OLUGBENGA AKINKUNMI GEORGE

SENTIMENT ANALYSIS APPLIED TO TOURISM: EXPLORING THE TOURIST GENERATED CONTENT IN THE CASE OF A WELLNESS TOURISM DESTINATION.

UAlg FE

UNIVERSIDADE DO ALGARVE

FACULDADE DE ECONOMIA

UNIVERSITY OF ALGARVE

FACULTY OF ECONOMICS

2023

OLUGBENGA AKINKUNMI GEORGE

SENTIMENT ANALYSIS APPLIED TO TOURISM: EXPLORING THE TOURIST GENERATED CONTENT IN THE CASE OF A WELLNESS TOURISM DESTINATION.

Masters in Tourism Organization Management

Dissertation

made under the supervision of:

Professor Célia Maria Quitério Ramos

UNIVERSITY OF ALGARVE

FACULTY OF ECONOMICS

2023

Work Authorship Declaration

I declare to be the author of this work, which is unique and unprecedented. Authors and works consulted are properly cited in the text and are included in the listing of references.

(Olugbenga Akinkunmi George)

_____

(signature)

Copyright

Dedication

This research work is dedicated to Almighty God: To Thee I give all praise, glory and adoration (Halleluyah). To my lovely wife, Roseline Komiyo George

Acknowledgements

ABSTRACT

Sentiment analysis applied to tourist-generated content through social media is one of the many sources of information for travellers in the case of a wellness tourism destination. The objectives of exploring tourist-generated content are to analyse the sentiments expressed on social media, in this research about wellness tourism destinations, based on destination reviews, and to identify the motivations that lead tourists to look for this kind of services in the Algarve destination The analysis of the hotel reviews exposes the tourist's opinion of the wellness destination and the grading of the infrastructural facilities and the acclaimed services in the categories of the tourist wellness destination. In attaining the objectives, quantitative approach methods were used to review 1,294 comments manually extracted from TripAdvisor and apply text mining algorithms to assess the sentiment expressed in the comments. The result shows that the three key extracted words frequently used by all categories of the component in the Algarve wellness destination are "bird", "day", "experience", "Faro", "fish", "great", "horse", "love", "market", "massage", "nice", "old", "park", "Pinetree", "place", "relax", "ride", "spa", "time", "town", "visit" and "walk", also revealed some categories of a hotel into the services of spas and massage in the tourist wellness destination mainly satisfy the desire of the tourist. Exploring generated content with sentiment analysis helped to stimulate opinion on destinations with wellness facilities, revealing how the wellness of individuals at tourism destinations are generally satisfied based on keywords that illustrated opinions, shows diverse opinion, which shows that expectations are being met from the comments in TripAdvisor. It also widens the destination management organisation's (DMO) knowledge of the sustainability of destination infrastructures due to tourist views and opinions on destinations. However, generated review content which may be changed due to implicit factors, sometimes gives a psychological time frame view of the individual reviewer.

KEYWORDS: Tourism Generated content, Sentiment Analysis, Tourism, Wellness destination, Text Mining.

RESUMO

A análise de sentimentos efetuada aos conteúdos gerados pelo turista nas redes sociais é uma das muitas fontes de informação consultadas pelos viajantes, permite visualizar os testemunhos de outros turistas que já visitaram os mesmos destinos.

Através de avaliações efetuadas nas plataformas de social mídia, sobre infraestruturas existentes nos destinos turísticos, como no caso de um destino de saúde e bem-estar, permite estudar os principais motivos pelos quais estas regiões suscitam um sentimento mais positivo que outros, associado à experiência turística, com ou sem valor acrescentado pelos gestores do destino.

A diversidade de opiniões, que por vezes é em função das expectativas que estão a ser satisfeitas, ou não, expressas através dos comentários gerados no TripAdvisor, podem contribuir para a sustentabilidade de infraestruturas de apoio ao turismo, que inclui, entre outras, instalações culturais tangíveis e intangíveis. Face ao exposto, os objetivos da presente dissertação são: (i) analisar o sentimento expresso nos conteúdos turísticos gerados pelos viajantes nas redes sociais sobre o destino turístico de saúde e bem-estar, (ii) avaliar o sentimento do turista sobre as componentes que caracterizam este tipo de destino e (iii) identificar as motivações que levam os viajantes a optar pelo Algarve, como um destino de turismo de saúde e bem-estar, que compreende dezasseis municípios com atrações e especificidades únicas.

Para analisar o sentimento expresso nos comentários, foi necessário recorrer a uma metodologia própria de mineração de texto, cujo processo envolveu a recolha de dados, pré-processamento, extração de características, aplicação do algoritmo de análise de sentimento e por fim avaliação dos resultados. As avaliações foram extraídas manualmente e categorizadas em Número da Entidade (*ID*), Nome (*Nome*), Cidade onde está localizada (*Cidade*), País (*País*), Categoria (*Categoria*) Número de estrelas (*Estrela*) e número de avaliações (*Número de avaliações*). No conjunto de entidades foram considerados hotéis e infraestruturas selecionados pelas suas características: classificação e serviços turísticos adicionais prestados na área.

De acordo com a revisão da literatura, as componentes consideradas e analisadas no que se refere a um destino de saúde e bem-estar foram dez, nomeadamente: Termas, Spas,

Cuidados de corpo e mente, Turismo Médico, Ambiente Natural, Espiritualidade, Cultura, Enogastronomia, Desporto e Eventos. Neste contexto, foram identificados os hotéis e infraestruturas considerados como representativos de cada tipo de instalação que caracteriza este tipo de turismo e em seguida foram extraídos comentários do TripAdvisor. A seleção recaiu sobre as seguintes instalações: Caldas de Monchique (Monchique), Zensation - Body & Soul (Albufeira), Mercado de Loulé (Loulé), Moinhos Velhos Juice Detox & Yoga Retreat (Lagos), Old Town Faro (Faro), Pinetrees Horse Riding (Almancil), Ria Formosa Natural Park (Olhão), Swan Day Spa (Alvor) e Zoomarine Algarve (Albufeira).

O pré-processamento envolveu a limpeza de elementos indesejados, transformação, tokenização, normalização e filtragem de texto.

Na avaliação do resultado, a análise de sentimento foi também aplicada a duas dimensões (palavras extraídas e associadas em cada categoria), que foram comparadas com o comentário correspondente para determinar qual o sentimento associado a cada entidade considerada na investigação. Para atingir os objetivos, foram utilizados métodos de abordagem quantitativa para rever 1,294 comentários extraídos manualmente do TripAdvisor e nos quais foram aplicados algoritmos de mineração de texto para avaliar o sentimento expresso nos mesmos. O resultado destacou como as palavras-chave de comentários positivos: "*bird*", "*day*", "*experience*", "Faro", "*fish*", "*great*", "*horse*", "*love*", "*market*", "*massage*", "*nice*", "*old*", "*park*", "Pinetree", "*place*", "*relax*", "*ride*", "*spa*", "*time*", "*town*", "*visit*" e "*walk*", Também, evidenciou que o sentimento associado a "spas", "luxo", "massage" e "sauna" obtiveram os valores mais elevados, no que se refere ao sentimento positivo.

Os resultados expressaram de forma clara a aceitabilidade e popularidade perante os turistas, do Algarve como destino de saúde e bem-estar, onde se destacaram os termos "spas" e "cuidados de corpo e mente", "ambiente natural" e "desporto" com desvio padrão de 0.4 e 0.3, com média de 4.7 por cada termo. Mostra que destinos de saúde e bem-estar com desporto (golfe), natureza e paisagem são características positivas neste tipo de destino. Além disso, as avaliações geradas durante os meses de inverno: janeiro, fevereiro, março, novembro e dezembro apresentaram valores mais altos. O estudo revelou, ainda, que algumas das categorias associadas ao destino de bem-estar turístico

satisfazem o desejo do turista, principalmente, no que refere aos serviços de spas e massagem oferecidos pelos hotéis.

A extração de comentários manualmente, apesar de ser uma limitação, concedeu ao investigador a possibilidade de confirmar questões relacionadas com a inclusão de taxas escondidas, com o tempo gasto no acesso à área de destino e às instalações turísticas, as atitudes do staff, prestação seletiva de serviços sem regras padrão, sem correlação entre o serviço anunciado e o prestado.

A exploração de conteúdos gerados pelos turistas, em conjunto com análise de sentimentos, permitiu adquirir conhecimento sobre a perceção dos turistas sobre o destino, assim como revelou que o bem-estar dos mesmos nos destinos turísticos é geralmente satisfatório e expresso através de comentários diversos nas redes sociais que manifestam sentimentos, o que mostra que as expectativas foram satisfeitas. Para além disso, também permitem à organização de gestão de destinos (DMO) conhecer as potencialidades e desafios identificados pelos turistas. As categorias de instalações associadas a um destino de saúde e bem-estar, podem ajudar a melhorar a sua viabilidade económica através da gestão do sector turístico, contribuindo para o PIB, bem como permite identificar áreas que necessitam de especial atenção por parte dos governantes, como os eventos e o turismo médico.

Contudo, o conteúdo gerado pelos turistas pode ser alterado devido a fatores implícitos, como a visão psicológica do estado do revisor individual ou o tempo que caracteriza o destino no momento. Assim, para futuros estudos, o processo de mineração de texto deverá ser efetuado de forma automática e em diversas fontes de informação, de modo que inclua mais instalações e constituído por mais comentários, especialmente no turismo médico que não foram analisados, uma vez que não existiam.

Palavras-chave: Conteúdos Gerados pelo Turista, Análise de sentimento, Turismo, Destino de saúde e bem-estar, Mineração de texto.

GENERAL INDEX

FIGURES INDEX

TABLES INDEX

CHAPTER ONE - INTRODUCTION

1.1. Introduction

Customers of tourism products relate more on social networks and search for online reviews to be informed on all kinds of goods, services and brands (Filieri and McLeay, 2013). The new means to quantify the quality of destinations in terms of in-depth attractions and wellness is the Web 2.0 applications that empower the users and influence the information gathering of travellers' destinations (Filieri and McLeay 2013; González-Rodríguez, Martínez-Torres, and Toral, 2016), and also credible informative data in the sense that people share their personal travelling experience about destination visited (Ren and Hong, 2017). Still, it is a subjective opinion (positive, negative) (Vu, Li, Law, and Zhang, 2019).

The motivating subjective opinion that enhances the choice of a destination can be through multiple sources: induce agents, magazines, articles, guidebook and WOM (Word-of-Mouth) or e-WOM (electronic Word-of-Mouth) or both (González-Rodríguez *et al.,* 2016), Twitter, Tumblr, and Instagram (El-Masri, Altrabsheh, and Mansour, 2017). The organic image through WOM's provided by family and friends, or individual travellers on the past visit destination image, the e-WOM are from social networks and online reviews and the inducing agents are made up of published magazines, articles, guide books, news, films, and Television promotions and documentaries by the destinations management organisations (DMO's) (González-Rodríguez *et al.*, 2016; Jiang, Chan, Eichelberger, Ma, and Pikkemaat, 2021).

In the choice of destination, the online review has a significant impact, and at the same time, it can hinder the choice of destination because of the volume of information which is subjective (González-Rodríguez *et al.*, 2016). In the tourism destination, a lot of attractions make a unit destination to be attractive to tourists. The image may be in natural resources, general infrastructure, tourism infrastructure, tourism leisure and recreation, culture, history and art, political and economic factors, natural environment, social environment, and atmosphere (Jiang *et al.*, 2021).

In determining the opinion about a wellness tourism destination in hotel review, consideration should be given to a safe/secure environment in both perception and reality, clean and sanitary infrastructure for both locals and visitors, quality-of-life for locals who benefit from tourism dollars—e.g. the creation of jobs within the industry and the creation of a market for locally made produce/products/services, natural assets such as hot springs/mountains/bodies of water/forests/resources for thalassotherapy or other natural assets within the confines of the destination and easily accessible to visitors (Arean, 2019).

Wellness Tourism and Wellness Travel also encompass wellness for the planet, the destination must have substantial sustainability policies and practices in place, the availability and accessibility of a wide range of wellness-professionals and practitioners, including those who offer holistic and alternative modalities, selection of hotel restaurants and independent restaurants offering healthful cuisine prepared by chefs committed to clean eating and who work in partnership with local growers, availability of a range of fitness-based activities and tours—e.g. yoga, hiking, cycling, fitness classes, kayaking, stand-up paddle boarding, a physical environment that is somewhat removed from the noise that has become "daily life" in the 21st century (Arean, 2019).

The review of the related literature will discuss the concept of destination, the historical background of wellness tourism and sentiment analysis. Sentiment is more of an emotion, feelings (positive, negative) of the customer. Sentiment analysis is also called opinion mining which studies people's opinions towards the product (Muthukrishnan, Selvi, Deivakani, Subashini, and Kumar, 2021).

Also, the importance of this study is that, it will be a guide to public policy managers on policy-making and execution that will further be improved the levels of sustainability of destinations, improve the patronage to the tourism places and also the private sectors, who were usually directly in contact with tourist that visit the wellness destinations.

1.2. Importance of the study

This study will help to identify the reasons why some destinations will stimulate positive sentiment because of the ambience created by nature with or without added value from destination organisational managers.

It helps to reveal the diversity of opinion, which is sometimes a function of expectations being met and the mindset about a destination, through the comments on TripAdvisor.

The study will widen the knowledge of the destination organisational managers (DMOs) on the sustainability of infrastructural facilities, which includes, among others, tangible and intangible cultural facilities.

It shows that the objectives of visiting a well-being destination are to meet the needs of the individual tourist, which are revealed through online reviews.

1.3. Objectives of the Study

To analyse the sentiments about wellness tourism destinations in products reviews related with this type of tourism.

To evaluate the opinion of tourists based on destination review, which is subdivided into specific objectives:

- To identify and analyse the main studies on the wellness tourist destination.

- To identify the components that characterise a wellness tourist destination.

- Collect comments posted on a social network to form the sample.

- Apply the different pre-processing (cleaning) techniques to the database.

- Apply Text Mining algorithms to assess the sentiment expressed in the comments.

- Considering the results of sentiment analysis, to identify the motivations that lead tourists to look for wellness tourist destinations in Algarve.

1.4. Statement of the problem

The general problem emanated from the perceptions, opinions and feeling on wellness tourism destinations through words of mouth (WOMs) and in hotel reviews in social media, which have greater effects on the components and the images of units of events in the area of tangible and intangible facilities and activities that form a destination.

