

A INFLUÊNCIA DA MÍDIA SOCIAL NO VALOR DA PETROBRAS

SOCIAL MEDIA INFLUENCE ON PETROBRAS' VALUE

O artigo foi aprovado e apresentado no XXI "USP International Conference in Accounting", realizado de 28/07 a 31/07 de 2021, em São Paulo (SP)"

RESUMO

A presente pesquisa visa contribuir com a discussão sobre o sentimento do investidor. O objetivo do trabalho foi analisar a relação entre o sentimento exposto no Twitter e o retorno da Petrobras. Desse modo, *tweets* sobre a Petrobras, no período de 2010 até 2020, foram analisados a fim de se verificar se o sentimento textual dessas mensagens impacta no retorno. O método escolhido para classificar as palavras foi o mapeamento perceptual conjuntamente com a média e desvio padrão da frequência dos termos em dias positivos ou negativos. Os resultados apontaram que a variação do sentimento total e a variação do sentimento negativo tem relação com o retorno. A variação do sentimento é significativa quando o mercado está pessimista, demonstrando que a variação negativa do sentimento intensifica o movimento pessimista do mercado e ameniza a queda do mercado quando a variação do sentimento é positiva. Os achados corroboram com a teoria de finanças comportamentais de que o sentimento possui relacionamento com o retorno, no caso do estudo, o sentimento textual. Por fim, foram analisadas diferentes amostras com a mesma metodologia. Com isso, foi possível verificar que os indivíduos falam mais sobre a estatal, de modo geral, do que sobre os *tickers* da ação ou com a própria empresa na sua página oficial.

Palavras-Chave: Sentimento Textual; Mídias Sociais; Sentimento do Investidor.

ABSTRACT

This research seeks to contribute to the discussion on investor sentiment. The objective of the research was to analyze the relationship between social media sentiment on Twitter and the return of Petrobras. Thus, tweets about Petrobras, from 2010 to 2020, were analyzed to verify whether the textual sentiment of these messages impacts company return. The method to classify the words was the perceptual mapping together with the mean and standard deviation of the frequency of the terms in positive or negative days. The results showed that the variation of the total sentiment and the negative sentiment are related to the return. The variation in sentiment is significant when the market is bearish, demonstrating that the negative variation in sentiment intensifies the bearish movement in the market, smoothing the market's decline when the variation in sentiment is positive. The findings corroborate the behavioural finance theory that sentiment is related to return, textual feeling in the case of the study. Finally, different samples were analyzed using the same methodology. With this, it was possible to verify that individuals talk more about the state-owned company than about the stock tickers or about the company itself on its official page.

Keywords: Textual Sentiment; Social Media; Investor sentiment.

Milene Dias Almeida

Mestre em Ciências Contábeis pela Universidade Federal do Rio de Janeiro (UFRJ). Especialista em Ciência de Dados e Big Data pela PUC-MG. Graduada em Ciências Econômicas pela Universidade Federal Rural do Rio de Janeiro (UFRRJ). E-mail: milene.dias.a@gmail.com. ORCID: <https://orcid.org/0000-0003-1460-2064>. Lattes: <https://lattes.cnpq.br/2002658090997360>

Vinicius Mothé Maia

Doutor e Mestre em Administração de Empresas pelo IAG/PUC-Rio. Docente do curso de graduação e pós-graduação em Ciências Contábeis na Universidade Federal do Rio de Janeiro (UFRJ). E-mail: vmaia@facc.ufrj.br. ORCID: <https://orcid.org/0000-0001-6156-3104>. Lattes: <http://lattes.cnpq.br/7523431181094741>

Roberto Tommasetti

Pós Doutor (pela Universidade Estadual do Rio de Janeiro), Doutor (pela Universidade Federal do Rio de Janeiro) e Mestre (pela Pontifícia Universidade Católica de São Paulo) em Ciências Contábeis. Docente do curso de graduação e pós-graduação em Ciências Contábeis na Universidade Federal do Rio de Janeiro (UFRJ). Contador CRC-RJ (RJ-127075/0-5). E-mail: roberto.tommasetti@gmail.com. ORCID: <https://orcid.org/0000-0003-2087-4074>. Lattes: <http://lattes.cnpq.br/0787595610081997>

A INFLUÊNCIA DA MÍDIA SOCIAL NO VALOR DA PETROBRAS

1. INTRODUÇÃO

A teoria das finanças comportamentais sugere que os sentimentos irracionais dos investidores impulsionam uma parcela dos preços dos ativos (McGurk, Nowak & Hall, 2020). As emoções e humores dos indivíduos afetam seus processos de tomadas de decisão, levando-os a uma relação entre sentimento e comportamento do mercado acionário (Nofsinger, 2010).

De acordo com Marschner & Ceretta (2021), as pesquisas nessa área ainda são recentes e incipientes no Brasil, sem esclarecer o que afeta o sentimento do investidor brasileiro e quais são os seus veículos.

A análise de sentimento pode ser realizada ao extrair emoções e sentimentos de informações textuais no conteúdo das mídias sociais. Diariamente, as pessoas escrevem *posts*, comentários, críticas e *tweets* sobre variados tópicos. Com isso, é possível verificar se produtos, marcas e pessoas são visualizados positiva ou negativamente na internet (Godsay, 2015). A mídia social é um importante meio de compartilhar as emoções e impactar na opinião pública geral, inclusive, nos mercados (Naylor, Lambertson, & West, 2012).

A mídia social passou a integrar o leque das tradicionais fontes de informação dos investidores, como: jornais, televisão e demonstrações contábeis. Os investidores passaram a recorrer a plataformas como Twitter e Stocktwit como meio de informação, pode-se assim presumir que a mídia social possa influenciar as escolhas do investidor (Sprenger, Sandner, Tumasjan, & Welpe, 2014).

A literatura nacional tem documentado evidências sobre o sentimento do investidor e sua relação com o risco e retorno das ações (Piccoli, da Costa Jr, da Silva, & Cruz, 2018), gerenciamento de resultados (Santana, Santos, Carvalho & Martinez, 2019) e efeito disposição (Prates, da Costa & Santos et al., 2019), porém algumas perguntas relevantes ainda não estão respondidas: quanto o comportamento do investidor brasileiro é afetado pelo sentimento textual? Este impacto é coerente em casos de textos negativos e positivos e em dias negativos e positivos de bolsa valores?

Considerando essa lacuna teórica, esta pesquisa analisou qual o impacto do sentimento textual expresso no Twitter no retorno da Petrobras? Com isso, o objetivo do estudo foi de analisar a relação entre o sentimento exposto no Twitter e o retorno das ações.

Para alcançar este objetivo optou-se por observar ao longo de um período extenso (10 anos no caso) uma única empresa (Petrobras no caso) tanto em termos de retorno acionário quanto de sentimento textual na mídia social.

Quanto à empresa, escolheu-se a Petrobras para participar do estudo de caso porque é uma empresa de referência nacional e está entre as marcas que mais foi percebida por alguma publicidade, por notícias ou propagandas “boca a boca” na internet, no ano de 2019, em relação ao ano anterior. A instituição foi a que teve o maior salto de popularidade (“Buzz Rankings de 2019: Brasil | YouGov - BrandIndex”, 2020). Além disso, a empresa teve uma liderança isolada por quatro anos, de 2016 até 2019, no Top of Mind, na categoria “Marca que representa o Brasil” (Pinto, 2020).

