

Análise exploratória da imagem dos lácteos em tempos de coronavírus*

Kennya Beatriz Siqueira

Thallys Silva Nogueira

Emerson Wendelin Alves Moreira Campos

Nedson Donato Soares

Emerson Augusto Priamo Moraes

Regina Maria Maciel Braga Villela

José Maria Nazar David

Priscila Vanessa Zabala Capriles Goliatt

O mundo está vivendo, no momento, um estado de pandemia, onde milhares de pessoas encontram-se confinadas em suas casas usando a internet e, principalmente, as redes sociais para se comunicar com o resto do mundo. Com isso, a internet tornou-se uma importante ferramenta de análise de comportamento e tendências.

De acordo com pesquisa da TIC Domicílios realizada em 2018, 70% dos brasileiros têm acesso à internet, o que inclui metade da população rural e das classes D e E. Com relação ao tempo que os brasileiros gastam na internet, as redes sociais se destacam, com média de 3 horas e 34 minutos (TIC Domicílios, 2018).

Os conteúdos compartilhados na internet, e em redes sociais, tendem a evidenciar características associadas ao perfil de cada consumidor, realçando seus interesses e opiniões relacionadas a diferentes assuntos. Nesse cenário, o Twitter destaca-se como uma das maiores redes sociais da atualidade, possuindo mais de 300 milhões de usuários ativos mensalmente. O Twitter é uma rede social que oferece aos seus usuários um espaço para conversação, compartilhamento de textos, fotografias e vídeo. De acordo com In Digital (2020) essa é a sexta rede social mais usada no Brasil, com

* Publicado em Revista Indústria de Laticínios

cerca de 48% da população utilizando essa plataforma. O Brasil tem o segundo maior número de usuários da rede, logo atrás dos Estados Unidos, com mais de 27,7 milhões de contas ativas (EMARKETER, 2016).

Segundo Pak & Paroubek (2010) e Araújo *et al.* (2014), analisar os conteúdos compartilhados em redes sociais pode auxiliar no entendimento da opinião das pessoas sobre diferentes assuntos, e, a partir desta análise, dentre diversas outras aplicações possíveis, empresas podem saber mais sobre o que os seus consumidores pensam sobre determinados produtos e serviços (Pushpam & Jayanthi, 2017).

Assim, as redes sociais são no momento, uma alternativa às pesquisas de mercado tradicionais, já que há um maior número de pessoas em casa, se comunicando, gerando maior tráfego de informações na web e principalmente nas redes sociais. De acordo com levantamento da Fundação Dom Cabral (FDC, 2020) realizado nas redes sociais, na última semana, os *posts* relacionados ao coronavírus ultrapassaram 226 mil/dia no Brasil. Isso indica um grande número de pessoas publicando e expressando suas opiniões na internet. Diante disso, torna-se oportuno avaliar esse conteúdo para analisar o que os consumidores de leite e derivados têm expressado no Twitter como forma de identificar comportamentos e tendências de mercado neste período de crise.

De acordo com o último post da ComScore (24 de março de 2020), foi possível observar que grandes marcas de produtos diminuíram suas ações em redes sociais em cerca de 9%, mostrando que as marcas não estão sabendo como engajar o público para os seus produtos durante esse período de pandemia. Mas como será que o mercado do leite e derivados está se saindo diante deste cenário?

O Twitter disponibiliza gratuitamente uma Application Programming Interface (API) para mineração de dados públicos criados por seus usuários, facilitando a obtenção dessas informações. Neste trabalho, foram minerados *posts* através de um *script* na linguagem de programação R (R Core Team, 2013), 5 grupos de palavras-chave combinadas como é mostrado abaixo no período de 20 a 27 de março de 2020:

1. quarentena + {leite}
2. coronavirus + {leite}
3. pandemia + {leite}
4. covid + {leite}
5. isolamento social + {leite}

em que {leite} refere-se aos seguintes termos: creme de leite, doce de leite, iogurte, lactobacilos, leite condensado, leite em pó, leite UHT, leite, manteiga, queijo, requeijão, sorvete, leite longa vida, bebida láctea, leiteinho, *ghee*, leite pasteurizado, leite de saquinho, leite barriga mole, lácteos, laticínios, derivados do leite, derivados lácteos, ricota, mussarela, muçarela, brie, gorgonzola, lactose, leite fermentado, *kefir*. Alguns nomes de queijos também foram inseridos nas buscas que são: provolone, queijo artesanal, queijo canastra, parmesão, *camembert*, *roquefort*, *cottage*, mascarpone, queijo prato, pecorino, *gruyere*, edam, estepe, *emmental*, queijo do reino e gouda.