1.5.   Boundaries of research

The scope of the research is limited to the geographical region of the Algarve. The components that characterize a wellness tourist destination were located in some of the districts of the Algarve, which include Faro, Albufeira, Loulé, Monchique, and Olhão.

1.6. A brief summary of the chapters' part of the work

The research work will be divided into six chapters.

- Chapters one will deal with introduction part, significance of the study, the objectives of the investigation, the statement of the problem that lead to the study and the boundaries that will be covered by the research.

- Chapter two will look at the review of related literature on the study. It comprises the theoretical concept of a tourism destination, the historical background of wellness tourism, the components that characterize wellness destinations, the concept of sentiment analysis in a user-generated content (UGC) and the conceptual research model for the present study.

- Chapter three will be for the methodology of the study, which will use the primary data of reviews and comments generated from the TripAdvisor on wellness tourism products of Algarve destination. The chapters will evaluate the result from the application of the sentiment algorithms with the use of data mining software to pre-process the extracted reviews and comments from TripAdvisor in order to extract the opinion of tourist in Algarve destination.

- Chapter four will show the analysis, interpretation of result and conclusion with graphical illustration for a deep analysis and a clear outcome interpretation.

- Chapter five covers the practical implications for the management of the tourism sector and the contributions to the scientific community generally.

- Chapter six will summarise the introductory statement, validate the result, clarify the boundary of the research and suggestions for further study.

CHAPTER TWO - LITERATURE REVIEW

2.1. Tourism Destination

A destination is the place where someone is going or where something is being sent or taken, worth making a special journey or, and it collocates with certain words which include attractive destination, and favourite destination. In the view of Medilik (2003), destinations are the main concentrations of tourist attractions, accommodation and other tourist facilities and services, where the main impacts of tourism– economic, social, and physical occurs. When the physical plant and its supporting amenities, facilities, activities, interests, and attractions are combined, the resulting geographic space, is called a tourist destination (Klein-Hewett, 2021). For Komilova, Usmanov, Safarova, Matchanova, and Murtazaeva (2021), a destination is a geographical region with particular attractiveness for tourists and must meet the following basic requirements:

i.   Availability of the necessary service complex for the reception of tourists;

ii. Availability of attractive objects and places of interest to tourists (one of the main factors of competition between destinations is attractiveness, which should be a unique aspect to attract tourists to the destination area);

iii. The availability of information systems is a necessary "tool" of destinations in the tourism market.

In the review of different definitions of "destination", Komilova *et al.* (2021) divide concept of "destination" into four groups, namely: territorial, economic, social, and administrative.

Territorial assessment of the concept of "tourist destination", to Komilova *et al.* (2021), is based primarily on the criterion of "dependence" on the specific geographical location of the region, which cannot be changed, and tourists are directed in that direction, also destination can be considered buildings that are not related to a specific area, that is, in the opinion of the author, the location of the destination may change.

In the economic aspect (Komilova *et al.*, 2021), the tourist destination is considered a specific destination visited by tourists. It is regarded as an economic category, the tourist demand, not only as an attractive area for tourists, but primarily have their own consumption value and as an object of market relations and a tourist product. The presence of tourist infrastructure (accommodation, catering, entertainment, vehicles) is also an important feature. So, it can be concluded, that a tourist destination not only plays the role of a commodity or product, but also a specific economic space.

The social aspect on tourist destination is interpreted as a specific socio-geographical place (Komilova *et al.*, 2021) which is also a place of life in the local community.

The administrative aspect considered "tourist destination" as a subject of descriptive governance control of any system by forces of a natural nature (Komilova *et al.*, 2021). In this opinion, a person as a subject of management, and the tourist destination are evaluated from the same point of view regarding management. They conclude that there is a management aspect in the definition of the concept of a "tourist destination" and its evaluation as a market object, so a tourism destination is interpreted as a controlled system. Arif, Nugroho, and Hariadi (2019) refer to a tourism destination as an area chosen by visitors to carry out tourism activities, covering all available facilities, including entertainment facilities, accommodation, lodging, restaurants, and so on.

Żemła (2016) viewed the concept of Tourism destination from four approaches: spatial approach, economic approach, managerial approach, systems approach, and network approach.

A spatial approach to the definition of a tourism destination is considered to be a geographical area or region, in which the visitor enjoys various types of travel experiences in the form of physical entities and intangible socio-cultural entities which motivate visiting tourists and encourage tourism activities (Żemła, 2016).

The economic approach considered by Żemła (2016), is divided to demand side and supply side approach. While demand side approach defined tourism destination as a choice for a tourist because it ensures satisfaction with the holiday's special needs in the area of accommodation, catering and entertainment, cultural background, the purpose of visit, educational level, past experience and their travel itinerary, but Krakover and Corsale (2021) identified time and money budgets as a constraint for repetition of the trip. The supply side approaches defined tourism destination as an area, country, region or

human settlement for which a significant offer of attractions and infrastructure of tourism with visitation experiences, it contains concentration of small/medium-sized companies which offer tourism products or services.

Managerial approaches to the concept of destination according to Żemła (2016), considered a destination as a product and a company as a metaphor of a destination. A destination as a product is an area with a different combination of tourism products or services that provide a unique experience, and as a company as a metaphor of a destination, it may be seen as strategic business units from the management point of view that provide products and offers for defined target groups and guest segments.

In according with Żemła (2016) in the definition of destination, the system approach considered a destination as a place of actors that co-operates in order to supply an integrated tourist product. The actors include private and public sectors.

In the area of network approach, also Żemła (2016) considered the definition of a tourism destination as a complex system, as an area bound to no administrative limitations, where tourist aspects are interrelated and integrated.

Based on reviews of various conceptualizations on destinations, Jovicic (2019) focuses on three destination concepts, which are the classical–traditional view of destinations, the systemic approach to destinations, and the concept of smart tourism destinations.

In the classical–traditional view of destinations, Jovicic (2019) focused his research on the characteristics that are geographical to define tourist destinations as tourism as an industry occurs at destination areas – areas with different natural and/or man-made features, which attract nonlocal visitors (or tourists) for a variety of activities. The tourism destination of the classical concept view that tourism destinations need to meet certain criteria in order to be considered a destination, and such criteria include having tourist attractions and accommodations as well as the transport to, from, and within the destination but did not emphasise the significance of cooperation within the destination nor the role of tourists as actors in the destination. As such, tourists are only seen as consumers of a destination's supply of services, while neglecting those changes in demand lead to changes in the structure of the destination. Also, in interpreting a specific destination and its market attractiveness, we must consider the perception of tourists, who experience a given destination in a subjective way, depending on their purpose of travel, educational and cultural level, past experience and so on.

The systemic approach to destinations, according to Jovicic (2019), implies a more comprehensive understanding of tourism destinations, where tourism should be considered in the context of other systems that interact with tourism, treats a destination as an open and flexible system, characterized by a high degree of interaction between its constituent elements, such as firms providing tourist services, residents of destinations, local authorities and tourists and also conceptualize a destination as a network of connected organisations/stakeholders whose productivity is essential for the functioning of the destination system.

In the concepts of the smart tourism destinations approach (Jovicic, 2019; Gelter, Lexhagen, and Fuchs, 2021), information and communication technologies (ICTs) development has changed the traditional frame for tourism development and brought smartness into destinations. A smart tourism destination is a knowledge-based destination, where ICTs are used to provide a technological platform on which information and knowledge relating to tourism could be instantly exchanged and suggest that there are several forms of ICTs which are vital for setting up smart tourism destinations, especially the Cloud Computing and Internet of Things (IoT).

According to Jovicic (2019), Cloud Computing services are designed to provide a convenient way to access solid web platforms and data storage through the certain networks. It means storing and accessing data and programme over the Internet instead of a computer's hard drive, or simply saying: if someone does not have enough money to buy an apartment or a car, then they can rent them. It stimulates information sharing that has a great significance for the operation of smart tourism destinations. Through a cloud computing system, a tour guide system could serve a massive number of tourists without being actually installed on any personal device, also tourists can get information irrespective of time and space by the use of mobile phones or portable devices from TripAdvisor, WhatsApp, Tripcast and HearPlanet. IoT technology generates automatic real-time interactions among real-world objects that connect to the Internet which reduces the gap between the real world and the digital realm. Implementation of IoT in tourism means that tourists could simply use their mobile phones to explore the destination attractions using in-situ data collection and reporting (Jovicic, 2019).

The unique influencing factor that encourage visitors to destinations is its attractiveness and it define the choice of a tourist destination, tourist frame of mind, expected well-being at the destination and the motivator in recommending the destination to others (Reitsamer

and Brunner-Sperdin, 2017). The well-being is affective or hedonic and a cognitive, while the hedonic aspect refers to the rate to which several desires a person experiences are pleasant, for example, how well one usually feels, and the cognitive aspect represents the rate at which a person perceives his or her desires to have been met (i.e., the extent to which one is contented with life in general) (Reitsamer *et al.*, 2017). An alternative conceptualisation views the well-being achievement of a person as an evaluation of the "wellness of the person's state of being", which depends on a person's actual ability to achieve various valuable functioning (i.e., parts of the state of a person, such as being adequately nourished or being in good health). Reitsamer *et al.* (2017), although to Milićević and Jovanović (2015), Wellness today represents the world's lifestyle.

2.2. Wellness tourism

2.2.1. Historical background of wellness tourism

According to Hankering (2021), the true origin of wellness tourism has many historical sources, from the ancient religious practice of about 3000 B.C.E to modern wellness that began around the 20th century.

In ancient practice, wellness can be found in every religion that focuses on the harmony of the spirits and the body. According to Miller (2021), Ayurvedic medicine ("Ayurveda" for short) is one of the world's oldest holistic ("whole-body") healing systems. Ayurvedic was developed more than 3,000 years ago in India and the earliest reference to healing waters of spas (Latin *sanitas per aqua*) is about 1,700 BC. Greeks, Romans and Ottomans have left a legacy of baths to the Europeans, also the Siamese (Thai) practised massage even before 100 BC, and the Japanese used natural therapy springs in 737 AD (Koncul, 2021).

It's based on the belief that health and wellness aim is to encourage healthy living and it all depends on the steadiness of the mind, body and spirit. However, the management is geared to a particular health issue (Miller, 2021).

Wellness development can be traced back to the Chinese system of medicine from Buddhism and Taoism which uses acupuncture and herbal medicines (Hankering, 2021).

People have been using spas for their well-being from ancient times as archaeological evidence for bathing in Egypt and Babylon points to nearly 3,000 BCE and idolization of antique Greek steam baths Laconia and, later, Roman baths of Tepidarium around the same period (Kazakov and Oyner 2021).

Tubergen and Linden, (2002) viewed that freshwater baths were more desirable by the Greeks from natural resources, especially from the sea(ocean) although, it was private bathing initially, confined to the wealthier people but later opened to the public and those bath place were dedicated to several deities and were considered sacred. The Roman early history of the primary use of curative baths was to heal the wound of roman soldiers and also as resting and recreation centres for healthy soldiers during the reign of Augustus from 27 B.C to 14 A.D. at this time, there were approximately 170 baths, known as a thermal in Rome and it did not take long before all the cities began to view baths as a form of rest relaxation, the Romans considered the baths more important for cleansing, exercises, socialising, relaxation, and worship, medical treatment was also applied extensively (Tubergen and Linden, 2002).

It was in 70 A.D that the Roman built a thermal baths spa around the hot springs at Bath, the first of its kind in Britain (The history of spas and spas treatment). In 1.326, Collin le Loup, an iron master from Liege, Belgium, discovers the chalybeate springs in the town of Spa, Belgium. A famous health resort eventually grew around these springs. The term "spa" came to refer to any health resort located near natural springs, with individual springs being associated with diseases they were thought to benefit from (Tubergen and Linden, 2002).

In the 20th century and beyond, according to Mueller and Lanz Kaufmann (2001) cited in Tubergen and Linden (2002), Europe (Britain, Germany, Austria, and Belgium) and North America witness the growth of hotels and guesthouse and the use of mineral waters during springs. Also, there were growth in the numbers of outstanding hotels with spa resort had its own theatre, casino, and promenades beside the bathing buildings. Tubergen et al. (2002) opined that the spa resorts turn out to be the elites' meeting centre and also a place of creativity for painters, writers, and composers.

Health preservation and promotion present an important issue in tourists' behavioural aspects (Damijanic, 2019; Mueller and Lanz Kaufmann, 2001), such as diet, fitness and stress management and in determining the connection concerning a healthy way of life

and intention to travel. So, wellness was characterised as a state of health which is connection of body, mind and spirit, with self-responsibility, physical fitness/beauty care, healthy nutrition/diet, relaxation/meditation, mental activity/education and environmental sensitivity/social contacts as fundamental elements (Damijanic, 2019).

Kazakov and Oyner (2021) emphasis that wellness tourism is now acceptable in the industrialised countries as a way of life and has become a more pleasure-seeking, and a big market in the twenty-first century. So, the whole wellness tourism concepts will result into providing various tourist services by merging health, wellbeing, hospitality and transportation into an industry.

2.2.2. The components that characterize a wellness tourist destination

Wellness tourism can be defined as "travel associated with the pursuit of maintaining or enhancing one's personal well-being" (Global Wellness institute, 2022). Also, Kathleen Lesage (Wellness Tourism Association, 2022) defined wellness tourism as offering personal growth, outdoor adventures, nutritious cuisine, spa treatments, bodywork and opportunities in preventative care, all while on vacation. The China National Tourism Administration defines it as the sum of all kinds of tourism activities that enable tourists to achieve a natural, harmonious and pleasing state in body, mind and spirit through various means such as physical and physical health, nutrition, mental health, caring for the environment, etc (Zhang, Li and Wu, 2021). Wellness is also the total interactions from a trip by tourists with the objectives of maintaining their health and well-being and who stay at a facility for a nigh which is explicitly designed to aid physical, psychological, spiritual and/or social well-being (Valentine 2016; Stará and Peterson, 2017). In enhancing psychological social well-being, Thal and Hudson (2017) study emphasised that, the service scape qualities standard and relationship between staff and guests are the best for creating a psychologically beneficial environment, so the basis of wellness tourism is related to travelling for maintaining well-being and lifestyle improvement (Han, Kiatkawsin, Koo, and Kim, 2020), and also the term "wellness" refers to "a special state of health, incorporating the harmony of body, mind, and spirit" and once lifestyle (Han et al., 2020)

Also, Pan, Yang, Han, Lu and Liu (2019), defined initially Wellness as a state with optimal physical, mental, and social well-being but with the new trend in the market due

to the boom in wellness tourism, with many holistic and individual perspectives, wellness tourism takes new dimensions which include: self-worth, exercise, spirituality, stress management, love, gender identity, realistic beliefs, leisure, humour, intelligence, nutrition, emotion, friendship, self-care, work, social identity, local safety, institutional concern, and personal attributions. It also has some specific activities, which include personal service, health promotion treatments, environment, healthy diet, relaxation, social activities, the experience of unique tourism resources, and mental learning (Pan, *et al.,* 2019).

