

## Article

# Unveiling Island Tourism in Cape Verde through Online Reviews

Cristina Oliveira <sup>1</sup>, Paulo Rita <sup>2,\*</sup>  and Sérgio Moro <sup>1</sup> 

<sup>1</sup> Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR, 1649-026 Lisboa, Portugal; cristinamarketing20@gmail.com (C.O.); scmoro@gmail.com (S.M.)

<sup>2</sup> NOVA Information Management School (NOVA IMS), Universidade Nova de Lisboa, 1070-312 Lisboa, Portugal

\* Correspondence: prita@novaims.unl.pt

**Abstract:** This paper is focused on research addressing a large quantity of data extracted from online reviews written by tourists visiting islands. These were extracted from TripAdvisor regarding island tourist destinations since there is a gap in the scientific literature using this approach on island tourism. The Islands of the Sun, Boa Vista and Sal, of Cape Verde, a Small Island Developing State (SIDS), were the targets of this investigation. After applying text mining to a large dataset, results are discussed, including from the perspectives of hotels, restaurants, and tourist attractions. For example, the beach is the main tourist attraction in both islands, but whereas in Boa Vista, tours on quad bikes constitute a major tourist activity, its equivalent in Sal is actually diving. The location of hotels near the beach is a big plus for tourists who also emphasize their human interaction with staff members in both hotels and restaurants.

**Keywords:** island tourism; online reviews; big data; text mining; Cape Verde; Atlantic Islands



**Citation:** Oliveira, C.; Rita, P.; Moro, S. Unveiling Island Tourism in Cape Verde through Online Reviews. *Sustainability* **2021**, *13*, 8167. <https://doi.org/10.3390/su13158167>

Academic Editor: J. Andres Coca-Stefaniak

Received: 30 June 2021  
Accepted: 17 July 2021  
Published: 21 July 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Small Island Developing States (SIDS) depend on the tourism industry as an important source for their Gross Domestic Product (GDP). Indeed, tourism is seen as a viable and sometimes the only means of economic growth [1], job creation, and earnings [2]. Lack of studies was found both in SIDS, including Cape Verde, in regard to analyzing customers perception and satisfaction, and even more so when considering online reviews as a proxy of tourism experience.

As a consequence of the market positioning strategy developed and communicated by Cape Verde's National Tourism Organization (NTO), it is justified to study whether the tourists' opinions perceived through the various experienced elements when visiting Cape Verde as a destination are in accordance with the current marketing strategy for these islands, or else if there is a mismatch between the islands' identities as formulated by managerial decision making from the respective NTO and the image perception from the perspective of tourists. In this way, it is possible to compare identity from the supply side, communication being conveyed, and experience-based image from the demand side. Ultimately, it is believed that this data will certainly be of great interest for Cape Verde's NTO as it may help to refine and adapt the way it deals and promotes tourism as well as improves and develops new products.

This research aims to challenge Cape Verde's NTO marketing communication approach by directly exploring the real data source of what are the most valued experiences that tourists have, if and how they are presented on the official website. More specifically, based on online reviews of previous visitors, this research was able to identify and analyze the specific features that are more relevant for an island tourism destination, Cape Verde, and evaluate whether its national tourism organization (NTO) is using them to promote its brand in an appropriate way.

The contributions of this investigation are:

1. To study tourists' experiences expressed through opinions in the form of reviews which are shared after they visit the islands of Cape Verde;
2. To determine what are the topics or themes which tourists most comment on and that may be the real identity of Cape Verde's activities, hotels, and restaurants;
3. To provide information that might be essential to NTO's marketing decision-makers to help them develop a more assertive promotional image as well as develop the necessary support instruments to assist and enhance the whole tourist experience;
4. To understand if the environmental issues emerge from visitors' experiences as they report them in the online reviews they write, which can help in guiding efforts from local authorities to a more sustainable tourism;
5. Finally, to contribute to the theoretical development of social media analysis via opinions and sentiments related to hospitality products, identifying features that influence the satisfaction of customers using lodging, attraction, and restaurant services.

## 2. Background

### 2.1. Island Tourism

Islands are usually associated with unique, exclusive, and different experiences granted to tourists who visit these destinations [3], thus emphasizing the importance of studying tourism in islands. These are usually characterized by: remoteness and insularity; susceptibility to natural disasters; institutional capacity constraints, resulting from the lack of human capital and the lack of economies of scale; limited economic diversification, which also involves a low degree of employment diversification; the small states economic pattern gives importance to the service sector, often tourism; openness which for small economies results in a high level of external dependency, socioeconomic and cultural vulnerability, and emigration; remittances remain a major ingredient of the small states growth model; income volatility and poverty [4,5]. Although there is no definite evidence of higher levels of poverty among small countries, their external dependency and vulnerability lend themselves to considerably uncertain income levels.

In the past, islands used to rely on agriculture, fishing and, eventually, handicraft as pillars of their economy [6]. The recognized limited economic diversification due to the poorness of resources turned attention to tourism as a solution to bring more diversification, employment, and better infrastructure [7]. Nevertheless, tourism activities have also brought negative impacts to islands, seen by their local communities [8]. Indeed, islands have also suffered from wearing out of their limited resources and fragile environment [9]. Furthermore, the use of sustainability criteria is not always integrated into the development options of interest groups [10].

It should also be stated that islands normally feature a relative shortage in transportation, food, and lodging, making them more expensive than they would be if located on the mainland [11]. Since island destinations are also attractive to tourists based on weather, climate, nature, and scenery, they end up being challenged by a permanent tension between attractiveness and scarcity for each dimension experienced by tourists [12].

Previous academic studies have researched tourism from several perspective layers: on the one hand, as individual islands, like Aruba [13], Barbados [14], Cyprus [15], and Mauritius [16], with most studied SIDS being Mauritius and Aruba; on the other hand, as archipelagos, such as the Azores [17], Canary Islands [18], and Madeira [19]; finally, as regions, like Asia/Pacific, the Caribbean, and Europe [20].

Individual islands have been studied to address a number of relevant issues. Following the recognition of limited literature on productivity management in small island hotels, Kilic and Okumus [15] developed an empirical research study to identify influencing factors from the perspective of hotel managers in Cyprus. Other authors studied the impact of foreign direct investment (FDI) on the economic growth of islands, such as Mauritius [16], comparing FDI with both domestic private and public investment.

