

## OBSERVATÓRIO DO CONSUMIDOR: POTENCIALIZANDO A INDÚSTRIA LÁCTEA COM BUSINESS INTELLIGENCE

### “OBSERVATÓRIO DO CONSUMIDOR”: EMPOWERING THE DAIRY INDUSTRY WITH BUSINESS INTELLIGENCE

Thallys da Silva Nogueira<sup>1</sup>, Kennya Beatriz Siqueira<sup>2</sup>, Priscila Vanessa Zabala Capriles Goliatt<sup>1</sup>

<sup>1</sup>Programa de Pós-Graduação em Modelagem Computacional, Universidade Federal de Juiz de Fora

<sup>2</sup>Empresa Brasileira de Pesquisa Agropecuária, Embrapa Gado de Leite

#### Resumo

Em um mercado dinâmico e competitivo, a indústria de laticínios enfrenta o desafio de acompanhar as demandas e tendências do consumidor. A falta de dados atualizados sobre o consumo, no entanto, pode dificultar a tomada de decisões estratégicas. Para superar essa barreira, surge o Observatório do Consumidor (OC), uma ferramenta inovadora que oferece informações e análises sobre o consumo de produtos lácteos no Brasil. O OC utiliza dados de redes sociais para extrair informações estratégicas, como padrões de consumo, preferências regionais e o sentimento do consumidor após a compra/consumo. Com essas informações detalhadas, as empresas podem tomar decisões mais assertivas, fortalecer suas estratégias de marketing e desenvolver novos produtos que atendam às necessidades do mercado, impulsionando um crescimento sustentável e eficaz.

**Palavras-chave** Business Intelligence, Análise de Dados, Mercado Consumidor

#### Introdução

A indústria de laticínios enfrenta desafios e oportunidades em um mercado competitivo e em constante evolução. Em um ambiente altamente dinâmico como este, compreender profundamente o consumidor é crucial para prever e seguir as tendências emergentes. No entanto, no Brasil, há uma carência de dados de consumo atualizados para auxiliar as empresas na tomada de decisão. A ausência de dados atualizados sobre o consumo de alimentos é uma lacuna preocupante, especialmente considerando que a última Pesquisa de Orçamentos Familiares (POF) do IBGE (2024) data de 2017-2018.

A POF é a pesquisa oficial mais detalhada sobre consumo no Brasil, mas a defasagem de seus dados dificulta uma análise precisa do mercado atual. Outra pesquisa oficial do mesmo órgão, a Pesquisa Industrial Anual (PIA), também apresenta uma defasagem de dois anos em suas atualizações. Essa falta de informações recentes limita a capacidade das empresas de laticínios de se adaptarem rapidamente às mudanças nas preferências dos consumidores e às condições do mercado.

Para preencher essa lacuna, surge o Observatório do Consumidor (OC), uma plataforma inovadora baseada nos conceitos de *Business Intelligence* (BI). Desenvolvida em parceria entre a Empresa Brasileira de Pesquisa Agropecuária - EMBRAPA Gado de Leite, a Universidade Federal de Juiz de Fora (UFJF) e o Instituto Federal de Educação, Ciência e Tecnologia do Sudeste de Minas - Campus Juiz de Fora (IF-Sudeste/MG), o OC visa oferecer informações dinâmicas e atualizadas sobre o comportamento do consumidor e as tendências de mercado do setor lácteo brasileiro (Nogueira, 2022).

O OC coleta, armazena, processa e analisa dados das redes sociais X (antigo Twitter) e YouTube, além de utilizar ferramentas de análise de tendências como o Google Trends. Para extrair essas informações de forma eficiente, o OC emprega técnicas avançadas de processamento de linguagem natural (PLN), mineração de dados e inteligência artificial. Essas técnicas são essenciais

para traduzir a linguagem humana em dados compreensíveis por máquinas, facilitando a extração eficiente de informações (Covington, 1994; Russell, 1995).