A seleção da empresa também se deu pelo fato de que o nome “Petrobras” não é homônimo, auxiliando na obtenção dos dados. Caso a escolha fosse por outra empresa, como a Oi, todos os *tweets* com a palavra “oi” seriam coletadas. Com isso, poderiam surgir mensagens que não são relacionadas com a operadora, e sim, com uma saudação.

Quanto à escolha da rede social, optou-se para Twitter, pois nele há uma circulação mais aberta de informações, já que os perfis são públicos, permitindo também uma coleta de dados mais acessíveis e, ao mesmo tempo, essa ferramenta valoriza a expressão do indivíduo, o modo como ele avalia e se posiciona diante dos acontecimentos e de informações publicadas (Silva, Maia, Haveroth & Tommasetti, 2021).

O diferencial da pesquisa foi a criação de uma lista de termos/palavras, a partir das publicações no Twitter, de modo a construir um novo dicionário de palavras em língua portuguesa associadas ao sentimento, criando assim um dicionário de palavras personalizado para a organização. Os dicionários existentes são genéricos e com menor qualidade. Segundo Li, Wu e Wang (2020), uma lista de palavras gerais não é aplicável a textos financeiros, pois um dicionário específico modela melhor do que um adaptado de outras áreas ou do que um genérico. Os termos usados para expressar opiniões são geralmente específicos e altamente correlacionados a um determinado domínio (Cruz, Ochoa, Rocha & Poncellet, 2017) e uma simples tradução de uma lista de conteúdo específico em outra língua poderia levar a um resultado de análise de sentimento errôneo (Banner, Pauls, & Walter, 2019).

2. FUNDAMENTAÇÃO TEÓRICA

2.1 Sentimento Textual do Investidor

O sentimento dos investidores pode promover desvios no valor de um ativo que, em teoria, deveria ser igual ao seu valor intrínseco. A teoria comportamental analisa os sentimentos que levam os investidores a tomar decisões distantes das ótimas, contrapondo a teoria de mercados eficientes. Uma mudança no sentimento do investidor pode gerar consequências de eventos, como novos padrões de compras (Baker & Wurgler, 2006).

O sentimento do investidor está embutido nas opiniões e nos discursos não verbais. Nesse sentido, é o sentimento que dá o tom mais otimista (positivo) ou pessimista (negativo) a um texto opinativo. A análise linguística das palavras pode captar o sentimento otimista ou pessimista de quem as escreveu (Beattie, 2014).

O sentimento da mídia social é importante preditor na determinação da avaliação do indivíduo, mas nem todas as mensagens de mídia social têm o mesmo impacto. O grupo de ativos no Twitter é a minoria. Entretanto, acontece uma influência desproporcional na população. Ainda assim, a minoria posta cada vez mais, pois assim possui maior probabilidade de se tornar um líder. No Twitter, as informações são diárias, visíveis e acessíveis, e os investidores são mais propensos a ler e a responder a essa mídia social. Esse efeito também acontece no mercado financeiro, que pode levar ao comportamento de manada, quando os indivíduos do grupo reagem todos da mesma forma (Mai, Shan, Bai, Wang & Chiang, 2018).

Sprenger *et al.* (2014) estudaram os fóruns de investimento do Twitter como fonte de informação alternativa do varejo e criou um mecanismo que exemplifica como funciona a demanda de informações. O mecanismo está exposto na Figura 1, reelaborada a partir desse estudo citado anteriormente com o trabalho de Santos, Laender e Pereira (2015).

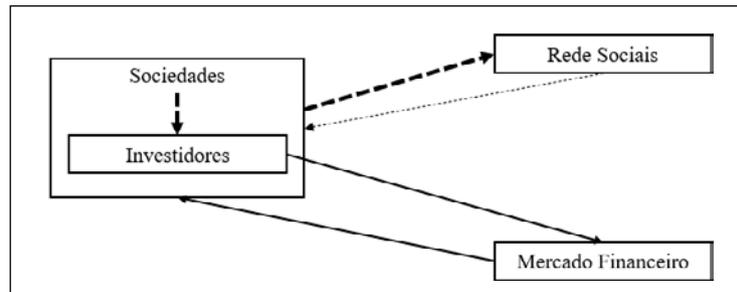


Figura 1. Mecanismos de transmissão – demanda de informações

Fonte: Elaboração própria baseada nos artigos de Sprenger *et al.* (2014) e de Santos *et al.* (2015)

É possível perceber, pela Figura 1, que os fóruns de investimento presentes nas mídias sociais se conectam com os investidores com uma discussão sobre ações e mercados. O mecanismo começa com os investidores e a sociedade postando suas emoções nas mídias sociais. Os investidores buscam informações para fazer sua tomada de decisão, no mecanismo de pesquisa do Google ou em fóruns de investimento em mídias sociais. Como a mídia social não é a única fonte de informação dos investidores, pode-se presumir um impacto, pelo menos parcial, da demanda de informações (Sprenger *et al.*, 2014) representado pela seta pontilhada dos investidores às mídias sociais mutuamente.

Por outro lado, o mercado define o valor do preço das ações e índice pela negociação dos investidores de forma agregada. Isso é indicado pela seta sólida. Por fim, o sentimento adquirido influencia na tomada de decisão do investidor, que se reflete no preço das ações (Sprenger *et al.*, 2014). Contudo, os movimentos das ações, eventos e notícias também conseguem impactar o Twitter e gerar picos de postagens, e a frequência das postagens acompanha o início das negociações e se mantém por três horas (Santos *et al.*, 2015).

2.2 Estudos anteriores e formulação das hipóteses

Pesquisas anteriores verificaram ser importante analisar o sentimento do investidor exposto no Twitter, para tentar prever o valor futuro das ações das empresas, mesmo quando em ambiente de assimetria informacional (Bartov, Faurel, & Mohanram, 2018). O sentimento agregado expresso em *tweets* podem transmitir informações sobre como estão avaliando as empresas (McGurk *et al.*, 2020; Carosia, Coelho, & Silva, 2019; Ranco, Aleksovski, Caldarelli, Grčar, & Mozetič, 2015) e ter capacidade de previsão (Souza, Lucena & Queiroz, 2019; Bartov *et al.*, 2018; Fang, Liang, Zou & Huang, 2017; Ito, Sakaji, Izumi, Tsubouchi & Yamashita, 2017).

Contas de Twitter de pessoas importantes para a sociedade também foram investigadas. Foi verificado que os *tweets* do então Presidente dos Estados Unidos Trump possuem influência no mercado acionário americano, aumentando assim o volume e a volatilidade dos preços, além de atrair maior atenção do investidor. No entanto, esse efeito se dissipa em poucos dias (Ge, Kurov & Wolfe, 2018). No Brasil, foi investigado o mesmo efeito no setor bancário para o Presidente da República, Bolsonaro. Nesse caso, como resultado, foi verificado que as postagens do Presidente no Twitter não possuem impacto no retorno ou no volume negociados das ações de empresas do setor bancário. Os autores do estudo argumentam que esse resultado foi provavelmente devido ao fato do setor bancário ser muito concentrado, pelo pouco tempo de governo do Presidente e pelo conteúdo das postagens, que podem ser interpretadas mais voltadas para o meio político do que ao mercado (Tommasetti, Maia, & Macedo, 2021).