On other side, Brandão, Liberato, Teixeira, and Liberato (2021) stress that wellness tourism concerns tourists who were looking for improvement in general well-being through involvement in activities that are self-care nature, such as yoga, tai chi, hot springs/spa, massage, and physical activities. Plzáková and Stupková (2019), declare that health tourism is an umbrella for medical tourism and wellness tourism and thus defined wellness tourism as "trips aiming at a state of health featuring the harmony of the body, mind and spirit, self-responsibility, physical fitness, beauty care, healthy nutrition, relaxation, meditation, mental activity, education, environmental sensitivity and social contacts as fundamental elements" (Smith and Puczkó, 2014: 208). The holistic perspective of Dini and Pencarelli (2021), provides a more detailed explanation of the wellness activities that make up each component of wellness tourism as follows: hot springs, spas, care of body and mind, medical tourism, natural environment, spirituality, culture, enogastronomy, sports, and events.

- **Hot springs:** This is a related wellness service that involved thermal treatments water termed Hydrotherapy treatment - baths, hot and cold shower baths and water massages better known as balneotherapy (Dini and Pencarelli, 2021; Rodrigues, Brochado and Troilo, 2020). Also, Brandão *et al.* (2021) define "hot springs" as a mineral water source with important therapeutic properties for the treatment of certain diseases that have existed for centuries.
- **Spas:** are organisations that offer water-based treatments – thermal (hydrotherapy) and / or marine (thalassotherapy) – for body care and their peculiar feature lies precisely in the use of water which becomes the central element behind each therapy Lemi (2021). Brandão *et al.* (2021) recognise spas as a type of health tourism associated with curative and/or preventive reasons, and also "the minimum requirements for a spa should include a swimming pool, sauna,

whirlpool or steam bath, facilities for physical fitness, relaxation, mental activity and healthy nutrition; there should also be at least one wellness professional (doctor, physiotherapist, wellness trainer or sports instructor) present to provide individual care and advice" (Dini and Pencarelli, 2021: 401).

- **Care of body and mind:** Physical fitness and beauty care are components of the care of the body, and also meditation, mental activity and education are components of the mind (Plzáková and Stupková, 2019). Dini and Pencarelli (2021: 401), opined that these "components - often tied to that of spas, but it can be proposed as a stand-alone offer by hotels or centres specialising in wellness services" and where it can be included "sauna and beauty service", "fitness", "cosmetics", "massages" (physical therapy and osteopathy), "rhythmic movement therapy (RMT)", "gyms", among others.

- **Medical tourism:** Health tourism is the combination of medical tourism and wellness tourism (Plzáková and Stupková, 2019) "Medical tourism" is a "healthy" component and is subdivided into three subcategories (Dini and Pencarelli, 2021):

  I. **Medical surgical** – "involving some type of operation/surgical procedure, e.g., for "tourists with cancer and other chronic diseases" (Dini and Pencarelli, 2021: 401),

  II. **Medical therapeutic** – "similar to surgical, but the therapy is repeated over time and requires a longer stay, e.g., nursing services for the elderly "aged nursing" (Dini and Pencarelli, 2021: 401), and

  III. **Medical wellness** – "services are provided by specialised medical personnel, but the aim is not to cure specific pathologies (e.g., "babymoon destination", "healthcare and wellness tourism" or "Ayurveda", as referred by Dini and Pencarelli (2021: 401).

Brandão *et al.* (2021), believes that health tourism integrates three dimensions: therapeutic or curative, preventive, and recovery or rehabilitation.

- **Natural environment:** The components of wellness tourism include: "nature" and the "environment" which can be "rural" or "therapeutic landscapes" or natural landscapes, 'for rural destinations can combine general well-being services with touristic resources already present and generate "social value". Also, the destinations included "coastal areas", "lakes", "adventure tourism/ecotourism", and an "arboretum", a place where trees, shrubs, and herbaceous plants are cultivated for scientific and educational purposes (Dini and Pencarelli, 2021).

- **Spirituality:** The "spiritual" experiences are also acknowledged as an important component of wellness tourism (Dini and Pencarelli, 2021; Zhang, et al., 2021; Han et al., 2020; Pan et al., 2019; Plzáková and Stupková, 2019; Stará and Peterson, 2017; Valentine (2016). This group, focus on the search for spirituality, by spending one's free time exploring mystical and religious experiences. In contemporary society, which is undergoing an ever more pronounced "spiritual revolution," there are growing numbers of tourists, looking for spiritual experiences via different meditative paths such as "Yoga" or "New Age" (Dini and Pencarelli, 2021).

- **Culture:** This type of component generally refers to a destination's cultural intangible heritage in form of "local culture" and "local way of life". It is an emerging sector component, as evidenced by recent studies highlighting the links between wellness tourism and culture (Dini and Pencarelli, 2021).

- **Enogastronomy:** Madeira, Correia and Filipe (2019: 8), describe in detail the elements that define enogastronomic experience as "the food and beverages products, eating practices, the art and customs of preparing and eating, the origins of food, the sensory elements (taste, smell, touch, look), the food preparation (techniques of cooking), the origins of food (organic food, ethnic cuisines, locally produced food, and so on.), forms of serving (fast food, slow food, street food, and so on.) and the context in which it is served and consumed (restaurants, bars, markets, food quarters, streets, and so on)" In according with Dini and Pencarelli (2021: 402), this "experiences are becoming a more significant experience component of wellness tourism products in the food and beverage enterprises".

- **Sports:** The sports experience "includes indoor and outdoor "sports tourism" activities designed to encourage the active participation of tourists. They range from "sports and recovery" to "horse riding" and "golf" (Dini and Pencarelli, 2021: 402).

- **Events**: This type of experience and offer is still considered a marginal component of the current wellness tourism scenario, "festivals and events" can positively contribute to tourists' well-being from a holistic perspective" (Dini and Pencarelli, 2021: 402).

Wellness tourism destination relates to a healthy way of life adoption. Because of the effects on the economy and globalisation, it implies an expanding tourism niche

encompassing individual or group travel to specialised resorts and destinations with the purpose of physical and mental health maintenance. Dini and Pencarelli (2021: 399-400) have ten categories, and words associated with each type, presented in Table 1.

Table 2.1- Wellness tourism dimensions/categories and words associated

Categories	Words associated	Sources
Hot springs	Mineral and thermal waters; Hot springs establishments/hotel; Thermal springs/centre/complex/; Health preservation bath.	Dini and Pencarelli (2021); Rodrigues *et al.* (2019); Brandão *et al.* (2021).
Spas	Hotels-spa; Spas; Spa tourism; Spas destination; Luxury spa.	Lemi (2021); Brandão *et al.* (2021); Dini and Pencarelli (2021).
Care of body and mind	Sauna and beauty service; Fitness; Cosmetics; Massages; Rhythmic movement therapy (RMT); Gyms.	Dini and Pencarelli (2021); Plzáková, and Stupková (2019).
Medical tourism	Medical tourism; Healthy; Tourists with cancer and other chronic diseases; Aged nursing; Babymoon destination; Healthcare and wellness tourism.	Brandão *et al.* (2021); Dini and Pencarelli (2021); Plzáková, and Stupková (2019).
Natural environment	Nature; Environment; Rural; Therapeutic landscapes; Social value; Coastal areas; Lakes; Adventure tourism/ecotourism; Arboretum.	Dini and Pencarelli (2021).
Spirituality	Spiritual; Spiritual revolution; Yoga; New age.	Dini and Pencarelli (2021); Zhang *et al.* (2021); Pan *et al.* (2019); Plzáková, and Stupková (2019); Stará and Peterson (2017); Valentine (2016).
Culture	China wellness culture; Longevity in Chinese culture; local culture; Local way of life.	Dini and Pencarelli (2021).
Enogastronomy	Food quality and/or food service fresh and local; Nutritional foods; Healthy food; Indigenous food; Dietary.	Dini and Pencarelli (2021); Madeira, Correia, and Filipe, (2019).
Sports	Sports tourism; Sports and recovery; Horse riding; Golf.	Dini and Pencarelli (2021).
Events	Festival and events.	Dini and Pencarelli (2021).

Source: Adapted from Dini and Pencarelli (2021: 339-340).

In travel decision-making and tourist behaviour tracking, User-generated content (UGC), comprising narrative text and photographs, becomes a means for collecting travelling information for potential tourists, and it is also an important channel for destination image perceptions (Huang, Han, Meng, Zeng, and Liao, 2022).

Wellness destination consumer searches specifically for lodging facilities that range from spa destinations to traditional hotels and resorts, and spending is on various products and services such as wellness tourism products, food and beverages, shopping, travelling, sightseeing, going on family vacations, and joining recreational activities, etc. (Rodrigues, Brochado, Troilo, & Mohsin, 2022). User-generated content (UGC), which is created and gives the opinion or semantic orientation is usually weighted heavily by tourists as more trusted and reliable travel information (Huang et al., 2022; Rodrigues et al., 2022).

2.3. Sentiment analysis in User Generated-Content (UGC)

Kirilenko, Stepchenkova, Kim, and Li (2017: 1013-1014) define "Sentiment analysis or opinion mining is the computational study of people's opinions, appraisals, attitudes, and emotions toward entities, individuals, issues, events, topics and their attributes". According to Alaei, Becken, and Stantic (2017) it aims to determine the overall contextual polarity of a text document, a review, an opinion, or an emotion expressed in online user general content UGC, whereby polarity can be positive, neutral, or negative - Semantic orientation. The information provided through text can be subjective, objective or neutral. To Alaei *et al.* (2017), subjective reviews are based on opinions, personal feelings, beliefs, and judgment about entities or events, while objective reviews are based on facts, evidence, and measurable observations.

In the view of Mukherjee and Bhattacharyya (2013), objective (facts), positive (denotes a state of happiness, bliss or satisfaction on part of the writer) or negative (denotes a state of sorrow, dejection or disappointment on part of the writer). In conceptualising, Alaei *et al.* (2017), classify sentiment polarity into binary, ternary, or ordinal, so based on customer subjective review, the binary is assumed to be either positive or negative and also it depends on where it is applicable and the domain. Since customer reviews will not

17

always be subjective, ternary contains the third category neutral, which gives the difference between subjective and objective sentences.

Ordinal classification of sentiments can further be given a score based on their degree of positivity, negativity or objectivity. Since the sentiment analysis dataset's class range is involved, it should not be limited to positive or negative. Agreed or disagreed, good or bad can also be used, and can also be quantified on a 5-point scale (strongly disagree, disagree, neutral, agree, or strongly agree) (Nandwanj and Verma, 2021; Alaei *et al.* 2017). Sentiment Analysis identifies what people like and dislike and helps build recommendation systems and targeted marketing campaigns (Muthukrishnan, *et al.,* 2021).

Sentiment analysis can be divided into single-modal sentiment analysis and multi-modal sentiment analysis (Guo, Yu, and Wang, 2021). In multi-modal data sentiment analysis, the content may have multiple forms such as text, video, pictures, etc. Multi-modal data can complement each other, but how to integrate multi-modal data and use the alignment information between different modal data to model the association between different modal data still faces great challenges (Guo *et al.*, 2021), and there are four types of sentiment analysis (Muthukrishnan *et al.,* 2021); fine-grained sentiment, emotion detection sentiment analysis, aspect-based sentiment analysis and intent sentiment analysis.

- **Fine-Grained Sentiment Analysis**, analyses sentiment in each of the sentence parts. It is used on a sub-sentence level, and it is intended to identify the focus of a feeling. In fine-grained sentiment, a particular sentence is broken into phrases or clauses, and connectivity in each part is analysed with the others (Altexsoft, 2018). According to Altexsoft (2018) and Muthukrishnan *et al.* (2021) fine-grained sentiment analysis gives an understanding of the response from the users of the product or service. Also, it describes sentiment intensity with the use of very positive and very negative categories, apart from the usual standard of three sentiment scores (negative, neutral, and positive) (Altexsoft, 2018). Though, the process of understanding the concept can be complex and cost intensive as compared to other types (Muthukrishnan *et al.,* 2021).

- **Emotion Detection Sentiment Analysis** is a means of identifying the emotion in a part of text (Muthukrishnan *et al.*, 2021). There are three types of sentiment and

emotion analysis techniques: lexicon based, machine learning-based, and deep learning-based (Nandwanj *et al.,* 2021). Lexicons and machine learning are used to determine the sentiment (Muthukrishnan *et al.*, 2021) and deep learning employed multiple processing layers to learn hierarchical representations of data (Young, Hazarika, Poria, and Cambria, 2018). Lexicons are lists of words or lexical features (Bonta, Kumaresh, and Janardhan, 2019) that are either positive or negative (Muthukrishnan *et al.,* 2021). This makes it easier to segregate the terms according to their sentiment (Muthukrishnan *et al*., 2021), although may fail for compositional methods to obtain the sentiment of larger text units, such as phrases and sentences (Li, Lu and Long, 2017). To Bonta *et al.* (2019), the manual approach, lexical or dictionary-based approach and corpus-based approach are the three types of lexicon-based approaches based on sentiment lexicon.

i.   <u>Manual approach</u> is the task of reading a pre-selected document for sentiment analysis and providing additional information in the pre-process and normalising the text which is to be analysed by manually removing punctuation marks and special characters, removal of stop words like "a", "an"," and", "for", etc, expansion of word contractions. E.g.," They'll" as "They will"; "Haven't" as "Have not" etc, and Stemming and Lemmatization – Converting a word into its base form. E.g., apples → apple; happiness → happy, among others. The process is time-consuming (Neves and Seva, 2021; Taj, Shaikh, and Meghji, 2019), it has a limited size of data (Li *et al.*, 2017), but it is also straightforwardly used for machine learning purposes (Neves and Seva, 2021).

ii.   <u>Lexical or dictionary-based</u> use a word stock dictionary with opinion words and match given set of words in a text for finding polarity, that is, each word in the dictionary is labelled as positive, negative, neutral sentiments along with polarity, parts of speech and subjectivity classifiers, mood, modality and a sentence is tokenised. Each token is matched with the available words in the model to find out its context and sentiment. The final prediction regarding the total text component is a function of the sum or average taken. The approach does not need to pre-process data, nor does it have to train a classifier (Taj *et al.*, 2019).

iii.   <u>Corpus-based</u> approach typically trains classifiers using existing labelled corpus (supervised classification) (Biltawi, Al-Naymat, and Tedmori, 2017).

