INSTITUTO
UNIVERSITÁRIO
DE LISBOA

The relationship between Brand Coolness, Brand Love, Loyalty and e-WOM: A text mining and sentiment analysis approach focused on a tech brand (Apple)

João Pedro de Sousa Apolinário

Master (MSc) in Management

Supervisor: PhD Catarina Maria Valente Antunes Marques, Associate Professor

ISCTE – University Institute of Lisbon

Co-Supervisor: PhD Sérgio Miguel Carneiro Moro, Full Professor

ISCTE – University Institute of Lisbon

October 2023

BUSINESS
SCHOOL

Department of Marketing, Operations and Management

The relationship between Brand Coolness, Brand Love, Loyalty and e-WOM: A text mining and sentiment analysis approach focused on a tech brand (Apple)

João Pedro de Sousa Apolinário

Master (MSc) in Management

Supervisor: PhD Catarina Maria Valente Antunes Marques, Associate Professor

ISCTE – University Institute of Lisbon

Co-Supervisor: PhD Sérgio Miguel Carneiro Moro, Full Professor

ISCTE – University Institute of Lisbon

October 2023

Acknowledgements

This dissertation would not have been possible without all the interest, support, contributions and excellence of my supervisors Professor Catarina Marques and Professor Sérgio Moro, to whom I express great gratitude.

My gratitude also goes out to ISCTE-IUL and to all the teachers and professors that have contributed to my learning journey throughout life, from elementary school through university.

Also want to thank my parents for always encouraging me to pursue further education studies.

Finally, I want to thank everyone else who contributed in any way to the completion of this dissertation.

Abstract

Brand coolness is a multidimensional construct that encompasses different characteristics consumers perceive in a brand they think is cool. More research has been dedicated to the theme in recent years, been recognized as being of great importance for companies and marketeers alike. Brand love is another construct, which encompasses many kinds of positive emotions a customer can have regarding a brand.

Previous research has been mostly based on prompted, pre-formatted surveys. Concurrently, the new wealth of data available online has brought along new technology able to extract and analyze said data, its value being widely recognized for many marketing and managerial purposes.

This dissertation contributed to existing research by proposing a different method to gather and analyze consumer online feedback regarding a tech brand, using text mining and sentiment analysis techniques. More than 2000 consumer reviews were extracted, cleaned, processed and analyzed and a model was tested using linear regression models, for the relationship between brand coolness, brand love, loyalty and e-WOM (measured in its volume).

Results showed brand coolness is rather present in consumer online feedback regarding a tech brand, and that of all the brand coolness subdimensions, useful/reliable, usability and aesthetic were the most represented; extraordinary, energetic, and original/innovative were the most positively evaluated. In addition, results also showed there is in fact a causation effect of brand coolness on brand love and of those two on loyalty. The causal relationship between brand love and loyalty with e-WOM, measured in its volume, was not statistically significant.

Keywords: text mining; brand love; brand coolness; brand loyalty; sentiment analysis.

JEL Classification System: M31, M39

Resumo

Brand coolness é um construto multidimensional que abrange diferentes características que os consumidores podem percecionar numa marca que consideram *cool*. Vários estudos têm sido dedicados ao tema nos últimos anos, que tem sido reconhecido como de grande importância para empresas e profissionais de *marketing*. *Brand love* é outro constructo que abrange vários tipos de emoções positivas que um consumidor pode ter em relação a uma marca.

Estudos anteriores basearam-se principalmente em questionários pré-formatados. Ao mesmo tempo, a riqueza dos dados online trouxe novas tecnologias para extrair e analisar esses dados, sendo o seu valor reconhecido para muitos fins de *marketing* e de gestão.

Esta tese contribui para a literatura, propondo um método diferente para recolher e analisar o *feedback online* do consumidor em relação a uma marca de tecnologia, utilizando técnicas de mineração de texto e análise de sentimento. Mais de 2000 críticas de consumidores foram extraídas, limpas, processadas e analisadas e as hipoteses foram testadas usando modelos de regressão linear, para a relação entre *brand coolness, brand love, loyalty* e *e-WOM*.

Os resultados mostraram que o *brand coolness* está presente no *feedback online* do consumidor, e que, de todas as subdimensões do *brand coolness*, útil/confiável, usabilidade e estética foram as mais representadas; extraordinário, enérgico e original/inovador foram avaliadas mais positivamente. Além disso, os resultados mostraram que existe um efeito causal de *brand coolness* no *brand love* e de ambos em *loyalty*. A relação causal entre *brand love* e *loyalty* com *e-WOM*, medido em volume, não foi significativa estatisticamente.

Palavras-chave: mineração de texto; *brand love*; *brand coolness*; *brand loyalty*; análise de sentimentos.

Sistema de Classificação JEL: M31, M39

Table of Contents

List of Tables

List of Figures

List of Acronyms and Abbreviations

AI – Artificial Intelligence

BA - Brand Awareness

BL - Brand Love

CFI - Comparative Fit Index

CSV – Comma Separated Values

e-WOM - Electronic word-of-mouth

HTML – Hypertext Markup Language

ID - Identifier

IDE - Integrated Development Environment

IQR – Inter Quartile Range

SEM - Structural Equation Modelling

UGC - User-generated content

WTP – Willingness-to-pay

WOM – Word-of-mouth

TOL - Tolerance

VIF - Variance Inflation Factor

1. Introduction

Technology has become overarching in today's world, with tech companies reaching their highest market valuations in recent years (Pisani, 2023). At the same time, competition is fierce in this market, and it has become extremely important for companies to invest in marketing activities, namely in brand building campaigns (Kato, 2021; Agamudainambhi et al.,2021; Tiwari et al.,2021).

One of the biggest goals for marketing is to achieve consumer loyalty, in all its different forms (Oliver, 1999). Two concepts that have come into discussion during the previous years, due to their potential connections to marketing, branding and to consumer loyalty, are brand love and brand coolness. Having a better understanding of such concepts has become key for marketing management and for achieving healthier business outcomes (Carroll & Ahuvia, 2006; Jiménez-Barreto et al., 2022; Attiq et al.,2022). Both concepts have been potentiated by the ability that consumers nowadays have to engage directly with the brand and with other consumers online through electronic word-of-mouth, e-WOM (Yodpram & Intalar, 2020; Agamudainambhi et al.,2021). Social networking sites and online platforms facilitate these interactions and expand the spheres of influence of consumers (Rosenbaum & Massiah, 2007; De Valck et al., 2009).

Despite the fact there is already some research around brand coolness and brand love (Carroll & Ahuvia, 2006; Jiménez-Barreto et al., 2022; Attiq et al.,2022; Kato, 2021; among others), there is still a lack of research regarding their importance for a tech brand. From the research analyzed during the literature review, only entirely Tiwari et al. (2021) focused on tech brands, with most research focusing on services and consumer goods brands. Hence, this dissertation provided a necessary contribution to explore such constructs within the tech brand market.

In addition, most of that research uses prompted and pre-formatted data, gathered in the form of survey/questionnaire methods and focus groups (Jiménez-Barreto et al., 2022). While it is important to investigate using those types of data, there is a persistent gap in previous research regarding the use of unprompted, unstructured data, such as online reviews and social media comments, which could provide a more authentic and cost-effective data source (Lee, 2018; Dahiya et al., 2021). Online user-generated content (UGC) is created by individual consumers directly expressing their perceptions. Previous research has shown how relevant UGC is for consumers shopping online (Smith & Anderson, 2016). For such reasons, it is extremely important for brands to measure and to have a better understanding of UGC and e-WOM. This dissertation offers a method that is an improvement on that limitation that previous studies had and can offer a broader perspective in terms of understanding consumer perception based on UGC (Lawrence et al., 2013; Smith & Anderson, 2014; Dahiya et al., 2021).

Consumer engagement in the digital world regarding brands has reached all-time highs, with one of the most common forms of customer engagement being precisely e-WOM (Greve, 2014). This is both a consequence of the impact a brand has and a stimulant of further brand awareness, working as a marketing channel in itself (Agamudainambhi et al., 2022; Rosenbaum & Massiah, 2007). Moreover, content that consumers create online through e-WOM is an important factor in their decision-making process since consumers sometimes consider it to be more reliable than traditional marketing methods (Rosenbaum & Massiah, 2007). Effective management of virtual communities was proven to be able to expand markets, enhance visibility, and improve profitability for firms, according to Verhoef et al. (2009).

Online and digital activity contains a wealth of data that companies are learning to harvest and analyze (Lee, 2018). These vast amounts of data are not always easy to analyze, but the value that can be extracted from there has led to the appearance of several technologies that allow for big data analysis, data mining, text mining and sentiment analysis (Tan et al.,2023; Lee, 2018).

Hence, in this dissertation, I proposed to conduct research, first using data scrapping tools to gather user generated content published online in the form of Amazon reviews about a technology brand, Apple, and then to analyze it using text mining and sentiment analysis techniques. A conceptual model relating brand coolness, brand love, brand loyalty and e-WOM is proposed and tested. More specifically, this study aimed to understand:

- How are brand coolness and brand love related to brand loyalty?
- How are the subdimensions of brand coolness represented in the data?
- How are these results in the context of a tech brand specifically?
- Can all of the above research questions be measured directly from the content consumers create online, through text mining and sentiment analysis?
- How do these constructs influence e-WOM, measured as the volume/length of feedback the consumer provides on a review?

To achieve this, a literature review was first conducted to understand all existing literature regarding the topics mentioned above, which helped settle the hypotheses. Within the literature review phase, several relationships between constructs and metrics are hypothesized.

Afterwards, methodology was defined, and data was scrapped. That data was then treated and classified, identifying themes and keywords in the data and comparing those to dictionaries built around the constructs through text mining. Data was put through further text mining and sentiment analysis techniques in R software (The R Project for Statistical Computing, 2023) to compute sentiment scores and other metrics such as sentence and word count. Those metrics were then analyzed with different statistical analysis. Finally, results were compared to what was found in previous research and further considerations were made.

The results of this research contribute to marketing and branding, particularly in terms of tech brands. The novelty is in proposing a new process, using text mining and sentiment analysis techniques to extract value from the vast amounts of data available online, to measure consumer perception, complex brand perception constructs and their interactions with consumer loyalty. This will help companies understand their consumers' perceptions about their brand, leading to improved marketing strategies and better business performance.

2. Literature Review and Hypothesis Development

2.1. Brand Coolness

The concept of cool is challenging to define, with numerous synonyms and varied descriptions. It is a multidimensional construct which can encompass different meanings (Warren et al., 2019). While there is no consensus on one specific definition, it is important to establish a grounded understanding of brand coolness.

Warren et al. (2019) argued that coolness is subjective and socially constructed, attributed to cultural objects, and perceived as appropriately autonomous. Attiq et al. (2022) suggested that brands can leverage coolness perceptions to foster positive consumer experiences and cultivate customer loyalty. Brand coolness has also been seen as a socially constructed concept with shared meanings and behavioral standards by peers of individuals that share similar experiences (Runyan et al., 2013; Chen et al., 2021). However, these shared meanings may differ across generations and contexts, making it challenging to define coolness conceptually as shown by Warren and Campbell (2014); the latter research also found that what younger generations (gen X, millennials, and gen Z) consider cool, differs from their older counterparts (baby boomers) due to differences in autonomous thinking behavior. More specifically, boomers perceive cool brands to be less cool compared to other generations.

Consumers have also emphasized the communal-brand connection, described by Jiménez-Barreto et al. (2022) as the social relations and communal spaces linked to a brand that provide an environment where consumers can form their self-identity and be recognized by other consumers. The authors found that consumers felt that attending a music festival perceived as cool enhanced their own coolness. These findings highlighted the interrelated nature of brand coolness and the communal-brand connection in highly experiential services.

Coolness has been seen as a way for brands to differentiate themselves in a competitive marketplace, being associated with positive consumer perceptions and social image (Chen et al., 2021). Warren et al. (2019) highlighted the importance of identifying the characteristics that differentiate cool from uncool brands, especially as they transition from niche cool to mass cool. Warren and Campbell (2014) defended measuring brand coolness should involve direct assessments of consumers' perceptions. The authors considered that although there is no clear working definition of coolness, people can easily recognize a cool brand when they see it, and coolness perceptions are shared within social communities.

Warren et al. (2019) also defended the subjective and dynamic nature of brand coolness: initially niche, these brands are considered cool by a knowledgeable insider group but remain relatively unfamiliar to the broader population. Over time, some transition to mass cool status, gaining wider adoption and being perceived as more popular and iconic, albeit with potentially reduced autonomy. The same authors, on the fourth experiment of their research, acknowledged that actual perceptions of brand coolness are formed over multiple exposures to marketing and social signals over time. On the other hand, some factors do not seem to differentiate brands considered cool from those considered uncool, such as cultural knowledge, emotional concealment, friendliness, and competence. These traits, although desirable in people, did not emerge as significant factors for cool brands. Chen et al. (2021), in a study about hotel brands, also introduced autonomy, defined as the willingness to follow one's own path regardless of norms, as a driver of coolness perceptions. Brands that display autonomy stand out by offering products or services that deviate from market norms. This study mentions creative brands like Apple and boutique hotels that utilize distinctive designs and features to differentiate themselves from competitors. Attiq et al. (2022) further characterized brand coolness as evoking intense positive emotions and passion towards brands which can contribute to customer well-being and delight. Their findings suggested that brands can enhance customer satisfaction by cultivating coolness perceptions and fostering brand love and engagement, since these have a determining influence on satisfaction.

Warren et al. (2019) elaborated on the characteristics of each dimension of brand coolness: useful/extraordinary, described as being perceived as high quality, offering tangible benefits or previously unheard-of capabilities; aesthetically appealing, as being elegant and showing visually appealing designs; energetic, being associated with activities that evoke positive emotions and remarkable experiences; high-status, seen as having a high social status, exclusivity, and sophistication; original, as being considered creative and ahead of the curve; authentic, as remaining true to their roots and behaving consistently; rebellious, being associated with rule-breaking, controversy, and being revolutionary; subcultural, as being linked to specific subcultures that provide a sense of belonging to those subcultures; iconic, since cool brands hold strong and valued meanings to consumers, symbolizing memories, identity traits, and cultural values; and finally being popular, which means being recognized and widely admired, and appealing to a variety of people. These descriptions were fundamental when building the brand coolness construct dictionary. In a study on brand coolness within the field of technology, Tiwari et al. (2021) have argued that the increasing similarity among technology products makes it difficult for consumers to differentiate and choose one product over another with coolness emerging as a crucial factor in product evaluation and differentiation. They found that coolness is viewed as an abstract concept, characterized as dynamic, constantly changing over time. In the same study six dimensions of perceived coolness were found, which contained some similarities with Warren et al. (2019) and Attiq et al. (2022), namely desirability, since consumers purchase socially desirable products to shape their self-image and attain a desirable social identity; the innovativeness of technology, being the creative and unique aspects that differentiate tech products from competitors and fulfill consumers' need for assimilation or differentiation; attractiveness, the visual aesthetics and socially acceptable style; rebelliousness, the divergence from established norms and an appropriate form of non-conformity; and also added some new dimensions such as usability, referring to the functionality and practicality of products; and reliability, mentioning a product's ability to keep its promises and perform consistently, playing a crucial role in maintaining the perception of coolness and preventing reputation deterioration. This study was particularly interesting for the present research as the subject matter was the same, technology brands, and it highlighted the impact a strong measurement of brand coolness can have to brands. In line with other research reviewed (Warren et al., 2019; Chen et al., 2021) their study also acknowledged that the understanding of what consumers consider cool may vary across cultures and economies, leading to different dimensions of coolness.

Jiménez-Barreto et al. (2022) brought a different perspective to the study of brand coolness by looking specifically at this concept within the context of service brands. Results revealed that when it came to events such as music festivals, participants associated brand coolness with characteristics such as exclusivity, excitement, rebellion, originality, popularity, and aesthetics.

On the other hand, when it came to fast-food restaurants, Jiménez-Barreto et al. (2022) also found that brand coolness was associated with being extraordinary, exciting, aesthetically appealing, original, authentic, popular, subcultural and iconic. In this study, there were no concepts related to high-status and rebelliousness, unlike other research (Warren et al., 2019; Tiwari et al., 2021). Some participants in this study also expressed that fast-food restaurants were not cool experiences due to a lack of originality in their offerings which seemed to go against coolness in mass service brands (Jiménez-Barreto et al., 2022).

Regarding the concept of brand coolness and its relationships with other brand-related concepts, Attiq et al. (2022) found coolness had a positive effect on brand love and that brands perceived as cool contribute to consumers' positive emotions towards the brand. There was also a positive association between brand coolness and brand engagement, suggesting consumers are more willing to engage actively with a brand they perceive as being cool. In addition, Jiménez-Barreto et al. (2022) found that there was a relationship between brand coolness, communal-brand connection and brand loyalty, where the communal-brand connection mediates the relationship between service brand coolness and brand loyalty. Warren et al. (2019), on the other hand, quantify the relationship between brand coolness and various dependent variables; the results indicated that the effects of brand coolness on brand attitude, willingness to pay (WTP), and word-of-mouth (WOM) were partially or fully mediated by brand love, but brand coolness directly influenced these dependent variables as well. Brand coolness was also found to be related to self-brand connections and brand familiarity (Warren et al., 2019).

