

Article

Experience Economy Perspective on Recreational Fishing Tourism Travelers' Reviews: A Data Science Approach

Georgia Yfantidou ^{1,*}, Sofia Yfantidou ² , Dimitra Karanatsiou ² , Angelos Lontakis ³ , Mairi Maniopoulos ⁴, Panagiota Balaska ⁵ , Alkistis Papaioannou ⁶  and Vassiliki Vassilopoulos ⁴

¹ Department of Physical Education and Sport Science, Democritus University of Thrace, 69100 Komotini, Greece

² School of Informatics, Aristotle University of Thessaloniki, 57001 Thessaloniki, Greece; syfantid@csd.auth.gr (S.Y.); dimitrakd@csd.auth.gr (D.K.)

³ Department of Agribusiness and Supply Chain Management, Agricultural University of Athens, 32200 Thiva, Greece; alontakis@aua.gr

⁴ Hellenic Center for Marine Research, Institute of Marine Biological Resources and Inland Waters, 19013 Anavyssos, Greece; maniopoulos@hcmr.gr (M.M.); celia@hcmr.gr (V.V.)

⁵ Department of Physical Education and Sport Science, Aristotle University of Thessaloniki, 57001 Thessaloniki, Greece; pmpalask@phed.auth.gr

⁶ School of Social Sciences, Hellenic Open University, 26335 Patras, Greece; papaioannou.alkistis@ac.eap.gr

* Correspondence: gifantid@phyed.duth.gr; Tel.: +30-25310-39709

Abstract: Fishing tourism, a niche yet burgeoning sector within the broader tourism industry, presents a unique opportunity to explore the dynamics of the Experience Economy. This study delves into the intricacies of fishing tourism through the lens of data science, leveraging web crawling techniques to gather extensive data from reviewer profiles on TripAdvisor. By employing natural language processing (NLP) techniques, the study investigates the correlation between Experience Economy dimensions and user profiling aspects within this domain. The findings of this study shed light on the predominant dimensions of the Experience Economy within fishing tourism. "Entertainment" emerges as the primary dimension, closely followed by "Aesthetic", "Educational", and "Escapist" elements. Notably, the study reveals frequent co-occurrences of certain dimension pairs, such as "Entertainment"–"Aesthetic" and "Educational"–"Entertainment", underscoring the multifaceted nature of the fishing tourism experience. The implications of this research extend beyond academic discourse to practical considerations for stakeholders in the fishing tourism sector. By highlighting the socioeconomic benefits inherent in fishing tourism for local communities and fishers, the study emphasizes the importance of governmental support. This support, in the form of infrastructure development, effective leadership, legislative measures, and financial backing, is deemed essential for fostering sustainable growth and development in this sector. This study stands as a pioneering endeavor within the realm of fishing tourism research, particularly in its focus on Experience Economy dimensions and user profiling. By drawing data from both business pages and user profiles on TripAdvisor, it provides a comprehensive understanding of the intricate interplay between tourist experiences and the broader economic and social landscape of fishing tourism destinations.

Keywords: fishing tourism; sustainable tourism; recreation; experience economy; TripAdvisor; user profiling; natural language processing



Citation: Yfantidou, G.; Yfantidou, S.; Karanatsiou, D.; Lontakis, A.; Maniopoulos, M.; Balaska, P.; Papaioannou, A.; Vassilopoulos, V. Experience Economy Perspective on Recreational Fishing Tourism Travelers' Reviews: A Data Science Approach. *Tour. Hosp.* **2024**, *5*, 354–380. <https://doi.org/10.3390/tourhosp5020023>

Academic Editor: Brian Garrod

Received: 28 February 2024

Revised: 31 March 2024

Accepted: 19 April 2024

Published: 28 April 2024



Copyright: © 2024 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

Travelers often seek experiences in different tourism types, for instance, food and wine tourism, sports tourism, or luxury tourism, based on specific purposes. Desirable experiences can positively affect emotions, remain in customers' memory, and influence their consequent behaviors. Hence, understanding customer experience in the tourism sector is crucial for building better products and services, one such alternative tourism

product being sports tourism. As such, we utilize the Experience Economy perspective [1] to explore different experience realms. The Experience Economy presents four realms, the so-called 4Es, of experiential value for a business. The 4Es add Educational, Aesthetic, Escapist, and Entertainment experiences to a business offering. In this line, to remain relevant in the future of experiential travel, tourism businesses need to transform while also keeping in mind their responsibility towards a more sustainable development of the tourism industry. Transformations of tourism businesses, particularly in the sphere of collaborative consumption, in turn can lead to the transformation of local communities which then impact global communities. Taking a prospective lens and critically reflecting on the future of experiential travel needs more attention, and this special issue is the first milestone in deepening the knowledge on the who, why and what behind transformative experiences in experiential travel. This is also essential in order to profit from the added value that is derived from creating experiences that are in demand [2]. The study of Mehmetoglu and Engen [3] suggests that an evaluation of an experience based on the senses of feeling, learning, being, and doing may be an appropriate tool. Moreover, there are implications for how organizations within the tourism industry should think in regard to creating and developing their products and services. In order to meet the market needs and demands, it is important to create and “stage” experiences so that they capture the essence of the four dimensions (feeling, learning, being, and doing). It is becoming increasingly imperative to provide experiences, but there is hardly a recipe for what the experience should include in order to meet the customers’ expectations. The formula will vary according to the context and content of the experience, and also in relation to who will have the experience, viz, the “experience-customers”. Creating the “right” experience demands specific knowledge of the content needed by the customers [3].

Fishing tourism is a developing form of recreational tourism that promotes fisheries and aquaculture, offering travelers an opportunity to enrich their activities. Fishing tourism can be defined as “a set of activities carried out by professionals in order to differentiate their incomes, promote and valorize their profession and socio-cultural heritage, and enhance sustainable use of marine ecosystems by means of boarding non-crew individuals on fishing vessels” [4]. According to a 2020 study by the Centre for the Promotion of Imports from Developing Countries of the Dutch Ministry of Foreign Affairs, “The European Potential for the Development of Sports Tourism”, sports tourism was the fastest growing form of tourism before the pandemic. After the pandemic we can expect a lot of opportunities, especially in certain sports that contribute to sustainable development. This study showed that the sports tourism market can be divided into nine smaller markets. Figure 1 illustrates a model describing the smaller focused sports tourism niche markets that contribute to sustainable development. The model refers to those groups of tourists who have a strong commitment to their sport of choice and common characteristics (high education and income, active in social networks, belong to smaller age groups, etc.) [5].

In line with the above, the main task of this project is to classify tourism dimensions and build tourist profiles based on text reviews for fishing tourism businesses on Tripadvisor. The project results will help to better understand this new form of eco-tourism and further bolster fishers’ livelihoods. Our first goal is to classify tourists’ reviews on TripAdvisor according to the Experience Economy perspective to increase our understanding of the tourist experience regarding fishing tourism and recreation, while exploring the topics that emerge within these dimensions. At the same time, we focus on detecting tourist profiles based on their reviews, not only about fishing tourism businesses but also other businesses they might have reviewed during their travels. This sort of analysis can provide the local island businesses, the competent national and regional authorities, and the Fisheries Local Action Groups (FLAGs), which are partnerships across EU countries that aim to promote sustainable fishing/coastal communities, with additional knowledge of their customer base and might enable them to engage in destination development through niche tourism. Niche tourism refers to “how a specific tourism product can be tailored to meet the needs of a particular audience/market segment” [6]. Specifically, our second goal is to examine the

feasibility of building reviews-based predictors of various user characteristics, including, but not limited to, gender, age group, marital status, and interests. The latter category is particularly interesting because it allows us to explore secondary tourist interests to create more holistic experiences.

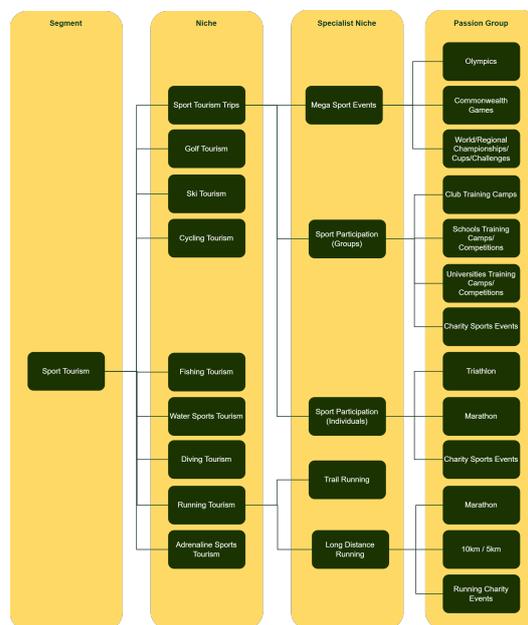


Figure 1. Niche markets of sport tourism, which include fishing tourism based on the research of CBI (figure courtesy of the Centre for the Promotion of Imports from developing countries [5]).

2. Related Work

Applying natural language processing (NLP) methods to the tourism domain is no new undertaking. Online platforms, such as Yelp and TripAdvisor, offer millions of publicly available reviews of businesses in the tourism domain, making them an attractive data source for research projects. However, to the best of our knowledge, this is the first study that aims to explore the Experience Economy dimensions in relation to user profiling aspects on fishing tourism.

2.1. Fishing Tourism

Marine tourism is the set of recreational activities and experiences that take place in the marine and coastal areas of a country to provide entertainment to tourists. Fishing tourism is also part of this category, which means that services are offered to visitors interested in alternative activities; thus, experience tourism is gaining ground in the preferences of tourists who love fishing and want to experience it. Fishing tourism is defined by the promotion and exploitation of fishing and aquaculture, activities with a rich traditional character in terms of employment and the means used but also in terms of the aquatic environment and aquatic life [7]. The term recreational includes any form of fishing which can be focused on entertainment, pleasure, competition, or even tourism. Moreover, some definitions of recreational fishing originate from the fishers' behavior and their fishing techniques. To conclude, in fishing tourism, tourists move from home to another place, where they stay overnight so as to participate in a fishing activity and receive tourist services. Thus, the fishing tourism combines two independent concepts: (1) fishing and (2) tourism, where 'fishing' answers to the consumer's question "What I want to do" and 'tourism' answers "Where can I do it" [8]. In particular, fishing tourism is defined as follows:

- The performance of the daily fishing process, accompanied by an explanation of the process to passengers;
- Encouraging the active, safe participation of visitors in the whole process of fishing and with the opportunity to be engaged at marine sport activities;

- Informing tourists about the fishing activity and fishing tradition;
- Visits to beaches, underwater caves, and boat trips;
- Possibility of diving for fishing and observation of marine flora and fauna;
- Contact with local flavors and traditional cooking of the catch;
- On-site tasting and sale of traditional fishing products;
- Overnight accommodation and catering services in fishers' houses or other "fisher's style" establishments.

Fishing tourism is developed in all aquatic destinations, including seas, lakes, rivers, and lagoons, where the fishing, farming, and breeding of aquatic organisms can be practiced. Furthermore, this form of tourism creates new infrastructure and jobs and is characterized by three forms: (a) active, (b) passive, and (c) shore fishing. In the active mode, the tourist actively participates in the fishing activity through a privately owned or chartered boat. In passive fishing, the tourist boards a professional fishing boat and watches the fishing activity as a spectator, thus coming into contact with the natural environment. Finally, shore fishing is a very popular recreational activity that involves long periods of inactivity that are ideal for rest and relaxation and also for periods of action in case the fish is hooked. Fishing tourists can join local fishing boats as paid guests, thus enabling them to enjoy the discreet beauty of the aquatic landscapes directly on top of the water. In Bangladesh, numerous options are available for recreational fishing, which can be highly beneficial to the nation and serve as a crucial tool for sustainable human development, including the eradication of poverty, the creation of jobs, and the improvement of rural areas [9].