Além disso, foi empregado o software Orange Canvas desenvolvido na Linguagem Python, aplicando técnicas de análise de sentimentos para analisar os sentimentos associados às postagens no Twitter. O Orange Canvas é um conjunto abrangente de software baseado em componentes para mineração de dados e aprendizagem de máquina, desenvolvido no Laboratório de Bioinformática, da Faculdade de Computação e Ciência da Informação da Universidade de Ljubljana, Eslovénia, juntamente com a comunidade de código aberto (ERJAVEC, 2013). O Orange Canvas trabalha com blocos de construção de fluxos chamados *Widgets* para análise de dados que são montados no ambiente de programação visual do sistema. Os *Widgets* são agrupados em classes de acordo com sua função.

A combinação das palavras visa identificar as opiniões de consumidores que publicam conteúdos lácteos associados à pandemia do novo coronavírus. A Figura 1 apresenta o total de *tweets*. No total foram buscados 14.537 *tweets* no intervalo de 20 a 27 de março de 2020 (dados puros). Realizando a remoção de *tweets* duplicados foi contabilizado um total de 12.185 *tweets* (dados tratados).

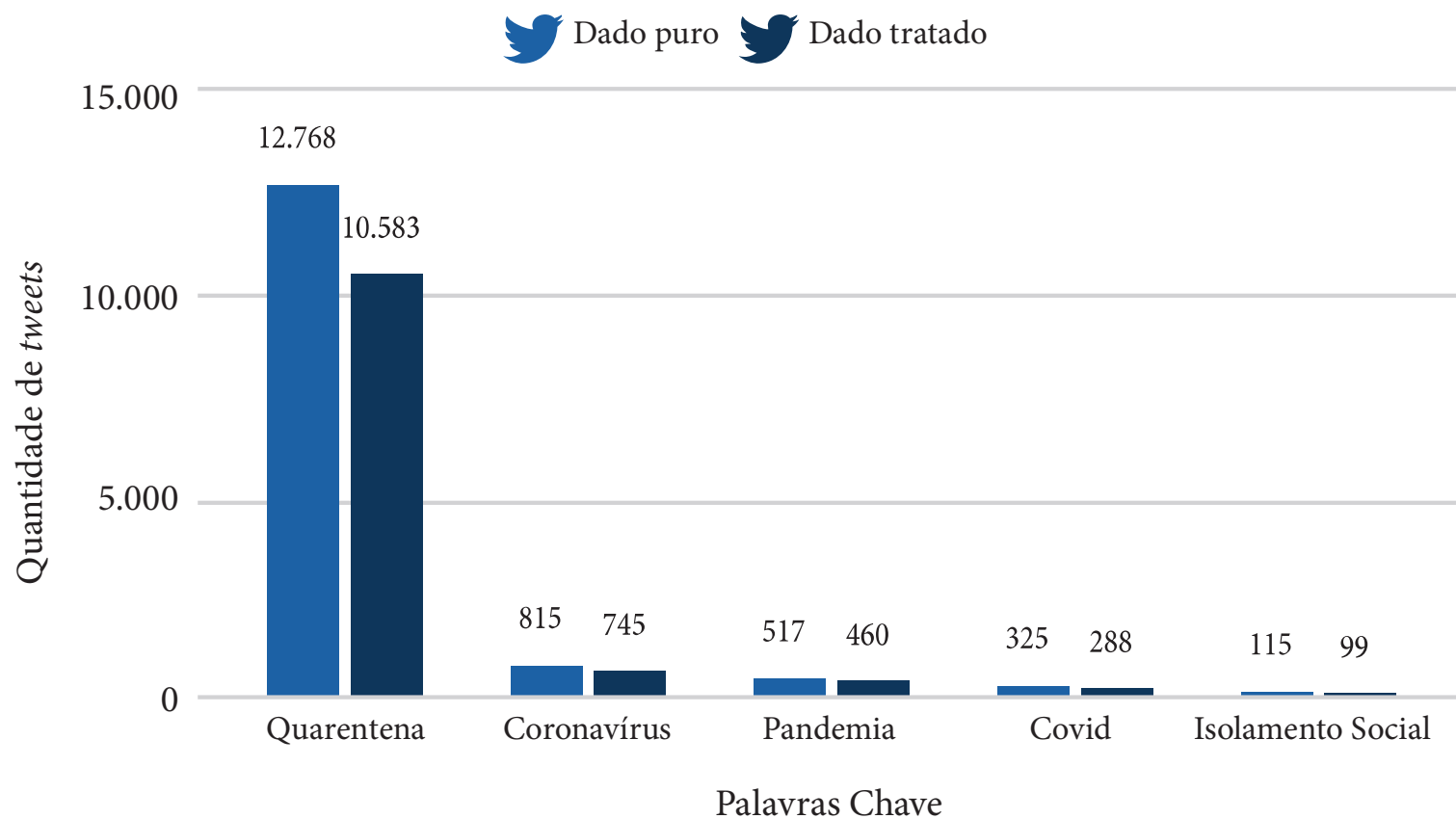


Figura 1. Total de *tweets* minerados no período de 20 a 27 de março de 2020.

Fonte: Elaborado pelos autores.

Pode-se observar que o maior número de resultados foi encontrado para a palavra quarentena associada à lácneos, o que já sugere uma maior preocupação dos brasileiros não apenas com a doença causada pelo coronavírus, mas sim com o isolamento e confinamento em casa. A Figura 2 apresenta os principais derivados lácneos mencionados no Twitter no período analisado.

