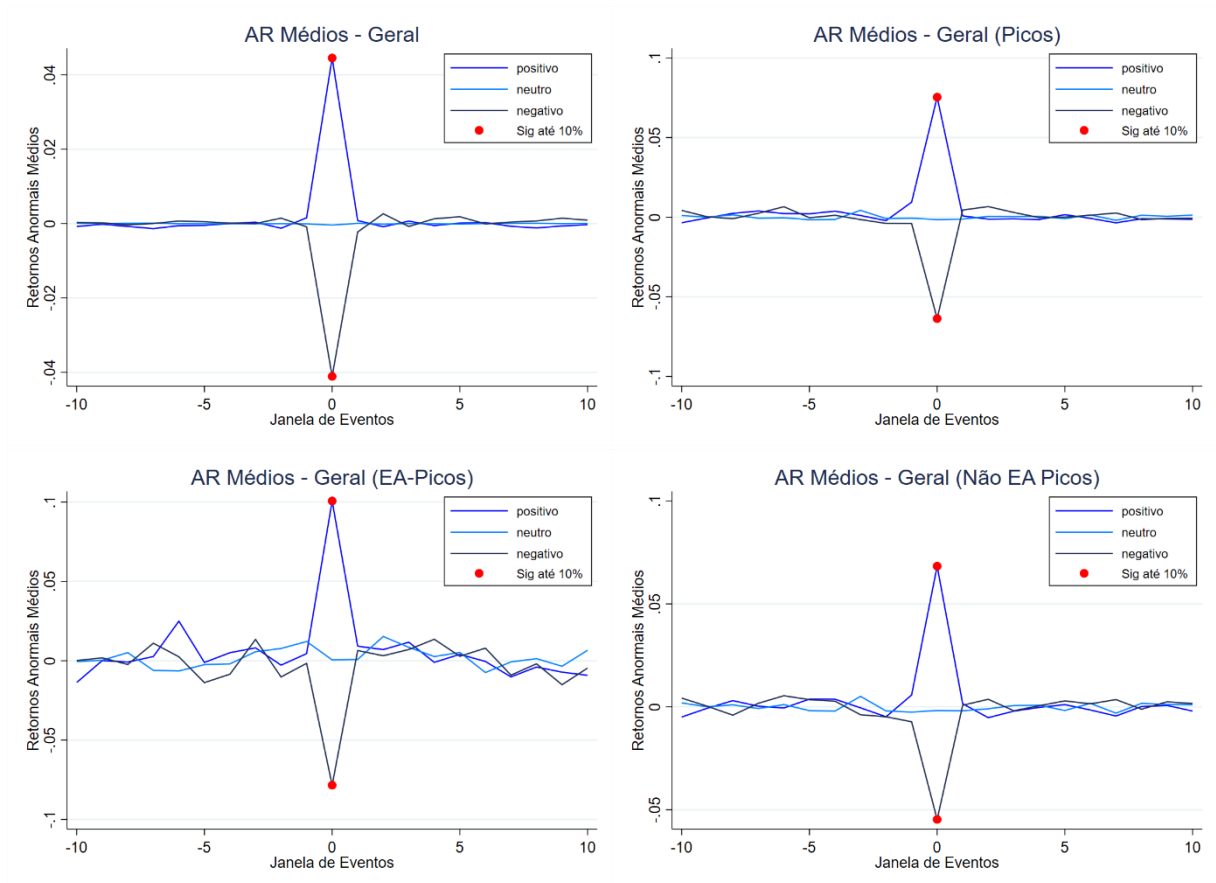
Apêndice D – Série diária do volume de *tweets*, picos de *tweets* e anúncios de resultados







Apêndice E – Figuras: Retornos Anormais Médios



Apêndice F – Sentimento Diário por Setor

Sentimento Diário por Setor	Média	Mediana	Desvio Padrão	Assimetria	Curtose	Obs	Mínimo	Máximo
Bens industriais	0,69620	1,00000	0,49715	-1,91952	6,35130	1.229	-1	1
Comunicações	0,79204	1,00000	0,45613	-2,17605	7,13986	252	-1	1
Consumo não cíclicos	0,61927	0,62500	0,38866	-0,74408	3,41362	1.837	-1	1
Financeiro	-0,27932	-0,29032	0,16477	0,82838	6,99551	1.323	-1	1
Materiais básico	0,73295	1,00000	0,45767	-1,97006	6,59446	2.097	-1	1
Petróleo, gás e biocombustíveis	0,57903	0,58397	0,43147	-0,94152	4,23841	968	-1	1
Utilidade pública	-0,01662	-0,07692	0,42693	0,51564	3,10920	1.809	-1	1

