JOURNAL OF THE FACULTY OF TOURISM AND HOTELS
UNIVERSITY OF SADAT CITY
Journal homepage:
https://mfth.journals.ekb.eg/

Leveraging Big Data for Content Analysis of User-Generated Content to Evaluate Visitor Satisfaction at Egyptian Cultural Heritage Sites

Iten ElRouby

Associate Professor - Tourism Studies Department - Faculty of Tourism & Hotels - Alexandria University

itenelrouby@alexu.edu.eg

Printed ISSN: 2537-0952
Online ISSN: 3062-5262
DOI: 10.21608/MFTH.2025.452620

ABSTRACT

Cultural heritage includes traditions, monuments, artifacts, and lifestyles that are transmitted from past generations and are significant to a community's identity. In recent years, the integration of big data analytics into the field of cultural heritage has opened new pathways for research, preservation, and engagement. Big data, which consists of large repositories of information generated by social media platforms, can be effectively utilized to extract valuable insights for both organizations and destinations, thereby increasing customer satisfaction and enhancing the quality of services offered.

This study aims to evaluate the cultural heritage experience at one of the world's most renowned heritage sites, the Giza Pyramids, by utilizing big data to analyze user-generated content from TripAdvisor.

In order to answer the research questions, NVivo14 was used to analyze customer reviews spanning the period from 2018 to 2024. NVivo 14 is an analysis software that works with unstructured qualitative data to analyze user sentiments. The software focuses on trends in word occurrences and identifies themes within the text.

The analysis with Nvivo14 revealed that positive sentiments predominated, as reflected in frequent references such as "perfect," "amazing," "breathtaking," and "great." The word frequency analysis conducted identified the most commonly used terms in the dataset, providing an understanding of the primary topics discussed by visitors. The word frequency query highlighted recurring words such as "pyramid(s)," "guide," "camel(s)," "great," "experience," "amazing," and

"ride(s)". The thematic analysis uncovered central themes emerging from the TripAdvisor reviews like "Visitor Experience", "Aesthetic and Cultural Appreciation," and "Construction and Grandeur". However, some reviews showed concerns, particularly regarding "overcrowding" and "persistent vendors."

The research suggests that the government should enforce solid measures to enhance visitor satisfaction. These measures may include regulating traffic at the site, managing vendors more effectively, and closely supervising camel and horse rides. Additionally, the study highlights the importance of utilizing big data to improve customer experience in cultural heritage sites.

KEYWORDS

Big Data, Content Analysis, User-Generated Content (UGC), Giza Pyramids, Heritage Sites, NVivo14, TripAdvisor.

الاستفادة من البيانات الضخمة في تحليل المحتوي الرقمي الُمنشأ بواسطة المستخدمين لتقييم رضا الزوار لمواقع التراث الثقافي المصرية

ايتن الروبي

استاذ مساعد- قسم الدراسات السياحية – كلية السياحة و الفنادق -جامعة الاسكندرية

itenelrouby@alexu.edu.eg

الترقيم الدولى الموحد للطباعة:
2537–0952
الترقيم الدولى الموحد الإلكترونى:
3062–2562
DOI:
10.21608/MFTH.2025.452620

الملخص

يشمل التراث الثقافي التقاليد والآثار والقطع الأثرية وأنماط الحياة المتوارثة من الأجيال السابقة، والتي تُشكل جزءًا لا يتجزأ من هوية المجتمع. و قد فتحت تحليلات البيانات الضخمة في مجال التراث الثقافي في السنوات الأخيرة آفاقًا جديدة للبحث والمحافظة علي الاثار. ويمكن الاستفادة بفاعلية من البيانات الضخمة، التي تتكون من مستودعات ضخمة من المعلومات الُمولّدة من منصات التواصل الاجتماعي، لاستخلاص رؤى قيّمة للمؤسسات والوجهات السياحية، مما يزيد من رضا العملاء ويعزز جودة الخدمات الُمقّمة.

تهدف هذه الدراسة إلى تقييم تجربة التراث الثقافي في أحد أشهر مواقع التراث العالمي، و هي أهرامات الجيزة، من خلال استخدام البيانات الضخمة لتحليل المحتوى الذي يُنشئه المستخدمون علي منصة TripAdvisor.