As an example, the article of Smith et al. [10] summarized the economic impact of the culturally important catch-and-release recreational flats fisheries in the Bahamas, Belize, and Florida Keys, with a combined estimated annual economic impact exceeding EUR 689 million. Moreover, in the Bahamas and Belize, fishing guides must be residents, and no foreign guides are allowed to be licensed, thus further decreasing leakage because the fishing guide income remains in the local community. In the Bahamas and Belize, flat fishing is a significant component of each country's gross domestic product (GDP). It is much less so in the diversified and highly developed economy of Florida. The experience of trying to catch the fish, the enjoyment of escaping from work and everyday life, is generally more important than the actual event of catching a large quantity of fish. In many areas, shore fishing has been organized to take the form of a tourist product. This tourist product, i.e., "shore fishing", involves firstly the sale of fishing licenses for a limited period of time, the renting of rooms in tourist accommodation in the area by tourists who come exclusively for the purpose of fishing, and the provision of guiding services for the fishing area. In order not to exceed the bearing capacity of the area, the number of amateur fishers is strictly controlled, and the amateur fishers follow the tactic of catch and release.

The terms and conditions for carrying out fishing tourism must be exercised by professional fishers and owners of professional vessels who wish to carry out fishing tourism alongside their professional fishing activities. Fishing vessels carrying out fishing tourism are required to abide by the following:

- They must have an overall length of up to 15 m.
- They must be equipped with a professional fishing license for fishing with gear other than bottom trawls with nets and boat-drawn gillnets.
- They must meet the requirements of professional tourist vessels under the relevant laws.
- They must carry up to 12 passengers.
- They must be equipped with a certificate of seaworthiness stating the number of passengers they can carry, the extent of the voyages, and the relevant "Orders-Instructions" without requiring the issue of special or other certificates.
- There shall be a special waiting area for all passengers to be safely accommodated during fishing operations without obstructing them.
- They must comply with the rules laid down by the legislation in force at the time concerning the safety of navigation, manning, hygiene, and the suitability of the fishing vessel for the embarkation of passengers.

When conducting fishing tourism, professional fishers or sponge fishers shall demonstrate fishing or sponge fishing techniques in accordance with the national and fisheries legislation in force, using the fishing methods and gear specified in the vessel's professional license, except for bottom trawling with gillnets and boat-towed gillnets. Furthermore, fishing gear shall be so arranged on board the vessel that it does not impede the free and safe movement of the passengers and any activity on board. Furthermore, tourists may fish only with fishing lines, trolling lines, and probes, which may be handled manually and not mechanically, and may participate, under the responsibility of the master of the vessel and during fishing activities, only in operations that do not endanger their safety. Fishing tourism, for fans and non-fans alike, is an unprecedented and exciting activity that goes beyond the usual and introduces you to a different form of tourism and presentation of the beauty of a destination accompanied with recreation.

2.2. The Experience Realm

Holbrook and Hirschman [11] and Pine et al. [12] are the first authors of business studies to interpret in their book on business management the categories of experience "4E". They present the nature of the experiences in terms of economic activities. The groundbreaking work of Pine et al. [12] illustrates the four ways in which customers (tourists) can become involved or engaged in tourism experiences. The coupling of the dimension "tourist participation" with the dimension "environmental relationship" defines the four "realms" of an experience:

1. **Entertainment:** Usually, this experience is passively gained, where the viewer is not directly involved in the "performance" of the entertainment, e.g., participation in the theater, cinema, concerts, parades, nightclubs, carnivals, and folklore festivals as spectators.
2. **Education:** This type is the result of active participation and absorption of the material element that a person has been exposed to. The presentation by speech at a conference of thematic modules that are simultaneously a professional, e.g., a doctor, can be considered an experience of this module classification.
3. **Aesthetics:** The category of this type of experience is based on both exciting and passive enjoyment. A classic example of this type of experience is the understanding and inner search for the stimuli evoked by a series of specific themes of artistic works that are exhibited in a gallery or in an exhibition of unique exhibits in a museum.
4. **Escape:** This type exists when you are immersed in an activity that is actively engaged in by stakeholders who are transported into a new state of experience, e.g., role-playing to enhance relationship building in a working group wherein partners take on the role of role models playing the role of experts to solve a crisis, for example, due to an epidemic.

The shift towards prioritizing customer experiences and emotions stems from a changing consumer landscape, where passive engagement is no longer sufficient. Instead, customers actively seek meaningful interactions and value-added propositions. This evolution necessitates an escape from a traditional sales approach towards fostering engagement and co-creation. Concepts such as imagination, participation, and co-creation serve as the bootstrap for delivering this new paradigm of value, where customers play an integral role in shaping the products and services they consume. The theoretical framing of this experiential development in marketing came through the [12] proposal of 4E theory, whereby experiences are classified between two axes, and the poles of these axes as shown in Figure 2 [13]:

- Active (act) and passive (accept) participation, respectively, in the first axis;
- Basic adaptation/absorption and total immersion, respectively, in the second axis.

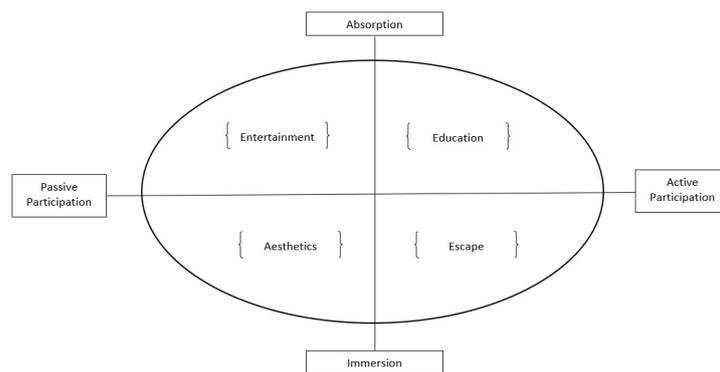


Figure 2. The experience realm (figure courtesy of Pikkemaat and Weiermair [13]).

2.3. Natural Language Processing and the Tourism Domain

Topic detection on travelers' reviews has been one of the main focuses of researchers in the tourism field [14,15]. Afzaal et al. [16] developed a multi-aspect classification approach, identifying aspects such as food, price, location, service, and ambiance on reviews from multiple online platforms. The same group [17] developed an alternative approach, including also the sentiment of the comment in the classification process. Sentiment analysis methods have also been applied in the field of online tourist reviews. Yu et al. [18] focused on foreign language sentiment analysis focusing on Japanese tourists' reviews. Marrese-Taylor et al. [19] introduced new domain-specific features for sentiment extraction, while Kirilenko et al. [20] provided a comparison between different machine learning algorithms for the task of sentiment analysis in the tourism domain. However, none of the above approaches have applied topic modeling and sentiment analysis approaches to examine the concept of user experience and niche tourism, as well as the tourists' perception of those concepts. More user-oriented research has focused on predicting reviews' usefulness [21,22], identifying suitable attractions' recommendations for users [23,24], and extracting certain user profiles [25]. While these works study individual user behavior, they do not examine the correlation between user profiles and tourist experiences to better understand different market segments. To the best of our knowledge, our work is the first to study user experiences, behaviors, and profiles in the domain of niche tourism.

2.4. User Profiling in Tourism-Related Platforms

Recommendation systems are mainly based on user choices and profiles to make accurate recommendations. These profiles are created in different ways regarding the field of application, for example, based on market basket [26] or user-generated content in the form of images, videos [27], or reviews [28]. In the tourism industry, recommendation engines on relevant platforms build individual tourist profiles to (a) suggest hotels, restaurants, attractions, or routes based on the shared ratings, reviews, photos, videos, or likes, and (b) provide businesses with insights about their customers segments. Recent studies have focused on tourist reviews to build user profiles. Kavitha et al. [29] exploit TripAdvisor reviews and social media profiles metadata to build a destination recommendation engine based on users' previously visited locations and matching user experiences about a destination. Moreover, Leal et al. [23] designed an algorithm for personalized destinations recommendation based on Expedia reviews by applying content-based filtering to topic-modeled tourists and locations. Except for the destination recommendation itself, researchers also focus on tourism-related services, i.e., restaurants and activities. The approach of Missaoui et al. [30] utilizes users' Yelp reviews to recommend the most relevant services by taking into account the opinions that this user has explicitly expressed through her/his previous reviews concerning other similar services following a language modeling approach. In addition, patterns on travelers' preferences extracted from reviews have also been identified by Fazzolari and Petrocchi [31], who propose methods that automatically analyze and summarize the reviews' features. However, none of the above methods infer

specific profiles from users' review history and platform metadata. To the extent we are aware of, our work is the first to build user profiles with targeted attributes categories (i.e., gender, age, marital/family status, and interest in activities) that successfully meet the needs of niche tourism businesses.

3. Materials and Methods

This sections discusses the process of acquiring and annotating the ground truth data prior to linguistic analysis.

3.1. Pipeline Overview

As a starting point, the HCMR researchers provide the "seed list" of fishing businesses' pages on TripAdvisor. Based on this list, the raw, unlabeled data are crawled from (1) fishing tourism businesses' pages on TripAdvisor, and (2) the respective user profiles on TripAdvisor using Python (<https://www.python.org/> accessed on 15 June 2021) programming language and the BeautifulSoup (<https://beautiful-soup-4.readthedocs.io/en/latest/> accessed on 15 June 2021) and Selenium (<https://www.selenium.dev/> accessed on 15 June 2021) packages for parsing web documents (i.e., dataset collection discussed in Section 3.2). After crawling the businesses' pages to obtain user reviews, we need to annotate a subset of the data according to the 4Es dimensions. Annotating a subset of the reviews as per the Experience Economy theory enables us to build models for automating this process for the remaining data, saving significant human effort (i.e., dataset annotation discussed in Section 3.3). Thus, authors manually annotate a subset of the dataset based on a pre-defined methodology that includes word frequencies of specific, pre-defined keywords. Then, for each collected review, we crawl the reviewer's profile data. These data contain information about previously visited locations and reviews as well as some demographic data (if available), such as the user's gender, age, interests, and permanent (home) location. Note that all data are pseudonymized upon collection. In addition, TripAdvisor assigns specific badges to users as a recognition for their contribution to the platform. These badges refer mainly to the platform use, for instance, if the user attracts readership attention or writes multiple reviews. However, some of these badges reflect the interests and behavior of the user while traveling, in particular expertise in restaurant reviews, in specific types of visited accommodation (luxury, resort, etc.), in attractions, in photography, and many more. All the aforementioned data (explicit profile) will be aggregated with specific characteristics mined from the reviews (implicit profile) to create holistic and informative tourist profiles through Python programming language and the scikit-learn library for machine learning (<https://scikit-learn.org/stable/> accessed on 15 June 2021) (i.e., model building discussed in Section 3.4).

3.2. Dataset Collection

TripAdvisor does not grant access to its content Application Programming Interface (API) for academic research purposes. Thus, to collect data for this project, we resort to web crawling as a means of mass data collection. A web crawler "is an Internet bot that systematically browses the World Wide Web, typically for the purpose of Web indexing" (https://en.wikipedia.org/wiki/Web_crawler accessed on 15 June 2021). In other words, it imitates the behavior of a user by visiting a set of pages and extracting data from them. Our web crawler operates per the *robots.txt* file of TripAdvisor; namely, it only requests pages that are permitted according to the file's directives.

For the scope of this project, we crawl the TripAdvisor profiles of 30 fishing businesses from various Greek islands (i.e., the "seed list"), containing 1062 English reviews from 1026 unique users. For each review, we store the reviewer's profile link and the date, rating, title, and full text of the review. Then, for each of the 1018 reviewers, we crawl their personal profiles, namely reviews, demographics, and badges. Precisely, we extract 11,862 user reviews for 9071 different businesses in 2497 locations worldwide. At the same time, for each of the 1018 users, we collect a sum of 11,562 badges of 27 unique types, such

as “Hotel Expert” or “New Photographer”. Finally, for 999 users whose demographics are fully or partially available, we collect data regarding gender (158 users), age (103 users), location (597 users), membership year (997 users), number of cities visited (706 users), number of contributions to TripAdvisor (687 users), and related tags (5 users).

