

Percepção das redes sociais sobre o consumo de tomate: um estudo exploratório com dados do twitter

Social media perception on tomato consumption: an exploratory study with twitter data

DOI: 10.34140/bjbv3n5-016

Recebimento dos originais: 04/03/2021

Aceitação para publicação: 30/06/2021

Laleska Moda

Estudante de mestrado em Administração pelo Departamento de Economia, Administração e Sociologia da Esalq/USP
Escola Superior de Agricultura “Luiz de Queiroz” – ESALQ/USP
Estrada Monte Alegre, Piracicaba - SP, 13400-970
E-mail: laleska.moda@cepea.org.br

Margarete Boteon

Professora do Departamento de Economia, Administração e Sociologia da Esalq/USP
Escola Superior de Agricultura “Luiz de Queiroz” – ESALQ/USP
Estrada Monte Alegre, Piracicaba - SP, 13400-970
E-mail: margarete.boteon@usp.br

Diego Henrique de Campos Rodrigues

Analista de dados do Centro de Estudos Avançados em Economia Aplicada (Cepea), Esalq/USP.
Centro de Estudos Avançados em Economia Aplicada (Cepea) – ESALQ/USP
Estrada Monte Alegre, Piracicaba - SP, 13400-970
E-mail: cepea@cepea.org.br

RESUMO

O presente trabalho avalia a percepção do consumo de tomate através da análise das mensagens do microblog *Twitter* entre janeiro e novembro de 2020. O programa R foi utilizado como método de coleta e análise das mensagens. Utilizou-se duas técnicas de análises: mineração de texto com análise de conteúdo baseada em aspectos e análise do sentimento. Apesar das limitações do público que acessa as redes sociais, normalmente mais jovem, os resultados indicam uma sensibilidade positiva da rede ao consumo do tomate. O assunto mais comentado é o seu uso como ingrediente culinário. É válido destacar que é necessário maior refinamento da metodologia de mineração de palavras para utilizar as redes sociais como uma ferramenta de análise de percepção de consumo.

Palavras-chave: Tomate, Redes Sociais, Twitter, Percepção de Consumo

ABSTRACT

The present work evaluates the perception of tomato consumption through an analysis of the messages from the Twitter microblog between January and November 2020. The methodology used was the mining of Twitter messages in order to obtain key information and trends in tomato consumption in Brazil. The R statistical software was used as a method of collecting and analyzing messages. We used two analyzes techniques: Aspect-Based Opinion Mining and Aspect Based Sentiment Analysis (ABSA). Despite the limitations of the public that accesses social media, usually a younger audience, the results indicate a positive sensitivity to tomato consumption. The most talked about subject is its use as a culinary ingredient. It should be highlighted that it is even more necessary refinement of the words of mining

methodology to use social media as a tool of analysis of perception consumption.

Keywords: *Tomato, Social Media, Twitter, Consumer Perception*

1 INTRODUÇÃO

Nas pesquisas relacionadas ao consumo de alimentos, usualmente, utiliza-se métodos de campo, que demandam muito tempo e recursos. Com o avanço da tecnologia de informação, entretanto, uma nova vertente de estudos relacionado ao consumo está surgindo, através do monitoramento das conversas em redes sociais, que permitem a avaliação de um maior número de informações de forma mais rápida. Ainda que existam algumas ressalvas na utilização dessas ferramentas, a análise das mensagens em redes sociais já tem se mostrado um grande indicativo de tendências e percepção de consumo, uma vez que os usuários expressam sua opinião mais abertamente sobre produtos e serviços do que em questionários tradicionais.

