



BUSINESS
SCHOOL

The brand equity of a destination: a user-generated content analysis

Ana Filipa Monteiro Rodrigues Cardoso

Master in Business Analytics

Supervisor:

Professor Sérgio Moro, PhD, Associate Professor with Habilitation at Iscte-iul

Co-Supervisor:

Professor Catarina Marques, PhD, Assistant Professor at Iscte-iul

August, 2022

Aos meus pais e ao meu irmão.

Acknowledgements

This dissertation would not be possible without all the support and excellence of my supervisors to whom I express my gratitude, Professor Sérgio Moro and Professor Catarina Marques. My gratitude also extends to Iscte and all professors that contributed to enhance my knowledge, with a special thanks to Professor Raul Laureano, director of this Masters.

Special thanks to my parents for the education given to me that makes me embrace the most difficult challenges and obstacles. Also, thanks to my brother for all the understanding and guidance to always do my best.

I would also like to thank my friends and boyfriend who have been present during this dissertation, for the support and positiveness that made me achieve my objectives.

Abstract

Customer-based brand equity can be defined as the differential effect of brand knowledge on consumer response to the marketing of the brand. Although there are already studies focused on the brand equity of destinations, most of them are based on surveys. Still, it is possible to extract and analyse visitors' opinion by their self-reported experience that is available on travel blogs, in which bloggers write in a way that best represents their experience. Tourism in Lisbon has been increasing over the last years and Portugal is one of the most visited countries in Europe.

The novelty of this dissertation is to propose a new methodology of measuring brand equity through text-mining of user-generated blog posts, based on the visitors' evaluations of Lisbon. For this purpose, 100 posts were collected to test a brand equity model measured by four constructs: brand awareness, brand image - decomposed in cognitive, affective, and unique image components, perceived quality, and brand loyalty.

Findings revealed that the components of destination image are highly related, with the cognitive image being strongly related to higher destination brand image. Regarding to the constructs, awareness and image are the most important to make the consumers loyal to Lisbon, with image and perceived quality representing the strongest relationship between constructs. Further, higher awareness can lead to negative perceived quality, but the image can work as a moderator of this relationship. Perceived quality was the only construct that did not reveal a significant relationship with destination loyalty.

Keywords: brand equity; text mining; brand image; Lisbon; sentiment analysis.

(JEL Classification System: L83, M39)

Resumo

O valor da marca baseado no cliente pode ser definido como o efeito diferencial do conhecimento da marca na resposta do consumidor ao marketing da marca. Embora existam estudos focados no valor da marca dos destinos, a maioria são baseados em questionários. Contudo, é possível extrair e analisar a opinião dos visitantes através do seu relato da experiência, disponível em *blogs* de viagens, nos quais os *bloggers* escrevem da forma que melhor representa a sua experiência. O turismo em Lisboa tem aumentado nos últimos anos e Portugal é um dos países mais visitados da Europa.

A novidade desta dissertação é propor uma nova metodologia para medir o valor da marca através da mineração de dados de publicações em *blogs*, baseada na avaliação dos visitantes de Lisboa. Para isso, 100 publicações foram recolhidas para testar um modelo de valor da marca medido por quatro construtos: notoriedade, imagem - decomposta em componentes cognitivos, afetivos e imagem única, qualidade percebida e lealdade.

Os resultados revelaram que as componentes da imagem do destino estão altamente relacionadas, com a imagem cognitiva fortemente relacionada à imagem da marca. Relativamente aos construtos, a notoriedade e a imagem são os mais importantes para fidelizar os consumidores a Lisboa, com a imagem e a qualidade percebida representando a relação mais forte entre eles. Além disso, maior notoriedade pode levar a qualidade percebida negativa, mas a imagem pode funcionar como moderadora desta relação. A qualidade percebida foi o único construto que não revelou relação significativa com a lealdade ao destino.

Palavras-chave: valor da marca; mineração de dados; imagem da marca; Lisboa; análise de sentimentos.

(JEL Classification System: L83, M39)

Table of Contents

Acknowledgements	i
Abstract	iii
Resumo	v
Table of Contents	vii
List of Figures	ix
List of Tables	xi
List of Acronyms and Abbreviations	xiii
1. Introduction	1
2. Literature Review	3
2.1. Systematic literature review protocol.....	3
2.2. Analysis of the literature.....	8
2.2.1. Destination as a brand.....	8
2.2.2. Customer-based brand equity.....	8
2.2.3. CBBE constructs and hypothesis development.....	9
2.2.4. CBBE models and constructs testing.....	14
2.2.5. Research methods.....	18
2.2.6. Limitations and existing gaps.....	19
2.2.7. Evaluation of the relevance of the articles.....	21
3. Methodology	23
3.1. Problem awareness.....	24
3.2. Suggestion.....	24
3.2.1. User-generated content.....	24
3.2.2. Study destination: Lisbon, Portugal.....	25
3.2.3. Text mining and sentiment analysis.....	26
3.3. Development.....	26
4. Results	33
4.1. Sample characterisation.....	33
4.2. Constructs and sentiment analysis.....	34
4.3. Structural equation modelling and hypothesis testing.....	37
4.4. Discussion.....	40
5. Conclusions	43
References	47
Appendix I	53
Appendix II	55

List of Figures

Figure 2.1: Flowchart of the PRISMA protocol	5
Figure 2.2: Word cloud of keywords of selected articles	7
Figure 2.3: Proposed conceptual model	18
Figure 3.1: General steps of DSR methodology.....	23
Figure 3.2: Number of overnight stays in Portugal	25
Figure 3.3: Sentiment analysis approach	27
Figure 4.1: Sample characterisation.....	33
Figure 4.2: Distribution of components of Lisbon brand equity	34
Figure 4.3: Total average sentiment per post	35
Figure 4.4: Average score per post.....	36
Figure 4.5: Path diagram for the destination brand equity model	38

List of Tables

Table 2.1: Research questions and objectives of the SLR	4
Table 2.2: Inclusion and exclusion criteria	5
Table 2.3: General information related to the studies included in the SLR	6
Table 2.4: Quality criteria for the evaluation of the articles	7
Table 2.5: Destination brand equity constructs.	9
Table 2.6: Summary of research specifications of the studies included in the SLR	15
Table 2.7: Research methodology.....	19
Table 2.8: Contributions and limitations.....	20
Table 2.9: Quality evaluation of the articles of SLR.....	21
Table 3.1: Data dictionary	30
Table 4.1: Descriptive statistics	35
Table 4.2: Correlation values.....	37
Table 4.3: Coefficient estimates of the model structural relationships.....	38
Table 4.4: Standardized coefficients for indirect and total effects	39
Table I.1: Keywords of the selected articles of SLR.....	54

List of Acronyms and Abbreviations

AI	Affective Image
ADL	Attitudinal Destination Loyalty
CBA	Cultural Brand Assets
CBBE	Customer-Based Brand Equity
CFI	Comparative Fit Index
CI	Cognitive Image
DBA	Destination Brand Awareness
DBE	Destination Brand Equity
DBI	Destination Brand Image
DBL	Destination Brand Loyalty
DBQ	Destination Brand Perceived Quality
DBS	Destination Brand Saliency
DI	Destination Image
DQP	Destination Quality Perception
DSR	Design Science Research
GFI	Goodness-of-Fit Index
IBA	Internal Brand Awareness
IBC	Internal Brand Commitment
IBI	Internal Brand Image
IBL	Internal Brand Loyalty
IFI	Incremental Fit Index
OBE	Overall Brand Equity
PQ	Perceived Quality
PRISMA	Preferred Reporting Items for Systematic review and Meta-Analysis
PV	Perceived Value
RMSEA	Root Mean Square Error of Approximation
RQ	Research Question
Sat	Satisfaction
SLR	Systematic Literature Review
SRMR	Standardized Root Mean Squared Residual
SSI	Social Self-Image
TI	Travel Intentions
UNWTO	United Nations World Tourism Organization
WoM	Word of Mouth

CHAPTER 1

Introduction

Tourism in Lisbon has been increasing through the last years and Portugal is one of the most visited countries in Europe (WorldData, 2022). Developing marketing activities that can increase the value of assets from tourist destinations can help achieve the goal of winning competition, increasing profits, and visitor loyalty (Kim et al., 2017). Developing a brand equity model for a destination offers marketers a performance instrument to evaluate and measure consumer perceptions of a destination brand (Pike & Bianchi, 2016). Further, measuring the effectiveness of destination branding upon visitors' perception is an important tool to evaluate intent to visit a place, thus, contributing for the improvement of the brand equity of its destination (Chi et al., 2020). The aim of this dissertation is to propose and test a brand equity model for the city of Lisbon.

Adaptations of Keller's (1993) customer-based brand equity concept in a tourism context have attempted to test the applicability of the customer-based brand equity model by using a four-dimensional structure incorporating brand awareness, perceived quality, brand image, and brand loyalty (Im et al., 2012; Kim et al., 2017). Most of the studies measuring brand equity of destinations have used empirical studies based on a cross-sectional survey, i.e., a single sample obtained at a single moment in time, thus, being a limitation of the existing approaches (e.g., Chi et al. (2020); Kim & Lee (2018); Kladou & Kehagias (2014) and Bianchi & Pike (2011)). The use of different methods of research through online recommendations or content available online to measure the constructs of destination brand equity can represent an improved contribution to existing approaches, since it allows to capture the real visitors experience from a broad perspective.

In order to measure brand equity of a destination through text mining and sentiment analysis of user-generated content, the following research question is addressed: How to measure brand equity of a destination through text mining of user-generated content?

The novelty of this dissertation is centred in a different method of research being proposed, specifically in the data collection method, in order to gather a more varied sample. Online recommendations are one of the most important factors in consumers' decision-making processes since consumers consider opinions of their peers more reliable than those posted by the service provider (Guerreiro & Rita, 2020). Therefore, this study proposes measuring the destination brand equity constructs using the user-generated content in blog posts through text mining and sentiment analysis. In addition, this dissertation contributes to the tourism destination branding literature by testing a conceptual model of destination brand equity, while attempting to understand the important branding constructs, models and concepts of destination brand equity through a systematic literature

review. Further, this dissertation contributes to the branding literature by defining the determinants of destination brand loyalty through structural equation modelling.

For such purpose, a review of existing literature through a systematic literature review is made in Chapter 2, starting by offering a comprehensive picture of destination as a brand, as well as the customer-based brand equity definition. The articles, selected through a protocol of systematic reviews, are evaluated according to their relevance. The developed hypotheses result in a conceptual model that is tested through sentiment analysis of a 100 blog posts sample, composed of blogs from Lisbon visitors. The Design Science Research methodology is followed, in which the methods to proceed the analysis will be detailed in Chapter 3. The obtained results are presented and discussed in Chapter 4 by using statistical descriptive analyses and structural equation modelling. Limitations of this study, managerial and theoretical implications and suggestions for future research are presented in Chapter 5.

CHAPTER 2

Literature Review

The aim of this dissertation is to study the brand equity of a specific destination so a full understanding of brand equity as a concept, and how it was measured through the years, is needed, hence, a Systematic Literature Review (SLR) is performed. It is a summary of the literature that uses explicit and reproducible methods to systematically search, critically appraise, and synthesize on a specific issue (Gopalakrishnan & Ganeshkumar, 2013). Due to the explicit methods used, SLR presents advantages on drawing reliable and accurate conclusions as well as easily providing the necessary information to researchers (Sánchez-Rebull et al., 2018; Gopalakrishnan & Ganeshkumar, 2013).

The complex characteristics of a destination make branding and measuring the brand equity a challenge (Boo et al., 2009). For that reason, the specific elements that compose a destination brand and the measurement methods being used are still under analysis (Ferns & Walls, 2012). The SLR allows to find the best articles for defining the required concepts and the appropriate dimensions for measuring the brand equity of a destination.

In the first part of this chapter, the SLR protocol is described and used to select the best articles. Based on the selected articles, the required concepts and the appropriate dimensions for measuring the brand equity of a destination are defined. Afterwards, it is mandatory to understand how the different dimensions are related, based on the existing models for measuring brand equity of destinations. Also, it is of relevance to analyse the obtained results to understand if they were well succeeded or not. Finally, evaluation criteria are defined, in order to understand how the selected articles are useful within the scope of this dissertation.

By identifying the concepts, relationships and measurement methods of brand equity, it is possible to define the most appropriate dimensions and model to measure the brand equity of the destination being addressed in this dissertation.

2.1. Systematic literature review protocol

This SLR aim to help answering the general research question of this dissertation: How to measure brand equity of a destination through text mining of user-generated content? In 2010, Pike et al. (2010) identifies the concept “brand equity” as a potential research gap in the context of tourism destinations marketing. Further, understanding how to measure brand equity of a destination contributes to build successful destination brands and manage the factors that determine brand equity for tourism destinations (Gómez et al., 2015). Due to the complexity of the nature of a destination brand, the following specific research questions (RQ) were formulated, Table 2.1.

Table 2.1: Research questions and objectives of the SLR

ID	Research question	Objective
RQ1	What is brand equity and customer-based brand equity and why is it useful to destinations?	To identify the contributions of brand equity to the improvement of a destination brand.
RQ2	What are the dimensions used to measure brand equity of a destination?	To identify the variety of constructs used to measure brand equity.
RQ3	How different dimensions are related and which models were developed?	To identify destination brand equity models used in other studies.
RQ4	Which research methods were used to measure brand equity (surveys, text mining of online reviews, ...)	To identify the methods commonly used and analyse a potential gap of research.

To answer each research question, the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) protocol was followed. It consists of a strategy of systematic reviews composed of four phases: identification, screening, eligibility and inclusion (Moher et al., 2009). PRISMA is aimed to help authors to improve the reporting of systematic reviews (Shamseer et al., 2015) in order to gather the studies included in this SLR. This review's PRISMA protocol is illustrated in Figure 2.1, where, for each phase, the number of remaining articles, n , is presented, e.g., from the identification phase, $n = 79$ articles remain for the next phases.

In the identification stage, Web of Science and Scopus databases were searched using a query applied to the title, abstract, and keywords: TITLE-ABS-KEY (("brand equity of a destination" OR "destination brand equity") AND ("model" OR "dimension*" OR "measure*")). These two databases were chosen since they are the most important multidisciplinary bibliographic data sources, providing metadata on scientific documents and on citation links between these documents (Visser et al., 2021).

The search was conducted between November 2021 and February 2022, in which 64 articles from Web of Science and 52 articles from Scopus were identified for further stages. Inclusion and exclusion criteria were applied in all stages and are presented in Table 2.2. The method to create the query was based on the inclusion criteria for the identification stage and, for next stages, inclusion and exclusion criteria were applied to the resulting 116 articles from the identification stage.

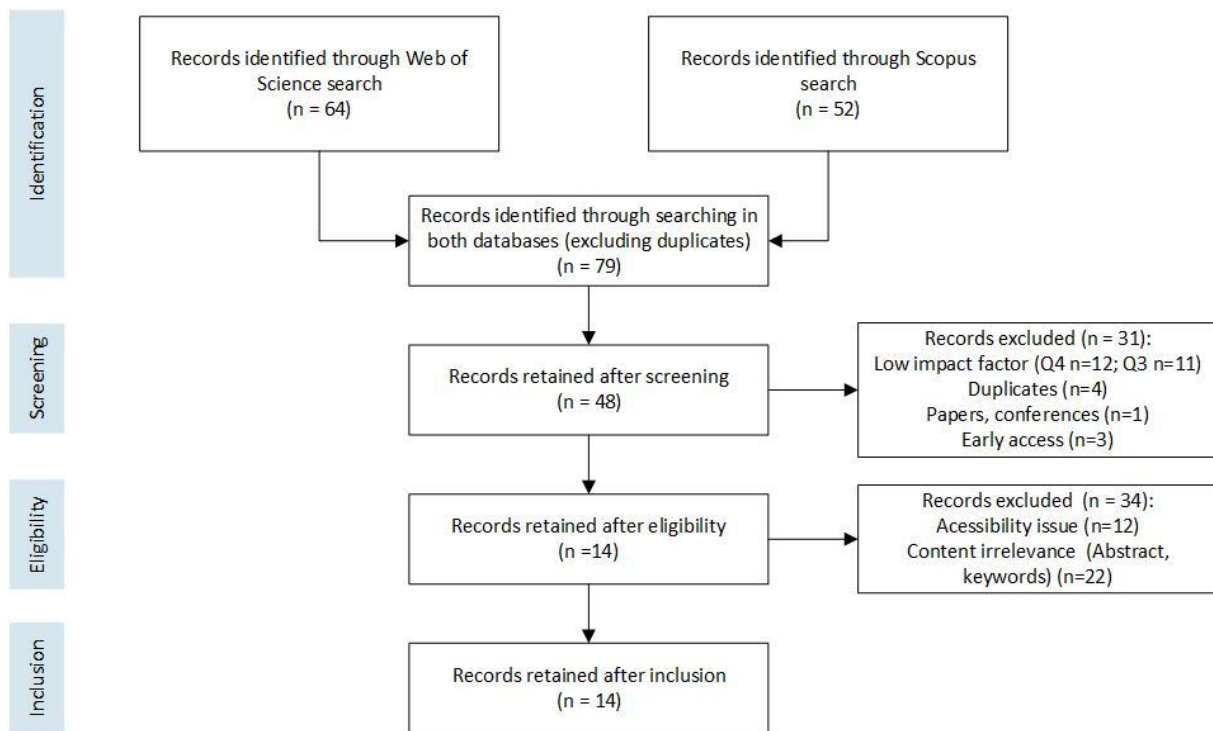


Figure 2.1: Flowchart of the PRISMA protocol

Table 2.2: Inclusion and exclusion criteria

Category	Stage	Criteria
Inclusion	Identification	Articles measuring brand equity of destinations.
		Articles developing models of destination brand equity.
		Articles studying the dimensions to analyse destination brand equity.
Inclusion	RQs were checked to see if all the themes of the literature had been addressed.	
Exclusion	Screening	Articles published in journals with low factor impact (Q3 or Q4).
		Duplicated articles.
		Articles not published in peer-reviewed journals (e.g., conference paper, working paper, or book chapter).
		Early access articles.
		Articles published before 2010.
	Eligibility	Articles with accessibility issue (unavailable to download).
Articles not relevant after reading the abstract and keywords.		