Summing up, we build a new dataset of user reviews and related user data focused on niche tourism and specifically fishing tourism by crawling (1) fishing tourism businesses’ pages on TripAdvisor and (2) the respective user profiles on TripAdvisor.

3.3. Dataset Annotation

The Experience Economy framework was previously tested and confirmed in quantitative studies that deal with tourism issues [32,33]. However, according to Quadri-Felitti and Fiore [33], qualitative research methods may be developed to enrich the theoretical knowledge of the Experience Economy within another destination-specific tourist context. A qualitative design can improve our understanding and provide a more meaningful analysis of all components of the wine tourism experience. To operationalize the study, we select netnography [34] as a research technique. Netnography analysis is based on the collection of consumers’ reviews containing detailed information about their experiences published on the Internet. Compared to other qualitative research techniques, the distinctive value of netnography is that it excels at telling the story, understanding complex social phenomena, and assists the researcher in developing themes from the consumers’ points of view [35,36]. We use the Tourist Role Preference Scale of Gibson and Yiannakis [37] so as to combine tourist behavior with the 4E theory.

Ground truth is required to apply supervised learning techniques to classify tourists’ reviews on TripAdvisor according to the Experience Economy perspective. To acquire such ground truth, we label a subset (200 comments) of the reviews for fishing tourism businesses, where each comment is annotated by two to three researchers. Each comment is labeled with one or more of the 4Es dimensions, making our task a multi-label classification problem. The labeling decision is based on the domain expertise of the researchers who have created a list of key phrases for each dimension of the Experience Economy. For example, the education dimension may contain concepts such as fishing techniques (e.g., casting, UK-style fishing, and kayak fishing) or responsible fishing, while the entertainment dimension may contain concepts such as swimming, eating, or cooking. Similar concepts have been identified for the remaining two dimensions. For example, the escape dimension may contain concepts such as meeting or being hosted by locals and storytelling about island life, while the Aesthetics dimension may contain concepts such as enjoying the marine ecosystem beauty (flora and fauna). For the second dataset, i.e., user profiles, we utilize the scraped demographics data, such as gender or age, as labels for supervised learning tasks.

Summing up, not only do we build a new dataset of raw user reviews and related user data focused on fishing tourism but we also provide access to annotated data, which can be utilized for machine learning tasks for diverse purposes (e.g., user profiling and Experience Economy dimensions’ inference).

3.4. Model Building

For creating holistic and informative tourist profiles, we aggregate demographic information about the users (explicit profile) with specific characteristics mined from the reviews (implicit profile). These characteristics are inferred through natural language processing tasks, including linguistic insights (discussed in Section 4.1.1), extracted sentiment and emotion (discussed in Section 4.1.2), and topics detected (discussed in Section 4.1.3). For the explicit user profiling, we utilize the available demographic information to build machine learning models for inferring demographics (i.e., gender, age, and marital status) for users not having provided this information (discussed in Section 4.2). Finally, bringing it all together, we utilize the collected user reviews and the holistic user profiles to build machine learning models for inferring the Experience Economy dimensions encountered in user comments (discussed in Section 4.3).

Summing up, we propose a novel methodological approach for tourism-related research, incorporating machine learning and data science methodologies, enabling data-driven knowledge extraction from publicly available data to enhance touristic product offerings, i.e., TripAdvisor reviews.

4. Results

In this section, we delineate our findings on extracting sentiments, emotions, and relevant insights from the reviews' linguistic cues. We also explore user profiling aspects in fishing tourism and how the 4Es translate to this niche tourism product.

4.1. Extracting Sentiment, Emotion, and Descriptive Insights of Tourists and Businesses from Linguistic Cues

To gain insights into customers' collective word of mouth for fishing businesses and the tourists' general interests, we perform natural language processing tasks on the extracted reviews. In particular, we first preprocess the text of the reviews to find the most occurring words and phrases and relevant linguistic insights. Moreover, we carry out sentiment and emotion analysis following mainly an unsupervised approach using relevant lexicons and relevant expressions. Finally, we apply text clustering and topic extraction to the reviews to elicit specific interests of tourists and commonly emerging user experiences.

4.1.1. Linguistics Insights

Text cleaning. The initial step is to clean the text of the reviews in order to extract noisy info and maintain a clear dataset that will be used for future tasks. We first convert the words in our corpus to lowercase. Then, we filter out external URLs, Unicode characters and emojis (symbols and pictographs), digits, and punctuation. After this step, all *stopwords*, the set of most commonly used in the language in general, are taken out of the corpus.

Words, stems, and lemmas. After cleaning the reviews corpus, we segment the text into word units using a pre-built tokenizer. We then apply stemming and lemmatization to the words' vocabulary to reduce the inflectional forms of each word into a common base or root. *Stemming* is the process of cutting off word endings to gain a common root for words. In contrast, *lemmatization* refers to the action of gaining a common root using morphological analysis of the words. To become more specific, the difference between the concepts is that in stemming, we obtain the root after applying a set of rules without bothering about the part of speech (POS) or the context of the word occurrence. At the same time, lemmatization deals with obtaining the root of a word after understanding the POS and the context [38]. We utilize Snowball Stemmer and WordNet lemmatizer from the "nltk" package for each task.

Words relevance and importance. Since not all words carry the same importance for the reviews, we apply the Term Frequency–Inverse Document Frequency (TF-IDF) statistical measure that reflects how important a word is in a document collection. Moreover, in the same spirit, we extract contiguous sequences of words, *ngrams* to account for phrases that appear often.

Collective insights for fishing businesses. As for businesses and collective word-of-mouth, we visualize (Figure 3a,b) the emerged insights that reveal the overall satisfaction and excitement about the fishing boat trip activity. In particular, we highlight the following insights:

- Tourists went on boat trips mostly for fishing and actively participated in the process (fishing experience, caught fish, etc.).
- The overall experience of fishing boat trips is highly recommended, as tourists mention that they had a "great time" and "fantastic day".
- Tourists appreciate the beauty of the natural environment by leaving positive comments about the "crystal waters", the sea, the fresh fish, etc.
- Tourists highlight the hospitality and the skills of the crew and business owners as underscored by the existence of the crew's names in the word clouds.



Figure 3. Words and concepts that appear most in fishing business reviews (source: authors' own creation).

These reviews reflect the positive tourist experience translated in star ratings from the tourists, shown in Figure 4. It is evident that the vast majority of tourists leave five-star-rated reviews.

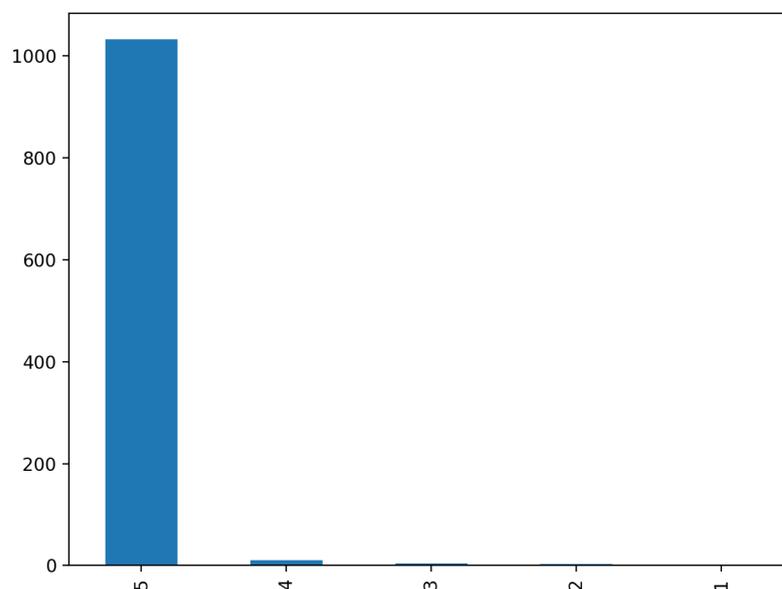


Figure 4. User ratings on a scale from 5 to 1 stars that appear in fishing businesses reviews. The majority of the users gave 5-star reviews. (source: authors' own creation).

Collective insights for tourists. As for tourists and their general interests as depicted in our dataset, we visualize (Figure 5a,b) the emerged topics revealing that tourists mostly comment about food, restaurants, hotels and services. In particular, we highlight the following insights:

- Tourists overall put emphasis on commenting on the offered services.
- Food and restaurants are at the top of tourists' attention.
- Tourists often make positive comments about services, with phrases such as “really good”, “really nice”, “well worth”, “great food”, etc.
- Tourists in practice write reviews for businesses in order to recommend services and experiences or not.

An interesting finding is that the rating distribution for overall user reviews (Table 1) significantly differs from the one identified previously in the fishing businesses' reviews (Figure 4). Specifically, while only 7% of the overall user reviews have a five-star rating, this percentage rises to almost 100% when it comes to fishing tourism.



Figure 5. Words and concepts that appear most in users' reviews (source: authors' own creation).

Table 1. General user ratings distribution and insights (source: authors' own creation).

Rating	Counts Percentage	Avg #Words
1	6087 (51.3%)	82.3
2	15 (0.1%)	59.1
3	46 (0.3%)	118.4
4	4795 (40.4%)	95.0
5	919 (7.0%)	81.8

Additionally, we have explored the TripCollective badges assigned by TripAdvisor to each user to highlight one's contribution to the community. Specifically, TripAdvisor assigns badges for the following categories:

- **Reviewer Badges:** These badges are graded starting from the "New Reviewer" (1 review) to the "Top Contributor" (more than 50 reviews). Figure 6a shows the distribution of tourists who reviewed fishing tourism business based on their "Reviewer Badges". Interestingly, approximately half the users belong to the "New Reviewer" category, meaning that they only joined the TripAdvisor platform to positively review the respective fishing tourism business.
- **Expertise Badges:** These badges showcase the unique knowledge of the users. For example, if a user publishes multiple reviews in a single category—hotels, restaurants, or attractions—they will be assigned the respective "Expertise Badge". Figure 6b shows the distribution of tourists who reviewed the fishing tourism business based on their "Expertise Badges". We notice that the users who review fishing tourism businesses also tend to review hotels and attractions, as well as luxury and boutique hotels, and B&B and Inns at a smaller scale.
- **Passport Badge:** This badge recognizes users for being world travelers. Once they have added reviews for places in at least two destinations, they start collecting such graded badges. Figure 7 shows that most users who have reviewed fishing tourism businesses have only reviewed destinations in limited locations. This is not surprising, as almost half of users are "New Reviewers" as mentioned previously.
- **Explorer Badge:** This badge is assigned to users who are amongst the first to review a hotel, restaurant, or attraction in a given language. Our results indicate that one out of three users who have reviewed a fishing tourism business own this badge, meaning that they are trailblazers in the tourism domain, seeking out-of-the-beaten-path experiences.

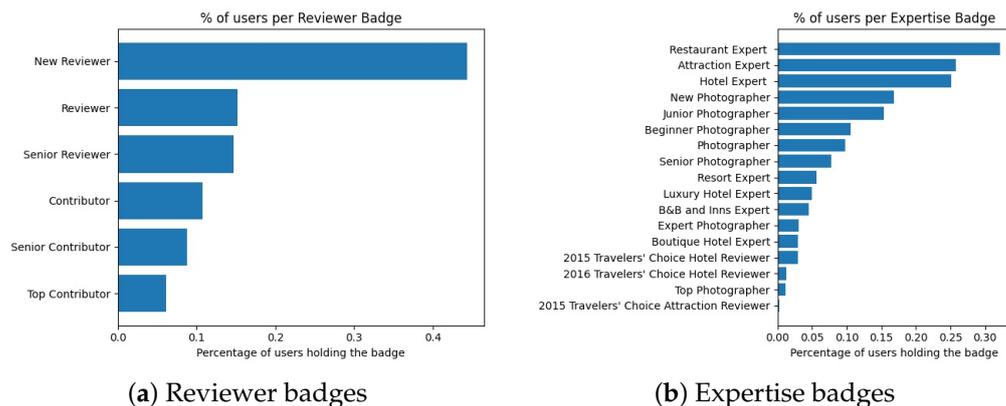


Figure 6. The distribution of different types of badges across all users who reviewed fishing tourism businesses (source: authors’ own creation).