- **Aspect-based sentiment analysis** this type of sentiment analysis is usually for one aspect of a service or product. Aspect extraction is based on the notion that an opinion aimed at no target (i.e. aspect) is of limited use. Thus, the goal of aspect-based sentiment analysis is to discover sentiments and the elements that those sentiments are aimed at (Alqaryouti, Siyam, Monem., and Shaalan, 2020). For example, a store that sells televisions can be used this type of sentiment analysis, and it could be for one aspect of televisions – such as brightness, or sound, among others. So, they can understand how customers feel about specific attributes of the product. Muthukrishnan *et al.* (2021) found that aspect-based sentiment analysis can be done through aspect identification and sentiment classification of the identified aspects into positive or negative.

  Under the aspect identification:

  i. <u>Implicit aspect</u> is difficult to identify (Muthukrishnan *et al.*, 2021; Rana, Cheah and Rana, 2020). The aspects are not expressed explicitly by the users for example: "The phone is light and small". The two aspects (targets) related to the phone, "weight" and "size", but no explicit terms were used to express these aspects. Though, "phone" is the only term related to both opinions "light" and "small" and can help to understand the nature of the sentence. So, due to the complex nature of the implicit aspect, few studies were done on it (Rana *et al.*, 2020).

  ii. <u>Co-referential aspects</u> are difficult to identify. When two or more expressions in a text refer to the same person or thing, that is, having the same synonym (Dictionary.com). For example, ecosystem and biosphere are co-referential because both refer to the environment (Muthukrishnan *et al.*, 2021).

  iii. <u>Infrequent features</u> that are difficult to identify because they are not used by many reviewers (Muthukrishnan *et al.*, 2021).

  In sentiment classification of the identified aspects, as presented in Figure 2.1, has to do with the classification of the multiple aspects reviews where information is gathered from online review, Twitter, Facebook or any other blog that provide review on wellness destination.

- **Intent Sentiment Analysis** is all about understanding the mindset of a customer on a product or service; the intention of a customer for contacting the company

on products or services. For example, a company can guess customer intends to use a product or not. It means the intentions of that exact customer can be traced, forming a pattern which is used for marketing and advertising (Muthukrishnan *et al.*, 2021), as presented in Figure 2.2.

Figure 2.1 - Process of opinion mining in multiple aspects
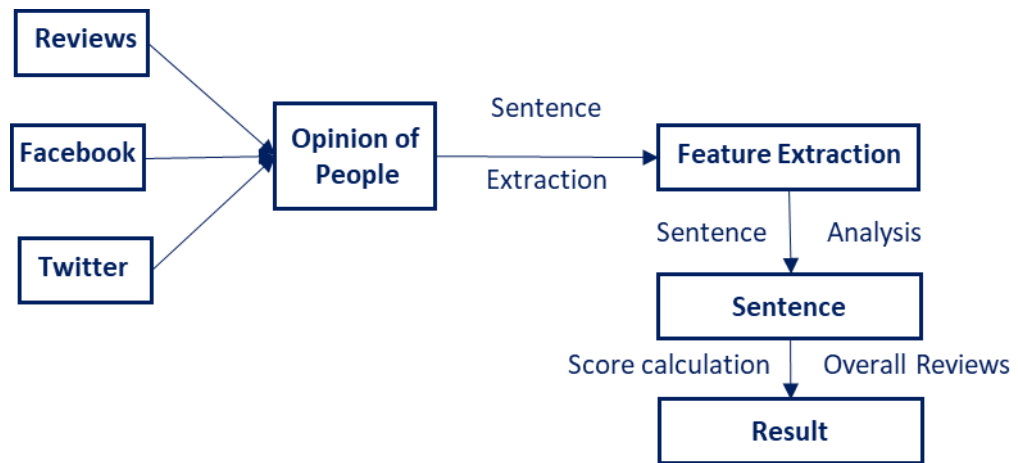
Source: Muthukrishnan *et al.* (2021: 5185).

Figure 2.2 - Semantic pattern model
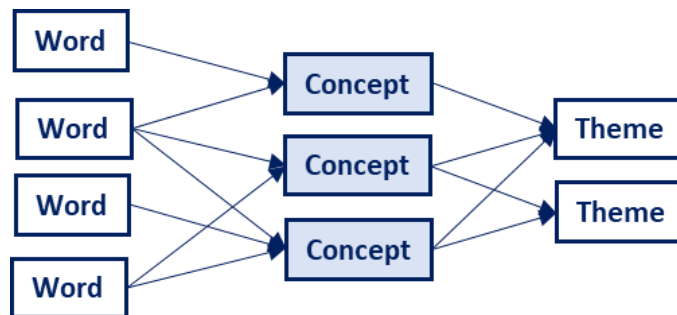
Source: Rodrigues, *et al.* (2020: 652)

To identify the key drivers of wellness tourism satisfaction and dissatisfaction, (Rodrigues, *et al.*, 2020) used a software called Leximancer feature "Sentiment Lens". The sentiment Lens will only apply the sentiment terms that identified words that reveal favourable sentiment terms and unfavourable or negative sentiment terms. For example, the phrase, "the service is unacceptable," is a negative description, sentiment analysis can provide further insights by organizing tourists' descriptions in terms of negative or

positive appraisals as relevant and used consistently within all the reviews (Rodrigues, *et al.*, 2020).

So, sentiment analysis is a method that seeks to identify the sentiment or opinion that an author of a comment or document, by obtaining the polarity associated with the written text (Rodrigues, *et al.*, 2020).

2.4. Present study and conceptual research model

In the introduction and the theoretical review, the destination concept is made up of units of attraction, whereby each unit can be categorised into a single service unit or multiple service units.

The present study aims to explore the Tourist UGC as related to Table 2.1 in the case of a wellness tourism destination and the usage of sentiment analysis to bring together the various views on wellness tourism by exploring the under-listed stated objectives of the study:

   I.   To analyse the feeling about wellness tourism destinations in reviews.

  II.   To evaluate the opinion of tourists based on destination review, which is subdivided into specific objectives:

- Identify the components that characterise a wellness tourist destination.
- Collect comments posted on a social network to form the sample.
- Apply the different pre-processing (cleaning) techniques to the database.
- Apply text mining algorithms to assess the sentiment expressed in the comments.
- Considering the sentiment analysis results, identify the motivations that lead tourists to look for wellness tourist destinations in Algarve.

The identified components in the Table 2.1 will form the structure of groups where opinions will be extracted through social networks (TripAdvisor) by the traditional manual system to ease the pre-processing and extraction of the necessary features. The pre-processing of the component unit gives a more precise idea of the destination, reviews the text to minimise errors, and progressively extracts the users' opinions to be positive or negative.

The sentiment analysis of the review text will be a single modal because it will involve text extractions from only single online review source (TripAdvisor), and machine learning will be used for the analysis as illustrated in Figure 2.3.

Figure 2.3 - Sentiment analysis process on product reviews
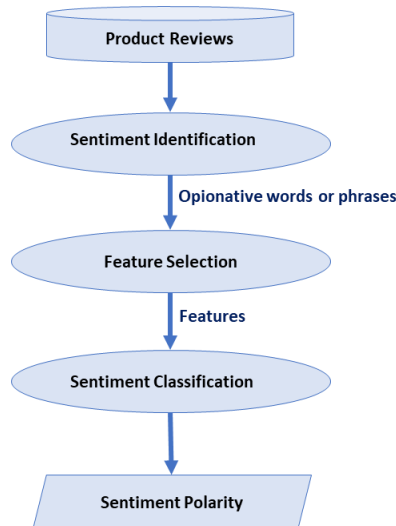
Source: Adapted from Medhat, Hassan, and Korashy (2014: 1094).

Social reviews are one of the generated content to analyse sentiments which serve as a source of information for travellers to wellness tourist destinations. Analysing the feeling about wellness tourism destinations in the reviews of the hotel, to evaluate the sentiment of tourist on the component that characterized a wellness tourist destination and identifying the motivations that lead tourists to look for wellness tourist destinations in the Algarve are the main aim of exploring tourist generated content. To achieve the objectives and to measure the sentiment expressed in the comments, text mining algorithms were applied to comments collected from a social media platform.

CHAPTER THREE - RESEARCH METHODOLOGY

3.1. Application of the method

Sentiment analysis can be applied for different purposes such as political campaigns and riots (El-Masri *et al.,* 2017), marketing, political and sociological (D'Andrea, Ferri, Grifoni, and Guzzo, 2015). However, it is most often applied to consumer reviews from different domains like movies, advertisements, products, cars, smartphones, tourism, and e-learning (El-Masri *et al.*, 2017), but with some limitations in tourism, which include among others: ability to obtain data, the methodology, and way results are interpreted, due to the study base and technologies used (Borrajo-Millán, Alonso-Almeida, Escat-Cortes, and Yi, 2021).

In the opinion of Borrajo-Millán *et al.*, (2021), researchers use a very limited number of tourism databases, which are namely, TripAdvisor (tool NLTK), Expedia, Booking, Airbnb, Twitter, Amazon, Facebook or Academic Yelp Dataset.

From limitations in the methodology, Borrajo-Millán *et al.*, (2021) cited the four most commonly used methodologies which are dictionary-based methods, machine learning methods, deep learning methods, and hybrids of previous methods and all required the use of multiple combined technologies, the use of large unstructured databases, and the difficulty in reduction and acquisition of results.

The third limitation is how results are interpreted, due to the study base and technologies used (Borrajo-Millán *et al.*, 2021).

There are five main steps to applying sentiment analysis (El-Masri *et al.*, 2017), Data Collection, Pre-processing, Feature extraction, Application of the sentiment analysis algorithm, Results evaluation, as presented below.

3.1.1 Data Collection

The data for this research work will be collected from domain-specified websites which is TripAdvisor, which is for tourism- related reviews. Although, the most common source

for collecting data is social media websites such as Facebook and Twitter (El-Masri *et al.,* 2017), blogs, forums, and social networks (Borrajo-Millán *et al.,* 2021).

Data collection was carried out between the months of February and March 2022 on the TripAdvisor platform and was collected manually, from TripAdvisor platform to an Excel file.

The region of Algarve comprises sixteen municipalities with unique attractions in each one. Algarve is the centre of the tourist attraction in Portugal and Europe, in general, because of the beaches, food, wines and ports, golfs, dolphin watching, Ponta da Piedade, Cabo de São Vicente, The Countryside, and Arts and Crafts (Wilson, 2019). The touristic image of Algarve was also boosted by the location of some of the best hotels in Portugal, like Epic Sana Algarve Hotel, Conrad Algarve, and Hotel Quinta do Lago.

The first objective of the research is the analysis of the sentiments about wellness tourism destinations in TripAdvisor reviews, and the reviews from Algarve destinations which were manually extracted, have a total of 1,294 reviews, as presented in Table 3.1. The reviews were categorised into Entity Number (*ID*), entity name (*Name*), city where is located (*City*), country (*Country*), the category in according to Table 2.1 (*Category*), number of stars (*Star*), number of reviews (*Number of reviews*)*,* which are represented in Table 3.1, and considered the rating attribute to the experience (*Review rating*) and the review content (*Full text*). The *ID* were manually arranged in an excel sheet for easy pre-processing.

These hotels and facilities were chosen because of their review rating, services, environment, and other additional tourist services provided by other complementary tourist organisations. In the case of Caldas de Monchique, which is rated 4.5, and provided Hot spring services, also there were about three to fourteen things to do in Monchique (Top picks Balloons), day trip and excursion, a full-day tour, sporting events, food, wine and nightlife, luxury and special occasions, ports of call tours, walking tours, museum ticket and passes, bus services and Jets boat rentals). Swan Day Spa has unique customer service from the point of stepping into the spa hotel; their services rejuvenate senses and were rated 5.0.

Zensation - Body & Soul was rated 4.5 and its location is in the centre of Albufeira; its massages and the styles of its spa using fish, express its unique service characteristics. Moinhos Velhos Juice Detox & Yoga Retreat was rated 4.5. In their yoga and wellness

services, they run juice fasting programme for spiritual edification as part of the health and wellness retreat.

Old Town Faro with rating 4.0 is not a hotel but was chosen because it shows the culture of the Algarve destination, and it enhances the ways of life by walking through the historic sites and cities. Pinetrees Horse Riding is a sporting activity, rated 5.0. The system renders the services of trail riding wish improved the wellness of riders.

Zoomarine Algarve, with a rating of 4.5, is a theme park with multiple events that promotes wellness of the whole family. It has an array of aquatic amusements and mechanical rides, including a mini-roller coaster, carousel, giant wheel, and so much more. Mercado de Loulé, with a rating of 4.5, is a wellness destination for enogastronomy with other complementary stalls like gift and craft stalls, and fresh fish vegetable stalls.

Tabela 3.1 - Algarve destination TripAdvisor

ID	Name	City	Country	Category	Star	No. of Reviews
1	Caldas de Monchique	Faro	Portugal	Hot springs	4.0	156
2	Zensation - Body & Soul	Albufeira	Portugal	Care of body and mind	4.5	142
3	Mercado de Loulé	Loule	Portugal	Enogastronomy	4.5	160
4	Moinhos Velhos Juice Detox & Yoga Retreat	Lagos	Portugal	Spirituality	4.5	50
5	Old Town Faro	Faro	Portugal	Culture	4.0	160
6	Pinetrees Horse Riding	Almancil	Portugal	Sport	5.0	156
7	Ria Formosa Natural Park	Abufeira	Portugal	Natural environment	4.5	160
8	Swan Day Spa	Alvor	Portugal	Spas	5.0	160
9	Zoomarine Algarve	Albufeira	Portugal	Event	4.5	150
					Total	1,294

Source: Author's elaboration.

The present selection of hotels and facilities were considered to analyse a representative entity of each category of health and wellness presented in Table 2.1, according to the literature review, to characterize a health and wellness destination.

3.1.2. Pre-processing

This involved cleaning the data from unwanted elements which is the process of preparing text data so that machines can use the same to perform tasks like analysis (Khanna, 2021), predictions and the pre-processing is as follows (El-Masri *et al.,* 2017):

I.   **Tokenization:** is the process of breaking a stream of text into meaningful words (stems), phrases, characters or symbols (El-Masri *et al.*2017; Vishwanathan, 2014).

