

Social media exploration for understanding food product attributes perception: the case of coffee and health with Twitter data

Social media
and food
product
attributes

3815

Antonella Samoggia
*Department of Agricultural and Food Sciences and Technologies,
University of Bologna, Bologna, Italy, and*
Bettina Riedel and Arianna Ruggeri
University of Bologna, Bologna, Italy

Received 13 March 2019
Revised 26 July 2019
12 September 2019
5 February 2020
Accepted 15 March 2020

Abstract

Purpose – Food companies and consumers are increasingly interested in healthy food and beverages. Coffee is one of the most commonly consumed beverages worldwide. There is increasing consensus that coffee consumption can have beneficial effects on human body. This paper aims at exploring Twitter messages' content and sentiment towards health attributes of coffee.

Design/methodology/approach – The research adopted a utilitarian and hedonic consumer behaviour perspective to analyse online community messages. A sample of 13,000 tweets, from around 4,800 users, that mentions keywords coffee and health was collected on a daily basis for a month in mid-2017. The tweets were categorized with a term frequency analysis, keyword-in-context analysis and sentiment analysis.

Findings – Results showed that the majority of tweets are neutral or slightly positive towards coffee's effects on health. Media and consumers are dynamic Twitter users. Findings support that coffee consumption brings favourable emotions, wellness, energy, positive state of mind and an enjoyable and trendy lifestyle. Many tweets have a positive perception of coffee health benefits, especially relating to mental and physical well-being.

Research limitations/implications – The high number of users and tweets analysed compensates the limited amount of time of data collection, Twitter messages' restricted number of characters and quantitative software analysis limitations.

Practical implications – The research provides valuable suggestions for food and beverage industry managers.

Originality/value – This work adds value to the literature by expanding scholars' research on food product attributes perception analysis by using social media as a source of information. Moreover, it provides valuable information on marketable coffee attributes.

Keywords Social media, Coffee, Health, Food, Perception, Sentiment analysis, Content analysis

Paper type Research paper

1. Introduction

Coffee is one of the most commonly consumed beverages in the world, with fast growing sales at global level ([International Coffee Organization \[ICO\], 2018](#)). The coffee industry is continuously investing to meet coffee consumption trends by offering new sensorial attributes, consumption experiences and occasions of consumption. There is a growing trend to highlight coffee's product of origin, sustainability, organic, fair trade and environmental friendliness ([Samoggia and Riedel, 2018](#)). Consumers are continually looking for new products in the coffee category, and the coffee industry has to keep up with product innovation.



© Antonella Samoggia, Bettina Riedel and Arianna Ruggeri. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licences/by/4.0/legalcode>

Coffee has seen a revolutionary growth over the recent decades from an ordinary commodity to a speciality product. One contributing factor to an increase of coffee consumption is consumer enthusiasm for gourmet coffee varieties (e.g. Espresso-based beverages, frozen-blended, cold brew, nitrogen-infused). Today's coffee trends include premiumisation, customisation, single origin and roast type (Panhuyzen and Pierrot, 2018). Other factors that contribute to an increase in coffee consumption are convenience and functional aspects (NCA, 2018a). Convenience means that high quality coffee can be purchased at speciality coffee shops or easily prepared at home. Functional means coffee with functional benefits. Recent studies found that consumers want coffee with benefits for metabolism, alert or energy boosting, for healthier skin or high in antioxidants (NCA, 2018a). There is increasing interest towards health-oriented coffee products, with specific focus on ready-to-drink coffee (Just-drinks.com, 2018). Coffee is an essential part of many people's daily routines globally. People drink coffee for pleasure because of its sensory properties (taste and aroma) and for stimulation due to its psychoactive properties (Nieber, 2017). This popularity and the possible beneficial health effects make coffee a product of public and scientific interest (Lofffield *et al.*, 2016), and of importance for the public health sector (Grosso *et al.*, 2017).

Consumers have contradictory perceptions of the effects of coffee consumption on their body and health (Samoggia and Riedel, 2019; Institute for Scientific Information on Coffee - ISIC, 2016). Some people perceive coffee as bad for their health and wish to drink less coffee. Others enjoy drinking coffee, perceiving coffee consumption as a pleasant moment in their daily routine, to be consumed alone or shared with others (Samoggia and Riedel, 2018). Yet, it is unclear whether coffee fits within the trend of indulgent treats, healthy lifestyles and diets, increasingly common among consumers (Euromonitor International, 2017). Food products enriched with healthier and tastier options are growing in popularity as they satisfy busy and gratification-seeking lifestyles (Jiang and Leung, 2012).

Coffee is already consumed as a functional drink, in particular as a source of energy in the morning for its caffeine content. However, its health impact has been questioned over the years due to consumers' concerns over caffeine effects on human body. Currently, there is limited understanding over consumers' perception on coffee health impact. Coffee is consumed at global level, and social media messages can capture worldwide consumers' opinions over coffee and health and provide insights on associated messages' content and sentiment. Considering current consumers' interest towards healthier food options, an explorative analysis of the perceptions on the relationship between coffee and health may provide advice on new market management potentials to the coffee industry.

The present study extends research approaches of food product attributes perception analysis by adopting social media as a source of information. This paper aims at exploring the use of Twitter as a tool to develop future marketing opportunities for the coffee industry. In particular, the paper focusses on exploring Twitter messages' content and sentiment towards health attributes of coffee.

2. Theoretical framework

The present research adopted a utilitarian and hedonic consumer behaviour perspective. Utilitarian value and hedonic value have different effects on consumer behaviour (Chaudhuri and Holbrook, 2001). Utilitarian products are "effective, helpful, functional, necessary and practical". The products are purchased for their practical uses and satisfy the consumer's needs (Holbrook, 2006; Lu *et al.*, 2016). Hedonic products are "fun, exciting, delightful, thrilling, and enjoyable". The value "arises from people's own pleasure in the consumption experience, appreciated for their own sake" (Holbrook, 2006; Lu *et al.*, 2016).

The present study explored how health, a utilitarian attribute of coffee, is perceived (Wang and Yu, 2016). The perception can influence consumer purchasing behaviour and companies'

marketing strategies to cater to that behaviour (Chen and Sun, 2014; Cronin *et al.*, 2000; Chen and Dubinsky, 2003). Past studies support that the understanding of the appreciation of product attributes serves as value cues. The consumers' value perception can be used to predict consumer behaviour.

Recent reviews on coffee consumer behaviour highlight the importance of the hedonic dimension in coffee consumption. The sensorial appreciation (e.g. aroma, taste) and the recreational (e.g. socialisation) aspects are key elements in coffee consumption (Samoggia and Riedel, 2018). Utilitarian values are also crucial in coffee consumption. Many coffee consumers are motivated by the stimulation and alertness that coffee can provide. Coffee drinking experience becomes functional when coffee, with its caffeine content, is consumed for the physiological benefits (e.g. increased energy, improved performance, diminished headache and increased blood pressure).

3. Literature review

3.1 Health properties of coffee

The scientific community is researching the health effects of coffee. Clinical research has studied the associations between coffee and cancer, cardiovascular diseases, neurological, gastrointestinal, as well as liver systems and effects on pregnancy. Several studies attempted to point out the risks and benefits related to coffee intake; however, there is no consensus to date on the effects of coffee on human health (Grosso *et al.*, 2017; Mitchell *et al.*, 2014; Nieber, 2017).

In the 1970s and 1980s, the first research findings concerning the health effects of coffee consumption connected coffee intake to an increased risk of cardiovascular problems. A study with 12,000 patients found drinking of one to five cups of coffee a day increased risk of heart attacks by 60% (Jick *et al.*, 1974). Another study from 1978 found a short-term rise in blood pressure after three cups of coffee (Robertson *et al.*, 1978). Recent studies concluded that drinking coffee was consistently connected with a lower risk of death from all causes. There are promising results on the beneficial effects of coffee on cancer, cardiovascular disease, as well as metabolic and neurological disorders (Grosso *et al.*, 2017; Poole *et al.*, 2017; Wikoff *et al.*, 2017). One of the most cited assessments of caffeine health effects in the peer-reviewed literature is published by Health Canada (Nawrot *et al.*, 2003). They concluded that caffeine consumption at amounts up to 400 mg/day for healthy adults, 300 mg/day for healthy pregnant women and 2.5 mg/kg-day for healthy children is not associated with adverse effects. As an update to this publication, Wikoff *et al.* (2017) provided a systematic review of data on adverse effects of caffeine and reached similar conclusions. The evidence supports that consumption of up to 400 mg caffeine/day in healthy adults is not associated with adverse cardiovascular, behavioural, reproductive effects or bone status.

At present, researchers agree that moderate consumption of coffee may help improve concentration and physical performances (Poole *et al.*, 2017). Others maintain that drinking coffee may negatively affect certain bodily functions (Wikoff *et al.*, 2017). Poole *et al.* (2017) analyse evidence of meta-analyses of 218 observational and interventional studies of coffee consumption and conclude, that "coffee consumption was more often associated with benefit than harm. Coffee drinking is safe within usual consumption patterns". A moderate amount of coffee intake is considered to be around 3–4 cups/day providing 300–400 mg/day of caffeine (European Food Safety Authority [EFSA], 2015; Poole *et al.*, 2017; Wikoff *et al.*, 2017). The largest reduction in relative risk of death is to be found at three cups a day, compared to non-coffee drinkers. Accordingly, Grosso *et al.* (2017) conclude that adverse effects of coffee drinking are limited to pregnancy and to female with increased risk of bone fracture.