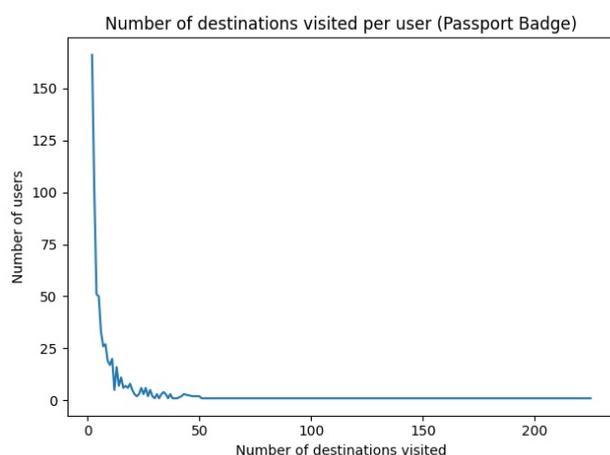


Figure 7. A plot of passport badges indicating the number of locations visited across all users who reviewed fishing tourism businesses (source: authors’ own creation).

4.1.2. Sentiment Analysis and Emotion Extraction

Sentiment analysis refers to the opinion mining task that aims to discover the attitude of the author towards the discussed entity expressed in their texts. The sentiment is mainly classified as positive, neutral, and negative. Emotion extraction consists of a focused sentiment analysis task that aims to extract specific emotions and not general attitudes. In our analysis, we employed the model of the six primary emotions defined by Ekman [39]: *joy, sadness, disgust, anger, fear, and surprise*.

Sentiment analysis. To extract the total sentiment for each review, we employed a pre-trained model for sentiment analysis, the Textblob [40].

- **Insights for tourists.** The distribution of the overall sentiments found in user reviews in our corpus reveals that there is a tendency to express neutral to positive comments about tourist venues and experiences. A surprising finding is that even for lower ratings, the detected sentiment polarity is rather positive than negative (see Figure 8a).
- **Insights for businesses.** The distribution of the sentiments found in reviews for fishing businesses, as shown in Figure 9, reflects the overall satisfaction of customers with the provided services. We expected a positive sentiment after the aforementioned linguistic insights and the high ratings that fishing tourism businesses received.

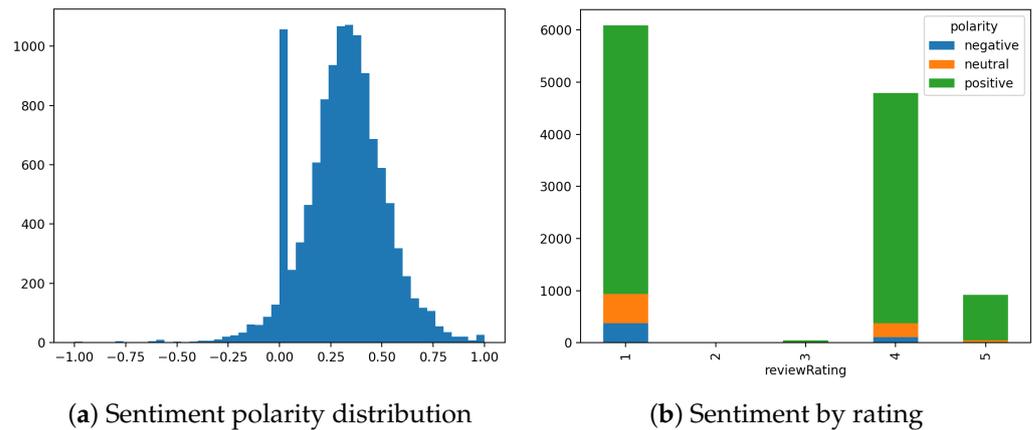


Figure 8. Overall users' sentiment polarity distribution and sentiment by rating for all reviews (source: authors' own creation).

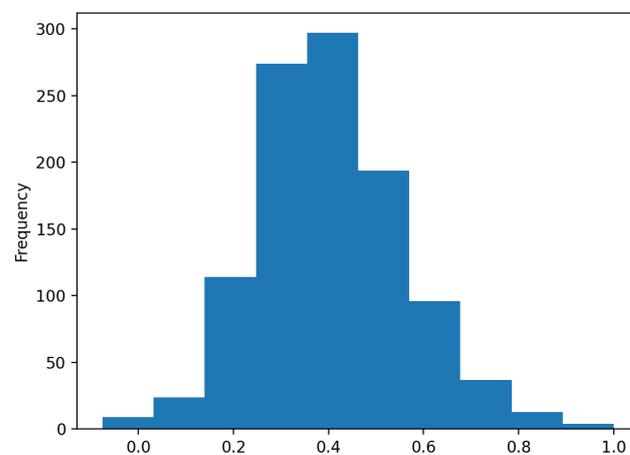


Figure 9. Fishing tourism businesses review sentiment polarity distribution (source: authors' own creation).

Emotion extraction. To extract specific emotions expressed in reviews, we employed an unsupervised approach by employing an affective lexicon, including affective emojis. In particular, we utilized “Wordnet Affect”, an extension of WordNet Domains, including a subset of synsets suitable to represent affective concepts correlated with affective words. We mapped words in reviews with the six primary emotions based on these affective concepts. If other specific emotions emerged, they were grouped into the general classes of positive and negative emotions. This approach works mainly for reviews written in English.

- **Insights for general tourist reviews.** We collected a set of 11,861 reviews from the profiles of users in our dataset. Given that a percentage of 90.8% of these reviews are written in English, we found 30,649 occurrences of affective concepts based on WordNet affect (see Table 2 for distributions). Emotion analysis shows that users are more likely to express positive emotions in their reviews, specifically surprise and joy, while negative emotions are expressed less frequently.
- **Insights for fishing tourism businesses.** Concerning the affective concepts found in users' reviews for fishing businesses, based on the linguistic and sentiment analysis preceded, we expected a high percentage of positive emotions in the total of 3506 emotion terms found (see Table 3 for distributions). Indeed, positive emotions and surprise dominate, indicating customer satisfaction above expectation (prevalence of surprise emotion).

Table 2. Emotions found in users' profiles (source: authors' own creation).

Emotion	Percentage (Counts)
positive emotion	56.5% (17,322)
negative emotion	4.8% (1482)
other emotion	5.8% (1798)
joy	7.3% (2239)
surprise	20.4% (6255)
anger	0.05% (176)
disgust	0.01% (32)
fear	1.1% (345)
sadness	3.2% (993)

Table 3. Emotions found in user reviews for fishing businesses (source: authors' own creation).

Emotion	Percentage (Counts)
positive emotion	53.4% (1875)
negative emotion	3.6% (128)
other emotion	5.0% (178)
joy	4.5% (161)
surprise	30.2% (1061)
anger	0.1% (6)
disgust	0.2% (8)
fear	0.5% (21)
sadness	1.9% (68)

4.1.3. Topic Detection

Understanding a customer basis emerges from knowing customers' needs and interests. This can be achieved by extracting topics and sub-topics of interests discussed or declared by users. For this task, we employed unsupervised learning techniques *k-means topic clustering* [41] and *Latent Dirichlet Allocation (LDA) topic modeling* [42]. Both algorithms define clusters of topics in a document collection given k , the predefined number of topics to be extracted. However, the main difference between the two methods is that k -means partitions the given documents into disjoint clusters (topics). At the same time, LDA assigns a document to a mixture of topics (one or more) with a representative percentage distribution [43]. We present the results at the collective fishing tourism business and user level to gain insights about the customer basis of the fishing tourism industry.

K -means topic clustering performed poorly in our data, as evaluated with elbow method and silhouette index. In particular, the elbow method reveals that the greater the number of k , the lower the sum of squared distances among clusters centroids, however, without converging to zero. Similar results are derived from the silhouette index evaluation, as the highest score achieved is lower than 0.5, indicating poor clustering, which is also evaluated manually. LDA typically requires a significant amount of text to identify well-defined topics. Despite the relatively small volume of reviews in our dataset, LDA-emerged topics seem more reasonable than the k -means approach. In our case, after manual experimentation and human judgment, we identified four distinct super topics as shown in Figure 10: hotel reviews and accommodation, food experience, general service quality, and fishing experiences. This reflects that fishing is a primary and general interest of users who visit fishing boat businesses. Another interesting insight regarding the third topic (fishing tourism) is that the included reviews are more personal than average since the most relevant term for this topic is actually the boat's owner's name. Regarding the remaining topics, it is worth mentioning that the personnel of a touristic entity is amongst the most relevant terms in both the first and second topics. This signifies the high importance of proper staffing, especially in the hotels and accommodation industry, where the term "staff" appears even higher than the hotel's amenities (rooms, pool, bar, restaurant, etc.).