و للإجابة على أسئلة البحث، تم استخدام برنامج NVivo14 لتحليل اراء العملاء للفترة من 2018 إلى 2024. وقد كشف التحليل باستخدام Nvivo14 عن هيمنة المشاعر الإيجابية، كما اتضح من تكرارات كلمات مثل "ممتاز" و"مذهل" و"رائع". واتاح تحليل تكرار الكلمات فهمًا للمواضيع الرئيسية التي ناقشها الزوار. سلّط استعلام تكرار الكلمات الضوء على كلمات متكررة مثل "هرم"، و"دليل"، و"جمل"، و"رائع"، و"تجربة"، و"مذهل" و كذا كشف التحليل الموضوعي للبيانات عن مواضيع رئيسية برزت في تقييمات TripAdvisor، مثل "تجربة الزائر"، و"التقدير الجمالي والثقافي"، و"البناء والفخامة". ومع ذلك، أبدت بعض التقييمات مخاوف، لا سيما فيما يتعلق بـ"الاكتظاظ" و"اصرار البائعين ". ويشير البحث إلى ضرورة اتخاذ الحكومة إجراءات صارمة لتعزيز رضا الزوار، و قد تشمل هذه التدابير تنظيم حركة المرور في الموقع التراث الاثري، وإدارة البائعين بكفاءة أكبر، والإشراف الدقيق على رحلات الجمال والخيول.إضافةً إلى ذلك، تُسلّط الدراسة الضوء على أهمية استخدام البيانات الضخمة لتحسين تجربة العملاء في مواقع التراث الثقافي.

الكلمات الدالة

البيانات الضخمة، تحليل المحتوى، المحتوى الذي ينشئه المستخدمون (UGC)، أهرامات الجيزة، المواقع التراثية، NVivo14، TripAdvisor

1. Introduction

The definition of "cultural heritage", according to UNESCO (2024), encompasses both tangible and intangible aspects. Tangible cultural heritage embraces movable, immovable, and underwater artifacts, while intangible cultural heritage comprises oral traditions, performing arts, and rituals. Cultural heritage embodies the material and social accomplishments of previous eras and countries, preserved for the benefit of contemporary and future generations. It may encompass historical monuments, groups of buildings, or sites of historical or aesthetic significance of universal value.

The term cultural heritage tourism, as stated by Burns et al. (2010), goes beyond mere visits to ancient sites and entails a profound engagement with the historical, traditional, and cultural facets of a destination. Reflecting upon the idea that every community has a unique and compelling story to share, cultural heritage tourism strives to facilitate an authentic encounter with local customs, traditions, art, history, ruins, relics, and culture that truly depict a specific location. It is based on the premise that the cultural experience for tourists must be genuinely distinctive and authentic.

The study by McKercher & Du Cros (2003) defined cultural tourists are individuals who travel to places outside their home country to engage specifically in activities related to history, arts, lifestyle, and other aspects that define a destination's culture. The study also identified five typologies of cultural tourists: serendipitous, purposeful, incidental, casual, and sightseeing cultural tourists. This categorization is based on the depth of experience pursued from visiting a destination and the importance of culture in the decision to visit that destination.

Nowadays, people are surrounded by all types of technological tools used in communicating, sharing, and buying. Customers' footprints can be easily traced by devices such as sensors, GPSs, cameras, mobiles, and computers. Huge amounts of data about customers, suppliers, and competitors are captured, which can be analyzed by businesses to create brand personas. This unstructured heterogeneous data that is generated in real-time and accumulated rapidly is referred to as "Big Data" (De Mauro et al., 2015).

The tourism industry is a service sector characterized by a close relationship between travelers and information. Travelers are constantly searching for multiple information sources to help them in their decision-making processes when planning trips. In this digital age, online data—such as reviews, posts, likes, and shares- plays an important role in determining their choices. This online sharing of experiences is often referred to as electronic word of mouth (eWOM). E-WOM has transformed how travelers perceive information, evolving from traditional word of mouth, which relied on personal recommendations and conversations among friends and family before the advent of the Internet. Nowadays, travelers can access abundant information from varied sources, allowing them to compare options, assess service quality, and ultimately make more informed booking decisions Camacho et al., 2020; Chittiprolu et al., 2021)

The integration of cultural heritage and big data offers vital possibilities for research, conservation, and community involvement. By leveraging big data analytics, cultural heritage organizations can deepen their insight into cultural practices by visitors in destinations, enrich visitors' experiences, and protect heritage for generations to come.

Nevertheless, it is essential to tackle the related challenges to fully capitalize on the benefits of this integration.