Diante disto, o presente trabalho avalia a percepção do consumo de tomate pelo brasileiro através de uma análise de mineração das mensagens do microblog Twitter entre janeiro e novembro de 2020. Esse estudo exploratório faz parte de um projeto desenvolvido pelo Centro de Estudos Avançados em Economia Aplicada (Esaq/USP) em parceria com a Comissão do Tomate de Mesa (CNTM) do Instituto Brasileiro de Horticultura (IBRAHORT). O estudo contempla a análise da percepção de consumo de tomate no Brasil, como também outros dois eixos de pesquisa: produção e comercialização

2 REDES SOCIAIS COMO FERRAMENTA DE PERCEPÇÃO DE CONSUMO

As inovações tecnológicas têm permitido cada vez mais o largo uso de recursos digitais e a possibilidade de armazenar dados para posterior análise, o que, por sua vez conflui para a existência dos chamados *Big datas*. Diversas técnicas têm sido empregadas para analisar esses conjuntos de dados de grande volume, alta variedade e alimentação contínua. Técnicas estas que antes se davam apenas na área da ciência da computação, mas que atualmente são usadas em outras diversas aplicações, como é o caso dos estudos com processamento de linguagem natural (PNL). E isso se deve, em parte, pelo crescimento das mídias sociais, como *Facebook*, *Twitter*, *blogs*, *microblogs*, *fóruns on-line*, entre outras, que possibilitaram a existência de um grande volume de dados opinativos registrados no âmbito digital e passíveis de serem coletados e analisados de forma muito mais dinâmica e acessível (LIU, 2012).

Ao se tratar de dados textuais, várias técnicas de análise de texto, ou mineração de texto, podem ser empregadas a fim de extrair informações desse tipo de banco de dados, que podem ser construídos a partir de dados coletados de redes sociais, *e-mails*, *fóruns on-line*, respostas de pesquisas e etc. Esse tipo de análise envolve avaliações estatísticas, linguística computacional e até mesmo *machine learning* (GANDOMI; HAIDER, 2015).

Dentre essas técnicas está a Análise de Sentimento, ou Mineração de Opiniões, que busca interpretar opiniões, sentimentos, atitudes, emoções e avaliações das pessoas a partir da linguagem escrita e então classificá-los em três direções: positivos, negativos ou neutros (LIU, 2012). Esse procedimento permite analisar o quão positivo ou negativa é a percepção das pessoas sobre um determinado tópico, produto, serviço, indivíduo ou qualquer outro assunto, a partir de um conjunto de opiniões coletadas em forma de textos. A Mineração de Opiniões já foi utilizada em estudos de diversas áreas como tecnologia (LI, et al., 2019), finanças (BOLLEN; MAO; ZENG, 2011), marketing, política e estudo de questões sociais (ÖZTÜRK; AYVAZ, 2018).

Isso pode ser uma ferramenta muito rica, uma vez que possibilita maior compreensão sobre a percepção de um consumidor em relação aos diferentes atributos de um produto ou serviço, permitindo, assim, às empresas diagnosticarem problemas e proporem soluções de forma mais assertiva (CHAMLERTWAT; et al., 2012).

3 CONSIDERAÇÕES METODOLÓGICAS

A metodologia empregada foi a mineração de mensagens do *Twitter* visando obter informações-chaves e tendências do consumo de tomate no Brasil. Como método de coleta e de análise, foi utilizado o programa R. Utilizou-se duas análises: técnicas de mineração de texto com análise de conteúdo baseada em aspecto e análise do sentimento (CHAMLERTWAT et al., 2012; R. M. et al., 2018; SAMMOGIA; RIEDEL; RUGGERI, 2020). A dinâmica empregada no trabalho, resumidamente, pode ser resumida em 4 etapas:

- 1) **Coleta de dados:** coleta de tweets contendo a palavra “tomate”, pelo programa R no período de janeiro a novembro de 2020, totalizando 495.659 tweets.
- 2) **Preparação:** remoção de links, imagens e frases contendo a palavra “tomate” fora de contexto de consumo. Um estudo sobre o banco de dados remanescente manualmente (1.000 tweets) permitiu organizar as palavras-chaves básicas e criação de categorias para o estudo.
- 3) **Análise de conteúdo:** através da análise manual (os assuntos mais recorrentes nos 1.000 tweets avaliados) e estudos anteriores sobre o consumo de tomate foram selecionadas 5 categorias, sendo 4 referentes ao consumo (culinária, saúde e nutrição, sustentabilidade e produção) e 1 referente a produtos “derivados” do tomate. A categoria “derivados” foi criada devido à grande menção de produtos que contém a palavra tomate, mas não são o produto *in natura*. Em cada uma das 5 categorias foi selecionado um grupo de palavras relevantes que se repetiam nos tweets, para que fosse possível a separação no banco de dados pelo programa.
- 4) **Análise de sentimento:** conhecida como *Apect Based Sentiment Analysis* (ABSA), busca identificar os aspectos (tópicos, assuntos, etc.) mencionados em um texto e determinar a polaridade do sentimento (positivo e negativo). Esse procedimento permite analisar o quão positivo ou negativa é a percepção das pessoas sobre um determinado tópico, produto, serviço, indivíduo ou qualquer outro assunto, a partir de um conjunto de opiniões coletadas em forma de textos.