O indicador de sentimento textual pode ser desenvolvido por análise morfológica das palavras (Ito *et al.*, 2017), por qualidade textual (Pengnate & Riggins, 2020) e por estatísticas (Alkubaisi, Kamaruddin, & Husni, 2018). As palavras podem ser classificadas como positivas, negativas ou neutras (Ranco *et al.*, 2015). A abordagem mais comum tem sido empregar o modelo de regressão linear em dados de séries temporais, que abrange o sentimento geral no nível do mercado e o desempenho do índice de ações (Kearney & Liu, 2014).

Diferentes dicionários de palavras financeiras foram comparados ao longo dos anos. Entre os dicionários de sentimento, o dicionário de sentimento específico do domínio de finanças modela melhor os sentimentos e possui uma relação mais forte com o comportamento do preço das ações (Li *et al.*, 2020; Mangee, 2018; Loughran & McDonald, 2015).

Listas foram criadas com base no dicionário Loughran-McDonald Financial, sendo que, adaptaram a lista ao idioma do texto em análise (Banner *et al.*, 2019) e outros trabalhos, em língua inglesa, aplicaram o mesmo dicionário, mas para amostras diferentes, como em documentos financeiros (Gandhi, Loughran, & McDonald, 2019; Hájek, 2018), em mensagens de mídias sociais (Houlihan & Creamer, 2017) e até em documentos que a análise textual conseguiu prever falência de bancos (Gupta, Simaan, & Zaki, 2017).

O modelo geral do estudo analisou a influência do sentimento textual do Twitter no retorno da Petrobras, tendo em vista que o retorno do ativo possui influência do retorno do mercado e do sentimento textual como foi visto nos estudos citados anteriormente. Dessa forma, têm-se as hipóteses de pesquisa:

H_1 : A variação do sentimento textual se relaciona positivamente com o retorno;

H_2 : As variações negativas do sentimento textual possuem maior impacto no retorno do que as variações positivas; e,

H_3 : A variação do sentimento textual impacta o movimento de mercado em relação ao retorno.

A segunda hipótese foi construída seguindo o conceito da teoria dos prospectos. Nesse teorema, os indivíduos com riscos e benefícios (ganho ou perda) iguais, tem comportamentos diferentes no processo de tomada de decisão. Com isso, surge o conceito de aversão à perda. Os indivíduos sentem muito mais a dor de uma perda do que o prazer de um ganho. Há uma assimetria entre ganhos e perdas. A função de utilidade, nesse modelo, demonstra que uma perda gera mais impacto do que um ganho na mesma magnitude (Kahneman & Tversky, 1979). Por isso, a variação do sentimento foi dividida em positiva e negativa. Esperou-se que as variações negativas do sentimento possuíssem maior peso do que variações positivas do sentimento no retorno.

Aprofundando o tema, na terceira hipótese foi investigada se a variação do sentimento textual possuiu algum efeito na relação entre retorno de mercado e retorno da Petrobras.

Este estudo preenche a lacuna teórica a respeito de quanto o comportamento do investidor brasileiro é afetado pelo sentimento textual e se este impacto é coerente em casos de textos negativos e positivos e em dias negativos e positivos de bolsa valores.

3. METODOLOGIA

A presente pesquisa visou criar uma lista de termos/palavras, a partir das publicações no Twitter, de modo a construir um novo dicionário de palavras em língua portuguesa associadas ao sentimento. O objetivo foi de criar um dicionário personalizado, que não fosse adaptado de outra língua ou de outra área, mesmo que ampla, pois um dicionário específico modela melhor do que um adaptado de outras áreas ou do que um genérico.

Dados do mercado de capitais e da mídia social foram coletados para entender melhor o comportamento dos investidores e o que os influenciaram nas suas tomadas de decisão. O modelo geral analisou a influência do sentimento textual do Twitter no retorno da Petrobras.

Para alcançar esse objetivo, foram coletados os dados sobre a Petrobras e as mensagens expostas no Twitter entre 2010 e 2020. O histórico da série diária de preços da Petrobras foi coletado na Bloomberg.

De acordo com Almeida *et al.* (2021), a metodologia foi desenvolvida em duas fases: construção da variável variação do sentimento e os testes das hipóteses. A Figura 2 ilustra o fluxo de trabalho utilizado no estudo.

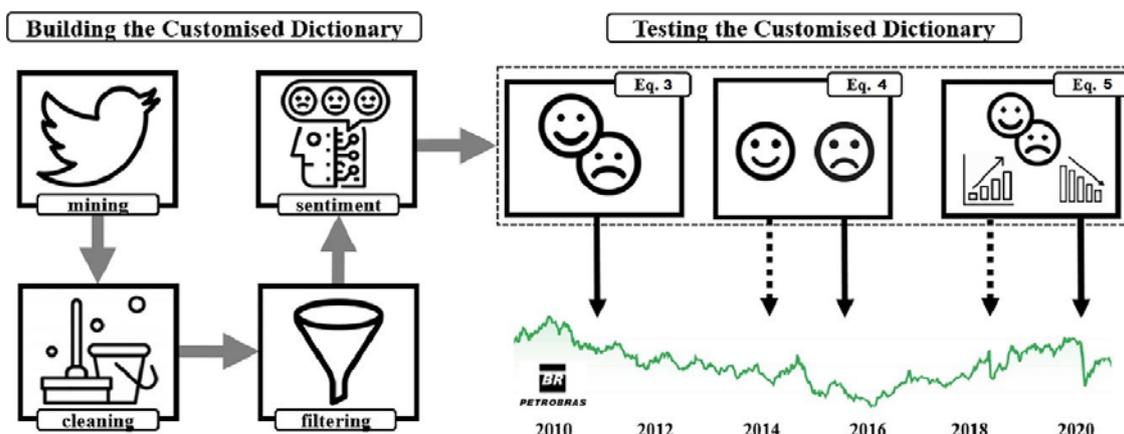


Figura 2 - Etapas para o Processamento da Linguagem Natural usados no estudo.

Fonte: Elaboração própria baseada no artigo de Almeida *et al.* (2021)

Para realizar a análise do sentimento exposto no Twitter, foi necessário coletar as postagens acerca da Petrobras em pré-processamento dos dados. Os *tweets* foram coletados por meio da biblioteca GetOldTweets3, no Python. Foram

selecionadas as postagens em língua portuguesa, que continham o nome da instituição (Petrobras), os *tickers* da ação (PETR3 e PETR4) e as respostas às postagens da empresa na sua conta oficial (@petrobras), entre 2010 e junho de 2020.

A escolha do banco de dados foi feita com o propósito de avaliar qual informação seria mais importante para o investidor: sobre a Petrobras, no geral; sobre a ação da instituição; ou sobre as respostas direcionadas à empresa através de sua conta oficial no Twitter. Esse parâmetro para captar determinados *tweets* também foi utilizado nos estudos de Souza *et al.* (2019), em que coletaram postagens, em português ou inglês, que continham nomes das instituições participantes do índice IBOVESPA e/ou termos escolhidos pelos autores: IBOVESPA, IBOV, Brasil, Bolsa, Balcão e Bolsa de Valores Brasileira.