Apêndice G – Rotina STATA 16 – Estudo de Eventos

```
clear all
edit

// importar e configurar as bases
import excel "D:\Beatriz_Stata\eventdata.xlsx", sheet("eventos") firstrow
save "D:\Beatriz_Stata\eventdata", replace
use "D:\Beatriz_Stata\eventdata", clear
    encode company , gen (companynum)
        gen company_id = companynum
        drop company
        ren companynum company
    gen event_date = date
        format event_date %tdnn/dd/YY
        recast int event_date, force
        recast int date, force
save "D:\Beatriz_Stata\eventdata2", replace
clear

import excel "D:\Beatriz_Stata\stockdata.xlsx", sheet("stock_data") firstrow
save "D:\Beatriz_Stata\stockdata", replace
use "D:\Beatriz_Stata\stockdata", clear
    encode company , gen (companynum)
        gen company_id = companynum
        drop company
        ren companynum company
        sort company_id date
        by company_id: gen year = year(date)
save "D:\Beatriz_Stata\stockdata2", replace

// Unindo as bases baseado nas empresas e datas

use "D:\Beatriz_Stata\stockdata2", clear
    merge m:m company_id date using "D:\Beatriz_Stata\eventdata2"
    tab _merge
    drop if _merge==2
save "D:\Beatriz_Stata\merged", replace
set more off

// Gerar as marcações para os retornos

use "D:\Beatriz_Stata\merged", clear
sort company_id date
by company_id: gen datenum=_n
by company_id: gen target=datenum if date==event_date
egen td=min(target), by(company_id)
drop target
summarize td
gen dif=datenum-td

// calculando retornos das ações e do índice de mercado

// RETORNOS

xtset company_id datenum
// Retorno das ações
    bysort company_id: gen stock_return = log(close/L.close)
    list company_id company date stock_return if company_id[_n] != company_id[_n-1]
```

```

// Retorno do do Ibovespa
    bysort company_id: gen market_return = log(ibov_close/L.ibov_close)
    list company_id company date market_return if company_id[_n] != company_id[_n-1]

save "D:\Beatriz_Stata\merged02", replace
use "D:\Beatriz_Stata\merged02", clear

// VOLATILIDADES

//Parkinson (1980)
    gen PK = (1 / (4 * ln(2))) * ((ln(high/low))^2)

//Garman e Klass (1980)
    gen GK = 0.5 * ((ln(high/low))^2) - (2*ln(2)-1)* ((ln(close/open))^2)

// Rogers e Satchell (1991)
    gen RS = (ln(high/close))*(ln(high/open))+(ln(low/close))*(ln(low/open))

// Yang e Zhang (2000)
    gen k = (0.34/(1.34+((30+1)/(30-1)))) // 30 equivale a "n" que representa o nº de períodos de
    dados históricos analisados, sendo > 1
    gen YZ =.
    forvalues i = 1(1)24 {
        l company_id if company_id==`i'
        replace YZ = ((ln(open[_n]/close[_n-1]))^2) + k * ((ln(close[_n]/open[_n]))^2) + (1-
k) * RS if company_id==`i'
    }

// Gerando Retorno Previsto, Retorno Anormal (AR) e Retorno Anormal Acumulado (CAR)

// RETORNOS PREVISTOS (Janela de estimação de 30 dias)
    bys company_id: asreg stock_return market_return, wind(datenum -31 -1) minimum(30) se rmse fit
    ren _b_cons intercept
    ren _b_market_return slope
    ren _rmse rmse
    drop _Nobs _R2 _adjR2 _se_market_return _se_cons _fitted_residuals
    gen p_return = intercept + slope * market_return
    by company_id, sort : count if p_return ==.

// RETORNO ANORMAL (AR)
    gen AR = stock_return - p_return if dif>=-20

// Teste Significância do AR
    gen T_AR = AR / rmse
    gen label_T_AR = " "
    replace label_T_AR = "NS" if abs(T_AR) < 1.64 & dif>=0
    replace label_T_AR = "*" if abs(T_AR) >= 1.64 & dif>=0
    replace label_T_AR = "***" if abs(T_AR) >= 1.95 & dif>=0
    replace label_T_AR = "****" if abs(T_AR) >= 2.58 & dif>=0
    // Obs: 1,64 para 90% (*), 1,96 para 95% (**) e 2,58 para 99% (***)

* RETORNO ANORMAL ACUMULADO (CAR -5;+5) JANELA DE 11 DIAS
    gen CAR_5_5 =.
    forvalues i = 1/'_N' {
        local event_diff = dif[`i']
        if dif[`i'] < 0 continue

```

```

+ 5)          quietly sum AR if company_id == company_id[`i'] & inrange(dif, `event_diff' - 5, `event_diff'

              // alterar a janela para [-5;+5]
              replace CAR_5_5 = r(sum) if _n == `i'
            }

            *Teste Significância do CAR(-5;+5)
            gen T_CAR_5_5 = CAR_5_5 / rmse
            gen label_T_CAR_5_5 = " "
            replace label_T_CAR_5_5 = "NS" if abs(T_CAR_5_5) < 1.64 & dif>=0
            replace label_T_CAR_5_5 = "*" if abs(T_CAR_5_5) >= 1.64 & dif>=0
            replace label_T_CAR_5_5 = "***" if abs(T_CAR_5_5) >= 1.95 & dif>=0
            replace label_T_CAR_5_5 = "****" if abs(T_CAR_5_5) >= 2.58 & dif>=0
            // Obs: 1,64 para 90% (*), 1,96 para 95% (**) e 2,58 para 99% (***)

            * exportando e salvando a base
            *outsheet using "D:\Beatriz_Stata\base_AR_e_CAR.csv", delimiter(";") names replace
            *export excel using "D:\Beatriz_Stata\base_AR_e_CAR.xls", firstrow(variables) replace
            save "D:\Beatriz_Stata\base_AR_e_CAR", replace
            *export excel using "D:\Beatriz_Stata\merged03.xls", firstrow(variables) replace