2.2. **Brand Love**

Brand love refers to the "passionate emotional attachment" that consumers have towards a brand (Robertson et al., 2022; p.652). The term was also used by Carroll and Ahuvia (2006) to describe the passionate, emotional, love-like attachment consumers have for a brand. It encompasses various dimensions of passion, attachment, positive evaluation, positive emotions, and declarations of love for the brand, thus distinguishing itself from satisfaction and simple brand affect. Brand love is, in addition, conceptually distinct from satisfaction because while satisfaction is typically a cognitive judgment, brand love has a stronger affective focus (Fournier & Mick, 1999). Moreover, according to Unal and Aydin (2013), brand love goes beyond satisfaction since it involves a deep connection and preference for a brand over the alternatives. Although brand love and satisfaction share similarities, brand love is developed over multiple interactions and signifies a stronger bond with the brand (Unal & Aydin, 2013).

The concept of brand love reflects a deep and enduring relationship that consumers develop with a brand, making it irreplaceable (Albert & Merunka, 2013). Consumers usually form a strong emotional bond with a brand when the brand image is positive and aligns with their social identity (Unal & Aydin, 2013). Moreover, brand love involves a willingness to declare love for the brand and integrates the brand into the consumer's identity (Carroll & Ahuvia, 2006). Some studies have also compared brand love to interpersonal love, emphasizing its long-term nature and the deep relationship between customers and brands (Fournier, 1998). This last author identifies brand love as a major relationship category that customers can develop with a brand, emphasizing its meaningful and long-lasting nature. These relationships are based on the fulfillment of consumers' needs and can lead to relationships where the brand becomes part of the individual's self-concept.

Brand love is particularly relevant for hedonic brands that offer symbolic benefit (Robertson et al., 2022). Findings from research by Anggraeni and Rachmanita (2015) indicated that there is a significant relationship between brand image, some elements of brand personality, and brand love. Additionally, brand love, brand personality and brand image have a significant positive relationship with WOM. Some authors have shown that brand love is associated with willingness to pay a premium, positive WOM, and brand loyalty (Carroll & Ahuvia, 2006; Albert & Merunka, 2013) and that it requires a holistic brand effort to evoke positive feelings that can be fostered through brand promotion, consumption experiences, and improvements in the consumer-brand relationship. Interestingly, brand love tends to exclude any negative feelings towards the brand and have its lower bound as the absence of emotional response instead (Oliver, 1999; Carroll & Ahuvia, 2006).

On the connection between brand love and loyalty, research has shown brand love plays a significant role in both behavioral and attitudinal loyalty (Robertson et al., 2022; Albert & Merunka, 2013; Carroll & Ahuvia, 2006). When consumers fall in love with a brand, their brand loyalty increases, leading to a stronger emotional connection and brand commitment (Albert & Merunka, 2013). Unique and memorable experiences also contribute to building brand loyalty and establishing a strong emotional bond with the brand (Robertson et al., 2022). Moreover, Robertson et al. (2022) found that brand love acted as an antecedent of brand loyalty and is considered a strong predictor of brand equity. The same study's results highlight the importance of considering both brand love and brand loyalty as separate dimensions in understanding the consumer-brand relationship for mass consumption prestige brands. This study challenges the idea that brand loyalty is the sole measure of success in consumer-brand relationships and argues that brand love may be even more important, particularly for brands that offer high hedonic, prestige, and symbolic value; furthermore, it shows that high brand love can exist even without brand loyalty. Findings from Anggraeni and Rachmanita (2015) also demonstrated that brand love can lead to positive WOM, corroborating the link between brand love and customer advocacy mentioned earlier.

The relationship between brand love and brand coolness seems apparent but some studies have shown that brand coolness is a distinct construct from brand love and other brand attitudes (Warren et al., 2019). Brand coolness is seen as a perceived attribute of a brand, while brand love is mostly considered a response to and a consequence of brand coolness. Attiq et al. (2022) found there was a mediating role of brand love and brand engagement in the relationships between brand coolness, customer delight, and customer psychological well-being.

Overall, the findings emphasize the importance of cultivating brand love among consumers to enhance brand equity and improve general brand perceptions. Taking into account the literature review about brand coolness and brand love, and the relationship between those two constructs, we hypothesize that:

- H1: Brand coolness positively influences Brand love (+)

2.3. Brand Loyalty

Oliver (1997) is one of the first authors to talk about consumer loyalty. The author describes loyalty as "a deeply held commitment to rebuy or patronize a preferred product/service consistently in the future, thereby causing repetitive same-brand or same brand-set purchasing, despite situational influences and marketing efforts having the potential to cause switching behavior" (p. 392). Oliver's (1997) framework proposes a sequence of loyalty phases that consumers go through: the cognitive loyalty phase, where consumers perceive one brand as preferable based on attribute information; the affective loyalty phase, when consumers develop a liking or positive attitude towards the brand through satisfying usage occasions; the final phase is action loyalty, where intentions are transformed into readiness to act and overcome obstacles to repurchasing. Taking this definition into account it is fair to say that even customers who have not purchased or used a product can display loyalty to a brand, specifically in the cognitive phase.

The relationship between loyalty and satisfaction has been defined as complex and asymmetrical (Oliver, 1999), since while loyal consumers are typically satisfied, satisfaction does not always lead to loyalty. The author's analysis revealed that satisfaction is a necessary step in the formation of loyalty but becomes less significant as loyalty develops through other mechanisms (Oliver, 1999). These mechanisms include personal determinism (described as "fortitude") and social bonding at both the institutional and personal levels. When these factors are considered, ultimate loyalty emerges as a combination of perceived product superiority, personal fortitude, social bonding, and their synergistic effects.

Kato (2021) found that the most influential factors on loyalty were the product characteristics, followed by the staff that represents the brand, and the brand image. Another study by Kiss et al. (2022), about chocolate brands, found that brand loyalty is influenced by previous experiences and by brand satisfaction, and plays a crucial role in consumers' preferences. If consumers have a positive experience with a particular brand, they are more likely to exhibit brand loyalty and continue repurchasing it, which confirmed similar conclusions from previous work by Oliver (1999), Johnson et al. (2006) and Kato (2021), among others. The conceptual model developed by Johnson et al. (2006) also suggests that loyalty intentions are influenced by perceived value, brand equity, and affective commitment. Johnson et al. (2006) also found that evaluations of perceived value, brand equity, affective commitment, and loyalty intentions are usually updated versions of evaluations the customer has previously made.

Oliver (1999) makes a distinction between simple repeat purchases and those driven by high loyalty. Factors such as low prices or proximity of stores can lead to spurious loyalty, where consumers easily switch brands. To capture loyalty accurately, both behavioral and psychological indexes should be considered (Aaker & Joachimsthaler, 2000). In a study by Johnson et al. (2006), which examines the evolution of customer loyalty intentions throughout the introduction and growth phases of a product's life cycle, the authors found that loyalty intentions are initially influenced by perceived value. Brand loyalty has also been characterized as consumers' intention to repurchase and their commitment to choosing a specific brand over others (Agamudainambhi et al., 2022) in a study which proposes direct relationships between brand loyalty and brand image with consumers' willingness to pay a premium price. More research by Yodpram and Intalar (2020) demonstrated the positive influence of brand loyalty on willingness to pay a premium price. Kato (2021), in a research paper about the factors of brand loyalty on the automotive sector, highlights the importance of customer loyalty in driving repeat purchases, price increases, and positive WOM. Many companies introduce loyalty programs, but those traditional approaches, which include discounts and customer points systems, were seen as ineffective. Still regarding the concept of brand loyalty, Kato (2021) noted that it is considered crucial to a brand's equity, with repeat purchases serving as behavioral indexes.

Customer engagement was also proven to have an impact on the brand image-brand loyalty relationship in a study about online customer engagement through social networking sites (Greve, 2014). Brand equity, which refers to the unique impact of brand knowledge on consumer responses to marketing efforts (Keller, 1993), seems to also be connected to loyalty and has been suggested as going beyond the value proposition and capturing the additional effect brands have on loyalty intentions, acting as a mediator between quality, satisfaction, and loyalty intentions to varying degrees (Keller, 1993; Johnson et al., 2006).

Considering the literature review about brand coolness, brand love and the antecedents of loyalty, we hypothesize that:

- H2: Brand coolness positively influences Loyalty (+)

- H3: Brand love positively influences Loyalty (+)

2.4. Word-of-mouth (WOM) / Electronic word-of-mouth (e-WOM)

Word-of-mouth (WOM) involves the informal sharing of opinions, information and comments about products, brands, and services among consumers with each other (Hawkins et al., 2004). It has also been defined as a non-commercial means of communication driven by satisfied customers who aim to endorse compatible brands (Arnett et al., 2003), although this definition encompasses only the positive information being shared by customers. WOM plays a significant role in influencing consumer decision-making and often holds more sway than traditional advertising as found by Nguyen and Romaniuk (2014). WOM is considered a part of customer engagement as engaged customers perform specific actions such as providing ideas, collaborating, purchasing, recommending, and providing feedback (Greve, 2014).

Initially, WOM was interpersonal communication influencing purchase behavior. With technological advancements, electronic word-of-mouth (e-WOM) emerged as unpaid online communication visible to a wider audience (Williams, 2019). e-WOM, is an extension of WOM into text-based communication, and encompasses various forms of computer-mediated communication, such as blogs, emails, and bulletin board systems (Hu & Ha, 2015). Hu and Ha (2015) further categorized e-WOM into four classes based on functions and communication forums: specialized e-WOM (on comparison-shopping or rating websites); affiliated e-WOM (affiliated with retail websites); social e-WOM (information related to brands/products exchanged on social networking sites) and miscellaneous e-WOM (brand/product information exchanged on other online social media platforms).

e-WOM can take diverse formats, including reviews, recommendations, social media posts, and blogs, and operates through various modes and timings, having different motivations, from self-interest to altruism, characterized by being unpaid and happening online (Williams et al., 2019). Rheingold (1993) first defined virtual communities as “social aggregations that form through online discussions, creating personal relationships in cyberspace” (p.5). These communities have evolved from early electronic boards to blogs and online commerce platforms, comment sections and social media, where members engage in textual conversations to find solutions, exchange best practices, and build expertise while forming meaningful social relationships (Mathwick et al., 2008). Within virtual communities, social bonding plays a crucial role in fostering ongoing participation, creating a home-like atmosphere for its members, and encouraging them to share feedback about brands, companies and products through electronic word-of-mouth (e-WOM) (Yodpram & Intalar, 2020; Chou & Sawang, 2015).

Other authors have recognized the importance of understanding consumer activity and interactions in virtual communities (Colliander & Wien, 2013; Verhoef et al., 2009). Interactions within virtual communities often generate influential information about firms, products, services, and consumption experiences, which can then influence consumers (Colliander & Wien, 2013). Kudeshia et al. (2017), in a study with 325 participants answering a questionnaire about their activity on Facebook regarding a smartphone brand, also found that social e-WOM had a significantly positive impact on brand attitude and a mediating effect on purchase intention.

Online customer engagement involves a cognitive and affective commitment to actively interact with the brand through digital platforms (Greve, 2014). User-generated content, a term encompassing media created by the public rather than professional copywriters, provides significant influence in consumers' decision-making processes (Greve, 2014; Lee, 2018). Social media platforms have become a popular outlet for individuals to express their self-identity through brand related user-generated content (UGC) which marketers have recognized as important for the credibility and authenticity of brands (Bernritter et al., 2017). Brand related UGC is perceived to be even more believable, authentic, and more engaging both emotionally and intellectually, compared to traditional advertising (Lawrence et al., 2013). Consumers tend to place higher trust in the opinions of their peers compared to those published by service providers (Lawrence et al., 2013). Positive online reviews function as cost-effective advertising for brands whereas negative comments in the digital sphere can have the opposite effect, detrimental to a company, brand image and reputation (Greve, 2014, Chen et al., 2022). The impact of e-WOM varies based on product types and user characteristics. Hu and Ha (2015) investigated the use of electronic word-of-mouth (e-WOM) and traditional word-of-mouth (WOM) among consumers for four product types. The findings indicated that consumers were more likely to seek external sources when considering purchasing books, movies, music, games, and electronic products compared to health, beauty, clothing, shoes, and jewelry products. In the case of electronic products, specialized e-WOM and expert opinions, either online or face-to-face, were the preferred sources for relevant information.

Roy et al. (2020), in their study about hotel brands, discusses the impact of e-WOM stimuli on perceived service quality, highlighting various characteristics such as valence, volume, and image-based reviews. e-WOM valence is the qualitative nature (positive, negative, or mixed) of e-WOM content which significantly affects various online behavioral outcomes, including credibility, product judgment, purchase intention, online sales, and brand attitude. e-WOM volume is related to the quantity, length or count of sentences and words within the e-WOM user generated content (UGC). In the context of hotel booking, customers often check both valence and volume to assess the credibility of reviews. High e-WOM volume and positive valence together have a higher impact on consumer perception of product/service quality. Visual information in image-based reviews is essential for customer consideration. Both e-WOM valence and volume were found be important when it comes to enhancing perceived service quality, referring to how a customer experiences a service in comparison to their expectations, and to reducing the customer gap, when the outcome quality matches customer perception and expectation (Roy et al., 2020). A similar conclusion was reached by Kudeshia et al. (2017) that showed that the "valence" or evaluative direction (positive, neutral, or negative) of reviews influences consumers' purchasing decisions. Positive reviews on e-WOM platforms lead to higher recommendations to friends compared to products with negative reviews (Kudeshia et al., 2017). Roy et al. (2020) also found that there is a connection between customer loyalty and recommendation intention: satisfied customers are more willing to recommend to others. The author suggests that in the online era, customers express preferences and loyalty by recommending online. Roy et al. (2020) further recommended managers to monitor both e-WOM valence and volume, considering the significant influence on new customers' perceptions.

According to research by De Valck et al. (2009), consumers have been shown to engage in WOM to assist others in their decision-making process and to alleviate their own uncertainties, sharing both positive and negative experiences. Another theory frequently approached regarding WOM and e-WOM research is the resource exchange theory which suggests that virtual communities allow users to engage in exchange activities, where they seek others' opinions and provide support by sharing resources and information (Rosenbaum & Massiah, 2007). As mentioned above, social networking sites and online platforms have further facilitated these interactions and expanded the spheres of influence of consumers (Rosenbaum & Massiah, 2007; De Valck et al., 2009).

The valence of e-WOM is part of the sentiment scores of reviews which is used to measure constructs within this dissertation. On the other hand, e-WOM volume was also shown as potentially having a positive impact for brands in the research mentioned above (Roy et al.,2020; Kudeshia et al.,2017), which makes it important to understand its causes.

For those reasons e-WOM is also measured in its volume, as the length of the review in number of words and sentences the consumer wrote regarding the brand and their experience with its products and services.

Following the literature review on WOM and e-WOM, we hypothesize that:

- H4a: Brand love positively influences e-WOM (in volume of words) (+)
- H4b: Brand love positively influences e-WOM (in volume of sentences) (+)
- H5a: Loyalty positively influences e-WOM (volume of words) (+)
- H5b: Loyalty positively influences e-WOM (volume of sentences) (+)

3. Methodological Approach

Figure 3.1 presents the proposed conceptual model and hypotheses regarding the relationships between the constructs and e-WOM metrics.

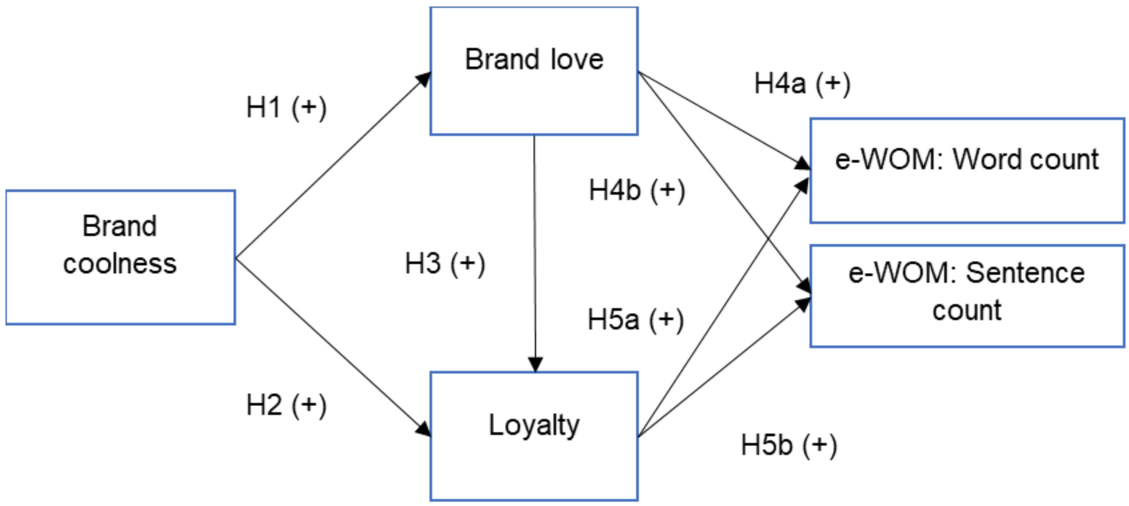

Figure 3.1: Model structure for the relationships between the selected constructs and metrics

3.1. Methodology

The literature review allowed to identify interesting avenues for research due to the increasing importance of such constructs as brand coolness, brand love and the more explored, brand loyalty together with e-WOM (Oliver, 1999; Johnson et al., 2006; Agamudainambhi et al., 2022; Greve, 2014 among others). In addition, the literature review showed much research is still dependent on surveys and is not extracting value from the feedback users create online, which can be a more direct and cost-effective method (Dahiya et al., 2021) of gathering data from consumers while also reflecting their perceptions of the brand in a more direct way (Lee, 2018). Following the literature review, the methodology was designed. In addition to the constructs' measurement through sentiment analysis and word/sentence counts, a model was created to measure the relationships between the different constructs and metrics. The main goal of the analysis is to have a data-driven approach based on text mining, sentiment analysis and UGC.

e-WOM volume was measured as the length of the review in number of sentences and number of words. As mentioned in the literature review this decision was made since the sentiment score measured for the constructs was a part of the overall e-WOM valence/sentiment already. There is a nested and dependent relationship between both variables since one integrates the other. Moreover, including the e-WOM volume metrics and measuring its relationship with the other constructs is novel to research among all the literature analyzed and can provide value since volume was found to influence consumer perception when receiving e-WOM from previous brand customers (Roy et al.,2020; Kudeshia et al.,2017). Hence, sentence and word counts were computed to conduct the e-WOM measurement as the length of the review and see if this is, in any way, influenced by the constructs under analysis. Carroll and Ahuvia (2006) found that WOM was impacted by brand love, however, the authors used a cross survey questionnaire asking about WOM intentions; the volume of the e-WOM communication itself is more a reflection of those intentions.