A construção do indicador de sentimento parte da mineração das mensagens postadas nas redes sociais. Como os textos são longos, foi necessário realizar a *tokenização*, que converte cada *tweet* em um vetor de palavras (“unigrama”) (Li *et al.*, 2020). Em seguida, foi necessário limpar o banco de dados removendo *stopwords*, palavras que não possuem valor semântico. Essa mineração de dados foi construída com algoritmos em Python e, contendo esse tratamento de limpeza nos dados, a biblioteca usada foi a *tm – Text Mining*.

A fase subsequente foi descartar as palavras neutras. Estas aparecem diariamente e na mesma frequência, independentemente da variação diária do preço das ações da empresa. O filtro utilizado foi o de palavras com diferença média e de desvio padrão entre dias de retorno positivo e negativo inferiores a 20%.

Posteriormente, as palavras filtradas foram inseridas no mapa perceptual. O dimensionamento multidimensional é um método para visualizar o grau de dissimilaridade entre os objetos. Com base nestes coeficientes de diferença, uma série de distâncias euclidianas é calculada. Os dados são geralmente plotados em dois espaços cartesianos multidimensionais (Gower & Legendre, 1986). Preliminarmente, é necessário calcular a semelhança, por meio do conjunto de objetos, para construir um mapa perceptivo. Suponha que as palavras “a” e “b” recebam valores semelhantes em comparação com outros objetos possíveis (“c”, “d”, “e”). Nesse caso, a escala multidimensional atribui uma distância euclidiana menor entre “a” e “b” do que entre os outros objetos.

A lista de palavras do estudo foi posicionada no mapa. Cada vetor do mapa é representado por uma palavra dessa lista. Os eixos representam os polos positivos e negativos, para que, quanto mais extrema a palavra for, mais positiva ou negativa ela será. Com isso, foi realizada a classificação de polaridade, conforme o eixo que o termo se encontra.

A abordagem de classificação utilizada por essa metodologia é o chamado “bag of words”, que separa os termos do dicionário em positivos e negativos. Assim, as palavras recebem escores numéricos, de acordo com sua polaridade, sendo positivo (+1) e negativo (-1) (Jiang, Lee, Martin, & Zhou, 2019).

Essa lista de palavras positivas e negativas relacionadas à empresa consistiu no dicionário personalizado. O valor de um sentimento de *tweet* foi estimado com base no peso das palavras positivas e negativas. A medida é o total de palavras positivas menos negativas, como segue na Equação 1.

$$STO_t = \sum PP_t + \sum PN_t \quad (1)$$

Onde:

STO_t : Sentimento Total dos *Tweets* j;

PP_t : Quantidade de palavras positivas (Sua polaridade é “+”);

PN_t : Quantidade de palavras negativas (Sua polaridade é “-”).

Visando realizar um teste de robustez da variável sentimento, foi construída a segunda variável de sentimento diário, sendo essa a média de sentimento dos *tweets* em determinado dia. Essa decisão teve o objetivo de investigar se haveria diferença nos resultados, caso a variável de sentimento fosse construída de diferentes formas. Foi feito um teste de robustez da classificação de polaridade dos termos do estudo. A medida é o sentimento médio, calculado como divisão do sentimento total pelo total de palavras. O mesmo indicador foi utilizado por Feuerriegel, Wolff e Neumann (2015). Essa rotulação foi realizada a partir do resultado da Equação 2.

$$SME_t = \frac{STO_t}{\sum PT_t} \quad (2)$$

Onde:

SME_t : Sentimento Médio do *Tweet* j;

PT_t : Quantidade de palavras totais.

A primeira diferença foi usada nas duas variáveis de sentimento (ΔSTO e ΔSME) para capturar a mudança no sentimento global em vez do status absoluto (positivo ou negativo). Por exemplo, se o índice passar de 0,90 para 0,75, continua positivo em termos absolutos, mas sua tendência - que é variável de foco - é negativa.

Para calcular o retorno observado no período da Petrobras, foi utilizada a função logarítmica que analisa o retorno sobre o preço da ação no dia presente e no dia anterior. A partir do modelo de mercado desenvolvido por Sharpe (1963), *Capital Asset Pricing Model* (CAPM), foi construída uma equação para verificar se o retorno da Petrobras está relacionado com o sentimento textual presente em *tweets*, além do movimento do mercado. O retorno do mercado foi o índice IBOVESPA. Com base nisso, a Equação 3 foi estimada para analisar se o sentimento textual se relaciona com o retorno do ativo.

$$Ret_t = \alpha + \beta_1 R_{m,t} + \beta_2 \Delta Sent_t + \varepsilon_t \quad (3)$$

Onde:

Ret_t : Retorno da Petrobras;

$R_{m,t}$: Retorno do mercado;

$\Delta Sent_t$: Variação do Sentimento Textual (STO ou SME);

É esperado que o sinal de β_2 seja positivo. A variação do sentimento impacta o retorno. Se a variação do sentimento for positiva, é gerado um efeito positivo no retorno. No caso de variação negativa do sentimento, o retorno recebe um efeito negativo.

A fim de aprofundar o tema, a variável sentimento foi dividida em sentimento positivo e sentimento negativo. Optou-se, então, por esse caminho para analisar se existe alguma diferença entre o efeito do sentimento separadamente como está presente na Equação 4.

$$Ret_t = \alpha + \beta_1 R_{m,t} + \beta_2 \Delta PST_t + \beta_3 \Delta NST_t + \varepsilon_t \quad (4)$$

Onde:

ΔPST_t : Variação Positiva do Sentimento Textual (STO ou SME);

ΔNST_t : Variação Negativa do Sentimento Textual (STO ou SME);

O comportamento do sentimento textual sobre o movimento do mercado foi dividido em duas variáveis, pois se espera comportamentos diferentes em cada uma delas. Esperou-se encontrar significância e sinal positivo em β_2 e β_3 . Com isso, a variação positiva do sentimento aumenta os retornos e a variação negativa do sentimento diminui os retornos.

Por fim, foi investigado se a variação do sentimento amplificou ou diminuiu os retornos (positivos ou negativos) do mercado. Conforme pode ser visualizado na Equação 5.

$$Ret_t = \alpha + \beta_1 R_{m,t} + \beta_2 \Delta Sent_t + \beta_3 RP_{m,t} \Delta Sent_t + \beta_4 RN_{m,t} \Delta Sent_t + \varepsilon_t \quad (5)$$

Onde:

$RP_{m,t}$: Retorno Positivo do mercado;

$RN_{m,t}$: Retorno Negativo do mercado;

O objetivo foi determinar se, quando o mercado está pessimista, o sentimento positivo pode moderar, suavizando o impacto no retorno da empresa, e o negativo, em exacerbar o efeito sobre os retornos. A expectativa oposta é quando o mercado está otimista, analisar se o sentimento positivo amplifica esse efeito ou se o sentimento negativo modera. Portanto, o sinal esperado das variáveis β_3 é positivo e β_4 é negativo. Assim, a variação do sentimento possui algum efeito no mercado.

O modelo de regressão linear para séries temporais com a estimativa por Mínimos Quadrados Ordinários (MQO) foi realizado para todas as regressões do estudo.

Executaram-se os seguintes testes para identificar possíveis problemas de distribuição residual: teste Jarque-Bera para normalidade, teste Breusch-Pagan para heteroscedasticidade, teste de Fator de Inflação da Variância (VIF) para multicolinearidade e teste Ljung-Box para autocorrelação. Quando um dos testes acima rejeitou a hipótese de distribuição residual, foram aplicados erros padrões consistentes de correção automática (HAC).