```

Apêndice H – Rotina STATA 16 – Identificação dos Picos

```
clear all

// importando e mesclando base de anuncio de resultados
import excel "D:\Beatriz_Stata\anuncios_de_resultados.xlsx", sheet("Planilha1") firstrow
save "D:\Beatriz_Stata\anuncios_de_resultados", replace

use "D:\Beatriz_Stata\base_AR_e_CAR.dta ", clear
rename _merge _merge1
merge m:m company_id date using "D:\Beatriz_Stata\anuncios_de_resultados"
tab _merge
drop if _merge==2
save "D:\Beatriz_Stata\Base_Completa", replace
use "D:\Beatriz_Stata\Base_Completa", clear

// ORGANIZANDO VARIÁVEIS
edit
gen month= date
format month %tdMon
gen data= date
format data %tdMon/YY
format %15.2gc ibov_qtits
format %15.2gc ibov_vol
destring q_neg, replace
encode setor, ge(setor1)
gen setor_dummy = setor1
encode subsetor, ge(subsetor1)
gen subsetor_dummy = subsetor1
encode segmento, ge(segmento1)
gen segmento_dummy = segmento1
encode governanca, ge(gov1)
gen gov_dummy = gov1
encode Rating , ge(rating1)
gen rating_dummy = rating1
encode tipo, ge(tipo1)
gen tipo_dummy = tipo1
gen d_event = !missing(SD) // 0 para quando não teve postagem e 1 para quando teve
gen date_EA = date if _merge==3
format date_EA %tdnn/dd/YY
gen dummy_EA = 1 if !missing(date_EA)

xtset company_id date

//-----//

// PICOS

// RANCO et al. (2015)

// IDENTIFICAÇÃO DOS PICOS DE TWEETS
sort company_id date
ssc install rangestat, replace
rangestat (median) Volume_median=Total_Cont, interval(datenum -5 5) by(company_id)
gen phi_d0 = (Total_Cont - Volume_median) / max(Volume_median, 5)
gen is_peak = (phi_d0 > 2) & !missing(phi_d0)
count if is_peak == 1
list company company_id date SD Rating Total_Cont if is_peak == 1
```

```

(EA) // VERIFICAÇÃO DE COINCIDÊNCIA ENTRE PICOS E DIVULGAÇÕES DE RESULTADOS

sort company_id date
gen is_EA_peak = is_peak == 1 & !missing(date_EA)
by company_id: summarize is_EA_peak if is_EA_peak == 1
count if is_EA_peak == 1
list company_id company date SD Rating Total_Cont if is_EA_peak == 1

// DETECÇÃO DE EVENTOS NÃO RELACIONADOS A DIVULGAÇÃO DE RESULTADOS (EA)
gen exclude = 0
replace exclude = 1 if dummy_EA == 1 | dummy_EA[_n-1] == 1 | dummy_EA[_n+1] == 1
count if is_peak == 1 & exclude == 0
list company company_id date SD Rating Total_Cont if is_peak == 1 & exclude == 0
gen non_EA_peak = is_peak == 1 & exclude == 0
count if non_EA_peak == 1

list company company_id date SD Rating Total_Cont if non_EA_peak == 1

save "D:\Beatriz_Stata\Base_Completa2", replace

//-----//

drop if year == 2021
drop if year == 2024

// GRÁFICOS

// Histogramas dos Sentimentos
//Geral
preserve
histogram SD, xscale(range(-1 1)) title("Sentimento Diário - GERAL")
graph export "D:\Beatriz_Stata\Imagens_Dissertação\SD_GERAL.png", as(png) replace
restore

// Individual
local company_ids "1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24"
local company_names "AMBEV" "ASSAI" "AZUL" "BRADESCO" "BRASKEM" "BRF" "CEMIG"
"COPEL" "COSAN" "ELETROBRAS" "EMBRAER" "GERDAU" "GOL" "ITAU" "NATURA" "PÃO DE
AÇÚCAR" "PETROBRAS" "SABESP" "SANTANDER" "SID NACIONAL" "SUZANO" "TELEF BRASIL"
"ULTRAPAR" "VALE"
local company_file_names "AMBEV ASSAI AZUL BRADESCO BRASKEM BRF CEMIG COPEL COSAN
ELETROBRAS EMBRAER GERDAU GOL ITAU NATURA PAO_DE_ACUCAR PETROBRAS SABESP
SANTANDER SID_NACIONAL SUZANO TELEF_BRASIL ULTRAPAR VALE"
local ncompanies : word count `company_ids'
forvalues i = 1/`ncompanies' {
    local company_id : word `i' of `company_ids'
    local company_name : word `i' of `company_names'
    local company_file_name : word `i' of `company_file_names'
    preserve
    keep if company_id == `company_id'
    histogram SD, xscale(range(-1 1)) title(`"Sentimento Diário - `company_name"')
    local filename
    "D:\Beatriz_Stata\Imagens_Dissertação\Histogramas_SD\SD_`company_file_name'.png"
    graph export ``" filename"', as(png) replace
    restore
}

// Volume de Tweets, Picos e Anuncios de Resultados
local company_ids "1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24"