3.2. Subject for the study

Apple Inc. is a tech giant of great importance in today's world. Its influence resonates across global markets, earning it a place among the most talked-about companies today (Statista, 2022; Beattie, 2021). Apple's significant impact on the consumer electronics industry ranges from the groundbreaking release of the iPhone in 2007, which revolutionized the smartphone market, to the consistent development of products such as the MacBook, iPad, Apple Watch, Apple TV, also extending to services such as Apple Music, iTunes, App Store and Apple TV+. Apple's reputation for innovation, design, and user-friendly technology has made it a household name (Beattie, 2021). According to Statista (2022), Apple held a 23.3% share of the global smartphone market in the fourth quarter of 2021, reinforcing its dominance in the tech industry. Forbes ranks Apple as one of the world's most valuable brands, highlighting its financial prowess and cultural significance (Kelly, 2023). Its prominence on the global stage as a tech innovator and its cultural impact through products like the iPhone and Mac are of great significance.

Having a long and strong impact on the world of technology created one of the biggest digital footprints and larger customer bases, which in turn allows for a more successful data collection process. All the reasons mentioned here made it the brand subject of choice for this dissertation.

3.3. Text mining and sentiment analysis

Text mining, as defined by Hearst (1999), involves leveraging large online text data to uncover insights and trends about a topic of research. It encompasses the process of finding patterns or knowledge from unstructured text documents. This process provides a structured framework to extract value from vast amounts of text allowing the identification of patterns and connections between resources, leading to insight discovery across diverse fields. Sentiment analysis represents a pivotal facet within the realm of natural language processing, directed at the classification of text into three fundamental emotional categories: positivity, negativity, and neutrality (Tan et al., 2023).

The surge in online platforms providing individuals with an open channel for articulating their viewpoints and thoughts, has underscored the growing significance of deciphering the emotional undercurrents that are reflected within these expressions (Greve,2014; Williams et al.,2019; Roy et al.,2020). The different sentiment analysis approaches can be categorized as machine learning, deep learning, or ensemble learning (Tan et al., 2023). These are becoming indispensable for organizations seeking to gather valuable insights and to make data-driven decisions based on large pools of textual data.

3.4. Data extraction and data cleaning process

To conduct the research, several potential sources of text data were investigated, among them: Twitter, Facebook, Instagram, Amazon and CNET. However, following the recent limitations that some platforms put up on data scrapping technologies (Instagram, 2023; Twitter, 2023) and considering the greater amount and diversity of data available, Amazon was chosen as the source to extract data from.

A total of 2319 reviews were extracted regarding five different Apple product lines (iPad, Apple Watch, Apple TV, Airpods/ Apple headphones and Macbook) from the US/Global Amazon website (amazon.com). The iPhone reviews were left out following a preliminary analysis which revealed iPhones sold through Amazon were bound to a mobile network carrier, which influenced feedback and reduced the relevance of data regarding the customer sentiment on Apple itself. All product pages containing at least 100 ratings were chosen to allow a bigger pool to extract data. After that, all ratings with written reviews were selected, sorted by the option “top comments”, which reflects comments that resonated more with the users (Amazon, 2023) and the first 100 were selected when available. Only “verified purchases” reviews were selected to provide the most reliable data possible, since those are from verified customers of the brand (Amazon, 2023). In addition, reviews there were not in the English language were translated using Google Translate. The process of data extraction and cleaning is shown in figure 3.2.

For the data extraction process, Instant Data Scrapper from Web Robots was used. Instant Data Scraper is a no-code, automated data extraction tool for any website that uses AI to analyze the HTML structure of pages, scrape specific sections of the page and save it to an excel or CSV file. It streamlines the process of gathering information from web pages by automating the extraction of text, images, links, and other content elements and is useful for various purposes, such as data collection for research, market analysis, lead generation, or any scenario where unstructured data from websites is required (Web Robots, 2023).

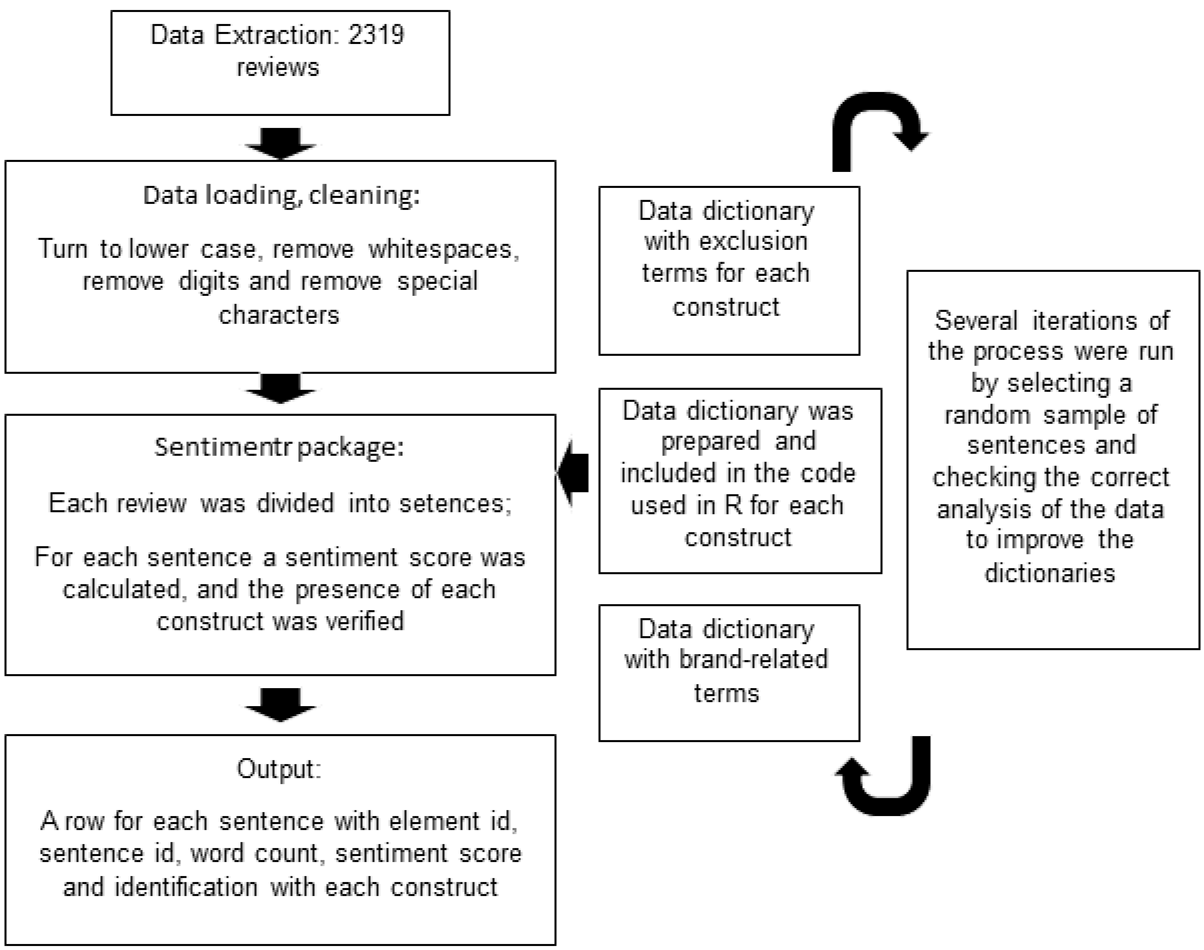

Figure 3.2: Data extraction, cleaning and analysis process

Following extraction, data was loaded onto R software (The R Project for Statistical Computing, 2023). R is an integrated development environment (IDE) for the R programming language, providing a wide range of features for data analysis and statistical modeling. It offers tools for coding, visualization, and data manipulation, making it a go-to choice for many data analysts, scientists and statisticians.

“Sentimentr” is an R package that specializes in sentiment analysis, allowing to assess the emotional tone of text data (Rinker, 2021). It considers various aspects of sentiment, including valence, arousal, and dominance. The package is a helpful tool for researchers to gain deeper insights into public opinion, customer feedback, or social media conversations (Rinker, 2021). It offers several functions that are tailored for conducting sentiment analysis, but within this study, two key functions, namely "sentiment" and "get_sentences" were employed to gauge text sentiment on a sentence-by-sentence basis, together with other R functions such as “strsplit”, “str_detect” “tolower” and “gsub” for the data cleaning process outlined in figure 3.2. The “sentimentr” function involves the examination of sentences within each post, seeking the presence of specific keywords associated with various sentiment categories. Upon discovering a match, a sentiment score is calculated, assigning negative values (down to –1) for negative sentiments, zero for neutrality, and positive values for positive polarities (up until +1). “Sentimentr” also has the ability to handle valence shifters effectively. Valence shifters are words or phrases that can change the polarity or sentiment of a sentence. For example, the sentence "I like it, but it's expensive" contains a valence shifter ("but") that changes the overall sentiment (Rinker, 2021).

After uploading the data into R, it was turned into lower case and cleaned from any whitespaces, digits and special characters. Furthermore, each review was divided into separate sentences, with each being checked against the constructs’ dictionaries prepared during the literature review phase. A word count was computed and the sentence id count was used as sentence count.

3.5. Data Dictionary Development

To form the construct dictionaries the following authors’ contributions were considered, as taken from the literature review phase. All dictionaries were subject to changes following several iterations of data analysis process to better adapt them to the collected data and ensure their efficacy in capturing the constructs correctly.

3.5.1. Brand coolness construct dictionary

Coolness is seen as a way for brands to differentiate themselves (Chen et al., 2021). Attiq et al. (2022) characterizes brand coolness by factors such as reliability, usability, and prestige, which can contribute to customer well-being and delight. Attiq et al. (2022) also mentions that coolness perception is associated with features like reliability, dependability, usability, and uniqueness, which evoke intense positive emotions and passion towards brands. Chen et al. (2021) introduces autonomy, defined as the willingness to follow one's own path regardless of norms, as a driver of coolness perceptions. Characteristics associated with coolness also include subcultural, attractive, original, unique, extraordinary, aesthetically appealing, energetic, high status, rebellious, authentic, iconic, and popular (Warren et al., 2019); Jiménez-Barreto et al. (2022) revealed that participants associated brand coolness with characteristics such as exclusivity, excitement, rebellion, originality, popularity, and aesthetics and found that brand coolness was also associated with being extraordinary, exciting, authentic, popular, subcultural, and iconic. Within the field of technology, Tiwari et al. (2021) found that brand coolness is viewed as an abstract concept, characterized as dynamic, constantly changing over time; Tiwari et al. (2021) also studied the following subdimensions of perceived coolness in tech: desirability, innovativeness of technology, attractiveness, visual aesthetics and socially acceptable style, rebelliousness, usability (referring to the functionality and practicality of products) and reliability.

This information was then organized and adapted to prepare the following dictionary for brand coolness seen in table 3.1.

Table 3.1: Keywords identified in the Brand coolness literature per subdimension

Construct	Subdimension	Keywords
Brand coolness	Useful	"useful", "helps people", "helpful", "valuable", "versatile", "functional", "reliab", "quality", "high-quality", "reliable", "working", "work", "convenient", "effective", "handy", "productive", "efficient", "as promised", "sturdy", "performance", "responsive"
	Usability	"usability", "practical", "simple to use", "simple configuration", "simple command", "simplified", "easy to use", "easy to operate", "easy to program", "easy to use", "user-friendly", "user friendly", "intuitive", "user-centric", "user centric", "easy-to-use", "accessibility", "effortless", "seamless", "seamless to use", "smooth", "user experience", "user interface", "navigable" , "interactive", "learnability", "human-centric design", "ux"
	Extraordinary	"exceptional", "superb", "valuable", "fantastic", "extraordinary", "astonishing", "set the standard", "outstanding"

	Energetic	"energetic", "outgoing", "lively", "vigorous", "excit","dynamic", "active", "cheerful", "joy", "vibrant", "thrilling", "enthus"
	Aesthetic	"looks good", "good looking", "nice look", "aesthetically appealing", "aesthetic", "attractive", "appearance", "nice appearance", "attractive design", "nice design", "cool design", "visual aesthetics", "style", "elegant","sophisticat", "design", "beautiful", "look cool", "looks cool", "gorgeous", "stunning", "captivating"
	Original	"original product", "original look", "originality", "original design", "original brand", "does its own thing", "innovative", "innovation", "unique", "a different product", "innovati", "innovative tech", "creativ", "it's ahead", "ahead of the curve", "distinctive", "singular", "unparalleled", "uncommon", "matchless", "with no match", "one-of-a-kind", "distinct brand", "distinct product", "distinct tv", "distinct watch", "distinct laptop", "one of a kind", "the only product", "pioneering", "trailblazing", "game-changer", "game changer", "new gen", "new generation", "rare", "cutting-edge", "avant-garde", "avant garde", "visionary"
	Authentic	"authentic", "true to its roots", "true product", "true brand", "artificial", "doesn't try to be something it's not", "real product", "real brand", "sincere product", "sincere brand", "bonafide product", "bonafide brand", "honest product", "honest brand"
	Rebellious	"rebellious", "rebel", "defiant", "not afraid to break rules", "break rules", "broke rules", "nonconformist", "independent", "rule-breaking", "against the rules", "controversy", "revolutionary", "revolution", "outside the norm", "outside the ordinary", "different from the norm", "stands apart", "against the norm"

	High-status	"high status", "high-status", "high-end", "high-end market","chic","glamorous","glam","sophisticated","sophistication","ritzy","prestige", "exclusivity", "exclusiv", "luxury", "investment"
	Popular	"liked by", "liked by most people", "popular", "widely liked", "it is cool", "it's cool", "cool product", "was a success", "trendy", "in vogue"
	Subcultural	"people who use it different from other people", "stand apart from others", "stand apart", "stand apart from the crowd", "unique people", "subcultural", "subculture"
	Iconic	"cultural symbol", "culture", "cultural", "icon", "iconic product", "iconic design", "iconic brand", "legendary", "symbolic", "recognizable", "famous", "distinctive", "trademark"

3.5.2. Brand love construct dictionary

Carroll and Ahuvia (2006) described it as the passionate emotional, “love-like” attachment consumers have towards a brand. It encompasses various dimensions of passion, attachment, positive evaluation, positive emotions, and declarations of love for the brand, distinguishing it from satisfaction and simple brand affect. According to Robertson et al. (2022), the conceptualization of brand love includes dimensions such as positive brand emotions, brand evaluations, brand passion, and declarations of love towards the brand. The brand love scale from Carroll and Ahuvia (2006) contained the following eight sentences (p. 84-85):

“This is a wonderful brand. / This brand makes me feel good. / This brand is totally awesome (...) / This brand makes me very happy. / I love this brand! (…) / This brand is a pure delight. / I am passionate about this brand. / I’m very attached to this brand.”

The brand love scale taken from Unal and Aydin (2013, p.82) has seven sentences related to brand love (in that case a sport shoe brand):

“This (…) is a perfect brand. / This brand makes me feel good. / This brand is completely a wonderful brand. / This (...) brand makes me happy. / I like this (...) brand. / I am passionately attached to this [product]. / I am like a whole with this (…) brand.”

Combining the characteristics of the two scales, table 3.2 presents the keywords for the Brand love dictionary.

Table 3.2: Keywords identified in the Brand love literature

Construct	Keywords
Brand love	"positive emotions", "emotions", "passion", "love", "adore", "declaration of love", "incredible", "wonderful", "awesome", "very good", "great", "very happy", "delight", "pure delight", "passion", "very attached", "adore", "cherish", "admire", "idolize", "treasure", "appreciate", "worship", "revel in", "romance", "romanticize", "esteem", "enamor", "dote on", "revere", "devote", "crave", "yearn for", "covet", "perfect", "amazing", "wonderful", "passion", "attached", "passionately attached", "favorite", "favorite brand"

3.5.3. Loyalty construct dictionary

Loyalty emerges as a combination of perceived product superiority, personal fortitude, social bonding, and their synergistic effects (Oliver, 1999). For consumers to become and remain loyal, they must hold the belief that the products offered by a particular company or brand continue to be the best choice available (Oliver, 1999).

Johnson et al. (2006) have measured the construct of loyalty including the recommendation behavior within it; their loyalty construct scale included the following four sentences (p. 127) about a cell phone product:

“Next time I will definitely buy this [product] (or its successor) again. / If I lose my [product], I will definitely buy it again. / (…) I recommend my [product] to other people. / I talk to other people about my [product].”