3.2 Social media and food product perception

Consumers' accessibility to information on the effects of coffee consumption is high, deriving from television, newspapers, websites and other social media. The practice of using social media to create and disseminate information and opinions is to be mined as a valuable resource for consumers' perceptions and sentiments in various fields of society and economy (Mostafa, 2013; Yu *et al.*, 2013).

Consumers are increasingly using the Internet for a number of purposes, from searching and reading information, to buying and selling goods. Social media in particular are the virtual place for the exchange of opinions and information. Thus, social media become a source of information to explore what consumers think (Singh *et al.*, 2016) and feel (Dodds *et al.*, 2011). Social media can contribute to consumers' marketing management (Chua and Banerjee, 2013) to gain insights on consumers' needs and perceptions of product attributes (Fileri, 2013). Consumers may also be better inclined in giving their opinions about products on social media, rather than through a traditional questionnaire (Rathore *et al.*, 2016).

By examining the content of tweets, companies can gain insights on consumers' attitudes about a specific issue or product. User-generated content is a valuable source of information to understand consumers' acceptance of innovations (Dubè *et al.*, 2018). Social media marketing is based on the perspective that social media content is a conversation initiated by consumers/audiences or a business/product/services (Dwivedi *et al.*, 2015). The communication created on social media allows companies and consumers to mutually learn about the experiential use of a given product (Alalwan *et al.*, 2017). A good practical example of its increasing importance is the "Observatory for social monitoring" (Nomisma, 2019) that aims at supporting industry and market research by listening and monitoring online conversations on products and brands to get consumer insight. Social media are the space where information exchange, communication and confrontation happen (Nomisma, 2019). The fact that this first observatory focusses on coffee underlines the importance of this type of data for consumer research in the coffee industry.

4. Materials and methods

Data gathering and elaboration were carried out in different steps (Figure 1).

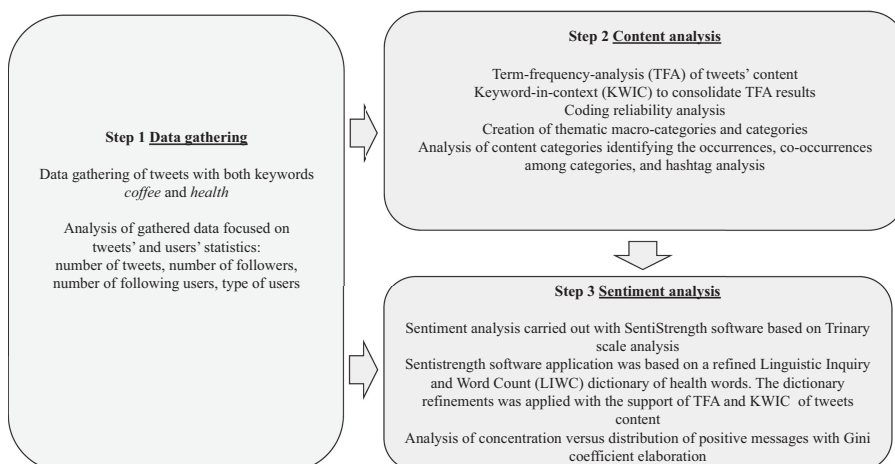


Figure 1.
Data gathering and elaboration

4.1 Step 1 - data gathering

A search of tweets was implemented between May 5, 2017 and June 4, 2017 on a daily basis. The search included tweets with both keywords coffee and health. To facilitate data elaboration and guarantee adequate coherence to the interpretation, tweets were included only if written in English (Graham *et al.*, 2014; Lamy *et al.*, 2016; Mostafa, 2018; Thelwall *et al.*, 2011). Tweets, retweets and repeated tweets were collected with support of Ncapture 10. Twitter users' information included total number of tweets, number of followers and number of following users.

4.2 Step 2 – content analysis

The user-generated content was analysed based on a content analysis, a methodology widely adopted to investigate and categorize the content of gathered messages (Krippendorff, 2018; Macnamara, 2005; Naaman *et al.*, 2010; Ruggeri and Samoggia, 2018). A term frequency analysis (TFA) adopting Nvivo software (version 10, QSR International Pty Ltd, Doncaster, Victoria, Australia) was used to define tweet contents. The TFA was applied to guarantee a bottom-up approach aimed to guarantee the maximum adherence to the analysis of word of mouth. The TFA included tweets, repeated tweets and retweets. The TFA defined a preliminary dictionary. This was consolidated with keyword-in-context analysis (KWIC). KWIC allowed to fully understand the words used in TFA and possibly confirm or exclude the unclear or misused words. Coded results were reviewed by two researchers. Coding reliability was checked by giving two trained researchers the results and asking them to confirm the machine-coded allocations to the categories. A high level of reliability was reached (91%). Coding reliability aimed at evaluating whether a coding exercise yields the same data within a tolerable margin of error. The higher reliability is, the more trustworthy and reproducible is the data set (Mikhaylov *et al.*, 2012; Landry, 2003; Hayes, 2005). Any inconsistencies pointed out by any researcher were then reviewed until consensus was reached. The categories' creation followed so as to analyse original tweets' contents according to the most common key topics included in the tweets. Categories were analysed by identifying the occurrences and co-occurrences among categories.

4.3 Step 3 - sentiment analysis

Sentiment analysis, also known as opinion mining, is a method used to understand the sentiment of social media users. The sentiment analysis shows if the respondent has a positive, negative or neutral opinion about a certain argument (Dubè *et al.*, 2018; Perera and Caldera, 2017). It allows businesses to collect, monitor and analyse user-generated data of their corporate social media sites and of their competitors. This allows to define upcoming trends, to explore what consumers want and think and to analyse the competitive business environment (He *et al.*, 2015). The up-to-date feedback obtained from social media is extremely valuable to marketing research companies and public opinion organizations (Mostafa, 2013). It is mainly applied to understand data of social networks, as Facebook, Twitter, Instagram and LinkedIn with the support of specific software programs (Kapoor *et al.*, 2017; Mäntylä and Graziotin, 2018; Rathore and Ilavarasan, 2017).

Twitter can be a reliable source of information in analysing consumers' attitudes and redesigning companies' marketing and advertising campaigns (Mostafa, 2013). Twitter provides a high number of big data, and the capability of analysing such amount of data is important. Sentiment analysis is a text analysis and mining method is a "distinct method to study people's opinions in terms of views, attitudes, appraisals, and emotions towards entities, individuals, issues, events, topics, and their attributes in a more thorough way" (Liu and Lopez, 2016; Pang and Lee, 2008; Stieglitz *et al.*, 2014).

What makes sentiment analysis on social media particularly valuable is that the data originate from a very heterogeneous group of social media users (Singh *et al.*, 2016). Therefore, companies have the opportunity to utilize data to gather information, perception and opinions for the creation of a new product and innovative product positioning that reflects the actual need of consumers. For this reason, this type of analysis is frequently used to analyse the launch of a new product (Rathore and Ilavarasan, 2017; Kim *et al.*, 2015; He *et al.*, 2013; Liu *et al.*, 2017; Rathore and Ilavarasan, 2017).

4.4 Sentiment analysis application

The sentiment analysis was carried out to assess social media users' perception and possible polarization of opinions (Ravaglia *et al.*, 2016). The content was analysed to identify the general perception over messages. A sentiment analysis was carried out to assess social media users' perception and possible polarization of opinions (Ravaglia *et al.*, 2016). Because tweets' content is often informal, poorly structured in terms of syntax and short to respect length limits set by Twitter rules, the literature identifies the lexical approach as the most suitable to implement sentiment analysis (Boiy and Moens, 2009; Mostafa, 2017; Perera and Caldera, 2017).

The sentiment analysis was applied to evaluate the positive, neutral or negative value of categories by tweets. This step included the use of SentiStrength software that applies a lexical approach and is among the most used free software to implement the analysis on informal conversations, providing accurate results for Tweets analysis (Chae, 2015; Martínez-Cámara *et al.*, 2012; Thelwall *et al.*, 2011; Thelwall, 2017). Moreover, SentiStrength is effective in evaluating short online texts, and it allows the researcher to modify the lexicon necessary for the present research to improve the accuracy of text analysis (Friedrich *et al.*, 2015; Ribeiro *et al.*, 2016).

SentiStrength is based on an adaptable LIWC dictionary (Tausczik and Pennebaker, 2010); sentences values range was selected on a trinary scale – positive/neutral/negative – between +5 as very positive and –5 as very negative (Booth and Matic, 2011; Thelwall, 2017). It is based on the psychological research, and it reports that positive and negative emotions can be experienced simultaneously and independently (Booth and Matic, 2011; Mostafa, 2013).