4. RESULTADOS

4.1 Análise das Amostras

Para investigar qual seria a amostra mais adequada para o estudo, optou-se por verificar se o impacto da variação do sentimento no retorno ocorreu com postagens coletadas diretamente do perfil da empresa no Twitter (@Petrobras), com alguma citação do termo “petrobras”, ou com a citação dos *tickers* da ação (PETR3 e/ou PETR4). O objetivo foi avaliar qual coleta da amostra seria mais fácil e menos custosa, pois poderiam ser avaliados menos *tweets* para obter o mesmo resultado. Por isso, foi realizada a mesma metodologia, porém, com amostras diferentes. Das três amostras colhidas para o estudo, apenas a que tinha o filtro “petrobras” gerou uma grande lista de termos.

Analisando a amostra, foi possível verificar que essa teve mais *tweets* em relação às outras: 3.826.463 *tweets*, enquanto as outras foram de 139.530 e 210.881. Também apresentou um conteúdo mais diversificado, com mais opiniões e destaques diários. As palavras das menores amostras foram muito frequentes, apareciam tanto em dias positivos quanto em negativos e não apresentaram grandes desvios padrões. Com isso, não seria interessante ter a conta oficial e/ou *tweets* com termos PETR3 e PETR4 como filtro.

Tabela 1 - Quantidade de *tweets*

Ano	Nº de Obs.
2010	147.951
2011	170.844
2012	213.645
2013	216.071
2014	833.109
2015	792.601
2016	350.400
2017	185.125
2018	638.913
2019	207.842
2020*	69.962
2010-2020	3.826.463
Ano	Nº de Obs.

Nota: *dados até junho de 2020

Fonte: Elaboração própria

Na Tabela 1, é possível observar que no ano de 2014, esse que foi um ano de eleição presidencial no Brasil, aconteceu uma alta das citações da empresa, principalmente nos meses de outubro e novembro.

O mesmo ocorreu em 2018. Os anos que tiveram maiores manifestações dos usuários sobre a Petrobras foram 2014 e 2015, respectivamente. No ano de 2014, a Petrobras foi citada pela primeira vez na operação Lava Jato da Polícia Federal. No ano seguinte, os escândalos de corrupção envolvendo o nome da empresa (Otoboni & Freire, 2020) continuaram. Possivelmente, essas foram as causas do aumento nesses anos.

Não há como comprovar que todos os *tweets* que continham o termo “petrobras” foram escritos por investidores.

A relação entre sentimento do investidor e sentimento textual da mídia social é complexa. O último inclui o primeiro. Contudo, é importante ressaltar que as mídias sociais demonstram as condições nas empresas e o mercado na visão da sociedade, incluindo os investidores e influenciam os mesmos.

Para avaliar o efeito do sentimento expresso, via Twitter, sobre os retornos da Petrobras, buscou-se, construir uma lista de termos presentes na rede social que exercem influência no retorno da ação e as estatísticas descritivas das variáveis. Depois da filtragem, a lista de termos foi incluída no mapeamento perceptual.

O mapa perceptivo demonstrou o posicionamento relativo de todas as palavras. Com a classificação da polaridade dos termos concluída, foi realizada a construção das variáveis do sentimento, total e médio.

Vale ressaltar que nessas foram avaliadas as variações e não os valores absolutos diários. Os gráficos da Figura 3

apresentam as oscilações que ocorreram nas variáveis em análise, de 2010 até 2020.

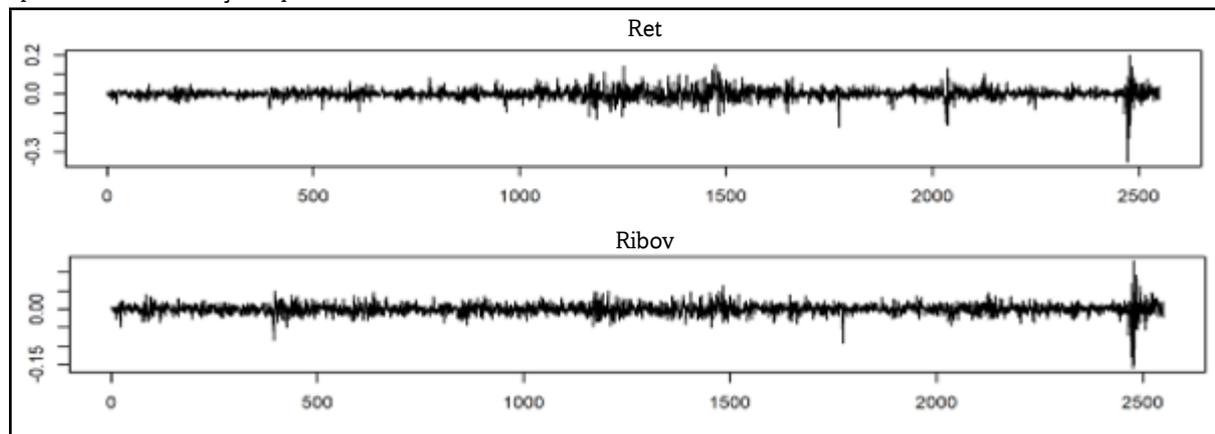


Figura 3 - Oscilações das variáveis ao longo do tempo

Fonte: Elaborado pelos autores a partir de dados da pesquisa.

A partir da Figura 3, é possível verificar que as séries de retornos tiveram momentos de maior e menor oscilação, evidenciando assim o comportamento comum em séries de retorno de formarem agrupamentos de volatilidade, além das oscilações entre as séries terem comportamentos parecidos, evidenciando que os retornos da Petrobras se relacionam em alguma medida aos retornos do mercado, como seria esperado. Na Tabela 2, estão dispostas as estatísticas descritivas de cada variável testada no estudo.

Tabela 2 - Estatística Descritiva do sentimento das variáveis do estudo

Variável	Média	Desvio Padrão	Mínimo	Máximo
ΔSTO	-55.401	891.69	-11491	10366
ΔSME	-0.00033	0.01774	-0.110922	0.07260
Ret	-0.00021	0.03022	-0.35236	0.20067
Ribov	0.00011	0.01612	-0.15993	0.13022

Fonte: Elaborado pelos autores a partir de dados da pesquisa

Nota-se que o valor do retorno do IBOVESPA e da ação da Petrobras tiveram média, respectivamente, positiva e negativa. Apesar de o mercado se manter positivo durante o período, as ações da Petrobras tiveram, em média, durante o período em análise, retornos negativos. As duas variáveis tiveram valor mínimo negativo e um valor máximo de positivo, demonstrando grandes oscilações no retorno do mercado, principalmente no retorno da Petrobras.

4.2 Análise Das Regressões

Após o tratamento dos dados e das variáveis, buscou-se testar as hipóteses do estudo por meio dos modelos de série temporal. Para cada indicador de variação de sentimento, foi realizado o cálculo da mesma equação. Com isso, foi possível a comparação dos resultados e a verificação se o modelo possuiu robustez. Os resultados do primeiro indicador, variação do sentimento total, estão demonstrados na Tabela 3.