At the screening stage, 31 articles were excluded according to the exclusion criteria. At the eligibility stage, the articles were read by title, abstract and keywords to assess their relevance to the topic of this review. As a result, 12 articles were excluded since they were unavailable to download and 22 articles were identified for content irrelevance, and 14 articles were identified for the inclusion stage, in which no article was identified to reinforce this review. Thus, 14 articles were included, as detailed in Table 2.3.

Table 2.3: General information related to the studies included in the SLR

ID	Name	Author, Year	Journal, Vol(Number), Pages
1	Destination Brand Equity: A Perspective of Generation Z on A World Heritage Site in Indonesia	Kusumaningrum, S.D., 2021	Journal of Asian Finance, Economics and Business, 8(2), 1071–1078
2	Elements of destination brand equity and destination familiarity regarding travel intention	Chi, H.K.; Huang, K.C.; Nguyen, H.M., 2020	Journal of Retailing and Consumer Services, 52, Article 101728
3	A structural analysis of destination brand equity in mountainous tourism destination in northern India	Kaushal, V.; Sharma, S.; Reddy, G.M., 2019	Tourism and Hospitality Research, 19(4), 452–464
4	An integrative model of destination brand equity and tourist satisfaction	San Martin, H.; Herrero, A.; de los Salmones, M.D., 2019	Current Issues in Tourism, 22(16), 1992–2013
5	Brand Equity of a Tourist Destination	Kim, H.K.; Lee, T.J., 2018	Sustainability, 10, 431
6	Destination Brand Equity for Australia: Testing a Model of CBBE in Short-Haul and Long-Haul Markets	Pike, S.; Bianchi, C., 2016	Journal of Hospitality & Tourism Research, 40(1), 114–134
7	A model of tourism destination brand equity: The case of wine tourism destinations in Spain	Gómez, M.; Lopez, C.; Molina, A., 2015	Tourism Management, 51, 210-222
8	Assessing destination brand equity: An integrated approach	Kladou, S.; Kehagias, J., 2014	Journal of Destination Marketing & Management, 3, 2–10
9	Enduring travel involvement, destination brand equity, and travelers' visit intentions: A structural model analysis	Ferns, B.H.; Walls, A., 2012	Journal of Destination Marketing & Management 1, 27–35
10	Tourist destination brand equity and internal stakeholders: An empirical research	Sartori A., Mottironi C., Corigliano M.A., 2012	Journal of Vacation Marketing, 18(4), 327–340
11	Conceptualizing Destination Brand Equity Dimensions from a Consumer-Based Brand Equity Perspective	Im H.H., Kim S.S., Elliot S., Han H., 2012	Journal of Travel & Tourism Marketing, 29(4), 385-403
12	Tourism Destination Brand Equity Dimensions: Renewal versus Repeat Market	Gartner, W.C.; Konecnik Ruzzier, M., 2011	Journal of Travel Research, 50(5), 471–481
13	Antecedents of Destination Brand Loyalty for a Long-Haul Market: Australia's Destination Loyalty Among Chilean Travelers	Bianchi C., Pike S., 2011	Journal of Travel & Tourism Marketing, 28(7), 736-750
14	Exploring relationships between Mongolian destination brand equity, satisfaction and destination loyalty	Chen C.F., Myagmarsuren O., 2010	Tourism Economics, 2010, 16(4), 981–994

Overall, 14 articles were identified for content analysis. To enhance the content, the keywords referring to the articles are shown in a word cloud in which stop words, numbers and special characters were removed for legibility, Figure 2.2. The table of frequencies that originated this word cloud is presented in Appendix I.



Figure 2.2: Word cloud of keywords of selected articles

The word cloud shows, as expected, that the most mentioned keywords are “destination”, “brand” and “equity”, once the query used to search the articles contained these words. It is also possible to observe which destinations were studied in these articles, such as Rome, Zealand and Australia, as well as references to statistical analyses such as structural equation modelling and regression analysis.

After the content analysis, the relevance of the selected articles is evaluated based on the quality criteria, Table 2.4. All the questions aim to infer if the selected articles were useful to answer the presented research questions. The results of this evaluation are detailed in Section 2.2.7.

Table 2.4: Quality criteria for the evaluation of the articles

Topic	Evaluation question
Destination brand equity concepts	Q1 - Is the evolution of concepts used to measure brand equity of a destination properly described?
	Q2 - Are new concepts introduced into the destination brand equity study?
Models of Destination brand equity measurement	Q3 - Are new models of destination brand equity tested?
	Q4 - Is the evaluation criteria of the model well defined?
	Q5 - Is the model well succeeded verifying all hypotheses?
Method of research	Q6 - Is the method of research well presented and appropriated?
	Q7 - Are different methods of research used?
	Q8 - Are the methods of research applied to different population samples?
Conclusions	Q9 - Are the contributions of the study described?
	Q10 - Are limitations of the study identified?

2.2. Analysis of the literature

2.2.1. Destination as a brand

According to Chi et al. (2020), destinations are seen as intangible products, thus, subjective, and depending upon the route of travel, culture, purpose of the visit, educational level and past experience of visitors. Branding a destination is defined as the process used to develop a unique identity and personality that is different from all competitive destinations (Kladou & Kehagias, 2014). As the choice of a destination is a multidimensional problem and can provide different experiences for tourists, destination branding is important as it plays a major role in facilitating tourists' pre-trip planning from its competitors (Chi et al., 2020). Further, the choice is determined by the brand value of each potential tourism destination, and this is a result of the consideration of the merits and attractions of the various options (Kim & Lee, 2018). Hence, destinations cannot be traded like usual products or services, nor is their ownership transferrable, so other destinations cannot directly determine their destination brand equity (Ferns & Walls, 2012).

Through the years, corresponding to the expectations of the consumers becomes a major concern for those trying to develop and support destination brands (Gartner & Ruzzier, 2011). From the tourism perspective, destination branding might support visitors in enhancing their awareness of a destination after their tourism experiences (Kladou & Kehagias, 2014). As well as conveying a promise about potential experiences in that place for tourists, the destination brand must contribute to enhancing and consolidating positive memories associated with a place (Chi et al., 2020). The aim of destination branding should be to stimulate the intention to visit and revisit, which are indicators of brand loyalty (Pike & Bianchi, 2016).

2.2.2. Customer-based brand equity

In a highly competitive tourism market, marketing managers are always seeking for solutions to strengthen the value of their destinations. They usually do this by concentrating on vital factors which can enhance their brand equity (Chi et al., 2020). Developing marketing activities that can increase the value of assets from tourist destinations can help achieve the goal of winning competition, increasing profits, and increasing visitor loyalty (Kim et al., 2017). Because of that, enhancing brand equity of a destination has been recognized as an important factor for creating competitive market advantages and marketing strategies of differentiation (Horng et al., 2012). Hence, destination marketers have realized the growing importance of brand equity in promoting their destination (Chi et al., 2020).

Brand equity was defined by Aaker (1991, p. 27) as "a set of brand assets and liabilities linked to a brand, its name, and symbol, which add to or subtract from the value provided by a producer, by a product or service to a firm and/or to that firm's customer". Keller (1993, p. 8) introduced the concept

of brand equity based on consumers, conceptualizing Customer-Based Brand Equity (CBBE) as “the differential effect of brand knowledge on consumer response to the marketing of the brand”. This concept offers destination marketers a performance instrument with which to evaluate and measure consumer perceptions of a destination brand (Pike & Bianchi, 2016). CBBE has been recognized for its applicability in the service industry (Kim & Lee, 2018) and, in the hospitality and tourism literature, it has been used to assess the brand equity of several tourism destinations (eg. Boo et al., 2009; Im et al., 2012; Kim et al., 2017; Kladou & Kehagias, 2014).

2.2.3. CBBE constructs and hypothesis development

Due to the multidimensional nature of CBBE, Keller (1993) introduced the use of multiple measures for the purpose of assessing brand equity. Gartner & Ruzzier (2011) suggested that the right way to analyse brand equity of tourism destination is to consider individual dimensions that cumulatively strengthen or weaken destination brand equity. Although there are disparities in the various conceptualizations of brand equity and its composition, most adaptations of Keller’s (1993) CBBE concept in a tourism context have attempted to test the applicability of the CBBE model by using a four-dimensional structure incorporating brand awareness, perceived quality, brand image and brand loyalty (Im et al., 2012; Kim et al., 2017). These four constructs are summarised in Table 2.5.

Table 2.5: Destination brand equity constructs.

Construct	Description	Reference
Brand awareness	It refers to the ability to recognize or recall that a brand is a member of a product category. In the context of tourist destinations, it can be conceived as the presence of a destination in the minds of people when a given travel context is considered.	Aaker (1996) (Herrero et al., 2017)
Brand image	It represents the set of associations attached to the destination, composed of a variety of individual perceptions relating to various attributes of the destination that may or may not reflect objective reality.	Aaker (1996), Keller (1993)
Perceived quality	It is the overall perception of customers about brilliance and quality of products or services in comparison with the rivalry offering. When the intangible attributes acquire high predominance, as is the case in tourism, the quality assessment depends almost exclusively on perceived quality.	Aaker (1991), Dias & Cardoso (2017)
Loyalty	It represents the core dimension of the CBBE concept and the main source of customer-based brand equity. In tourism, loyalty is usually considered as the intention to revisit the destination and the willingness to recommend it to other people (Word of Mouth intentions).	Chen & Myagmarsuren (2010), Pike et al. (2010)

It is highly relevant the fact that all brand equity constructs are closely related among them (Dias & Cardoso, 2017). Konecnik & Gartner (2007) have identified a number of relationships between these constructs and stated that the image is a central concept in destination branding. Kim & Kim (2005) underline that CBBE falls into two groups: consumer perception equity (brand awareness, brand image and perceived quality) and consumer behaviour equity (brand loyalty), that can be treated as one of the consequences of perceptual equity.

Destination brand awareness is a key element in creating the value of the destination as a favourable destination brand in the minds of travellers (Im et al., 2012). A place must be known to the consumer, before it can even be considered as a potential destination (Gartner & Ruzzier, 2011). Hence, Kaushal et al. (2019) advised that destination managers duly consider brand awareness as a vital element in the development of brand equity of distinct geographical regions. Konecnik and Gartner (2007, p.403) referred awareness as “what someone knows and or thinks they know about a destination” being linked with the destination selection process. According to Kim & Lee (2018), brand awareness is a combination of whether a specific brand is stored in the memory of consumers and their ability to recall a specific brand within a product category. For that matter brand equity can only occur if the consumer is familiar with and holds some favourable associations in memory (Ferns & Walls, 2012).

Once, brand awareness influences the formation of associations, it is to be expected that a greater awareness of a destination will enhance the perception of its brand image (San Martín et al., 2019). Hence, Keller (1993) argued that brand awareness can also be viewed in conjunction with brand image, where both the elements of brand equity are together viewed as the overall brand knowledge. All in all, awareness of a tourism destination is key to encourage tourists to visit the destination and thus increase their familiarity with and fondness for the brand (Gómez et al., 2015).

Destination brand image has been widely defined as impressions or perceptions of a place (Marques et al., 2021). Image refers to the attributes one expects a destination to possess (Gartner & Ruzzier, 2011), hence, favourable destination image and strong differentiation from competitors influence tourists’ choice behaviour (Kaushal et al., 2019). Besides, the more unique and favourable images the consumer holds in his/her memory, the stronger can be the connection a consumer has with the destination (Chi et al., 2020). Although image is not the sole component of destination branding, destination image is central to the formation of a destination branding model (Konecnik and Gartner, 2007). It has three components: cognitive, affective and unique (Qu et al., 2011; Marques et al., 2021).

Cognitive image refers to the individual’s own knowledge and beliefs about the destination while the affective image is associated with emotions and feelings about it (Sánchez-Rebull et al., 2018). The cognition also results of one’s comprehensive evaluation of the utility of a product/service rooted in

the perceptions of what is obtained and given (Han & Hwang, 2018). Although the concepts of cognitive and affective images components are different, the formation and structure of the affective image may rely on the cognitive image, so they are interrelated (Xu et al., 2018). It was already shown that the cognitive evaluation of an image is a more dominant factor among non-visitors, and the affective evaluation becomes stronger once a tourist visits the place (Marques et al., 2021).

The unique image is measured by features which sets a place apart from all the other. These unique characteristics, such as local gastronomy and tradition, grouped together form the unique image, a set of both cognitive and affective characteristics which are unique and typical for a certain place and which help to differentiate it from other destinations in tourists' mind (Qu et al., 2011; Marques et al., 2021). Unique image helps building destination identities and marketing differentiation strategies that upgrade a destination's competitiveness (Lin & Kuo, 2018). There are a small number of studies abording this concept; Qu et al. (2011) has shown that unique image has a stronger influence on forming the overall image of a destination than the affective cue and that is important as a key differentiator of the destination proposition.

Accordingly, this dissertation adopts an approach of three components for measuring destination image (cognitive-affective-unique) when establishing the interrelationships between this construct and other variables such as awareness.

Destinations use images extensively in their promotional literature to gain awareness for the attributes that set them apart from competitors (Gartner & Ruzzier, 2011). As well as Kaushal et al., (2019), Kusumaningrum (2021) has shown a positive influence of destination brand awareness on destination brand image advising that making marketing promotions, for example, through social media, would be valuable for the success of the destination. Also, Kim & Lee (2018) showed a positive influence of awareness on brand image, but that advertisement does not affect brand awareness, recommending the use of Word of Mouth (WoM). According to the destination brand equity model proposed by Konecnik and Gartner (2007), the awareness dimension affects the cognitive image component, thus affecting the destination brand image (Gartner & Ruzzier, 2011). Further, tourist's awareness of destination influences the positive perceptions of it, manipulating the expectation and actual experience of the destination's perceived service quality (Chen & Myagmarsuren, 2010).

Therefore, the following hypotheses were formulated:

- H1a: Tourists who have a better destination brand image have better destination cognitive image;
- H1b: Tourists who have a better destination brand image have better destination affective image;
- H1c: Tourists who have a better destination brand image have better destination unique image;

- H2: Tourists who have a better destination brand awareness have better destination brand image.

Perceived quality is another important dimension of brand equity in the sense that it provides value to consumers by giving them a reason to buy and by differentiating the brand from competitor brands (Aaker, 1991). Destination brand perceived quality refers to perceptions of quality of the facilities and nonphysical aspects of the destinations (Pike & Bianchi, 2016). Also, Gartner & Ruzzier (2011) argued that destination experience is not required to assess destination quality because it is inherent to the individual and may be refined through experience with similar products. In conceptualizing a destination brand equity model, perceived quality is one of the constructs frequently used by tourism researchers (e.g., Boo et al., 2009; Konecnik & Gartner, 2007; Pike et al., 2010; Gartner & Ruzzier, 2011). There is a growing consensus over the positive impacts of perceived quality on the intention of visitors to keep loyal to the same destination brand (Kaushal et al., 2019). Further, perceived quality of brands plays a significant role in increasing value to brand loyalty and consumers' purchases (Chi et al., 2020) but quality is often viewed as simply meeting or exceeding expectations (Gartner & Ruzzier, 2011). However, perceived quality is a vital attribute of brand equity in the sense that it creates value for consumers by differentiating the brand from competitors and giving consumers a reason to buy (Chi et al., 2020). The awareness of brand can influence quality assessment in terms of price paid to obtain the product or service, also that the higher quality evaluations are predicated on the brand awareness of customers (Kaushal et al., 2019).

In 2021, Kusumaningrum (2021) studied the Borobudur destination in Indonesia and the results showed that the more successful Borobudur is in creating awareness to tourists, the more positive the impression of tourists on the image and quality of Borobudur, enhancing positive relationships between awareness and image with perceived quality. However, Kaushal et al. (2019) also studied these relationships and found a non-significant relationship between destination image and perceived quality, when studying the brand equity of a tourism destination in northern India. Still, the relationship between awareness and perceived quality was also not verified in the studies of Kim & Lee (2018) and Chen & Myagmarsuren (2010). It was concluded that a tourist's awareness, which influences the positive perceptions of the destination (Chen & Myagmarsuren, 2010), leads the expectation and actual experience of the destination's perceived service quality. In the same study, the image aspect is considered a direct determinant of tourists' perceived quality. According to Gartner & Ruzzier (2011), quality should be a central feature of destination awareness marketing to the renewal market as well as featured prominently in communications with repeat visitors. Further, image and quality combine as the central organizing dimensions that should be used to create brand awareness and build loyalty with the repeat market.

The following hypotheses were formulated:

- H3: Tourists who have a better destination brand awareness have better destination brand perceived quality;
- H4: Tourists who have a better destination brand image have better destination brand perceived quality.

Keller (1993) distinguished destination brand loyalty as the main source of customer-based brand equity. It is a vital factor for achieving repeat visitation and positive WoM among visitors (Pike & Bianchi, 2016). Further, loyalty is an expression of customers' perpetual engagement with a brand, which manifests over an extended time span of exchange rather than in a short duration (Kaushal et al., 2019). Furthermore, it is suitable for application with prospective visitors as well as previous visitors (Pike & Bianchi, 2016). For that reason, to cultivate brand loyalty of a tourism destination, efforts should be made to improve tourism destination perceived quality and brand image (Kim & Lee, 2018). Also, the intention to revisit a destination and the willingness to recommend it to other people are positively affected by tourist satisfaction (San Martín et al., 2019). This means that the destination brand loyalty depends on the level of tourism satisfaction. Organizations, in particular, are aiming to forge a unique and efficient image and identity that will encourage tourist travel intentions, with the aim of making those tourists into loyal visitors (Horng et al., 2012). Also, Kaushal et al. (2019) considered the correspondence among tourists' self-image and destination image as an imperative antecedent of destination brand loyalty.

The impact of awareness, image and perceived quality on destination brand loyalty was studied through the years. The first one was tested by Im et al. (2012) that confirmed a positive impact. Also, Kladou & Kehagias (2014) and Kim & Lee (2018) studied the brand equity of Rome and Greater China, respectively, and the results showed that brand quality has a positive impact on brand loyalty. However, Kaushal et al. (2019), Pike and Bianchi (2016), Bianchi & Pike (2011) and Chen & Myagmarsuren (2010) concluded that the brand quality is not significantly related to destination brand loyalty. Regarding to the relationship between destination brand image and loyalty, all these studies plus Kim & Lee (2018) confirmed this relationship except Chen & Myagmarsuren (2010). This happens due to the fact that perceived service quality with an expectation and service performance are both specific beliefs of a future experience, which may mediate the impact of the image (Chen & Myagmarsuren, 2010). This study also concluded that an increase in perceived quality leads to a customer's overall satisfaction, but perceived quality does not influence loyalty directly.