Reconhecendo a importância da informação para decisões mais direcionadas, o objetivo deste trabalho é mostrar o potencial do sistema de BI do Observatório do Consumidor para delinear o perfil dos consumidores de queijos (um dos derivados lácteos mais consumidos no País), identificando as características mais marcantes desses consumidores no Brasil. Este estudo permite uma visão geral sobre o interesse dos brasileiros nos queijos, além de fornecer informações detalhadas sobre o produto de maior interesse, demonstrando como o OC pode auxiliar a indústria de laticínios a alcançar seus objetivos, utilizando uma ferramenta de BI que proporciona percepções valiosas para a tomada de decisões estratégicas.

## Material e Métodos

Para viabilizar todo o procedimento metodológico do OC, três componentes principais são essenciais: a infraestrutura computacional, as fontes de dados e os resultados alcançados pela ferramenta. A Figura 1 resume os componentes que sustentam o OC.

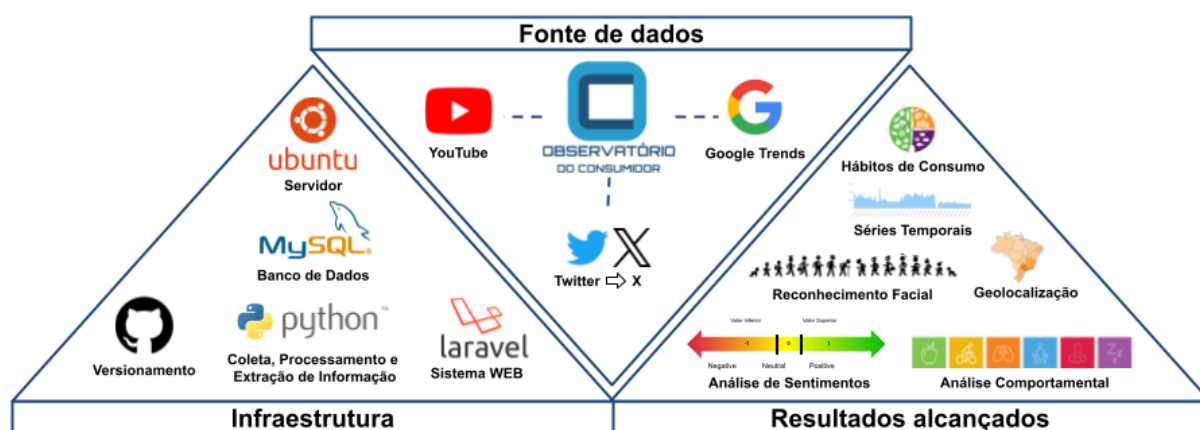


Figura 1 - Componentes do Observatório do Consumidor.  
Fonte: autores.

A infraestrutura computacional utilizada pela ferramenta para implementar o processo metodológico do sistema de BI conta com diversas tecnologias avançadas. A primeira tecnologia é o banco de dados. Utilizamos o MySQL para armazenamento eficiente dos resultados. Cada rede social utilizada como fonte de dados possui suas respectivas tabelas, garantindo a organização e acessibilidade das informações coletadas.

Com o banco de dados estruturado, a extração de informações é realizada utilizando a linguagem de programação Python. Escolhemos Python por sua ampla utilização em aplicações de ciência de dados e pela sinergia com diversas análises que realizamos. Python oferece uma vasta gama de bibliotecas e ferramentas que facilitam o processamento e a análise de grandes volumes de dados, incluindo aqueles provenientes de redes sociais.

Para a etapa de visualização dos resultados, desenvolvemos um sistema computacional utilizando o *framework* Laravel, escrito em PHP. Essa ferramenta oferece numerosos recursos que facilitam a criação de aplicações web robustas e escaláveis, permitindo uma apresentação clara e interativa dos dados analisados. Todos os códigos desenvolvidos são versionados no GitHub, garantindo controle de versão e colaboração eficiente entre os desenvolvedores. Para disponibilizar o sistema online, utilizou-se um servidor Ubuntu, que oferece um ambiente estável e seguro para a hospedagem da aplicação.