The dictionary was integrated and reassessed by the authors both in respect of terms and measures to meet the objective of the study (Thelwall, 2017). In particular, based on the results of the preliminary TFA implemented in step one, Linguistic Inquiry and Word Count (LIWC) dictionary was integrated and consolidated to avoid results' distortion. The dictionary refining focussed on LIWC dictionary on health words. Given the KWIC results, words involving diseases or health status resulted as with an informative or descriptive purpose, assuming a neutral purpose. The original LIWC dictionary needed to be reassessed. Several words are assumed in a common jargon, within an informal conversation, as with a negative connotation (e.g. cancer). Moreover, some slang terms had to be integrated to follow common jargon used on Twitter. Before implementing any dictionary change, the word was double-checked with a key word in text analysis and validated by the authors. A list of emoticon, as provided by the original dictionary, was also used, as Nvivo software is able to catch and codify all available emoticons present on the Internet. Finally, the research aimed at investigating the possible concentration versus distribution of the positive sentiment on health and coffee among different types of users. A Gini coefficient elaboration was carried out on positive tweets.

5. Results

5.1 Analysis of tweets

Tweets focussing on coffee and health are 12,894 including the following: 3,933 original tweets and 8,957 repeated and retweeted messages (40% repeated and 60% retweeted

tweets). Messages included 1,639 different hashtags, mentioned 13,688 times (with an average use of 8.4 times per hashtag). 39.2% of messages use hashtags (2.7 hashtags per message).

There are 4,714 users that mentioned health and coffee in their Tweets, meaning that each of them sent an average of 2.7 messages on health and coffee. Active users are calculated based on the number of tweets (original tweets and retweets) and visible users based on the number of followers. These are not necessarily the same users (Table 1). Results show that some of the most active users are also the most visible users. Yet, there are differences among the typologies of the top visible and active users (Table 1).

The most visible users mentioning the link between coffee and health are the media, whereas the most active users are single persons, closely followed by media. This shows that the dissemination of the coffee and health issue is of main interest for media (magazine, journals, newspaper, etc.). Media messages content can be expected to mostly spread information and let readers and their Twitter follower know about recent scientific development. A relevant number of single persons play a role in increasing the visibility of the issue.

The visibility of users differs. There are users that are extremely visible with more than 6.5 million followers. If one of the most visible users tweets, there are millions of people that can be exposed to the content of the message. When one of these users is active, the impact of a message can be relevant. Similarly, when less visible users are very active and frequently tweet on a topic, they may have a high impact and give relevance to a topic. The highest number of health and coffee tweets sent by the same user was around 440. In one month three, among the most visible and most active users possibly reached around 1.5 *bn* visualizations on health and coffee.

5.2 Content analysis

The word grouping resulted in the definition of main categories that synthesized the main themes (Table 2). These were grouped as a thematic category on food and beverage habits and preferences (time, meal, food categories, typologies of coffee, sensorial), consumers' types and living habits (consumers' groups, lifestyle), information type (science-based) and a category clearly referring to health and nutrition (see Appendix 1 for a detailed description). Messages mainly focus on the time of the day, week or season (25%), body parts (17%), food categories (16%) and lifestyles (10%). This shows that messages on coffee and health refer to coffee consumption habits, mentioned together with other food and in relation with parts of the body. Messages also refer to lifestyle and occasions of consumption within meals.

	Most visible users (%)	Most active users (%)
Media	80	40
Single persons	8	48
Journalists	4	4
Associations	8	8
Total	100	100

Note(s): The highest number of followers by one user: 6,639,130; The highest number of tweets on health and coffee by one user: 440

List of the most visible and active media: CNN International; Men's Health Mag; Women's Health; Health; Zee News; Dylan Sprouse; New Scientist; Today's Parent; The Independent; Daily Mail Online; Runner's World; Harvard Health; This Morning; FITNESS Magazine; NBC News Health; Men's Fitness; Daily Mirror; TIME Health

Table 1.
Typologies of the 300
most visible or
active users

Thematic categories					
		Meal		867	8%
		Time		2,589	25%
		Food categories		1,651	16%
		Lifestyle		1,087	10%
		Consumer group		771	7%
		Body parts		1,772	17%
		Science-based		542	5%
		Sensorial		156	1%
		Typology of coffee		1,071	10%
		<i>Total</i>		<i>10,506</i>	<i>100%</i>

				Main co-occurring category type	Co-occurring highest category (%)
Nutrition	Nutrition	423	7%	Meal	27
	Diet	597	10%	Types of coffee	47
	Overconsumption	399	7%	Consumers group	72
Wellness	Body wellness	1,392	24%	Lifestyle	41
	Mental wellness	606	10%	Consumers group	40
	energy				
	Mental wellness health	426	7%	Meal	41
Health Diseases	Health diseases	2,074	35%		
	<i>Alzheimer</i>	180	9%	Body parts	2
	<i>Anxiety</i>	37	2%	Food categories	22
	<i>Cancer</i>	1,029	50%	Body parts	95
	<i>Dementia</i>	216	10%	Meal	15
	<i>Depression</i>	17	1%	Time, types of coffee	73
	<i>Diabetes</i>	45	2%	Body parts	34
	<i>Disease</i>	203	10%	Body parts	8
	<i>Headache</i>	45	2%	Science based	98
	<i>Lymphoma</i>	21	1%	-	0
	<i>Pain</i>	209	10%	Time	37
	<i>Parkinson</i>	14	1%	Body parts	82
	<i>Stroke</i>	58	3%	Body parts	4
	<i>Total</i>	<i>5,917</i>	<i>100%</i>		

Table 2. Categories' occurrences and co-occurrences

The health and nutrition categories describe health issues from different perspectives. These have been categorized according to specific health diseases, nutrition-based disease or disorders, body and mental wellness (see [Appendix 2](#) for detailed description). One-third of tweets on the health and nutrition macro-category mentions a health disease category and one-fourth refer to body wellness (24%), mental energy and diet (10% both) ([Table 2](#)). The most mentioned health categories are cancer (50%), dementia, disease and pain (all 10%).

There are relevant co-occurrences between health and nutrition categories and general categories ([Table 2](#)). Tweets on health, coffee and headache make reference to science-based information, cancer and Parkinson also focus on body parts, depression with types of coffee and time of day, overconsumption of coffee and health mentions specific consumer groups.

In addition to health and coffee, the most mentioned hashtags are fitness, caffeine, beauty, travel, food, fashion, style, summer and wellness ([Table 3](#)). There is limited use of hashtags with health disease words. The most frequent ones are as follows: #prostate, #headache, #prostatecancer, #mentalhealth, #dementia, #brainhealth, #diabetes, and #livercancer. These were mentioned between 67 and 16 times. The use of hashtags allows to group topics so that users can easily reach related topics using the search function. Twitter users' final aim is to become visible and potentially create a trending status on a given topic.

Hashtag	Frequency	Social media and food product attributes
#health	2,920	3823
#coffee	2,732	
#fitness	536	
#caffeine	405	
#beauty	342	
#travel	308	
#food	221	
#fashion	186	
#style	175	
#summer	168	
#wellness	154	
#cancer	151	
#cars	151	
#coffeeandcars	151	
#nutrition	142	
#uk	130	
#sports	127	
#diet	119	
#tea	116	
#fit	110	
#weightloss	108	
#hair	107	
#energy	103	
#valentus	104	

Table 3. Hashtags mentioned more than 100 times

Hashtags allow to identify the most active users on a given topic. Research results confirm that users aim at creating a trend related to coffee and health mainly related to an enjoyable way of life. Users may not be interested in creating a trend on a specific medical focus, when referring to coffee and health.

5.3 Sentiment analysis

The sentiment analysis was carried out on all messages (Figure 2) and specifically on messages mentioning a health category (Figure 3). The first sentiment analysis shows that in general tweets are neutral or positive (around 45–47%). This result is coherent with previous studies (Mostafa, 2018) that argued that the negative opinions on social media are generally fewer than positive ones. The high presence of indifferent opinions can be explained by the high presence of media among social media users. Media content is mostly informative,

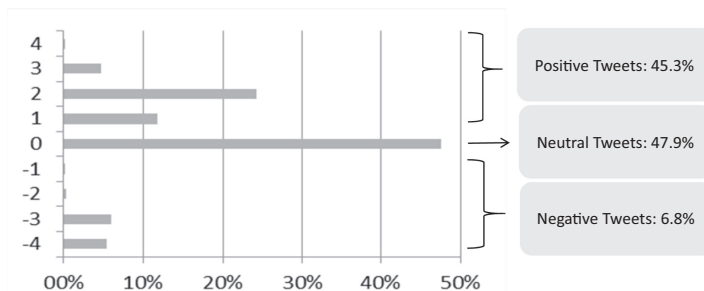


Figure 2. Sentiment analysis of all tweets

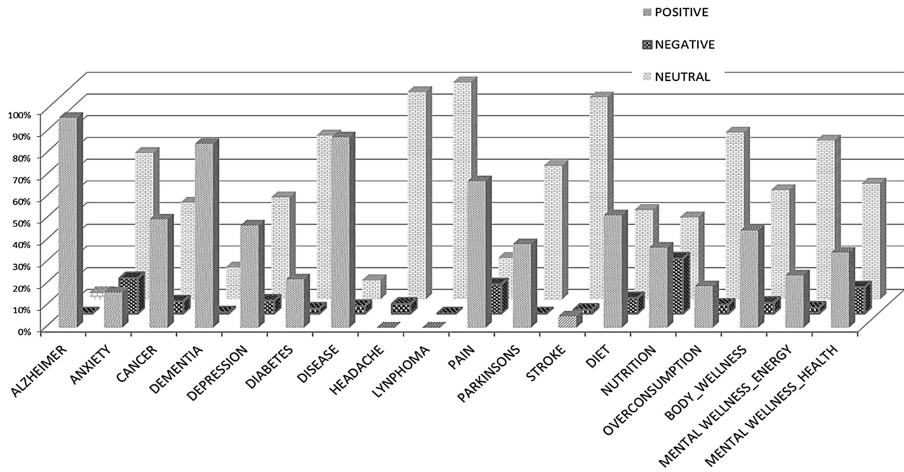


Figure 3. Sentiment analysis of tweets mentioning a health category

therefore maintaining an explanatory focus, rather than expressing feelings on news. This is confirmed by the analysis of the concentration of positive messages per categories of users (Table 4). There is a slightly higher concentration of positive messages among single persons (Gini coefficient equal to 0.51), compared to the other types of users.