Many SIDS have shifted their economies towards the tourism industry. Since the sustainability of tourism can be potentially affected by climate change, Cashman and Moore [14] examined this issue in Barbados. Indeed, this SIDS might be negatively affected by an increase in the consciousness of tourists from its long-haul source markets of their carbon footprint. Hotels on the island are also likely to have to put up with increasing operating and cooling costs as well as higher insurance premiums as a consequence of climate change impacts [21].

There is a clear research gap in the literature for two main reasons: (1) despite their importance for SIDS development, there are only a few studies on non-Caribbean and Oceania small islands developing states; (2) in addition, there are no studies using User Generated Content (UGC) on SIDS.

## 2.2. Online Reviews

Tourists' expectations of a destination are influenced by a number of information available, whether through pamphlets, tourism agencies, word of mouth, NTO online promotion information, and more recently through online reviews, the two latter being currently the most used. Online reviews are posted after the tourists experience the various elements available during their visit [22]. This mean can influence consumers' purchase intentions. Nowadays, there is an abundance of diversified unstructured data available on a number of social media sites that are growing exponentially in relevance and quantity every day. They are considered unstructured because they do not fit neatly into a database. These data are very valuable as a way to obtain market and customer intelligence. Therefore, business processes need to be fueled with this information as well as analyze it; otherwise, they will get rapidly obsolete and risk losing market share by not understanding, adapting, or predicting customer needs as well as their perceptions of the brand.

Consumers' intention of leaving a review was studied by Cheung and Lee [23], whereby among other motivations, they found enjoyment in helping other consumers. Fu et al. [24] found that positive electronic word of mouth (eWOM) was more driven by attitude, whereas negative eWOM was more driven by subjective norms. In restaurant experience, Jeong et al. [25] found that positive reviews were triggered by satisfaction experience with food, staff, appearance, and ambience, while the price was not a key driver for positive reviews. Text mining extracts and analyzes the useful information and knowledge that is hidden in text content. Furthermore, government strategies for tourism may consider an analysis of online reviews [26] whereby NTOs can use the proposed data analytics approach to understand what drives tourists' satisfaction, assisting in shaping a country's strategy.

## 2.3. Small Island Developing States

Jafari [27] stated that at that time, almost every nation in the world realized the potential and importance of tourism as an economic and prosperous industry. Although there is a number of well established, fully functioning and popular tourist islands in the world, such as Honolulu, Santorini, Crete, Cyprus, Capri, Bora Bora, Bali, Phuket, Cozumel, Oahu, and Kauai, there is also a number of small islands that are still going through development. SIDS were first recognized as a distinct group of developing countries at the UNCED [28]. A list of SIDS divided by geographic regions appears on the UNCTAD website: <https://sustainabledevelopment.un.org/topics/sids/list> (accessed on 30 June 2021).

Although islands suffer from general scarcity and the idea of being out of reach, difficult access makes it an attractive and almost exclusive destination. As so, islands are becoming an even more attractive destination. The number of international tourists visiting SIDS has increased by almost 300% in the decade up to 2011, from over 12 million to 41 million [29]. The UNWTO, in its 2012 World Report (Rio +20) as well as later [30], explained that international tourism is an essential economic source that promotes growth, foreign exchange, income redistribution, poverty alleviation, and job opportunities in SIDS.

As they are becoming more known around the globe, SIDS' strategic development needs to ensure their sustainable growth. UNWTO [29] argued that tourism development planned and managed in a sustainable way would bring huge economic benefits as well as development to small islands. This report also stated that tourism had already helped Cape Verde and the Maldives to move up position from their previous category of the least developed countries. They have a number of challenges to deal with, whether environmental and climate change, water shortage, natural disasters, external economic shocks, or dependency on imports. To support their development, international summits and conferences were established. Furthermore, according to UNWTO [30], the Barbados Program of Action was developed to assist SIDS in their sustainable development efforts. Also, an action platform was developed to help and monitor the implementation of pledges and commitments of partnerships.

The 2030 United Nations agenda is a balanced economic, social, and environmental plan of action for sustainable development where a worldwide collaborative partnership was established. Among other things, the plan seeks to eradicate poverty as the basis for sustainable development. All countries and stakeholders will make part of it in a collaborative partnership. They will give special attention to SIDS in pursuing sustainable development (<https://sustainabledevelopment.un.org/post2015/transformingourworld> (accessed on 30 June 2021)).

#### 2.4. Cape Verde

Cape Verde was classified as a SIDS back in 1992 [28]. Historically speaking, Cape Verde was discovered and colonized by Portuguese explorers who, in the 15th century, established the first European settlement in the tropics. Only in 1975, Cape Verde achieved its independence from Portugal. This archipelago with ten islands of volcanic origin is located in West Africa. It is part of the Macaronesia eco-region in the central Atlantic Ocean along with the Portuguese Madeira and the Spanish Canary Islands. As of 2018, the total population of Cape Verde was just over half a million (circa 553 thousand), with nearly 64% of the inhabitants living in urban areas. The total land area of this country is 4033 km<sup>2</sup>, with a density of 136 people per km<sup>2</sup> ([www.worldpopulationreview.com](http://www.worldpopulationreview.com) (accessed on 30 June 2021)). Due to the lack of natural resources, its developing economy is mostly service-oriented. However, there has been a growing focus on tourism and on foreign investment [30].

In this considered SIDS (UN, 2016), there are four islands (Brava, Fogo, Santo Antão, and São Nicolau) being promoted by its national tourism organization (NTO) as Islands of the Senses; three (Boa Vista, Maio, and Sal) are promoted as Islands of the Sun, and two (Santiago and São Vicente) are communicated as Islands of the Essence; one (Santa Luzia) is an unoccupied island ([www.guidadecaboverde.cv](http://www.guidadecaboverde.cv) (accessed 15 July 2020)). According to the Instituto Nacional de Estatística [31], the most visited islands are by far the Islands of the Sun—Sal (372k) and Boa Vista (241k)—followed by the Islands of the Essence—Santiago (95k) and São Vicente (48k)—and then the Islands of the Senses. The main tourists generating markets are the UK (196k), Germany (92k), France (85k), Portugal (80k), and Benelux (80k).