This empirical study investigates user satisfaction by analyzing User-Generated Content (UCG) created by travelers who visited Egyptian heritage sites. The goal is to improve customer satisfaction and provide valuable insights for potential enhancements, ultimately enriching the way visitors engage with and appreciate Egypt's rich cultural heritage. The research addresses the following questions:

- *What are the overall visitor sentiments regarding their experiences at the Giza Pyramids?*
- *What specific aspects of the visit (e.g., guided tours, accessibility, facilities) contribute most to positive or negative sentiments?*
- *How do visitors perceive the cultural significance of the Giza Pyramids, and how does this perception affect their overall sentiment?*
- *What suggestions do visitors have for improving their experience at the Giza Pyramids?*

2. LITERATURE REVIEW

2.1 Cultural Heritage and Big Data

Today, Big Data can be used in the tourism sector to uncover important information about key players in the tourism industry. As an industry focusing on human beings, analyzing data footprints before, during, and after the tourist experience can enhance the services and develop business approaches to meet tourist desires and needs (Li et al., 2018).

Using Big Data in tourism is still in its infancy; despite that, using Big Data can support the tourism industry towards high-quality products and higher levels of customization. Some researchers tackled this important field of research; still, practitioners need more in-depth implementation, a better understanding of the benefits of Big Data in analyzing patterns, and a smooth inclusion into business operations (Centobelli & Ndou, 2019; Mariani, 2020).

The capabilities of Big Data have evolved. It started with analyzing structured data that could be presented in tables or interactive databases. As data generated has been developing into unstructured data like photos, audios, and videos, big data analytics has developed accordingly (Li et al., 2020).

Recommender systems are becoming an indispensable tool in the tourism industry to guide travelers in their search and trip planning. These systems offer recommendations to travelers for products and services based on prior purchases or products that have been bought by other customers with similar preferences. The integration of big data into tourism recommender systems enables the analysis of large-scale traveler behaviors, preferences, and feedback, thereby enhancing the accuracy and personalization of tourism recommendations (Solano-Barliza et al., 2024)

Big data technologies for deep learning, like classification, association rules, collaborative filtering, and neural networks, can be used to analyze tourist information gathered from different sources. During the three phases, before/during/after the trip,

machine learning technologies analyze generated data in real time to better serve travelers' needs and provide a decision support system for better choices (Lu, 2022).
Big data sources in tourism can be categorized into three main types: 1. Visitor-generated data, which includes text, videos, and audio content. 2. Device data, such as WIFI, GPS, Bluetooth, and roaming information. 3. Transaction data, which encompasses web searches, online bookings, payment records, and browsing activity (Salas-Olmedo et al., 2018). The data derived through data analytics techniques can vary from simple spreadsheets for less tech-savvy users to complicated relational databases (Wallace et al., 2023).
Significant data repositories for big data mining in cultural heritage include social media platforms, online gaming, transaction processing systems, customer databases, emails, internet clickstream logs, and mobile applications (Amato et al., 2017).
The main areas of technology addressing cultural heritage informatics to boost customer experience are digitalization, big data, semantic analysis, crowdsourcing, VR, and IoT (Internet of Things). Cultural heritage attractions have taken major leaps in the area of technology and innovation (Poulopoulos & Wallace, 2022).
The illusion projected by virtual, augmented, and mixed or extended reality that makes the physical or online visitor feel as if he is physically at the site is considered one of the technological milestones that transmitted cultural visitors from the real to the virtual world. In this context, technology serves either as a mediator of the experience or becomes an integral part of the experience itself (Innocente et al., 2023).
Cultural informatics has experienced a huge transformation in the last decades. It has shifted towards a more people-centric approach, away from an approach that basically focused on objects. The cultural heritage sites developed from purely serving as information providers to becoming places that prioritize personalization, enhance user experiences, and fully immerse visitors in the cultural spirit through data analysis and extraction techniques (Kontogiannis, et al., 2020; Kosmopoulos & Styliaras, 2018).
A vibrant new wave of applications dedicated to museums and cultural heritage sites is emerging, embracing customization and storytelling in the visitor experience, creating a seamless journey for users. Furthermore, the research by Vassilakis et al. (2017) identifies important links between user profiles and their museum experiences. Today, many innovative applications have made a significant mark on cultural heritage, tailoring each user's journey to ensure a personalized experience.

2.2 Cultural Heritage and Sentiment Analysis

Sentiment analysis in tourism using text mining and language processing software is utilized to examine and identify emotions and attitudes conveyed by tourists towards various aspects of their experience. This is captured through feedback data assessing tourism service providers such as tourism heritage attractions, hotels, and restaurants. This feedback can be reflected in likes, comments, and shares in response to a new product, destination, or service quality. By analyzing this feedback, the sentiment analysis helps to uncover valuable insights into tourists' perceptions and satisfaction levels, ultimately allowing businesses and stakeholders to boost the overall travel experience through a more cognitive and data-driven decision-making process. Managers and marketers can make constant adjustments based on visitor feedback to attract more visitors. Therefore, Sentiment Analysis has drawn the attention of

researchers and practitioners in the tourism industry (Escandon-Barbosa & Salas-Paramo, 2021).