The second sentiment analysis shows that the perception of messages on coffee and health with specific focus on a health category greatly varies depending on the single health category (Figure 3). Users positively perceive some of the health categories, whereas others are perceived negatively and neutrally. There is an overall positive perception towards the connection between coffee and health with Alzheimer, dementia, diseases, pain, diet and body wellness. There is negative perception towards the connection between coffee and health with lymphoma, Parkinson’s, headache, stroke, overconsumption, wellness and energy. For the following categories, there is a good balance between positive messages and neutral messages: cancer, depression, nutrition, body wellness and mental health. Appendix 3 provides an example of the content of positive and negative tweet messages.

6. Discussion

The present research explains the results according to the utilitarian and hedonic theory. This paper highlights the relationships between the consumer-perceived beliefs of utilitarian functional attributes and hedonic value of coffee product attribute. The utilitarian value perception is considered as health-related, pertaining to consumer value perceptions of coffee as beneficial or detrimental to health. The hedonic motivation suggests a coffee consumption experience based on a consumer’s positive response associated with product consumption, including enjoyment and social appreciation.

Table 4. Positive sentiment concentration per type of user

Type of users	Gini’s coefficient
Media	0.33
Companies	0.25
Persons	0.51

The present study contributes to the literature by enriching research work on coffee product attribute beliefs and value perceptions. Previous studies suggested that coffee consumers' perceived utilitarian value is predominantly driven by functional and economic attribute perceptions, whereas perceived hedonic value is dominated by content sensory attribute beliefs of coffee (Wang and Yu, 2016; Samoggia and Riedel, 2018). Limited attention has been given to specific coffee health attribute as key element of consumer-perceived utilitarian and hedonic coffee consumption attitude.

Research results show that consumers mainly have a hedonic approach to coffee drinking. Coffee arouses emotions and creates pleasant experiences. It is perceived as an enjoyable drinking habit and part of a trendy lifestyle. The way social media users select hashtags shows that they aim at drinking coffee in a fashionable and stylish manner. It fits with beauty, fitness, travel and summer. Coffee brings favourable emotions, wellness and positive state of mind. Past research findings support this outcome (Samoggia and Riedel, 2018; Bhumiratana *et al.*, 2014; Agoston *et al.*, 2018). Coffee is considered nutritionally good, contributing to weight loss within a low-calorie diet and bringing energy to the drinkers. It provides a hedonic, indulgent and light drinking experience. The overall coffee image is trendy, beneficial for body and mind and gratifying. There is limited interest in focussing on specific medical health effects of coffee. Moreover, in addition to the many indifferent consumers, there are also many positively inclined messages towards the potential positive health effects of coffee. These consumers associate the health attribute of coffee as a utilitarian value. The utilitarian value derives from the health benefits that coffee provides. These consumers provide a rational explanation to coffee consumption. Thus, results support that coffee can be positioned in consumers' mind for its health benefits. This type of positioning strategy would aim at satisfying a consumers' utilitarian need. The insights offered by the present study complement past research findings supporting the uncertainty consumers have on coffee health effects (ISIC, 2016; Corso *et al.*, 2018; Corso and Benassi, 2015).

Furthermore, this research shows that the issue of coffee and health is of interest for the wide public and for media institutions. Media institutions, through their social media accounts and the high number of followers, can exert great influence on public perception of coffee effects on health. This is due to the fact that consumers' accessibility and exposure to media content is higher, compared to scientific publications. As supported by past studies, consumers may more easily trust media messages that echo science-based information and research findings, rather than companies' communication and promotional messages (Weingart and Guenther, 2016).

The perception of consumers towards some health effects of coffee is positive. Tweets' sentiments towards coffee and health suggest that consumers can be receptive to a utilitarian health-oriented product positioning of coffee products. This is supported by past studies. Coffee is already positioned in consumers' minds to be a beverage that can boost their energy during a stressful workday and facilitate a relaxing afternoon among friends (Samoggia and Riedel, 2018). As past studies on consumers' behaviour support, consumers are still uncertain about the health benefits of coffee. Medical studies' findings on positive effects of coffee on health are shared on Twitter. However, the present research findings support that Twitter users use the hashtags cautiously. In particular, they tend not to expand their visibility pointing on the positive relation between coffee and medical health. They mostly adopt an emotional and hedonic communication style.

Furthermore, the present research highlights the value of social media as sources of information for research studies. Social media users can be considered communities of food consumers useful to extrapolate information on food product attribute perception. Social media users post, comment on and assess innovative products and food attributes (Dong and Wu, 2015). As supported by past studies, Twitter users' messages are rich in information and provide preliminary cues on consumers' perceptions that are worth exploring to

conceptualize food product attributes. There are thousands of tweets posted daily, and they picture consumers' sentiment towards a specific topic (Montoyo *et al.*, 2012; Mostafa, 2018).

Companies may consider Twitter and Facebook messages as an electronic form of the traditional word-of-mouth marketing technique (Mostafa, 2018). The present research shows that social media can be used as an extensive focus group useful for marketing and consumer research. By examining the content of tweets, companies can gain insights on consumers' attitudes about a specific issue or product (Simeone and Russo, 2017). Social media information can be used to explore the consumer-perceived utilitarian and hedonic value of a specific food product. Past research studies have examined social media to predict consumer behaviour (Alalwan *et al.*, 2017; Hamilton *et al.*, 2016; Erkan and Evans, 2016; Patino and Pitta, 2012; Wang *et al.*, 2012; Zhu and Chen, 2015). Others focussed on Twitter data to analyse the launch of a new product and to explore consumers' opinions (Rathore and Ilavarasan, 2017). Another study analysed how consumers reacted to existing products and gave suggestions for new products (Martini *et al.*, 2013). The present research uses social media to explore the users' utilitarian and hedonic perception of a specific food product attribute. It is an application of social media content and sentiment analysis and provides results valuable for implementing a utilitarian or hedonic benefits marketing strategy of coffee product.

7. Conclusion

The investigation highlights the potential of social media as a source of information for marketing purposes, and it explores how communication contents could drive food and beverage companies in developing innovative positioning strategies for coffee products. Companies may pay attention to social media content for a better understanding of consumers' perception or reaction to a given message. Social media are used to achieve successful branding and as part of wider advertising campaigns. They should be exploited also to establish a dialogue with consumers. They may provide scientifically informed evidence on recent research updates on the effects of coffee on health. Social media should become an integral part of the overall proactive marketing strategy of any company, especially if pursuing innovation.

Furthermore, the present research provides a direction on what aspects of coffee and health a communication strategy should focus on. The focus should be on the coffee effects on state of mind or diseases related to brain functioning. If coffee's beneficial effects on health are scientifically proved, the coffee industry will have an opportunity to further develop its marketing strategy. This could value consumers' interest towards health and wellness beverages and a trendy lifestyle.

Research results support the effectiveness of the utilitarian and hedonic consumer behaviour theory approach to analyse the consumers' perception of coffee health benefits gathered through social media data. Results also support that consumers' coffee experience is driven by a complex set of factors, and there is need to adequately capture the different dimensions. Results further show the proximity of utilitarian and hedonic values of consumers' perception of coffee.

7.1 Managerial implications

There is increasing perception on the positive effects of coffee on health. However, it is not yet the mainstream opinion that the consumers have. The coffee industry needs to invest on disseminating information and knowledge about coffee as part of a healthy lifestyle. Consumers must be educated and made aware of the most recent scientific findings. In this perspective, national coffee associations and independent research institutes worldwide increasingly point out the significant healthful properties of coffee. They emphasize that it is important to provide the consumers with accurate information about the health effects of coffee (Institute for Scientific Information on Coffee - ISIC, 2016; NCA, 2018b).

Consumers are interested in food and beverages that can increase their health and wellness. The coffee industry should invest to cater to this trend. As much as many other food and drinks, coffee beneficial effects could be further strengthened by adding nutrients that can add health benefits (GlobalData, 2018). This can expand the coffee appeal for the consumers. In the last decade, the coffee industry has strongly invested in coffee pods, new formats, roastery coffee and ready-to-drink products and by adding attributes such as sustainability and speciality coffee. Some of these products are reaching the maturity stage, and an increasing number of competitors have entered the market. Other coffee products can further expand their sales (ICO, 2018; NCA, 2018a).