This paper reports on the two most visited Islands of the Sun, Boa Vista and Sal. In terms of arrivals, these two islands account for three quarters (74.9%) of all tourists coming to the country (Sal: 45.5%; Boa Vista: 29.4%). Regarding nights, the figures are even more impressive, totaling 89.9% (Sal: 51.9%; Boa Vista: 38.0%). Understandably, the supply numbers also show a strong predominance of these two islands, with Sal (47.4%) and Boa Vista (26.7%) taking 74.1% of the overall lodging capacity in Cape Verde. Finally, these two islands alone employ 77.6% of the workforce in tourism (Sal: 52.9%; Boa Vista: 24.7%) of the whole nine islands of the country [31].

Those islands of the sun are located closest to Africa and the Sahara desert, with more sand, better suited to beach tourism. They are communicated as offering beaches of white sand to lose sight on the horizon, some deserted virgins in a pure state. Waters are tepid and translucent, sometimes a turquoise blue to remind tourists of true gemstones.

Visitors are able to dive with fish and let the body go to the waves while the sun shines high, practically 365 days a year. They provide activities aimed at travelers who do not dispense unique moments by the sea or long walks by the sand while picking up shells. Tourists may surf waves, paddle, or venture into windsurfing or kitesurfing. These islands are targeted to those in love with the sun and the sea.

Cape Verde is present on online social media platforms, such as TripAdvisor. Nowadays, it is vital for a national tourist office to keep updated useful tourism information as well as to be able to make more assertive tourism destination online promotions. There are several international chains (e.g., Iberostar, Riu, Decameron) operating in both Sal and Boa Vista islands. The information generated from data analysis of extracted textual online reviews shared by visitors of island tourist destinations is a must for foundational decision making by destination marketing organizations as well as businesses. Hence, this rationale justifies the current study and its addressable identified research gap found in the scientific literature.

### 3. Materials and Methods

This study adopted a holistic approach by encompassing all feedback gathered on Boa Vista and Sal islands from TripAdvisor, the largest online reviews' tourism platform [32], among others such as Expedia and Yelp [33]. According to Statista [34], this American internet-based travel firm established in the year 2000 accounts for more than 884 million user reviews, covering over eight million listings of attractions, hotels, and restaurants. This early adopter of UGC is a publicly traded company in the Nasdaq stock exchange with a market capitalization of nearly USD 5 billion [35].

Automated data collection used the R statistical tool [36], an open-source software characterized by its wide range of packages published in the CRAN (Comprehensive R Archive Network) website: <https://cran.r-project.org/> (accessed on 15 July 2020). A web-scraping procedure was performed [37]. Specifically, a web-scraping script was developed to efficiently and automatically gather all online reviews from three major tourist categories, according to TripAdvisor's format of information: attractions, hotels, and restaurants. The result was a total of almost 38,000 online reviews from 324 units available in the above mentioned two islands (Table 1). Moreover, considering the land area of both islands (Boa Vista: 620 km<sup>2</sup>; Sal: 216 km<sup>2</sup>), these figures highlight the relevance of tourism to both islands.

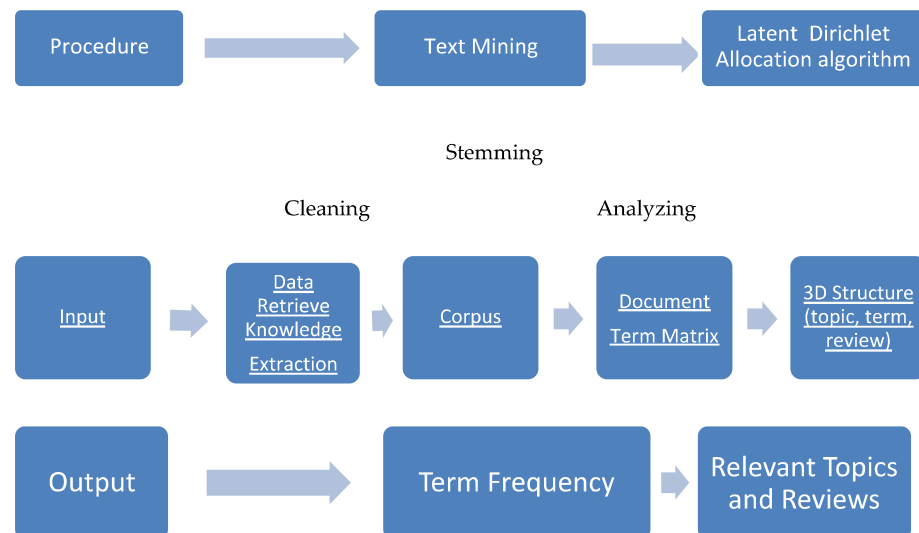
**Table 1.** Total units and reviews collected.

		Units	%	Reviews	%
Boa Vista	Attractions	38	11.7%	3052	8.0%
	Hotels	25	7.7%	7951	20.9%
	Restaurants	42	13.0%	1903	5.0%
Sal	Attractions	90	27.8%	4189	11.0%
	Hotels	49	15.1%	12,817	33.8%
	Restaurants	80	24.7%	8050	21.2%
		324	100.0%	37,962	100.0%

The analysis of such a volume of data requires automated tools that enable the extraction of the relevant patterns of information from text. Thus, we applied a text mining approach (Figure 1), taking into account methods previously used by Calheiros et al. [38], and Moro and Rita [39].

Text mining normally involves two processes for building the corpus of reviews: (1) cleaning the text of irrelevant words, like articles and adverbs; (2) stemming, i.e., reducing words to a single root word (e.g., "diving" is reduced to "dive"). First, a word frequency count took place to quantify the information, which was stored in a document-term matrix. In this step, we did not use an input dictionary because we aimed at an unguided approach without restricting the word frequency to a set of terms. To match

variations of words for the same family, a stemming procedure was conducted based on Guerreiro and Moro [40]. Finally, the document–term matrix was used as an input to obtain a coherent set of topics characterized by words used in the reviews. In our case, we adopted the Latent Dirichlet Allocation (LDA) algorithm since it offers a simple yet effective approach to aggregate topics [41].