On the other hand, Kiss et al. (2022) brand loyalty scale (p. 7) is composed by the six following sentences:

“I prefer famous and reputable brands to less known brands. / Brands are very important to me. / My favorite brand never causes disappointment. / I regard my favorite brand as my friend. (...) / I will continue to buy my favorite brand in the future as well. (...) / I am happy to recommend my favorite brand to others."

Table 3.3 presents the keywords created for Brand loyalty:

Table 3.3 Keywords identified in the Brand loyalty literature

Construct	Keywords
Loyalty	"loyalty to", "loyal to", "continue to buy", "continuity", "continue buying", "repurchase", "reorder", "favorite", "buy again", "will buy again", "always buy", "share this", "part of my choice" ,"part of my identity", "part of my life", "part of my routine", "recommend", "recommend it", "share it with others", "my preference", "choice of brand", "choice of company", "the brand i choose", "the product i choose", "my first choice", "always choose","i'm a loyal fan of","i'm a loyal customer", "stick with", "a customer for years", "a customer since", "proud supporter", "apple supporter", "i'm committed to", "i've tried others but i always come back", "i've tried others, but I always come back","i always come back", "i've tried others", "i tried others", "relationship with brand", "relationship with company", "fidelity to", "devotion to", "would buy again", "would buy this again", "would buy it again", "apple fan", "fan of apple", "convert to"

For each construct, some synonyms were also compiled for the keywords extracted from the literature review phase, using the Cambridge Thesaurus, Merriam-Webster and Dictionary.com online dictionaries. Several iterations of this process were conducted, mainly to improve the data dictionary quality, and to fine tune the data cleaning process. Throughout all those iterations, the constructs' dictionaries were continuously modified, expanded or reduced, taking into account the quality of the data categorization and the amount of constructs the model was able to capture. These modifications were always done considering the previous research studies analyzed in the literature review and mentioned above.

Amazon reviews are unstructured and free form. The reviews extracted for this dissertation contained between 1 and 169 sentences, and between 2 and 3184 words. Titles (which differ from the review body) were appended to all reviews as their first sentence. In addition to the dictionary that helps identify the constructs, an exclusion dictionary for each construct was also prepared (as seen in Appendix A, table A.1). The reason for this was that during the different iterations it was found some sentences where the user mentioned a competitor brand would be wrongly attributed to the brand under analysis in this study. For instance, the sentence “My Sony headphones have much better sound quality” would be positively classified by the sentiment analysis algorithm and would be flagged as containing the coolness construct (useful subdimension), but the sentence would be about another brand, in this case Sony and not Apple. Multiple expressions were added to the dictionaries considering the different constructs and the results of the different data analysis iterations conducted.

In addition, a brand-related dictionary was also created (seen in Appendix A, table A.2) and applied to each sentence to guarantee that the user mentioned the brand (or any brand related terms) in the sentence being analyzed. Only in case this was positive would that sentence have a construct attributed to it.

3.6. Data analysis process

The final output contained the review id (element id on the output), sentence id, sentiment score and the confirmation if the sentence belonged to any construct. In addition, it also contained a word and sentence count (computed from the sentence id) for each review which will be used to measure e-WOM volume (Greve, 2014). To measure the brand coolness construct, all coolness subdimensions (twelve, adapted from Warren et al., 2019 and Tiwari et al., 2021) were measured separately and from there an arithmetic average was calculated as the sentiment attributed to brand coolness overall. This was the variable used in the regression models to test the impact of brand coolness in brand love and loyalty.

Some reviews lacked certain constructs (especially for some of brand coolness’ subdimensions). This created a lot of missing values. A first test was done to model fit, replacing missing values as zero, using the Lavaan package in R for Structural Equation Modelling (SEM), and having brand coolness as a latent variable predicted by all twelve subdimensions. This model proved not having a good fit with a test statistic of 410.443 with 64 degrees of freedom and a p-value (chi-square) of 0.000 which indicated that there was a significant difference between the observed covariance matrix and the expected matrix under the fitted SEM. In structural equation modeling, the null hypothesis is that the model fits the data well, and the alternative hypothesis is that the model does not fit the data well (Boslaugh & McNutt, 2008).

To test the hypotheses, linear regression models were used instead to estimate the coefficient of each relationship. The first analysis regarding the descriptive statistics was done using Microsoft Excel (Microsoft, 2023). This analysis helped to summarize data and to further characterize the sample. Mean, median, standard deviation, skewness and kurtosis were calculated for each construct. The Pearson correlation coefficient was also calculated to examine associations between constructs, revealing which variables and subdimensions were linked to each other. The correlation coefficient quantifies the inclination of two variables to change concurrently, either increasing or decreasing together and ranging from -1 to 1, with 0 indicating no correlation, 1 denoting complete positive correlation, and -1 representing a negative correlation (Stewart, 2023). The linear regression analyses were conducted using IBM's SPSS Version 29.0.1. IBM SPSS is a software program used for statistical analysis and data management. It offers a wide range of tools for researchers and analysts to perform data analysis and generate insights (IBM, 2023). As previously mentioned, to analyze e-WOM, the metrics of sentence count and word count were also used.

4. Results

4.1. Sample data

The main sample of 2319 reviews was extracted from product pages in Amazon. The data was then processed using R to remove extra whitespaces, numbers and special characters that did not represent any interest for the analysis. Reviews were divided into sentences using the “get_sentences” function as mentioned before. This resulted in a total of 15555 sentences. The dictionaries were then run for each of those sentences looking for constructs. Whenever the same construct was present multiple times in a review, an arithmetic average was created for that construct in that review. Out of the total 2319 reviews, the constructs under analysis were found in 1731 of those, which formed the final sample data used in this dissertation. As mentioned in the previous section regarding model fit, reviews where no construct was found were removed from the analysis.

4.2. Descriptive statistics

The average review score was 4.7 stars out of 5 on the Amazon website itself. In order to protect consumer privacy, demographic data about the user who posted the review is not made available by Amazon (Amazon, 2023), but all reviews were filtered as coming from the United States, with 98% being in English and the remainder in Spanish (translated to English before analysis as mentioned in the Methodology section). In terms of percentual distribution of the main constructs found in the data, as seen from figure 4.1 below, the most represented was brand love (39.6%) followed by brand coolness with 28% and loyalty with around 7%. It is noteworthy that some sentences/reviews could contain more than one construct. For instance, the sentence “(…) love the mac book, it helps with day-to-day tasks and with better organization for work and home” contains both brand love and brand coolness constructs.

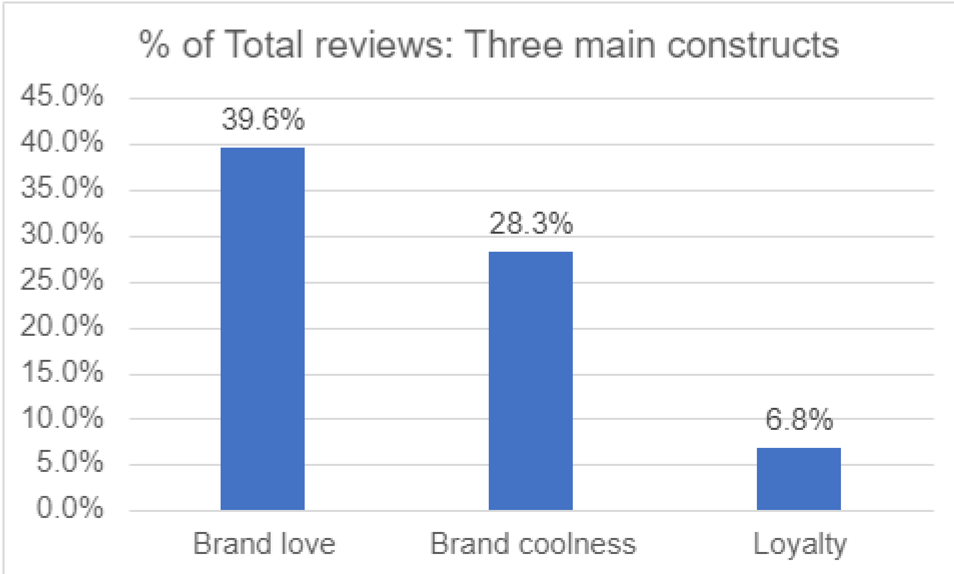

Figure 4.1: Construct representation in the data

In terms of sub dimensional representation within the reviews that contain the brand coolness construct (figure 4.2), there is a clear impact of the brand coolness/ useful subdimension in 74.7% of those reviews, followed by brand coolness/ usability with 22.6%, brand coolness/ aesthetic with 18.8%, brand coolness/ extraordinary with 10.2% and brand coolness/ energetic with 7.6%. All the remaining seven subdimensions of coolness selected for the present research ranged between 4.6% and 0.2%. Particularly, the rebellious, authentic, popular, subcultural and iconic subdimensions had very low representation.

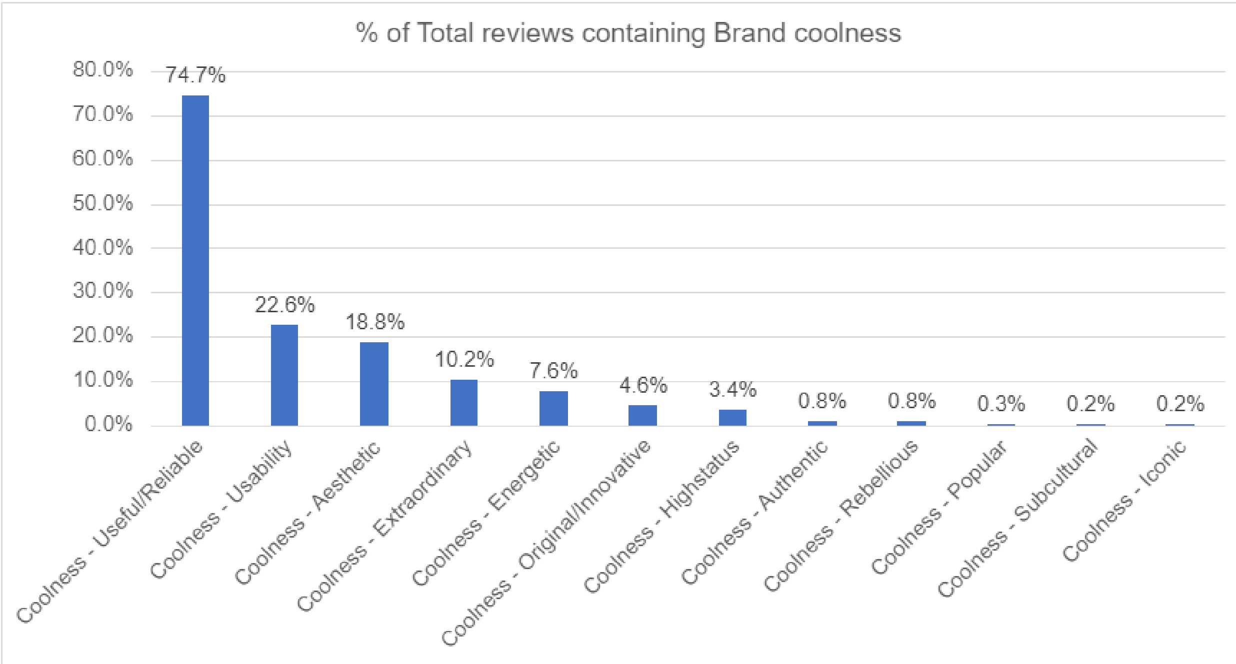

Figure 4.2: Brand coolness subdimensions representation in the data

In general terms, the most predominant construct within the reviews Apple customers left on Amazon, was brand love followed by brand coolness. Within coolness, the subdimensions of useful/reliable, usability and aesthetic are the most mentioned ones. This was consistent with the general Apple image reflected on several media regarding its attributes (Pisani, 2023; Beattie, 2021; McGee, 2023).

Regarding the distribution of sentiment score computed for each of the three main constructs, the sentiment was generally positive with an average of 0.270 overall (table 4.1). Looking specifically at the construct statistics and boxplots we saw that for sentences containing the loyalty construct (table 4.1 and figure 4.3 below), the mean was 0.373 vs. 0.356 in terms of median. This means the distribution is close to being symmetrical with a slight skewness to the right. The first and third quartiles, which concentrate the 50% more centrally distributed occurrences were situated between 0.190 and 0.497.

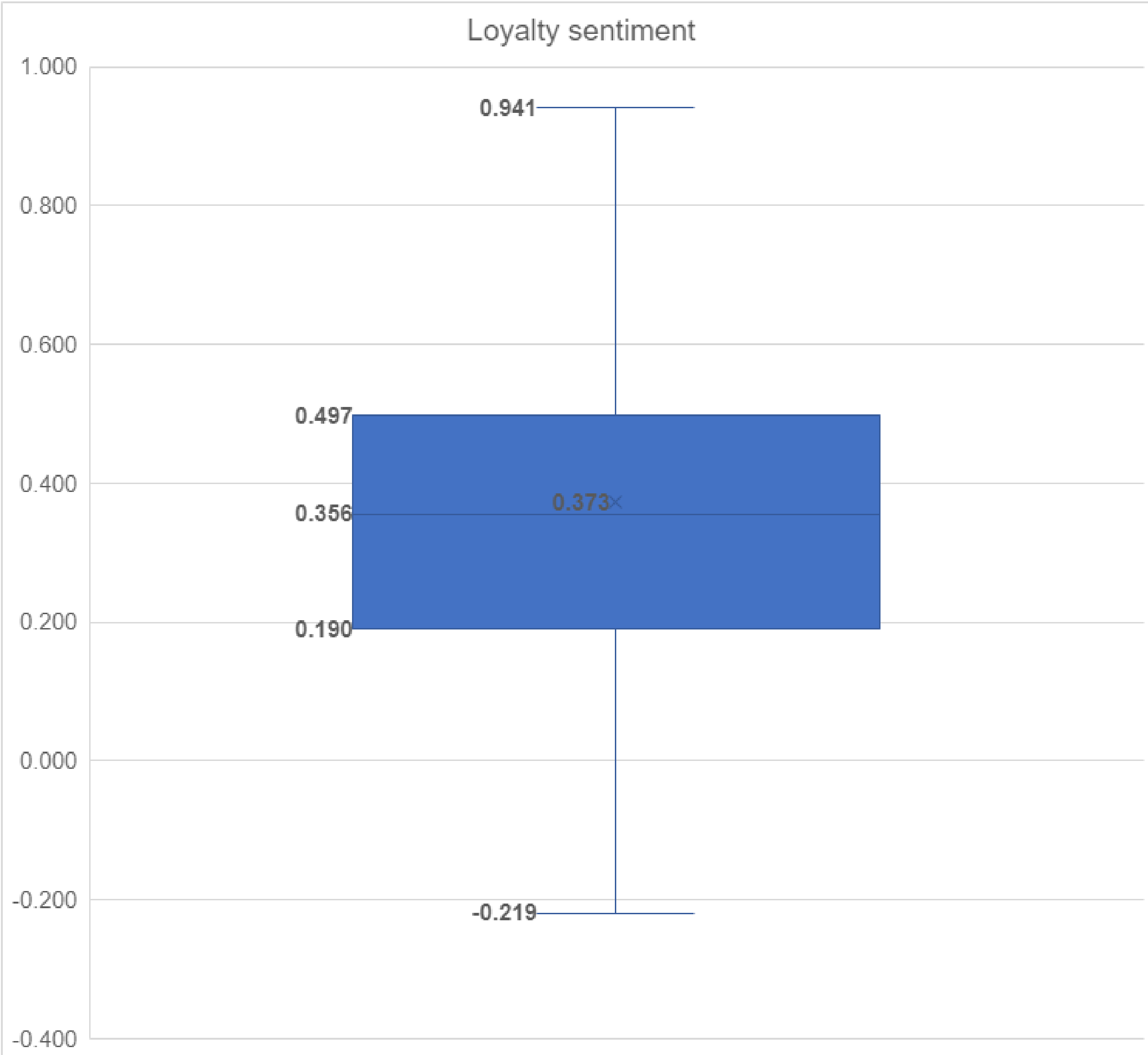

Figure 4.3: Sentiment per post: Loyalty construct (x marks the mean)

Considering the group of reviews where the coolness construct was identified (in any of its subdimensions) we can see that the mean was 0.370 vs a median of 0.375. This means the distribution was very symmetrical close to normality and with only minor skewness. The first and third quartiles range between 0.181 and 0.564 indicating a generally positive feeling towards the brand regarding the coolness construct. However, the minimum and first quartile were located at -0.365 and 0.181, respectively, indicating a generally neutral to negative feeling for the 25% reviews with the lowest sentiment scores (table 4.1 and figure 4.4).