```

```

local company_names ` "AMBEV" "ASSAI" "AZUL" "BRADESCO" "BRASKEM" "BRF" "CEMIG"
"COPEL" "COSAN" "ELETROBRAS" "EMBRAER" "GERDAU" "GOL" "ITAU" "NATURA" "PÃO DE
AÇÚCAR" "PETROBRAS" "SABESP" "SANTANDER" "SID NACIONAL" "SUZANO" "TELEF BRASIL"
"ULTRAPAR" "VALE"
local company_file_names "AMBEV ASSAI AZUL BRADESCO BRASKEM BRF CEMIG COPEL COSAN
ELETROBRAS EMBRAER GERDAU GOL ITAU NATURA PAO_DE_ACUCAR PETROBRAS SABESP
SANTANDER SID_NACIONAL SUZANO TELEF_BRASIL ULTRAPAR VALE"
local ncompanies : word count `company_ids'
forvalues i = 1/`ncompanies' {
    local company_id : word `i' of `company_ids'
    local company_name : word `i' of `company_names'
    local company_file_name : word `i' of `company_file_names'
    preserve
    twoway ///
        (line Total_Cont date if company_id == `company_id', lcolor(blue) lwidth(thin)) ///
        (scatter Total_Cont date if is_peak == 1 & company_id == `company_id', mcolor(green)
msize(medium) mlcolor(black) mlwidth(thin) msymbol(triangle)) ///
        (scatter Total_Cont date if dummy_EA == 1 & company_id == `company_id', mcolor(red)
msize(medsmall) mlcolor(black) mlwidth(thin) msymbol(circle)) ///
        , ///
        title(`"'company_name"') ///
        xlabel(#8, format(%tdMon/YY)) ///
        legend(order(1 "Volume de Tweets" 2 "Picos" 3 "Anúncios de Resultados")) ///
        ylabel(0 100) ///
        ytitle("Volume de Tweets") xtitle("Data")
    local filename
"D:\Beatriz_Stata\Imagens_Dissertação\Volume_Tweets_Picos_EA\VT_`company_file_name'.png"
    graph export `"'filename"', as(png) replace
    restore
}

// SD & PICOS DE TWEETS
count if SD < 0 & is_peak == 1
local total_neg = r(N)
count if SD == 0 & is_peak == 1
local total_neu = r(N)
count if SD > 0 & is_peak == 1
local total_pos = r(N)

histogram SD if is_peak == 1, width(0.04) frequency ///
    xtitle("Polaridade do Sentimento") ///
    ytitle("Número de Picos") ///
    xscale(range(-1 1)) ///
    ylabel(0(10)100) ///
    fcolor(blue) lcolor(black) lwidth(vthin) ///
    xline(0, lcolor(red) lwidth(medium)) ///
    text(95 -0.7 "Total Negativo = `total_neg'", size(small)) ///
    text(95 0.7 "Total Positivo = `total_pos'", size(small)) ///
    text(95 0 "Total Neutro = `total_neu'", size(small))
graph export "D:\Beatriz_Stata\Imagens_Dissertação\Distribuicao_Sentimento_Picos.png", as(png)
replace