**Figure 1.** Text Mining Analysis Procedure.

The reviews and each of the terms considered constitute the two dimensions of the document matrix, whereby each of the cells contains the frequency each term occurs in each of the reviews. As outputs, a table of frequencies counts the number of occurrences of each term.

The document term matrix is used as input for the LDA topic modelling that is a three-level hierarchical Bayesian modelling process grouping collections of items in topics defined by identified words/terms and the probability that each of them characterizes the topic [42]. This allows the analysis of the relative relevance of each term using the  $\beta$  distribution value that characterizes the relationship between the topic and the given term. A stronger relationship between a term and its corresponding topic is represented by a  $\beta$  closer to zero. The LDA final output is a tridimensional matrix encompassing terms, reviews, and topics.

#### 4. Results and Discussion

In this section, the results are shown divided by the two islands, Boa Vista and Sal, and the three categories (attractions, hotels, and restaurants). Thus, six-word frequency tables and six-topic tables are shown, enabling to enrich the discussion through the comparison of the results. The number of topics was tuned to seven topics according to the recommendation by Moro et al. [43].

##### 4.1. Term Frequency: Boa Vista and Sal Islands

“Tour” constitutes the most frequent word in Attractions implying that for consumers touring Boa Vista island is a cornerstone motivation for their visit (Table 2). This is also reinforced by the relevance of “trip”, the fifth most used word. Since this is one of the Islands of the Sun in the Cape Verde archipelago, it is understandable why “beach” comes second in the ranking. Moreover, “quad” (8th) shows up in the top ten as quad bikes are often used as means of transport for tourists to enjoy a “great” (4th) experience and reach “beautiful” (7th) scenery.

**Table 2.** Word frequencies for Boa Vista.

Attractions		Hotels		Restaurants	
Word	Frequency	Word	Frequency	Word	Frequency
tour	3299	room	13,159	beach	946
beach	2535	pool	10,548	great	930
day	2517	beach	10,511	place	838
great	2475	day	9721	good	812
trip	2074	staff	9564	love	676
book	1607	good	9286	bar	609
beautiful	1490	time	6871	visit	574
quad	1383	one	6748	staff	542
back	1358	bar	6674	drink	528
thank	1189	love	6426	friend	504

In hotels, “room” stands out as number one, underlying the importance given to hotels in fulfilling their accommodation function. Both “swimming pool” (2nd) and “beach” (3rd) have similar importance. This happens because hotels invest significantly in swimming pool areas (e.g., Riu, Iberostar). The importance of managing human resources in hotels, namely those that interact with guests, is stressed by the frequent use of the term “staff” (5th). Furthermore, the relevance of hotels providing bar services is also denoted in the appearance of “bar” (9th). “Good” (6th) and “love” (10th) observe overall sentiments of appreciation shared by guests as far as their experience in hotels of this island.

A large number of restaurants are located by the “beach” (1st), and this is recognized by the tourists when writing their reviews. Due to characteristics involving the service encounter, “staff” (8th) is also listed as one of the most frequently used words. Interestingly, “bar” (6th) and “drink” (9th) are also emphasized in the online reviews about restaurants in Boa Vista Island. There seem to be overall positive sentiments judging by the frequent use of the words “great” (2nd), “good” (4th), and “love” (5th).

Comparing the word frequencies for the island of Boa Vista across Attractions, Hotels, and Restaurants, the importance of “beach” is striking (2nd, 3rd and 1st positions, respectively). Since “staff” appears both in hotels (5th) and restaurants (8th) but not in tourist attractions, would it make sense to invest more in staff associated with tour activities? As expected, “bar” is significant for both hotels (9th) and restaurants (6th). Some words work as a proxy for customer satisfaction, such as “great”, “good”, and “love”. Interestingly, only Restaurant reviews include all those three (ranked 2nd, 4th, and 5th respectively), whereas online reviews of Hotels include two, “good” (6th) and “love” (10th). Reviews of tourist Attractions use only one, i.e., “great” (4th), but this is accompanied by another word, “beautiful” (6th), that is indeed more suitable to be used for Attractions than for Hotels or Restaurants.

In the tourist Attractions of Sal island, “dive” (3rd), “beach” (4th), and “sea” (10th) show how important is this Island of the Sun as a beach and sea destination where diving is placed at the forefront of tourist activities (Table 3). In addition, tourist experiences seem to be rewarded by the use of “great” (2nd), “good” (5th), and “love” (6th) adjectives.

The focus on Hotels starts with “room” (1st). Both “beach” (2nd) and “pool” (5th) seem to be perceived by guests as must services associated with Hotels. “Staff” (3rd) is very important in hotel units while “bar” (7th) is also a service valued by customers. The fact that “good” (4th), “great” (9th), and “love” (10th) appear in the list suggests an overall quite positive assessment by guests.

Restaurants in Sal Island receive attention regarding their “service” (7th) and as a “place” (3rd) to be with “friends” (5th). “Staff” (4th) is important in Restaurants, and, interestingly, “drink” (10th) is in the list but not “food”, thus underlining just the “B”

of “F&B”, i.e., “Food & Beverage”. Overall, satisfaction with Restaurants seems to be significant since the following four words are included: “good” (1st), “great” (2nd), “nice” (6th), and “love” (8th).

**Table 3.** Word frequencies for Sal.

Attractions		Hotels		Restaurants	
Word	Frequency	Word	Frequency	Word	Frequency
day	1912	room	19,499	good	4564
great	1739	beach	16,154	great	3294
dive	1473	staff	14,725	place	3133
beach	1463	good	14,582	staff	2439
good	1156	pool	13,896	friend	2339
love	1137	day	12,901	nice	2193
time	1115	bar	11,948	service	2157
back	1113	one	10,472	love	2092
see	1091	great	10,342	visit	2063
sea	1084	love	9749	drink	1812

“Beach” seems to be more relevant for Attractions (4th) and Hotels (2nd) than for Restaurants (not in the top ten). Conversely, “staff” receives more attention in Hotels (3rd) and Restaurants (4th). In general, all types of services are perceived positively by tourists with the words “great”, “good”, and “love” appearing in the top 10 lists of Attractions, Hotels, and Restaurants, with the latter also adding “nice” (6th).