Como brevemente mencionado, o OC utiliza duas redes sociais, X e YouTube, além do Google Trends como fontes de dados. O Google Trends é integrado ao sistema do OC por meio de *widgets*, permitindo que os usuários explorem as tendências temporais relacionadas a dez categorias de

produtos lácteos<sup>1</sup>. Após a coleta, armazenamento, processamento e extração de informações, o OC apresenta os resultados das análises realizadas no sistema computacional desenvolvido<sup>2</sup>.

Com a infraestrutura do OC, seis perguntas essenciais foram desenvolvidas neste trabalho para delinear o perfil dos consumidores de queijos no Brasil, visando construir uma persona detalhada. As perguntas são: (i) quem são os consumidores de queijo?; (ii) quando ocorreu o consumo desses produtos?; (iii) qual foi o sentimento relacionado esse consumo?; (iv) quais tipos de queijos de maior destaque?; (v) quais os principais acompanhamentos?; e (vi) em quais regiões o queijo foi mais consumido?

## Resultados e Discussão

Desde o início do projeto, o OC já coletou, armazenou e processou mais de 28.465.878 publicações nas redes sociais X e YouTube. Especificamente em relação aos dados sobre queijos, sendo o foco deste estudo de caso, foram identificadas 8.711.860 publicações, representando 30,6% dos dados totais sobre lácteos. Este trabalho concentra-se exclusivamente nas publicações relacionadas ao consumo de queijo.

Para analisar o consumo de queijo no Brasil, aplicou-se uma análise de dependência sintática aos dados. Essa técnica estabelece relações entre as palavras em uma frase, representando como elas dependem umas das outras. Cada relação de dependência conecta uma palavra principal (cabeça) a uma palavra dependente. Deste modo foram mapeadas todas as ocorrências de verbos associados ao consumo de alimentos ligados aos nomes dos queijos (substantivos) estudados pelo OC. Após esse processamento, o universo de dados para todas as análises subsequentes consistiu em 230.939 postagens.

Para identificar o perfil do consumidor de queijo, o OC aplicou análise facial e mineração de dados por meio de casamento de padrões, extraindo informações como sexo, geração e renda dos usuários que publicaram sobre consumo de queijos. Foi constatado que 74,7% das postagens sobre o consumo de queijos foram feitas por homens, enquanto 25,3% foram feitas por mulheres. Entre os homens, 71,3% pertencem à geração Y, nascidos entre 1981 e 1995, enquanto, entre as mulheres, essa porcentagem foi de 77,2%. Em relação à renda, para ambos os sexos, a classe de renda predominante foi a classe D, representando 41,7% dos consumidores.

Para analisar a evolução do consumo, o OC realizou uma análise temporal das menções a verbos relacionados ao consumo de alimentos, mapeando as ações mais comuns dos indivíduos. A Figura 2 ilustra a evolução do consumo ao longo do tempo.