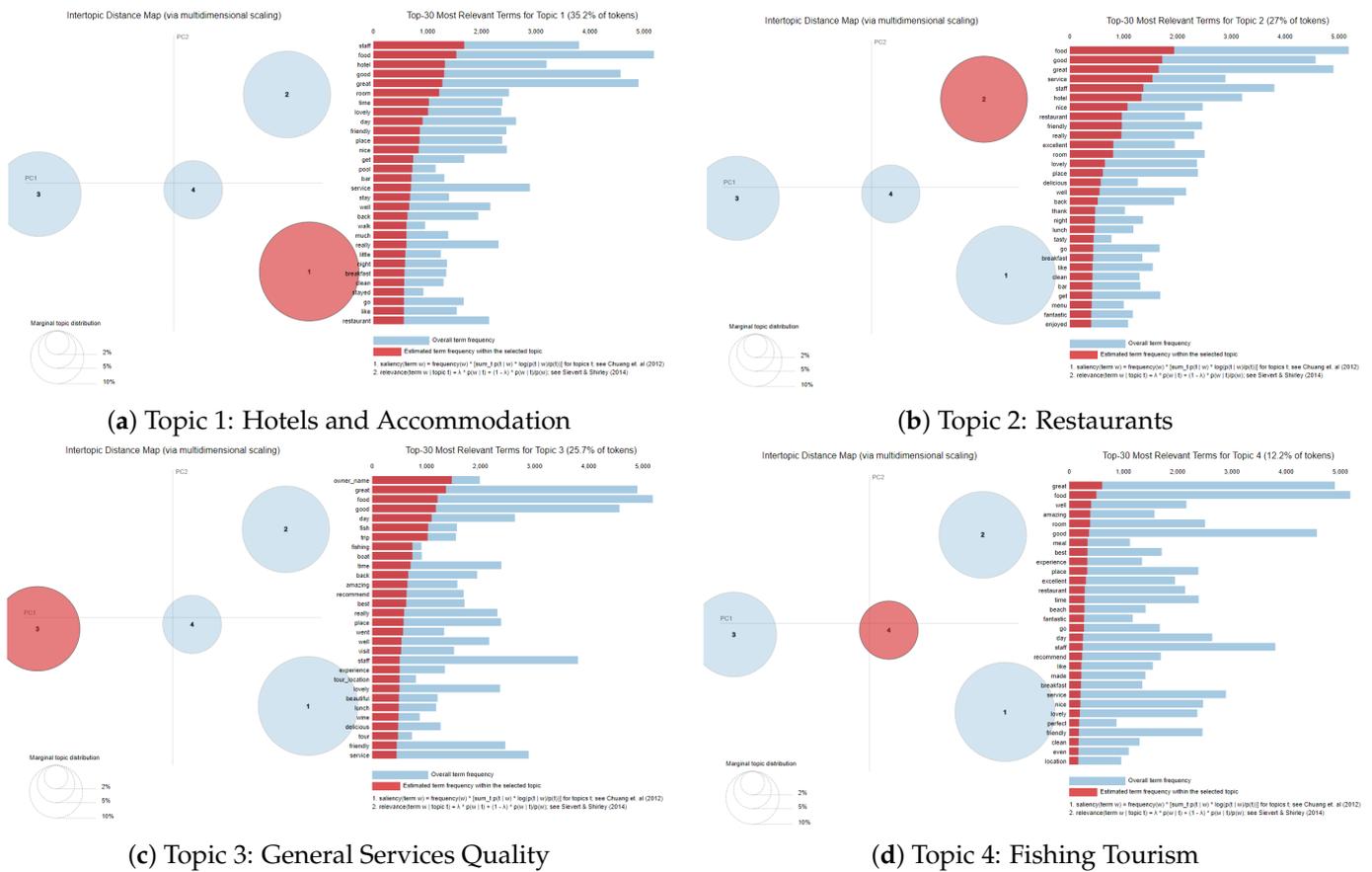


Figure 10. Intertopic distance map (via multidimensional scaling) [44,45] and top 30 most relevant terms for each topic as resulted from the LDA analysis (source: authors’ own creation). There are four topics identified in user reviews, namely, hotels and accommodation in (a), restaurants in (b), general services quality in (c), and fishing tourism in (d). Each topic is represented by a circle. The size of the circle refers to the prevalence of the topic in the data. Next to the topics, the words refer to the top 30 most relevant terms for each topic (colored in red).

4.2. User Profiling Aspects: Gender, Age and Marital Status

Demographic information of customers is a crucial aspect for any business to identify its audience and their specific needs and customize its products or/and services to maximize consumer satisfaction. Based on the information provided on users’ profiles, we were able to extract the explicitly declared information about their gender, age, and marital status. With these ground truth data for a subset of our dataset, we train classifiers to infer this demographic information for the rest of the users.

4.2.1. Gender Classification

Following the generic methodology presented in Figure 11, we describe the steps and the obtained results for the gender prediction task.

Ground truth. Individuals’ gender in our dataset could take one out of two labels: woman or man. Thus, our challenge is to build a binary classification model. As seen in Figure 12, in our dataset, 28.2% of our users are women, 18.4% are men, and 53.3% have no label available (nan). For this subset, we build the gender classifier to infer their gender.

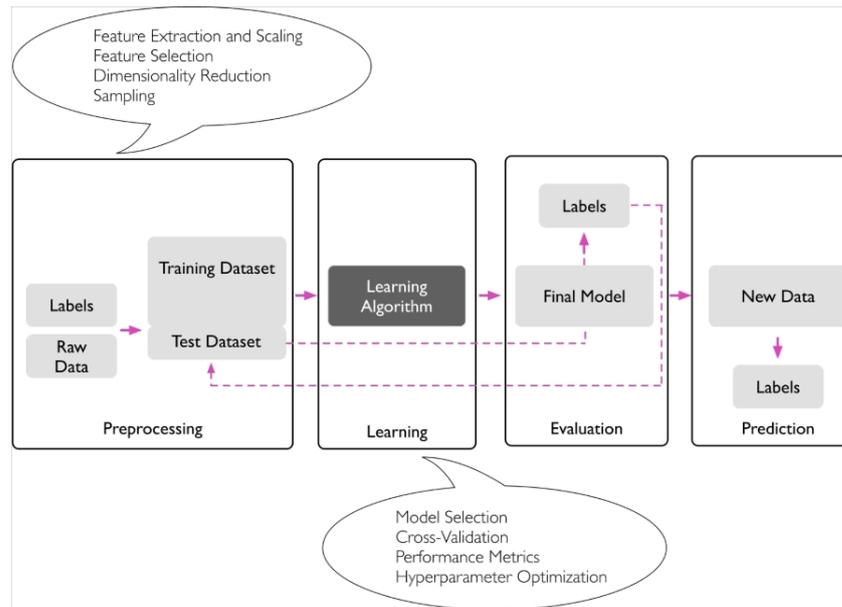


Figure 11. Typical machine learning process for predictive modeling (Figure courtesy of Raschka and Mirjalili [46]).

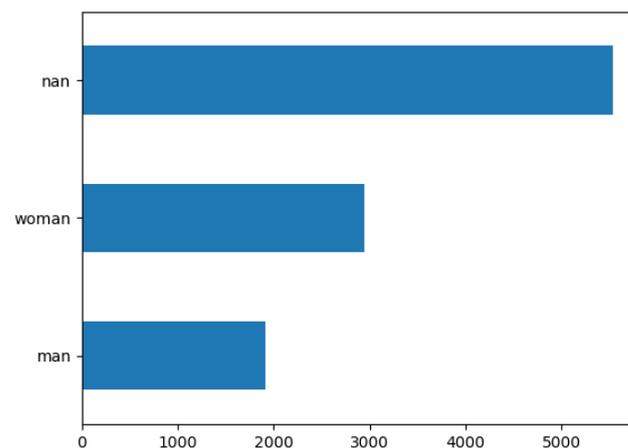


Figure 12. Gender labels in our dataset. The term “nan” signifies missing data (source: authors’ own creation).

Features extraction. In order to train our gender classifier, we need to extract the most appropriate features arising from related studies [47,48]. These can be summarized as follows:

- **Gender estimation based on name:** We train a naive Bayes model in order to obtain an estimation of each user’s gender based on their username according to names lexicons with female and male names. This is assumed to be a powerful feature for gender detection.
- **Sentiment score:** The aggregated score of reviews sentiment.
- **Syntax features:** Number of part of speech tags (adjectives, nouns, verbs, and adverbs).
- **Language vectors:** The frequency vectors of words used by each user. We construct TF-IDF and n-grams vectors reflecting the importance of words and phrases in a collection of documents based on preprocessed textual data.

Model training. We experiment with different types of classifiers, as mentioned above, such as logistic regression (lr), random forest (rf), and stochastic gradient descent (sdg). We split the ground truth to train and test sets while keeping 70% for training and 30% for testing each model. We follow cross-validation with 10 iterations in order to check

model accuracy. We evaluate the performance of our models with accuracy and f1 score. The results are presented in Table 4.

Table 4. Gender classifier performance on training and testing data. Performance is assessed via the accuracy score (ACC.) and the F1-Score (f1). The maximum performance is achieved by the logistic regression model (lr) in bold. (source: authors' own creation).

Classifier		Training	Testing
lr	acc.	0.84	0.73
	f1	0.83	0.72
rf	acc.	1.0	0.71
	f1	1.0	0.68
sgd	acc.	0.40	0.42
	f1	0.23	0.25

The best results are achieved with the logistic regression classifier achieving an accuracy score of 73%. In detail, the model is confused and is most likely to predict the man label falsely. After training, we apply the model to our dataset to annotate it and obtain labels for the nan cases. The gender-labeled data are used to extract insights about the 4Es.

4.2.2. Age Classifier

Following the same methodology, we describe the steps and the obtained results for the age prediction task.

Ground truth. Individuals' age exists in a total of 32% in our dataset and falls in one of the classes 18–24 (0.25%), 25–34 (4.6%), 35–49 (14.9%), 50–64 (12.3%). In Figure 13, we notice that there is a high imbalance in class frequency, and as a result, to build a valid classifier, we need to apply class imbalance machine learning techniques.

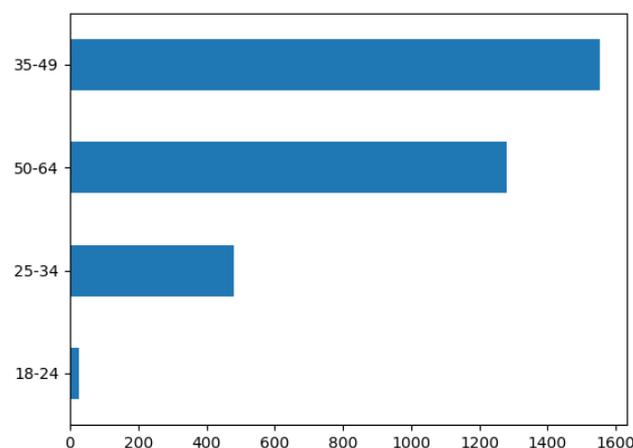


Figure 13. Age labels in our dataset (source: authors' own creation).

Features extraction. Similarly, with the case of the gender classifier, we need to extract the most appropriate features arising from related studies [48] to build the age classifier. These can be summarized as follows:

- **Structure features:** Refer to the structural use of language. These features include the number of words in each review, the number of characters, the number of words in a sentence, and the number of exclamation marks.
- **Syntax features:** Refer to the number of parts of speech tags (adjectives, nouns, verbs, and adverbs).
- **Sentiment score:** The aggregated score of reviews sentiment.
- **Readability features:** Refer to the level of the text complexity. We include (a) Flesch reading ease, indicating how easy is a text to read, (b) Smog index, estimating the years

of education needed to understand a piece of writing, (c) Flesch–Kincaid grade, indicating the average student in that grade level that can read the text, (d) Coleman–Liau index, gauging the understandability of a text, (e) automated readability index, assessing the understandability of a text, (f) Dale–Chall readability score, providing a numeric gauge of the comprehension difficulty that readers come upon when reading a text, (g) difficult words, indicating how many difficult words are used in a text, and (h) gunning fog, estimating the years of formal education a person needs to understand the text on the first reading.

- **Language vectors:** The frequency vectors of words used by each user. We construct TF-IDF and n-grams vectors reflecting the importance of words and phrases in a collection of documents based on preprocessed textual data.

Imbalance learning. As already mentioned, our ground-truth dataset is imbalanced. We experiment with different techniques to handle this challenge as oversampling and undersampling, use of synthetic examples, etc., and we conclude that the Synthetic Minority Oversampling Technique (SMOTE), which synthesizes new examples for the minority class, gives the best results. Specifically, SMOTE works by selecting examples that are close to the feature space, drawing a line between the examples in the feature space, and drawing a new sample at a point along that line. A random example from the minority class is first chosen. Then, k of the nearest neighbors for that example is found (typically $k = 5$). A randomly selected neighbor is chosen, and a synthetic example is created at a randomly selected point between the two examples in feature space [49]. We present our results with and without the use of the SMOTE technique. Our sampling strategy follows the rule of oversampling the minority class up to 10% of the majority class.

Model training. We experiment with different types of classifiers as mentioned above, such as logistic regression (lr), random forest (rf), and stochastic gradient descent (sgd). We split the ground truth to train and test sets while keeping 70% for training and 30% for testing each model. We follow cross-validation with 10 iterations in order to check the model’s accuracy. We evaluate the performance of our models with the accuracy and f1 score. The results are present in Table 5.

Table 5. Age classifier performance on training and testing data. Performance is assessed via the accuracy score (ACC.) and the F1-Score (f1). The maximum performance is achieved by the random forest model (lr) in bold. (source: authors’ own creation).

Classifier		Training	Testing
lr	acc.	0.47	0.46
	f1	0.42	0.42
rf	acc.	1.0	0.62
	f1	1.0	0.59
sgd	acc.	0.44	0.46
	f1	0.32	0.34

The best results are achieved with the random forest classifier, achieving an accuracy score of 62%. Specifically, the model is confused and is most likely to falsely predict the age categories 35–49 and 50–64. After training, we apply the model to our dataset to annotate it and obtain labels for the unlabeled cases. The age-labeled data are used to extract insights about the 4Es.