Both for Boa Vista and Sal Islands “beach” is a key attraction, making them part of the cluster of the Islands of the Sun. However, whereas Boa Vista is more associated with “tour” and “trip” in quad bikes, Sal is more characterized by “diving” activities. Unlike tourist attractions where we can find some differences between the two islands, regarding hotels these islands seem to be similar since both receive online reviews focused on hotel “rooms”, “beach”, “pool”, “staff”, and “bar”, also with overall positive evaluations by guests. Finally, in restaurants, the common denominators across the islands are “staff” and “drink”. Curiously, the term “food” does not show in either of them.

#### 4.2. Relevant Topics: Attractions, Hotels, Restaurants

The second level of analysis used LDA parametrizing seven topics for attractions, hotels, and restaurants in each of the islands that are shown in Tables 4–9. Topics are represented in horizontal lines whereby the column labelled “Topic” shows the most relevant terms. The distribution values in each row refer to their respective topic, with the dominant (first) term (word) being the one that has a value closer to 0. The four most relevant terms are shown for each topic. Very interesting results were obtained from the analysis performed in the next six tables, whereby analysis of representative online reviews per topic is also provided in the text.

##### 4.2.1. Attractions

“Beach” in Boa Vista is clearly the main attraction, being the first term in topic 1, mentioned in 981 reviews, and associated with positive words such as “great” and “best” (Table 4); it is also the first word in topic 4 being addressed in 380 reviews. Online reviews mentioned: “we headed off road to a nearby beautiful *beach* and got some *great* pictures”; “we went to a beautiful *beach* only accessible off road, which I can say is the *best* I have ever seen (even in the Caribbean!!!)”.

The importance of “tour” can be assessed by the fact that it appears five times: two times as the first word (topics 5 and 6, with 342 and 291 reviews); two times as



the third word (topics 2 and 3, with 456 and 448 reviews); and once as the fourth word (topic 4, with 380 reviews). Online reviews included: “I enjoyed the *tour* in quad”; “we had a *tour* with a very small group (4 quads in total) nicer than the big groups of tourists going out with 20 quads”.

Whereas quad bikes are important means of transport when performing some tours (“quad” is the second word in both topics 5 and 6), diving is a relevant activity (“dive” is the first word in topic 3). Online reviews reported: “great day on the *quad* bike back and take his other quad tour!!”; “I was able to *dive* through the waves whilst making sure not to get out of my depth, have had a really enjoyable swim”.

The topics show three types of attractions standing out: beach, dive, and quad bike.

Three topics emphasize attractions during the day (topics 1, 3 and 7) and one at night (topic 2).

**Table 4.** Topics for Boa Vista’s attractions.

Topic	Nr. Reviews	1st Word		2nd Word		3rd Word		4th Word	
		Word	$\beta$	Word	$\beta$	Word	$\beta$	Word	$\beta$
1	981	beach	3.33	great	4.01	day	4.02	best	4.26
2	456	trip	2.58	night	3.40	tour	3.57	saturday	3.60
3	448	dive	3.28	day	3.56	tour	3.59	amazing	3.67
4	380	beach	3.25	book	3.26	beautiful	3.37	tour	3.55
5	342	tour	2.79	quad	3.30	thank	3.33	back	3.80
6	291	tour	3.03	quad	3.24	bike	3.73	safe	3.83
7	154	great	2.92	day	3.33	can	3.55	see	3.79

“Beach” is mentioned three times in tourist attractions of Sal Island and is the first term in topic three with 684 reviews (Table 5). Online reviews described: “the *beach* and sea are both beautiful”; “the *beach* was beautiful and free entertainment watching surfers and kite surfers enjoying themselves in the big rolling waves”.

Moreover, while “dive” is the first word in topic 2, “surf” appears in two topics (2 and 6). Online reviews refer: “I have *dived* all over the world and this is one of my favorite destinations”; “plenty of activity going on with some very talented local *surfers*”.

Tourist experience assessed by online reviews show positive impressions such as “great”, “love”, “nice”, and “like”.

Thus, for both islands, the relevance of “beach” is paramount. Although in the topic analysis, “diving” also emerges for Boa Vista Island, this activity is more pronounced in Sal Island together with “surf”, whereas in the former tours in quad bikes are more emphasized.

**Table 5.** Topics for Sal’s attractions.

Topic	Nr. Reviews	1st Word		2nd Word		3rd Word		4th Word	
		Word	$\beta$	Word	$\beta$	Word	$\beta$	Word	$\beta$
1	1308	see	3.84	day	3.98	trip	3.98	shark	4.24
2	1039	dive	3.23	great	3.75	time	4.20	surf	4.20
3	684	beach	3.40	horse	3.47	love	3.68	clean	3.95
4	485	salt	3.59	people	3.82	can	3.83	nice	3.85
5	312	plastic	3.54	day	3.74	beach	3.75	bus	3.78
6	208	activities	3.59	surfer	3.63	beach	3.96	first	4.00
7	153	salt	2.89	place	3.74	think	3.88	like	3.98

#### 4.2.2. Hotels

In Table 6, “beach” is the first word in topic 1 for hotels in Boa Vista but also appears as the fourth word in topic 5 and the second word in topic 6. Online reviews impressions: “beautiful hotel, stunning view of the Atlantic ocean and untouched *beaches*”; “the *beach* day is fantastic as they take you to a *beach* where the sea is turquoise and calm”.

Swimming “pools” are important for hotels as these are the leading words in three topics (4, 5, and 7). Online reviews refer: “a *pool* to swim near the room”; “loved the *pool* overlooking the ocean”; “round the clock entertainment around the *pool*”.

The relevance of “room” is witnessed by the fact it is not only the first word in topic 3 but it is also included in topics 1, 4 and 7. Online reviews: “*rooms* are 5stars or plus”; “I love this hotel, the *rooms* are clean and spacious”.

Both “staff” (topics 5 and 6) and “bar” (topics 2 and 3) appear twice. Online reviews critics: “*staff* were really friendly and helpful right across the whole hotel”; the *staff* make the hotel what is it, they are fantastic”; “good choice of drinks at *bars*—better choice than some all-inclusive”; “DJ was good and *bar* staff were excellent”.