II. **Remove stop words:** These are the most common words in any language (like articles, prepositions, pronouns, conjunctions, etc) and do not add much information to the text. Examples of a few stop words in English are "the", "a", "an", "so", "what" but the removal of stop words is highly dependent on the task we are performing and the goal we want to achieve (Khanna, 2021).

III. **Remove or identify punctuation**: This is the removal of punctuation such as full stops and commas as they are irrelevant to the polarity detection (El-Masri *et al.*, 2017). There are 14 punctuation marks listed in English grammar**.** They are the period (full stop), question mark, exclamation point/mark, comma, semicolon, colon, dash, hyphen, parentheses, brackets, braces, apostrophe, quotation marks, and ellipses (Mahdu, 2021).

IV. **Convert text to lower or upper case**: Converting the letters into upper or lower case. If the word is all in capitals, it sometimes suggests strong sentiment/emotion (El-Masri *et al.*, 2017), and it helps to maintain the consistent flow during the natural language processing (NLP) tasks and text mining. The lower function makes the whole process quite straightforward (Singhal, 2020).

V.   **Stemming the word:** This is the process of reducing a word to its word stem that affixes to suffixes and prefixes or the roots of words, known as a lemma. Stemming is important in natural language understanding (NLU) and (NLP) (TechTarget Contributor, 2018).

VI. **Check the Spelling:** Data from social media and reviews contain many spelling mistakes, such as missing letters or extra letters. Spelling can be corrected by removing extra letters also NLP corrected spelling mistakes manually (El-Masri *et al.*, 2017).

The pre-processing with the use of text mining, involved: transformation, tokenization, normalization and filtration. The words were transformed into lower case, especially

words in a phrase or a sentence. Tokenization of the full-text by breaking the stream of words, phrases, and symbols and removal of punctuations (El-Masri *et al.,* 2017; Vishwanathan, 2014). The stop words, numbers and regular expressions were filtered, and the text was normalised by removing prefixes or suffixes in word sentences.

The pre-processing of data, through the corpus, leads to key extractions (See Table 3.2) of some cloud of words, which are peculiar to hotels or facilities in the Algarve destination and which express sentiment. It also identifies a cloud of words and exact keywords in each of the nine categories in the Algarve destination.

Table 3.2 - Popular hotels and its category visited by tourists.

Entity	Caldas de Monchique	Swan Day Spa	Zensation - Body & Soul		Ria Formosa Natural Park	Moinhos Velhos Juice Detox & Yoga Retreat	Old Town Faro	Mercado de Loulé	Pinetrees Horse Riding	Zoomarine Algarve
Category	Hot springs	Spas	Care of body and mind	Medical Tourism	Natural environment	Spirituality	Culture	Enogastronomy	Sport	Event
thermal	11									
health	30									
spa		248								
luxury		4								
massage			199							
sauna			18							
healthy				10						
medical				5						
nature					55					
landscape					9					
spiritual						9				
yoga						35				
culture							8			
wellness							2			
food								127		
nutrition								9		
sport									6	
golf									5	
festival										4
event										9
Count	41	252	217	15	64	44	10	136	11	13

Source: Author's elaboration.

The data set which represents *ID* are the components categories of: Hot springs, Spas, Care of body and mind, Medical tourism, Natural environment, Spirituality, Culture, Enogastronomy, Sports, and Events, though, there are no generated reviews for ID for Medical tourism.

3.1.3. Feature extraction

This is known as the process of deriving a set of discriminative, informative and non-redundant values to numerically represent a review or text (Alaei, *et al.,* 2017). Still, in this research, the domain features, semantic (positive/negative) features will be used (El-Masri *et al.*, 2017).

3.1.4. Application of the sentiment analysis algorithm

Sentiment analysis, in Orange Data Mining, can be operationalised from different sentiment classification models. However, before the analysis, the following algorithms were considered according to Hutto and Gilbert (2014):

- Naive Bayes (NB),
- Maximum Entropy (MaxEnt, or ME),
- Support Vector Machines (SVMs),
- SVM-Classification,
- SVM-Regression (SVM-R) models.

Naive Bayes (NB) (Hutto and Gilbert, 2014), NB classifier is a simple classifier that relies on Bayesian probability and the naive assumption that feature probabilities are independent of one another.

Maximum Entropy (MaxEnt, or ME), It is a general-purpose machine learning technique belonging to the class of exponential models using multinomial logistic regression.

Support Vector Machines (SVMs) are non-probability classifiers which operate by separating data points in space using one or more hyperplanes (centrelines of the gaps separating different classes).

SVM-Classification is a fast and dependable classification algorithm that performs very well with a limited amount of data to analyse.

SVM-Regression (SVM-R), Support Vector Regression, is a supervised learning algorithm that predicts discrete values. Support Vector Regression uses the same principle as the SVMs. The basic idea behind SVR is to find the best fit line. In SVR, the best fit line is the hyperplane that has the maximum number of points (Raj, 2020).

According to Hutto and Gilbert (2014), the difference to NB is that in ME, there is no restricted independence statement between features and thereby accounts for information entropy (feature weightings). Support Vector Machines (SVMs) differ from both NB and ME models in that SVMs are non-probability classifiers which operate by separating data points in space using one or more hyperplanes (centrelines of the gaps separating different classes). We use the Python-based machine learning algorithms from scikit-learn.org for the NB, ME, SVM-Classification (SVM-C) and SVM-Regression (SVM-R) models. The final workflow elaborated in Orange is presented in the Appendix B, which include the steps considered in the methlogy.

3.1.5. Evaluation of Results

The sentiment analyses were also analysed with two dimensions of words extracted from Words associated to each category (Table 2.1). The words were compared with the corresponding *ID* and *Full text* comments in the sentiment result worksheet to determine opinion for each of the hotels and facilities. Based on the choices of word categorisation from the Algarve destination tourist in Table 3.2, it clearly shows the motivations for the selection of hotels or facilities by tourists, which results were presented in the next sections.

CHAPTER FOUR - RESULTS

4.1. Data description and Sample characterization

The empirical analysis employed for this research is the reviews of the TripAdvisor on Algarve destination. The destination hotels are ten categories with words associated with each category, as presented in Table 2.1. A total of 1,294 reviews was manually extracted from nine categories and medical tourism has no reviews within the period the researcher collect data from TripAdvisor, which is shown in Table 3.1.

The objectives of the research stated were as follows:

I.   To analyse the sentiments about wellness tourism destinations in hotel and facilities reviews.

II.   To evaluate the opinion of tourists based on destination review, which is subdivided into specific objectives:

- Identify the components that characterise a wellness tourist destination.

- Collect comments posted on a social network to form the sample.

- Apply the different pre-processing (cleaning) techniques to the database.

- Apply text mining algorithms to assess the sentiment expressed in the comments.

- Considering the sentiment analysis results, identify the motivations that lead tourists to look for wellness tourist destinations in Algarve.

To analyse the feeling about wellness tourism destinations in the reviews, Figure 4.1 and Figure 4.2 and Table 3.1 and Table 4.1, presents the sentiments about the full text. Also, the evaluating of the opinions based on destination reviews can be analysed from Graph 4.3 to Graph 4.13, also considering the content of Table 3.2.

The Figure 4.1 shows the relation between the average of the review rating and the average of sentiment normalised by country, associated with all comments collected on TripAdvisor. Also, Figure 4.3 presents the words that appear most in the reviews.

Figure 4.1 - Sentiment by author country
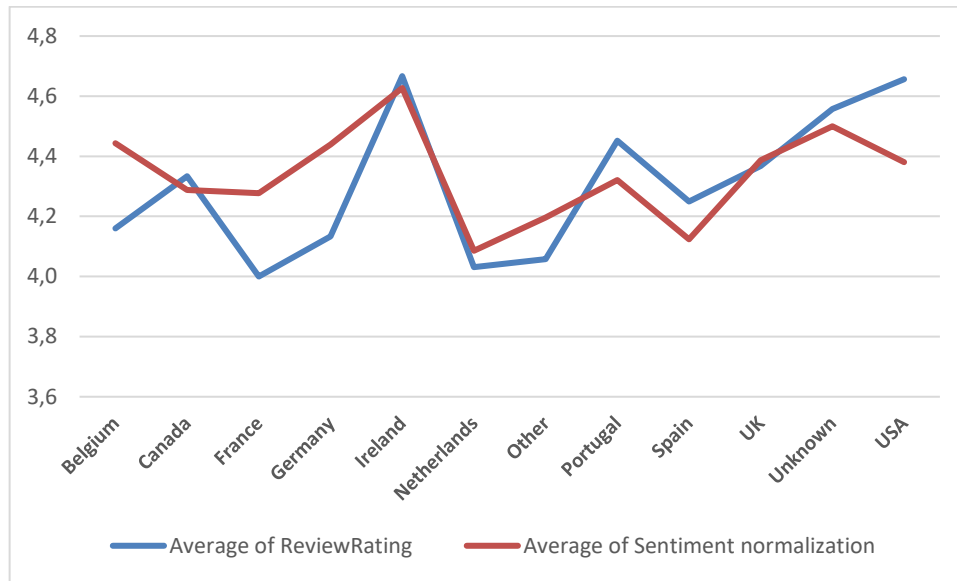
Source: Author's elaboration.

However, the sentiment value obtained by the Vader algorithm (Hutto and Gilbert, 2014) presents a result between -1 and 1, which is difficult to compare with the ratings expressed in the reviews, which are defined between 1 and 5, so we opted to normalize the values obtained from the sentiment. expressed on the same scale as ratings to facilitate analysis and comparison of values. To carry out the normalization, the formula presented by Chen, Xu, Zheng, Yu, & Luo (2020: 9) was considered, and presented in equation 1.

$$x'_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{1}$$

The Graph 4.2 shows the relation between the average of the review rating and the average of sentiment normalised by hotels and facilities, associated to all comments collected in TripAdvisor. Also, Figure 4.3 presents the words that appear most in the reviews.
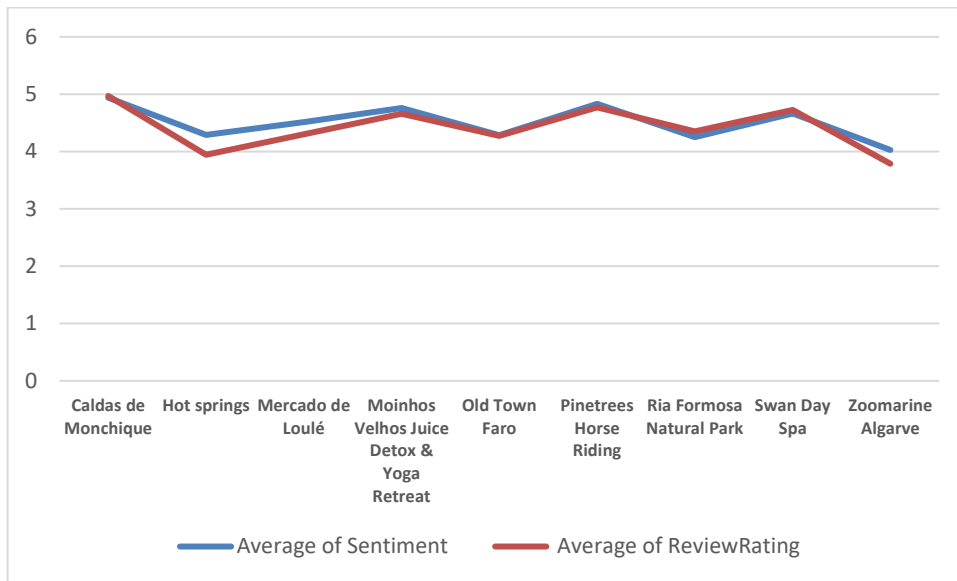
Figure 4.2 - Sentiment by Hotel and facilities

Source: Author's elaboration.

Figure 4.3 - A word cloud of all comments

Source: Author's elaboration.

In general and taking into consideration the word cloud presented in the Figure 4.3, the words that most represent the Algarve as a wellness destination is "love", "great", "visit".

4.2. Sentiment score by wellness tourism category

Taken into consideration the categories presented in Table 3.1, and the data collected, it is possible to analyse Table 4.1, where the evaluation is presented by the wellness tourism category entity represented and in Table 4.2 wellness tourism category.

Table 4.1 - Sentiment comments and keyword analysis about the wellness tourism destination

ID	Average ReviewRating	Average of sentiment normalised. (1-5)	Sentiment Average (-1 to +1)	Keywords of positive comments	Positive words most used in negative sentiments	Negative words most used in negative sentiments
1	3.94	4.29	0.67	Nice Place spa	Spa Place Nice	Place Spa water
2	4.97	4.94	0.93	Massage Spa Relax	Massage Spa Relax	Massage Spa Relax
3	4.73	4.66	0.82	Fish Massage Spa	Fish Massage Spa	Massage Back time
4	4.35	4.25	0.67	Park Bird Walk	Park Bird Walk	See Walk Visit
5	4.66	4.76	0.89	Experience Time Great	Experience Great Place	Staff Time
6	4.28	4.28	0.67	Old Town Faro	Old Town Faro	Town Old Faro
7	4.77	4.83	0.90	Ride Horse Pinetree	Ride Horse Pinetree	Day Ride Park
8	3.79	4.03	0.59	Day Ride Park	Day Ride Great	Ride People Park
9	4.3	4.51	0.75	Market Fish Visit	Market Fish Visit	Market Visit Local
Total	4.40	4..48	0.76	Visit love great	Love Great Visit	Time Place Visit

Source: Author's elaboration.

Table 4.1 illustrates the statistics of keywords of positive comments, positive words most used in negative sentiments, and negative words most used in negative feelings in the wellness destinations. The keywords extracted from positive comments associated to all

wellness tourism categories are: "bird", "day", "experience", "Faro", "fish", "great", "horse", "love", "market", "massage", "nice", "old", "park", "Pinetree", "place", "relax", "ride", "spa", "time", "town", "visit" and "walk".

In *ID=1*, most tourists describe the destination as being a nice place, and the spa facilities are also nice, but not too special. And some of the few negative sentiments were the issue of the spring not flowing.

In *ID=2*, keywords are massage, spa and good for relaxing. In *ID=3*, the destination positive comment is on the unique fish massage spa, and the negative comment is for the timing for the pedicure and massage.

The impression of the tourists for *ID=4* was that the facilities were just like a garden with the sound of birds and also a nice place for relaxation, but not much to see or visit in the destination. The tourist's experience was nice and great, but the negative comments on the attitude of *ID=5* staff.

And on *ID=6* positive keywords old town and Faro were frequently used. The tourist in *ID=7* enjoyed horse riding in Pinetrees.