Tabela 3 - Impacto da variação do sentimento total no retorno da Petrobras (2,550 Observações)

Regressores	Sinal	Equação 4	Equação 5	Equação 6
<i>Ribov</i>	+	1.4466e+00 *** ($< 2.2e-16$)	1.4435e+00*** ($< 2.2e-16$)	1.3961e+00*** ($< 2.2e-16$)
$\Delta Sent$	+	4.7366e-06 *** (1.642e-06)	-	3.3927e-06* (0.01156)
$\Delta Pos Sent$	+	-	3.4063e-07 (0.71833)	-

Regressores	Sinal	Equação 4	Equação 5	Equação 6
$\Delta Neg Sent$	+	-	7.1590e-06*** (3.554e-08)	-
$RPibov*\Delta Sent$	+	-	-	-1.6263e-04 (0.4407)
$RNibov*\Delta Sent$	-	-	-	-2.9479e-04*** (3.78e-06)
α	na	-1.2957e-04 (0.7122)	1.1165e-03 (0.00592)	1.5154e-04 (0.66403)
R^2 ajustado		0.6334	0.6407	0.6507
Normalidade p-value		< 2.2e-16	< 2.2e-16	< 2.2e-16
Heteroscedasticidade p-value		< 2.2e-16	< 2.2e-16	5.531e-12
Multilinearidade VIF		Ribov:1.007288 Sent: 1.007288	Ribov: 1.07650 $\Delta Pos Sent$:1.011903 $\Delta Neg Sent$: 1.017039	Ribov:1.074142 Sent:2.166817 $RPibov*\Delta Sent$:1.604941 $RNibov*\Delta Sent$:1.659181
Autocorrelação p-value		0.9429	0.7624	0.6099

Nota: Normalidade: teste Jarque-Bera; Heteroscedasticidade: teste Breusch-Pagan. Teste Ljung-Box para autocorrelação (10 lags). Método de estimação: MQO com Autocorrelation-Consistent Standard Errors (HAC). *p < 0.10; ** 0.01 < p < 0.05; *** p < 0.01.

Os resultados demonstraram concordar com a teoria clássica, no que tange o modelo CAPM, e a teoria comportamental em relação ao sentimento textual. O retorno do mercado foi significativo e positivo, da mesma forma explicitada no modelo de Sharpe (1963). O indicador de variação de sentimento textual teve significância e positivo em relação ao retorno. Com isso, foi possível identificar padrões de comportamento nas mídias sociais ligados às oscilações do mercado. O sentimento textual exposto no Twitter influenciou o investidor na sua tomada de decisão. A variação de postagens positivas no Twitter sobre a Petrobras incentivou as pessoas a investirem nessa empresa e, conseqüentemente, os preços das ações aumentaram e quando a variação foi negativa, ocorreu o inverso, o preço da ação diminuiu.

A primeira hipótese da pesquisa foi confirmada: a variação do sentimento textual se relaciona com o retorno. Esses achados corroboram com os resultados dos estudos de Bartov *et al.* (2018), Carosia *et al.* (2019) e McGurk *et al.* (2020), que também verificaram influências do sentimento via Twitter com o retorno.

Na segunda regressão (Equação 5), disposta na Tabela 3, o objetivo foi analisar a segunda hipótese de pesquisa. Buscou-se verificar se a variação positiva e a variação negativa dos sentimentos, separadamente, possuem impacto no retorno. Nesse caso, somente a variação negativa dos sentimentos possuiu impacto negativo no retorno. Esse resultado vai ao encontro à teoria do prospecto, que afirma haver uma assimetria entre ganhos e perdas (Kahneman & Tversky, 1979).

No presente estudo, foi possível verificar que a variação negativa do sentimento possuiu maior impacto nos indivíduos do que a variação positiva do sentimento. As pessoas se sentem mais sensibilizadas pela perda do que pelo ganho, mesmo que seja na mesma magnitude. É um comportamento de defesa ao seu patrimônio. A segunda hipótese de pesquisa foi confirmada, visto que, apenas a variação negativa do sentimento textual demonstrou ter relação com o retorno, indo ao encontro da teoria do prospecto que trata da aversão ao risco dos indivíduos.

Na linha dos resultados da terceira regressão, a variação do sentimento textual total na Equação 6 teve efeito somente com o mercado negativo. O retorno de mercado foi significativo e positivo, da mesma forma que ocorreu nas equações passadas. A variação do sentimento obteve significância e, com efeito positivo em relação ao retorno. Esse achado demonstra que no cenário de o mercado estar em queda, com retornos negativos, e se a variação do sentimento também for negativa, o impacto no retorno da Petrobras será intensificado negativamente. No mesmo cenário, caso a variação do sentimento na rede social seja positiva, o impacto no retorno da estatal será suavizado. Esses resultados concordam com a linha teórica de finanças comportamentais. O sentimento promoveu desvios no valor do ativo. A terceira hipótese da presente pesquisa foi parcialmente confirmada: a variação do sentimento textual impactou o movimento de mercado em relação ao retorno, porém, somente quando o mercado está negativo.

Os resultados também mostram que a proporção da variância na dependente explicada pelas variáveis independentes, representadas por R^2 , aumenta da Equação 4, onde a variação do sentimento total é o regressor observado, para a Equação 5, onde o regressor acima é dividido em variações positivas e negativas do sentimento, e finalmente para a Equação 6, onde a mesma variável interage com dias de mercado favoráveis e desfavoráveis. Esses resultados representam uma confirmação adicional dos comportamentos assimétricos dos investidores nos cenários de ganhos ou perdas, com este último gerando mais impacto do que o primeiro, sendo as pessoas serem avessas ao risco. É um comportamen-

to natural defender seus bens (Jiang *et al.*, 2019).

Visando realizar o teste de robustez do indicador variação do sentimento, os mesmos modelos foram calculados para a variável variação do sentimento médio. Os resultados encontram-se na Tabela 4.

Tabela 4 - Impacto da variação do sentimento médio no retorno da Petrobras (2,550 Observações)

Regressores	Sinal	Equação 4	Equação 5	Equação 6
<i>Ribov</i>	+	1.457e+00 *** ($< 2e-16$)	1.453e+00 *** ($< 2e-16$)	1.432e+00*** ($< 2.e-16$)
$\Delta Sent$	+	1.579e-16 * (0.04854)	-	4.759e-17 (0.2760)
$\Delta Pos Sent$	+	-	-3.436e-17 (0.31805)	-
$\Delta Neg Sent$	+	-	2.036e-16 * (0.02601)	-
<i>RPibov</i> * $\Delta Sent$	+	-	-	-8.054e-15 (0.2569)
<i>RNibov</i> * $\Delta Sent$	-	-	-	-4.911e-15*** ($<2e-16$)
α	na	-2.823e-04 (0.43293)	-1.843e-04 (0.61690)	-2.575e-04 (0.4719)
R ² Ajustado		0.6185	0.6199	0.6268
Normalidade p-value		$< 2.2e-16$	$< 2.2e-16$	$< 2.2e-16$
Heteroscedasticidade p-value		4.12e-10	1.251e-10	0.8977
Multicolinearidade VIF		<i>Ribov</i> :1.00953; <i>Sent</i> : 1.0095	<i>Ribov</i> : 1.011323 $\Delta Pos Sent$:1.0001 $\Delta Neg Sent$: 1.011316	<i>Ribov</i> :1.029258 <i>Sent</i> :1.738738 <i>RPibov</i> * $\Delta Sent$:1.297859 <i>RNibov</i> * $\Delta Sent$:1.465194
Autocorrelação p-value		0.928	0.9002	0.6641

Normalidade: teste Jarque-Bera; Heteroscedasticidade: teste Breusch-Pagan. Teste Ljung-Box para autocorrelação (10 lags). Método de estimação: MQO com Autocorrelation-Consistent Standard Errors (HAC). * $p < 0.10$; ** $0.01 < p < 0.05$; *** $p < 0.01$.