The most promising innovation for functional coffee revolves around energy enhancement, building on coffee's traditional strength. In particular, the coffee industry is investing in the following benefits: mental energy enhancement, brain power enhancement, weight management benefits and body health enhancer suitable for the gym. Currently, there are current coffee products positioned with a health benefit. In particular, there are ready-to-drink coffee products (i.e. packaged liquid coffee designed to be consumed instantaneously when opened without any additional steps), prebiotic fibre-enriched ground coffees with digestive health benefits and with antioxidant-enriched ground coffees (Samoggia and Riedel, 2019). To fully exploit the potential of coffee health benefits, it should be scientifically proven that coffee consumption is beneficial. Thus, the coffee industry should support independent scientific research on coffee and health and promote the active dissemination of balanced coffee and health scientific evidence and knowledge to a broad range of stakeholders. The food industry may promote a new wave of coffee consumption. It could be based on new ways to approach the consumers and take market shares belonging to other food and beverage industries (GlobalData, 2018).

Furthermore, latest data show that the average age of coffee drinkers is diminishing, and increasingly young consumers drive coffee sales (NCA, 2018a). These already are a key consumer target and represent the next generation of everyday coffee drinkers. Educating them to a responsible and quality coffee drinking experience, without exceeding the recommended daily amount, can be in the interest of the coffee industry. Recent studies show that for many millennials, drinking coffee is an emotional experience (NCA, 2018a). Coffee should be fresh, affordable and socially and environmentally respectful in how coffee beans are grown and harvested (Samoggia and Riedel, 2018). Given the high use of social media by the younger generations, this communication channel can be potentially useful for innovative marketing strategies to achieve this target public. The communication content could focus on the health benefits of coffee.

7.2 Research limitations and future research

The set of collected data shows that there was a significant tweeting and retweeting activity by a high number of users. This compensates the definite amount of time of data collection. The high number of users identified provides information on the perception of a wide audience of people. This compensates Twitter messages' limited number of characters that may hinder the full representation of the sentiment towards a specific topic. Moreover, the authors' preliminary work on messages analysis targeted the computer-limited capability to fully understand the meaning of some messages, puns, ironic content or word double meaning. This allowed the creation of a robust set of data.

Future studies may extend the analysis on the temporal pattern of the tweets' sentiment during the day. Furthermore, social media analysis provides preliminary insights over consumers' perceptions on coffee health attributes. The present research results can provide a basis for future qualitative research studies, based on surveys, consumers' interviews and observations. These may provide a more detailed and nuanced understanding of consumers'

perceptions. Finally, future studies based on social media data may opt for a netnographic approach (Kozinets, 2002, 2015), which, similar to the present research, can use data generated from online community messages. Indeed, a netnographic research approach may capture dialogues among social media users, by immersing in the net community and providing members with opportunities for feedback.

References

- Ágoston, C., Urbán, R., Király, O., Griffiths, M.D., Rogers, P.J. and Demetrovics, Z. (2018), "Why do you drink caffeine? The development of the motives for caffeine consumption questionnaire (MCCQ) and its relationship with gender, age and the types of caffeinated beverages", *International Journal of Mental Health and Addiction*, Vol. 16, pp. 981-999, doi: [10.1007/s11469-017-9822-3](https://doi.org/10.1007/s11469-017-9822-3).
- Alalwan, A.A., Rana, N.P., Dwivedi, Y.K. and Algharabat, R. (2017), "Social media in marketing: a review and analysis of the existing literature", *Telematics and Informatics*, Vol. 34 No. 7, pp. 1177-1190, doi: [10.1016/j.tele.2017.05.008](https://doi.org/10.1016/j.tele.2017.05.008).
- Bhumiratana, N., Adhikari, K. and Chambers, E. (2014), "The development of an emotion lexicon for the coffee drinking experience", *Food Research International*, Vol. 61, pp. 83-92, doi: [10.1016/j.foodres.2014.03.008](https://doi.org/10.1016/j.foodres.2014.03.008).
- Boiy, E. and Moens, M. (2009), "A machine learning approach to sentiment analysis in multi-lingual Web texts", *Information Retrieval*, Vol. 12, pp. 526-558, doi: [10.1007/s10791-008-9070-z](https://doi.org/10.1007/s10791-008-9070-z).
- Booth, N. and Matic, J.A. (2011), "Mapping and leveraging influencers in social media to shape corporate brand perceptions", *Corporate Communications*, Vol. 16 No. 3, pp. 184-191, doi: [10.1108/13563281111156853](https://doi.org/10.1108/13563281111156853).
- Chae, B. (2015), "Insights from hashtag #supplychain and Twitter analytics: considering Twitter and Twitter data for supply chain practice and research", *International Journal of Production Economics*, Vol. 165, pp. 247-259, doi: [10.1016/j.ijpe.2014.12.037](https://doi.org/10.1016/j.ijpe.2014.12.037).
- Chaudhuri, A. and Holbrook, M.B. (2001), "The chain of effects from brand trust and brand affect to brand performance: the role of brand loyalty", *Journal of Marketing*, Vol. 65 No. 2, pp. 81-93.
- Chen, Z. and Dubinsky, A.J. (2003), "A conceptual model of perceived customer value in e-commerce: a preliminary investigation", *Psychology and Marketing*, Vol. 20 No. 4, pp. 323-347.
- Chen, H. and Sun, T.-H. (2014), "Clarifying the impact of product scarcity and perceived uniqueness in buyers' purchase behavior of games of limited-amount version", *Asia Pacific Journal of Marketing and Logistics*, Vol. 26 No. 2, pp. 232-249.
- Chua, A.Y.K. and Banerjee, S. (2013), "Customer knowledge management via social media: the case of Starbucks", *Journal of Knowledge Management*, Vol. 17 No. 2, pp. 237-249, doi: [10.1108/1367327131131519](https://doi.org/10.1108/1367327131131519).
- Corso, M.P. and Benassi, M.D.T. (2015), "Packaging attributes of antioxidant-rich instant coffee and their influence on the purchase intent", *Beverages*, Vol. 1, pp. 273-291.
- Corso, M., Kalschne, D. and Benassi, M. (2018), "Consumer's attitude regarding soluble coffee enriched with antioxidants", *Beverages*, Vol. 4 No. 4, p. 72, doi: [10.3390/beverages4040072](https://doi.org/10.3390/beverages4040072).
- Cronin, J.J., Brady, M.K. and Hult, G.T. (2000), "Assessing the effects of quality, value, and customer satisfaction on consumer behavioral intentions in service environments", *Journal of Retailing*, Vol. 76 No. 2, pp. 193-218.
- Dodds, P.S., Harris, K.D., Kloumann, I.M., Bliss, C.A. and Danforth, C.M. (2011), "Temporal patterns of happiness and information in a global social network: hedonometrics and twitter", *PLoS One*, Vol. 6 No. 12, p. e26752, doi: [10.1371/journal.pone.0026752](https://doi.org/10.1371/journal.pone.0026752).
- Dong, J.Q. and Wu, W. (2015), "Business value of social media technologies: evidence from online user innovation communities", *The Journal of Strategic Information Systems*, Vol. 24 No. 2, pp. 113-127, doi: [10.1016/j.jsis.2015.04.003](https://doi.org/10.1016/j.jsis.2015.04.003).