In *ID*=8, facilities provide a satisfying day ride, but there were too many people every day to ride in the park. In ID=9, negative comments were on the market being local, but it is a market for different types of fresh fish, and it is worth visiting.

Table 4.2, with a standard deviation of 0.8, and 0.6 with an average rating of 4.6 and 4.7 respectively for spas and care of body and mind, shows that a lot of sentiment was attached to spas, luxury, massage and sauna. It clearly expressed the acceptability and popularity to tourists of wellness destinations with spas and care for body and mind. Also, natural environment and sport, with a standard deviation of 0.4 and 0.3, with an average of 4.7, show that wellness destinations with sport (golf), nature and landscape are acceptable to tourists as a wellness destination.

The author's sentiment *Full text, p*resented in Figure 4.1 and 4.2, with a total of 1,294, as shown in the Table 4.1 comments review, 24% of the total is a positive comment, while 3% of the sentiment comments are negative. The *ID* 1,2,3,6,7 and,8, has more excellent reviews because they might have had good services previously. Still, the change in service experience due to covid experience influences their sentiment review rating vis-à-vis average positive and negative sentiments (Figure 4.2 and Figure 4.5).

Table 4.2 - Wellness tourism category and keywords about wellness tourism destination

| Category	Features	No of reviews	Minimum	maximum	Average	Standard Deviation
Hot springs	Thermal health	41	0.3	5	4.5	1.1
Spas	spa luxury	251	0.6	5	4.6	0.8
Care of body and mind	massage sauna	210	0.2	5	4.7	0.6
Medical Tourism	healthy medical	14	0.1	5	3.9	2
Natural environment	nature landscape	63	2.7	5	4.7	0.4
Spirituality	spiritual yoga	36	0.1	5	4.7	1.1
Culture	culture welness	10	0.3	5	4.1	1.7
Enogastronomy	food nutrition	135	0.1	5	4.3	1.1
Sport	sport golf	11	4.2	5	4.7	0.3
Event	festival event	13	0.4	5	3.2	1.9

Source: Author's elaboration.

In relating the *ReviewRating* with positive and negative sentiment, (Figure 4.2) illustrated *ReviewRating* are greater than positive and negative sentiment. Also, comparing sentiment normalisation with positive and negative sentiment, Figure 4.7 illustrated hotels yearly average reviews with Zoomarine Algarve reviews was the highest between 2019 and 2022. Reviews were at their highest during the winter months of January, February, March, November and December (Figure 4.4).

The Table 3.2 illustrates the opinions of the tourist on each of the hotels. Hot spring with categories of words, thermal (11) and health (30), in total 41. Spas categories words are spas (248) and luxury (4), in total 252. Care of body and mind categories are massage (199) and sauna (18), in total 217. Healthy (10) and medical (5) for Medical tourism, in total 15. Natural environment were nature (55) and landscape (9), in total 64. Spirituality has spiritual (9) and yoga (35), in total 44. Culture categories were culture (8) and wellness (2), in total 10. Enogastronomy with food (127) and nutrition (9), in total (136). Sport has total of 11, from sport (6) and golf (5). And Event has festival (4) and event (9), in total 13. Also, Table 4.1 identifies the frequency-inverse document frequency (TF-IDF) of the motivating words, used by the reviewers of the nine hotels in the Algarve

destination. In all the nine hotel and facilities reviews, the reviewer frequently used the words great, love and visit. These TF-IDF words are motivating words which explicitly express positive sentiment to the Algarve destination.

In consideration to the keywords of positive comments of the reviewers (Table 4.1), (Rodrigues, *et al.,* 2020) reiterated that writers and speakers of some topics and attitudes are determined by sentiment. Also, (Żemła, 2016; Jovicic, 2019) definition of a tourist destination, the Algarve destination has natural and man-made features which attract non-local visitors as illustrated in Table 3.2. It has helped in maintaining well-being and lifestyle improvement (Han *et al.*, 2020), also to support the view of Thal and Hudson (2017) from the reviews, the relationship between staff and guests has created a psychologically beneficial environment and also, optimal physical, mental, and social well-being (Pan, *et al.,* 2019). The table keyword positive comments such as nice, place, great, experience, etc, which simply means Algarve wellness destinations worth to re-visit. Some demotivating factors were discovered when extracting reviews manually, which include hidden charges and time spent in accessing the destination area or/ and accessing the tourism facilities due to high volume of patronage at a time, attitudes of destination staff, selective service delivery, no standard rules, no correlation between the service advertised and service delivery, according to Krakover and Corsale (2021), it affects time and money budgets and also discourage revisiting to the wellness destination.

Figure 4.4 - Destination Sentiments versus year with winters effect
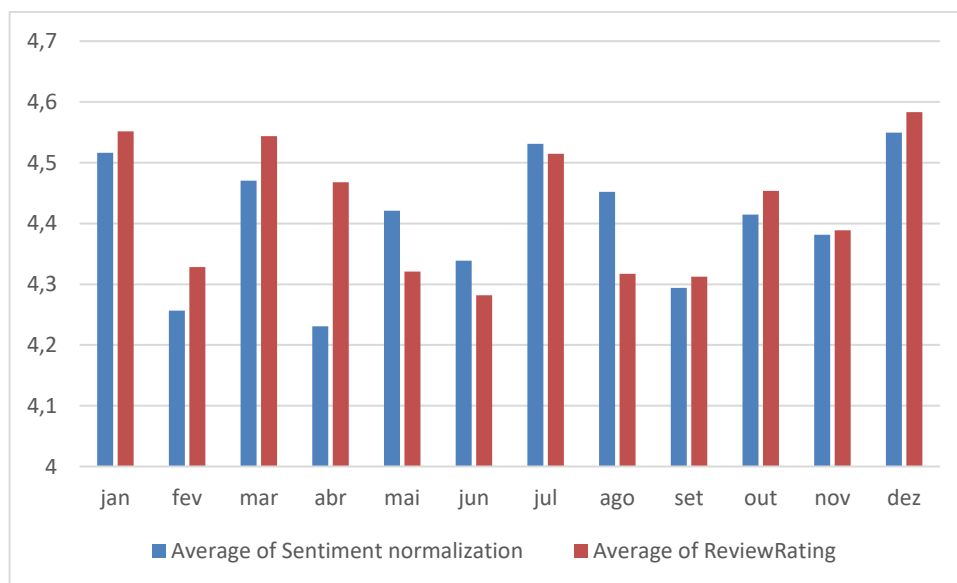
Source: Author's elaboration.

Figure 4.5 - sentiment with covid effects
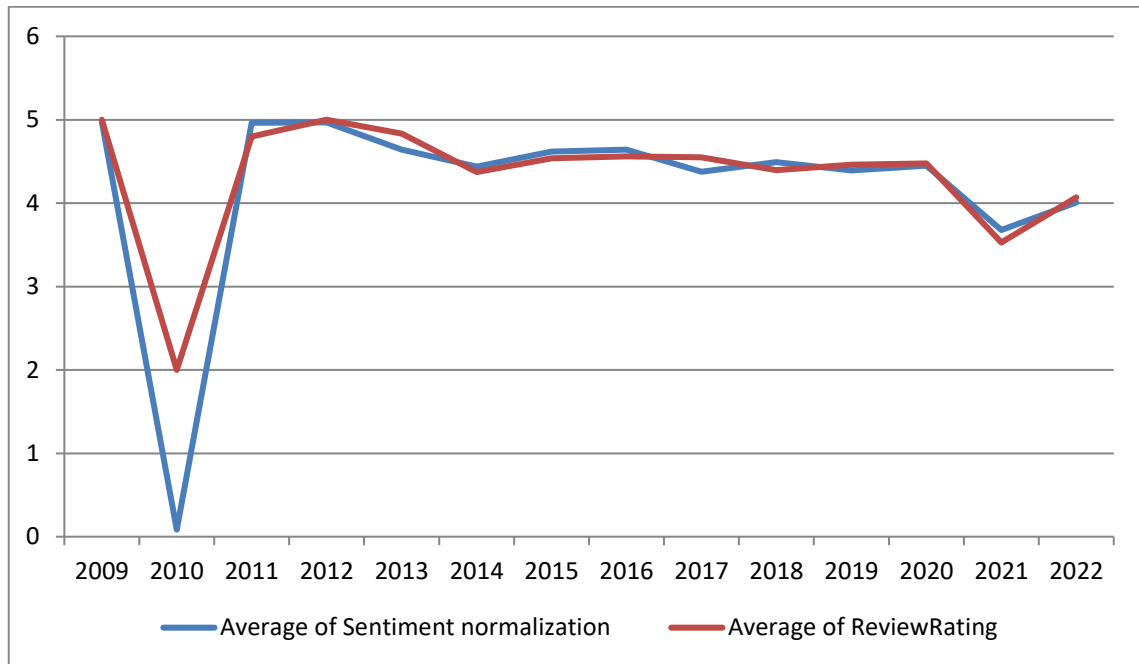
Source: Author's elaboration.

As presented in the Figure 4.4, during the year the average of sentiments is higher in the summer months, when compared with the average of ratings.

In Figure 4.5, can be analysed the average of the sentiment and ratings over the years and it is observed that during the pandemic period the average of sentiments and ratings dropped, which can be explained by the feeling of insecurity of tourists, but from the summer of 2021, both values rose again, although they have not yet reached pre-pandemic period values.

In the Figure 4.6, The Ireland is the country where the ratings and the sentiments is higher, on the other hand, France is the one with the lowest values, in terms of ratings, followed by Netherlands, however in the first country the average of sentiments is high and has a marked distance from the ratings.

Figure 4.6 - Author country versus sentiment
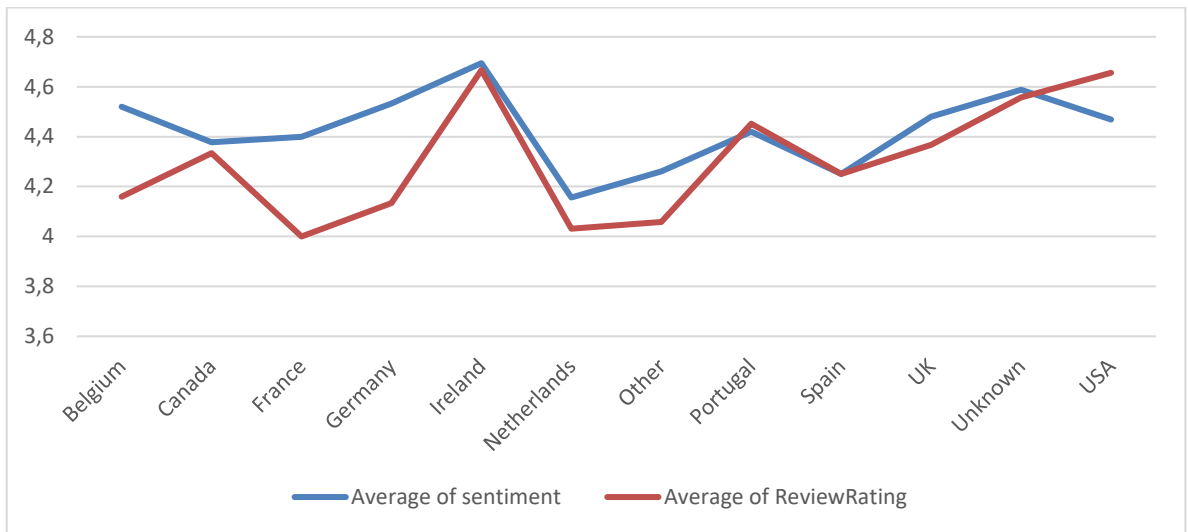
Source: Author's elaboration.

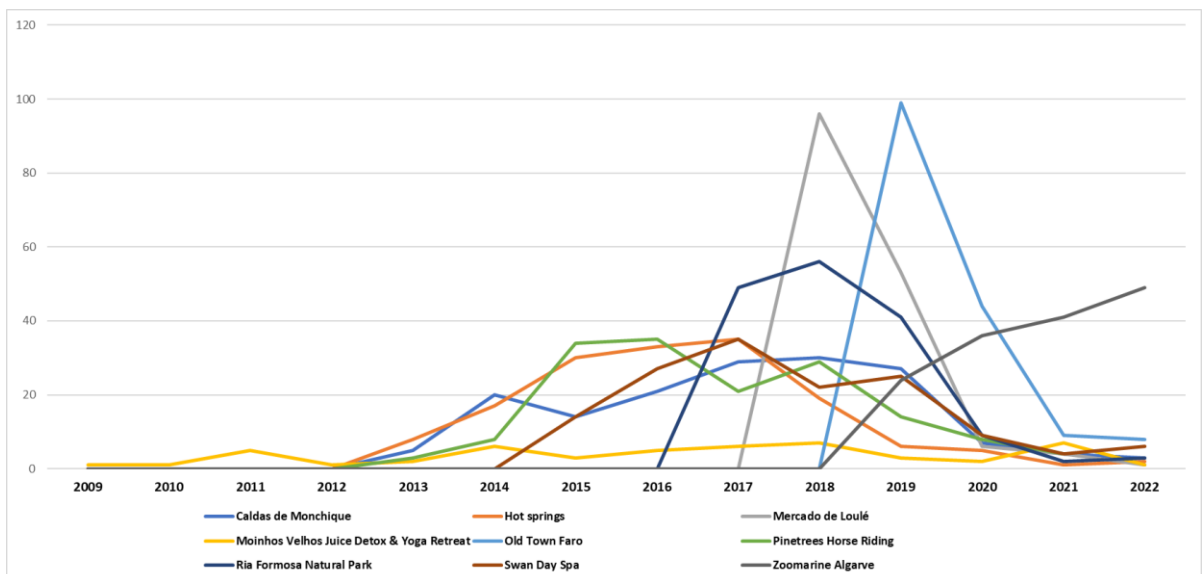In the Figure 4.7 presents the number of reviews for each hotel and facilities, where it can be seen that this number has grown, noting a slight decrease in the pandemic period. In other words, more and more tourists share their experience on social networks.

Figure 4.7 - Number of the reviews by hotel and facilities

Source: Author's elaboration.

Table 4.3 - Average of sentiment and average of review ratings by wellness category

Category	Average of sentiment	Average of ReviewRating
Hot Springs	4,6	4,1
Spas	4,7	4,6
Care of Body and mind	4,8	4,9
Medical Tourism	4,0	3,9
Natural Environment	4,8	4,6
Spirituality	4,7	4,6
Culture	4,2	4,3
Enogastronomy	4,4	4,2
Sport	4,7	4,5
Event	3,2	3,2

Source: Author's elaboration.