```

Apêndice I – Rotina STATA 16 – Retornos Anormais Médios

```
clear all
use "D:\Beatriz_Stata\Base_Completa2", clear

//----- AR (-10 A +10) -----//
*-10
gen CAR_10n= .
forvalues i = 1/'= _N' {
    local event_diff = dif['i']
    if dif['i'] < 0 continue
    quietly sum AR if company_id == company_id['i'] & inrange(dif, `event_diff' - 10, `event_diff' - 10)

replace CAR_10n = r(sum) if _n == `i'
}
gen CAR_10 = CAR_10n
drop CAR_10n

*-9
gen CAR_9n= .
forvalues i = 1/'= _N' {
    local event_diff = dif['i']
    if dif['i'] < 0 continue
    quietly sum AR if company_id == company_id['i'] & inrange(dif, `event_diff' - 9, `event_diff' - 9)

replace CAR_9n = r(sum) if _n == `i'
}
gen CAR_9 = CAR_9n / 9
drop CAR_9n

*-8
gen CAR_8n= .
forvalues i = 1/'= _N' {
    local event_diff = dif['i']
    if dif['i'] < 0 continue
    quietly sum AR if company_id == company_id['i'] & inrange(dif, `event_diff' - 8, `event_diff' - 8)
    replace CAR_8n = r(sum) if _n == `i'
}
gen CAR_8 = CAR_8n
drop CAR_8n

*-7
gen CAR_7n= .
forvalues i = 1/'= _N' {
    local event_diff = dif['i']
    if dif['i'] < 0 continue
    quietly sum AR if company_id == company_id['i'] & inrange(dif, `event_diff' - 7, `event_diff' - 7)
    replace CAR_7n = r(sum) if _n == `i'
}
gen CAR_7 = CAR_7n
drop CAR_7n

*-6
gen CAR_6n= .
forvalues i = 1/'= _N' {
    local event_diff = dif['i']
```

```

    if dif[i] < 0 continue
    quietly sum AR if company_id == company_id[i] & inrange(dif, `event_diff' - 6, `event_diff' - 6)
replace CAR_6n = r(sum) if _n == i
}
gen CAR_6 = CAR_6n
drop CAR_6n

*-5
gen CAR_5n = .
forvalues i = 1/'=_N' {
    local event_diff = dif[i]
    if dif[i] < 0 continue
    quietly sum AR if company_id == company_id[i] & inrange(dif, `event_diff' - 5, `event_diff' - 5)
replace CAR_5n = r(sum) if _n == i
}
gen CAR_5 = CAR_5n
drop CAR_5n

*-4
gen CAR_4n = .
forvalues i = 1/'=_N' {
    local event_diff = dif[i]
    if dif[i] < 0 continue
    quietly sum AR if company_id == company_id[i] & inrange(dif, `event_diff' - 4, `event_diff' - 4)
replace CAR_4n = r(sum) if _n == i
}
gen CAR_4 = CAR_4n
drop CAR_4n

*-3
gen CAR_3n = .
forvalues i = 1/'=_N' {
    local event_diff = dif[i]
    if dif[i] < 0 continue
    quietly sum AR if company_id == company_id[i] & inrange(dif, `event_diff' - 3, `event_diff' - 3)
replace CAR_3n = r(sum) if _n == i
}
gen CAR_3 = CAR_3n
drop CAR_3n

*-2
gen CAR_2n = .
forvalues i = 1/'=_N' {
    local event_diff = dif[i]
    if dif[i] < 0 continue
    quietly sum AR if company_id == company_id[i] & inrange(dif, `event_diff' - 2, `event_diff' - 2)
replace CAR_2n = r(sum) if _n == i
}
gen CAR_2 = CAR_2n
drop CAR_2n

*-1
gen CAR_1n = .
forvalues i = 1/'=_N' {
    local event_diff = dif[i]
    if dif[i] < 0 continue
    quietly sum AR if company_id == company_id[i] & inrange(dif, `event_diff' - 1, `event_diff' - 1)
replace CAR_1n = r(sum) if _n == i
}
gen CAR_1 = CAR_1n

```

```

drop CAR_1n

* CAR para o dia 0
gen CAR_0 = AR

*+1
gen CAR_1p = .
forvalues i = 1/'=_N' {
    local event_diff = dif['i']
    if dif['i'] < 0 continue
    quietly sum AR if company_id == company_id['i'] & inrange(dif, `event_diff' + 1, `event_diff' + 1)
    replace CAR_1p = r(sum) if _n == `i'
}
gen CAR1 = CAR_1p
drop CAR_1p

* +2
gen CAR_2p = .
forvalues i = 1/'=_N' {
    local event_diff = dif['i']
    if dif['i'] < 0 continue
    quietly sum AR if company_id == company_id['i'] & inrange(dif, `event_diff' + 2, `event_diff' + 2)
    replace CAR_2p = r(sum) if _n == `i'
}
gen CAR2 = CAR_2p
drop CAR_2p

* +3
gen CAR_3p = .
forvalues i = 1/'=_N' {
    local event_diff = dif['i']
    if dif['i'] < 0 continue
    quietly sum AR if company_id == company_id['i'] & inrange(dif, `event_diff' + 3, `event_diff' + 3)
    replace CAR_3p = r(sum) if _n == `i'
}
gen CAR3 = CAR_3p
drop CAR_3p

* +4
gen CAR_4p = .
forvalues i = 1/'=_N' {
    local event_diff = dif['i']
    if dif['i'] < 0 continue
    quietly sum AR if company_id == company_id['i'] & inrange(dif, `event_diff' + 4, `event_diff' + 4)
    replace CAR_4p = r(sum) if _n == `i'
}
gen CAR4 = CAR_4p
drop CAR_4p

* +5
gen CAR_5p = .
forvalues i = 1/'=_N' {
    local event_diff = dif['i']
    if dif['i'] < 0 continue
    quietly sum AR if company_id == company_id['i'] & inrange(dif, `event_diff' + 5, `event_diff' + 5)
    replace CAR_5p = r(sum) if _n == `i'
}
gen CAR5 = CAR_5p