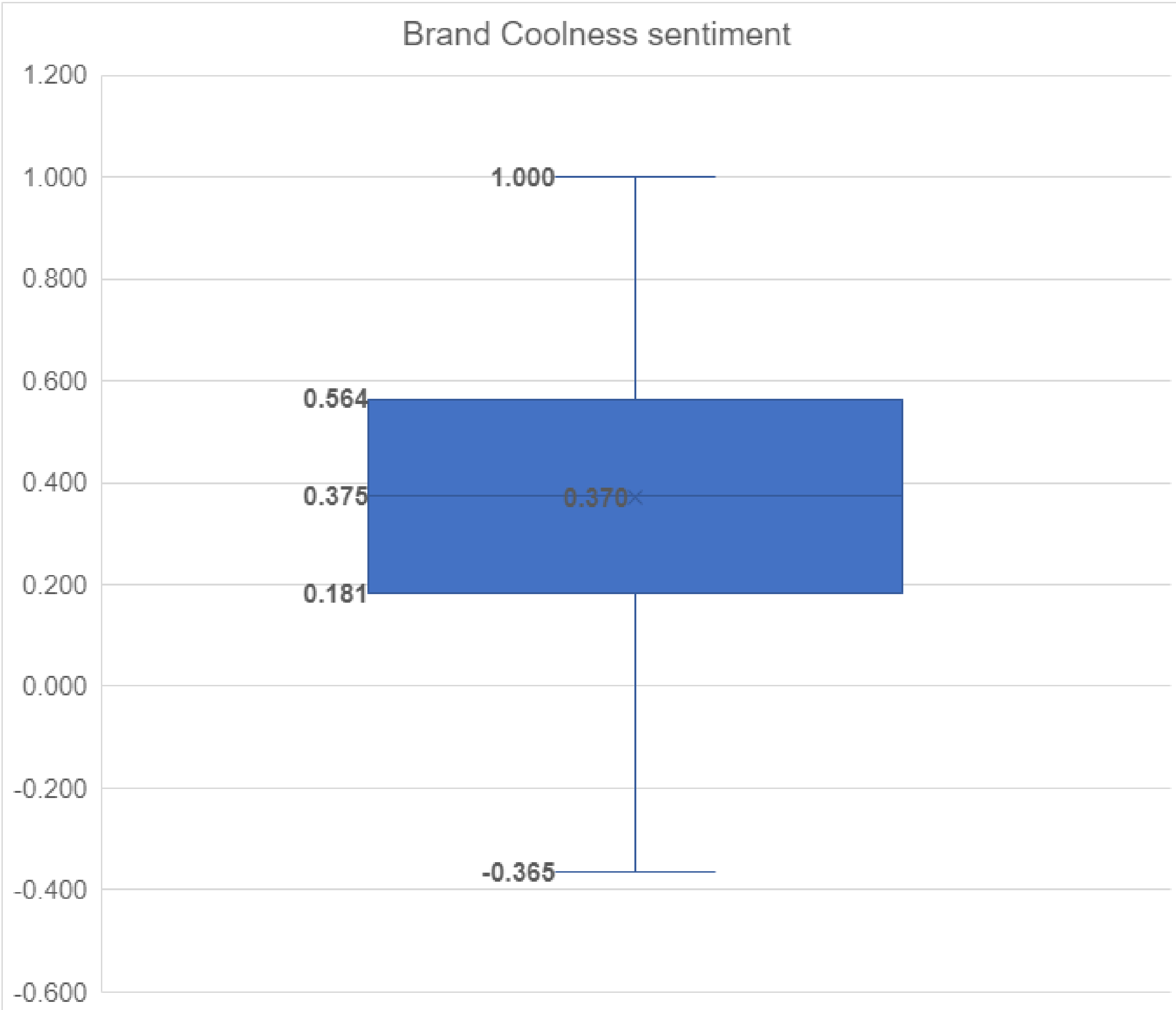

Figure 4.4: Sentiment per post - Brand coolness construct (x marks the mean)

Regarding the sentiment towards the reviews containing the brand love construct seen in figure 4.5, we had a close mean (0.388) to median (0.355), with the inter quartile range (IQR) between 0.250 and 0.518. The minimum value is –0.143 and the maximum is 0.917.

Figure 4.5: Sentiment per post - Brand love construct (x marks the mean)

In general terms, all three main constructs have a similar mean sentiment (0.373 for loyalty vs 0.370 on average for brand coolness vs 0.388 for brand love). However, the minimum value differs significantly with a minimum of –0.362 measured for loyalty compared to –0.949 for brand coolness and -0.889 for brand love. In terms of the subdimensions of coolness (and considering only the top seven subdimensions with the highest percentage of occurrences in the data), we see that extraordinary, original/innovative and energetic are the subdimensions with the higher sentiment ranging between 0.442 and 0.537. Skewness and kurtosis values help identify the normality of the data distribution. According to Hair et al. (2010), normality was confirmed for all sentiment measurements considering commonly accepted values of skewness within a maximum of 2 (and ranging between -2 and +2) and for kurtosis a maximum of 7 (between -7 and +7) (table 4.1 below).

Table 4.1: Descriptive statistics for all the constructs/subdimensions and e-WOM metrics

Descriptive statistics								
	Total occurrences	**Mean**	**Median**	**Standard deviation**	**Minimum**	**Maximum**	**Skewness**	**Kurtosis**
Overall sentiment	2319	0.270	0.257	0.218	-0.640	1.000	0.133	.695
Constructs (Sentiment scores)								
Loyalty	157	0.373	0.356	0.287	-0.362	1.000	0.220	0.151
Brand Coolness	656	0.370	0.375	0.298	-0.949	1.000	-0.308	.734
Coolness: Useful/Reliable	490	0.375	0.372	0.312	-0.688	1.000	-0.150	0.016
Coolness: Usability	148	0.386	0.415	0.311	-0.688	1.000	-0.526	.770
Coolness: Aesthetic	123	0.357	0.363	0.314	-0.949	1.000	-0.621	.323
Coolness: Extraordinary	67	0.537	0.530	0.242	-0.165	1.000	-0.065	.388
Coolness: Energetic	50	0.442	0.436	0.330	-0.265	1.000	-0.051	0.792
Coolness: Original/Innovative	30	0.461	0.414	0.290	-0.194	1.000	0.005	0.139
Coolness: High-status	22	0.387	0.287	0.396	-0.226	1.000	0.172	1.386
Coolness: Authentic	5	0.375	0.344	0.249	0.125	0.650	0.165	2.944
Coolness: Rebellious	5	0.388	0.439	0.307	0.090	0.835	0.600	0.365
Coolness: Popular	2	0.237	0.237	0.360	-0.017	0.491		
Coolness: Subcultural	1	0.048	0.048		0.048	0.048		
Coolness: Iconic	1	-0.707	-0.707		-0.707	-0.707		
Brand Love	918	0.388	0.355	0.256	-0.889	1.000	-0.244	2.011
Other metrics (Absolute values)								
Sentences per review	-	6.7	4	8.8	1	169	6.8	82.3
Word count per review	-	94.2	43	178.1	2	3184	7.0	80.6

Other metrics used in this dissertation's research consisted of the number of sentences and words per review. For the first one, sentences per review, the reviews had an average of 6.7 sentences, ranging from a minimum of 1 to maximum of 169. In terms of words per review, the mean was 94 ranging between 2 and 3184. The distribution of these variables was not normal considering the high skewness and kurtosis levels of 6.8, 82.3 respectively, for sentences and 7 and 80.6 respectively, for the word count. These metrics were standardized prior to the data analysis.

Looking at other descriptive statistics, namely correlation scores (Stewart, 2023), it is interesting to try and understand what constructs are correlated with each other. Three correlation matrices were prepared, one looking at the correlations between the three main constructs, a second one looking at the correlation between the brand coolness subdimensions with brand love and loyalty, and among themselves and a third table looking at the relationship between the three main constructs and the e-WOM volume metrics.

The first matrix shows us that all three main constructs are moderately to highly correlated to each other. The values range from 0.415 between loyalty and brand coolness, 0.529 between loyalty and brand love and 0.492 between brand love and brand coolness. All correlations between the constructs were statistically significant at the 95% significance level (table 4.2).

Table 4.2: Pearson correlation scores for the three main constructs

Correlations: Three main constructs				
		Loyalty	Brand Coolness	Brand Love
Loyalty	Pearson Correlation	**1**	**.415***	**.529***
	Sig. (2-tailed)		<.001	<.001
Brand Coolness	Pearson Correlation	**.415***	**1**	**.492***
	Sig. (2-tailed)	<.001		<.001
Brand Love	Pearson Correlation	**.529***	**.492***	**1**
	Sig. (2-tailed)	<.001	<.001	

* Correlation is significant at the 0.05 level (2-tailed).

From the second correlation matrix (table 4.3 below), between brand coolness subdimensions among themselves and with the loyalty and brand love constructs, we can see that most constructs and subdimensions are positively correlated; the only exceptions to this being the correlation scores between the energetic and high-status subdimensions with a moderate negative score of -0.408 but with a very high significance level, meaning there is no statistical evidence this result accurately represents reality. The same is true between the coolness - aesthetic and coolness - original/innovative subdimensions with a weak negative correlation of -0.138. Again, the lower sample size in some of the coolness subdimensions might help explain the higher p-values, as mentioned above.

Looking just at statistically significant correlation scores we see loyalty is moderately positively correlated with coolness: useful/reliable at 0.366 and coolness: aesthetic with 0.411. Coolness - useful/reliable is the construct with the most significant correlations with other constructs and subdimensions: 0.535 with usability, 0.470 with extraordinary, 0.646 (which is considered a high correlation score within the social sciences field; Senthilnathan, 2019) with energetic, 0.519 with aesthetic and 0.483 with original/innovative. The useful/reliable subdimension is significantly correlated with all other constructs and subdimensions except high-status (there is a weak to moderate correlation there, but which is not statistically significant). Another noticeable correlation was seen between aesthetic and high-status at 0.713 which can be considered a very high positive correlation and that is statistically significant at the 95% level (sig.=0.009). In terms of the correlations between brand love and other constructs/subdimensions, all correlation scores are positive but at different levels, with only some reaching statistical significance. Just looking at those significant relationships, brand love shows a moderate to high positive correlation with coolness - energetic at 0.582. There is also a moderate to high significant correlation with coolness - useful/reliable (0.507) and more moderate to low with coolness - aesthetic (0.376) and coolness - usability (0.242).

Table 4.3: Pearson correlation scores for brand coolness subdimensions

Correlations (Coolness subdimensions)										
		Loyalty	Coolness: Useful/Reliable	Coolness: Usability	Coolness: Extraordinary	Coolness: Energetic	Coolness: Aesthetic	Coolness: Original/Innovative	Coolness: High-status	Brand Love
Loyalty	Pearson Correlation	**1**	**.366***	**0.164**	**0.208**	**0.219**	**.411***	**0.006**	**0.576**	**.529***
	Sig. (2-tailed)		<.001	0.282	0.408	0.368	0.006	0.984	0.064	<.001
Coolness: Useful/Reliable	Pearson Correlation	**.366***	**1**	**.535***	**.470***	**.694***	**.519***	**.483***	**0.208**	**.507***
	Sig. (2-tailed)	.001		<.001	0.004	<.001	<.001	0.023	0.564	<.001
Coolness: Usability	Pearson Correlation	**0.164**	**.535***	**1**	**0.374**	**0.303**	**0.209**	**0.229**	**0.126**	**.242***
	Sig. (2-tailed)	0.282	<.001		0.066	0.207	0.208	0.360	0.746	0.015
Coolness: Extraordinary	Pearson Correlation	**.208**	**.470***	**0.374**	**1**	**0.198**	**0.333**	**-0.085**	**0.546**	**0.059**
	Sig. (2-tailed)	0.408	0.004	0.066		0.517	0.140	0.755	0.103	0.72
Coolness: Energetic	Pearson Correlation	**0.219**	**.694***	**0.303**	**0.198**	**1**	**0.472**	**0.179**	**-0.408**	**.582***
	Sig. (2-tailed)	0.368	<.001	0.207	0.517		0.056	0.621	0.422	<.001
Coolness: Aesthetic	Pearson Correlation	**.411***	**.519***	**0.209**	**0.333**	**0.472**	**1**	**-0.138**	**.713***	**.376***
	Sig. (2-tailed)	0.006	<.001	0.208	0.14	0.056		0.598	0.009	<.001
Coolness: Original/Innovative	Pearson Correlation	**0.006**	**.483***	**0.229**	**-0.085**	**0.179**	**-0.138**	**1**	**0.027**	**0.029**
	Sig. (2-tailed)	0.984	0.023	0.36	0.755	0.621	0.598		0.973	0.901
Coolness: High-status	Pearson Correlation	**0.576**	**0.208**	**0.126**	**0.546**	**-0.408**	**.713***	**0.027**	**1**	**0.399**
	Sig. (2-tailed)	0.064	0.564	0.746	0.103	0.422	0.009	0.973		0.101
Brand Love	Pearson Correlation	**.529***	**.507***	**.242***	**0.059**	**.582***	**.376***	**0.029**	**0.399**	**1**
	Sig. (2-tailed)	<.001	<.001	0.015	0.720	<.001	<.001	0.901	0.101	

* Correlation is significant at the 0.05 level (2-tailed).

Finally, Pearson coefficient scores were calculated for the association between the main constructs under analysis (loyalty, brand coolness and brand love) and the e-WOM volume metrics. Looking at the matrix in table 4.4 below, we can see that sentence and word count have a weak negative correlation with brand coolness and brand love between -0.148 and –0.191. This means there is a slight tendency for more negative reviews mentioning those constructs to be slightly bigger in terms of length, even though the correlation score is low. The correlation between the e-WOM volume metrics and loyalty was very weak but deemed not significant. One explanation for this might be the lower sample size of 157 for the loyalty construct compared to the coolness (656) and love (918) constructs, which can amplify the variability found in the data and increase the significance value (Field, 2009).

Table 4.4: Pearson correlation scores for the three main constructs with e-WOM volume metrics

Correlations						
		Loyalty	Brand Coolness	Brand Love	e-WOM: Sentence count	e-WOM: Word count
e-WOM: Sentence count	Pearson Correlation	**.024**	**-.156***	**-.148***	**1**	**.895***
	Sig. (2-tailed)	.761	<.001	<.001		<.001
e-WOM: Word count	Pearson Correlation	**.040**	**-.191***	**-.167***	**.895***	**1**
	Sig. (2-tailed)	.615	<.001	<.001	<.001	

* Correlation is significant at the 0.05 level (2-tailed).

4.3. Inferential statistics and regression models

Following the correlation analysis, it was time to test the model and the causal relationships between the constructs, coolness subdimensions and e-WOM metrics. A first test was done on calculating the regression coefficients between the subdimensions of brand coolness and the constructs of brand love, and loyalty; however, for a big percentage of the coolness subdimensions only a few cases were present, which would yield models with irregular fit and lacking statistical significance (Field, 2009); hence the decision was made to include in the regression models only the main brand coolness construct together with the other aforementioned constructs. To test brand coolness as a variable in the model, an arithmetic average was created from all the coolness subdimensions present in each review. This was then used as the sentiment value for brand coolness. All regression models were validated having tolerance (TOL) and variance inflation factor (VIF) values above 0.1 and below 10, respectively, showing acceptable levels of collinearity (Field, 2009; Regorz, 2020).

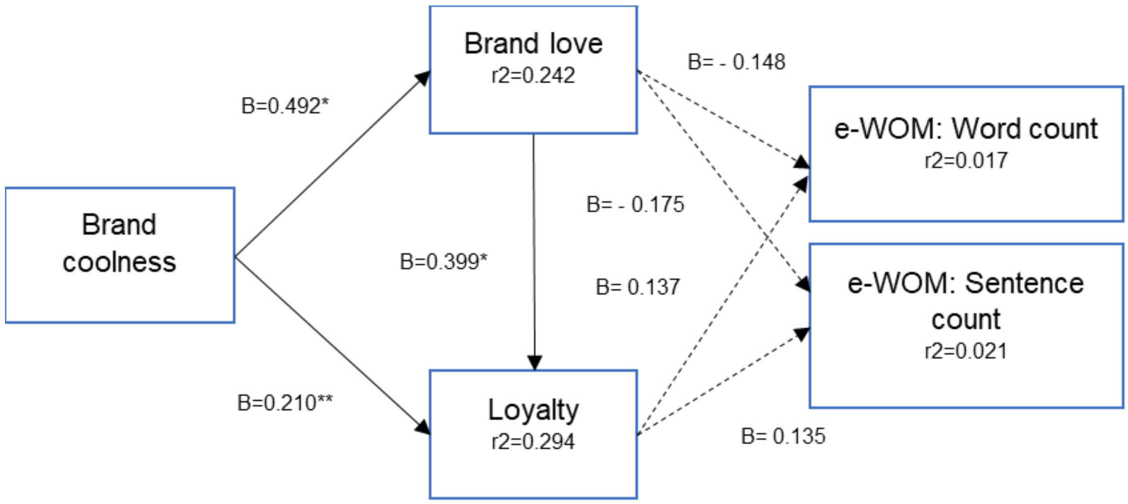

Figure 4.6: Model with the corresponding R square values and standardized beta coefficients (relationships were significant at a 5% or 10%** level except where in dashed line)*

The first hypothesis, "H1: Brand coolness positively influences Brand love" (equation model: Brand Love = B0 + B1 Brand Coolness + E) was confirmed by the regression results. The r square value of 0.242 suggests that 24.2% of the variation in brand love could be explained by the variation in brand coolness. Looking at the beta coefficient we can see that its value of 0.492 represents a moderate positive linear relationship between brand coolness (independent variable) and brand love (dependent variable), meaning that as the independent variable increases by one standard deviation, the dependent variable is expected to increase by approximately 0.492 standard deviations, assuming the linear relationship. The significance level below 0.001 means this conclusion is statistically significant and therefore H1 is supported (tables 4.5 and 4.6).