“Good” (four times) and “love” (twice) are also worth noting in terms of the positive feelings they represent.

**Table 6.** Topics for Boa Vista’s hotels.

Topic	Nr. Reviews	1st Word		2nd Word		3rd Word		4th Word	
		Word	$\beta$	Word	$\beta$	Word	$\beta$	Word	$\beta$
1	1337	beach	3.43	good	4.10	room	4.12	love	4.26
2	1270	good	3.92	love	4.43	bar	4.47	friend	4.50
3	1171	room	3.83	good	4.22	bar	4.29	people	4.45
4	1132	pool	3.93	room	4.01	main	4.30	stay	4.35
5	1039	pool	4.07	staff	4.17	even	4.19	beach	4.34
6	1017	room	4.02	beach	4.09	staff	4.09	day	4.38
7	985	pool	3.75	room	3.96	good	4.31	one	4.48

The importance of “room” regarding hotels in Sal is clear since it appears in four of the seven topics and leads in topic 3 (Table 7). Online reviews opinions: “the building *rooms* and general layout are just as favor”; “*bedrooms* are well equipped with sea view”.

Location near the “beach” seems to be valued by guests as it is shown in three topics, in two of which it is the first term (topics 5 and 6). Online reviews evaluation: “you could go to the *beach* and sun loungers are provided there”; “this is a lovely hotel perfectly situated on the *beach* in its own grounds”.

Swimming “pool” is also of relevance, being presented twice (topics 1 and 7). Online reviews image: “great breakfasts taken outside by the *pool*”; “the *pool* area was great and there was always plenty of sun loungers free any time of the day”.

Both “staff” resources and “bar” service are worth noticing as these words show up twice in the top seven topics generated from analyzing online reviews. Online reviews: “hotel *staff* was very good”; “all hotel staff was brilliant, any minor complaints were dealt with quickly”; “we enjoyed the snack *bar* food”.

Finally, in more than one fifth of the topics x words, the terms “good” (four times) and “love” (twice) are present.

As was seen in the analysis of term frequency, hotels in both islands seem to share similar features in online reviews, with an emphasis being given to “beach” and “pool”, “room” and “bar”, and “staff”. Nevertheless, “room” is more stressed in Sal while “beach” and “pool” are more underlined in Boa Vista.

Table 7. Topics for Sal's hotels.

Topic	Nr. Reviews	1st Word		2nd Word		3rd Word		4th Word	
		Word	$\beta$	Word	$\beta$	Word	$\beta$	Word	$\beta$
1	2495	good	3.60	pool	3.66	love	3.97	beach	4.09
2	2159	staff	3.74	room	3.86	day	4.07	clean	4.39
3	2047	room	3.64	staff	3.72	bar	3.89	stay	4.14
4	1578	one	3.94	room	4.09	good	4.35	Bar	4.42
5	1563	beach	4.04	night	4.19	good	4.23	time	4.45
6	1495	beach	3.88	room	4.17	holiday	4.28	good	4.30
7	1480	one	4.02	pool	4.04	even	4.41	love	4.57

#### 4.2.3. Restaurants

The importance of “beach” for Restaurants in Boa Vista island is clearly shown by its first position in topics 3, 4, and 5 and the third one in topic 6 (Table 8). Online reviews: “a superb *beach* location made for a very pleasant afternoon”; “this is a terrific place to have a relaxed meal and then rest in the sunshine on a beautiful *beach*”.

Although far away, “bar” and “drink” are also noteworthy. Online reviews: “this *bar* is a great place to spend a night out!”; “loved this *bar*! Such a gem to the island! Playing relaxing music, good food and *drinks*”.

It is striking the relevance of adjectives with a positive valence such as “love” (appears four times), “great” (three times), “good”, “amazing”, and “beautiful” (once each), totaling ten appearances, i.e., more than one third of the total.

Table 8. Topics for Boa Vista's restaurants.

Topic	Nr. Reviews	1st Word		2nd Word		3rd Word		4th Word	
		Word	$\beta$	Word	$\beta$	Word	$\beta$	Word	$\beta$
1	430	good	2.73	great	3.25	place	3.33	drink	3.80
2	275	friend	3.65	night	3.73	bar	3.83	time	3.86
3	255	beach	3.44	amazing	3.96	visit	4.01	even	4.02
4	245	beach	3.78	visit	3.80	love	3.95	staff	4.14
5	244	beach	3.49	great	3.53	place	3.63	love	3.79
6	233	place	3.31	love	3.83	beach	4.11	beautiful	4.16
7	221	great	3.36	bar	3.96	love	3.99	friend	3.99

Online reviews of Restaurants from Sal reveal the importance of their location as “place” is seen four times in Table 9. Online reviews illustrate: “great *place* to watch football, good food”; “this is a great *place* to eat”.

“Service” and “staff” are also given some importance, although apparently much less. Online reviews: “delicious food, great *service* and lovely atmosphere”; “nice friendly *service*”; “the best thing about this place is the *staff*”.

What is striking is the number of times adjectives qualifying tourist experience are present in the top seven topics generated from online reviews: “good” is the first word in an impressive number of four topics (topics 3, 5, 6, and 7), “great” in one, appearing three times in total, “nice” is also present three times, “love” is shown twice and “delicious” once, reaching a total of over 45% of share in the four top words of the seven topics.

By comparing results from Boa Vista and Sal, the location of restaurants is valued (“beach”, “place”) in both islands, but whereas the reviews for the first also give some

relevance to “bar” and “drink”, the second receives more reviews focused on “service” and “staff”.