It is worth mentioning that by delving into content created by cultural heritage visitors and shared through different social media platforms like Twitter, TripAdvisor, and Facebook, cultural heritage officials can gain perceptions of the dynamics of cultural heritage tourism and tailor experiences for visitors. New developments in cultural informatics paved the way for a better understanding of visitors' behaviors in cultural heritage sites. It also facilitated the way we can maneuver with a large volume of data sets gathered from different social media platforms, digitized cultural content, and Internet of Things (IoT) appliances, which are stored in data lakes. The primary goal of analyzing this huge amount of extracted data is to provide visitors visiting cultural heritage sites with personalized experiences (Wallace et al., 2023).

In the study by Deligiannis et.al (2020), an online data lake was introduced to store, share, and analyze resourceful data sets. This framework enables people responsible for museums and other cultural heritage sites to use data acquisition services like web crawlers and social media scrapers to get the needed data, store it efficiently, share it with stakeholders, filter, and analyze data with user-friendly graphical tools, and additionally control access to stored data by data management tools.

According to Spiliotopoulos et.al. (2020), data from a music event on Twitter were extracted and analyzed to develop user personas. The findings showed that using a data analytics technique for topic-based modeling can create clear and consistent personas. Nevertheless, the research also highlighted the importance of human intervention to fine-tune results

As highlighted in the work conducted by Konstantakis et al. (2020), a recommender system was proposed that creates propositions for cultural sites for visitors based on an analysis of cultural visitor typologies and user preferences. The research revealed that the suggestions of recommender systems can be enhanced by incorporating additional information about visitors' cultural backgrounds.

The study by Amato et al. (2017) presents a project called SCRABS that utilizes Big Data algorithms to provide more personalized experiences for visitors at cultural heritage sites and museums. It integrates data such as user preferences, geographical locations, and cultural content to deliver customized recommendations in real-time. The application is capable of accomplishing tasks such as gathering, filtering, managing, and analyzing data. The application was tested on a group of users and demonstrated its effectiveness in providing recommendations for cultural heritage sites of interest.

As indicated in the study by Drivas et. al. (2020), a framework was developed to enhance organic traffic optimization. The research studied 171 cultural websites in terms of user experience and navigational behaviors. It aimed to create fine-tuned Big Data analytics strategies to improve visibility and provide user-friendly navigation options to cultural heritage website visitors, which can be measured through click-throughs, user engagement, and conversion rates.

According to Antoniou (2017), in order to provide accurate recommendations, both explicit and implicit information is gathered from the user. Explicit information includes direct data obtained through questionnaires, ratings, or comments, as well as

indirect data gathered through gamification. Implicit information may include tracking browsing activity, checking into locations, and monitoring social network interaction. Combining these types of information can assist in conveying more precise suggestions.

In the study by Julia & Junshean (2015), tweets were analyzed to capture visitors' attitudes towards Lake Toba in Indonesia. Tourists' opinions about variables such as hotel service, accommodation, transportation, and other areas were examined. The analysis showed that most of the data fell into the neutral category with 409 neutral, 304 positive, and 37 negative tweets.

The study by Khomsah, et al. (2021) examined tourists' sentiments towards eight tourism destinations in Purwokerto, Central Java, Indonesia. The researchers used lexicons for automatically classifying the sentiments of the travelers. Word cloud analysis was also used to extract keywords that reflect the travelers' experiences of tourist destinations. The results showed that visitors had a positive sentiment towards five destinations with scores ranging from 77% to 91%.

Atabay & Çizel (2020) examined tourists' experiences with hotels offering all-inclusive service in three popular tourist destinations: Antalya, Majorca, and Sharm El Sheikh. Customer-generated content from top hotels in these destinations from an online review site was analyzed by the R program. The analysis revealed diverse emotions about service components in hotels in the related destinations.

In the research conducted by Widyawati et al. (2021), big data datasets extracted from TripAdvisor were utilized to examine visitors' emotions and perceptions regarding service quality in Borobudur. Techniques such as sentiment analysis, classifications, and text network analysis were employed to reveal patterns of visitors' emotions towards TOURQUAL components related to the destination. The findings of the research indicate that Borobudur showed a significantly higher positive rating for service quality, with 83% of visitors expressing positive feelings compared to only 17% who reported negative experiences.