The following hypotheses were formulated:

- H5: Tourists who have a better brand awareness about a tourism destination have better destination brand loyalty;

- H6: Tourists who have a better brand image about a tourism destination have better destination brand loyalty;
- H7: Tourists who have a better perceived quality about a tourism destination have better destination brand loyalty.

2.2.4. CBBE models and constructs testing

As already mentioned, the complex characteristics of a destination make branding and measuring the brand equity a challenge (Boo et al., 2009). For that reason, the specific elements that compose a destination brand and the measurement methods being used are still under analysis (Ferns & Walls, 2012). While no universally accepted CBBE model has emerged, new constructs and relationships between them have been established in order to gather a more complete analysis. Some studies included new constructs to the CBBE measurement such as destination brand value (Boo et al., 2009; Pike & Bianchi, 2016); destination familiarity and travel intentions (Horng et al., 2012; Chi et al., 2020); satisfaction (Chen & Myagmarsuren, 2010; San Martín et al., 2019) and destination brand salience (Pike et al., 2010; Bianchi et al., 2014).

Boo et al.'s (2009) results show that there is a positive relationship between the perceived value of a product's brand and future behavioural intentions characterized as repurchase intention, and customer value plays an important role in creating customer loyalty. Moreover, Pike & Bianchi (2016) also studied this construct and the results reveal weaker effects of brand value for the long-haul travellers compared with the short-haul travellers.

Destination familiarity significantly impacts travel intentions and can potentially play an important role in the travelling decisions (Chi et al., 2020). Moreover, the results of the study showed that brand equity can be seen as an important indicator and a key determinant of travel intentions (Chi et al., 2020). Also, Horng et al. (2012) concluded that destination familiarity might increase foreign tourist brand awareness and brand image for travel intentions, but without significant evidence. On the other hand, destination familiarity increases perceived quality and brand loyalty for travel intentions among foreign tourists (Horng et al., 2012).

Satisfaction, which is a tourist's cognitive-affective state derived from the experience at the destination, evidenced a strong link between "quality-satisfaction-loyalty" regardless of the origin of tourists (San Martín et al., 2019).

Destination brand salience, conceptualized as the strength of awareness of the destination in the mind of an individual when a given travel situation, is related to the same purpose of brand awareness study (Pike et al., 2010). In that way it can be considered the same construct. It was suggested as an indirect relationship between destination brand salience and attitudinal loyalty for short-haul destination context and it has been verified positively (Bianchi et al., 2014).

From the analysed articles, the constructs that were studied, the hypotheses that were formulated and the corresponding findings are summarized in Table 2.6 where the symbol (+) stands for direct and positive relationship between the constructs.

Table 2.6: Summary of research specifications of the studies included in the SLR

ID	Constructs	Hypotheses	Findings
1	Destination Brand Awareness (DBA) Destination Brand Image (DBI) Destination Brand Loyalty (DBL) Destination Quality Perception (DQP) Destination Brand Equity (DBE)	H1: DBA → DBI (+) H2: DBA → DQP (+) H3: DBI → DQP (+) H4: DBI → DBL (+) H5: DQP → DBL (+) H6: DBI → DBE (+) H7: DBA → DBE (+) H8: DQP → DBE (+) H9: DBL → DBE (+)	<ul style="list-style-type: none"> • The more successful the destination in creating awareness to tourists, the more positive is the impression of tourists on the quality of the destination. • The destination should be able to build an effective and unique impression that suits the personality and style of the tourists. • H1, H2, H3, H4, H5, H8 and H9 were positively validated.
2	Destination Brand Awareness (DBA) Destination Brand Image (DBI) Destination Brand Loyalty (DBL) Destination Brand Perceived Quality (DBQ) Travel Intentions (TI) Destination Familiarity (DF)	H1: DBQ → TI (+) H2: DBI → TI H3: DBL → TI (+) H4: DBA → TI (+) H5: DF has a positive moderating effect on the previous relationship hypothesised	<ul style="list-style-type: none"> • The TI of tourists will be influenced by brand equity. • The moderator, DF, has an impact on TI and brand equity. • Brand equity can be seen as an important indicator and a key determinant of travel intentions. • H1, H2, H3 and H4 were positively validated. • It was confirmed that DF has a positive moderating effect in H1 and H4.
3	Destination Brand Awareness (DBA) Destination Image (DI) Social Self-Image (SSI) Destination Brand Loyalty (DBL) Perceived Quality (PQ) Perceived Value (PV)	H1: DBA → PQ (+) H2: DBA → DI (+) H3: DBA → PV (+) H4: DI → PQ (+) H5: DI → DBL (+) H6: DI → DPV (+) H7: SSI → DBL (+) H8: PQ → DBL (+) H9: PQ → PV (+) H10: PV → DBL (+)	<ul style="list-style-type: none"> • Destination quality indicators did not find to have impacted tourist behaviour. • The presence of favourable SSI may imply that visitors hold the state attractions in higher esteem. • The price perception created by visitors seemed vital in determining their future association with the destination brand. • H1, H2, H5, H7, H9 and H10 were positively validated.
4	Destination Brand Awareness (DBA) Affective Image (AI) Cognitive Image (CI) Destination Perceived Quality (DBQ) Satisfaction (Sat) Destination Brand Loyalty (DBL)	H1: DBA → CI (+) H2: DBA → AI (+) H3: CI → DPQ (+) H4: AI → DPQ (+) H5: CI → AI (+) H6: DPQ → Sat (+) H7: Sat → DBL (+)	<ul style="list-style-type: none"> • The causal relationships between DBA, DBI, DBQ AND DBL are sequential. • Both dimensions of destination image are interrelated (in particular, AI is positively influenced by). • Once tourists have visited the destination, Sat appears as a key variable for brand equity. • All hypotheses were validated except H3.

Table 2.6: Summary of research specifications of the studies included in the SLR (cont.)

ID	Constructs	Hypotheses	Findings
5	Destination Brand Perceived Quality (DBQ) Destination Brand Awareness (DBA) Destination Brand Image (DBI) Destination Brand Loyalty (DBL)	H1: Tourism influencing factors → DBQ (+) H2: Tourism influencing factors → DBA (+) H3: Tourism influencing factors → DBI (+) H4: DBA → DBQ (+) H5: DBA → DBI (+) H6: DBQ → DBI (+) H7: DBQ → DBL (+) H8: DBI → DBL (+)	<ul style="list-style-type: none"> • The impact of brand equity influencing factors on tourism destination brand awareness shows that price, promotion, and WoM have an important influence. • WoM is the more influential. • To cultivate DBL of a tourism destination, efforts should be made to improve tourism DBQ and DBI. • H1, H2 and H3 were partially validated. • H5, H6, H7 and H8 were confirmed.
6	Destination Brand Salience (DBS) Destination Brand Image (DBI) Destination Brand Quality (DBQ) Destination Brand Value (DBV) Attitudinal Destination Loyalty (ADL)	H1: DBS → ADL (+) H2: DBQ → ADL (+) H3: DBI → ADL (+) H4: DBV → ADL (+)	<ul style="list-style-type: none"> • Brand salience, brand image, and brand value are positively related to brand loyalty. • DBS and value have stronger effects on DBL for short-haul travellers. • H1, H3 and H4 were positively validated.
7	Destination Brand Image (DBI) Destination Image (DI)	H1: DBI → Wine Tourism Destination Brand Equity (+) H2: DI → Wine Tourism Destination Brand Equity (+) H3: DI → DBI (+)	<ul style="list-style-type: none"> • The images of destinations can help strengthen the brand equity of wine tourism destinations. • Developing and reinforcing favourable, strong, and unique associations is important for increasing wine tourism destination brand equity. • H1, H2 and H3 were confirmed.
8	Cultural Brand Assets (CBA) Destination Brand Awareness (DBA) Destination Brand Associations (DBAss) Destination Brand Quality (DBQ) Destination Brand Loyalty (DBL) Cultural Destination Brand Equity	H1: Brand Equity Dimensions → Cultural Destination Brand Equity (+) H2: CBA → DBA (+) H3: DBA → DBAss (+) H4: DBA → DBQ (+) H5: DBAss → DBL (+) H6: DBQ → DBL (+)	<ul style="list-style-type: none"> • All five brand equity dimensions have a statistically significant impact on cultural destination brand equity. • Assets have a positive impact on both awareness and quality, while awareness has an impact on both quality and associations. • Associations should be connected with quality in order to achieve a bigger impact of perceived associations on loyalty. • All hypotheses were positively validated.
9	Enduring Travel Involvement Destination Brand Perceived Quality (DBQ) Destination Brand Awareness (DBA) Destination Brand Image (DBI) Destination Brand Loyalty (DBL)	H1: Enduring Travel Involvement → DBI (+) H2: Enduring Travel Involvement → DBA (+) H3: Enduring Travel Involvement → DBQ (+) H4: Enduring Travel Involvement → DBL (+)	<ul style="list-style-type: none"> • Brand experience (image and quality) was found to have the strongest relationship to travel involvement. • A person's degree of interest in travel does indeed have a direct impact on their understanding and perceptions of the brand-equity of a destination. • Enduring travel involvement does not influence destination loyalty. • H1, H2 and H3 were confirmed.

Table 2.6: Summary of research specifications of the studies included in the SLR (cont.)

ID	Constructs	Hypotheses	Findings
10	Internal Brand Awareness (IBA) Internal Brand Image (IBI) Internal Brand Commitment (IBC) Internal Brand Loyalty (IBL)	H1: IBA → IBI (+) H2: IBI → IBC (+) H3: IBC → IBL (+)	<ul style="list-style-type: none"> • Recursive relationship between the external and internal performance of the destination brand. • It is important to have a participative approach to the branding process in order to enhance its overall internal equity. • All hypotheses were positively validated.
11	Destination Brand Awareness (DBA) Destination Brand Image (DBI) Destination Brand Loyalty (DBL) Destination Brand Associations (DBAss) Overall Brand Equity (OBE)	H1: DBA → DBL (+) H2: DBI → DBL (+) H3: DBAss → DBL (+) H4: DBA → OBE (+) H5: DBI → OBE (+) H6: DBAss → OBE (+) H7: DBL → OBE (+)	<ul style="list-style-type: none"> • DBA is a pivotal element in creating the value of the destination as a favourable destination brand in the minds of travellers. • The influence of image in OBE is through DBL. • High brand equity is affected by enhancing and strengthening the customer's perception of brand awareness, brand loyalty, brand quality, and brand attitude relating to the brand. • H2, H3, H4, H6 and H7 were supported.
12	Destination Perceived Quality (DBQ) Destination Brand Awareness (DBA) Destination Brand Image (DBI) Destination Brand Loyalty (DBL) CBBE for a tourism destination	H1: DBA → DBQ (+) H2: DBA → DBI (+) H3: DBQ → DBI (+) H4: DBQ → DBL (+) H5: DBI → DBL (+) H6: DBA → DBL (+) H7: The importance of each dimension for the second-order factor, CBBE for a tourism destination, is not the same.	<ul style="list-style-type: none"> • The quality dimension showed as much strength for both the renewal and repeat markets as the image dimension. • Destination experience is not required to assess destination quality since the quality is a subjective judgment that is held closely by the individual. • The most important dimensions for CBBE were quality and image and there was confirmed relationships between dimensions.
13	Destination Brand Salience (DBS) Destination Brand Image (DBI) Destination Brand Quality (DBQ) Destination Brand Value (DBV) Destination Brand Loyalty (DBL)	H1: DBS → ADL (+) H2: DBQ → ADL (+) H3: DBI → ADL (+) H4: DBV → ADL (+)	<ul style="list-style-type: none"> • The perceptions of past visitors were significantly more positive than those of non visitors for brand salience, brand quality, and brand image. • Although brand quality is the best performing dimension, the results show that it was not significantly related to brand loyalty. • H1, H3 and H4 were positively validated.
14	Destination Perceived Quality (DBQ) Destination Brand Awareness (DBA) Destination Brand Image (DBI) Destination Brand Loyalty (DBL) Satisfaction (Sat)	H1: DBA → DBI (+) H2: DBA → DBQ (+) H3: DBI → DBQ (+) H4: DBI → Sat (+) H5: DBQ → Sat (+) H6: DBI → DBL (+) H7: DBQ → DBL (+) H8: Sat → DBL (+)	<ul style="list-style-type: none"> • The image dimension is a direct determinant of tourists' perceived quality. • DBA dimension has no significant influence on DBQ • DBI does not affect tourist satisfaction and loyalty • Sat is a key antecedent of DBL. • H1, H3, H5, H8 were supported.

From the SLR, the articles mostly address the relationships of brand awareness, brand image and perceived quality with destination brand loyalty in which all confirmed positive impacts of brand awareness, perceived quality and image on destination brand loyalty for different destinations and travel intentions (e.g., Chen & Myagmarsuren, 2010; Bianchi et al., 2011; Im et al., 2012; Kim & Lee, 2018). Brand loyalty is at the top of the CBBE hierarchies proposed in the work of Aaker (1991) and Keller (1993), and as such, destination brand loyalty is the dependent variable in the measurement of destination brand equity in this dissertation.

Figure 2.3 shows the proposed conceptual model and hypotheses used to measure the destination brand equity.

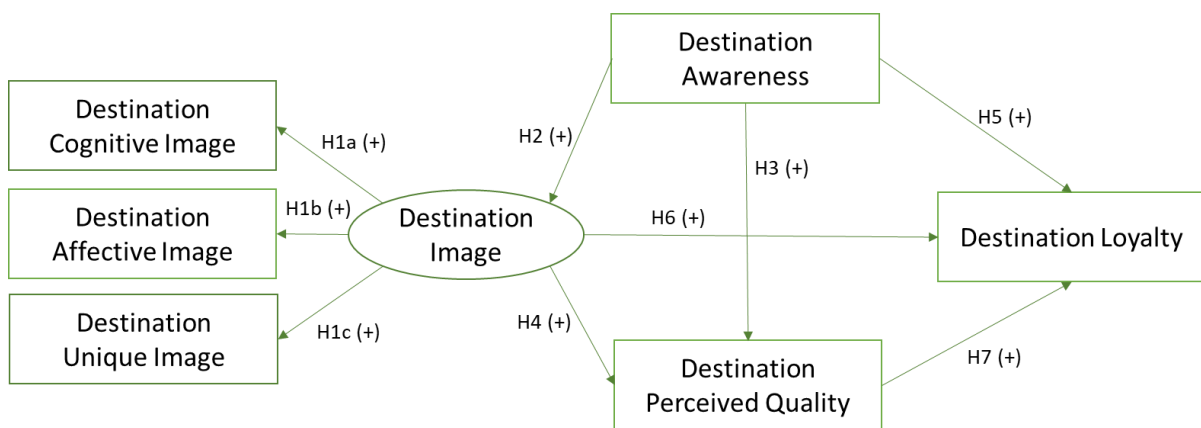


Figure 2.3: Proposed conceptual model

2.2.5. Research methods

Most of the studies have used empirical studies based on a cross-sectional survey, i.e., a single sample obtained at a single moment in time. Table 2.7 presents a summary of the research methods applied in the 14 articles. However, some studies tested the survey on different samples, such as national and foreign tourists (e.g., Pike & Bianchi, 2016; Sánchez-Rebull et al., 2018) to find differences among them. The use of different methods of research through online recommendations or content available online to measure the constructs of destination brand equity can represent an improved contribution to existing approaches, since it allows to capture the real visitors experience from a broad perspective.

Online recommendations are one of the most important factors in consumers' decision-making processes, as consumers believe that the opinions of their peers are more reliable than those posted by the service provider (Guerreiro & Rita, 2020). The novelty of this dissertation is to measure the destination brand equity from the visitor's experience, as a whole, already available from online recommendations in tourist blogs.

Regarding to the modelling techniques, mostly the 14 articles used structural equation modelling, a multivariate technique to test and evaluate multivariate causal relationships (San Martín et al., 2019). The main advantages of this method include simultaneous estimation of multiple and interrelated dependence relationships, representing an unobservable concept, and accounting for measurement errors in the estimation process (Ferns & Walls, 2012).

Table 2.7: Research methodology

ID	City/Country of Study	Research method (Sample Size)	Modelling techniques
1	Borobudur, Indonesia	Purposive sampling (167)	Structural Equation Modelling
2	Vietnam	Survey (531)	Partial Least Squares
3	Uttarakhand, India	Survey (580)	Structural Equation Modelling
4	Cantabria, Spain	Two samples for national and international tourists (667)	Structural Equation Modelling
5	Seoul, South Korea	Survey (385)	Structural Equation Modelling
6	Australia	Survey (New Zealand sample – 858 and Chile sample – 845)	Structural Equation Modelling
7	Spain	Survey (817)	Partial Least Squares-based Multi-Group Analysis
8	Rome, Italy	Survey (450)	Path Analysis
9	Travel region in a US Midwest state	Survey (195)	Structural Equation Modelling
10	South Tyrol's, Italy	Survey (306)	Exploratory Factor Analysis and descriptive statistics
11	Malaysia	Survey (326)	Structural Equation Modelling
12	Slovenia	Telephone interviews (376)	Confirmatory Factor Analysis
13	Australia	Survey (341)	Structural Equation Modelling
14	Mongolia	Survey (128)	Structural Equation Modelling

2.2.6. Limitations and existing gaps

The last quality criteria of this SLR are related to the contributions and limitations identified in the articles, which help identifying existing gaps in the literature. Table 2.8 summarises the contributions and limitations of the 14 articles of the SLR. Most studies addressed brand equity models in one destination in which the method of research is a survey. Although, surveys represent an effective method in measuring destination brand equity, they have some identified limitations. Kusumaningrum (2021) and Chi et al. (2020) argue that, due to surveys being only answered by tourists, the number of respondents is limited and different sociodemographic variables are not considered. Also, San Martín et al. (2019) alerts for the generalization of the results when focusing on small regions that cannot be transpose to the country brand equity. Another limitation that must be taken in account is the origin of the tourists, i.e., the country they are, as well as if they are repeat visitors or first-time travellers (Chi et al., 2020; Kim & Lee, 2018; Pike & Bianchi, 2016).