```

```

drop CAR_5p

* +6
gen CAR_6p = .
forvalues i = 1/'_N' {
    local event_diff = dif[`i']
    if dif[`i'] < 0 continue
    quietly sum AR if company_id == company_id[`i'] & inrange(dif, `event_diff' + 6, `event_diff' + 6)
    replace CAR_6p = r(sum) if _n == `i'
}
gen CAR6 = CAR_6p
drop CAR_6p

* +7
gen CAR_7p = .
forvalues i = 1/'_N' {
    local event_diff = dif[`i']
    if dif[`i'] < 0 continue
    quietly sum AR if company_id == company_id[`i'] & inrange(dif, `event_diff' + 7, `event_diff' + 7)
    replace CAR_7p = r(sum) if _n == `i'
}
gen CAR7 = CAR_7p
drop CAR_7p

* +8
gen CAR_8p = .
forvalues i = 1/'_N' {
    local event_diff = dif[`i']
    if dif[`i'] < 0 continue
    quietly sum AR if company_id == company_id[`i'] & inrange(dif, `event_diff' + 8, `event_diff' + 8)
    replace CAR_8p = r(sum) if _n == `i'
}
gen CAR8 = CAR_8p
drop CAR_8p

* +9
gen CAR_9p = .
forvalues i = 1/'_N' {
    local event_diff = dif[`i']
    if dif[`i'] < 0 continue
    quietly sum AR if company_id == company_id[`i'] & inrange(dif, `event_diff' + 9, `event_diff' + 9)
    replace CAR_9p = r(sum) if _n == `i'
}
gen CAR9 = CAR_9p
drop CAR_9p

* +10
gen CAR_10p = .
forvalues i = 1/'_N' {
    local event_diff = dif[`i']
    if dif[`i'] < 0 continue
    quietly sum AR if company_id == company_id[`i'] & inrange(dif, `event_diff'+10, `event_diff' + 10)
    replace CAR_10p = r(sum) if _n == `i'
}
gen CAR10 = CAR_10p
drop CAR_10p

gen polaridade_AR = ""
replace polaridade_AR = "NEUTRO" if label_T_AR == "NS" & dif > -20
replace polaridade_AR = "POSITIVO" if label_T_AR != "NS" & AR > 0 & dif > -20

```

```

replace polaridade_AR = "NEGATIVO" if label_T_AR != "NS" & AR < 0 & dif>-20

drop if year == 2024
drop if year == 2021

egen std_AR = sd(AR) // .02289873
egen Média_AR = mean(AR) // 5.22e-06

save "D:\Beatriz_Stata\Base_Completa3", replace

//-----AR MÉDIO e CAR -----//

use "D:\Beatriz_Stata\Base_Completa3", clear
edit

replace dummy_EA = 0 if dummy_EA ==.

// GERAL

// RETORNOS ANORMAIS MÉDIOS (AR Médios)

* keep if is_peak == 1
* keep if is_EA_peak == 1
* keep if non_EA_peak == 1
* keep if dummy_EA == 1
* keep if Rating == "Positivo"
* keep if Rating == "Negativo"
* keep if Rating == "Neutro"
keep if SD !=.

collapse (mean) CAR_10 (mean) CAR_9 (mean) CAR_8 (mean) CAR_7 (mean) CAR_6 (mean) CAR_5 (mean)
CAR_4 CAR_3 ///
              (mean) CAR_2 (mean) CAR_1 (mean) CAR_0 (mean) CAR1 (mean) CAR2 (mean) CAR3
(mean) CAR4 (mean) CAR5 ///
              (mean) CAR6 (mean) CAR7 (mean) CAR8 (mean) CAR9 (mean) CAR10, by(polaridade_AR)

xpose, clear varname
ren v1 negativo
ren v2 neutro
ren v3 positivo
drop if _varname == "polaridade_AR"

gen window = _n - 11

gen T_ARneg = (negativo - 0.00000522) / 0.02289873
gen label_negativo = " "
    replace label_negativo = "NS" if abs(T_ARneg) < 1.64
    replace label_negativo = "*" if abs(T_ARneg) >= 1.64
    replace label_negativo = "***" if abs(T_ARneg) >= 1.95
    replace label_negativo = "****" if abs(T_ARneg) >= 2.58

gen T_ARneu = (neutro - 0.00000522) / 0.02289873
gen label_neutro = " "
    replace label_neutro = "NS" if abs(T_ARneu) < 1.64
    replace label_neutro = "*" if abs(T_ARneu) >= 1.64
    replace label_neutro = "***" if abs(T_ARneu) >= 1.95
    replace label_neutro = "****" if abs(T_ARneu) >= 2.58

gen T_ARpos = (positivo - 0.00000522) / 0.02289873
gen label_positivo = " "