4.2.3. Marital Status Detection

Extracting the marital status of each reviewer is a detection task since we do not have a ground truth to train a machine learning model and infer the status of the travelers in our dataset. As a result, based on a lexicon approach, we define vocabulary sets to detect if the reviewer uses words or phrases that indicate if they travel with their *family* or their

partner. In the cases where there is not any relevant information available, we assign the class *unknown*.

We define the marital classes as follows: (a) family travelers (those who travel with kids)—use of words like children, child, kid(s), son, daughter, dad, father, mom, mum, and mother; (b) couple travelers (those who travel with their partner, no kids)—use of words like husband, wife, spouse, girlfriend, boyfriend, partner; and (c) unknown, for the rest of the cases, which can be solo travelers, a group of people traveling together, or any other case.

As seen in Figure 14, we detect family travelers at 22.9%, couple travelers at 7.3%, and the unknown class at 70%. We present the tourism preferences of different marital statuses as well as the relationship with demographics and 4E classes.

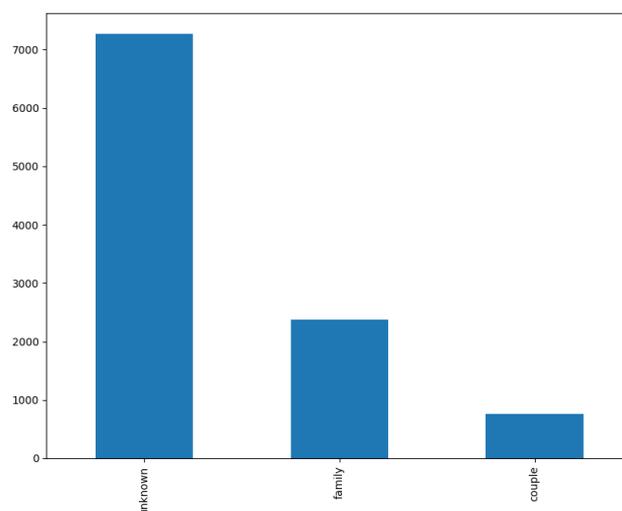


Figure 14. Marital status distribution (source: authors' own creation).

4.3. The 4Es of Experience Economy

In this section, we dive into the different experience realms of the fishing tourism industry as seen through the lens of TripAdvisor tourist reviews. To acquire information regarding the experience dimensions, i.e., “Educational”, “Entertainment”, “Aesthetic”, “Escapist”, expressed in each review, we manually annotate 240 user reviews. In other words, human annotators read and annotate a number of reviews each, according to the four dimensions of the Experience Economy.

The heat map in Figure 15 visualizes the results of this annotation process in the form of a frequency matrix for the four dimensions. We notice that “Entertainment” is by far the prevalent dimension in the tourists’ reviews (88% of reviews), followed by “Aesthetic” (34% of reviews), “Educational” (32% of reviews), and “Escapist” (19% of reviews). Not surprisingly, when it comes down to the most prevalent pairs of dimensions, the pairs of “Entertainment”–“Aesthetic” and “Educational”–“Entertainment” are the most frequent, co-existing in 30% and 28% of the reviews, respectively. However, to obtain a better idea of the true co-existence of dimensions that is not biased by the frequency of appearance of the individual experience realms, we visualize the Jaccard similarity index for each dimension pair in Figure 16. The Jaccard similarity index “compares members for two sets to see which members are shared and which are distinct”. It is “a measure of similarity for the two sets of data, with a range from 0 to 1” (<https://www.statisticshowto.com/jaccard-index/> accessed on 20 July 2021). The higher the percentage, the more similar the two populations. In the case of tourist reviews, the higher the Jaccard index, the more the reviews are shared among the two dimensions. Now, apart from the pairs we discussed previously, another common pair emerges, specifically the “Aesthetic”–“Escapist” combination (Jaccard index of 0.26), denoting the co-existence of the “Aesthetic” and “Escapist” experiences in the fishing tourism reviews.

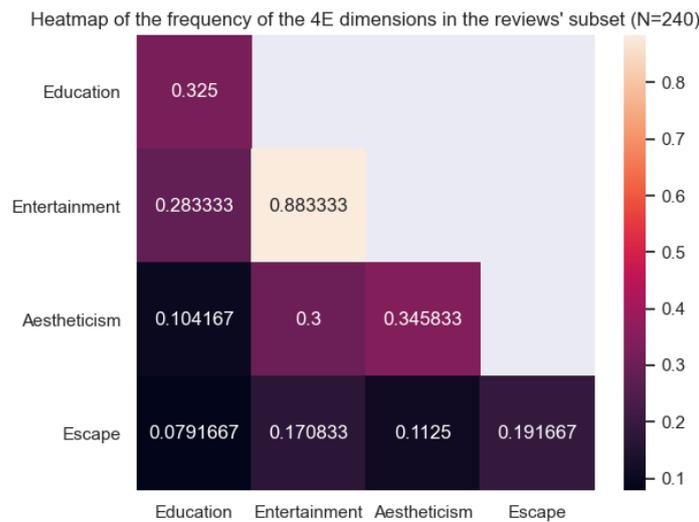


Figure 15. Frequency matrix for the four dimensions of the Experience Economy (source: authors' own creation).

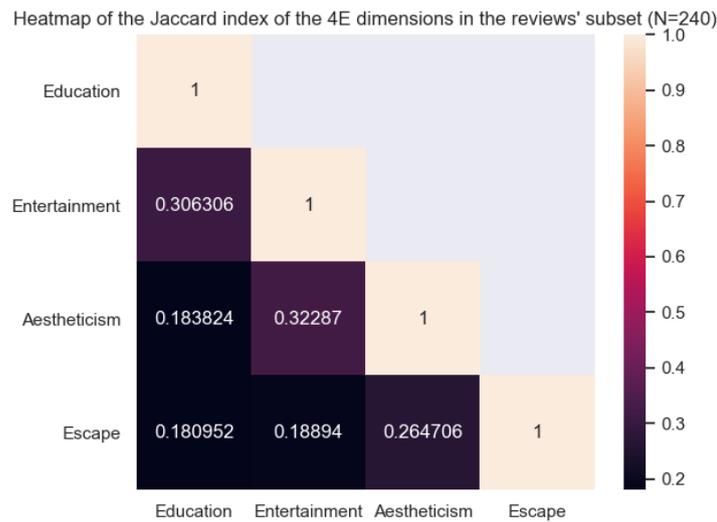


Figure 16. Jaccard similarity matrix for the four dimensions of the Experience Economy (source: authors' own creation).

With regards to the demographics of the users within each experience realm, we visualize the age, gender, and marital status of the reviewers in Figure 17, Figure 18 and Figure 19, respectively. Regarding age, the distribution of age groups appearing in the different experience realms does not vary significantly. There is only a slight difference in the Entertainment realm, where the oldest age group (50–64) has a larger percentage compared to the other realms, potentially declaring that fishing tourism is particularly enjoyable for older adults. Similarly, the age distribution does not vary significantly between the different dimensions of the Experience Economy, with most reviews coming from female users, as seen in Figure 18. Lastly, concerning marital status, the percentage of reviews coming from couples is almost steady across all dimensions. However, when it comes to family users, the “Aesthetic” and “Escapist” dimensions show higher percentages, meaning that these experiences are possibly felt more intensely by this user group.

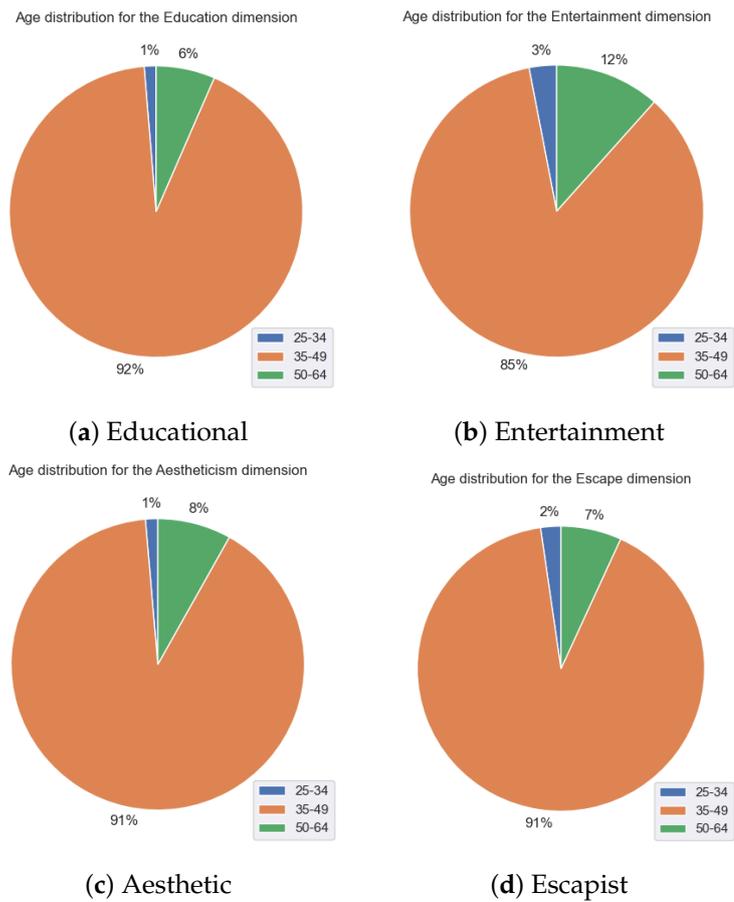


Figure 17. The age distribution of reviewers per experience realm (source: authors' own creation).

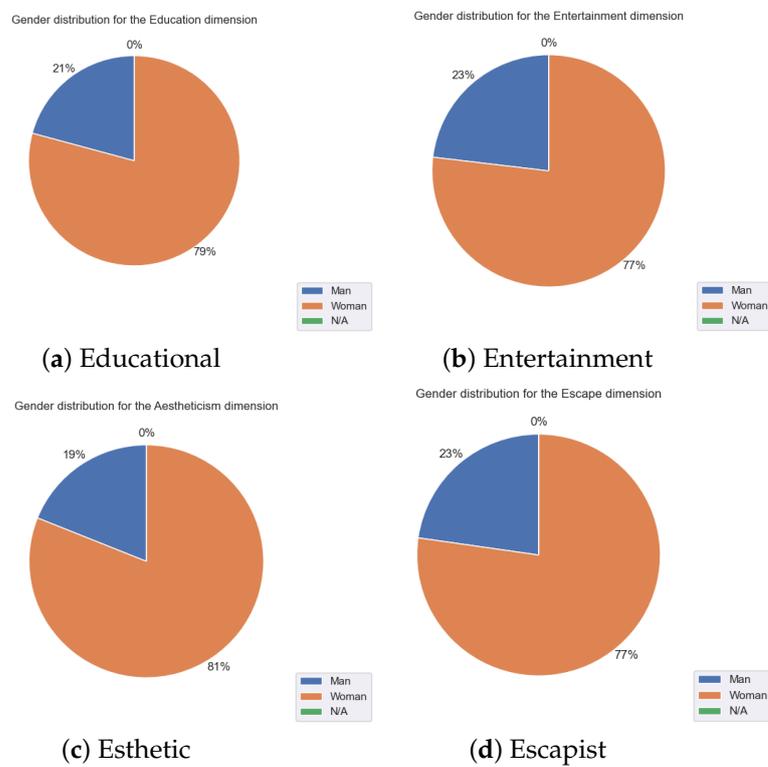


Figure 18. The gender distribution of reviewers per experience realm (source: authors' own creation).