Taking into account the wellness destination, Table 4.3 presents the average of sentiments and rating associated with each category defined in Table 3.1 for the Algarve destination, complemented with the graph shown in Figure 4.8, where it is highlighted that the "Care of Body and mind " accompanied by "Natural Environment" are the categories ranked with the highest sentiments, followed closely by the "Spa", "Spirituality" and "Sports" categories. The "Events" category has the lowest rating.

Figure 4.8 - Average of sentiment and average of review ratings by wellness category
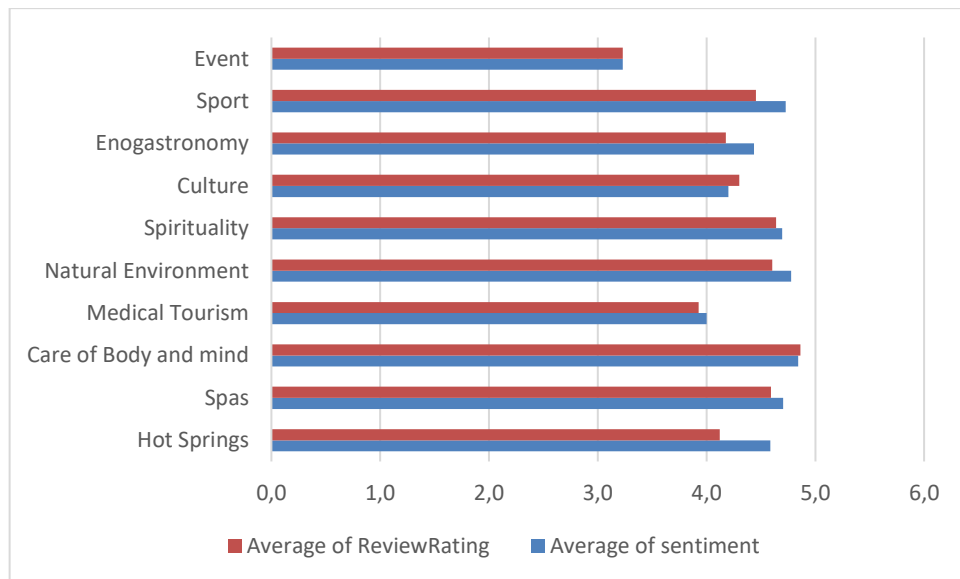
Source: Author's elaboration.

Regarding the average of the ratings, the category "Care of body and mind" has the highest rating and "Events" the lowest. The evaluation between the events at the sentiment or rating level is the same, it does not change in the two evaluations, on the contrary, "Hot Springs" presents higher values of difference between the feelings provided and the rating evaluation, which may mean that it helps to the tourist's well-being.

In terms of the keywords considered for each category of wellness destination, Table 4.4 presents the average of sentiment and ratings evaluated by tourists, complemented by the graph presented in Figure 4.9, which highlights that the experiences associated with the luxurious environment of spas, the well-being provided by culture and the nutritional part of the Enogastronomy category has the highest ratings in terms of feelings, in terms of rating they are also the highest with the exception of nutritional experience which has a low value when compared to feelings.

Table 4.4 - Average of sentiment and average of review ratings by words considered in the wellness category

	Word	Average of sentiment	Average of ReviewRating
Hot Springs	Thermal	4,5	3,7
	Health	4,6	4,3
Spas	Spa	4,7	4,6
	Luxury	5,0	5,0
Care of Body and mind	Massage	4,8	4,9
	Sauna	4,9	4,8
Medical Tourism	Healthy	4,6	4,3
	Medical	3,0	3,4
Natural Environment	Nature	4,8	4,6
	Landscape	4,7	4,9
Spirituality	Yoga	4,7	4,6
	Spiritual	4,4	4,4
Culture	Culture	4,0	4,1
	Welness	5,0	5,0
Enogastronomy	Food	4,4	4,2
	Nutrition	5,0	4,4
Sport	Sport	4,7	4,2
	Golf	4,8	4,8
Event	Festival	4,5	4,5
	Event	2,7	2,7

Source: Author's elaboration.

In the analysis of Table 4.4 e Figure 4.9, it is highlighted that the experience associated with events arouses low feelings, so it should be a category that Algarve destination managers should pay attention to in the future, followed by a category associated with medical services.

Figure 4.9 - Average of sentiment and average of review ratings by words of wellness category
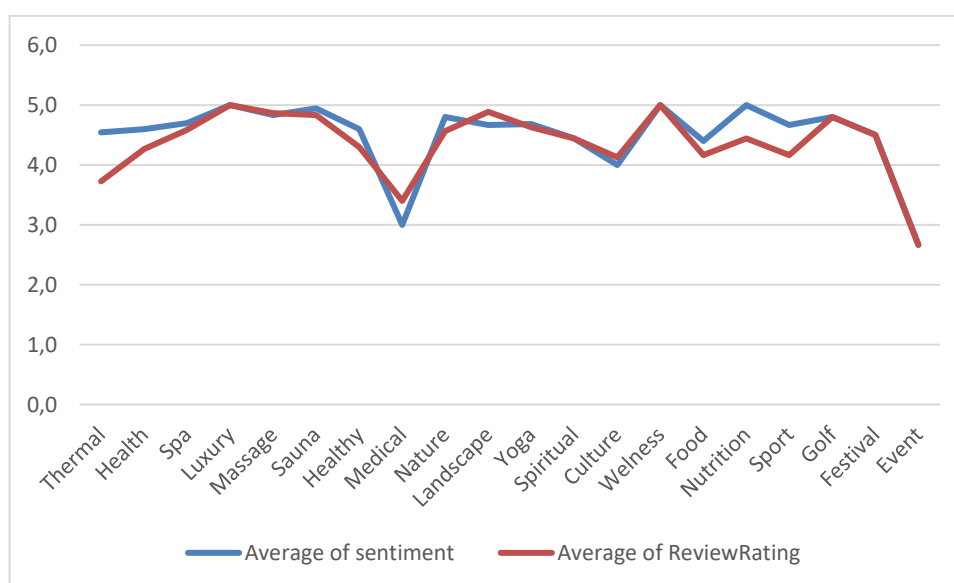
Source: Author's elaboration.

Table 4.5 - Average of sentiment by country of the author's review considered by the wellness category.

	Belgium	Canada	France	Germany	Ireland	Netherlands	Other	Portugal	Spain	UK	Unknown	USA
Hot Springs	5,0				5,0	5,0	5,0	5,0		4,3	5,0	5,0
Spas	4,0	4,0	4,0	4,2	5,0	4,0	4,9	4,6	5,0	4,6	4,9	4,5
Care of Body and mind	5,0	5,0	5,0	4,5	5,0	5,0	4,3	4,9	5,0	4,8	4,9	4,5
Medical Tourism	5,0									3,6	5,0	5,0
Natural Environment	5,0	5,0	4,0	4,0	5,0	4,3	4,7	5,0		4,8	5,0	5,0
Spirituality	5,0						5,0	5,0		4,5	5,0	5,0
Culture	5,0							4,0		3,3	5,0	
Enogastronomy		5,0	5,0	4,5	4,0	4,5	3,7	4,7	5,0	4,4	4,5	4,3
Sports					4,0	5,0		5,0		4,8		
Event	5,0						5,0	1,0	3,0	3,0	2,7	5,0

Source: Author's elaboration.

Source: Author's elaboration.

In terms of countries of origin of tourists and the analysis of sentiments, the Portuguese themselves stand out as negatively evaluating the Algarve in terms of events, followed by Spain and the United Kingdom. The category of spirituality obtained the maximum score in the feelings expressed by tourists, as well as the Hot Springs in almost all countries evaluated, except the United Kingdom, as presented in Table 4.5 and in Figure 4.10.

CHAPTER FIVE – DISCUSSION

5.1. Practical implications for the management of the tourism sector.

Sentiment analysis applied to tourism set to explore the opinion of tourist from the generated content, in essence, it helped to stimulate opinion on destination with wellness facilities, it reveals how wellness of individual at tourism destinations are generally satisfied based on keywords that illustrated opinions. It also widens the knowledge of the destination management organisation (DMO) on the sustainability of destination infrastructures due to tourist views and opinions on destinations. However, sentimental analysis of autonomous generated content still has some implications for managing the tourism sector.

The generated review content may give a psychological time frame view of the individual reviewers, which are subject to change as a result of implicit factors. The psychological views that are implicit may be general nature of the reviewer to be oval in given right judgments of services received and feedback become ambiguous and does not help in the growth of the sector.

The *Fulltext, Reviewrating* and the sentiment analysis result are sometimes not congruent. For example, the *Fulltext* "Bring your hiking shoes if visiting this area. Plan to stay for a few days at least, this was our biggest regret that we were not staying longer", Review rating is 5, sentiment normalization is 1 and sentiment opinion is negative.

Another example *Fulltext*: "We had a very enjoyable holiday in a bubble resort in the Algarve. But that is the best way to experience the area. Venture outside your manicured paradise and you'll find a fairly characterless, soulless place, typified by the 'historic' old town in Faro. This was really only a collection of cafes and restaurants in a block or two of cobbled streets where cars freely drove. It was an endurance test for our family and toddler. But it was also a reminder that this is holiday resort land and not full of beautiful, picturesque towns. If you're thinking about visiting, don't"

The *Reviewrating* is 1, sentiment normalization is 5 and positive opinion analysis. Although, the larger volumes of *Fulltext* are congruent with *Reviewrating*, sentiment

normalization and sentiment opinion, the outliers can still have effects on the management of the tourist sectors.

5.2. Contributions to the scientific community

The contribution of this research study is in the area of theory, methodology and practice. The theoretical study examined tourist destinations as an entity and the characteristics of wellness tourist destinations. The components of the wellness destinations (Dini and Pencarelli, 2021) were theoretically examined in furtherance of other studies. Unlike earlier studies that give descriptive analysis deducted from other literature, this study further applied the scientific method through data mining to analyse the sentiment of tourists toward each of the components in the wellness destinations. The sentimental analysis used a corpus graph to extract most keywords (visit, give, and love) common to all categories in the component.

The study from the sentimental analysis exposed the wellness categories in the destinations the tourists mostly visited, and which shows the economic viability of those categories to the management of the tourist sector as another contributor to the GDP (Gross domestic product). Also, it shows the weakness in other categories that will need the attention of the management of the tourist sector.

CHAPTER SIX – CONCLUSIONS

6.1.   Conclusion

Social networks and online reviews were the sources of information for tourism and all kinds of goods and services (Filieri and McLeay, 2013). Also, they are the sources of credible information data because the personal experiences of travellers are shared for public consumption (Ren and Hong, 2017).

In the analysis of the results with Text Mining algorithms to assess the sentiment expressed in the comments, the three key extracted words frequently used by each of the categories of the component in the Algarve wellness destination are Nice, Place, spa(Hot springs), Massage, Spa, Relax(Care of body and mind), Fish, Massage, Spa(Spirituality), Park, Bird, Walk(Natural environment), Experience, Time, Great(Event), Old, Town, Faro(Culture), Day, Ride, Park(Natural environment), Ride, Horse, Pinetree (Sport) and Market, Fish, Visit (Enogastronomy), The level of opinion on the popular hotels and its category visited by tourists in Algarve wellness destinations shows that Care of body and mind(massage and sauna), Enogastronomy (food and nutrition) and Natural environment(nature) were most acceptable to the tourist on wellness destination. The less popular destinations for wellness tourists include Culture, Sports, and Events.

The boundary of the study was limited to Algarve regions in Portugal, which include: Faro, Albufeira, Loule, Monchique, and Olhao and within these boundaries sited hotels and centre, namely: Caldas de Monchique, Zensation - Body & Soul, Mercado de Loulé, Moinhos Velhos Juice Detox & Yoga Retreat, Old Town Faro, Pinetrees Horse Riding, Ria Formosa Natural Park, Swan Day Spa, and Zoomarine Algarve.

The supposed components are ten: Hot springs, Spas, Care of body and mind, Medical tourism, Natural environment, Spirituality, Culture, Enogastronomy, Sports, and Events. At the time of extracting reviews manually from TripAdvisor on the Algarve wellness destination, there was no review on medical tourism for Algarve. However, it is suggested that for further study, the Process of opinion mining be multiple, which are reviews, Twitter and Facebook (Muthukrishnan et al. 2021) or Google reviews, so that there will be wider sources of information, especially in medical tourism.

6.2. Limitations

One of the limitations was having to collect manually, despite having made it possible to become aware of other situations that occurred in the establishments referred to in the text. If the software that allowed automatic collection had been used, more comments would have been considered for analysis.

The second limitation was having considered only one entity for each health and wellness tourism category. However, it allowed balancing the number of cases for each category, as not all have such a wide diversity of establishments as Spas.

More than two words could also have been considered for each health and wellness tourism dimension, but this also allowed for balancing the number of words analysed.

6.3. Future Work

As future work, it will be interesting to continue this line of investigation, include more comments, and compare with other tourist destinations that offer services and products and health and wellness tourism, which compete directly with the Algarve. In the analysis, each dimension can be considered in a more in-depth way, considering more words that represent the pertinent aspects with interest both for stakeholders and for the governance of destinations, which also allows for developing sensitivity to sustainability issues.

BIBLIOGRAPHY

Alaei, A. R, Becken, S. & Stantic, B. (2017) Sentiment Analysis in Tourism: Capitalizing on Big Data. *Journal of Travel Research*, 58 (2), 175–191. Doi: 10.1177/0047287517747753

Alqaryouti, O., Siyam, N., Monem, A. A., & Shaalan, K. (2020) Aspect-based sentiment analysis using smart government review data. *Applied Computing and Informatics*. Doi: 10.1016/j.aci.2019.11.003

Altexsoft (2018) *Sentiment Analysis: Types, Tools, and Use Cases.* Available at: https://www.altexsoft.com/blog/business/sentiment-analysis-types-tools-and-use-cases/ (Accessed, April 12, 2022).

Arean, M. (2019) *9 Requirement of wellness destination."recommended' helping travel advisors sell well.* Available at: https://recommend.com/get-inspired/wellness-travel/9-requirements-wellness-destination/. (Accessed September 30, 2021).

Arif, Y. M., Nugroho M. S., & Hariadi M. (2019) Selection of Tourism Destinations Priority using 6AsTD Framework and TOPSIS. In *2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*. IEEE, 346-351.

Biltawi, M., Al-Naymat, G., & Tedmori, S. (2017) Arabic sentiment classification: a hybrid approach. In *2017 international conference on new trends in computing sciences (ICTCS)*. IEEE, 104-108. Doi 10.1109/ICTCS.2017.24

Bonta, V., & Janardhan, N. K. N. (2019) A comprehensive study on lexicon based approaches for sentiment analysis. *Asian Journal of Computer Science and Technology*, 8 (S2), 1-6.