```

```

replace label_positivo = "NS" if abs(T_ARpos) < 1.64
replace label_positivo = "*" if abs(T_ARpos) >= 1.64
replace label_positivo = "***" if abs(T_ARpos) >= 1.95
replace label_positivo = "****" if abs(T_ARpos) >= 2.58

twoway ///
  (line positivo window, lcolor(blue) lpattern(solid)) ///
  (line neutro window, lcolor(midblue) lpattern(solid)) ///
  (line negativo window, lcolor(dknavy) lpattern(solid)) ///
  (scatter neutro window if label_neutro == "*" | label_neutro == "***" | label_neutro == "****", mcolor(red)
msymbol(circle) msize(medium)) ///
  (scatter positivo window if label_positivo == "*" | label_positivo == "***" | label_positivo == "****",
mcolor(red) msymbol(circle) msize(medium)) ///
  (scatter negativo window if label_negativo == "*" | label_negativo == "***" | label_negativo == "****",
mcolor(red) msymbol(circle) msize(medium)), ///
  title("AR Médios - Geral (Dia com Postagem)") ///
  xtitle("Janela de Eventos") ytitle("Retornos Anormais Médios") ///
  legend(order(1 "positivo" 2 "neutro" 3 "negativo" 4 "Sig até 10%")) ///
  position(2) ring(0) size(small) cols(1) keygap(2)) ///
  graphregion(color(white))
graph export
"D:\Beatriz_Stata\Imagens_Dissertação\AR_CAR_Médios\AR_Médios_Geral\AR_Médio_Geral_(Dia com
Postagem).png", as(png) replace

// RETORNOS ANORMAIS MÉDIOS ACUMULADOS (CAR Médios)
gen CAR_neg = sum(negativo)
gen CAR_neu = sum(neutro)
gen CAR_pos = sum(positivo)

gen T_CARneg = (CAR_neg - 0.00000522) / 0.02289873
gen label_Cnegativo = " "
  replace label_Cnegativo = "NS" if abs(T_CARneg) < 1.64
  replace label_Cnegativo = "*" if abs(T_CARneg) >= 1.64
  replace label_Cnegativo = "***" if abs(T_CARneg) >= 1.95
  replace label_Cnegativo = "****" if abs(T_CARneg) >= 2.58

gen T_CARneu = (CAR_neu - 0.00000522) / 0.02289873
gen label_Cneutro = " "
  replace label_Cneutro = "NS" if abs(T_CARneu) < 1.64
  replace label_Cneutro = "*" if abs(T_CARneu) >= 1.64
  replace label_Cneutro = "***" if abs(T_CARneu) >= 1.95
  replace label_Cneutro = "****" if abs(T_CARneu) >= 2.58

gen T_CARpos = (CAR_pos - 0.00000522) / 0.02289873
gen label_Cpositivo = " "
  replace label_Cpositivo = "NS" if abs(T_CARpos) < 1.64
  replace label_Cpositivo = "*" if abs(T_CARpos) >= 1.64
  replace label_Cpositivo = "***" if abs(T_CARpos) >= 1.95
  replace label_Cpositivo = "****" if abs(T_CARpos) >= 2.58

twoway ///
  (line CAR_pos window, lcolor(blue) lpattern(solid)) ///
  (line CAR_neu window, lcolor(midblue) lpattern(solid)) ///
  (line CAR_neg window, lcolor(dknavy) lpattern(solid)) ///
  (scatter CAR_neu window if label_Cneutro == "*" | label_Cneutro == "***" | label_Cneutro == "****",
mcolor(red) msymbol(circle) msize(medium)) ///
  (scatter CAR_pos window if label_Cpositivo == "*" | label_Cpositivo == "***" | label_Cpositivo ==
"****", mcolor(red) msymbol(circle) msize(medium)) ///

```

```

(scatter CAR_neg window if label_Cnegativo == "*" | label_Cnegativo == "***" | label_Cnegativo == "***",
mcolor(red) msymbol(circle) msize(medium)), ///
title("CAR Médios - Geral (Dia com Postagem)") ///
xtitle("Janela de Eventos") ytitle("Retornos Anormais Médios Acumulados") ///
legend(order(1 "positivo" 2 "neutro" 3 "negativo" 4 "Sig até 10%") ///
size(small) cols(2) keygap(2)) ///
graphregion(color(white))
graph export
"D:\Beatriz_Stata\Imagens_Dissertação\AR_CAR_Médios\CAR_Médios_Geral\CAR_Médio_Geral_(Dia com
Postagem).png", as(png) replace