Table 2.8: Contributions and limitations

ID	Contributions	Limitations
1	<ul style="list-style-type: none"> • Testing an existing model on a different destination. 	<ul style="list-style-type: none"> • Testing a model in only one destination; • Small number of respondents.
2	<ul style="list-style-type: none"> • Use of conception of destination brand equity to measure travel intentions; • Introduce destination familiarity as a moderator of the relationship between destination brand equity and travel intentions. 	<ul style="list-style-type: none"> • Not considering the roles of different sociodemographic variables; • Respondents were only tourists; • Not comparing prospective, first-time, and repeat visitors from the viewpoint of destination branding.
3	<ul style="list-style-type: none"> • Adding a new variable to brand equity study: social self-image; • Testing a model in a latent tourism destination. 	<ul style="list-style-type: none"> • Not considering the variables' moderating and mediating effects among the set of constructs of destination brand equity.
4	<ul style="list-style-type: none"> • Considering national and international tourists; • Including tourism satisfaction and brand image components: cognitive and affective. 	<ul style="list-style-type: none"> • Focusing on a small region that could limit the generalization of the results.
5	<ul style="list-style-type: none"> • Testing the effect of influencing factors (price, advertising, promotion, and word on mouth) on destination brand equity constructs. 	<ul style="list-style-type: none"> • Focusing on only Chinese tourists can limit objectivity and generalization; • Not including marketing strategies designed to enhance the brand equity.
6	<ul style="list-style-type: none"> • Considering tourists from two different countries; • Compare short and long-haul travellers. 	<ul style="list-style-type: none"> • Limited to tourists from two countries; • Only considering attitudinal loyalty and not behavioural loyalty.
7	<ul style="list-style-type: none"> • Studying the impact of designation of origin brand equity on a wine tourism destination; • Includes two samples (managers and visitors); • Testing a reflective-formative type model. 	<ul style="list-style-type: none"> • Not considering stakeholders such as local residents on the analysis; • Focusing on only five wineries in one country.
8	<ul style="list-style-type: none"> • Including five dimensions of brand equity in an integrated model; • Following an integrated approach for describing the developed relationships. 	<ul style="list-style-type: none"> • Limited list of cultural and destination assets taken into consideration; • Excluding assets' impact on brand equity dimensions in terms of positioning.
9	<ul style="list-style-type: none"> • Testing a model in a dynamic process of information search using data collected from potential travellers before their trips occur. 	<ul style="list-style-type: none"> • Small sample size; • Considering only one destination.
10	<ul style="list-style-type: none"> • Providing an internal stakeholder-based destination brand equity overview. 	<ul style="list-style-type: none"> • The concepts and scales are specific and sensitive to the destination.
11	<ul style="list-style-type: none"> • Exploring multidimensionality of consumer based brand equity within a tourism context. 	<ul style="list-style-type: none"> • Limited to the dimensions constructed for the specific destination.
12	<ul style="list-style-type: none"> • Studying the dimensions of brand equity and their effect on repeat and renewal markets. 	<ul style="list-style-type: none"> • Sample size and analysis focus on one group of present and potential tourists.
13	<ul style="list-style-type: none"> • Including past and non-visitors to identify perceptual and attitudinal differences. 	<ul style="list-style-type: none"> • Sample is biased in gender and education; • Limited to tourists from one country.
14	<ul style="list-style-type: none"> • Identifying the causal relationship between perceptual brand equity and behavioural brand equity; • Including tourism satisfaction as an antecedent of destination loyalty. 	<ul style="list-style-type: none"> • Survey designed and used only in English.

Further, there were identified limitations related to the conceptual model that was tested. For example, Kaushal et al. (2019) did not considered the variables' moderating and mediating effects among the set of constructs of destination brand equity and Pike & Bianchi (2016) only have considered attitudinal loyalty and not behavioural loyalty. However, the main contributions of the articles are related to testing a model in a different destination and adding new variables to already tested models.

The method of research, due to the limitations identified that can influence the analyses, constitutes a research gap that will be addressed in this dissertation, hence, contributing with new approaches to measure destination brand equity while trying to mitigate the described limitations.

2.2.7. Evaluation of the relevance of the articles

In order to understand the relevance of the selected articles within the scope of this dissertation, evaluation criteria, already presented in Table 2.4, are applied, allowing to rate the articles according to its relevance. For each quality criterion, a rating is given: 0 if does not answer the question; 0.5 if incompletely answered and 1 if fully answered (Abrantes, 2020). The ratings of each article and the quality criteria are described in Table 2.9.

Table 2.9: Quality evaluation of the articles of SLR

ID	Destination Brand equity concepts		Models of Destination BE measurement			Method of research			Conclusions		Total ¹
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	
1	0.5	0.0	0	1.0	1.0	0.5	0.0	0.5	1.0	1.0	5.5
2	1.0	1.0	1.0	1.0	1.0	1.0	0.0	1.0	1.0	1.0	9.0
3	1.0	1.0	1.0	0.5	1.0	1.0	0.0	1.0	1.0	1.0	8.5
4	0.5	1.0	1.0	1.0	0.5	1.0	0.5	1.0	1.0	1.0	8.5
5	1.0	1.0	0.5	1.0	1.0	1.0	0.5	0.5	0.5	0.5	7.5
6	1.0	0.5	0.5	1.0	1.0	0.5	0.5	1.0	1.0	1.0	9.0
7	0.5	1.0	1.0	1.0	1.0	1.0	0.5	0.5	1.0	1.0	8.5
8	1.0	0.5	0.0	1.0	1.0	1.0	0.0	0.0	1.0	1.0	6.5
9	1.0	1.0	0.5	1.0	1.0	1.0	1.0	0.0	1.0	0.5	8.0
10	1.0	1.0	1.0	1.0	0.5	1.0	1.0	0.0	1.0	0.0	7.5
11	1.0	0.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	1.0	6.0
12	1.0	0.0	0.0	0.5	1.0	1.0	0.5	0.0	1.0	1.0	6.0
13	1.0	1.0	1.0	1.0	0.5	0.5	0.0	0.0	1.0	1.0	7.0
14	1.0	0.5	1.0	1.0	0.5	1.0	0.0	0.5	1.0	0.5	6.5
Total²	12.5	9.5	8.5	13.0	12.0	12.5	4.5	5.0	13.5	12.5	

Note: Total¹: total score of the article; Total²: total score of the quality criterion.

From the selected articles of the SLR, all articles scored more than 5, in which the article with the worst score is the most recent article with 5.5/10, (Kusumaningrum, 2021), and there are two best scored articles, with 9/10 (Pike & Bianchi, 2016; Chi et al., 2020).

On one hand, the total scores of the quality criteria show that the best criteria are the contributions presented, followed by the limitations, identification, and concepts description. This result allows concluding that all the selected articles are useful as the basis for a complete literature review, as well as to identify limitations in the current studies and gaps needed to be addressed in future research.

On the other hand, the worst scores of the quality criteria are related to the methods of research. This is due to the fact that, in most studies, surveys are applied to a single sample at a single moment in time.

CHAPTER 3

Methodology

In this study, the Design Science Research (DSR) methodology is used, considering that it brings both practical relevance, via its emphasis on useful artifacts, and scientific rigor, via the formulation of design theories (Baskerville et al., 2018). It involves a rigorous process to design artifacts to solve observed problems, to make research contributions, to evaluate the designs, and to communicate the results to appropriate audiences (Hevner et al., 2004). This means that this methodology is oriented to problem solving where such artifacts may include constructs, models, methods, and instantiations. Vaishnavi & Kuechler (2004) describes the procedures of the DSR methodology through five steps, as illustrated in Figure 3.1. Although it suggests a sequential order, there is an iterative process between the steps, especially in the development and evaluation phase, where several iterations can also take place (Helms et al., 2010), e.g., if the evaluation metrics are under the expected, the development phase has to be improved to upgrade the model.

The first three steps of the DSR methodology are presented in this chapter. Since the proposed model is a confirmatory model of a theoretical model, the evaluation step is not executed. The conclusion step is presented in chapters 4 and 5 in which the model results are presented and discussed.

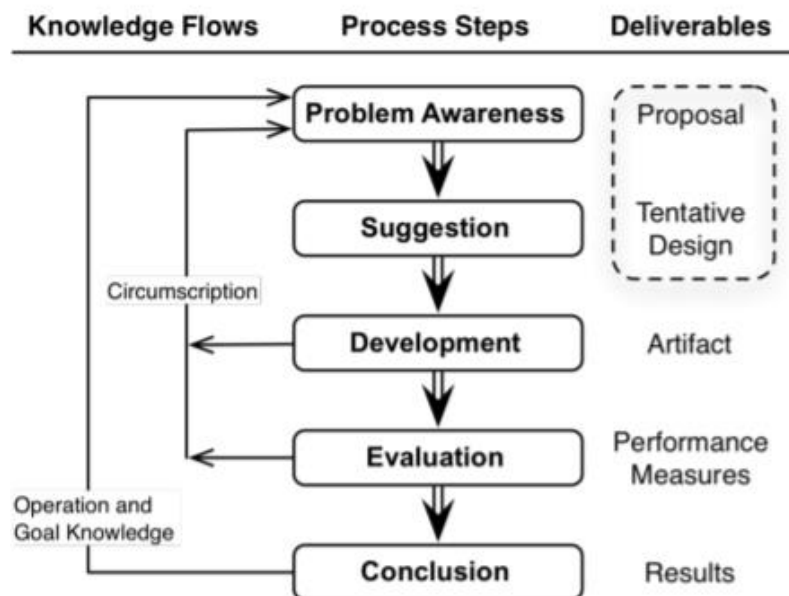


Figure 3.1: General steps of DSR methodology

Source: Vaishnavi & Kuechler (2004)

3.1. Problem awareness

The first step concerns awareness about a relevant business problem (Helms et al., 2010). The systematic literature review allowed to identify information about the destination brand equity constructs, mostly obtained from cross-sectional surveys, which is considered a limitation of some existing approaches. As already mentioned, although surveys represent an effective method in measuring destination brand equity, there are other available research methods that can be used to provide a different point of view for destination brand equity measurement, e.g., interviews or online reviews.

3.2. Suggestion

Suggestion is a creative step wherein new functionality is envisioned based on a novel configuration of either existing or new and existing elements (Vaishnavi & Kuechler, 2004). The suggestion presented in this dissertation is to develop a destination brand equity measurement through a new method of research. The new method being proposed in this dissertation is text mining through sentiment analysis of user-generated content, available in blogs of tourists who visit a destination. The study destination chosen is Lisbon, Portugal.

3.2.1. User-generated content

User-generated content is broadly defined as media content created or produced by the general public rather than by paid professionals (Poch & Martin, 2015). As previously mentioned, online recommendations are one of the most important factors in consumers' decision-making processes, as consumers believe that the opinions of their peers are more reliable than those posted by the service provider (Guerreiro & Rita, 2020). If the online reviews are positive, they work as free advertising for destinations and tourism providers, however, negative online comments can have the opposite effect and damage the image and reputation of a company/destination (Chen & Law, 2016).

Tourism blogs may be considered a type of WoM communications, since the current definitions of WoM includes not only all types of informal interpersonal communications from people that the consumers know but also sources from online platforms and other influencers which are not related to the brand or the sellers (Confente, 2015). Jalilvand et al. (2012) empirical results confirmed that electronic WoM positively affects the destination image, tourist attitude, and travel intention. Specifically, a satisfactory experience of tourists with a tourism destination can lead to a high possibility of repeating visits and recommendation of the destination to others (Chen & Law, 2016).

3.2.2. Study destination: Lisbon, Portugal

In Portugal, the tourism sector is a main economic activity contributing to wealth and employment and it has been increasing through the last years (Turismo de Portugal, 2022). According to the United Nations World Tourism Organization (UNWTO), Portugal is one of the most visited countries in the Europe and mostly visited by citizens of United Kingdom, Spain, France, Germany, Brazil and United States (UNWTO, 2022).

From 2010 to 2019, Portugal has registered an average annual growth rate of 7.2% in overnight stays, which translates into an increase from 45 million overnight stays in 2010 to 70 million overnight stays in 2019 (the highest value on record), Figure 3.2 (Statista, 2022). In 2020, because of the COVID-19 pandemic, Portugal registered, along with other world destinations, a sharp drop in demand, hence, it was not taken in account in this study. Since, starting at 2021, the tourism sector is recovering, it is expected that in the incoming years Portugal will recover from the impact of the COVID-19 pandemic, thus, experiencing similar figures as the ones before COVID-19 pandemic.

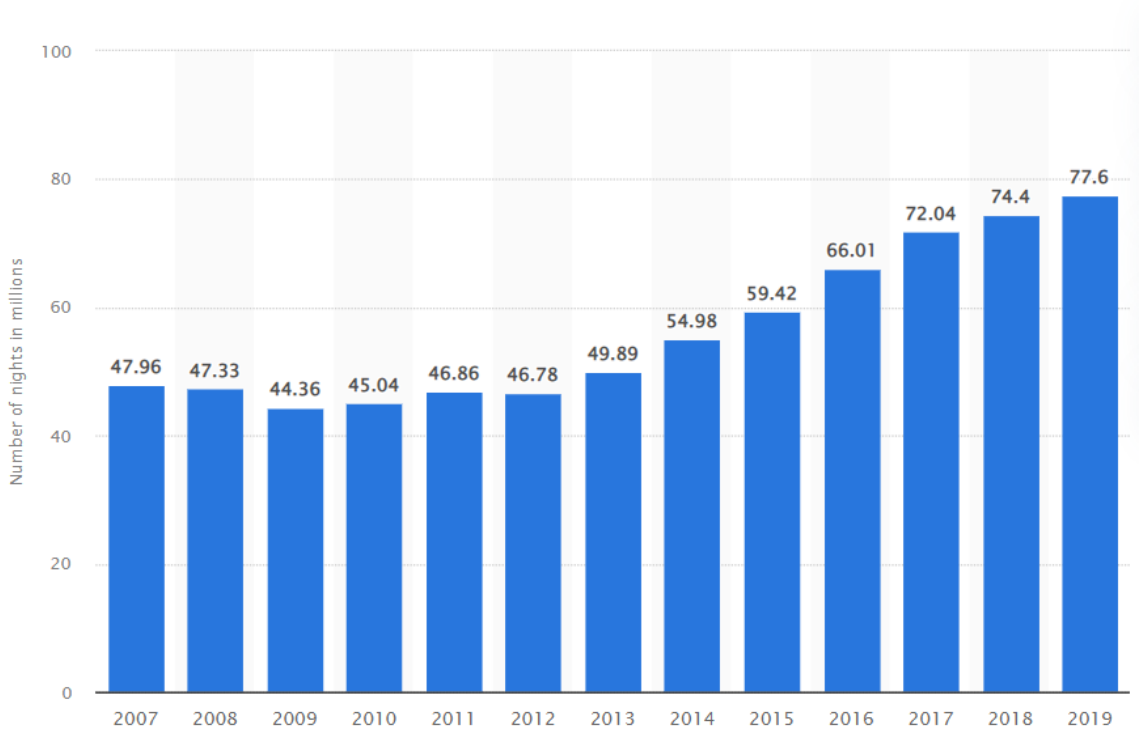


Figure 3.2: Number of overnight stays in Portugal

Source: Statista (2022)

Lisbon is the capital and the largest city of Portugal with a lively nightlife scene, festive markets, and vibrant museums. With over 20 centuries of History, this city located on the “7 hills” has more than 500,000 inhabitants. The wide variety of landscapes and heritage makes this destination a must to visit (Visit Portugal, 2022). With beaches, natural parks, cultural routes and accommodation for all

tastes, actually Lisbon is the most visited region of Portugal. In 2019, it registered about six million international tourist guests, especially from Spain, United States, Brazil, France and Germany (Turismo de Portugal, 2022).

3.2.3. Text mining and sentiment analysis

After online communications have emerging, text mining and sentiment analysis has been frequently applied into analysing electronic WoM (Kim et al., 2018). Text mining was defined by Hearst (1999, p. 8) as “the use of large online text collections to discover new facts and trends about the world itself”. Generally, it refers to the process of extracting interesting and non-trivial patterns or knowledge from unstructured text documents (Tan, 1999). It provides a framework to maximize the value of information within large quantities of text in which researchers can discover patterns and links between resources, providing more meaningful answers to complex research questions, and even support scientific discovery in various domains (Hassani et al., 2020).

Sentiment analysis is the task of determining the emotional state of a writer based on their written texts by considering the polarity of keywords used in the writing (Li et al., 2014). The polarity of keywords can be positive, negative, or neutral. From the perspective of text mining, sentiment analysis is done by finding a correlation between the writer’s style and emotional state and by categorizing keywords into emotional levels (Argamon et al., 2007). For instance, sentiment analysis is often used in opinion mining which focuses on extracting authors’ opinions and sentiments from user-generated content such as customer reviews, forum messages and blogs (Cambria et al., 2013).

Methods for extracting sentiment and opinion from text can be divided into two approaches: the machine learning approach and the lexicon-based approach in which a lexicon of opinion words is required along with the sentiment score (Mukhtar et al., 2018). In general, dictionaries have been preferred than manual sentiment dictionaries because of their high coverage and reliability, however, a dictionary generated for certain purposes can be more efficient for predicting the emotional opinions (Kim et al., 2018).

3.3. Development

In the development step the whole content is extracted from individual blogs available online, i.e., all the data is user-generated content. Due to the huge amount of textual data available in Web, text-mining of user-generated content has proved to be an effective method, as it not only makes access to information easier and faster but also reflects vivid visitor experiences (Marques et al., 2021).

The dataset to be used for analysis purposes must be large enough while providing the required diversity in order to ensure its statistical relevance. In this dissertation a total of 100 blog posts,

extracted from different international visitor blogs, written in English, Deutsch, French, Spanish and Portuguese languages are used. This approach ensures proper quality reviews for opinion extraction towards tourism destinations (Irawan et al., 2019). According to the theme of this dissertation, the type of blogs being considered are on travelling experiences, being based on the experience and personal opinion of the author of each individual blog. Hence, the posts can include guides, tips, and recommendations about what to do while visiting the city from the author’s perspective.