O teste de robustez com a variável de variação do sentimento médio apresentou, em geral, os mesmos resultados que o indicador de variação do sentimento total. A Tabela 4 demonstra resultados com as mesmas significâncias do que a Tabela 3, no geral. A diferença está na Equação 6 das duas tabelas: a variação do sentimento só foi significativa com a construção do sentimento total. Com isso, conclui-se que a construção do dicionário personalizado apresentou significância, independente da forma que a variável sentimento textual foi gerada.

O estudo se mostrou relevante, pois as palavras são componentes importantes dos sistemas de linguagem natural. A análise de sentimentos oferece às organizações uma maneira rápida de monitorar os sentimentos do público em relação à marca, aos negócios, diretores, etc. Isso auxiliará não apenas para atrair novos investidores ou clientes, mas também a manter os atuais.

5. CONCLUSÃO

O objetivo deste trabalho foi analisar a relação entre o sentimento exposto no Twitter e o retorno das ações. Para esse fim, os *tweets* sobre a Petrobras foram analisados, de modo a verificar se o sentimento textual dessas mensagens tem relação com o retorno da empresa.

A pesquisa partiu da premissa que no conteúdo da *web* de determinadas páginas podem ser extraídas informações que denotam emoções e sentimentos. As pessoas escrevem posts, comentários, críticas e *tweets* sobre variados tópicos diariamente (Godsay, 2015). Com isso, é possível verificar se a Petrobras é vista de forma positiva ou negativa na internet.

Todos os indicadores de variação do sentimento tiveram resultados positivos e significativos. A variação do sentimento textual dos *tweets* tem relação direta com o retorno da Petrobras. O resultado corrobora com os estudos de Bartov *et al.* (2018), Carosia *et al.* (2019) e McGurk *et al.* (2020).

A variação negativa do sentimento possuiu maior impacto nos indivíduos do que a variação positiva do sentimento. As pessoas possuem comportamento avesso ao risco, sentem mais pela perda do que pelo ganho, mesmo que seja na mesma magnitude.

Somente a variação do sentimento negativo mostrou resultados negativos e significativos no movimento de mercado. Desse modo, a variação do sentimento negativo potencializa o movimento negativo do mercado e ameniza o movimento positivo do mercado. Esse resultado corrobora com a teoria das finanças comportamentais, que afirma haver uma assimetria entre ganhos e perdas. Uma perda gera mais impacto do que um ganho na mesma magnitude (Kahneman & Tversky, 1979).

A contribuição do estudo foi personalizar o dicionário para organizações individuais, enquanto os dicionários existentes são genéricos e, portanto, com menor qualidade. Dessa forma, o indicador de sentimento particular pode mitigar ou melhorar o movimento de mercado dessa empresa. A desconfiança do mercado sobre alguns aspectos pode diminuir o valor das empresas. Com a construção desses mapas, pode surgir uma poderosa ferramenta para gestão de marketing e análise do sentimento sobre a organização.

Apesar de três amostras terem sido coletadas, optou-se pela que apresentou mais sentimento, diversidade de palavras e quantidade de *tweets*, que foi a que continha apenas a palavra “petrobras” como filtro de seleção. Essa amostra possui uma lista de termos que apresentam desvios relacionados aos retornos diários. Por isso, o presente estudo optou por aprofundar-se na lista de termos formada por essa amostra.

Os achados aqui documentados fornecem contribuições teóricas, gerenciais e sociais, com possível impacto nas áreas de finanças, economia e contabilidade. Os resultados documentados nesta pesquisa contribuem das seguintes formas: (i) ampliam a literatura e ajudam na compreensão teórica dos efeitos do sentimento textual na tomada de decisão do investidor, fenômeno até então pouco explorado no Brasil; (ii) com relação aos estudos anteriores, além de corroborar seus pressupostos, promovem uma melhoria metodológica nas estimativas realizadas ao usar um modelo econométrico capaz de capturar relações pontuais entre o ativo e o sentimento do investidor sobre *aquele* ativo; e (iii) considerando o sentimento um indicador útil para antecipar os rumos da economia (Marschner & Ceretta, 2021), formuladores de políticas e autoridades monetárias podem tomar diferentes medidas em resposta as diferentes mudanças nesse indicador. Todavia, essa é, ainda, uma pesquisa inicial e, havendo um campo promissor a ser explorado. Além da utilização de dados secundários, as abordagens experimentais podem contribuir de forma valiosa, pois ampliam o escopo de variáveis que podem ser consideradas determinantes para compreender o impacto das mídias sociais no sentimento do investidor.

A pesquisa apresentou as limitações de ter coletado somente os dados da mídia social Twitter e da escolha da organização, por não ser um nome homônimo, o que auxiliou na obtenção das informações. Esses resultados são relevantes para os gestores e agentes de políticas públicas, visto que é importante entender a opinião pública sobre a empresa Petrobras, para realizar uma tomada de decisão, assim como para os investidores na alocação dos seus recursos. Os achados demonstraram a relação entre os padrões de comportamento da mídia social e as oscilações do mercado. Para pesquisas futuras, sugere-se mudar a amostra para outras empresas de capital aberto, analisar como o movimento da mídia social Twitter e/ou outras se relacionam com o movimento de mercado e a inclusão de uma ponderação dos termos.

REFERÊNCIAS

- Alkubaisi, G. A. A. J., Kamaruddin, S. S., & Husni, H. (2018). Stock Market Classification Model Using Sentiment Analysis on Twitter Based on Hybrid Naive Bayes Classifiers. *Computer and Information Science*, 11 (1).
- Almeida, M. D., Maia, V. M., Tommasetti, R., & de Oliveira Leite, R. (2021). Sentiment analysis based on a social media customised dictionary. *MethodsX*, 8, 101449.
- Baker, M., & Wurgler, J. (2006) Investor sentiment and the cross-section of stock returns. *Journal of Finance*, 61 (4), 1645–1680.
- Bannier, C., Pauls, T., & Walter, A. (2019) Content analysis of business communication: introducing a German dictionary. *Journal of Business Economics*, 89 (1), 79–123.
- Bartov, E., Faurel, L., & Mohanram, P. S. (2018) Can twitter help predict firm-level earnings and stock returns? *The Accounting Review*, 93 (3), 25–57.
- Beattie, V. (2014). Accounting narratives and the narrative turn in accounting research: Issues, theory, methodology, methods and a research framework. *British Accounting Review*, 46 (2), 111–134.
- Bukovina, J. (2016). Social media big data and capital markets-An overview. *Journal of Behavioral and Experimental Finance*, 11, 18–26.
- Buzz Rankings de 2019: Brasil | YouGov - BrandIndex. (2020, June 13). Retrieved from: <<https://www.brandindex.com/ranking/brazil/2019-buzz>>.
- Carosia, A. E. O., Coelho, G. P., & Silva, A. E. A. (2019). Analyzing the Brazilian Financial Market through Portuguese Sentiment Analysis in Social Media. *Applied Artificial Intelligence*, 34 (1), 1–19.