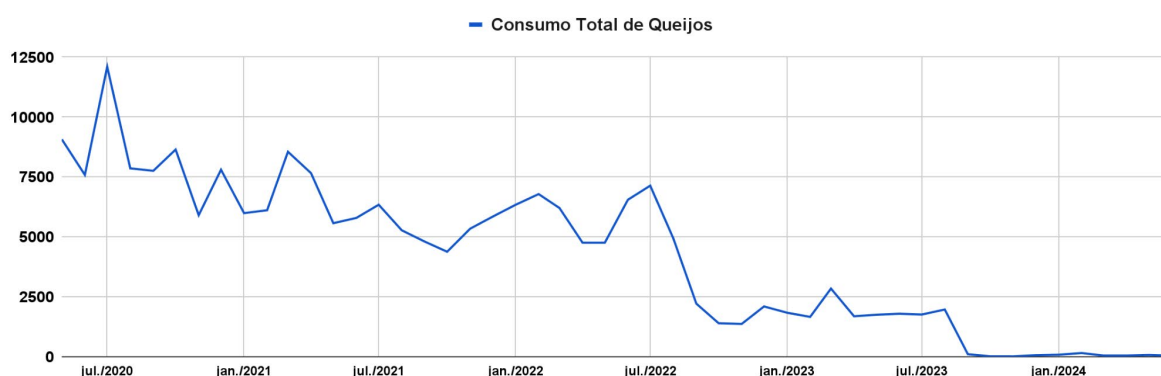


Figura 2 - Resultado da análise temporal das publicações sobre o consumo de queijos.

Fonte: OC/Elaborado pelos autores.

De modo geral, é possível observar picos de publicações sobre queijos no mês de julho, indicando um aumento de consumo neste mês. O período de maior consumo de queijos ocorreu

<sup>1</sup> Bebidas Lácteas, Creme de Leite, Doce de Leite, Iogurte, Leite, Leite Condensado, Leite Fermentado, Manteiga, Queijos e Sorvete.

<sup>2</sup> O sistema pode ser acessado no endereço eletrônico <https://observatorioconsumidor.cnpq.embrapa.br>.

durante a pandemia, quando as pessoas estavam em casa devido às medidas de isolamento social. No entanto, é importante salientar que, embora as postagens sobre queijos estejam se reduzindo, mudanças na política de compartilhamento de dados da rede social X podem ter impactado a coleta de dados pelo OC, afetando consequentemente a análise temporal.

No que diz respeito à avaliação do consumo, o OC empregou a análise de sentimentos (Nogueira, et al., 2024) para classificar as publicações sobre o consumo como negativas, neutras ou positivas. Os resultados revelaram que 78,5% das publicações após o consumo demonstraram sentimento positivo, indicando uma elevada satisfação com o queijo.

Ao analisar os tipos de queijos mais mencionados, o queijo coalho foi o mais frequentemente citado, com 6.125 menções no período analisado. Este tipo de queijo é popular em várias regiões do Brasil devido à sua versatilidade e sabor característico, frequentemente utilizado em churrascos e pratos típicos nordestinos. Em segundo lugar, o requeijão aparece com 4.694 menções, conhecido por sua textura cremosa e uso comum em pães e receitas culinárias diversas. O queijo manteiga e o queijo minas seguem, com 2.776 e 2.631 menções respectivamente, ambos apreciados por sua qualidade e variedade de usos na culinária brasileira. Outros tipos como o queijo ralado, mussarela, prato, cheddar, sem lactose e canastra também são mencionados, destacando-se por suas características específicas que atendem a diferentes preferências e necessidades dietéticas.

Identificar os principais acompanhamentos dos queijos pode fornecer informações valiosas para o setor. Saber quais produtos são frequentemente consumidos com os queijos permite a criação de estratégias de mercado eficazes, como o posicionamento estratégico de produtos relacionados nas prateleiras dos supermercados, incentivando vendas combinadas. Além disso, essas informações podem orientar campanhas de marketing direcionadas, promoções e até o desenvolvimento de novos produtos que complementem o consumo de queijo, aumentando a satisfação do consumidor e potencializando as vendas. A Figura 3 mostra em detalhes a *word cloud* dos acompanhamentos mais citados com os queijos.



Figura 3 - *Word cloud* dos principais acompanhamentos dos queijos.

Obs: em azul-claro, produtos categorizados como in natura e/ou minimamente processados; em azul-escuro, produtos ultraprocessados; em laranja, produtos processados; e em amarelo, ingredientes culinários, segundo o Guia Alimentar para a População Brasileira (Brasil, 2014).

Fonte: OC/Elaborado pelos autores.