Borrajo-Millán, F., Alonso-Almeida, M. D. M., Escat-Cortes, M., & Yi, L. (2021) Sentiment Analysis to Measure Quality and Build Sustainability in Tourism Destinations. *Sustainability*, 13 (11), 6015.

Brandão, F., Liberato, D., Teixeira, A.S., & Liberato, P. (2021) Motives for Thermal Tourism: An Application to North and Central Portugal. *Sustainability,* 132, 2688. Doi: 10.3390/su132212688

Chen, W., Xu, Z., Zheng, X., Yu, Q., & Luo, Y. (2020) Research on sentiment classification of online travel review text. *Applied Sciences,* 10 (15), 5275.

D'Andrea., Ferri, F., Grifoni, P., & Guzzo, T. (2015) Approaches, tools and applications for sentiment analysis implementation. *International Journal of Computer Applications*, 125(3), 0975-8887.

Damijani, A. T. (2019) Wellness and healthy lifestyle in tourism settings. *Tourism review,* 74 (4), 978-989.

Dini, M. & Pencarelli, T. (2021) *Wellness tourism and the components of its offer system: a holistic perspective*. Emerald Publishing Limited.

El-Masri. M, Altrabsheh. N, & Mansour. H. (2017) Successes and challenges of Arabic sentiment analysis research:a literature review. *Social Network Analysis and Mining*, *7* (1), 1-22.

Filieri, R. & McLeay, F. (2013) E-WOM and Accommodation: An Analysis of the Factors That Influence Travelers' Adoption of Information from Online Reviews. *Journal of travel research*, 53 (1), 44-57

Gelter, J., Lexhagen, M., & Fuchs, M., (2021) A meta-narrative analysis of smart tourism destinations: implications for tourism destination management. *Current Issues in Tourism*, 24 (20), 2860-2874. Doi: 10.1080/13683500.2020.1849048.

Global Welness Institute (2022) *How the wellness industry is taking over travel.* Available at: https://www.bbc.com/worklife/article/20200203-how-the-wellness-industry-is-taking-over-travel (Accessed September 20, 2022).

González-Rodríguez, M. R., Martínez-Torres, R., & Toral, S. (2016) Post-visit and pre-visit tourist destination image through eWOM sentiment analysis and perceived helpfulness. *International Journal of Contemporary Hospitality Management,* 28 (11), 2609-2627. Doi: 10.1108/IJCHM-02-2015-0057

Guo, X., Yu, W., & Wang, X. (2021) An Overview on Fine-grained Text Sentiment Analysis: Survey and Challenges. *Journal of Physics: Conference Series*, 1757 (1), 012038. Doi:10.1088/1742-6596/1757/1/012038

Han, H., Kiatkawsin, K., Koo, B., & Kim, W. (2020) Thai wellness tourism and quality: Comparison between Chinese and American visitors' behaviors. *Asia Pacific Journal of Tourism Research*, 25 (4), 424-440.

Hankering (2021) *The History of Wellness, Hankaring for History.* Available at: https://hankeringforhistory.com/-thehistory-of-wellness/ (Accessed September 15, 2022).

Huang, X., Han, Y., Meng, Q., Zeng, X., & Liao, H. (2022) Do the DMO and the Tourists Deliver the Similar Image? Research on Representation of the Health Destination Image Based on UGC and the Theory of Discourse Power: A Case Study of Bama, China. *Sustainability*, 14 (2), 953.

Hutto, C., & Gilbert, E. (2014, May) Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the international AAAI conference on web and social media*, 8 (1), 216-225.

Jiang, Q., Chan, C. S., Eichelberger, S., Ma, H., & Pikkemaat, B. (2021) Sentiment analysis of online destination image of Hong Kong held by mainland Chinese tourists. Current Issues in Tourism, 24 (17), 2501-2522.

Jovicic, D. Z. (2019) From the traditional understanding of tourism destination to the smart tourism destination. *Current Issues in Tourism*, 22 (3), 276-282. Doi: 10.1080/13683500.2017.1313203.

Kazakov, S. & Oyner, O. (2021) Wellness tourism: a perspective article. *Tourism Review,* 76 (1), 58-63.

Kirilenko, A. P, Stepchenkova, S. O., Kim, H. & Li, X. (2017) Automated Sentiment Analysis in Tourism: Comparison of Approaches. *Journal of Travel Research*, 57 (8), 1012– 1025. Doi: 10.1177/0047287517729757

Khanna, C. (2021) Text pre-processing: Stop words removal using different libraries: A handy guide about English stop words removal in Python .*Towards Data Science*. Available at: https://towardsdatascience.com/text-pre-processing-stop-words-removal-using-different-libraries-f20bac19929a (Accessed April 14, 2022).

Klein-Hewett, H. (2021) Design as an Indicator of Tourist Destination Change: The Concept Renewal Cycle at Watkins Glen State Park. *Land*, 10 (4), 367.

Komilova, N. Kh., Usmanov, M. R., Safarova, N. I., Matchanova, A. E., & Murtazaeva, G. I., (2021) Tourist destination as an object of research of social and economic geography**.** *Psychology and Education*, 58 (1), 2058-2067.

Koncul, N. (2021) Wellness: A New Mode of Tourism. *Economic Research - Ekonomska Istraživanja*, 25 (2), 525-534.

Krakover, S., & Corsale, A. (2021) Sieving tourism destinations: Decision-making processes and destination choice implications. *Journal of Tourism, Heritage & Services Marketing (JTHSM)*, 7 (1), 33-43.

Lemi (2021) *Difference between spa, hot springs and wellness centres.* Available at: www.lemispa.com/difference-between-spa-hot-springs-and-wellness-centers/ (Accessed September 21, 2022).

Li, M., Lu, Q., & Long, Y. (2017) Are manually prepared affective lexicons really useful for sentiment analysis. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing*; 2 (SP), 146-150.

Madeira, A., Correia, A. & Filipe, J.A. (2019) *Understanding Memorable Enogastronomic Experiences: a qualitative approach*. PhD Dissertation, ISCTE.

Mahdu, N. (2021) *What Are The 14 Punctuations Mark You Need To Know*? University of the People. Available at: https://www.uopeople.edu/blog/punctuation- (Accessed April 14, 2022).

Medhat, W., Hassan, A., & Korashy, H. (2014) Sentiment analysis algorithms and applications: A survey. *Ain Shams engineering journal,* 5 (4), 1093-1113.

Medilik. S., (2003) *Dictionary of Travel, Tourism and Hospitality,* Third edition. Butterworth-Heinemann.

Milićević. S, & Jovanović, D. (2015) Wellness Tourism - Competitive Basis of European Health Tourism Destination. *DIEM: Dubrovnik International Economic Meeting*, 2 (1), 851-863.

Miller, K. (2021) *What Is Ayurveda?* Available at: https://www.webmd.com/balance/guide/ayurvedic-treatmen (Accessed March 30, 2022).

Mueller, H. & Kaufmann, E.L. (2001) Wellness tourism: market analysis of a special health tourism segment and implications for the hotel industry. *Journal of Vacation Marketing,* 7 (1), 5-17.

Mukherjee, S. & Bhattacharyya, P. (2013) *Sentiment Analysis: A Literature Survey. Master dissertation.* Indian Institute of Technology, Bombay. Available at: https://www.researchgate.net/publication/236203597 (Accessed July 12, 2022).

Muthukrishnan, H., Selvi, C. T., Deivakani, M., Subashini, V., & Kumar, S. G. (2021) Aspect-Based Sentiment Analysis for Tourist Reviews. *Annals of the Romanian Society for Cell Biology*, 5,183-5194.

Nandwani, P., & Verma, R. (2021) A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining*, 11 (1), 1-19.

Neves, M. & Seva, J. (2021) An extensive review of tools for manual annotation of documents. *Briefings in Bioinformatics*, 22 (1), 146–163. Doi: 10.1093/bib/bbz130.

Pan, X., Yang, Z., Han, F., Lu, Y., & Liu, Q. (2019) Evaluating potential areas for mountain wellness tourism: A case study of Ili, Xinjiang Province. *Sustainability*, 11 (20), 5668.

Plzáková, L., & Stupková, C. (2019) Environment as a key factor of health and well-being tourism destinations in five European countries. *IBIMA Business Review,* 611983. Doi: 10.5171/2019.611983.

Raj, A. (2020) *Unlocking the True Power of Support Vector Regression.* Available at: https://medium.com/towards-data-science/unlocking-the-true-power-of-support-vector-regression-847fd123a4a0 (Accessed July 20, 2022).

Rana, T. A., Cheah, Y. N., & Rana, T. (2020) Multi-level knowledge-based approach for implicit aspect identification. *Applied Intelligence*, 50 (12), 4616-4630.

Reitsamer. B. F, & Brunner–Sperdin, A. (2017) Tourist destination perception and well-being: What makes a destination attractive? *Journal of Vacation Marketing,* 23 (1), 55–72. Doi: 10.1177/1356766715615914.

Ren, G., & Hong, T. (2017) Investigating online destination images using a topic-based sentiment analysis approach. *Sustainability*, 9 (10), 1765.

Rodrigues, H., Brochado, A., & Troilo, M. (2020) Listening to the murmur of water: essential satisfaction and dissatisfaction attributes of thermal and mineral spas. *Journal of Travel & Tourism Marketing*, 37 (5), 649-661.

Rodrigues, H., Brochado, A., Troilo, M., & Mohsin, A. (2022) Wellness comes in salty water: thalassotherapy spas and service level of satisfaction. *International Journal of Spa and Wellness*, 5 (1), 71-90.

Singhal, G. (2020) *Importance of Text Pre-processing*. Available at: https://www.pluralsight.com/guides/importance-of-text-pre-processing (Accessed June 16, 2022).

Smith, M., K., & Puczkó, L. (2014) *Health, Tourism and Hospitality: Spas, Wellness and Medical Travel*. Routledge, 508

Stará, J., & Peterson, C. (2017) Understanding the Concept of Wellness for the Future of the Tourism Industry: A Literature Review. *Journal of Tourism & Services*, 8 (14), 18-119.

Taj, S., Shaikh, B. B., & Meghji, A. F. (2019) Sentiment analysis of news articles: a lexicon based approach. In 2019 2nd *international conference on computing, mathematics and engineering technologies* (iCoMET) IEEE, 1-5.

TechTarget Contributor (2018) *Stemming*. Available at: https://www.techtarget.com/searchenterpriseai/definition/stemming (Accessed September 20, 2022).

Thal, K. I., & Hudson, S. (2019) A conceptual model of wellness destination characteristics that contribute to psychological well-being. *Journal of Hospitality & Tourism Research*, 43 (1), 41-57.

Tubergen, A., & Linden, S. (2002) A brief history of spa therapy. *Annals of the Rheumatic Diseases*, 61 (3), 273.

Valentine. N. A. (2016) Wellness Tourism: Using Tourists' Preference to Evaluate the Wellness Tourism Market in Jamaica. *Review of Social Science*, 1 (3), 25-44.

Vishwanathan, S. (2014) Sentiment Analysis for Movie Reviews. In *Proceeding of 3rd IRF International Conference* (pp. 80-82), 10th May-2014, Goa, India, ISBN: 978-93-84209-15-5.

Vu, H. Q., Li, G., Law, R., & Zhang, Y. (2019) Exploring tourist dining preferences based on restaurant reviews. *Journal of Travel Research*, 58 (1), 149-167.

Wellness Tourism Association (2022) *Wellness Tourism Association.* Available at: https://wellnesstourismassociation.org/member_directory/kathleen-lesage/ (Accessed September 20, 2022).

Wilson, B. (2019) 9 Reasons to Book a Trip to the Algarve, Portugal, Tripsavvy editorial guideline. Available at: https://www.tripsavvy.com/reasons-to-book-a-trip-to-the-algarve-portugal-4142009 (Accessed May 25, 2022).

Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018) Recent trends in deep learning based natural language processing. *ieee Computational intelligenCe magazine*, 13 (3), 55-75.

Żemła, M. (2016) Tourism destination: The networking approach. *Moravian Geographical Reports,* 24 (4), 2-14.

Zhang, J., Li, J., & Wu, X. (2021) Research About the Development Strategies of Wellness Tourism Products—Take Phoenix Valley Forest as an Example. *International Conference on Education, Humanity and Language, Art (EHLA 2021)*. ISBN: 978-1-60595-137-9

Appendix A – Average of *Reviewratings* and Sentiments by country

Lines Labels	Average of ReviewRating	Average of Sentiment normalization	No. Of Reviews
Australia	4,1	4,1	8
Austria	2,5	2,6	2
Bahrain	4,0	4,9	1
Belgium	4,2	4,4	25
Brazil	3,0	3,1	2
Bulgaria	5,0	4,9	1
Canada	4,3	4,3	45
China	4,5	4,8	2
Colombia	4,0	4,9	1
Czech Republic	5,0	4,7	2
Denmark	3,0	4,5	1
Finland	4,2	4,6	6
France	4,0	4,3	10
Germany	4,1	4,4	15
Gibraltar	4,0	3,4	2
Greece	5,0	4,9	1
Hungary	4,5	4,8	2
Iceland	4,0	3,9	1
India	5,0	3,9	1
Ireland	4,7	4,6	72
Italy	4,1	4,6	7
Japan	5,0	5,0	1
Kenya	5,0	4,9	1
Luxembourg	4,0	4,5	1

Lines Labels	Average of ReviewRating	Average of Sentiment normalization	No. Of Reviews
Malaysia	4,0	4,2	1
Mexico	3,0	1,5	2
Netherlands	4,0	4,1	32
Newzealand	3,0	2,6	3
Norway	4,7	4,4	3
Peru	3,0	2,5	1
Porto Rico	5,0	4,8	1
Portugal	4,5	4,4	83
Romania	4,5	3,4	2
Russia	5,0	4,9	1
Scotland	2,0	4,9	1
Singapore	4,0	4,5	1
Slovakia	4,0	3,4	4
South Africa	4,0	4,6	1
Spain	4,3	4,1	12
Sweden	4,0	4,3	2
Switzweland	4,3	4,5	6
Thailand	4,3	4,6	3
UAE	3,5	4,5	4
UK	4,4	4,4	620
Unknown	4,6	4,5	269
USA	4,7	4,4	32
Total Geral	**4,4**	**4,4**	**1,294**

Appendix B –Research Flow Chart Developed in Orange