- Dubè, L., Du, P., Mcrae, C., Sharma, N., Jayaraman, S. and Nie, J.-Y. (2018), "Convergent innovation in food through big data and artificial intelligence for societal-scale inclusive growth", *Technology Innovation Management Review*, Vol. 8 No. 2, pp. 49-66, doi: [10.22215/timreview/1139](https://doi.org/10.22215/timreview/1139).
- Dwivedi, Y., Kapoor, K. and Chen, H. (2015), "Social media marketing and advertising", *The Marketing Review*, Vol. 15 No. 3, pp. 289-309, doi: [10.1362/146934715X14441363377999](https://doi.org/10.1362/146934715X14441363377999).
- Erkan, I. and Evans, C. (2016), "The influence of eWOM in social media on consumers' purchase intentions: an extended approach to information adoption", *Computers in Human Behavior*, Vol. 61, pp. 47-55, doi: [10.1016/j.chb.2016.03.003](https://doi.org/10.1016/j.chb.2016.03.003).
- Euromonitor International (2017), *Repositioning Indulgence to Remain Relevant in the Healthy Snacking Era*, London, available at: <https://www.euromonitor.com/repositioning-indulgence-to-remain-relevant-in-the-healthy-snacking-era/report>.
- European Food Safety Authority [EFSA] (2015), "Scientific opinion on the safety of caffeine", *EFSA Journal*, Vol. 13 No. 5, p. 4102, doi: [10.2903/j.efsa.2015.4102](https://doi.org/10.2903/j.efsa.2015.4102).
- Filieri, R. (2013), "Consumer co-creation and new product development: a case study in the food industry", *Marketing Intelligence and Planning*, Vol. 31 No. 1, pp. 40-53, doi: [10.1108/02634501311292911](https://doi.org/10.1108/02634501311292911).
- Friedrich, N., Bowman, T.D., Stock, W.G. and Haustein, S. (2015), "Adapting sentiment analysis for tweets linking to scientific papers", *Paper Presented at the 15th International Society of Scientometrics and Informetrics Conference*, Istanbul, Turkey.
- GlobalData (2018), *Functional Coffee Is Closer to Breaking Out as the Next Trendy Beverage Sector*, GlobalData PLC, London.
- Graham, M., Hale, S.A. and Gaffney, D. (2014), "Where in the world are you? Geolocation and language identification in Twitter", *The Professional Geographer*, Vol. 66 No. 4, pp. 568-578, doi: [10.1080/00330124.2014.907699](https://doi.org/10.1080/00330124.2014.907699).
- Grosso, G., Godos, J., Galvano, F. and Giovannucci, E.L. (2017), "Coffee, caffeine, and health outcomes: an Umbrella Review", *Annual Review of Nutrition*, Vol. 37 No. 1, pp. 131-156, doi: [10.1146/annurev-nutr-071816-064941](https://doi.org/10.1146/annurev-nutr-071816-064941).
- Hamilton, W.L., Clark, K., Leskovec, J. and Jurafsky, D. (2016), "Inducing domain-specific sentiment lexicons from unlabeled corpora", *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 595-605, doi: [10.18653/v1/D16-1057](https://doi.org/10.18653/v1/D16-1057).
- Hayes, A.F. (2005), *Statistical Methods for Communication Science*, Lawrence Erlbaum Associates, Mahwah, NJ.
- He, W., Wu, H., Yan, G., Akula, V. and Shen, J. (2015), "A novel social media competitive analytics framework with sentiment benchmarks", *Information and Management*, Vol. 52 No. 7, pp. 801-812, doi: [10.1016/j.im.2015.04.006](https://doi.org/10.1016/j.im.2015.04.006).
- He, W., Zha, S. and Li, L. (2013), "Social media competitive analysis and text mining: a case study in the pizza industry", *International Journal of Information Management*, Vol. 33, pp. 464-472, doi: [10.1016/j.ijinfomgt.2013.01.001](https://doi.org/10.1016/j.ijinfomgt.2013.01.001).
- Holbrook, M. (2006), "Consumption experience, customer value, and subjective personal introspection: an illustrative photographic essay", *Journal of Business Research*, Vol. 59 No. 6, pp. 714-725, doi: [10.1016/j.jbusres.2006.01.008](https://doi.org/10.1016/j.jbusres.2006.01.008).
- Institute for Scientific Information on Coffee [ISIC] (2016), "The good things in life: coffee as part of a healthy diet and lifestyle", Roundtable Report, available at: https://www.coffeeandhealth.org/wp-content/uploads/2016/03/Roundtable-report_Coffee-as-part-of-a-healthy-diet.pdf.
- International Coffee Organization [ICO] (2018), *Monthly Coffee Market Report*, ICO, London, available at: <http://www.ico.org/documents/cy2017-18/cmr-0918-e.pdf>.
- Jiang, Q. and Leung, L. (2012), "Lifestyles, gratifications sought, and narrative appeal: American and Korean TV drama viewing among Internet users in urban China", *International Communication Gazette*, Vol. 74 No. 2, pp. 159-180, doi: [10.1177/1748048511432601](https://doi.org/10.1177/1748048511432601).
- Jick, H., Vessey, M.P., Slone, D., Shapiro, S., Heinonen, O.P. and Miettinen, O.S. (1974), "Coffee drinking and myocardial infarction", *Journal of the American Medical Association*, Vol. 227 No. 7, pp. 801-802, doi: [10.1001/jama.1974.03230200059021](https://doi.org/10.1001/jama.1974.03230200059021).

- Just-drinks.com (2018), available at: https://www.just-drinks.com/analysis/how-rtd-coffee-could-be-the-next-functional-drink-boom-area-npd-trends_id126604.aspx (accessed 6 June 2019).
- Kapoor, K.K., Tamilmani, K., Rana, N.P., Patil, P., Dwivedi, Y.K. and Nerur, S. (2017), "Advances in social media research: past, present and future", *Information Systems Frontiers*, Vol. 20 No. 3, pp. 531-558, doi: [10.1007/s10796-017-9810-y](https://doi.org/10.1007/s10796-017-9810-y).
- Kim, Y., Kang, S., Kim, T., Choi, Y., Park, M. and Choi, S. (2015), "Sentiment analysis of consumer opinion in blogs: a case study in Ramen market", in Leung, C.K. and Nasridinov, A. (Eds), *BigDAS'15, ACM International Conference Proceeding Series*, ACM The Association for Computing Machinery, New York, pp. 250-253, doi: [10.1145/2837060.2837107](https://doi.org/10.1145/2837060.2837107).
- Kozinets, R.V. (2002), "The field behind the screen: using netnography for marketing research in online communities", *Journal of Marketing Research*, Vol. 39 No. 1, pp. 61-72.
- Kozinets, R.V. (2015), *Netnography: Redefined*, Sage, Los Angeles, CA.
- Krippendorff, K. (2018), *Content Analysis: An Introduction to its Methodology*, 4th ed., Sage, Los Angeles, CA.
- Lamy, F., Daniulaityte, R., Sheth, A., Nahhas, R., Martins, S., Boyer, E. and Carlson, R. (2016), "Those edibles hit hard: exploration of Twitter data on cannabis edibles in the US", *Drug and Alcohol Dependence*, Vol. 164, pp. 64-70, doi: [10.1016/j.drugalcdep.2016.04.029](https://doi.org/10.1016/j.drugalcdep.2016.04.029).
- Landry, R. (2003), "Investigacao social", in Gauthier, B. (Ed.), *A analise de conteudo*, Lusociencia, Loures, Portugal, pp. 345-370.
- Liu, Y. and Lopez, R.A. (2016), "The impact of social media conversations on consumer brand choices", *Marketing Letters*, Vol. 27 No. 1, pp. 1-13, doi: [10.1007/s11002-014-9321-2](https://doi.org/10.1007/s11002-014-9321-2).
- Liu, X., Burns, A.C. and Hou, Y. (2017). "An investigation of brand-related user-generated content on Twitter", *Journal of Advertising*, Vol. 46 No. 2, pp. 236-247, doi: [10.1080/00913367.2017.1297273](https://doi.org/10.1080/00913367.2017.1297273).
- Lofffield, E., Freedman, N.D., Dodd, K.W., Vogtmann, E., Xiao, Q., Sinha, R. and Graubard, B.I. (2016), "Coffee drinking is widespread in the United States, but usual intake varies by key demographic and lifestyle factors", *Journal of Nutrition*, Vol. 146 No. 9, pp. 1762-1768, doi: [10.3945/jn.116.233940](https://doi.org/10.3945/jn.116.233940).
- Lu, J., Liu, Z. and Fang, Z. (2016), "Hedonic products for you, utilitarian products for me", *Judgment and Decision making*, Vol. 11 No. 4, pp. 332-341.
- Macnamara, J.R. (2005), "Media content analysis: it's uses, benefits and best practice methodology", *Asia Pacific Public Relations Journal*, Vol. 6 No. 1, pp. 1-34, available at: <https://opus.lib.uts.edu.au/handle/10453/10102>.
- Mäntylä, M.V. and Graziotin, D. (2018), "The evolution of sentiment analysis: a review of research topics, venues, and top cited papers", *Computer Science Review*, Vol. 27, pp. 16-32, doi: [10.1016/j.cosrev.2017.10.002](https://doi.org/10.1016/j.cosrev.2017.10.002).
- Martínez-Cámara, E., Martín-Valdivia, M.T., Ureña López, L.A. and Montejó-Ráez, A. (2012), "Sentiment analysis in twitter", *Natural Language Engineering*, Vol. 20 No. 1, pp. 1-28, doi: [10.1017/S1351324912000332](https://doi.org/10.1017/S1351324912000332).
- Martini, A., Massa, S. and Testa, S. (2013), "The firm, the platform and the customer: a 'double mangle' interpretation of social media for innovation", *Information and Organization*, Vol. 23 No. 3, pp. 198-213, doi: [10.1016/j.infoandorg.2013.07.001](https://doi.org/10.1016/j.infoandorg.2013.07.001).
- Mikhaylov, S., Laver, M. and Benoit, R. (2012), "Coder reliability and misclassification in the human coding of party manifestos", *Political Analysis*, Vol. 20, pp. 78-91, doi: [10.1093/pan/mpr047](https://doi.org/10.1093/pan/mpr047).
- Mitchell, D.C., Knight, C.A., Hockenberry, J., Teplansky, R. and Hartman, T.J. (2014), "Beverage caffeine intakes in the US", *Food Chem. Toxicology*, Vol. 63, pp. 136-142, doi: [10.1016/j.fct.2013.10.042](https://doi.org/10.1016/j.fct.2013.10.042).
- Montoyo, A., Martiniz-Barco, P. and Balahur, A. (2012), "Subjectivity and sentiment analysis: an overview of the current state of the area and envisaged developments", *Decision Support Systems*, Vol. 53, pp. 675-679, doi: [10.1016/j.dss.2012.05.022](https://doi.org/10.1016/j.dss.2012.05.022).