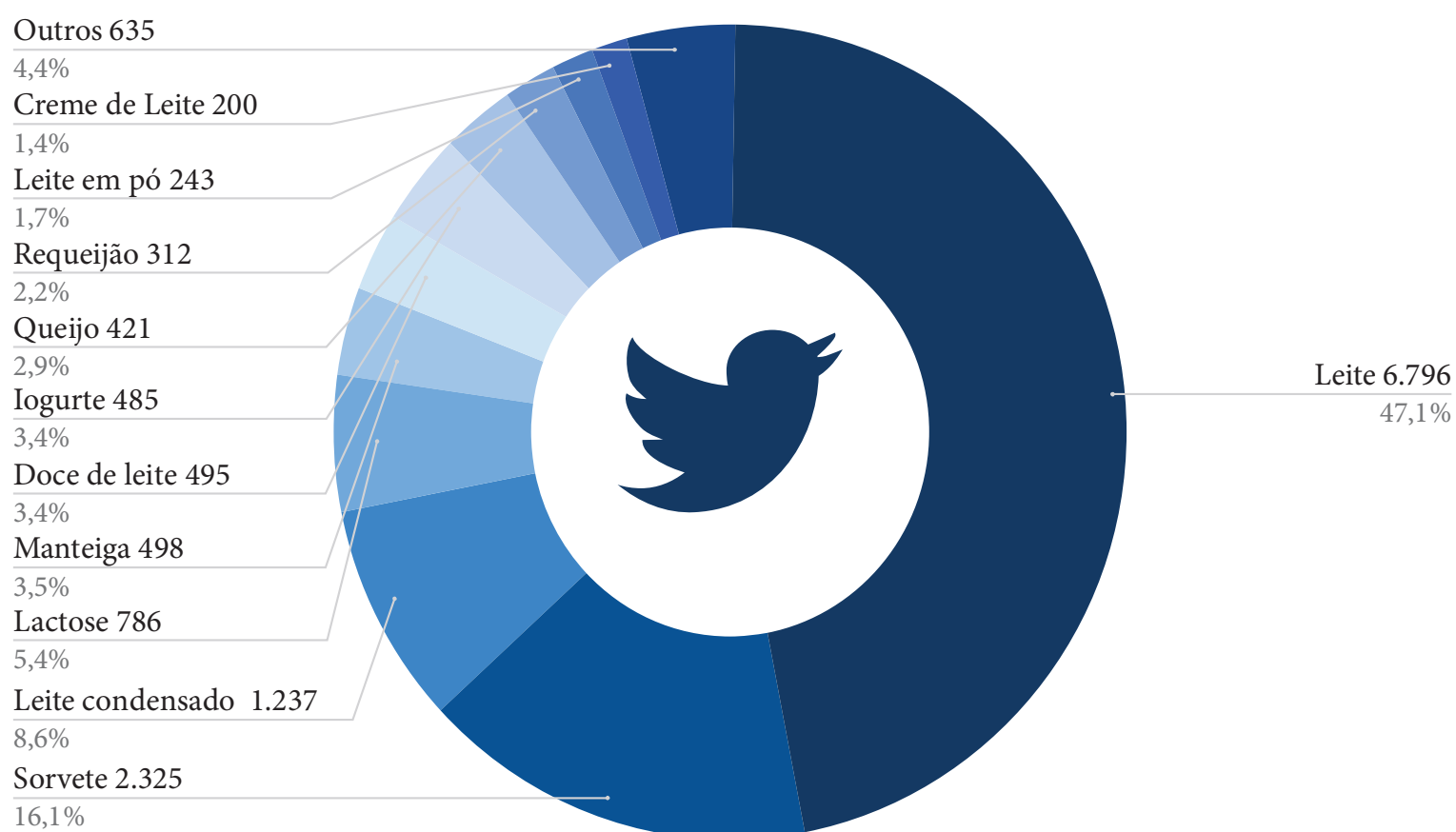


Figura 2. Quantidade de menções dos derivados do leite nos *tweets*.

Fonte: Elaborado pelos autores.

A palavra leite foi retirada da análise por retornar muitos *tweets* com conteúdo não relacionado, o que corrompe os resultados. Assim, os cinco derivados do leite mais comentados em *tweets* relacionados ao coronavírus foram: sorvete, leite condensado, lactose, manteiga e doce de leite.

É interessante notar que três desses derivados são produtos reconhecidamente indulgentes: sorvete, leite condensado e doce de leite. Alimentos indulgentes ou com apelo à indulgência são aqueles consumidos pura e simplesmente por prazer. De acordo com a ciência, isso é natural de ocorrer, pois o hormônio cortisol, conhecido como hormônio do estresse, é liberado em situações de emergência, medo e perigo. Ele gera um aumento de um neurotransmissor chamado NPY que estimula o apetite. Além disso, há a redução de dopamina e serotonina que fornecem a sensação de bem estar e prazer. Assim, a consequência em situações de estresse como a que estamos vivendo no momento é a busca por alimentos que forneçam essa sensação de prazer, ou seja, os alimentos indulgentes.

Com os dados coletados do Twitter foi gerada uma nuvem de palavras mais mencionadas nos *posts* no intervalo de tempo considerado (Figura 3).



Figura 3. Nuvem de palavras gerada com base nos *tweets* analisados.

Fonte: Elaborado pelos autores.

A nuvem de palavras evidencia novamente muitas palavras relacionadas ao momento de estresse e à busca por prazer. Além dos lácteos indulgentes, a nuvem de palavras apresenta outros produtos com apelo à indulgência: brigadeiro, pipoca, Nescau, pastel, biscoito. Além disso, é possível notar que, ao falar sobre lácteos e coronavírus, poucos consumidores consideram os nutrientes e benefícios que os derivados do leite podem trazer para o sistema imune. Ou seja, tudo indica que neste primeiro momento, o estresse está falando mais alto do que a preocupação com a saúde e a busca por lácteos tem sido mais para atender ao prazer do que à prevenção e aumento da imunidade. A análise de sentimentos realizada também corrobora com esse resultado (Figura 4).