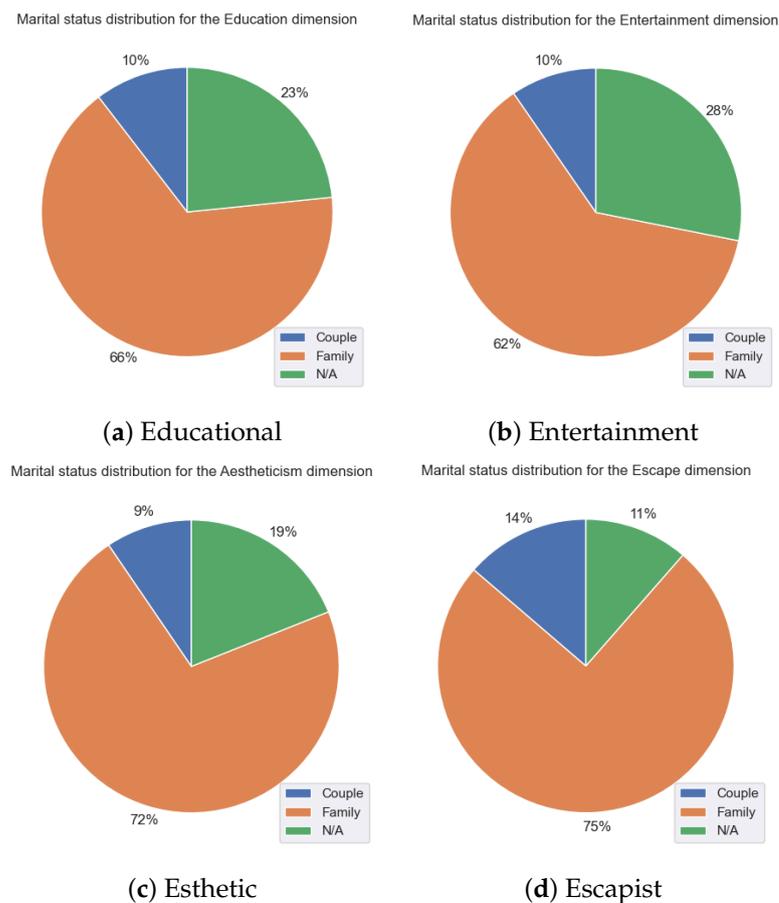


Figure 19. The marital distribution of reviewers per experience realm (source: authors' own creation).

Finally, the linguistic exploration of reviews belonging to different realms did not produce any significant results.

5. Discussion and Conclusions

Tourists went on boat trips mostly for fishing and actively participated in the process (fishing experience, caught fish, etc). The overall experience of fishing boat trips is highly recommended, as tourists mention that they had a “great time” and “fantastic day”. Tourists appreciate the beauty of the natural environment by leaving positive comments about the “crystal waters”, the sea, the fresh fish, etc. They highlight the hospitality and the skills of the crew and business owners. These reviews reflect the positive tourist experience translated into star ratings from the tourists. It is evident that the vast majority of tourists leave 5-star rated reviews. Tourists overall put emphasis on commenting on the offered services. Emotion analysis shows that users are more likely to express positive emotions in their reviews, specifically surprise and joy, while negative emotions are expressed less frequently. Indeed, positive emotions and surprise dominate, indicating customer satisfaction above expectation.

Interestingly, approximately half the users belong to the “New Reviewer” category, meaning that they only join the TripAdvisor platform to positively review the respective fishing tourism business. The users who review fishing tourism businesses also tend to review hotels and attractions, as well as luxury and boutique hotels, and B&B and Inns at a smaller scale. Fishing is a primary and general interest of users who visit fishing boat businesses. Another interesting insight regarding fishing tourism is that the included reviews are more personal than average since the most relevant term for this topic is actually the boat owner’s name.

“Entertainment” is by far the prevalent dimension in the tourists’ reviews, followed by “Aesthetic”, “Educational” and “Escapist”. Not surprisingly, when it comes down to the most prevalent pairs of dimensions, the pairs of “Entertainment”–“Aesthetic” and “Educational”–“Entertainment” are the most frequent. Another common pair emerges, specifically the “Aesthetic”–“Escapist” combination, denoting the co-existence of the “Aesthetic” and “Escapist” experiences in the fishing tourism reviews. To truly transfer value as a business or organization, we need to understand and act on the perceptions of customers about quality and value, the process of creating that value, and efficient and effective management of the same time resources in order to create this value [50].

Nowadays, fishing trip organizers list, on specialized websites, the fishing package or packages they intend to offer, i.e., detailing what they will provide to the client (tourist), in which areas, and at what cost. The client can make an on-the-spot online booking and payment, and the fisher will be informed immediately about the booking by email or SMS. Many fishing trip organizers have already signed contracts, and several are in the process of signing. These websites also cover tourists who are interested not only in fishing but also in being shown a way of fishing or even being trained in these types of fishing. In particular, in Greece, all the forms of fishing that are possible are included: (a) fishing from boats, (b) kayak fishing, (c) fishing from the shore of all types (Casting, English, Spinning, etc.), (d) fishing in lakes and rivers, and (e) spearfishing. Each new tourist activity goes through a period of evolution and adaptation within the community, earning the participants’ loyalty and reinforcing repeated tourism, which contributes to the development and improvement of the local tourism market. Fishing tourism as a new tourist activity provides potential benefits from fishing tourism to local communities. Therefore, the necessary attention should be given to developing a strategy in this tourism sector. This development will bring new opportunities and challenges; it can offer professional opportunities to fishers and will increase the rural economy if governments provide adequate infrastructure, leadership, legislative, and financial support that will set the foundation for sustainable development in the long term [51]. The combination of fishing tourism and marine reserves emerges as the optimal strategy, and the presence of visitors in these areas generates larger profits than if only fishing is considered [52].

Theoretical implications on fishing tourism can encompass a wide range of considerations that relate to the interactions between tourism and fishing activities. These implications may arise from various academic disciplines, including economics, ecology, sociology, and environmental science. Here are some theoretical implications to consider: the combination of the dimensions and the opportunity of developing value on recreational fishing tourism. Moreover, there are many impacts which are now open for research in various areas. Economic impact: Fishing tourism can have multiplier effects on local economies. Tourists’ expenditures on accommodations, equipment rental, guides, and other services can generate indirect and induced economic impacts in the host community. Ecological Impact: Fishing tourism can lead to overfishing if not managed sustainably. The theoretical implications involve the need for effective regulations and strategies to prevent the over-exploitation of fish populations and protect marine ecosystems. Ecotourism: Sustainable fishing tourism models can contribute to the conservation of natural resources by fostering awareness and appreciation for aquatic ecosystems.

There is no single magic solution to the crisis facing small-scale fisheries: action is needed on many fronts to make fisheries sustainable. However, sustainable fishing tourism is an increasingly popular activity around the world. Sustainable fishing tourism has different names in different countries—Pescaturismo, Pescaturisme, Pêchetourisme, Pesca Vivencial, Experiential Fishing, Ribolovni turizam, etc.—but the concept remains the same: it is only intended for professional fishers, allowing the diversification of their activities while continuing their traditional trade. This alternative income stream should reduce the intensity of fishing activities, contribute to sustainable management of fishery resources, and promote the cultural heritage of artisanal fishing (https://www.wwfmmi.org/what_we_do/fisheries/transforming_small_scale_fisheries/sustainable_fishing_tourism/ accessed on

20 July 2021). Ecotourism should be further supported by policy arrangements in order to promote sustainability of the marine resources. Indeed, in Greek waters, FT is considered to add to the income of small-scale fishers; however, as shown by existing research [53], only a small fraction of fishers can be involved in this activity, while the vast majority will keep their traditional fishing practice. Thus, in order for FT to really contribute to fishing communities in a just and equitable way, policy provisions should exclude those gaining from FT from commercial fishing operations during the days in which they act as FT operators. This seems to be a fair solution that would limit the effect of fishing efforts on the fishery resources and lower competition with the rest of the fishers who will not be able to increase their income through their involvement in FT operations.

Sociocultural Impact:

- **Cultural Exchange:** Fishing tourism can facilitate cultural exchange between tourists and local fishing communities, leading to mutual understanding and preservation of traditional practices. This implies the importance of promoting respectful interactions and cultural sensitivity.
- **Community Livelihoods:** The theoretical implications involve examining how fishing tourism affects the livelihoods and well-being of local communities. Sustainable fishing tourism can enhance income diversification and improve quality of life.
- **Tourism Management:** Theoretical discussions revolve around determining the carrying capacity of fishing tourism destinations to ensure that environmental and social impacts are kept within sustainable limits.
- **Stakeholder Engagement:** Effective stakeholder engagement is crucial for managing fishing tourism. The implications include the need for collaboration among governments, local communities, tour operators and conservation organizations in order to promote a fair transition to more 'green' occupations through the change in legislation.
- **Environmental Ethics:** Theoretical implications extend to ethical discussions about catch-and-release practices, the welfare of targeted fish species, and the broader ecological consequences of fishing tourism.
- **Conservation and Research:** Fishing tourism can provide opportunities for scientific research, such as studying fish populations, migration patterns, and ecosystem dynamics. Theoretical implications emphasize the role of fishing tourism in advancing marine conservation efforts.
- **Education and Outreach:** Fishing tourism can serve as a platform for educating tourists and the public about the importance of marine conservation, fostering a sense of responsibility and support for preserving aquatic ecosystems.
- **Climate Change Adaptation:** Theoretical implications may explore how fishing tourism destinations need to adapt to changing climate conditions, such as shifts in fish distribution and abundance, and how these changes could affect tourism experiences and local economies.

It has been pointed out that the provision of services related to commercial fishing is to the detriment of the marine ecosystem mainly due to overfishing [54]. Rural tourism development is struggling with competition and change due to fragmented structures, uncooperative short-term businesses, numerous small enterprises with lagging infrastructure, a lack of governance system, and a lack of control between man-made development and management of nature, in addition to several complexities, such as the need to implement fisheries-based tourism as a means of subsistence, the need for infrastructural development, the lack of investment support, the risk of environmental damage, the utter lack of policies to regulate and promote the industry, and the requirement to consider the concerns of various stakeholders [9]. It is important to note that the actual implications of fishing tourism will depend on factors such as destination characteristics, management strategies, regulatory frameworks, and the behaviors of tourists and local communities. Researchers and practitioners in various fields continue to explore these implications to ensure that fishing tourism contributes positively to both the environment and the well-being of host communities. The promotion activities of some actions, such as sports tourism

and especially fishing tourism, should be related to cohesion policy/community-led local development, smart specialization strategies. In particular, as far as smart specialization is concerned, research/academia stakeholders based on CLLD/RIS3 can conduct targeted research that provides added value in effectively advising towards achieving the sustainable development goals of the 2030 Agenda.

Author Contributions: Conceptualization, G.Y., S.Y., M.M. and A.L.; methodology, S.Y., G.Y. and D.K.; software, S.Y. and D.K.; validation, G.Y., S.Y., M.M. and A.L.; formal analysis, G.Y., S.Y., M.M. and A.L.; investigation, G.Y., S.Y., M.M., V.V. and A.L.; resources, G.Y. and P.B.; data curation, S.Y. and D.K.; writing—original draft preparation, G.Y., S.Y. and D.K.; writing—review and editing, G.Y., P.B. and A.P.; visualization, S.Y. and D.K.; supervision, G.Y. and S.Y.; project administration, G.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable, as the study did not involve any data acquired through intervention or interaction with individual human subjects, or any identifiable private information—all data collected were anonymous.

Informed Consent Statement: Subject consent was waived due to the study not involving any intervention or interaction with individual human subjects.