- Mostafa, M.M. (2013), "More than words: social networks' text mining for consumer brand sentiments", *Expert Systems with Applications*, Vol. 40 No. 10, pp. 4241-4251, doi: [10.1016/j.eswa.2013.01.019](https://doi.org/10.1016/j.eswa.2013.01.019).
- Mostafa, M.M. (2017), "Mining and mapping halal food consumers: a geo-located Twitter opinion polarity analysis", *Journal of Food Products Marketing*, Vol. 23, pp. 1-22, doi: [10.1080/10454446.2017.1418695](https://doi.org/10.1080/10454446.2017.1418695).
- Mostafa, M.M. (2018), "Clustering halal food consumers: a Twitter sentiment analysis", *International Journal of Market Research*, doi: [10.1177/1470785318771451](https://doi.org/10.1177/1470785318771451).
- Naaman, M., Boase, J. and Lai, C.H. (2010), "Is it really about me?", in Inkpen, K. and Gutwin, C. (Eds), *CSCW'10 Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*, Savannah, Georgia, ACM The Association for Computing Machinery, New York, pp. 189-192.
- National Coffee Association [NCA] (2018a), *National Coffee Drinking Trends Report*, New York, available at: <https://www.ncausa.org/Industry-Resources/Market-Research/NCDT>.
- National Coffee Association [NCA] (2018b), "Scientific evidence is clear: drinking coffee is good for you", [NCA Newsroom articles] available at: http://www.ncausa.org/Portals/56/PDFs/Communication/20180816_NCANews_scienceisclear.pdf?ver=2018-08-16-122429-123.
- Nawrot, P., Jordan, S., Eastwood, J., Rotstein, J., Hugenholtz, A. and Feeley, M. (2003), "Effects of caffeine on human health", *Food Additives and Contaminants*, Vol. 20 No. 1, pp. 1-30, doi: [10.1080/0265203021000007840](https://doi.org/10.1080/0265203021000007840).
- Nieber, K. (2017), "The impact of coffee on health", *Planta Medica*, Vol. 83 No. 16, pp. 1256-1263, doi: [10.1055/s-0043-115007](https://doi.org/10.1055/s-0043-115007).
- Nomisma (2019), *Osservatorio Caffè*, available at: <http://www.datalytics.it/osservatoriocaff%C3%A8.html>.
- Pang, B. and Lee, L. (2008), "Opinion mining and sentiment analysis", *Foundations and Trends in Information Retrieval*, Vol. 2 Nos 1/2, pp. 1-135, doi: [10.1561/15000000011](https://doi.org/10.1561/15000000011).
- Panhuysen, S. and Pierrot, J. (2018), *Coffee Barometer 2018*, Hivos, The Hague, available at: <https://hivos.org/assets/2018/06/Coffee-Barometer-2018.pdf>.
- Patino, A. and Pitta, D.A. (2012), "Social media's emerging importance in market research", *Journal of Consumer Marketing*, Vol. 29 No. 3, pp. 233-237, doi: [10.1108/07363761211221800](https://doi.org/10.1108/07363761211221800).
- Perera, I. and Caldera, H.A. (2017), "Aspect based opinion mining from restaurant reviews", in *Proceedings of the 2nd IEEE International Conference on Computational Intelligence and Applications (ICCA)*, IEEE Institute of Electrical and Electronics Engineers, Piscataway, NJ, pp. 542-546.
- Poole, R., Kennedy, O.J., Roderick, P., Fallowfield, J.A., Hayes, P.C. and Parkes, J. (2017), "Coffee consumption and health: umbrella review of meta-analyses of multiple health outcomes", *BMJ*, Vol. 359, p. j5024, doi: [10.1136/bmj.j5024](https://doi.org/10.1136/bmj.j5024).
- Rathore, A.K. and Ilavarasan, P.V. (2017), "Social media analytics for new product development: case of a pizza", in Agarwal, P.K. and Gupta, M. (Eds.), *Proceedings of the International Conference on Advances in Mechanical, Industrial, Automation and Management Systems*, IEEE Institute of Electrical and Electronics Engineers, Danvers, MA, pp. 213-219, doi: [10.1109/AMIAMS.2017.8069214](https://doi.org/10.1109/AMIAMS.2017.8069214).
- Rathore, A.K., Ilavarasan, P.V. and Dwivedi, Y.K. (2016), "Social media content and product creation: an emerging paradigm", *Journal of Enterprise Information Management*, Vol. 29 No. 1, pp. 7-18, doi: [10.1108/JEIM-06-2015-0047](https://doi.org/10.1108/JEIM-06-2015-0047).
- Ravaglia, V., Zanazzi, L. and Mazzoni, E. (2016), "Network text analysis and sentiment analysis: an integration to analyze word-of-mouth in the digital marketplace", in Rathore, S. and Panwar, A. (Eds), *Capturing, Analyzing, and Managing Word-of-Mouth in the Digital Marketplace*, IGI Global, Hershey, PA, pp. 137-153.
- Ribeiro, F.N., Araújo, M., Gonçalves, P., Gonçalves, M.A. and Benevenuto, F. (2016), "SentiBench - a benchmark comparison of state-of-the-practice sentiment analysis methods", *EPJ Data Science*, Vol. 5 No. 1, pp. 1-29, doi: [10.1140/epjds/s13688-016-0085-1](https://doi.org/10.1140/epjds/s13688-016-0085-1).

- Robertson, D., Frolich, J.C., Carr, R.K., Watson, J.T., Hollifield, J.W., Shand, D.G. and Oates, J.A. (1978), "Effects of caffeine on plasma renin activity, catecholamines and blood pressure", *New England Journal of Medicine*, Vol. 298, pp. 181-186.
- Ruggeri, A. and Samoggia, A. (2018), "Twitter communication of agri-food chain actors on palm oil environmental, socio-economic, and health sustainability", *Journal of Consumer Behaviour*, Vol. 17 No. 1, pp. 75-93, doi: [10.1002/cb.1699](https://doi.org/10.1002/cb.1699).
- Samoggia, A. and Riedel, B. (2018), "Coffee consumption and purchasing behavior review: insights for further research", *Appetite*, Vol. 129, pp. 70-81, doi: [10.1016/j.appet.2018.07.002](https://doi.org/10.1016/j.appet.2018.07.002).
- Samoggia, A. and Riedel, B. (2019), "Consumers' perception of coffee health benefits and motives for coffee consumption and purchasing", *Nutrients*, Vol. 11, Article: 653, pp. 1-21, doi: [10.3390/nu11030653](https://doi.org/10.3390/nu11030653).
- Simeone, M. and Russo, C. (2017), "The growing influence of social and digital media", *British Food Journal*, Vol. 119 No. 8, pp. 1766-1780, doi: [10.1108/BFJ-05-2017-0283](https://doi.org/10.1108/BFJ-05-2017-0283).
- Singh, A., Shukla, N. and Mishra, N. (2016), "Social media data analytics to improve supply chain management in food industries", *Transportation Research Part E: Logistics and Transportation Review*, Vol. 114, pp. 398-415, doi: [10.1016/j.tre.2017.05.008](https://doi.org/10.1016/j.tre.2017.05.008).
- Stieglitz, S., Dang-Xuan, L., Bruns, A. and Neuberger, C. (2014), "Social media analytics", *Business and Information Systems Engineering*, Vol. 6 No. 2, pp. 89-96, doi: [10.1007/s12599-014-0315-7](https://doi.org/10.1007/s12599-014-0315-7).
- Tausczik, Y.R. and Pennebaker, J.W. (2010), "The psychological meaning of words: LIWC and computerized text analysis methods", *Journal of Language and Social Psychology*, Vol. 29 No. 1, pp. 24-54, doi: [10.1177/0261927X09351676](https://doi.org/10.1177/0261927X09351676).
- Thelwall, M. (2017), "The heart and soul of the web? Sentiment strength detection in the social web with SentiStrength", in Janusz, A.H. (Ed.), *Cybermotions. Collective Emotions in Cyberspace*, Springer International Publishing, Switzerland, pp. 119-134.
- Thelwall, M., Buckley, K. and Paltoglou, G. (2011), "Sentiment in twitter events", *Journal of the American Society for Information Science and Technology*, Vol. 62 No. 2, pp. 406-418.
- Wang, E.S.T. and Yu, J.R. (2016), "Effect of product attribute beliefs of ready-to-drink coffee beverages on consumer-perceived value and repurchase intention", *British Food Journal*, Vol. 118 No. 12, pp. 2963-2980, doi: [10.1108/BFJ-03-2016-0128](https://doi.org/10.1108/BFJ-03-2016-0128).
- Wang, X., Yu, C. and Wei, Y. (2012), "Social media peer communication and impacts on purchase intentions: a consumer socialization framework", *Journal of Interactive Marketing*, Vol. 26 No. 4, pp. 198-208, doi: [10.1016/j.intmar.2011.11.004](https://doi.org/10.1016/j.intmar.2011.11.004).
- Weingart, P. and Guenther, L. (2016), "Communicating science: trust and the new media ecology. Science communication and the issue of trust", *Journal of Science Communication*, Vol. 15 No. 5, pp. 1-11, doi: [10.22323/2.15050301](https://doi.org/10.22323/2.15050301).
- Wikoff, D., Welsh, B.T., Henderson, R., Brorby, G.P., Britt, J., Myers, E., Goldberger, J., Lieberman, H.R., O'Brien, C., Peck, J., Tenenbein, M., Weaver, C., Harvey, S., Urban, J. and Doepker, C. (2017), "Systematic review of the potential adverse effects of caffeine consumption in healthy adults, pregnant women, adolescents, and children", *Food and Chemical Toxicology*, Vol. 109, pp. 585-648, doi: [10.1016/j.fct.2017.04.002](https://doi.org/10.1016/j.fct.2017.04.002).
- Yu, Y., Duan, W. and Cao, Q. (2013), "The impact of social and conventional media on firm equity value: a sentiment analysis approach", *Decision Support Systems*, Vol. 55 No. 4, pp. 919-926, doi: [10.1016/j.dss.2012.12.028](https://doi.org/10.1016/j.dss.2012.12.028).
- Zhu, Y.-Q. and Chen, H.-G. (2015), "Social media and human need satisfaction: implications for social media marketing", *Business Horizons*, Vol. 58 No. 3, pp. 335-45, doi: [10.1016/j.bushor.2015.01.006](https://doi.org/10.1016/j.bushor.2015.01.006).