Table 4.5: Brand coolness > Brand love regression model fit

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.492[a]	.242	.240	.22499870	2.054

a. Predictors: (Constant), Brand Coolness

b. Dependent Variable: Brand Love

Table 4.6: Brand coolness > Brand love regression model coefficient estimates

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	.226	.020		11.281	<.001		
	Brand Coolness	.460	.042	.492	10.930	<.001	1.000	1.000

a. Dependent Variable: Brand Love

Regarding the model results for “H2: Brand coolness positively influences Loyalty” and “H3: Brand love positively influences Loyalty” (equation model: Loyalty = B0 + B1 Brand Coolness + B2 Brand love + E), we can conclude that these two constructs have a moderate influence on Loyalty. At 0.294, the r square shows that brand coolness and brand love explain around ~29% of the variation of loyalty (table 4.7). The betas show those relationships are of moderate intensity when it comes to the impact of brand love on loyalty, with a standardized beta coefficient of 0.399 and a lower intensity for the impact of brand coolness on loyalty, with a beta of 0.210 (table 4.8). Results were statistically significant at a 5% level for the first relationship with a p-value of 0.002, although only at a 10% significance level for the second one, with a p-value of 0.093. Still, considering 10% the significance level, we can say that both H2 and H3 are supported.

Table 4.7: Brand coolness and Brand love > Loyalty regression model fit

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.542[a]	.294	.272	.23369971	1.866

a. Predictors: (Constant), Brand coolness, Brand love

b. Dependent Variable: Loyalty

Table 4.8: Brand coolness and Brand love > Loyalty regression model coefficient estimates

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	.106	.064		1.661	.101		
	Brand love	.440	.136	.399	3.242	.002	.708	1.413
	Brand coolness	.237	.139	.210	1.707	.093	.708	1.413

a. Dependent Variable: Loyalty

Some investigation was done regarding the standardization of data. Standardization ensures that variables with different units or scales do not dominate the regression model's optimization process. It is especially important when there are variables measured in different units or with significantly different ranges (Field, 2009). When we center the independent variable, the intercept represents the expected value of the dependent variable when the centered independent variable is equal to zero. This can make it easier to explain the model's predictions. For these reasons, data was standardized regarding the regression models involving the sentiment scores for the constructs and the volume metrics about e-WOM.

Looking at the results (equation model: e-WOM/ word count = B0 + B1 Brand love + B2 Loyalty + E) we can say that brand love and loyalty seem to have little to no impact on the word count e-WOM volume metric. R square is very low (0.017) and the standardized beta coefficient is -0.184 for the impact of brand love and 0.137 for the impact of loyalty on word count, denoting a low negative impact for the first independent variable and a low positive impact for the second. Overall, these results indicate a low influence of brand love and loyalty on the length of the review in the number of words that the customer decides to share with others. However, it is important to note results were not significant (sig.= 0.225 and 0.260, respectively). Taking all of this into account we can say that H4a: Brand Love positively influences e-WOM/ word count and H5a: Loyalty positively influences e-WOM/ word count are not supported (tables 4.9 and 4.10).

Table 4.9: Brand love and Loyalty > Word count regression model fit

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.129[a]	.017	-.002	2.42316412	1.935

a. Predictors: (Constant), Loyalty, Brand love

b. Dependent Variable: Word count

Table 4.10: Brand love and Loyalty > Word count regression model coefficient estimates

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	1.218	.236		5.168	<.001		
	Brand love	-.358	.293	-.148	-1.221	.225	.652	1.533
	Loyalty	.352	.311	.137	1.131	.260	.652	1.533

a. Dependent Variable: Word count

The conclusions are very similar when looking at the influence of brand love and loyalty on the sentence count e-WOM volume metric (equation model: e-WOM/ sentence count = B0 + B1 Brand love + B2 Loyalty + E). There is an r square of 0.021 showing a low impact of both independent variables on the dependent variable. Standardized beta coefficients are -0.175 and 0.135 for the impact of brand love and loyalty on the sentence counts of each review, respectively. This shows that there is a low negative impact of brand love and a low positive impact from loyalty as seen with the word count analysis. Again, these correlation scores were both not significant (significance levels of 0.150 and 0.266) which means there is no strong evidence to suggest that brand love and loyalty have any meaningful impact on the dependent variable sentence count. Likewise, we do not support H4b, that brand love positively influences e-WOM/ sentence count and H5b, that loyalty positively influences e-WOM/ sentence count (tables 4.11 and 4.12 below).

Table 4.11: Brand love and Loyalty > Sentence count regression model fit

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.145[a]	.021	.002	2.37423849	1.958

a. Predictors: (Constant), Loyalty, Brand love

b. Dependent Variable: Sentence count

Table 4.12: Brand love and Loyalty > Sentence count regression model coefficient estimate

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	1.116	.231		4.832	<.001		
	Brand love	-.417	.287	-.175	-1.450	.150	.652	1.533
	Loyalty	.340	.304	.135	1.117	.266	.652	1.533

a. Dependent Variable: Sentence count

5. Conclusion

5.1. Discussion

This study aimed to understand how brand coolness, brand love, brand loyalty and e-WOM (in volume) relate to each other in the context of a tech brand, using user-generated content. A conceptual model is proposed, and data was scrapped from Amazon in the form of customer reviews regarding Apple. Text-mining and sentiment analysis techniques in R software (The R Project for Statistical Computing, 2023) were used to compute sentiment scores and other metrics such as sentence and word count. Those metrics were then analyzed with statistical analysis techniques and several regression models were tested.

In general terms, the most predominant construct within the reviews Apple customers left on Amazon, was brand love followed by brand coolness. Within coolness, the subdimensions of useful/reliable, usability and aesthetic are the most mentioned ones. This was also consistent with the general Apple image reflected on several media regarding its attributes (Pisani, 2023; Beattie, 2021; McGee, 2023).

Looking at other results found in this dissertation we see that they point towards similar conclusions to those found by Tiwari et al. (2021) and others, in the sense that brand coolness is seen as having a positive influence on brand love. That same study, however, found a stronger influence having a measured r square of 0.780 and a beta coefficient of 0.880 (measured through confirmatory factor analysis), when compared to what was found in the present research (0.242 in terms of r square and 0.492 in terms of standardized beta, measured through linear regression modeling). Still, and although that paper served as source to build the coolness construct dictionary, it was based on an online survey, measuring the construct brand love through a 7-point Likert scale, shared through a Facebook advert. Respondents' ages ranged from 18 to 35 years. In the present research data was unstructured, free format, and unprompted, taken from Amazon reviews in the US and with no age group defined. Most importantly, the present research took data gathered from customer direct feedback, instead of using a questionnaire. In the end, we can still say both studies found a definitive causal positive effect of brand coolness on brand love.

Some of the brand coolness subdimensions were not very present in the data for this dissertation, namely the popular, subcultural, iconic and high-status subdimensions, which corroborates Warren et al. (2019) on their brand coolness subdimensions analysis. In the latter study, which also served as a base to prepare the construct dictionary, we see the same subdimensions had lower factor loadings.

Other authors investigated the connections between brand love and loyalty, showing brand love plays a significant role, in both behavioral and attitudinal loyalty (Robertson et al., 2022; Albert & Merunka, 2013). Robertson et al. (2022) found that brand love acted as an antecedent of brand loyalty. These findings were corroborated by this dissertation’s results. Still regarding the loyalty construct and its determinants, a study by Farahdiba (2023) conducted through a survey in Indonesia about coffee shops, found that brand love has a direct effect on loyalty. Again, this is corroborated by this dissertation’s since we found ~29% influence considering the linear regression model including brand love and coolness as the independent variables, and a standardized beta coefficient of 0.399 in terms of the relationship between brand love and loyalty. Research from Jiménez-Barreto et al. (2022) found that the connection between brand coolness and loyalty also encompassed some causality with standardized path coefficients of 0.580, in their study about music festival brands, and 0.590, for fast the food restaurant brands study, which compare to this dissertation’s result of 0.210 for a tech brand; it is noteworthy that the relationship between brand coolness and loyalty was only significant at the 10% level in this dissertation, on the upper limit of what is commonly accepted in terms of statistical significance testing. Jiménez-Barreto et al. (2022) study was based on a focus group with 22 participants, which again differs from the method pursued in the present research of text-mining and sentiment analysis. Still, it is interesting that the results for the relationship between brand coolness and loyalty in Jiménez-Barreto et al. (2022) were weaker than the relationship between brand love and loyalty in the other research mentioned before (Tiwari, 2021; Farahdiba, 2023). This is coherent with the results found in this dissertation where both relationships (brand love > loyalty and brand coolness > loyalty) were analyzed at the same time and the relationship between brand love and loyalty was stronger than the relationship between brand coolness and loyalty. Carroll and Ahuvia (2006) found that brand love contributes to 17% of loyalty in terms of r square, with a path coefficient of 0.250, all factors being considered, which points towards the same direction as the conclusions reached in the present study, although in our case the r square pointed towards a stronger combined influence (with brand coolness) of 27% and a path coefficient of 0.399.

5.2. Contribution

This dissertation aimed at proposing a different measurement approach for such constructs as brand coolness, brand love and their relationship to loyalty, which is a much-desired result in brand management and marketing across different business sectors (Oliver, 1999). A secondary topic was also introduced regarding the impact of these constructs on e-WOM volume. The results of this research contribute to the marketing study of brands, and particularly tech brands. The novelty of this research was in proposing a new process, using text mining and sentiment analysis techniques to extract value from the vast amounts of data available online, to measure consumer perception, complex brand perception constructs and their interactions with consumer loyalty. This will help companies understand their consumers' perceptions about their brand, leading to improved marketing strategies and better business performance. The same methodology used in this dissertation can be adapted by companies wanting to achieve the same goals.

Fostering brand coolness and brand love has been defined as one of the very important actions for marketers and brand strategists (Warren, 2019; Farahdiba, 2023) and its measurement is one of the key parts of marketers' work, for which this dissertation contributes to. Warren and Campbell (2014) defended measuring brand coolness should involve direct assessments of consumers' perceptions, for which this dissertation proposes a more direct method from what was seen across research analyzed during the literature review stage. In their study within the field of technology, Tiwari et al. (2021) have argued that the increasing similarity among technology products makes it difficult for consumers to differentiate and choose one product over another with "coolness" emerging as a crucial factor in product evaluation and differentiation. In addition to that research, this dissertation allowed to measure which subdimensions of brand coolness were most mentioned in reviews by tech consumers.

More broadly, the literature review allowed to identify two major gaps that were covered with the present dissertation: first, the common reliance on surveys to provide data for measurement of such constructs, ignoring the wealth of data the current online landscape and consumer habits allow to collect (Lee, 2018); second, the lack of research around these constructs on tech brands specifically. Both these gaps were met with the present dissertation. In addition to these two main gaps, we can add the combined measurement of brand coolness, brand love and loyalty and their relationships at the same time, and the use of e-WOM volume metrics, which were not present in any research analyzed during the literature review stage.

On this dissertation, we do not see a causation relationship between the sentiment the consumer has in terms of brand love and loyalty and the volume of e-WOM, which is partly corroborated by De Valck et al. (2019).

5.3. Managerial Implications

The positive influential relationship of brand coolness on loyalty suggests that marketing efforts should be directed towards enhancing the perceived coolness of brands. For tech brands specifically, this might involve emphasizing usefulness and reliability, usability, aesthetical and innovation-related features, as these were identified as being the factors that contribute the most to the brand's coolness image.

Similarly, since brand love is also positively related to loyalty, brand management should focus on building emotional connections with customers. This could involve creating meaningful brand stories, engaging in community-building activities, or product features that resonate emotionally. The impact of coolness on brand love was also identified, prompting companies to invest in the features that can make their brand appear cool in the customer’s mind.

Generally speaking, customer relationship management emerges as crucial to build long-term loyalty, which confirms the importance of this dissertation’s research and of looking into online customer-brand interactions, and then analyzing its content, which can present a great amount of value for the marketing strategy of a product or service (Greve, 2014; Kato, 2021). Management should closely monitor customer reviews for both positive and negative sentiments. Addressing negative sentiments promptly and positively can contribute to brand reputation management. Consumer preferences and perceptions can change over time; therefore, it is essential for management to continually monitor customer feedback, market trends, and competitors. This allows for adaptive strategies that respond to evolving customer needs and expectations. The research conducted in this dissertation presents a method of how to monitor consumer perceptions and sentiment towards a brand, specifically a tech brand, focusing on specific brand constructs. More investment should be made by companies in using and developing data analysis activities, namely text mining and sentiment analysis techniques combined with better data engineering and tooling resources. Those will increase the potential impact of a model like the one presented here and amplify its benefits.

The word and sentence count in reviews (e-WOM volume) did not show a strong impact from brand love and loyalty. In addition, those results were not statistically significant.

5.4. Limitations and Future Research

The construct dictionaries were built specifically for data related to tech brands, and also adapted to Apple, in order to capture the constructs more accurately. This, however, makes it difficult to use the exact same dictionaries in other business sectors and even in other tech brands without some prior adaptation being conducted. In addition, the data gathered concerned free-form reviews consumers made regarding Apple as a brand but also Apple's products. Future research might try to filter data in order to include only brand-related reviews to evaluate the constructs studied on this dissertation.

Linear regression models were used to test the causation effects of the constructs on each other and on e-WOM volume metrics. It was determined to be the model with the best fit for the data that was gathered considering there was not enough data to build more complex models, due to some technical limitations such as the limits to data scrapping from Twitter and Instagram created in late 2022/early 2023 and the lack of proper data engineering and data cleaning tools. These would facilitate the process of data extraction from multiple sources, in multiple formats and its preparation and integration into one big data set for analysis. Future research could test the same relationships using more complex modelling techniques by gathering a larger volume of data.

The analytics technology used in this dissertation consisted of the R analysis software free packages; however, this field is in constant technological development (Dahiya et al., 2021) which can provide new tools to perform similar analysis to this one in the future and which should be explored.

Some authors have shown that brand love and loyalty are associated with willingness to pay a premium price and brand personality (Albert & Merunka, 2013; Carroll & Ahuvia, 2006; Agamudainambhi et al., 2022; Yodpram & Intalar, 2020). These relationships could then be explored using the same text mining and sentiment analysis techniques. An interesting part of the research conducted by Kiss et al. (2022) was that higher prices lead to decreased consumer utility, indicating price sensitivity in specific product ranges, decreasing brand loyalty, in chocolate brands. Although being a different product from tech, it would still be interesting to see if the same can be found in comments or UGC where price is specifically mentioned, considering sometimes tech products (including Apple's) are seen as expensive (Schroeder, S., 2023; Kelly, 2023).

Johnson et al. (2006) had found that evaluations of perceived value, brand equity, affective commitment and loyalty intentions are not completely new each time but rather updated versions of previous evaluations. Given the design of the present dissertation, the effects of time on the brand constructs were not measured but this can become an avenue for future research, seeking to understand how consumer perceptions of the different constructs and the relationships between them evolve over time.

Research on product extensions and their impact on brand loyalty (Zhang et al., 2023) found that extension products with high perceived fit to the brand positively enhance brand loyalty, while those with low perceived fit have a negative impact on brand loyalty. Data regarding the different Apple product lines was combined and analyzed as one dataset for the purpose of this dissertation, but a future study could analyze a brand considering its different product lines and extensions. Regarding brand loyalty, Oliver (1999) mentioned the concept of consumer self-isolation, where the consumer actively blocks or screens out competitor information to their preferred brand. Hence, a future study could make the comparison between different brands and see how the constructs are perceived between them by consumers.