For the extraction process, *WebHarvy* is used, which is a web scraping software that can be used to easily scrape data from any website (WebHarvy, 2022). After the data extraction, due to the existence of multiple languages, all the data is translated to English using the Yandex Translate and the translation of names of tourist attractions are checked to make sure that they are written in the same way (Yandex, 2022).

The dataset is then loaded to *RStudio*, an integrated development environment for R, which is an open-source software for data analysis supported by a broad community of contributors for packages to perform numerous tasks, including sentiment analysis (RStudio, 2022). The proposed approach, detailed in Figure 3.3, aimed at discovering and classifying sentences towards the blogger's feelings regarding the constructs identified through keywords.

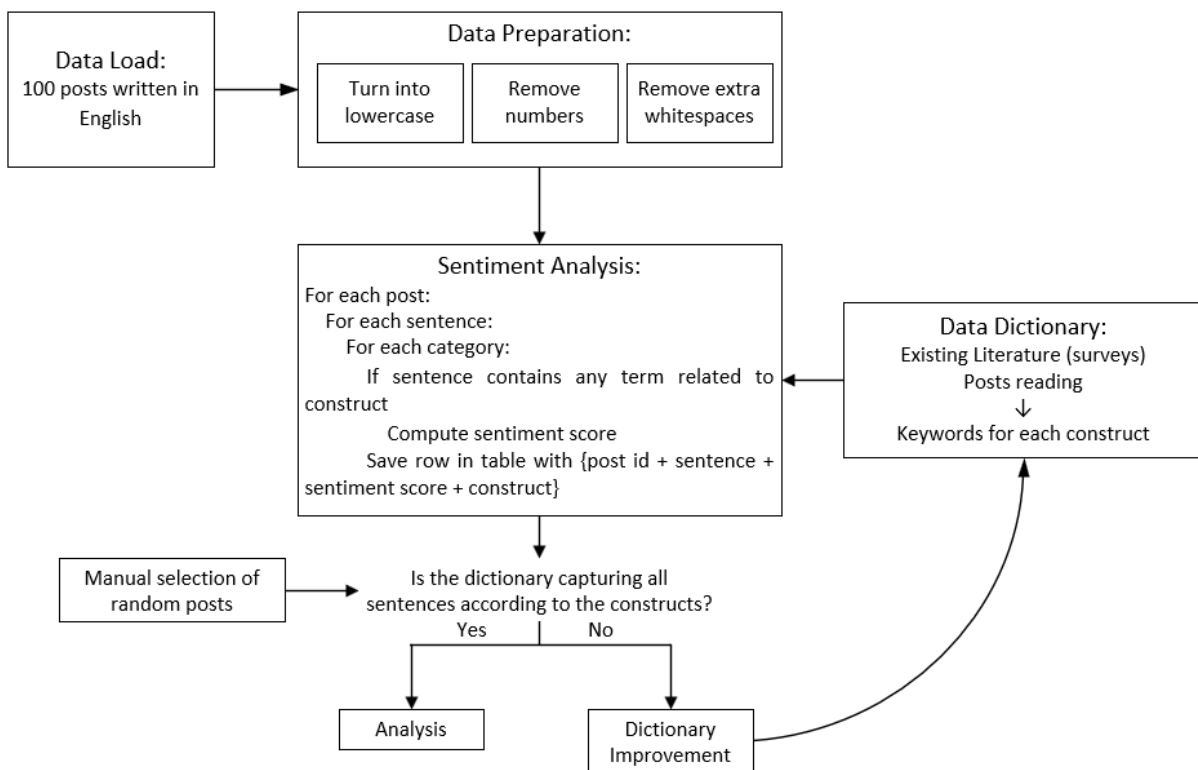


Figure 3.3: Sentiment analysis approach

Before the algorithm is executed, there is a data preparation phase. It consists of the following steps: removing numbers that are useless to the analysis, turning all data to lowercase, avoiding case sensitive issues when identifying the keywords, and finally removing extra whitespaces, thus, allowing to properly separate the sentences that include text. To execute the algorithm, a data dictionary is mandatory to make the connection between the constructs and the sentences.

In order to allow incremental improvement of the dictionary from previous iterations to achieve better results, the proposed approach is cyclic, since the data dictionary is manually developed based on existing surveys on the literature and on manual posts reading. First, the studies that analysed the constructs through surveys are mainly adapted from the questions of Boo et al. (2009). Questions from Boo et al. (2009), Im et al. (2012), Herrero et al. (2017) and Chi et al. (2020) are taken as the initial references to identify the keywords, as follows:

Brand Awareness:

- This destination has a good name/ reputation (Boo et al., 2009);
- This destination is very famous (Herrero et al., 2017; Boo et al., 2009);
- The characteristics of this destination come to tourist's mind quickly (Chi et al., 2020; Boo et al., 2009);
- I recognise this destination as a travel destination among other destinations (Chi et al., 2020; Im et al., 2012);
- I know what this destination looks like (Im et al., 2012);
- I can picture what the destination looks like in my mind (Chi et al., 2020);
- When I am thinking about travelling, this destination comes to my mind immediately (Chi et al., 2020);
- This destination is a clearly recognisable tourist destination (Herrero et al., 2017).

Brand Image:

- Characteristics of tourism facilities and attractions (Advanced technology and economy, high quality of infrastructure, shopping, events, food) (Im et al., 2012);
- Characteristics of environmental, natural and cultural resources (Safe, beautiful...) (Im et al., 2012; Herrero et al., 2017);
- Hospitality and amusement: (friendly locals, wide choice of accommodations and entertainment...) (Im et al., 2012);
- Convenience and comfort (language barriers, costs, weather, cleanliness) (Im et al., 2012);
- Leisure and recreation activities (sport, adventure, ...) (Im et al., 2012; Herrero et al., 2017).

Perceived Quality:

- The destination provides tourism offerings of consistent quality and attractive (Chi et al., 2020; Boo et al., 2009);
- The destination provides quality experiences and excellent products and services (Chi et al., 2020; Boo et al., 2009);
- From this destination's offerings, I can expect superior performance (Chi et al., 2020; Boo et al., 2009);
- This destination is a quality tourist destination (Herrero et al., 2017);
- This destination performs better than other similar destinations (Chi et al., 2020; Boo et al., 2009);
- Tourist products and services are excellent (Herrero et al., 2017).

Brand Loyalty:

- This destination would be my preferred choice for a vacation (Chi et al., 2020; Boo et al., 2009; Konecnik & Gartner, 2007);
- I recommend other people to visit this destination (Chi et al., 2020; Herrero et al., 2017; Im et al., 2012; Konecnik & Gartner, 2007);
- I still consider a trip to this destination even if the travel cost increased (Im et al., 2012);
- I intent to visit in the future (Konecnik & Gartner, 2007);
- I would intend to visit this destination in the next 5 years (Herrero et al., 2017; Im et al., 2012);
- I enjoy visiting this destination (Boo et al., 2009).

After a first iteration, if only a few sentences containing any of the constructs are found, the data dictionary has to be improved. This happens because the loaded posts are written in an informal language that is not captured in the initial dictionary matching terms of the questions to each construct. Because of that, a reading of random posts is performed and new keywords are added to the dictionary. In each iteration, the analysis is re-run and a small set of sentences are evaluated to assure if those contained any word that is not captured as belonging to a given construct. The data dictionary is presented in Table 3.1. The words "cit" and "activit" are used in order to stand for the singular and plural forms of city and activity, respectively.

Table 3.1: Data dictionary

Construct	Keywords
Brand Awareness	“famous destination”, “favorite destination”, “favourite destination”, “trendy”, “always knew”, “most beautiful destination”, “prettiest cit”, “ideal european cit”, “perfect cit”, “best place”, “perfect place”, “attractive destination”, “best european cit”, “good name”, “to my mind”, “icon”, “well-known”, “first choice”, “best for”, “vacation destination”, “looking for”, “perfect destination”
Brand Image	Cognitive Image: “street”, “cultural”, “technology”, “technological”, “weather”, “safe”, “budget”, “monument”, “museum”, “park”, “landscape”, “event”, “nightlife”, “sport”, “adventure”, “shopping”, “restaurant”, “environment”, “accommodation”, “hotel”, “food”, “comfort”, “cheap”, “economic”, “language”, “expensive”, “attraction”, “things to do”, “inexpensive”, “crowd”, “lifestyle”, “locals”, “transport”, “art”, “restaurant”, “view”, “climate”, “shops”, “neighbourhood”, “pickpocketing”, “music”, “hostel”, “meal”, “treatment”, “activit”
	Affective Image: “pleasant”, “fun”, “amazing” “beautiful”, “friendly”, “clean”, “messy”, “dirt”, “old” “affordable”, “instagrammable”, “welcoming”, “romantic”, “unique”, “underrated”, “overrated”, “sustainable”, “accommodating”, “modern”, “authentic”, “dangerous”, “lovely”, “charm”, “gorgeous”, “unbelievable”, “magnific”, “lively”
	Unique Image: “fado”, “coast”, “custard tart”, “seven hills”, “ocean”, “wine”, “sun”, “tiles”, “azulejo”, “tram 28”, “calçada”, “Portuguese pavement”, “bacalhau”, “wave”, “cork”, “colourful”, “colorful”, “cod”, “pastéis de belém”, “ginjinha”, “tagus river”
Perceived Quality	“quality”, “disappointed”, “surprised”, “attractive”, “performance”, “product”, “compared to”, “consistent”, “perfect”, “worse than”, “better than”, “expected more”, “fantastic”, “fabulous”
Brand Loyalty	“recommend”, “revisit”, “visit again”, “future”, “suggest”, “stop at this”, “return”, “come to visit”, “must visit”, “is a must”, “can't miss”, “should visit”, “worth visiting”, “must-visit”, “need to visit”, “want to go back”, “will go back”, “visit lisbon”, “advise you”

Finally, with the data loaded and the dictionary created, the sentiment analysis algorithm is executed by using the *R sentimentr* package (Rinker, 2021) that calculates text polarity sentiment.

This package contains functions that allows to perform sentiment analysis. In this work, the functions “sentiment” and “get sentences” are used to approximate the sentiment (polarity) of text by sentence. This consists of searching if a sentence contained any of the keywords related to each category in every post and every sentence in each post. If a match is found, the sentiment score, which is a polarity score, is computed where sentiments less than zero are negative, 0 are neutral, and greater

than zero positive polarity. To calculate the sentiment score, the words in each sentence are searched and compared to a dictionary of polarized words in which they are tagged with +1 and -1 denoting positive or negative score, respectively. This package takes into account valence shifters to deliver a better analysis of the content, with more reliable values of sentiment. Valence shifters are words that affect the polarity (e.g., “I do not like it”) and they can flip the meaning, increase or decrease the impact or overrule (e.g., “I like it but it is not worth it.”) a polarized word. The package also includes other parameters to boost performance. The functions description is available in Appendix II.

The final dataset contains the post identification, the sentence number, the calculated sentiment and an indicator of which construct it relates to. In case, there are sentences without any word, this means that a sentence does not include any contribution for the analysis. Hence, these cases are considered worthless and taken as missing values. For this reason, they are omitted from the dataset, following a method for handling missing data called listwise deletion. It consists of removing all data for a case that has one or more missing values (Elliott & Hawthorne, 2005).

The next step consists of analysing the dataset to understand if it is useful and representative. For this purpose, the dataset is imported to *Microsoft Excel*, in which an exploratory analysis of the data through descriptive statistics techniques is made. This analysis helps to interpretate the results while providing a useful strategy for summarising data in order to characterise the sample (Fisher & Marshall, 2009). Further, the Pearson correlation coefficient (r), is used to analyse the correlations between the constructs in order to understand which variables have associations with others. The correlation coefficient measures the tendency of two variables to change in value together either increasing or decreasing (Boslaugh & Andrew Watters, 2008). The correlation value ranges from -1, to 1, with 0 corresponding to no correlation, 1 to total positive correlation, and -1 to total negative correlation.

During the exploratory analysis, there may be posts where one or more constructs are not included. These cases are taken as missing values since the author has probably not considered a particular construct as being relevant to talk about. To deal with these missing values, a value of zero is considered for all cases, since, as the construct is not addressed, it is assumed that the author has no relevant opinion about it, or it is not clear enough to be reported.

To test the hypotheses, structural equation modelling is conducted in order to test the destination brand model with observed variables and image as a latent variable. Structural equation modelling is a powerful, multivariate technique to test and evaluate multivariate causal relationships (Fan et al., 2016). This technique is useful to examine situations in which there are final dependent variables, and those in which there are chains of influence, some of them being latent. For example, models in which a variable A influences variable B, which in turn affects variable C and at least one of them is latent (Streiner, 2005). AMOS Version 28 was used to estimate the model (IBM, 2022).

The proposed model is validated through model fitting indices based on Hu and Bentler (1999) thresholds. The ratio between the chi-square (χ^2) and the degrees of freedom (*df*) is below 3 with a non-significant *p*-value, the Comparative Fit Index (CFI) is above 0.90, the Root Mean Square Error of Approximation (RMSEA) is lower than 0.06 with a non-significant *p*-value and a 90% confidence level, the Standardized Root Mean Squared Residual (SRMR) is lower than 0.08, and the Goodness-of-Fit Index (GFI) and the Incremental Fit Index (IFI) are above 0.95 and 0.90, respectively.

In order to support the established hypotheses, the regression weight coefficients have to be significant, established by a *p*-value smaller than 0.05. A *p*-value is a statistical measure used to validate a hypothesis against observed data in which the smaller the value, higher the significance (Fisher, 1925). Also, t-test is calculated which is a statistical hypothesis test in which the test statistic follows a Student's t-distribution under the null hypothesis. If t-value is greater than +2 or less than – 2 the coefficient is acceptable, hence, the higher the t-value, the greater the confidence in the coefficient as a predictor (Cramér, 1999).

The standardized coefficient (std. coef.) represents whether there is a positive or negative correlation between the variables and, once there are only positive relationships hypothesized, the coefficients have to be positive to support the hypotheses. The correlations between the variables result from direct and indirect effects so it is important to analyse those effects to better understand the influence of each variable on the dependent variables.

CHAPTER 4

Results

4.1. Sample characterisation

The sample is composed by 100 blog posts, written by different bloggers. For the set of blogs being considered, the blog text size ranges between 306 and 5853 words, in which the average is 2431 words.

The dataset also includes the date of the post and the gender and nationality of the blogger, in order to characterise the sample, when available. For the blogs written by more than one person, the information of the author of the post was collected and the nationality is related to the country of birth, except in some cases, where only the current living country is available.

The sociodemographic characterisation of the sample is illustrated in Figure 4.1, comprising 58.6% males and 41.4% females from 16 different countries. Most of the bloggers are from France, Italy, Spain, United Kingdom and United States, corresponding to 67.8% of the total sample (i.e., including the not available biographic information), which is in good agreement with what was referred in Section 3.2.2, since most of the tourists visiting Portugal are from United Kingdom, Spain, France, Germany, and United States (Turismo de Portugal, 2022). The selected posts were written between 2014 and 2022, distributed along all quarters of the year.

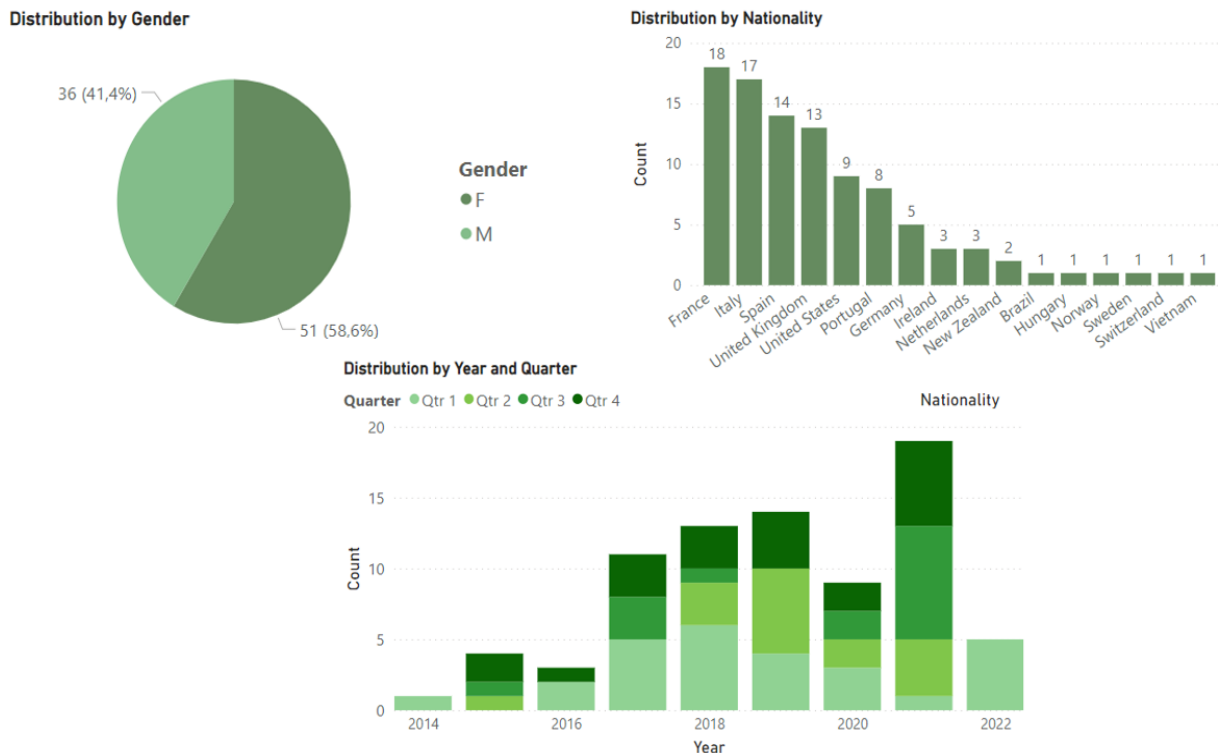


Figure 4.1: Sample characterisation

4.2. Constructs and sentiment analysis

From the sentiment analysis algorithm executed in *R*, 11862 sentences were identified and 8600 occurrences with information about the constructs were recorded. These occurrences can be related to the same sentence, which means that a sentence can contain opinions about more than one construct. For example, the sentence “*with seven hills that overlook the river tejo, lisbon has a stunning location and an amazing light that immediately captivates visitors*” refers to both the affective and the unique images. The percentual distribution of the constructs found in the sentences is presented in Figure 4.2. Blogs content are mainly focusing on the three components of image which in total make up 85.8% of posts. The cognitive image is the most addressed component with 53.1% of the occurrences and awareness is the less addressed with 3.2%, corresponding to 272 occurrences.