- Clement, J. (2020, March 15). Countries with most Twitter users 2020. *Statista*. Retrieved from: <<https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/>>.
- Cruz, L., Ochoa, J., Roche, M., Poncelet, P. (2017) Dictionary – Based Sentiment Analysis Applied to Specific Domain. In: Lossio-Ventura J., Alatrística-Salas H. (eds) Information Management and Big Data. SIMBig 2015, SIMBig 2016. Communications in Computer and Information Science, vol 656. Springer, Cham. https://doi.org/10.1007/978-3-319-55209-5_5.
- Silva, S. O., Maia, V. M., Haveroth, J., & Tommasetti, R. (2021). Papel da Auditoria: uma visão a partir do Twitter. *Revista Catarinense da Ciência Contábil*, 20, e3183-e3183.
- Dařena, F., Petrovský, J., Píichystal, J., & Žizka, J. (2018). Machine learning-based analysis of the association between online texts and stock price movements. *Inteligencia Artificial*, 21 (61), 95–110.
- Fang, B., Liang, S., Zou, Q., & Huang, W. (2017). Research on sentiment analysis of financial texts. *7th IEEE International Conference on Electronics Information and Emergency Communication*, (ICEIEC), 248-252.
- Feuerriegel, S., Wolff, G., & Neumann, D. (2015). Information processing of foreign exchange news: Extending the overshooting model to include qualitative information from news sentiment. *SSRN Electronic Journal*.
- Gandhi, P., Loughran, T., & McDonald, B. (2019). Using Annual Report Sentiment as a Proxy for Financial Distress in U.S. Banks. *Journal of Behavioral Finance*, 20 (4), 424–436.
- Ge, Q., Kurov, A., & Wolfe, M. H. (2018). Stock Market Reactions to Presidential Statements: Evidence from Company-Specific Tweets. *Economics Faculty Scholarship*, 29.
- Godsay, M. (2015) The Process of Sentiment Analysis: A Study. *International Journal of Computer Applications*, 126 (7), 26–30.
- Gower, J.C. & Legendre, P. (1986). Metric and Euclidean properties of dissimilarity coefficients. *Journal of Classification*, 3, 5-18.
- Gupta, A., Simaan, M., & Zaki, M. J. (2017). Investigating bank failures using text mining. *2016 IEEE Symposium Series on Computational Intelligence*, SSCI.
- Ito, T., Sakaji, H., Izumi, K., Tsubouchi, K., & Yamashita, T. (2017). Development of an interpretable neural network model for creation of polarity concept dictionaries. *IEEE International Conference on Data Mining Workshops*, ICDMW.
- Jiang, F., Lee, J., Martin, X., & Zhou, G. (2019). Manager Sentiment and Stock Returns. *Journal of Financial Economics*, 132 (1), 126-149.
- Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Journal of the Econometric society*, 47 (2), 263–292.
- Kearney, C., & Liu, S. (2014). Textual sentiment in finance: A survey of methods and models. *International Review of Financial Analysis*, 33, 171–185.
- Li, X., Wu, P., & Wang, W. (2020). Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong. *Information Processing and Management*, 57 (5).
- Loughran, T., & McDonald, B. (2015). The Use of Word Lists in Textual Analysis. *Journal of Behavioral Finance*, 16 (1), 1–11.
- Mai, F., Shan, Z., Bai, Q., Wang, X. S., & Chiang, R. H. L. (2018). How Does Social Media Impact Bitcoin Value? A Test of the Silent Majority Hypothesis. *Journal of Management Information Systems*, 35 (1), 19–52.
- Mangee, N. (2018). Stock Returns and the Tone of Marketplace Information: Does Context Matter? *Journal of Behavioral Finance*, 19 (4), 396–406.
- Marschner, P. F., & Ceretta, P. S. (2021). Sentimento do investidor, incerteza econômica e política monetária no Brasil. *Revista Contabilidade & Finanças*, 32, 528-540.
- Martins, R., Pereira, A., & Benevenuto, F. (2015) An approach to sentiment analysis of web applications in Portuguese. *WebMedia 2015 - Proceedings of the 21st Brazilian Symposium on Multimedia and the Web*, 105–112.
- McGurk, Z., Nowak, A., & Hall, J. C. (2020). Stock returns and investor sentiment: textual analysis and social media. *Journal of Economics and Finance*, 44 (3), 458–485.
- Naylor, R. W., Lamberton, C. P., & West, P. M. (2012). Beyond the “like” button: The impact of mere virtual presence on brand evaluations and purchase intentions in social media settings. *Journal of Marketing*, 76 (6), 105–120.
- Nofsinger, J. (2010). *The Psychology of Investing*. 4. ed. New Jersey: Prentice Hall.
- Otoboni, J., & Freire, D. (2020). Seis anos da Lava Jato: relembre todas as fases da operação. Retrieved from: <<https://www.cnnbrasil.com.br/politica/2020/03/07/seis-anos-da-lava-jato-relembre-todas-as-fases-da-operacao>>.
- Pengnate, S. F., & Riggins, F. J. (2020). The role of emotion in P2P microfinance funding: A sentiment analysis approach. *International Journal of Information Management*, 54.
- Piccoli, P., da Costa Jr, N. C., da Silva, W. V., & Cruz, J. A. (2018). Investor sentiment and the risk–return tradeoff in the Brazilian market. *Accounting & Finance*, 58, 599-618.
- Pinto, A. E. De S. (2020, October 11). Pelo 4o ano seguido, Petrobras é a marca que representa o Brasil. *Folha de S. Paulo*. Retrieved from: <<https://top-of-mind.folha.uol.com.br/2019/10/pelo-4o-ano-seguido-petrobras-e-a-marca-que-representa-o-brasil.shtml>>.
- Prates, W. R., da Costa, N. C., & Santos, A. A. (2019). Efeito disposição: Propensão à venda de investidores individuais e institucionais. *Revista Brasileira de Economia*, 73, 97-119.

- Ranco, G., Aleksovski, D., Caldarelli, G., Grčar, M., & Mozetič, I. (2015). The effects of twitter sentiment on stock price returns. *PLoS ONE*, 10 (9), 1–21.
- Santana, C. V. S., Santos, L. P. G. D., Carvalho, C. V. D. O., & Martinez, A. L. (2019). Sentimento do investidor e gerenciamento de resultados no Brasil. *Revista Contabilidade & Finanças*, 31, 283-301.
- Santos, H. S., Laender, A., & Pereira, A. C. (2015). Uma Visão do Mercado Brasileiro de Ações a partir de Dados do Twitter. *IV Brazilian Workshop on Social Network Analysis and Mining*, 10 (3), 369–394.
- Sharpe, W. F. A. (1963). Simplified Model For Portfolio Analysis. *Magazine Science*, 9 (2), 277–293.
- 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 Congresso USP*, 1–19.
- 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 and Accounting*, 41 (7–8), 791–830.
- Tarde, G. (1902). *Psychologie Économique*. Paris: Bibliothèque de philosophie contemporaine.
- Tommasetti, R., Maia, V. M., & Macedo, M. A. da S. (2021). Twitter as 'bully pulpit': Brazilian banking sector case. *Applied Economics Letters*, 28 (1), 61–64.
- Yadav, A., Jha, C. K., Sharan, A., & Vaish, V. (2020). Sentiment analysis of financial news using unsupervised approach. *Procedia Computer Science*, 167, 589-598.