References

Aaker, D. A., & Joachimsthaler, E. (2000). Brand Leadership. (Illustrated ed.). Free Press. ISBN 0684839245,9780684839240

Agamudainambhi, C. A., Mamun, A. A., Jayashree, S., Naznen, F., & Abir, T. (2022). Modelling the significance of social media marketing activities, brand equity and loyalty to predict consumers' willingness to pay premium price for portable tech gadgets*. Heliyon, 8*, e10145. https://doi.org/10.1016/j.heliyon.2022.e10145

Albert, N., & Merunka, D. (2013). The role of brand love in consumer-brand relationships. *Journal of Consumer Marketing*, 30(3), 258–266. https://doi.org/ 10.1108/07363761311328928

Amazon (2023). "Amazon.com Privacy Notice." Amazon Help Center, https://www.amazon.com/gp/help/customer/display.html?nodeId=GX7NJQ4ZB8MHFRNJ

Amazon (2023). "Understanding Customer Reviews and Ratings." Amazon Help Center, https://www.amazon.com/gp/help/customer/display.html?nodeId=G8UYX7LALQC8V9KA

Anggraeni, A., & Rachmanita. (2015). Effects of Brand Love, Personality, and Image on Word of Mouth: The Case of Local Fashion Brands Among Young Consumers. *Procedia - Social and Behavioral Sciences*, 211, 442-447. https://doi.org/10.1016/j.sbspro.2015.11.058

Arnett, D., German, S., & Hunt, S. (2003). The identity salience model of Relationship Marketing success: The case of Nonprofit Marketing. *Journal of Marketing*, 67 (4), 89-105. https://doi.org/10.1509/jmkg.67.2.89.1861

Attiq S., Abdul Hamid A. B., Khokhar M. N., Shah H. J. & Shahzad A. (2022) “Wow! It’s Cool”: How Brand Coolness Affects the Customer Psychological Well-Being Through Brand Love and Brand Engagement. *Front. Psychol.* 13:923870. https://doi.org/10.3389/fpsyg.2022.923870

Beattie, A. (2021, September 10). How Did Apple Get So Big? The Story Behind Apple's Success. https://www.investopedia.com/articles/personal-finance/042815/story-behind-apples-success.asp

Bernritter, S. F., Loermans, A. C., Verlegh, P. W. J., & Smit, E. G. (2017). ‘We’ are more likely to endorse than ‘I’: The effects of self-construal and brand symbolism on consumers’ online brand endorsements. *International Journal of Advertising*, 36(1), 107–120. https://doi.org/10.1080/02650487.2016.1186950

Boslaugh, S., & McNutt, L-A. (2008). "Structural Equation Modeling". *Encyclopedia of Epidemiology*. doi 10.4135/9781412953948.n443, ISBN 978-1-4129-2816-8

Cambridge Dictionary Thesaurus (2023). In Cambridge Dictionary [Online]. Available at: https://dictionary.cambridge.org/thesaurus/

Carroll, B., & Ahuvia, A. (2006). Some antecedents and outcomes of brand love. *Marketing Letters*, 17(2), 79–89. https://doi.org/10.1007/s11002-006-4219-2

Chen, F., Quadri-Felitti, D., & Mattila, A. S. (2021). Generation Influences Perceived Coolness But Not Favorable Attitudes Toward Cool Hotel Brands. *Cornell Hospitality Quarterly*, 1-9. doi: 10.1177/19389655211031442

Chen, T., Samaranayake, P., Cen, X., Qi, M., & Lan, Y. (2022). The Impact of Online Reviews on Consumers' Purchasing Decisions: Evidence From an Eye-Tracking Study. *Frontiers in Psychology*, 13, Article 865702. https://doi.org/10.3389/fpsyg.2022.865702

Chou, C. Y., & Sawang, S. (2015). Virtual community, purchasing behavior, and emotional well-being. *Australasian Marketing Journal*, 23(3), 161-167. https://doi.org/10.1016/j.ausmj.2015.06.001

Colliander, J., Wien, A.H., 2013. Trash talk rebuffed: consumers' defense of companies criticized in online communities. *Eur. J. Mark*. 47 (10), 1733–1757. https://doi.org/10.1108/EJM-04-2011-0191

Dahiya, A., Gautam, N., & Gautam, P. K. (2021). Data Mining Methods and Techniques for Online Customer Review Analysis: A Literature Review. *Journal of System and Management Sciences*, 11(3), 1-26. https://doi.org/10.33168/JSMS.2021.0301

De Valck, K., van Bruggen, G. H., & Wierenga, B. (2009). Virtual communities: A marketing perspective. *Decision Support Systems*, 47(3), 185-203. https://doi.org/10.1016/j.dss.2009.02.008

Dictionary.com (2023). In Dictionary.com [Online]. Available at: https://www.dictionary.com/

Farahdiba, D. (2023). Quo Vadis Brand Love? Role of Cognition-Affection-Behavior Model for Local Coffee Shops in Indonesia. *Business Management Analysis Journal* (BMAJ), 6(1), 90-108. DOI:10.24176/bmaj.v6i1.9741

Field, A. (2009). Discovering Statistics Using SPSS: (and sex and drugs and rock 'n' roll) (3rd ed.). Sage Publications. ISBN 978-1-84787-906-6, ISBN 978-1-84787-907-3.

Fournier, S. (1998). Consumers and their brands: developing relationship theory in consumer research. *Journal of Consumer Research*, 24, 343- 373. https://doi.org/10.1086/209515

Fournier, S., & Mick, D.G. (1999). Rediscovering satisfaction. *Journal of Marketing*, 63, 5–23. http://dx.doi.org/10.2307/1251971

Greve, G. (2014). The Moderating Effect of Customer Engagement on the Brand Image – Brand Loyalty Relationship. *Procedia - Social and Behavioral Sciences*, 148, 203-210. https://doi.org/10.1016/j.sbspro.2014.07.035

Hair, J.F., Black, W.C., Babin, B.J. and Anderson, R.E. (2010) Multivariate Data Analysis (7th ed.). Pearson, New York. ISBN 0138132631, 978-0138132637

Hawkins, D. I., Best, R. J., & Coney, K. A. (2004). Consumer Behavior: Building Marketing Strategy, Volume 2 (9th ed.). McGraw-Hill Irwin. ISBN 0072536861, 9780072536867

Hearst, M. (1999). Untangling Text Data Mining. School of Information Management & Systems, University of California, Berkeley. *Proceedings of ACL'99: the 37th Annual Meeting of the Association for Computational Linguistics*. Retrieved from http://www.sims.berkeley.edu/-hearst

Hu, X., & Ha, L. (2015). Which Form of Word-Of-Mouth is more important to Online Shoppers? A Comparative Study of WOM Use between General Population and College Students. *Journal of Communication and Media Research*, 7(2), 15–35.

IBM SPSS Software (2023). https://www.ibm.com/spss

Instagram (2023). Instagram Help Center. https://help.instagram.com/740480200552298

Instant Data Scrapper (2023). Web Robots. https://webrobots.io/

Jiménez-Barreto, J., Loureiro, S. M. C., Rubio, N., & Romero, J. (2022). Service brand coolness in the construction of brand loyalty: A self-presentation theory approach. *Journal of Retailing and Consumer Services*, 65, 102876. ISSN 0969-6989. https://doi.org/10.1016/j.jretconser.2021.102876.

Johnson, M. D., Herrmann, A., & Huber, F. (2006). The evolution of loyalty intentions. *Journal of Marketing*, 70(2), 122-132. https://doi.org/10.1509/jmkg.70.2.122

Kato, T. (2021). Factors of loyalty across corporate brand images, products, dealers, sales staff, and after-sales services in the automotive industry. *25th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems* (pp. 1411-1419). https://doi.org/10.1016/j.procs.2021.08.144

Keller, K.L. (1993). Conceptualizing, measuring, and managing customers-based brand equity. *Journal of Marketing. Management*. 57 (1), 1–12. https://doi.org/10.1177/002224299305700101

Kelly, G. (2023, August 11). Apple iPhones, iPads And MacBooks Are Getting More Expensive. Forbes. https://www.forbes.com/sites/gordonkelly/2023/01/25/apple-iphones-ipads-and-macbooks-are-getting-more-expensive/

Kiss, M., Czine, P., Balogh, P., & Szakaly, Z. (2022). The connection between manufacturer and private label brands and brand loyalty in chocolate bar buying decisions – A hybrid choice approach. *Appetite,* 178, 106145. https://doi.org/10.1016/j.appet.2022.106145

Kudeshia, C., Kumar, A. (2017). Social e-WOM: does it affect the brand attitude and purchase intention of brands. Manag. Res. Rev. 40 (3), 310–330. https://doi.org/10.1108/MRR-07-2015-0161

Lawrence, B., Fournier, S., & Brunel, F. (2013). When companies Do not make the ad: A multimethod inquiry into the differential effectiveness of consumer-generated advertising. *Journal of Advertising*, 42(4), 292– 307. https://doi.org/10.1080/00913367.2013.795120

Lee, In. (2018). Social media analytics for enterprises: Typology, methods, and processes. *Business Horizons*, 61(2), 199-210. https://doi.org/10.1016/j.bushor.2017.11.002

Mathwick, C., Wiertz, C. & De Ruyter, K. (2008). Social capital production in a virtual P3 community. J. Cons. Res. 34 (6), 832–849. https://doi.org/10.1086/523291

McGee, P. (2023, May 17). How Apple captured Gen Z in the US — and changed their social circles. Financial Times. https://www.ft.com/content/8a2e8442-449e-4dbd-bd6d-2656b4503526

Merriam-Webster Dictionary (2023). In Merriam-Webster [Online]. https://www.merriam-webster.com

Microsoft Office: Microsoft Excel (2023). https://www.microsoft.com/en-us/microsoft-365/excel

Nguyen, C. and Romaniuk, J. (2014) Pass it on: A framework for classifying the content of word-of-mouth. *Australian Marketing Journal* Vol. 22, Issue 2, 117-124. https://doi.org/10.1016/j.ausmj.2013.12.014

Pisani, B. (2023, May 10). Apple versus the world: The iPhone maker is bigger than almost any stock market in the world. https://www.cnbc.com/2023/05/10/apple-vs-the-world-apples-bigger-than-entire-overseas-stock-markets-.html

Oliver, R. L. (1997), Satisfaction: A Behavioral Perspective on the Consumer. New York: Irwin/McGraw-Hill. (1st ed.) ISBN: 0070480257, 9780070480254

Oliver, R. L. (1999). Whence Consumer Loyalty? *Journal of Marketing, 63(Fundamental Issues and Directions for Marketing),* 33-44. http://www.jstor.org/stable/1252099

Regorz, A. (2020). How to interpret a Collinearity Diagnostics table in SPSS. Regorz Statistik/Regorz Statistic. http://www.regorz-statistik.de/en/collinearity_diagnostics_table_SPSS.html

Rheingold, H. (1993). The Virtual Community: Homesteading on the Electronic Frontier. Addison-Wesley, Boston, MA. https://doi.org/10.16997/wpcc.206

Rinker, T. (2021). Package “sentimentr”: Calculate Text Polarity Sentiment. https://cran.r-project.org/web/packages/sentimentr/sentimentr.pdf

Robertson, J., Botha, E., Ferreira, C., & Pitt, L. (2022). How deep is your love? The brand love-loyalty matrix in consumer-brand relationships. *Journal of Business Research*, 149, 651-662

Rosenbaum, M.S., Massiah, C.A. (2007). When customers receive support from other customers: exploring the influence of intercustomer social support on customer voluntary performance. J. Serv. Res. 9 (3), 257–270. https://doi.org/10.1177/10946705062958

Roy, G., Datta, B., Mukherjee, S., Basu, R. (2020). Effect of e-WOM stimuli and e-WOM response on perceived service quality and online recommendation. Tour. Recreat. Res. 46 (4), 457–472. https://doi.org/10.1080/02508281.2020.1809822

Runyan, R. C., Noh, M., & Mosier, J. (2013). What is cool? Operationalizing the construct in an apparel context. *Journal of Fashion Marketing and Management*, 17(3), 322–340

Schroeder, S. (2023, June 7). Apple Vision Pro is incredibly expensive, but not for the reason you think. Mashable. https://mashable.com/article/apple-vision-pro-why-so-expensive

Senthilnathan, S. (2019). Usefulness of Correlation Analysis. https://ssrn.com/abstract=3416918 or http://dx.doi.org/10.2139/ssrn.3416918

Smith, A. & Anderson, M. (2016, December 19). Online Shopping and E-Commerce: Online Reviews. Pew Research Center. https://www.pewresearch.org/internet/2016/12/19/online-reviews

Statista (2023, August 9). Apple's revenue share by operating segment 2012-2023, by quarter. https://www.statista.com/statistics/382260/segments-share-revenue-of-apple/

Statista (2023, August 12). Global smartphone market share from 4th quarter 2009 to 1st quarter 2023 (by vendor). https://www.statista.com/statistics/271496/global-market-share-held-by-smartphone-vendors-since-4th-quarter-2009/

Stewart, K. (2023). Pearson's correlation coefficient. Encyclopedia Britannica, https://www.britannica.com/topic/Pearsons-correlation-coefficient

Tan, K., Lee, C., & Lim, K. (2023). A Survey of Sentiment Analysis: Approaches, Datasets, and Future Research. Appl. Sci., 13(7), 4550. https://doi.org/10.3390/app13074550.

The R Project for Statistical Computing (2023). https://www.r-project.org/

Tiwari, A., Chakraborty, A., & Maity, M. (2021). Technology product coolness and its implication for brand love. *Journal of Retailing and Consumer Services,* 58, 102258. https://doi.org/10.1016/j.jretconser.2020.102258

Twitter. (2023) "Update on Twitter's Limited Usage.". Twitter Business Blog. https://business.twitter.com/en/blog/update-on-twitters-limited-usage.html

Unal, S., & Aydın, H. (2013). An Investigation on the Evaluation of the Factors Affecting Brand Love. *Lumen International Conference Logos Universality Mentality Education Novelty*. https://doi.org/10.1016/j.sbspro.2013.08.640

Verhoef, P.C., Lemon, K.N., Parasuraman, A., Roggeveen, A., Tsiros, M. & Schlesinger, L.A. (2009). Customer experience creation: determinants, dynamics and management strategies. J. Retailing 85 (1), 31–41. https://doi.org/10.1016/j.jretai.2008.11.001

Warren, C., & Campbell, M. (2014). What makes things cool? How autonomy influences perceived coolness. *Journal of Consumer Research*, 41(2), 543–563

Warren, C., Batra, R., Loureiro, S. M. C., & Bagozzi, R. P. (2019). Brand coolness. *Journal of Marketing*, 83(5), 36-56. Retrieved from journals.sagepub.com/home/jmx. https://doi.org/10.1177/0022242919857698.

Williams, N.L., Ferdinand, N. & Bustard, J. (2019). From WOM to e-WOM – the evolution of unpaid influence: a perspective article. Tour. Rev. 75 (1), 314–318. https://doi.org/10.1108/TR-05-2019-0171

Yodpram, S., Intalar, N. (2020). Conceptualizing e-WOM, brand image, and brand attitude on consumer's willingness to pay in the low-cost airline industry in Thailand. *Proceedings* 39 (1), 1–4. https://doi.org/10.3390/proceedings2019039027

Zhang, P., Shi, X., Liu, W., Li, K., Zhao, L. & Zhou, J. (2023). The impact of extended product fit on brand loyalty: The road to durability and success for long-established enterprise. *Economic Analysis and Policy*, 77, 1055-1075. https://doi.org/10.1016/j.eap.2023.01.006

Appendix A

Table A.1: Exclusion dictionaries

Construct	Construct/Sub dimension	Keywords
Loyalty		"seem preferable", "preferred content","sony", "android", "sennheiser", "samsung", "galaxy", "windows", "google", "microsoft", "linux", "recommend a case", "recommend the case", "recommend charging", "recommend other shows", "recommended by apple", "cleaned the drive as recommend", "clean the drive as recommed", "it's recommendation", "it's recommended"
Brand coolness	Useful	"sony", "android", "sennheiser", "samsung", "galaxy", "windows", "google", "microsoft", "network"
	Usability	"sony", "android", "sennheiser", "samsung", "galaxy", "windows", "google", "microsoft", "network", "lux", "aux", "seamless design"
	Extraordinary	"sony", "android", "sennheiser", "samsung", "galaxy", "windows", "google", "microsoft"
	Energetic	"active noise", "sony", "android", "sennheiser", "samsung", "galaxy", "windows", "dynamic headphones", "enjoy", "google", "microsoft"
	Aesthetic	"graphic design", "pictures", "sony", "android", "sennheiser", "samsung", "galaxy", "windows", "lifestyle", "google", "microsoft"
	Original	"rarely", "sony", "android", "sennheiser", "samsung", "galaxy", "windows", "creative canvas", "creatives", "google", "microsoft"
	Authentic	"sony", "android", "sennheiser", "samsung", "galaxy", "windows", "google", "microsoft"
	Rebellious	"independently", "sony", "android", "sennheiser", "samsung", "galaxy", "windows", "google", "microsoft"
	High-status	"exclusively", "order status", "tracking status", "tracking my status", "tracking order status", "battery life status",

		"battery status", "record the status", "sony", "android", "sennheiser", "samsung", "galaxy", "windows", "wireless nature", "work in investment", "high end sound system", "investment in", "google", "microsoft"
	Popular	"cool down", "sony", "android", "sennheiser", "samsung", "galaxy", "windows", "popular belief", "google", "microsoft"
	Subcultural	N/A
	Iconic	"silicon", "tv icon", "app icon", "sony", "android", "sennheiser", "samsung", "galaxy", "windows", "icons", "the icon", "google", "microsoft"
Brand love		"sony", "android", "sennheiser", "samsung", "galaxy", "google", "windows"

Table A.2: Apple brand-related terms dictionary

Construct	Keywords
Apple brand-related terms	"apple", "mac", "ipad", "pad", "airpod", "product", "brand", "laptop", "pc", "headphones", "phones", "computer", "system", "app", "earplugs", "earbuds", "tablet", "gadget", "tech", "tv", "equipment", "watch", "fire stick", "interface", "wireless"

Appendix B

Sentimentr package documentation

get_sentences *Get Sentences*

Description

get_sentences - Get sentences from a character vector, sentiment, or sentiment_by object.

Usage

```
get_sentences(x, ...)
```

Arguments

x	A character vector, sentiment, or sentiment_by object.
...	Other arguments passed to split_sentence.

Value

Returns a list of vectors of sentences.

Examples

```
dat <- data.frame(
    w = c('Person 1', 'Person 2'),
    x = c(paste0(
        "Mr. Brown comes! He says hello. i give him coffee.  i will ",
        "go at 5 p. m. eastern time.  Or somewhere in between!go there"
    ), "One more thought for the road! I am going now.  Good day."),
    y = state.name[c(32, 38)],
    z = c(.456, .124),
    stringsAsFactors = FALSE
)
get_sentences(dat$x)
get_sentences(dat)
```

`sentiment`	*Polarity Score (Sentiment Analysis)*

Description

Approximate the sentiment (polarity) of text by sentence. This function allows the user to easily alter (add, change, replace) the default polarity an valence shifters dictionaries to suit the context dependent needs of a particular data set. See the `polarity_dt` and `valence_shifters_dt` arguments for more information. Other hyper-parameters may add additional fine tuned control of the algorithm that may boost performance in different contexts.