Ao analisar a *word cloud*, é possível observar que, embora os acompanhamentos in natura e/ou minimamente processados sejam mais frequentes, há uma variedade significativa de opções processadas e ultraprocessadas. Isso reflete a realidade da alimentação contemporânea, onde praticidade e sabor frequentemente prevalecem sobre a escolha por alimentos mais nutritivos. Essa análise permite identificar hábitos e comportamentos de consumo, fornecendo informações que podem contribuir para tomadas de decisão mais assertivas no setor.

Ao analisar a distribuição regional das menções ao consumo de queijo, observou-se que a região Sudeste lidera com aproximadamente 53,13% do total. A região Nordeste segue em segundo lugar, com cerca de 21,33%, enquanto a região Sul fica em terceiro, com aproximadamente 15,04%. A região Centro-Oeste representa cerca de 7,70%, e a região Norte possui menos menções, com aproximadamente 3,92% do total. Vale ressaltar que um dos estados que mais consome queijos, Pernambuco, está no Nordeste, onde se destaca o grande consumo de queijos artesanais, especialmente o queijo coalho, que foi o mais citado.

Esses dados revelam importantes tendências de mercado, destacando a necessidade de estratégias diferenciadas para cada região. Enquanto o Sudeste pode se beneficiar de campanhas de fidelização e lançamentos de novos produtos, o Norte representa uma oportunidade de expansão, demandando iniciativas para aumentar a penetração e a acessibilidade dos queijos. Essa análise regional é fundamental para orientar decisões estratégicas das empresas do setor, visando maximizar o alcance e a eficácia de suas ações de marketing e distribuição.

### Conclusão

Os resultados obtidos pelo OC oferecem uma visão detalhada e atualizada sobre o comportamento dos consumidores de queijos no Brasil. A análise revelou que a maioria das publicações após o consumo expressa um sentimento positivo, refletindo uma alta satisfação com o produto. Essas informações são essenciais para orientar estratégias de mercado mais precisas e eficazes para a indústria de laticínios, permitindo não apenas entender as preferências dos consumidores, mas também adaptar-se rapidamente às mudanças nas demandas do mercado. Com o auxílio do OC, as empresas do setor podem se preparar melhor para tomar decisões informadas que impulsionem o crescimento e a competitividade no mercado de queijos no Brasil.

### Referências Bibliográficas

IBGE - Instituto Brasileiro de Geografia e Estatística. Pesquisa de Orçamentos Familiares 2017-2018. 2024. Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/livros/liv101670.pdf>. Acesso em: 3 de julho de 2024.

NOGUEIRA, T. S., MOURO, V. A., SIQUEIRA, K. B., GOLIATT, P. V. Z. C. Analysis of the Brazilian Artisanal Cheese Market from the Perspective of Social Networks. In: ABRAHAM, A., GANDHI, N., HANNE, T., HONG, T. P., NOGUEIRA, R., T., DING, W. (eds) Intelligent Systems Design and Applications. ISDA 2021. Lecture Notes in Networks and Systems, vol 418. Springer, Cham. [https://doi.org/10.1007/978-3-030-96308-8\\_84](https://doi.org/10.1007/978-3-030-96308-8_84). 2022.

COVINGTON, M. NLP for Prolog Programmers, Prentice-Hall. 1994.

RUSSELL, S.; NORVIG, P. Artificial Intelligence - a modern approach, Prentice Hall. 1995.

NOGUEIRA, T. S., SIQUEIRA, K. B., GOLIATT, P. V. Z. C. Construction of a training dataset for a sentiment analysis model of dairy products tweets in Brazil. Social Networks Analysis and Mining. 14, 85. <https://doi.org/10.1007/s13278-024-01254-5>. 2024.

BRASIL. Ministério da Saúde–MS. Guia Alimentar para a População Brasileira, Brasília: MS; 2014.

Autor(a) a ser contatado: Thallys da Silva Nogueira, Doutorando em Modelagem Computacional, Rua Antônio Ribeiro de Novais, 92, Ewbank da Câmara. [thallys.nogueira@estudante.ufjf.br](mailto:thallys.nogueira@estudante.ufjf.br).