```

Apêndice J – Rotina STATA 16 – Correlação e Granger

```
clear all
use "D:\Beatriz_Stata\Base_Completa3.dta", clear
edit

//-----//

// CORRELAÇÃO DE PEARSON
    pwcorr SD stock_return, sig
    pwcorr Total_Cont stock_return, sig
    pwcorr SD q_neg, sig
    pwcorr Total_Cont q_neg, sig

    bysort company_id: pwcorr SD stock_return, sig
    bysort company_id: pwcorr Total_Cont stock_return, sig
    bysort company_id: pwcorr SD q_neg, sig
    bysort company_id: pwcorr Total_Cont q_neg, sig

//-----//

// CAUSALIDADE DE GRANGER

// 1 teste de Dickey-Fuller aumentado (ADF) dentro do xtunitroot para séries temporais em painel
    xtunitroot fisher SD, dfuller lags(1) drift // estacionária
    xtunitroot fisher stock_return, dfuller lags(1) drift // estacionária
    xtunitroot fisher Total_Cont, dfuller lags(1) drift // estacionária
    xtunitroot fisher q_neg, dfuller lags(1) drift // estacionária

// testando cada empresa

//Sentimento Diário e Retorno das ações
levelsof company_id, local(ids)
foreach id in `ids' {
    preserve
    keep if company_id == `id'
    display "Rodando para empresa: " `id'
    var SD stock_return, lags(4)
    vargranger
    restore
}

//Volume de Tweets e Retorno das ações
levelsof company_id, local(ids)
foreach id in `ids' {
    preserve
    keep if company_id == `id'
    display "Rodando para empresa: " `id'
    var Total_Cont stock_return, lags(4)
    vargranger
    restore
}

//Sentimento Diário e Quantidade de Negócios
levelsof company_id, local(ids)
foreach id in `ids' {
    preserve
    keep if company_id == `id'
```

```

        display "Rodando para empresa: " `id'
        var SD q_neg, lags(4)
        vargranger
        restore
    }

//Volume de Tweets e Quantidade de Negócios
levelsof company_id, local(ids)
foreach id in `ids' {
    preserve
    keep if company_id == `id'
    display "Rodando para empresa: " `id'
    var Total_Cont q_neg, lags(4)
    vargranger
    restore
}

```

ANEXOS

Anexo 1 – Ferramenta simplificada de coleta dos *tweets* via *Twscrape*

```
#Importação das dependencias e inicializacao da api.
import asyncio
from twscrape import API, gather
from twscrape.logger import set_log_level
import pandas as pd

api = API()

#Teste e/ou login das contas.
await api.pool.stats()
#await api.pool.accounts_info()
#await api.pool.login_all()

#Definindo strings de busca
twitter_data = []

empresa_string = '( "GRUPO NATURA" OR "NATURA &CO" OR "NTCO3" )'
nome_empresa = 'natura'

#Definindo filtros
string_search = f'{empresa_string} lang:pt until:2023-12-31 since:2023-01-01 -
filter:links -filter:replies'
response = await gather(api.search(string_search, limit=0))
for tweet in response:
    twitter_data.append(tweet.dict())

#Exportando
df = pd.DataFrame(twitter_data)
df.to_csv(f"{nome_empresa}.csv")
```

Anexo 2 – Ferramenta de atribuição do sentimento via *Google Natural Language API*

```
# Instalar as bibliotecas
!pip install pandas
!pip install openpyxl
!pip install seaborn
!pip install matplotlib
!pip install google-cloud-bigquery
!pip install google-cloud-language

# Importar as bibliotecas
import time
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from google.cloud import bigquery
from google.cloud import language_v1
from google.oauth2 import service_account

# Carregar o arquivo xlsx
file_name = "C:/Users/beatr/OneDrive/Documentos/sentimento_tweets.xlsx"

# Ler o arquivo xlsx
df = pd.read_excel(file_name, engine='openpyxl')
df.head()

credentials = service_account.Credentials.from_service_account_file(
    "C:/Users/beatr/OneDrive/Documentos/sentimento_tweets\beatriz-xxxxxx.json"
)

# Verificar se tem conteúdo em formatos que não seja texto
nan_rows = df[df['conteudo'].isna()]
print(nan_rows)

# Utilizando a API
score_list = []
magnitude_list = []

client = language_v1.LanguageServiceClient(credentials=credentials)

for index, row in df.iterrows():
    text = row['conteudo']
    time.sleep(0.05)

    document = language_v1.Document(
        content=text,
        type_=language_v1.Document.Type.PLAIN_TEXT,
        language='pt'
    )
    sentiment = client.analyze_sentiment(
        request={"document": document}
    )

    score_list.append(sentiment.document_sentiment.score)
    magnitude_list.append(sentiment.document_sentiment.magnitude)
```

```

# Criar as colunas
df['Score'] = score_list
df['Magnitude'] = magnitude_list

# Criando as faixas de valores da coluna de Score
conditionlist = [
    (df['Score'] > 0),
    (df['Score'] == 0),
    (df['Score'] < 0)
]

choicelist = ['Positivo', 'Neutro', 'Negativo']
df['score_clase'] = np.select(conditionlist, choicelist, default='Not
Specified')

# Mostrar a tabela
print(df)

# Criar o nome do arquivo de saída
output_file_name = file_name.replace('.xlsx', '_sentimento_analisado.xlsx')

# Exportar o DataFrame para um arquivo Excel
df.to_excel(output_file_name, index=False, engine='openpyxl')

# Baixar arquivo
import shutil
download_path = "C:/Users/beatr/Downloads"
shutil.copy(output_file_name, download_path)

```