Os resultados da análise de conteúdo e de sentimento estão descritos na Seção 3.

4 RESULTADOS E DISCUSSÃO

4.1 ANÁLISE DE CONTEÚDO: PALAVRAS MAIS COMENTADAS POR CATEGORIA

A tabela 1 mostra as principais palavras-chaves mais citadas em cada categoria. Através da análise anterior de 1.000 tweets, também foi possível observar o contexto inserido relacionado às palavras, dentro de cada categoria. Os resultados indicam “Derivados” e “Culinária” se destacam entre as categorias avaliadas, e as palavras mais citadas revelam que, independentemente da receita, o tomate está na cozinha do consumidor, seja ele em forma de molho, sobre as pizzas, compondo as saladas ou acompanhando o tradicional arroz no prato brasileiro.

Vale destacar que, foi muito recorrente a associação do tomate ao consumo de saladas nas refeições. É válido notar que a “salada” se relaciona tanto com a culinária como com a saúde, sendo destaque das dietas e apelo saudável. Na categoria saúde, as menções se relacionam especialmente com a salada, aliada a dietas de emagrecimento ou veganas.

Houve também forte menção ao miojo de tomate e, apesar de não ser um subproduto, expressa a busca por conveniência. Isso foi ressaltado devido ao perfil principal de usuários do Twitter (geração millenials, adultos entre 24 e 39 anos). Quanto à categoria “preço”, apesar de menos presente que as outras categorias, muitos usuários relacionam o produto como caro.

Tabela 1. Palavras mais comentadas por categoria no Twitter a respeito do tomate.

Categoria	Ranking	Palavra-chave	Frequência
Derivados	1 ^a	Molho	70.000
	2 ^a	Miojo	39.000
	3 ^a	Pizza	13.000
	4 ^a	Queijo	12.000
	5 ^a	Macarrão	11.000
Culinária	1 ^a	Molho	22.600
	2 ^a	Arroz	19.000
	3 ^a	Salada	18.500
	4 ^a	Macarrão	15.100
	5 ^a	Pão	12.500
Produção e Sustentabilidade	1 ^a	Colher	2.600
	2 ^a	Plantar	2.500
	3 ^a	Molho	1.650
	4 ^a	Feira	1.500
	5 ^a	Vou	1.100
Preço	1 ^a	Reais	2.500
	2 ^a	Molho	2.400
	3 ^a	Falta	1.450
	4 ^a	Caro	1.400
	5 ^a	Arroz	1.150
Saúde e Nutrição	1 ^a	Saudável	1.300
	2 ^a	Molho	1.400
	3 ^a	Alface	1.350
	4 ^a	Dieta	1.050
	5 ^a	Vegano	1.050

Obs: o período do estudo das categorias/palavras no Twitter foi de janeiro a novembro de 2020.

4.2 ANÁLISE DE SENTIMENTOS

A tabela 2 mostra a distribuição (%) de percepção entre positivo/negativo e neutro por categoria e no geral. A análise de sentimentos em toda a base de dados (janeiro a novembro) indicou prevalência de *tweets* positivos. As palavras com polaridade positiva mais frequentes foram “fresco”, “bom” e “melhor”, enquanto as negativas foram “cara”, “pouco” e “ruim”.

Quanto aos *tweets* negativos, alguns se relacionam à preferências pessoais (“não gostode tomate”, “odeio molho de tomate”, etc), dietas não equilibradas no caso do “miojo de tomate” e, especialmente ao preço do produto, como percebido pela maior participação de *tweets* negativos na categoria “preço”. As palavras negativas mais frequentes foram “caro/a”, “ruim”, “pouco” e “pior”. Na análise de sentimentos por trimestre, não foi encontrado uma grande mudança no resultado de *tweets* positivos, negativos ou neutros.