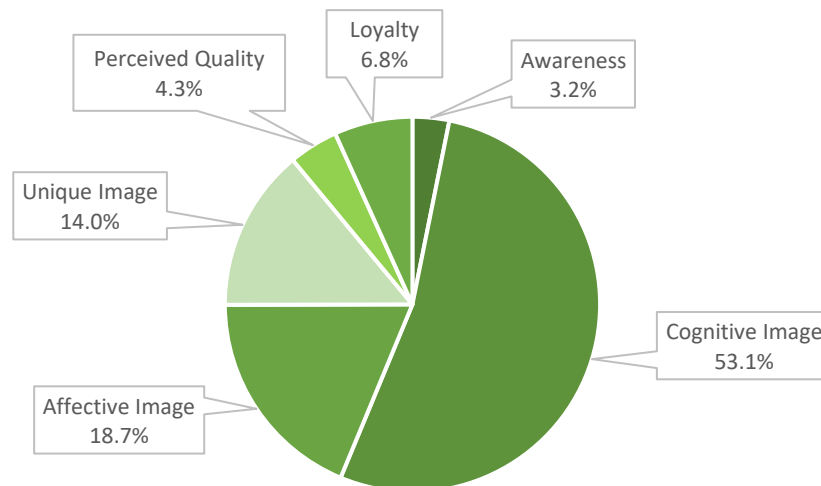


Figure 4.2: Distribution of components of Lisbon brand equity

This distribution shows that bloggers wrote mostly about the cognitive image, which comprises the own knowledge and beliefs that they achieved about the Lisbon image. Since cognitive image, affective image and unique image are components of destination brand image construct, it is possible to conclude that more than 85% of each post is related to the brand image of Lisbon. It is worth remembering that not addressing opinion about a construct does not mean that the opinion is negative, but it means that the blogger has probably not considered a particular construct as being relevant to talk about.

The opinions about each construct were measured through sentiment analysis. The average sentiment of each post is presented in Figure 4.3. It is possible to observe that the sentiment score is always positive with an average score of 0.18. The first quartile, i.e., the value under which 25% of data

points are found when they are arranged in increasing order, is 0.15 and that the mean is close to the median.

A summary of the descriptive statistics of sentiment value per construct is presented in Table 4.1, in which the sentiment refers to the values of sentiment per blog. Globally, all posts have a positive total sentiment which means that, on average, all blogs correspond to a positive opinion about Lisbon. The average sentiment per blog is 0.18 with 53% of them being above the average.

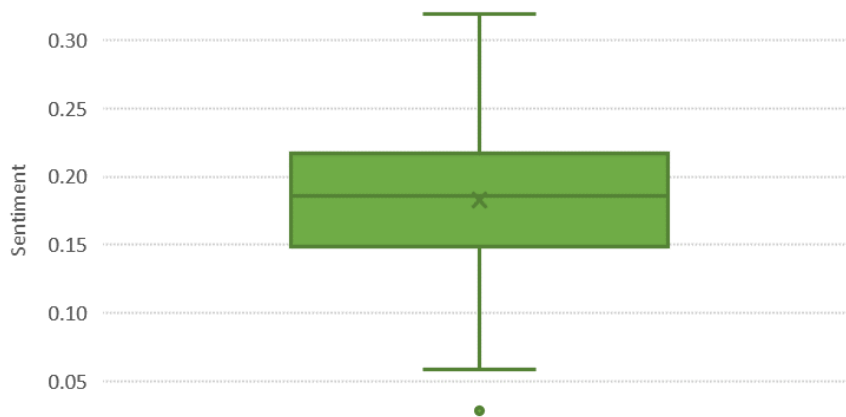


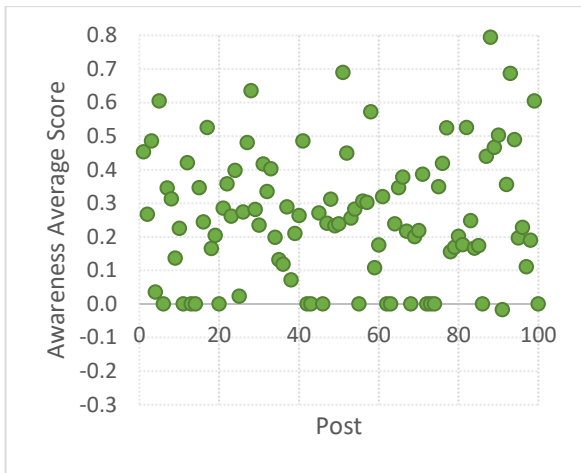
Figure 4.3: Total average sentiment per post

Table 4.1: Descriptive statistics

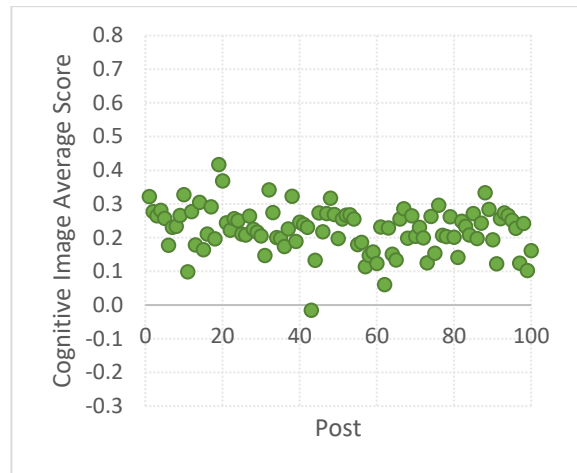
Construct	Minimum	Maximum	Mean	Standard Deviation	Skewness	Kurtosis
Awareness	-0.370	0.794	0.255	0.199	0.116	0.307
Cognitive image	-0.016	0.416	0.224	0.067	-0.405	1.347
Affective image	-0.031	0.727	0.315	0.111	-0.090	2.126
Unique image	-0.044	0.742	0.246	0.123	1.547	5.355
Perceived quality	-0.346	0.874	0.287	0.209	0.025	0.381
Loyalty	-0.203	0.726	0.239	0.167	0.137	0.087
Sentiment	0.029	0.319	0.183	0.052	-0.185	0.510

Regarding to the constructs, the statistical analysis shows that, on average, affective image is the most positive addressed component. This means that Lisbon visitors get positive emotions and feelings about the image of Lisbon. Furthermore, all constructs have a positive average, being above 0.224. Perceived quality and destination awareness are the constructs with both the best (regarding its maximum value) and worst opinions (regarding its minimum value), while, having the major dispersion, Figure 4.4. All constructs have at least one negative opinion.

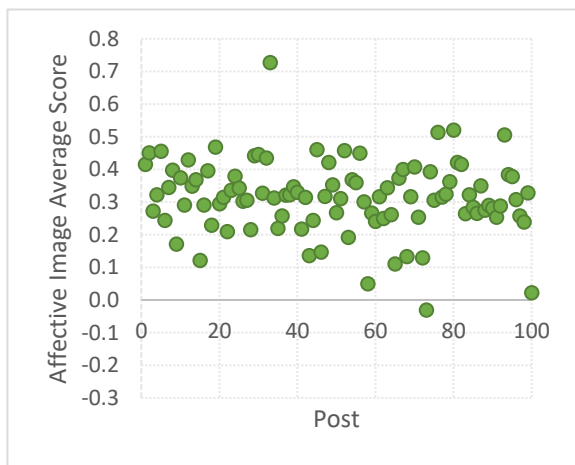
Skewness and kurtosis indexes were used to identify the normality of the data. According to Yadav & Pathak (2017), the deviation of data from normality was not severe since the values of skewness and kurtosis index were below 3 and 10, respectively.



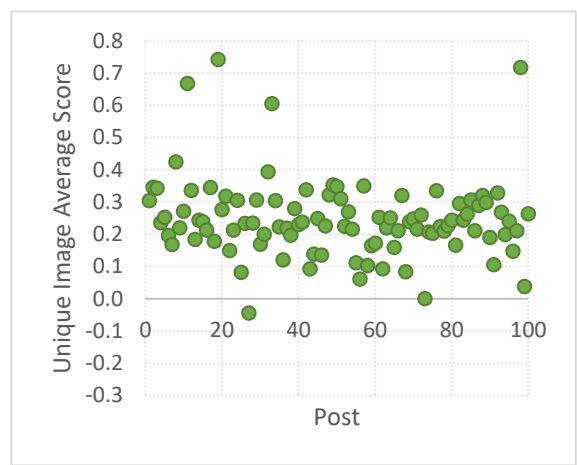
a) Awareness



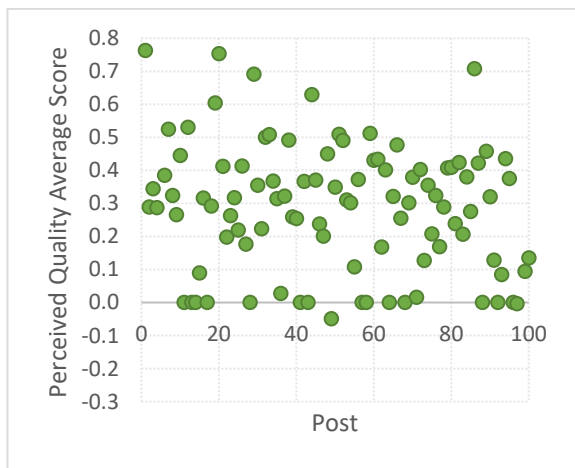
b) Cognitive image



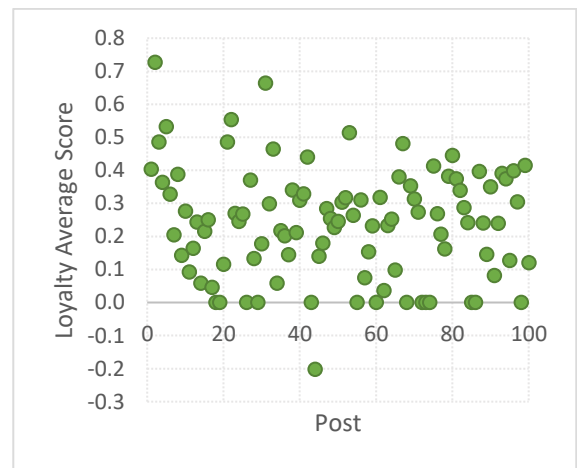
c) Affective image



d) Unique image



e) Perceived quality



f) Loyalty

Figure 4.4: Average score per post

To understand how the different constructs influence the intention of revisiting or recommending Lisbon, it is important to analyse what makes bloggers loyal to Lisbon. Looking at the correlation values (r), as presented in Table 4.2, it is observed that all variables have positive correlation, with the exception of the relationship between perceived quality and loyalty which is negative but not significant, i.e., these two variables are not correlated.

Table 4.2: Correlation values

	Awareness	Cognitive image	Affective image	Unique image	Perceived quality	Loyalty
Awareness	1.000	0.299	0.272	0.136 (ns)	-0.127 (ns)	0.439
Cognitive Image	0.299	1.000	0.471	0.429	0.340	0.238
Affective Image	0.272	0.471	1.000	0.373	0.272	0.354
Unique Image	0.136 (ns)	0.429	0.373	1.000	0.292	0.026 (ns)
Perceived Quality	-0.127 (ns)	0.340	0.272	0.292	1.000	-0.081 (ns)
Loyalty	0.439	0.238	0.354	0.026 (ns)	-0.081 (ns)	1.000

(ns) - correlation is not significant at the 0.05 level.

On one hand, the relationship between cognitive image and affective image is the strongest one ($r = 0.471, p < 0.05$), followed by the relationship between awareness and loyalty ($r = 0.439, p < 0.05$). On the other hand, the relationship between cognitive image and loyalty ($r = 0.238, p < 0.05$), is the weakest correlation, which means that the variables increase simultaneously, but with a lower effect than in other relationships. Note that the components of destination brand image are highly related.

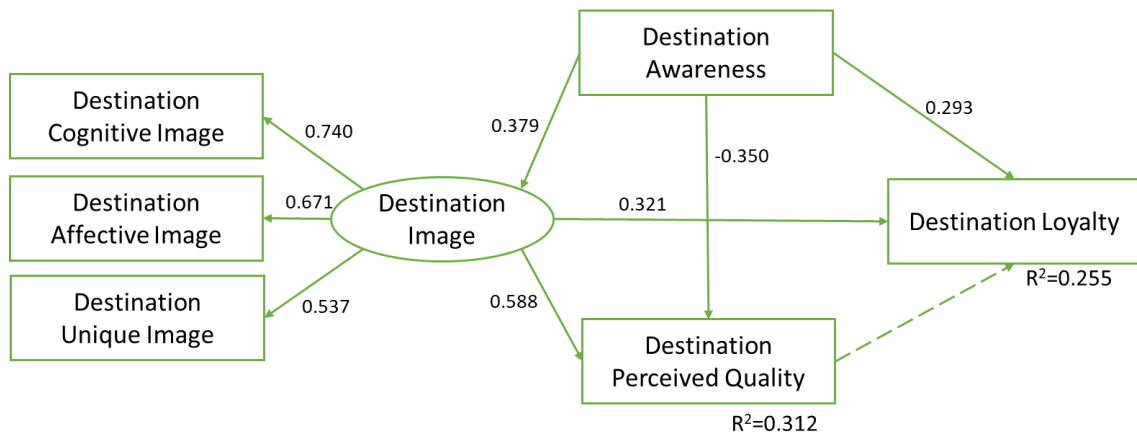
Based on the results from the statistical analysis, it is concluded that the set of data being considered is reliable to estimate a structural equation model to test the suggested hypotheses.

4.3. Structural equation modelling and hypothesis testing

To test the suggested hypotheses, structural equation modelling is conducted by incorporating destination brand image as a latent variable. All the other components were included in the model as observed variables.

Overall, the model has acceptable fit indices ($\chi^2 = 8.481, df = 6, p\text{-value} = 0.205$; CFI = 0.976; RMSEA = 0.065, $p\text{-value} = 0.341$, 90%CI = [0.000; 0.156]; SRMR = 0.0493; GFI = 0.973; IFI = 0.978) and it is depicted in Figure 4.5 showing all standardized coefficients and the squared multiple correlations (R^2) of the endogenous variables.

Destination Brand Equity Model*



*Significant standardized regression coefficients are placed close to arrows; non-significant relationships are marked by dashed arrows.

Figure 4.5: Path diagram for the destination brand equity model

Results in Figure 4.5 show that the value of R^2 is 0.255 and 0.312 for loyalty and perceived quality. These results mean that awareness and image as predictors explain a moderate proportion of variance in perceived quality, and that loyalty is explained by 25.5% of awareness, image and perceived quality.

The results of the standardised and unstandardised coefficient estimates of the hypothesised direct effects between constructs, as presented in Table 4.3, provide substantive support for the research hypotheses.

Table 4.3: Coefficient estimates of the model structural relationships

Relationship	Unstandardized coefficient	Standardized coefficient	Standard error	t-test	p-value
Awareness → Image	0.094	0.379	0.029	3.217	0.001
Awareness → Perceived Quality	-0.368	-0.350	0.109	-3.375	0.000
Image → Perceived Quality	2.497	0.588	0.602	4.150	0.000
Awareness → Loyalty	0.245	0.293	0.095	2.591	0.010
Perceived Quality → Loyalty	-0.152	-0.191	0.095	-1.595	0.111
Image → Loyalty	1.087	0.321	0.533	2.039	0.041
Image → Affective Image	1.507	0.671	0.292	5.160	0.000
Image → Unique Image	1.339	0.537	0.305	4.393	0.000
Image → Cognitive Image	1.000	0.740	-	-	0.000

The current study supports hypotheses H1a, H1b and H1c, corresponding to the influence of the destination brand image on the components of brand image. It was found that cognitive image is significantly related to higher destination brand image (std. coef. = 0.740, $p < 0.001$), meaning that the destination brand image strongly affects the cognitive image of Lisbon. The other components, affective image and unique image are still highly related to destination brand image. The relationship between awareness and image supports H2, which means that the higher the destination brand awareness, the higher the image perception of Lisbon (std. coef. = 0.379, $p = 0.001$). Also, the relationship between image and perceived quality represents the strongest influence (std. coef. = 0.588, $p < 0.001$), meaning that H4 is supported. However, the hypothesis H3 is not supported since the effect of awareness on perceived quality is negative (std. coef. = -0.350, $p < 0.001$). Regarding to the influence in destination brand loyalty, higher awareness is related to higher destination brand loyalty, supporting H5, although this influence is the weakest (std. coef. = 0.293, $p = 0.01$). Further, higher image (std. coef. = 0.321, $p < 0.05$) is related to higher destination brand loyalty, supporting H6. However, it is not possible to conclude that perceived quality influences destination brand loyalty since the relationship is not significant (std. coef. = -0.191, $p > 0.05$), meaning that H7 is not supported.

In addition to the direct effects, in this model there are indirect effects that affect the variables. For example, awareness affects perceived quality through image. Table 4.4 presents the standardised estimates of the indirect and total effects.

Table 4.4: Standardized coefficients for indirect and total effects

From	To	Cognitive image	Affective image	Unique image	Image	Perceived quality	Loyalty
Awareness	Indirect	0.281*	0.254*	0.204*	-	0.223*	0.146**
	Total	0.281*	0.254*	0.204*	0.379*	-0.127	0.439**
Image	Indirect	-	-	-	-	-	-0.112
	Total	0.740*	0.671*	0.537*	-	0.588*	0.209
Perceived Quality (PQ)	Indirect	-	-	-	-	-	-
	Total	-	-	-	-	-	-0.191

* $p < 0.01$; ** $p < 0.05$

Results show that (1) awareness influences all the components of destination brand image; (2) image works as a moderator of the relationship between awareness and perceived quality since the direct relationship is negative, but the indirect through image is positive, resulting in a non-significant relationship. A non-significant relationship represents that there is no evidence of influence, which means that the relation is, in the total effect, not negative; (3) Regarding to the indirect and total

effects in loyalty, awareness has a positive indirect effect that contributes to a strong positive total effect and image has not influence on loyalty.