**Table 9.** Topics for Sal’s restaurants.

Topic	Nr. Reviews	1st Word		2nd Word		3rd Word		4th Word	
		Word	$\beta$	Word	$\beta$	Word	$\beta$	Word	$\beta$
1	1424	service	3.31	love	3.62	visit	3.87	delicious	4.01
2	1262	great	3.06	place	3.58	staff	3.71	nice	4.05
3	1258	good	2.80	nice	4.07	table	4.20	also	4.21
4	1204	friend	3.36	place	3.67	love	3.90	nice	3.97
5	1037	good	3.42	place	3.47	great	3.51	friend	3.53
6	998	good	3.37	drink	4.13	will	4.23	table	4.26
7	867	good	3.45	place	3.75	great	4.11	beach	4.11

## 5. Conclusions

The image of a tourist destination is in the eye of the beholder, who in tourism marketing is undoubtedly the consumer. Existing tourists share their experience via writing online reviews, which influence the perception of potential tourists when they read them during the consumer journey leading to the choice and purchase decision. Tourism has become a key activity sector for many SIDS, which feel the pressure coming from lack of resources while seeing an opportunity in tourism to foster job creation, generate revenue, and create conditions for economic growth. This is underlined by the fact that islands portray uniqueness and show attractiveness to prospective customers.

Online reviews on tourist attractions in both islands stress the importance of the beach (after all, these are islands of the sun). However, whereas in Boa Vista, tours by quad bikes throughout the dunes are frequently talked about, in Sal, diving is more prominent as a tourist activity. Hotels and restaurants are perceived as a more common denominator across islands. On the one hand, hotel guests, besides focusing on rooms, pay attention to their location near the beach, their swimming pool, as well as to their offer in terms of bar and interaction with staff. On the other hand, restaurant customers emphasize their staff and drinks, but interestingly, food is mostly absent from their reviews. Since food is not a common term regarding restaurants, maybe Cape Verde’s National Tourism Organization should work on this issue putting in value the local gastronomy. Food is important to a large part of tourists, and it seems that it does not stand out as much as the site’s characteristics in this case. The term “beach” is present in attractions, hotels, and restaurants in Boa Vista, but not in restaurants in Sal, while “staff” is addressed in both islands regarding hotels and restaurants but not in attractions. Most reviewers show appreciation by using adjectives such as “good”, “great”, and “love” in all three dimensions, plus “nice” in restaurants and “beautiful” in attractions.

This research was able to identify the more relevant features for accommodation (hotels), restaurant and tourist attraction services in both Cape Verde Islands of the Sun, i.e., Boa Vista and Sal. These results are bound to support decision making for service management not only by the industry players but also by the umbrella organization leading tourist promotion of these islands to international tourist generating markets. Indeed, national tourism organizations should take into account the specific results derived from applying analytics to online reviews via text mining in order to develop more sound marketing and promotion strategies based on real-world consumer-led data.

One important contribution from our study is the lack of clear references (i.e., prominent words) related to environmental issues. Thus, words such as “environment”, “ecology”, and “nature” did not emerge from online reviews. We argue that the fact that tourism is mostly concentrated in resorts belonging to large international chains in the two analyzed islands (Boa Vista and Sal), promoted by the NTO as “Islands of the Sun”, somehow dis-

places visitors from the island context. Evidence from the words retrieved support our claims, with words such as beach, pool, bar, and drinks dominating the comments, with the most reviews being written related to hotels. Such context suggests that balancing sustainability with the needed income in islands highly dependent on tourism may be difficult to achieve. However, we pinpoint that attractions able to drag tourists from resorts to sea may be a good starting point to increase sustainable conscience among visitors, as sea activities emerge from online reviews about attractions. Nevertheless, local efforts need to be leveraged to bring tourists to experience local cuisine and culture and to allow a greater attachment to the islands beyond traditional beach tourism.

Our study contributes to research by showing how the experience of tourists visiting islands as destinations can be useful to understand the characteristics of the services that most impact visitors in a comprehensive approach incorporating attractions, hotels, and restaurants. Future research should consider applying a similar analysis to other SIDS islands as tourist destinations, non-SIDS and even consider a comparison between mainland archipelagos, always from the perspective of consumers via online reviews as a proxy of their experiences as tourists.

**Author Contributions:** Conceptualization, C.O. and S.M.; methodology, C.O.; software, C.O.; validation, C.O., P.R. and S.M.; formal analysis, C.O.; investigation, C.O.; writing—original draft preparation, C.O. and P.R.; writing—review and editing, C.O. and P.R.; visualization, P.R.; supervision, S.M.; funding acquisition, S.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work is partially funded by national funds through FCT—Fundação para a Ciência e Tecnologia, I.P., under the project FCT UIDB/04466/2020.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Pratt, S. The economic impact of tourism in SIDS. *Ann. Tour. Res.* **2015**, *52*, 148–160. [[CrossRef](#)]
2. Lee, C.C.; Chang, C.P. Tourism development and economic growth: A closer look at panels. *Tour. Manag.* **2008**, *29*, 180–192. [[CrossRef](#)]
3. Nunkoo, R.; Gursoy, D. Residents' support for tourism: An identity perspective. *Ann. Tour. Res.* **2012**, *39*, 243–268. [[CrossRef](#)]
4. Bardolet, E.; Sheldon, P.J. Tourism in archipelagos: Hawaii and the Balearics. *Ann. Tour. Res.* **2008**, *35*, 900–923. [[CrossRef](#)]
5. Lim, C.C.; Cooper, C. Beyond sustainability: Optimizing island tourism development. *Int. J. Tour. Res.* **2009**, *11*, 89–103. [[CrossRef](#)]
6. Markwick, M.C. Tourism and the development of handicraft production in the Maltese islands. *Tour. Geogr.* **2001**, *3*, 29–51. [[CrossRef](#)]
7. Sharpley, R. Tourism modernization on and development on the island of Cyprus: Challenges and policy responses. *J. Sustain. Tour.* **2003**, *11*, 246–265. [[CrossRef](#)]
8. Ko, D.W.; Stewart, W.P. A structural equation model of residents' attitudes for tourism development. *Tour. Manag.* **2002**, *23*, 521–530. [[CrossRef](#)]
9. Harrison, D. *Tourism and the Less Developed World: Issues and Case Studies*; Cabi: Oxon, UK, 2001.
10. Briassoulis, H. Crete: Endowed by nature, privileged by geography, threatened by tourism? *J. Sustain. Tour.* **2003**, *11*, 97–115. [[CrossRef](#)]
11. Mitchell, R.E.; Reid, D.G. Community integration: Island tourism in Peru. *Ann. Tour. Res.* **2001**, *28*, 113–139. [[CrossRef](#)]
12. Klint, L.M.; Wong, E.; Jiang, M.; Delacy, T.; Harrison, D.; Dominey-Howes, D. Climate change adaptation in the Pacific Island tourism sector analyzing the policy environment in Vanuatu. *Curr. Issues Tour.* **2012**, *15*, 247–274. [[CrossRef](#)]
13. Rivera, M.; Croes, R.; Zhong, Y. Developing mobile services: A look at first-time and repeat visitors in a small island destination. *Int. J. Contemp. Hosp. Manag.* **2016**, *28*, 2721–2747. [[CrossRef](#)]
14. Cashman, A.; Moore, W. A market-based proposal for encouraging water use efficiency in a tourism-based economy. *Int. J. Hosp. Manag.* **2012**, *31*, 286–294. [[CrossRef](#)]
15. Kilic, H.; Okumus, F. Factors influencing productivity in small island hotels: Evidence from Northern Cyprus. *Int. J. Contemp. Hosp. Manag.* **2005**, *17*, 315–331. [[CrossRef](#)]
16. Khadaroo, J.; Seetanah, B. Transport infrastructure and tourism development. *Ann. Tour. Res.* **2007**, *34*, 1021–1032. [[CrossRef](#)]
17. Moniz, A.I. A dynamic analysis of repeat visitors. *Tour. Econ.* **2012**, *18*, 505–517. [[CrossRef](#)]