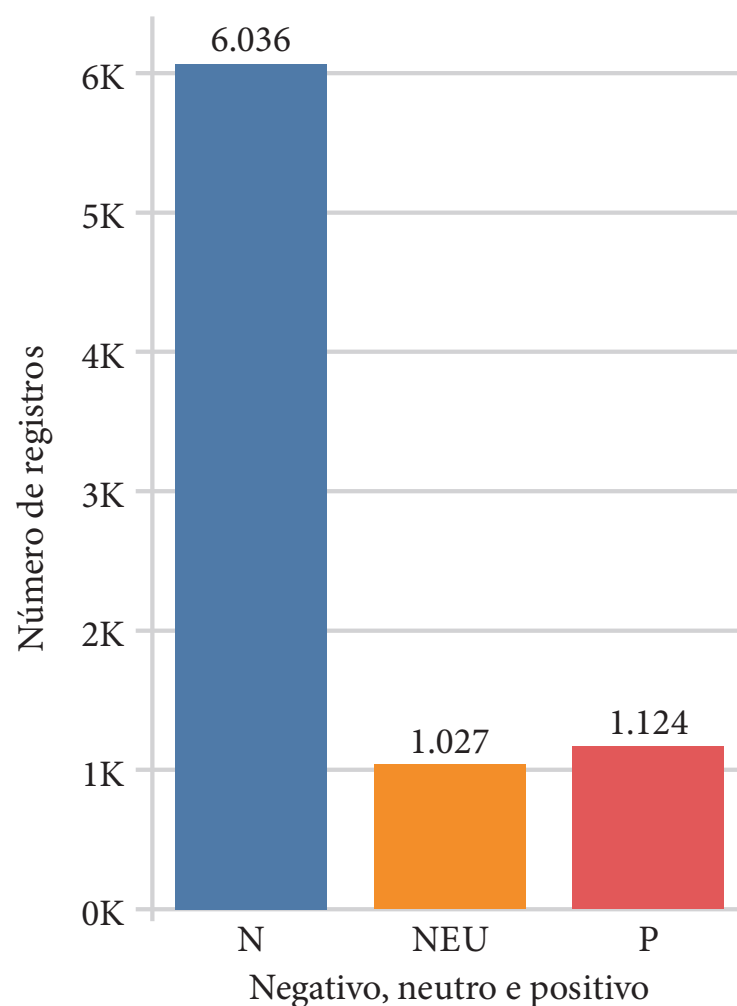


Figura 4. Gráfico de polaridade com a distribuição dos sentimentos Negativo, Neutro e Positivo.

Fonte: Elaborado pelos autores.

A Figura 4 evidencia que o conteúdo postado sobre lácteos e a pandemia reflete muito o lado negativo da situação, o que, reforça os resultados da situação de estresse e busca por satisfação e prazer.

Portanto, pode-se concluir que, num primeiro momento de isolamento social e expansão dos casos de novo coronavírus no Brasil, os lácteos mais procurados e desejados são aqueles que têm o poder de acalmar e tranquilizar o organismo humano. Muito mais do que nutrientes, saudabilidade e prevenção, os brasileiros, inicialmente, estão precisando de alimentos para aliviar o estresse provocado por toda essa turbulência. No entanto, mesmo os alimentos indulgentes não estão intactos à crise. Esse estudo reflete apenas o comportamento do consumidor nos primeiros dias de contato com a pandemia. Após os ânimos se acalmarem, a busca por saudabilidade que já era tendência antes da crise, pode ganhar força, especialmente se isso for bem administrado pelos agentes do setor.

Referências Bibliográficas

ARAÚJO, M., GONÇALVES, P., CHA, M., & BENEVENUTO, F. iFeel. In: International Conference on World Wide Web - WWW, 23., 2014. **Proceedings...** ACM Press, 2014. 14 Companion.

EMARKETER. **Twitter's User Base to Grow by Double Digits This Year**. 22 julho 2016. Disponível em: <<https://www.emarketer.com/Article/Twitter-User-Base-Grow-by-Double-Digits-This-Year/1014243>>. Acesso em: 18 fev. 2020.

ERJAYEC, A. Orange: data mining toolbox in Python; **Journal of Machine Learning Research**, p. 2349-2353, 2013.

FDC - Fundação Dom Cabral. **Report Diário Covid-19**. Março 2020.

PAK, A., & PAROUBEK, P. Twitter as a Corpus for Sentiment Analysis and Opinion Mining. In: International Conference Language Resources and Evaluation (LREC'10), 7., 2010. **Proceedings...** Valleta, Malta, , 2010.

PUSHPAM, C. A., & JAYANTHI, J. G. Overview on Data Mining in Social Media. **International Journal of Computer Sciences and Engineering**, v. 5, p. 147-157, 11 de 2017.

R CORE TEAM 2013. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>. 2013.

RELATÓRIO IN DIGITAL 2020. Disponível em: <<https://datareportal.com/reports/digital-2020-brazil>>. Acesso em: 26 fev. 2020.

TIC DOMICÍLIOS 2018. **Pesquisa sobre o uso das tecnologias de informação e comunicação nos domicílios brasileiros**. São Paulo : Comitê Gestor da Internet no Brasil, 2019. PDF.