4.4. Discussion

The formulated relationship hypotheses between cognitive, affective, and unique images with destination brand image (H1a, H1b, and H1c) were all verified. These results are in agreement with the ones obtained by Huete-Alcocer et al. (2019) and Qu et al. (2011), who also found that the three components are influenced by brand image. However, Huete-Alcocer et al. (2019) have evidenced that unique image has a stronger effect on overall image than cognitive image. Thus, this study highlights the idea that in an emerging destination all the image components contribute in different ways to image and branding destination (Qu et al. 2011).

Regarding to the impact of brand awareness on brand image and perceived quality (H2 and H3), the results provide support just for H2 referring to brand image. This positive relationship is in good agreement with the studies by Kaushal et al., (2019), Kusumaningrum (2021), Kim & Lee (2018), Herrero et al. (2017), Gartner & Ruzzier (2011) and Chen & Myagmarsuren (2010). The hypothesis regarding the relationship between awareness and perceived quality was not supported since the effect of awareness on perceived quality is negative. This agrees with Kim & Lee (2018) and Chen & Myagmarsuren (2010) that concluded that a tourist's awareness which influences the positive perceptions of the destination, leads to higher expectations and influences actual experience of the destination's perceived service quality. To sum up, high expectations require high levels of quality to avoid disappointment. However, these results are not consistent with other studies, for instance, Herrero et al. (2017), Kaushal et al. (2019) and Kladou & Kehagias (2014) that found a positive relationship between the constructs. According to Herrero et al. (2017), the better the image and awareness attributed by tourists to the destination, the better its perceived quality will be.

Regarding to the impact of brand image on perceived quality, results showed a positive relationship between these two constructs (H4). This agrees with Kusumaningrum (2021), Gartner & Ruzzier (2011) and Chen & Myagmarsuren (2010). Gartner & Ruzzier (2011) concluded that image and quality combine as the central organizing dimensions that should be used to create brand awareness and build loyalty. Hence, image aspect is a direct determinant of tourists' perceived quality (Chen & Myagmarsuren, 2010). Kaushal et al. (2019) also studied these relationships but have found a non-significant relationship. Further, Kim & Lee (2018) studied the reverse relationship, i.e., the impact of perceived quality on brand image and the obtained results have confirmed this relationship.

The last proposed hypotheses were related to the impact of awareness, image and perceived quality on destination brand loyalty (H5, H6 and H7). The findings support just two of the proposed

relationships (H5 and H6). Firstly, the relationship between awareness and loyalty is supported, agreeing with Im et al. (2012) that concluded that destination preferences are strongly associated with the perceived likelihood of intention to visit the place, thus, brand awareness is a pivotal element in creating value in the mind of travellers. The support of H6 agrees with Kaushal et al. (2019), Kim & Lee (2018), Pike and Bianchi (2016), Bianchi & Pike (2011). However, Chen & Myagmarsuren (2010) found a non-significant relationship between brand image and brand loyalty. This happens due to the fact that perceived service quality with an expectation and service performance are both specific beliefs of a future experience, which may mediate the impact of the image (Chen & Myagmarsuren, 2010). H7 is not supported, regarding to the impact of perceived quality, since the relationship was not found to be significant. These results are consistent with Kaushal et al. (2019), Pike & Bianchi (2016), Bianchi & Pike (2011) and Chen & Myagmarsuren (2010) studies that also rejected this hypothesis. They suggested that travellers assume that having a developed tourist destination with high quality facilities does not have impact on the intentions to recommend or revisit. This study also concluded that an increase in perceived quality leads to a customer's overall satisfaction, but perceived quality does not influence loyalty directly. However, Kladou & Kehagias (2014) and Kim & Lee (2018) showed that brand quality has a positive impact on brand loyalty.

Regarding to the indirect and total effects in loyalty, awareness reveals a positive indirect effect that contributes to a strong positive total effect and image. The results obtained by Herrero et al. (2017) show that more recognizable or renowned destinations will be perceived as being of higher quality, which intuitively shows a signal of destination brand quality. This way, awareness indirectly influences tourist loyalty towards the destination. Moreover, Herrero et al. (2017) also considered destination image as the main determinant of perceived quality of the tourist destination, exerting an indirect effect on tourist loyalty.

CHAPTER 5

Conclusions

This dissertation aimed at proposing a destination brand equity measurement approach based on text mining through sentiment analysis of user-generated content, available in blogs of tourists who visit Lisbon, in Portugal. The purpose of destination branding should be to stimulate intent to visit and revisit. Further, enhancing brand equity of a destination has been recognized as an important factor for creating competitive market advantages and marketing strategies of differentiation.

Lisbon was the chosen destination, since the tourism sector has been increasing through the last years and registered, from 2010 to 2019, an average annual growth rate of 7.2% in overnight stays (Statista, 2022). Furthermore, Portugal is one of the most visited countries in the Europe with Lisbon being the most visited city in Portugal, which registered about six million international tourist guests in 2019 (Turismo de Portugal, 2022). Enhancing the effect of brand equity in a tourism destination represents an important factor for creating competitive market advantages and marketing strategies of differentiation (Horng et al., 2012) so, the results in this dissertation are aimed at contributing for the development of the tourism sector of Lisbon.

The literature review was performed through a systematic literature review that helped to organise the existing literature to find possible gaps in the research. The PRISMA protocol was conducted in order to perform a methodical approach for finding the best articles to answer the research questions. Fourteen articles obtained from Web of Science and Scopus searches were chosen and evaluated using specific criteria to rate its relevance for the purpose of this dissertation. The selected articles allowed to define the general concepts of destination brand equity measurement, to identify the models being tested and which ones had success. Also, the research methods used to measure destination brand equity were identified where it was found gaps in the existing research. Finally, the evaluation criteria were applied to the articles and it was concluded that every article had a score better than 5 out of 10 and that the best criteria were the conclusions and contributions of each article.

The SLR had several contributions for this dissertation. It allowed to identify a gap in the knowledge of brand equity, namely in the brand equity of tourist destinations, that was addressed in this dissertation. This gap is related to the research method used to measure destination brand equity, which is surveys to a single sample at a single moment in time. Also, it allowed to define different dimensions that were tested in order to find an acceptable destination brand equity model. Overall, the main contribution of the SLR was providing a methodology for the literature review, allowing to analyse all the articles in order to answer the same research questions, being the core strategy to identify gaps and limitations of existing research.

Throughout the SLR, it was possible to define the general concepts of destination brand equity measurement and to identify the conceptual models tested. The destination brand image was defined through three components, namely, cognitive, affective, and unique image. This study includes unique image. Unique image, measured by features which sets a place apart from all the other, was recently started being studied revealing being strongly influenced by overall image, becoming a key differentiator of the destination proposition (Qu et al. 2011). Through the identified models and the defined concepts, the conceptual model was theoretically built.

There were established relationships among the destination brand equity constructs. Destination brand image influences the cognitive, affective and unique constructs of image. Relationships between brand awareness, brand image and perceived quality were established as well as these constructs with brand loyalty. Hypotheses were tested in order to examine Lisbon brand equity and to identify the most important constructs that influence the destination loyalty to Lisbon. This will help to develop marketing activities that can increase the value of Lisbon's assets. Results from a sample of 100 blog posts provided empirical support to most hypotheses formulated.

To test the model, data was needed, specifically blog posts. The type of blogs considered was on travelling experiences, based on the personal opinion of the author. This includes guides, tips and recommendations about what to do while visiting Lisbon. The sociodemographic info about the blogs was extracted to characterise the sample, encompassing 100 blogs, from 2014 to 2022. Overall, 58.6% male bloggers were identified, from 16 different countries, which in most of them were from France, Italy, Spain, United Kingdom and United States. The data was extracted in the original language and translated to English.

The algorithm was executed, and 8600 occurrences were identified through the created data dictionary as containing any info about a construct. More than 80% occurrences were related with image components, especially with cognitive image with more than 4000 occurrences. Awareness was the least addressed construct what does not mean that the opinion is negative, but the blogger has probably not considered this construct as being relevant. Overall, the sentiment was positive, with an average of 0.183 in which 53% posts were above the posts' sentiment average. The correlations between cognitive image and affective image were the strongest.

Regarding to the hypothesis testing, the relationships between each component of image and brand image were verified, in which cognitive image was the most significantly related to higher destination brand image, with a standardized coefficient of 0.740. This result agrees with Huete-Alcocer et al. (2019) and Qu et al. (2011) who found that the three components are influenced by brand image, but they found that unique image is the most significantly related to brand image. Through the indirect effects, it was possible to observe that awareness influences all the components of destination brand image. The relationships between awareness and both image and loyalty were

verified, however, between awareness and perceived quality was not. According to Kim & Lee (2018) and Chen & Myagmarsuren (2010), a tourist's awareness which influences the positive perceptions of the destination leads to higher expectations and influences the actual experience of quality and for that reason, the relationship is negative. Hence, a positive opinion about the awareness influences negatively the perceived quality of Lisbon. To sum up, high expectations require high levels of quality to avoid disappointment. However, these results are not consistent with other studies, for instance, Herrero et al. (2017), Kaushal et al. (2019) and Kladou & Kehagias (2014) that found a positive relationship between the constructs.

Further, the strongest influence is verified between image and perceived quality (with a standardized coefficient of 0.588). This means that the higher the image perception of Lisbon, the higher the perceived quality. In addition, image works as a moderator of the relationship between awareness and perceived quality, contributing to mitigate the negative direct effect to a null total effect as the total effect of awareness on perceived quality is not significant. Chen & Myagmarsuren (2010) concluded that image aspect is a direct determinant of tourists' perceived quality, in which they, Kusumaningrum (2021) and Gartner & Ruzzier (2011) also confirmed this relationship. However, Kaushal et al. (2019) also studied these relationships but found a non-significant relationship. Moreover, both awareness and image influence destination brand loyalty, however perceived quality revealed a non-significant relationship with loyalty. This means that it was not possible to conclude any influence between perceived quality and loyalty. These findings are consistent with Kaushal et al. (2019), Pike & Bianchi (2016), Bianchi & Pike (2011) and Chen & Myagmarsuren (2010) studies that also rejected this hypothesis. They suggested that travellers assume that having a developed tourist destination with high quality facilities does not have impact on the intentions to recommend or revisit. However, Kladou & Kehagias (2014) and Kim & Lee (2018) showed that brand quality has a positive impact on brand loyalty. Still, the results about the indirect and total effects revealed positive indirect effects of awareness on positive total effects of loyalty and image showed a negative indirect effect on loyalty, due to perceived quality, making the relationship not significant.

This dissertation enables to draw both theoretical and managerial implications for the tourism of Lisbon.

Firstly, this dissertation contributes for the destination brand equity literature in the sense of presenting a brand equity model with a decomposed brand image construct into its cognitive, affective and unique components, highlighting that there are a small number of studies abording the unique image concept. It is also worthwhile to consider that little research exists on brand equity in tourism's context in the city of Lisbon. Further, the proposed model allows to understand the impact of awareness, image and perceived quality on loyalty, which means that knowledge about what makes visitors become loyal and revisit Lisbon can be taken from this model. Further, collecting data in

different languages allowed to test the brand equity model on a global sample, not being restricted to the English language. Moreover, a new method of research is tested, introducing text mining and sentiment analysis in the destination brand equity evaluation. The approach proposed in this dissertation allowed to extract a broad perspective of the tourists through the real experience report and the personal opinion of the author. Once Jalilvand et al. (2012) empirical results confirmed that electronic WoM positively affects the destination image, tourist attitude, and travel intention, this method takes advantage of the emerging online communications, in which text mining and sentiment analysis has been frequently applied into analysing electronic WoM (Kim et al., 2018).

Secondly, the proposed model can help managers in making marketing decisions about what to do to ensure that visitors become loyal to recommend and revisit Lisbon. It was found that awareness and image are the most important constructs to make the consumers loyal to Lisbon. Further, the tourists of Lisbon wrote mostly about the cognitive image, but the affective image was the most positive addressed component. It is possible to conclude that the investments on facilities and attractions, as well as the hospitality and convenience make the tourists retain positive emotions and feelings about the image of Lisbon. Overall, all the constructs being studied had a positive average, which means that the overall opinion of the city of Lisbon is positive. However, perceived quality had the worst opinion and the major dispersion of opinions. Due to the non-significant effect of perceived quality on loyalty, significant effort should be directed to this construct to improve the performance at this level.

This study also has some limitations: (1) the data dictionary used to extract the sentiment score about the constructs was shaped to the city of Lisbon, which means that cannot be integrally transposed to other cities. In future research in another cities, this data dictionary has to be modified especially in the unique image component since it refers to the unique aspects of Lisbon; (2) Due to the fact that the keywords used to identify the constructs were obtain from manual reading and adapted from the questions of the survey used by Boo et al. (2009), there are a large number of missing values on the constructs identification because each author has his/her own way to express its opinion. In future research, effort should be done in the way of including synonyms in the data dictionary to provide a more complete analysis. Further, future research should be focused on the study of other cities through text mining and sentiment analysis to provide insights on how the brand equity knowledge about a destination can differ from the results obtained from surveys.

References

- Aaker, D. (1991). *Managing Brand Equity*. New York, USA: Free Press.
- Aaker, D. (1996). *Building Strong Brands*. New York, USA: Free Press.
- Abrantes, C. (2020). Os modelos preditivos do sucesso de candidaturas a fundos europeus: o papel da manipulação de resultados [Dissertação de mestrado] (in portuguese). Iscte - Instituto Universitário de Lisboa.
- Ah-Hwee Tan. (1999). Text Mining: The state of the art and the challenges. *Kent Ridge Digital Labs*. <https://www.researchgate.net/publication/2471634>
- Argamon, S., Bloom, K., Esuli, A., & Sebastiani, F. (2007). Automatically Determining Attitude Type and Force for Sentiment Analysis. *3rd Language and Technology Conference*, 369–373. Poznan, Poland. https://doi.org/10.1007/978-3-642-04235-5_19
- Baskerville, R., Baiyere, A., Gregor, S., Hevner, A. & Rossi, M. (2018). Design Science Research Contributions: Finding a Balance between Artifact and Theory. *Journal of the Association for Information Systems*, 19(5). <http://dx.doi.org/10.17705/1jais.00495>
- Bianchi, C., & Pike, S. (2011). Antecedents of Destination Brand Loyalty for a Long-Haul Market: Australia's Destination Loyalty Among Chilean Travelers. *Journal of Travel and Tourism Marketing*, 28(7), 736–750. <https://doi.org/10.1080/10548408.2011.611742>
- Bianchi, C., Pike, S., & Lings, I. (2014). Investigating attitudes towards three South American destinations in an emerging long haul market using a model of consumer-based brand equity (CBBE). *Tourism Management*, 42, 215–223. <https://doi.org/10.1016/j.tourman.2013.11.014>
- Boo, S., Busser, J., & Baloglu, S. (2009). A model of customer-based brand equity and its application to multiple destinations. *Tourism Management*, 30(2), 219–231. <https://doi.org/10.1016/j.tourman.2008.06.003>
- Boslaugh, S., & Andrew Watters, P. (2008). *Statistics in a Nutshell: A Desktop Quick Reference*. Sebastopol, CA: O'Reilly Media.
- Cambria, E., Schuller, B., Xia, Y., & Havasi, C. (2013). New avenues in opinion mining and sentiment analysis. *IEEE Intelligent Systems*, 28(2), 15–21. <https://doi.org/10.1109/MIS.2013.30>
- Chen, C. F., & Myagmarsuren, O. (2010). Exploring relationships between mongolian destination brand equity, satisfaction and destination loyalty. *Tourism Economics*, 16(4), 981–994. <https://doi.org/10.5367/te.2010.0004>

- Chen, Y. F., & Law, C. H. R. (2016). A Review of Research on Electronic Word-of-Mouth in Hospitality and Tourism Management. *International Journal of Hospitality and Tourism Administration*, 17(4), 347–372. <https://doi.org/10.1080/15256480.2016.1226150>
- Chi, H. K., Huang, K. C., & Nguyen, H. M. (2020). Elements of destination brand equity and destination familiarity regarding travel intention. *Journal of Retailing and Consumer Services*, 52. <https://doi.org/10.1016/j.jretconser.2018.12.012>
- Confente, I. (2015). Twenty-five years of word-of-mouth studies: a critical review of tourism research. *International Journal of Tourism Research*, 17, 613–624. <http://dx.doi.org/10.1002/jtr.2029>
- Cramér, H. (1999). *Mathematical Methods of Statistics*. Princeton, USA: Princeton Univ. Press
- Dias, F., & Cardoso, L. (2017). How can brand equity for tourism destinations be used to preview tourists' destination choice? An overview from the top of Tower of Babel. *Tourism & Management Studies*, 13(2), 13–23. <https://doi.org/10.18089/tms.2017.13202>
- Elliott, P., & Hawthorne, G. (2005). Imputing missing repeated measures data: how should we proceed?. *Australian & New Zealand Journal of Psychiatry*, 39(7).
- Fan, Y., Chen, J., Shirkey, G., John, R., Wu, S. R., Park, H., & Shao, C. (2016). Applications of structural equation modeling (SEM) in ecological studies: an updated review. *Ecological Processes*, 5(1). <https://doi.org/10.1186/s13717-016-0063-3>
- Ferns, B. H., & Walls, A. (2012). Enduring travel involvement, destination brand equity, and travelers' visit intentions: A structural model analysis. *Journal of Destination Marketing and Management*, 1(1–2), 27–35. <https://doi.org/10.1016/j.jdmm.2012.07.002>
- Fisher, M. J., & Marshall, A. P. (2009). Understanding descriptive statistics. *Australian Critical Care*, 22(2), 93–97. <https://doi.org/10.1016/j.aucc.2008.11.003>
- Fisher, R. A. (1925). *Statistical methods for research workers* (11th ed.). Edinburgh, UK: Oliver and Boyd.
- Gartner, W. C., & Ruzzier, M. K. (2011). Tourism destination brand equity dimensions: Renewal versus repeat market. *Journal of Travel Research*, 50(5), 471–481. <https://doi.org/10.1177/0047287510379157>
- Gómez, M., Lopez, C., & Molina, A. (2015). A model of tourism destination brand equity: The case of wine tourism destinations in Spain. *Tourism Management*, 51, 210–222. <https://doi.org/10.1016/j.tourman.2015.05.019>
- Gopalakrishnan, S., & Ganeshkumar, P. (2013). Systematic reviews and meta-analysis: Understanding the best evidence in primary healthcare. *Journal of Family Medicine and Primary Care*, 2(1), 9. <https://doi.org/10.4103/2249-4863.109934>
- Guerreiro, J., & Rita, P. (2020). How to predict explicit recommendations in online reviews using text mining and sentiment analysis. *Journal of Hospitality and Tourism Management*, 43, 269–272. <https://doi.org/10.1016/j.jhtm.2019.07.001>