18. Santana-Gallego, M.; Ledesma-Rodríguez, F.; Pérez-Rodríguez, J.V. Tourism and trade in small island regions: The case of the Canary Islands. *Tour. Econ.* **2011**, *17*, 107–125. [CrossRef]
19. Almeida, A.M.M.; Correia, A.; Pimpão, A. Segmentation by benefits sought: The case of rural tourism in Madeira. *Curr. Issues Tour.* **2014**, *17*, 813–831. [CrossRef]
20. Park, S.Y. Examining potential for cruise business in the American-affiliate Pacific islands. *Asia Pac. J. Tour. Res.* **2011**, *16*, 1–19. [CrossRef]
21. Kelman, I. Critiques of island sustainability in tourism. *Tour. Geogr.* **2021**, *23*, 397–414. [CrossRef]
22. Brochado, A.; Oliveira, C.; Rita, P.; Oliveira, F. Shopping centres beyond purchasing of luxury goods: A tourism perspective. *Ann. Leis. Res.* **2019**, *22*, 484–505. [CrossRef]
23. Cheung, C.M.K.; Lee, M.K.O. What drives consumers to spread electronic word of mouth in online consumer-opinion platforms. *Decis. Support Syst.* **2012**, *53*, 218–225. [CrossRef]
24. Fu, J.-R.; Ju, P.-H.; Hsu, C.-W. Understanding why consumers engage in electronic word-of-mouth communication: Perspectives from theory of planned behavior and justice theory. *Electron. Commer. Res. Appl.* **2015**, *14*, 616–630. [CrossRef]
25. Jeong, E.; Jang, S. Restaurant experiences triggering positive electronic word-of-mouth (eWOM) motivations. *Int. J. Hosp. Manag.* **2011**, *30*, 356–366. [CrossRef]
26. Moro, S.; Rita, P.; Oliveira, C.; Batista, F.; Ribeiro, R. Leveraging national tourist offices through data analytics. *Int. J. Cult. Tour. Hosp. Res.* **2018**, *12*, 420–426. [CrossRef]
27. Jafari, J. The socio-economic costs of tourism to developing countries. *Ann. Tour. Res.* **1974**, *2*, 237–245. [CrossRef]
28. UNCED. In Proceedings of the United Nations Conference on Environment and Development, New York, NY, USA, 13 June 1992.
29. UNWTO. *Tourism: A Development Opportunity for Small Island States*; UNWTO Report; PR No: PR12038; World Tourism Organization: Madrid, Spain, 2012.
30. UNWTO. In Proceedings of the Conference on Small Island Developing States, Apia, Samoa, 1–4 September 2014.
31. INE. Instituto Nacional de Estatística. Available online: <https://ine.cv/en/> (accessed on 15 July 2020).
32. Moro, S.; Rita, P.; Oliveira, C. Factors influencing hotels' online prices. *J. Hosp. Mark. Manag.* **2018**, *27*, 443–464. [CrossRef]
33. Xiang, Z.; Du, Q.; Ma, Y.; Fan, W. A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism. *Tour. Manag.* **2017**, *58*, 51–65. [CrossRef]
34. Statista. The Statistics Portal. 2020. Available online: <https://www.statista.com/topics/3443/tripadvisor/> (accessed on 20 March 2020).
35. Yahoo. Yahoo Finance. 2021. Available online: <https://finance.yahoo.com/quote/TRIP/> (accessed on 19 July 2021).
36. Hornik, K.; Grün, B. Topicmodels: An R package for fitting topic models. *J. Stat. Softw.* **2011**, *40*, 1–30.
37. Mitchell, R. *Web Scraping with Python: Collecting More Data from the Modern Web*; O'Reilly Media, Inc.: Newton, MA, USA, 2018.
38. Calheiros, A.C.; Moro, S.; Rita, P. Sentiment classification of consumer-generated online reviews using topic modeling. *J. Hosp. Mark. Manag.* **2017**, *26*, 675–693. [CrossRef]
39. Moro, S.; Rita, P. Brand strategies in social media in hospitality and tourism. *Int. J. Contemp. Hosp. Manag.* **2018**, *30*, 343–364. [CrossRef]
40. Guerreiro, J.; Moro, S. Are Yelp's tips helpful in building influential consumers? *Tour. Manag. Perspect.* **2017**, *24*, 151–154. [CrossRef]
41. Moro, S.; Rita, P.; Cortez, P. A text mining approach to analyzing Annals literature. *Ann. Tour. Res.* **2017**, *66*, 208–210. [CrossRef]
42. Blei, D.M. Probabilistic topic models. *Commun. ACM* **2012**, *55*, 77–84. [CrossRef]
43. Moro, S.; Cortez, P.; Rita, P. Business intelligence in banking: A literature analysis from 2002 to 2013 using text mining and latent Dirichlet allocation. *Expert Syst. Appl.* **2015**, *42*, 1314–1324. [CrossRef]