Tabela 2. Distribuição (%) da análise de sensibilidade (positivo/neutro e negativo) geral eentre as categorias analisadas.

Categoria	Polaridade	1º Tri/202º	Tri/203º	Tri/204º	Tri/20*	2020 (jan-nov)
Culinária	Neutro	33%	33%	35%	35%	34%
	Positivo	46%	47%	44%	44%	45%
	Negativo	21%	20%	22%	21%	21%
Preço	Neutro	28%	29%	33%	28%	27%
	Positivo	40%	39%	39%	40%	40%
	Negativo	32%	32%	28%	32%	34%
Sustentabilidade	Neutro	34%	32%	34%	35%	34%
	Positivo	41%	43%	40%	38%	41%
	Negativo	25%	24%	26%	27%	25%
Saúde / Nutrição	Neutro	33%	32%	32%	34%	33%
	Positivo	45%	48%	45%	43%	45%
	Negativo	22%	20%	22%	23%	22%
Derivados	Neutro	32%	32%	32%	31%	32%
	Positivo	43%	49%	48%	49%	48%
	Negativo	25%	19%	20%	20%	20%
Geral	Neutro	26%	29%	32%	28%	29%
	Positivo	54%	51%	46%	53%	50%
	Negativo	20%	20%	22%	20%	21%

*Obs: Os *Tweets* foram avaliados de janeiro a novembro. O último trimestre só apresentadados até novembro.

5 CONSIDERAÇÕES FINAIS

O presente estudo, de caráter exploratório, indicou uma sensibilidade positiva das redessociais (do Twitter) em relação ao tomate. O assunto mais comentado é o seu uso como ingrediente culinário em diversas receitas e, especialmente, na salada e no molho. Através da análise de conteúdo e de sentimentos também foi possível verificar a relevância dos produtos processados a partir do tomate, especialmente o miojo, o que mostra o interesse dos internautas em produtos práticos e indulgentes, também refletindo

a participação dos millenials na amostra. Ainda que a associação geral da categoria ainda tenha sido positiva, não devemos destacar o aumento dos tweets negativos relacionados ao tomate. Grande parte dos internautas ainda vê o produto como caro.

Apesar dos achados no presente artigo, é válido destacar que ainda é necessário maior refinamento da metodologia de mineração de palavras para utilizar as redes sociais como uma ferramenta de análise de percepção de consumo.

REFERENCIAS

BOLLEN, Johan; MAO, Huina; ZENG, Xiaojun. Twitter mood predicts the stock market. *Journal Of Computational Science*, v. 2, n. 1, p. 1-8, mar. 2011.

CHAMLERTWAT, Wilas; BHATTARAKOSOL, Pattarasinee; RUNGKASIRI, Tippakorn; HARUECHAIYASAK, Choochart. Discovering Consumer Insight from Twitter via Sentiment Analysis. *Journal of Universal Computer Science*. p. 973-992. abr. 2012.

GANDOMI, Amir; HAIDER, Murtaza. Beyond the hype: big data concepts, methods, and analytics. *International Journal Of Information Management*. p. 137-144. abr. 2015.

LI, Xin; XIE, Qianqian; JIANG, Jiaojiao; ZHOU, Yuan; HUANG, Lucheng. Identifying and monitoring the development trends of emerging technologies using patent analysis and Twitter data mining: the case of perovskite solar cell technology. *Technological Forecasting And Social Change*, v. 146, n. 0, p. 687-705, set. 2019.

LIU, Bing. *Sentiment Analysis and Opinion Mining*. Toronto: Synthesis Lectures On Human Language Technologies, 2012. 184 p.

ÖZTÜRK, Nazan; AYVAZ, Serkan. Sentiment analysis on Twitter: a text mining approach to the syrian refugee crisis. *Telematics And Informatics*. p. 136-147. abr. 2018.

SAMOGGIA, A.; RIEDEL, B.; RUGGERI, A. Social media exploration for understanding food product attributes perception: the case of coffee and health with Twitter data. *British Food Journal*. 2020.