- Han, H., & Hwang, J. (2018). Investigating healthcare hotel travelers' overall image formation: Impact of cognition, affect, and conation. *Tourism and Hospitality Research*, 18(3), 346–356. <https://doi.org/10.1177/1467358416670936>
- Hassani, H., Beneki, C., Unger, S., Mazinani, M. T., & Yeganegi, M. R. (2020). Text mining in big data analytics. *Big Data and Cognitive Computing*, 4(1), 1–34. <https://doi.org/10.3390/bdcc4010001>
- Hearst, M. A. (1999). Untangling Text Data Mining. *37th Annual Meeting of the Association for Computational Linguistics*. Maryland, USA. <http://dx.doi.org/10.3115/1034678.1034679>
- Helms, R., Giovacchini, E., Teigland, R., & Kohler, T. (2010). A Design Research Approach to Developing User Innovation Workshops in Second Life. *Journal For Virtual Worlds Research*, 3(1). <https://doi.org/10.4101/jvwr.v3i1.819>
- Herrero, Á., San Martín, H., Garcia de los Salmones, M. del M., & Collado, J. (2017). Examining the hierarchy of destination brands and the chain of effects between brand equity dimensions. *Journal of Destination Marketing and Management*, 6(4), 353–362. <https://doi.org/10.1016/j.jdmm.2016.05.001>
- Hevner, A., March, S., Park, J., & Ram, S. (2004) Design Science in Information Systems Research. *MIS Quarterly*, 28(1), 75–105. <https://doi.org/10.2307/25148625>
- Horng, J. S., Liu, C. H., Chou, H. Y., & Tsai, C. Y. (2012). Understanding the impact of culinary brand equity and destination familiarity on travel intentions. *Tourism Management*, 33(4), 815–824. <https://doi.org/10.1016/j.tourman.2011.09.004>
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Huete-Alcocer, N., Martinez-Ruiz, M. P., López-Ruiz, V. R., & Izquiedo-Yusta, A. (2019). Archeological tourist destination image formation: Influence of information sources on the cognitive, affective and unique image. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2019.02382>
- IBM. (2022, April). *Downloading IBM SPSS Amos 28* <https://www.ibm.com/support/pages/downloading-ibm-spss-amos-28>
- Im, H. H., Kim, S. S., Elliot, S., & Han, H. (2012). Conceptualizing Destination Brand Equity Dimensions from a Consumer-Based Brand Equity Perspective. *Journal of Travel and Tourism Marketing*, 29(4), 385–403. <https://doi.org/10.1080/10548408.2012.674884>
- Irawan, H., Akmalia, G., & Masrury, R. A. (2019). Mining tourist's perception toward Indonesia tourism destination using sentiment analysis and topic modelling. *4th International Conference on Cloud Computing and Internet of Things*, 7–12. <https://doi.org/10.1145/3361821.3361829>

- Jalilvand, M.R., Samiei, N., Dini, B., & Manzari, P.Y. (2012). Examining the structural relationships of electronic word of mouth, destination image, tourist attitude toward destination and travel intention: an integrated approach. *Journal of Destination Marketing and Management*, 1, 134–143. <https://doi.org/10.1016/j.jdmm.2012.10.001>
- Kaushal, V., Sharma, S., & Reddy, G. M. (2019). A structural analysis of destination brand equity in mountainous tourism destination in northern India. *Tourism and Hospitality Research*, 19(4), 452–464. <https://doi.org/10.1177/1467358418781442>
- Keller, K. L. (1993) Conceptualizing, Measuring, and Managing Customer-Based Brand Equity. *Journal of Marketing*, 57, 1–22. <http://dx.doi.org/10.2307/1252054>
- Kim, H. B., & Kim, W. G. (2005). The relationship between brand equity and firms' performance in luxury hotels and chain restaurants. *Tourism Management*, 26(4), 549–560. <https://doi.org/10.1016/j.tourman.2004.03.010>
- Kim, H. K., & Lee, T. J. (2018). Brand equity of a tourist destination. *Sustainability*, 10(2), 1–21. <https://doi.org/10.3390/su10020431>
- Kim, S. S., Schuckert, M. R., Im, H. H., & Elliot, S. (2017). An interregional extension of destination brand equity: From Hong Kong to Europe. *Journal of Vacation Marketing*, 23(4), 277–294. <https://doi.org/10.1177/1356766716672278>
- Kim, Y., Kang, M., & Jeong, S. R. (2018). Text mining and sentiment analysis for predicting box office success. *KSII Transactions on Internet and Information Systems*, 12(8), 4090–4102. <https://doi.org/10.3837/tiis.2018.08.030>
- Kladou, S., & Kehagias, J. (2014). Assessing destination brand equity: An integrated approach. *Journal of Destination Marketing and Management*, 3(1), 2–10. <https://doi.org/10.1016/j.jdmm.2013.11.002>
- Konecnik, M., & Gartner, W. C. (2007). Customer-based brand equity for a destination. *Annals of Tourism Research*, 34(2), 400–421. <https://doi.org/10.1016/j.annals.2006.10.005>
- Kusumaningrum, S. D. (2021). Destination Brand Equity: A Perspective of Generation Z on A World Heritage Site in Indonesia. *Journal of Asian Finance, Economics and Business*, 8(2), 1071–1078. <https://doi.org/10.13106/jafeb.2021.vol8.no2.1071>
- Li, J., Fong, S., Zhuang, Y., & Khoury, R. (2014). Hierarchical classification in text mining for sentiment analysis. *International Conference on Soft Computing and Machine Intelligence*, 46–51. New Delhi, India. <https://doi.org/10.1109/ISCM.2014.37>
- Lin, C. H., & Kuo, B. Z. L. (2018). The moderating effects of travel arrangement types on tourists' formation of Taiwan's unique image. *Tourism Management*, 66, 233–243. <https://doi.org/10.1016/j.tourman.2017.12.001>
- Marques, C., Vinhas da Silva, R., & Antova, S. (2021). Image, satisfaction, destination and product post-visit behaviours: How do they relate in emerging destinations? *Tourism Management*, 85. <https://doi.org/10.1016/j.tourman.2021.104293>

- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *BMJ*, *339*(7716), 332–336. <https://doi.org/10.1136/bmj.b2535>
- Mukhtar, N., Khan, M. A., & Chiragh, N. (2018). Lexicon-based approach outperforms Supervised Machine Learning approach for Urdu Sentiment Analysis in multiple domains. *Telematics and Informatics*, *35*(8), 2173–2183. <https://doi.org/10.1016/j.tele.2018.08.003>
- Pike, S., & Bianchi, C. (2016). Destination Brand Equity for Australia: Testing a Model of CBBE in Short-Haul and Long-Haul Markets. *Journal of Hospitality and Tourism Research*, *40*(1), 114–134. <https://doi.org/10.1177/1096348013491604>
- Pike, S., Bianchi, C., Kerr, G., & Patti, C. (2010). Consumer-based brand equity for Australia as a long-haul tourism destination in an emerging market. *International Marketing Review*, *27*(4), 434–449. <https://doi.org/10.1108/02651331011058590>
- Poch, R., & Martin, B. (2015). Effects of intrinsic and extrinsic motivation on user-generated content. *Journal of Strategic Marketing*, *23*(4), 305–317. <https://doi.org/10.1080/0965254X.2014.926966>
- Qu, H., Kim, L. H., & Im, H. H. (2011). A model of destination branding: Integrating the concepts of the branding and destination image. *Tourism Management*, *32*(3), 465–476. <https://doi.org/10.1016/j.tourman.2010.03.014>
- Rinker, T. (2021). *Package “sentimentr”: Calculate Text Polarity Sentiment*. <https://cran.r-project.org/web/packages/sentimentr/sentimentr.pdf>
- RStudio. (2022, April). *RStudio*. <https://www.rstudio.com/>
- San Martín, H., Herrero, A., & García de los Salmones, M. del M. (2019). An integrative model of destination brand equity and tourist satisfaction. *Current Issues in Tourism*, *22*(16), 1992–2013. <https://doi.org/10.1080/13683500.2018.1428286>
- Sánchez-Rebull, M. V., Rudchenko, V., & Martín, J. C. (2018). The antecedents and consequences of customer satisfaction in tourism: A systematic literature review. *Tourism and Hospitality Management*, *24*(1), 151–183. <https://doi.org/10.20867/thm.24.1.3>
- Sartori, A., Mottironi, C., & Corigliano, M. A. (2012). Tourist destination brand equity and internal stakeholders: An empirical research. *Journal of Vacation Marketing*, *18*(4), 327–340. <https://doi.org/10.1177/1356766712459689>
- Shamseer, L., Moher, D., Clarke, M., Gherzi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A., Altman, D. G., Booth, A., Chan, A. W., Chang, S., Clifford, T., Dickersin, K., Egger, M., Gøtzsche, P. C., Grimshaw, J. M., Groves, T., Helfand, M., ... Whitlock, E. (2015). Preferred reporting items for systematic review and meta-analysis protocols (prisma-p) 2015: Elaboration and explanation. *BMJ*, *349*. <https://doi.org/10.1136/bmj.g7647>
- Statista. (2022, April). *Statista*. <https://www.statista.com/>
- Streiner, D. L. (2005). Finding Our Way: An Introduction to Path Analysis. *Can J Psychiatry*, *50*(2)

- Turismo de Portugal. (2022, April). *Turismo de Portugal*.
<https://www.turismodeportugal.pt/pt/Paginas/homepage.aspx>
- UNWTO. (2022, April). *United Nations World Tourism Organization*.
<https://www.unwto.org/statistics>
- Vaishnavi, V., & Kuechler, B. (2004). Design Science Research in Information Systems. *Ais*, 1–45
- Visit Portugal (2022, April). *Visit Portugal*. <https://www.visitportugal.com/en/destinos/lisboa-regiao>
- Visser, M., van Eck, N. J., & Waltman, L. (2021). Large-scale comparison of bibliographic data sources: Scopus, web of science, dimensions, crossref, and microsoft academic. *Quantitative Science Studies*, 2(1), 20–41. https://doi.org/10.1162/qss_a_00112
- WebHarvy. (2022, April). *Web Harvy*. <https://www.webharvy.com/>
- WorldData. (2022, March). WorldData.
<https://www.worlddata.info/europe/portugal/tourism.php>
- Xu, J., Chan, E., & Pratt, S. (2018). Destination Image of Taiwan From the Perspective of Hong Kong Residents: Revisiting Structural Relationships Between Destination Image Attributes and Behavioral Intention. *International Journal of Hospitality and Tourism Administration*, 19(3), 289–310. <https://doi.org/10.1080/15256480.2017.1324339>
- Yadav, R., & Pathak, G. S. (2017). Determinants of Consumers' Green Purchase Behavior in a Developing Nation: Applying and Extending the Theory of Planned Behavior. *Ecological Economics*, 134, 114–122. <https://doi.org/10.1016/j.ecolecon.2016.12.019>
- Yandex. (2022, May). *Yandex*. <https://translate.yandex.com/>

Appendix I

Table of frequencies of the keywords of selected articles for SLR

Table I.1: Keywords of the selected articles of SLR

Word	Frequency	Word	Frequency
destination	20	structure	1
brand	19	modelling	1
equity	14	seoul	1
image	5	chinese	1
loyalty	5	frequency	1
branding	4	new	1
equation	3	zealand	1
awareness	3	wine	1
quality	3	designation	1
marketing	3	origin	1
visitors	3	partial	1
consumer-based	3	least	1
structural	2	squares	1
model	2	regression	1
travel	2	multi-group	1
intention	2	culture	1
perceived	2	rome	1
dimensions	2	enduring	1
tourist	2	involvement	1
satisfaction	2	visit	1
analysis	2	management	1
australia	2	stakeholders	1
tourism	2	modeling	1
internal	2	customer	1
customer-based	1	perspective	1
generation	1	renewal	1
world	1	repeat	1
heritage	1	chile	1
familiarity	1	latin	1
indian	1	america	1
state	1	mongolia	1

Appendix II

Sentimentr package to calculate the sentiment score

Package 'sentimentr'

October 12, 2021

Title Calculate Text Polarity Sentiment

Version 2.9.0

Maintainer Tyler Rinker <tyler.rinker@gmail.com>

Description Calculate text polarity sentiment at the sentence level and optionally aggregate by rows or grouping variable(s).

Depends R (>= 3.4.0)

Suggests testthat

Imports data.table, ggplot2, graphics, grid, lexicon (>= 1.2.1),
methods, stats, stringi, syuzhet, textclean (>= 0.6.1),
textshape (>= 1.3.0), utils

License MIT + file LICENSE

Encoding UTF-8

LazyData TRUE

RoxygenNote 7.1.2

URL <https://github.com/trinker/sentimentr>

BugReports <https://github.com/trinker/sentimentr/issues>

NeedsCompilation no

Author Tyler Rinker [aut, cre]

Repository CRAN

Date/Publication 2021-10-12 08:30:02 UTC

Description

Approximate the sentiment (polarity) of text by sentence. This function allows the user to easily alter (add, change, replace) the default polarity and 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

<code>text.var</code>	The text variable. Can be a <code>get_sentences</code> object or a raw character vector though <code>get_sentences</code> is preferred as it avoids the repeated cost of doing sentence boundary disambiguation every time <code>sentiment</code> is run.
<code>polarity_dt</code>	A data.table of positive/negative words and weights with <code>x</code> and <code>y</code> as column names. The lexicon package has several dictionaries that can be used, including: <ul style="list-style-type: none"> • <code>lexicon::hash_sentiment_jockers_rinker</code> • <code>lexicon::hash_sentiment_jockers</code> • <code>lexicon::emojis_sentiment</code> • <code>lexicon::hash_sentiment_emojis</code> • <code>lexicon::hash_sentiment_huliu</code> • <code>lexicon::hash_sentiment_loughran_mcdonald</code> • <code>lexicon::hash_sentiment_nrc</code> • <code>lexicon::hash_sentiment_senticnet</code> • <code>lexicon::hash_sentiment_sentiword</code> • <code>lexicon::hash_sentiment_slagsd</code> • <code>lexicon::hash_sentiment_socal_google</code> Additionally, the <code>as_key</code> function can be used to make a sentiment frame suitable for <code>polarity_dt</code> . This takes a 2 column <code>data.frame</code> 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 <code>lexicon::hash_sentiment_huliu</code> to obtain the old behavior.
<code>valence_shifters_dt</code>	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 <code>x</code> and <code>y</code> as column names.
<code>hyphen</code>	The character string to replace hyphens with. Default replaces with nothing so 'sugar-free' becomes 'sugarfree'. Setting <code>hyphen = " "</code> would result in a space between words (e.g., 'sugar free'). Typically use either " " or default "".
<code>amplifier.weight</code>	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.
<code>n.before</code>	The number of words to consider as valence shifters before the polarized word. To consider the entire beginning portion of a sentence use <code>n.before = Inf</code> .
<code>n.after</code>	The number of words to consider as valence shifters after the polarized word. To consider the entire ending portion of a sentence use <code>n.after = Inf</code> .

<code>question.weight</code>	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 evaluation 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.
<code>adversative.weight</code>	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 $1 + \text{number of occurrences of the adversative conjunctions before the polarized word times the weight given } (1 + N_{\text{adversative conjunctions}} * z_2$ where z_2 is the <code>adversative.weight</code>). Conversely, an adversative conjunction found after the polarized word in a context cluster down weights the cluster $1 - \text{number of occurrences of the adversative conjunctions after the polarized word times the weight given } (1 + N_{\text{adversative conjunctions}} * -1 * z_2)$. 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.
<code>neutral.nonverb.like</code>	logical. If TRUE, and 'like' is found in the <code>polarity_dt</code> , 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.
<code>missing_value</code>	A value to replace NA/NaN with. Use NULL to retain missing values.
<code>retention_regex</code>	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 <code>\\p{L}</code> for <code>[^[:alpha:];; , \\'</code> may retain more alpha letters but will likely decrease speed. See examples below for how to test the need for <code>\\p{L}</code> .
<code>...</code>	Ignored.

Details

The equation used by the algorithm to assign value to polarity of each sentence first utilizes the sentiment dictionary to tag polarized words. Each paragraph ($p_i = \{s_1, s_2, \dots, s_n\}$) composed of sentences, is broken into element sentences ($s_i, j = \{w_1, w_2, \dots, w_n\}$) 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 ($c_{i,j,l} = \{pw_{i,j,k-nb}, \dots, pw_{i,j,k}, \dots, pw_{i,j,k-na}\}$), 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. An adversative conjunction before the polarized word ($w_{adversative\ conjunction}, \dots, w_{i,j,k}^p$) up-weights the cluster by $1 + z_2 * \{|w_{adversative\ conjunction}|, \dots, w_{i,j,k}^p\}$ (.85 is the default weight (z_2)). An adversative conjunction after the polarized word down-weights the cluster by $1 + \{w_{i,j,k}^p, \dots, |w_{adversative\ conjunction}|\} * -1$ 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:

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

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

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

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

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

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

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

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/syuzhet>
- 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/jeffreymbreen/r-by-example-mining-twitter-for>
- <http://hedonometer.org/papers.html> Links to papers on hedonometrics

<code>get_sentences</code>	<i>Get Sentences</i>
----------------------------	----------------------

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)
```