Data Availability Statement: Data are unavailable due to privacy restrictions.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

API	Application Programming Interface
DUTH	Democritus University of Thrace
HCMR	Hellenic Centre for Marine Research
HTML	HyperText Markup Language
LDA	Latent Dirichlet Allocation
NLP	Natural Language Processing
POS	Part Of Speech
SMOTE	Synthetic Minority Oversampling Technique
TF-IDF	Term Frequency-Inverse Document Frequency
XML	Extensible Markup Language

References

1. Pine, B.J.; Pine, J.; Gilmore, J.H. *The Experience Economy: Work is Theatre & Every Business a Stage*; Harvard Business Press: Boston, MA, USA, 1999.
2. Seeler, S.; Schänzel, H.A. From the fluidity of the experience economy to transformative experiences as a catalyst for positive change: The future of experiential travel. *J. Tour. Futur.* **2019**, *5*, 111–113. [[CrossRef](#)]
3. Mehmetoglu, M.; Engen, M. Pine and Gilmore's concept of experience economy and its dimensions: An empirical examination in tourism. *J. Qual. Assur. Hosp. Tour.* **2011**, *12*, 237–255. [[CrossRef](#)]
4. Jafari, J. *Encyclopedia of Tourism*; Routledge: London, UK, 2002.
5. Centre for the Promotion of Imports from Developing Countries. *The European Market Potential for Sports Tourism*; Ministry of Foreign Affairs of the Netherlands: The Hague, The Netherlands, 2021.
6. Ali-Knight, J. The Role of Niche Tourism Products in Destination Development. Ph.D. Thesis, Edinburgh Napier University, Edinburgh, UK, 2010.
7. Riddick, C.C.; Russell, R.V. Research in Recreation, Parks, Sport, and Tourism. *J. Leis. Res.* **2011**, *43*, 148.
8. Tsafoutis, D.; Metaxas, T. Fishing tourism in Greece: Defining possibilities and prospects. *Sustainability* **2021**, *13*, 13847. [[CrossRef](#)]
9. Azmi, E.; Che Rose, R.A.; Awang, A.; Abas, A. Innovative and competitive: A systematic literature review on new tourism destinations and products for tourism supply. *Sustainability* **2023**, *15*, 1187. [[CrossRef](#)]
10. Smith, M.; Fedler, A.J.; Adams, A.J. Economic assessments of recreational flats fisheries provide leverage for conservation. *Environ. Biol. Fishes* **2023**, *106*, 131–145. [[CrossRef](#)]
11. Holbrook, M.B.; Hirschman, E.C. The experiential aspects of consumption: Consumer fantasies, feelings, and fun. *J. Consum. Res.* **1982**, *9*, 132–140. [[CrossRef](#)]

12. Pine, B.J.; Gilmore, J.H. *Welcome to the Experience Economy*; Harvard Business Review Press: Cambridge, MA, USA, 1998; Volume 76.
13. Pikkemaat, B.; Weiermair, K. The aesthetic (design) orientated customer in tourism-implications for product development. In Proceedings of the EIASM 10th International Product Development Management Conference, EIASM, Brussels, Belgium, 10–11 June 2003; Volume 825, p. 839.
14. Menner, T.; Höpken, W.; Fuchs, M.; Lexhagen, M. Topic detection: Identifying relevant topics in tourism reviews. In *Information and Communication Technologies in Tourism 2016*; Springer: Cham, Switzerland, 2016; pp. 411–423.
15. Rossetti, M.; Stella, F.; Zanker, M. Analyzing user reviews in tourism with topic models. *Inf. Technol. Tour.* **2016**, *16*, 5–21. [[CrossRef](#)]
16. Afzaal, M.; Usman, M.; Fong, A.C.; Fong, S. Multiaspect-based opinion classification model for tourist reviews. *Expert Syst.* **2019**, *36*, e12371. [[CrossRef](#)]
17. Afzaal, M.; Usman, M.; Fong, A. Predictive aspect-based sentiment classification of online tourist reviews. *J. Inf. Sci.* **2019**, *45*, 341–363. [[CrossRef](#)]
18. Yu, C.; Zhu, X.; Feng, B.; Cai, L.; An, L. Sentiment Analysis of Japanese Tourism Online Reviews. *J. Data Inf. Sci.* **2019**, *4*, 89–113. [[CrossRef](#)]
19. Marrese-Taylor, E.; Velásquez, J.D.; Bravo-Marquez, F. A novel deterministic approach for aspect-based opinion mining in tourism products reviews. *Expert Syst. Appl.* **2014**, *41*, 7764–7775. [[CrossRef](#)]
20. Kirilenko, A.P.; Stepchenkova, S.O.; Kim, H.; Li, X. Automated sentiment analysis in tourism: Comparison of approaches. *J. Travel Res.* **2018**, *57*, 1012–1025. [[CrossRef](#)]
21. Fang, B.; Ye, Q.; Kucukusta, D.; Law, R. Analysis of the perceived value of online tourism reviews: Influence of readability and reviewer characteristics. *Tour. Manag.* **2016**, *52*, 498–506. [[CrossRef](#)]
22. Shin, S.; Du, Q.; Xiang, Z. What's vs. how's in online hotel reviews: Comparing information value of content and writing style with machine learning. In *Information and Communication Technologies in Tourism 2019*; Springer: Cham, Switzerland, 2019; pp. 321–332.
23. Leal, F.; González-Vélez, H.; Malheiro, B.; Burguillo, J.C. Semantic profiling and destination recommendation based on crowd-sourced tourist reviews. In Proceedings of the International Symposium on Distributed Computing and Artificial Intelligence, Porto, Portugal, 21–23 June 2017; Springer: Cham, Switzerland, 2017; pp. 140–147.
24. Guerreiro, J.; Rita, P. How to predict explicit recommendations in online reviews using text mining and sentiment analysis. *J. Hosp. Tour. Manag.* **2020**, *43*, 269–272. [[CrossRef](#)]
25. Dong, X.; Li, T.; Song, R.; Ding, Z. Profiling users via their reviews: An extended systematic mapping study. *Softw. Syst. Model.* **2020**, *20*, 49–69. [[CrossRef](#)]
26. Madke, N.B.; Kulkarni, M.Y.; Mule, M.R.; Lakade, M.A.; Kulkarni, M.S. User Profile Based Behavior Identification Using Data Mining Technique. *Int. Res. J. Eng. Technol. (IRJET)* **2018**, *5*, 326–331.
27. Farnadi, G.; Tang, J.; De Cock, M.; Moens, M.F. User profiling through deep multimodal fusion. In Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining, Marina Del Rey, CA, USA, 5–9 February 2018; pp. 171–179.
28. Alkan, O.; Daly, E. User Profiling from Reviews for Accurate Time-Based Recommendations. *arXiv* **2020**, arXiv:2006.08805.
29. Kavitha, S.; Jobi, V.; Rajeswari, S. Tourism recommendation using social media profiles. In *Artificial Intelligence and Evolutionary Computations in Engineering Systems*; Springer: Cham, Switzerland, 2017; pp. 243–253.
30. Missaoui, S.; Viviani, M.; Faiz, R.; Pasi, G. A language modeling approach for the recommendation of tourism-related services. In Proceedings of the Symposium on Applied Computing, Marrakech, Morocco, 3–7 April 2017; pp. 1697–1700.
31. Fazzolari, M.; Petrocchi, M. A study on online travel reviews through intelligent data analysis. *Inf. Technol. Tour.* **2018**, *20*, 37–58. [[CrossRef](#)]
32. Oh, H.; Fiore, A.M.; Jeoung, M. Measuring experience economy concepts: Tourism applications. *J. Travel Res.* **2007**, *46*, 119–132. [[CrossRef](#)]
33. Quadri-Felitti, D.L.; Fiore, A.M. Destination loyalty: Effects of wine tourists' experiences, memories, and satisfaction on intentions. *Tour. Hosp. Res.* **2013**, *13*, 47–62. [[CrossRef](#)]
34. Thanh, T.V.; Kirova, V. Wine tourism experience: A netnography study. *J. Bus. Res.* **2018**, *83*, 30–37. [[CrossRef](#)]
35. Kozinets, R.V. The field behind the screen: Using netnography for marketing research in online communities. *J. Mark. Res.* **2002**, *39*, 61–72. [[CrossRef](#)]
36. Rageh, A.; Melewar, T.; Woodside, A. Using netnography research method to reveal the underlying dimensions of the customer/tourist experience. *Qual. Mark. Res. Int. J.* **2013**, *16*, 126–149. [[CrossRef](#)]
37. Gibson, H.; Yiannakis, A. Tourist roles: Needs and the lifecourse. *Ann. Tour. Res.* **2002**, *29*, 358–383. [[CrossRef](#)]
38. Jivani, A.G. A comparative study of stemming algorithms. *Int. J. Comp. Tech. Appl.* **2011**, *2*, 1930–1938.
39. Ekman, P.; Sorenson, E.R.; Friesen, W.V. Pan-cultural elements in facial displays of emotion. *Science* **1969**, *164*, 86–88. [[CrossRef](#)] [[PubMed](#)]
40. Loria, S. Textblob Documentation. *Release 0.15* **2018**, *2*, 269.
41. Likas, A.; Vlassis, N.; Verbeek, J.J. The global k-means clustering algorithm. *Pattern Recognit.* **2003**, *36*, 451–461. [[CrossRef](#)]

42. Ramage, D.; Hall, D.; Nallapati, R.; Manning, C.D. Labeled LDA: A supervised topic model for credit attribution in multi-labeled corpora. In Proceedings of the 2009 conference on Empirical Methods in Natural Language Processing, Singapore, 6–7 August 2009; pp. 248–256.
43. Kelaiaia, A.; Merouani, H.F. Clustering with probabilistic topic models on arabic texts: A comparative study of LDA and K-means. *Int. Arab J. Inf. Technol.* **2016**, *13*, 332–338.
44. Chuang, J.; Manning, C.D.; Heer, J. Termite: Visualization techniques for assessing textual topic models. In Proceedings of the International Working Conference on Advanced Visual Interfaces, Capri Island, Italy, 21–25 May 2012; pp. 74–77.
45. Sievert, C.; Shirley, K. LDAvis: A method for visualizing and interpreting topics. In Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces, Baltimore, MD, USA, 27 June 2014; pp. 63–70.
46. Raschka, S.; Mirjalili, V. *Python Machine Learning*, 2nd ed.; Packt Publishing: Birmingham, UK, 2017.
47. Marquardt, J.; Farnadi, G.; Vasudevan, G.; Moens, M.F.; Davalos, S.; Teredesai, A.; De Cock, M. Age and gender identification in social media. *Proc. CLEF 2014 Eval. Labs* **2014**, *1180*, 1129–1136.
48. Weren, E.R.; Kauer, A.U.; Mizusaki, L.; Moreira, V.P.; de Oliveira, J.P.M.; Wives, L.K. Examining multiple features for author profiling. *J. Inf. Data Manag.* **2014**, *5*, 266.
49. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [[CrossRef](#)]
50. Grönroos, C. Creating a relationship dialogue: Communication, interaction and value. *Mark. Rev.* **2000**, *1*, 5–14. [[CrossRef](#)]
51. Yfantidou, G.; Matarazzo, M. The future of sustainable tourism in developing countries. *Sustain. Dev.* **2017**, *25*, 459–466. [[CrossRef](#)]
52. Falcó, C.; Moeller, H.V. Optimal spatial management in a multiuse marine habitat: Balancing fisheries and tourism. *Nat. Resour. Model.* **2022**, *35*, e12309. [[CrossRef](#)]
53. Lontakis, A.; Vassilopoulou, V. Exploring fishing tourism sustainability in North-Eastern Mediterranean waters, through a stochastic modelling analysis: An opportunity for the few or a viable option for coastal communities? *Ocean Coast. Manag.* **2022**, *221*, 106118. [[CrossRef](#)]
54. Hyder, K.; Radford, Z.; Prellezo, R.; Weltersbach, M.S.; Lewin, W.C.; Zarauz, L.; Ferter, K.; Ruiz, J.; Townhill, B.; Mugerza, E.; et al. *Research for PECH Committee-Marine Recreational and Semi-Subsistence Fishing-Its Value and Its Impact on Fish Stocks*; European Parliament, Policy Department for Structural and Cohesion Policies: Brussels, Belgium, 2017.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.