Usage

```
sentiment(
  text.var,
  polarity_dt = lexicon::hash_sentiment_jockers_rinker,
  valence_shifters_dt = lexicon::hash_valence_shifters,
  hyphen = "",
  amplifier.weight = 0.8,
  n.before = 5,
  n.after = 2,
  question.weight = 1,
  adversative.weight = 0.25,
  neutral.nonverb.like = FALSE,
  missing_value = 0,
  retention_regex = "\\d:\\d|\\d\\s|[^[:alpha:]',;: ]",
  ...
)
```

Arguments

`text.var` — The text variable. Can be a `get_sentences` object or a raw character vector though `get_sentences` is preferred as it avoids the repeated cost of doing sentence boundary disambiguation every time `sentiment` is run.

`polarity_dt` — A **data.table** of positive/negative words and weights with x and y as column names. The **lexicon** package has several dictionaries that can be used, including:

- `lexicon::hash_sentiment_jockers_rinker`
- `lexicon::hash_sentiment_jockers`
- `lexicon::emojis_sentiment`
- `lexicon::hash_sentiment_emojis`
- `lexicon::hash_sentiment_huliu`
- `lexicon::hash_sentiment_loughran_mcdonald`
- `lexicon::hash_sentiment_nrc`
- `lexicon::hash_sentiment_senticnet`
- `lexicon::hash_sentiment_sentiword`
- `lexicon::hash_sentiment_slangsd`
- `lexicon::hash_sentiment_socal_google`

Additionally, the `as_key` function can be used to make a sentiment frame suitable for `polarity_dt`. This takes a 2 column data.frame with the first column being words and the second column being polarity values. Note that as of version 1.0.0 **sentimentr** switched from the Liu & HU (2004) dictionary as the default to Jocker's (2017) dictionary from the **syuzhet** package. Use `lexicon::hash_sentiment_huliu` to obtain the old behavior.

`valence_shifters_dt` — A **data.table** of valence shifters that can alter a polarized word's meaning and an integer key for negators (1), amplifiers [intensifiers] (2), de-amplifiers [downtoners] (3) and adversative conjunctions (4) with x and y as column names.

`hyphen` — The character string to replace hyphens with. Default replaces with nothing so 'sugar-free' becomes 'sugarfree'. Setting hyphen = " " would result in a space between words (e.g., 'sugar free'). Typically use either " " or default "".

`amplifier.weight` — The weight to apply to amplifiers/de-amplifiers [intensifiers/downtoners] (values from 0 to 1). This value will multiply the polarized terms by 1 + this value.

`n.before` — The number of words to consider as valence shifters before the polarized word. To consider the entire beginning portion of a sentence use `n.before = Inf`.

`n.after` — The number of words to consider as valence shifters after the polarized word. To consider the entire ending portion of a sentence use `n.after = Inf`.

`question.weight` — The weighting of questions (values from 0 to 1). Default is 1. A 0 corresponds with the belief that questions (pure questions) are not polarized. A weight may be applied based on the evidence that the questions function with polarized sentiment. In an opinion tasks such as a course evalaution the questions are more likely polarized, not designed to gain information. On the other hand, in a setting with more natural dialogue, the question is less likely polarized and is likely to function as a means to gather information.

`adversative.weight`

The weight to give to adversative conjunctions or contrasting conjunctions (e.g., "but") that overrule the previous clause (Halliday & Hasan, 2013). Weighting a contrasting statement stems from the belief that the adversative conjunctions like "but", "however", and "although" amplify the current clause and/or down weight the prior clause. If an adversative conjunction is located before the polarized word in the context cluster the cluster is up-weighted 1 + number of occurrences of the adversative conjunctions before the polarized word times the weight given ($1 + N_{adversative\,conjunctions} * z_2$ where z_2 is the `adversative.weight`). Conversely, an adversative conjunction found after the polarized word in a context cluster down weights the cluster 1 - number of occurrences of the adversative conjunctions after the polarized word times the weight given ($1 + N_{adversative\,conjunctions} * -1 * z_2$). These are added to the deamplifier and amplifier weights and thus the down weight is constrained to -1 as the lower bound. Set to zero to remove adversative conjunction weighting.

`neutral.nonverb.like`

logical. If `TRUE`, and 'like' is found in the `polarity_dt`, when the word 'like' is preceded by one of the following linking verbs: `"'s"`, `"was"`, `"is"`, `"has"`, `"am"`, `"are"`, `"'re"`, `"had"`, or `"been"` it is neutralized as this non-verb form of like is not likely polarized. This is a poor man's part of speech tagger, maintaining the balance between speed and accuracy. The word 'like', as a verb, tends to be polarized and is usually preceded by a noun or pronoun, not one of the linking verbs above. This hyper parameter doesn't always yield improved results depending on the context of where the text data comes from. For example, it is likely to be more useful in literary works, where like is often used in non-verb form, than product comments. Use of this parameter will add compute time, this must be weighed against the need for accuracy and the likeliness that more accurate results will come from setting this argument to `TRUE`.

`missing_value` A value to replace `NA/NaN` with. Use `NULL` to retain missing values.

`retention_regex`

A regex of what characters to keep. All other characters will be removed. Note that when this is used all text is lower case format. Only adjust this parameter if you really understand how it is used. Note that swapping the `\\p{L}` for `[^[:alpha:];:,\']` may retain more alpha letters but will likely decrease speed. See examples below for how to test the need for `\\p{L}`.

`...` Ignored.

Details

The equation used by the algorithm to assign value to polarity of each sentence fist utilizes the sentiment dictionary to tag polarized words. Each paragraph ($p_i = \{s_1, s_2, ..., s_n\}$) composed of sentences, is broken into element sentences ($s_i, j = \{w_1, w_2, ..., w_n\}$) where w are the words within sentences. Each sentence (s_j) is broken into a an ordered bag of words. Punctuation is removed with the exception of pause punctuations (commas, colons, semicolons) which are considered a word within the sentence. I will denote pause words as cw (comma words) for convenience. We can represent these words as an i,j,k notation as $w_{i,j,k}$. For example $w_{3,2,5}$ would be the fifth word of the second sentence of the third paragraph. While I use the term paragraph this merely represent a complete turn of talk. For example t may be a cell level response in a questionnaire composed of sentences.

The words in each sentence ($w_{i,j,k}$) are searched and compared to a dictionary of polarized words (e.g., Jockers (2017) dictionary found in the **lexicon** package). Positive ($w_{i,j,k}^{+}$) and negative ($w_{i,j,k}^{-}$) words are tagged with a $+1$ and -1 respectively. I will denote polarized words as pw for convenience. These will form a polar cluster ($c_{i,j,l}$) which is a subset of the a sentence ($c_{i,j,l} \subseteq s_i, j$).

The polarized context cluster ($c_{i,j,l}$) of words is pulled from around the polarized word (pw) and defaults to 4 words before and two words after pw) to be considered as valence shifters. The cluster can be represented as ($c_{i,j,l} = \{pw_{i,j,k-nb}, ..., pw_{i,j,k}, ..., pw_{i,j,k-na}\}$), where nb & na are the parameters `n.before` and `n.after` set by the user. The words in this polarized context cluster are tagged as neutral ($w_{i,j,k}^{0}$), negator ($w_{i,j,k}^{n}$), amplifier [intensifier]] ($w_{i,j,k}^{a}$), or de-amplifier [downtoner] ($w_{i,j,k}^{d}$). Neutral words hold no value in the equation but do affect word count (n). Each polarized word is then weighted (w) based on the weights from the `polarity_dt` argument and then further weighted by the function and number of the valence shifters directly surrounding the positive or negative word (pw). Pause (cw) locations (punctuation that denotes a pause including commas, colons, and semicolons) are indexed and considered in calculating the upper and lower bounds in the polarized context cluster. This is because these marks indicate a change in thought and words prior are not necessarily connected with words after these punctuation marks. The lower bound of the polarized context cluster is constrained to $\max\{pw_{i,j,k-nb}, 1, \max\{cw_{i,j,k} < pw_{i,j,k}\}\}$ and the upper bound is constrained to $\min\{pw_{i,j,k+na}, w_{i,jn}, \min\{cw_{i,j,k} > pw_{i,j,k}\}\}$ where $w_{i,jn}$ is the number of words in the sentence.

The core value in the cluster, the polarized word is acted upon by valence shifters. Amplifiers (intensifiers) increase the polarity by 1.8 (.8 is the default weight (z)). Amplifiers ($w_{i,j,k}^{a}$) become de-amplifiers if the context cluster contains an odd number of negators ($w_{i,j,k}^{n}$). De-amplifiers (downtoners) work to decrease the polarity. Negation ($w_{i,j,k}^{n}$) acts on amplifiers/de-amplifiers as discussed but also flip the sign of the polarized word. Negation is determined by raising -1 to the power of the number of negators ($w_{i,j,k}^{n}$) + 2. Simply, this is a result of a belief that two negatives equal a positive, 3 negatives a negative and so on.

The adversative conjunctions (i.e., 'but', 'however', and 'although') also weight the context cluster. A adversative conjunction before the polarized word ($w_{adversative\,conjunction}, ..., w_{i,j,k}^{p}$) up-weights the cluster by $1 + z_2 * \{|w_{adversative\,conjunction}|, ..., w_{i,j,k}^{p}\}$ (.85 is the default weight (z_2)). An adversative conjunction after the polarized word down-weights the cluster by $1 + \{w_{i,j,k}^{p}, ..., |w_{adversative\,conjunction}| * -1\} * z_2$. The number of occurrences before and after the polarized word are multiplied by 1 and -1 respectively and then summed within context cluster. It is this value that is multiplied by the weight and added to 1. This corresponds to the belief that an adversative conjunction makes the next clause of greater values while lowering the value placed on the prior clause.

The researcher may provide a weight z to be utilized with amplifiers/de-amplifiers (default is .8; de-amplifier weight is constrained to -1 lower bound). Last, these weighted context clusters ($c_{i,j,l}$) are summed ($c'_{i,j}$) and divided by the square root of the word count ($\sqrt{w_{i,jn}}$) yielding an **unbounded polarity score** (δ) for each sentence.

$$\delta = \frac{c'_{i,j}}{\sqrt{w_{i,jn}}}$$

Where:

$$c'_{i,j} = \sum \left((1 + w_{amp} + w_{deamp}) \cdot w_{i,j,k}^{p}(-1)^{2+w_{neg}}\right)$$

$$w_{amp} = (w_b > 1) + \sum (w_{neg} \cdot (z \cdot w^a_{i,j,k}))$$

$$w_{deamp} = \max(w_{deamp'}, -1)$$

$$w_{deamp'} = (w_b < 1) + \sum (z(-w_{neg} \cdot w^a_{i,j,k} + w^d_{i,j,k}))$$

$$w_b = 1 + z_2 * w_{b'}$$

$$w_{b'} = \sum (|w_{adversative\,conjunction}|, ..., w^p_{i,j,k}, w^p_{i,j,k}, ..., |w_{adversative\,conjunction}| * -1)$$

$$w_{neg} = \left(\sum w^n_{i,j,k}\right) \bmod 2$$

Value

Returns a **data.table** of:

- element_id - The id number of the original vector passed to `sentiment`
- sentence_id - The id number of the sentences within each `element_id`
- word_count - Word count
- sentiment - Sentiment/polarity score (note: sentiments less than zero is negative, 0 is neutral, and greater than zero positive polarity)

Note

The polarity score is dependent upon the polarity dictionary used. This function defaults to a combined and augmented version of Jocker's (2017) [originally exported by the **syuzhet** package] & Rinker's augmented Hu & Liu (2004) dictionaries in the **lexicon** package, however, this may not be appropriate, for example, in the context of children in a classroom. The user may (is encouraged) to provide/augment the dictionary (see the `as_key` function). For instance the word "sick" in a high school setting may mean that something is good, whereas "sick" used by a typical adult indicates something is not right or negative connotation (**deixis**).

References

Jockers, M. L. (2017). Syuzhet: Extract sentiment and plot arcs from text. Retrieved from https://github.com/mjockers/syuzh

Hu, M., & Liu, B. (2004). Mining opinion features in customer reviews. National Conference on Artificial Intelligence.

Halliday, M. A. K. & Hasan, R. (2013). Cohesion in English. New York, NY: Routledge.

https://www.slideshare.net/jeffreybreen/r-by-example-mining-twitter-for

http://hedonometer.org/papers.html Links to papers on hedonometrics

See Also

Original URL: https://github.com/trestletech/Sermon-Sentiment-Analysis

Other sentiment functions: `sentiment_by()`

Examples

```
mytext <- c(
   'do you like it?  But I hate really bad dogs',
   'I am the best friend.',
   "Do you really like it?  I'm not a fan",
   "It's like a tree."
)

## works on a character vector but not the preferred method avoiding the
## repeated cost of doing sentence boundary disambiguation every time
## `sentiment` is run.  For small batches the loss is minimal.
## Not run:
sentiment(mytext)

## End(Not run)

## preferred method avoiding paying the cost
mytext <- get_sentences(mytext)
sentiment(mytext)
sentiment(mytext, question.weight = 0)

sam_dat <- get_sentences(gsub("Sam-I-am", "Sam I am", sam_i_am))
(sam <- sentiment(sam_dat))
plot(sam)
plot(sam, scale_range = TRUE, low_pass_size = 5)
plot(sam, scale_range = TRUE, low_pass_size = 10)

## Not run: ## legacy transform functions from suuzhet
plot(sam, transformation.function = syuzhet::get_transformed_values)
plot(sam, transformation.function = syuzhet::get_transformed_values,
    scale_range = TRUE, low_pass_size = 5)

## End(Not run)

y <- get_sentences(
    "He was not the sort of man that one would describe as especially handsome."
)
sentiment(y)
sentiment(y, n.before=Inf)

## Not run: ## Categorize the polarity (tidyverse vs. data.table):
library(dplyr)
sentiment(mytext) %>%
as_tibble() %>%
    mutate(category = case_when(
        sentiment < 0 ~ 'Negative',
        sentiment == 0 ~ 'Neutral',
```

```
        sentiment > 0 ~ 'Positive'
    ) %>%
    factor(levels = c('Negative', 'Neutral', 'Positive'))
)

library(data.table)
dt <- sentiment(mytext)[, category := factor(fcase(
        sentiment < 0, 'Negative',
        sentiment == 0, 'Neutral',
        sentiment > 0, 'Positive'
    ), levels = c('Negative', 'Neutral', 'Positive'))][]
dt

## End(Not run)

dat <- data.frame(
    w = c('Person 1', 'Person 2'),
    x = c(paste0(
        "Mr. Brown is nasty! He says hello. i give him rage.  i will ",
        "go at 5 p. m. eastern time.  Angry thought in between!go there"
    ), "One more thought for the road! I am going now.  Good day and good riddance."),
    y = state.name[c(32, 38)],
    z = c(.456, .124),
    stringsAsFactors = FALSE
)
sentiment(get_sentences(dat$x))
sentiment(get_sentences(dat))

## Not run:
## tidy approach
library(dplyr)
library(magrittr)

hu_liu_cannon_reviews %>%
   mutate(review_split = get_sentences(text)) %$%
   sentiment(review_split)

## End(Not run)

## Emojis
## Not run:
## Load R twitter data
x <- read.delim(system.file("docs/r_tweets.txt", package = "textclean"),
    stringsAsFactors = FALSE)

x

library(dplyr); library(magrittr)

## There are 2 approaches
## Approach 1: Replace with words
x %>%
    mutate(Tweet = replace_emoji(Tweet)) %$%
```

```
    sentiment(Tweet)

## Approach 2: Replace with identifier token
combined_emoji <- update_polarity_table(
    lexicon::hash_sentiment_jockers_rinker,
    x = lexicon::hash_sentiment_emojis
)

x %>%
    mutate(Tweet = replace_emoji_identifier(Tweet)) %$%
    sentiment(Tweet, polarity_dt = combined_emoji)

## Use With Non-ASCII
## Warning: sentimentr has not been tested with languages other than English.
## The example below is how one might use sentimentr if you believe the
## language you are working with are similar enough in grammar to for
## sentimentr to be viable (likely Germanic languages)
## english_sents <- c(
##     "I hate bad people.",
##     "I like yummy cookie.",
##     "I don't love you anymore; sorry."
## )

## Roughly equivalent to the above English
danish_sents <- stringi::stri_unescape_unicode(c(
    "Jeg hader d\\u00e5rlige mennesker.",
    "Jeg kan godt lide l\\u00e6kker is.",
    "Jeg elsker dig ikke mere; undskyld."
))

danish_sents

## Polarity terms
polterms <- stringi::stri_unescape_unicode(
    c('hader', 'd\\u00e5rlige', 'undskyld', 'l\\u00e6kker', 'kan godt', 'elsker')
)

## Make polarity_dt
danish_polarity <- as_key(data.frame(
    x = stringi::stri_unescape_unicode(polterms),
    y = c(-1, -1, -1, 1, 1, 1)
))

## Make valence_shifters_dt
danish_valence_shifters <- as_key(
    data.frame(x='ikke', y="1"),
    sentiment = FALSE,
    comparison = NULL
)

sentiment(
    danish_sents,
    polarity_dt = danish_polarity,
    valence_shifters_dt = danish_valence_shifters,
    retention_regex = "\\d:\\d|\\d\\s|[^\\p{L}',;: ]"
)

## A way to test if you need [:alpha:] vs \p{L} in `retention_regex`:

## 1. Does it wreck some of the non-ascii characters by default?
sentimentr:::make_sentence_df2(danish_sents)

## 2.Does this?
sentimentr:::make_sentence_df2(danish_sents, "\\d:\\d|\\d\\s|[^\\p{L}',;: ]")

## If you answer yes to #1 but no to #2 you likely want \p{L}

## End(Not run)
```