Appendix

Meal #lunch, lunch #meals, meal, meals #morning, #mornings, morn', morning, morning', mornings, #break, break, breaks, #coffeetime, #breakfast, breakfast #coffeetime dinner, #dinner afternoon

Time #day, day, days, daily, #everyday, everyday friday, fridays week, weekly, weeks #summer, summer month, monthly, months #saturday, Saturday #monday, monday, Mondays

Food categories #food, food, foods #sugar, sugar, sugars #water, @waters, water #chocolate, #chocolates, chocolate, chocolates #milk, milk, milks #coconut, coconut oil, oils salt #pastry, pastry #cinnamon, cinnamon #herbal, #herbal', herbal, 'herbal' #beer, beer #nut, nut, nuts muffin, muffins #wine, wine #cake, cake, cakes #alcohol, alcohol spice, spiced, spices #tea, tea, teas #vegan, @vegan, vegan, vegans #beverage, #beverages, beverage, beverages #drinks, drinks #sweetener, sweeteners organic #organic #smoothie, #smoothies, smoothie, smoothies, smoothy mushroom, butter, berry, berries, grapefruit* jackfruit, sorghum, charcoal, whey, bitter melon, flour, ginger, banana

Lifestyle #travel, travel #style, style, styles #fashion, fashion #work, #working, work, worked, working, works #sport, #sports, sports #habit, habit, habits #live, #living, live, lived, lives, living #active, activated, active, activities #natural, #naturally, #nature, natural, naturally, nature traditional #beach, beach #sea, seas #smoking, smoke, smokes, smoking #walk, #walking, walk, walked, walking, walks #chat, chat, chatted, chatting #lifestyle, #lifestyles, lifestyle #job, job, jobs conversation, conversations #yoga, yoga cigarette, cigarettes #paris, paris #healthyliving #healthtip, #healthtips #healthlife

Consumer group #aging, age, aged, ageing, aging teen, teens, boy, boys, #boy, #boys, #girls, girl, girls, #teenager, teenager #women, women, women' #men, men couple, coupled #friends, friend, friends #girls worker, workers #parenting, #parents, parental, parenting, parents, #mother, #mothers, mother #sex, sex

Body parts synonym #liver, liver #hair, hair #prostate, prostate #brain, @brain, brain head, heads #heart, heart stomach #teeth, teeth belly #body, bodies, body #skin, skin

Science-based aware, awareness #science, @science, science, 'science, sciences #scientists, scientist, scientists, bioscience* report, reported, reportedly, reporter, reporters, reporting, reports, #healthnews #doctor, doctor, doctors biosciences #healthnews #medicinal, #medicine, medicinal, medicine #medical, @medical, medical, medicate, medication, medications, medics

Sensorial taste, tastes #fresh, fresh, freshly flavor, flavored, flavors, flavour, flavoured, flavours #free, free #eat, #eating, eat, eating #consumers, consume, consumed, consumer, consumers, consumes, consuming #consumption, consumption #drink, #drinking, drink, drinking, 'drinking

Typology of coffee dark ice, iced, icing black #decaf, decaf creamy #espresso, @espresso, espresso instant, instantly #irish, irish #italian, italian, italians #green #greencoffe #greencoffee #latte, latte, latte', latte', lattes #brew, brew, brewed, brewing, brews #bean, #beans, bean, beans, "green coffee" "green coffee beans" "Green Coffee Bean"

BFJ
122,12

3834

Body wellness #gym, gym #fit, #fitness, fit, fitness #beauty, beauty #well, #wellness, wellness #workout, #workouts, workout, workouts #clarity, clarity #fitnessmotivation #scrub, scrub, scrubs treatment #skincare, skincare #exercise, exercise, exercises #wealth, wealth #longevity, longevity #performance, perform, performance, performing coffeescrub #coffeescrub, #coffeescrubs

Nutrition #nutrition, nutrition, nutritional #protein, protein #toxic, toxic acid, acidic, acidity, acids #ingredients, ingredient, ingredients #antioxidant, #antioxidants, antioxidant, antioxidants #digestion, #digestive, digest, digestive #detox, detox #healthiness, #healthy, healthy, healthy'

Diet intake #calorie, calorie, calories #diet, diet, dieting #weight, weight, weights #weightloss, weightloss

Overconsumption die, died, dies, dying #death, death, deaths #overdose, overdose, overdosed, overdoses #addiction, addiction, addicts

Mental wellness_energy #energy, energy energy, energize, energized, energizer, energizing power, powered, powerful, powering #energydrink, #energydrinks fatigue #superfood

Mental wellness_health #mentalhealth #brainhealth #happiness, #happy, happiness #stress, stress #sleep, sleep, sleeping #mental, mental #mind, mind, minds

Health diseases

Alzheimer

#alzheimers, Alzheimer

Anxiety

#anxiety, anxiety

Cancer

#cancer, cancer, cancer', cancer', cancers, #prostatecancer, #livercancer

Dementia

#dementia, dementia

Depression

#depression, depressed, depressing, depression

Diabete

#diabetes, diabetes, diabetic, diabetics

Disease

#disease, disease, diseases, sick, sickness

Headache

#headache, headache, headaches

Lymphoma

#lymphoma lymphoma

Pain

#pain, pain, #chronicpain #chronic

Parkinson

#parkinsons, parkinson, parkinsons

Stroke

#stroke, stroke, strokes

Appendix 2.
Final dictionary of
health-related
categories and
occurrences

Positive

Just #Health & #Wealth ~ Coffee for weight loss. Coffee help lose weight
 11 ways #caffeine can affect your well-being. #coffee
 This is why anyone who wants to burn belly fat should switch to dark roast coffee
 Coffee scrub with cinnamon 150 g #HealthEmbassy #coffeescrub #cellulite #body #scrub #health
 Drinking coffee for Health and Weight Loss
 Increase ur #focus, #clarity n #energy #superfood #coffee #health
 For long-term sustainability of coffee farms, physical and mental health of labourers is as important as crop health
 More good news about health benefits of coffee
 Drinking Italian-style Coffee may cut #prostatecancer risk in half. #prevention #healthtips #health
 Coffee can be good for your health!
 The many wonderful #health benefits of #coffee. #HealthyLiving #wellness #nutrition #energy #diabetes #depression #coffee
 Organic Coffee Co #health #shop
 Did you know that drinking 3 cups of #coffee daily can help prevent Alzheimer's?
 Is the cure for #chronicpain in #coffee? Why #caffeine helps treat chronic pain
 Athletes, never fear – you do not need to abstain from caffeine before a big event
 Drinking #coffee may help prevent the most common form of #liver cancer -
 More reasons to love coffee! Five cups a day may cut risk of liver cancer by half
 Protein Iced Coffee (Low Carb, Sugar Free, THM-S). A creamy iced coffee full of health-boosting benefits

Negative

World Health Organization to drop coffee's status as possible carcinogen
 Health fact: Your daily cup of coffee may be ageing your skin
 Coffee bad for your health
 I had to give up coffee three years ago also for health reasons. . . If I now taste it, it tastes awful
 Me realising I have to quit coffee because my health canno't deal. I need a win guys. I keep losing
 Summer is here and so is my mom's dire warning that coffee in summer is really bad for health
 Teenager died from caffeine OD after quaffing coffee, soda, energy drink
 Okay so health scare is not so much of a scare as it is a wakeup call, going to avoid spicy food, coffee, booze for a while
 #Coffee and hormones: How coffee really affects your health
 I always crave coffee at night and that's probably a very bad thing for my health
 Sigh I'm back to surviving on coffee and cigarettes. . . . I know it's awful for my health but I cannot bring myself to eat proper meals. yikes
 I drink a lot coffee it's almost everyday and now I'm worried about my future health
 Ei, stop drinking so much coffee. It's bad for your health
 Your Morning Coffee Is Missing This Key Ingredient - And It Might Be Harming Your Health
 Coffee is not good food
 Coffee is not beneficial to life
 Why is coffee legal? It's a drug. It's addictive. It's bad for my health. I drink WAY too much coffee
 Coffee is not the only thing staining your teeth! Watch out for these 7 other tricky culprits

Appendix 3.
 Example list of positive
 and negative tweet
 messages on coffee and
 health

Corresponding author

Antonella Samoggia can be contacted at: antonella.samoggia@unibo.it

For instructions on how to order reprints of this article, please visit our website:

www.emeraldgrouppublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com