As of Q2, 2024, TripAdvisor reported having revenue of $497 million, reflecting year-over-year growth of 1% with net income of $24 million (TripAdvisor, 2024). TripAdvisor is one of the largest sources of user-generated content, with over 1.5 billion reviews and opinions available on the site. As of 2023, approximately 400 million new reviews were posted on TripAdvisor annually, a key feature of the site that drives much of its engagement (Tripadvisor, 2023). In 2024, TripAdvisor ranks second among the most popular travel and tourism websites worldwide based on the share of visits (Statista, 2024a). In September 2024, the TripAdvisor platform received over 130 million visits, showing a significant user engagement (Statista, 2024b).

In the research by Jangid et al. (2021), the researchers went beyond simply categorizing users according to their sentiments. They identified four distinct groups: loyal fans, unsatisfied users, silent followers, and cheerleaders who can serve as catalysts for brand identity. Loyal fans, for instance, can spread positive WOM and are highly appreciated by businesses and rewarded. Unhappy users can, on the other hand, have a negative impact on the company's reputation and should be addressed immediately and compensated. The research highlighted that companies should make an effort to engage Quiet followers, transforming them into advocates for the brand.

Finally, the research identified cheerleaders who are loyal supporters of the brand and should be kept inspired and informed to continue spreading positive WOM. The research used "Word Cloud" as a data visualization tool to extract data that shows users' attitudes and their problems, thereby providing frameworks to address them effectively.

The research by Lama (2024) showcased how marketers can use big data extracted from social media platforms in implementing sentiment analysis to obtain valuable insights. These insights can assist them in designing marketing campaigns that align with user behaviors and attitudes. The study concludes that marketing strategies should be tailored to meet the diverse needs and preferences of different customers. Furthermore, it emphasizes the importance of assessing social media platforms in terms of their performance and effectiveness in maintaining customers.

In the study by Farisi et al. (2019), the researchers used the Multinomial Naïve Bayes Classifier to classify the positive and negative opinions of hotel visitors on social media platforms. The first scenario employed a frequency-based model, which eliminated words with the lowest frequency, while the second scenario eliminated the features that showed a minimum variance of positive and negative probability values. The results from Scenario 1 were more accurate than those from Scenario 2.

In the study by Amanatidis et al. (2020), the relationship between cultural heritage sites, such as museums, and social media platforms was explored. The findings showed that only one museum in Greece maintains an Instagram account. This account features content generated by visitors who tag the museum in their personal photographs featuring the exhibits. The study highlights the untapped potential of social media and the gains of maintaining active social media platforms. In addition to increasing the visibility of the heritage sites, user-generated content in the form of text, photos, and videos creates big data repositories that could be utilized in conducting useful analyses and conveying useful insights. The research also indicated that cultural heritage sites should engage more in social media platforms and adopt a more committed and proactive approach to leverage social media potential and consequently tailor experiences based on visitors' preferences.

Sentiment analysis was utilized by Orea-Giner et al. (2022) to identify tourists' sentiments based on their interaction with service robots. Additionally, this technique was applied to evaluate visitor experiences at Cambodian heritage sites (Baniya et al., 2021) and to examine the sentiments of Chinese tourists regarding Australian destinations in Liu et al. (2019).

Wang & Moriarty (2018) outline challenges to big data implementation, including privacy, data security, reliability, and technical challenges, especially in sensitive applications like those involving personal or cultural data.

In their study, Teles da Mota & Pickering (2020) evaluated nature-based tourism from Flickr reviews. Additionally, Lu & Zheng (2021) focused on guest sentiments regarding cruise tourism during the COVID-19 pandemic.

As UCG websites have gained popularity for exchanging experiences, researchers have utilized the data obtained from these platforms to study several topics. One notable study by Park et al. (2020) investigated customer feedback on an online hotel reservation platform.

In the museum sector, Surugiu et al. (2022) utilized Facebook comments to evaluate visitor opinions on Romanian museum activities. The quality of Airport service levels was also examined by Gitto & Mancuso (2017), who relied on data from the Skytrax blog. Zhang et al. (2020) and Zhang et al. (2022) explored the effect of air pollution on tourist experiences in Beijing based on Weibo posts, as well as customer opinions on virtual tourism during COVID-19 from Weibo posts, respectively. Gharzouli et al. (2022) analyzed the quality of hotel services in India based on guest reviews, while Zhu et al. (2021) evaluated tourist satisfaction with Airbnb accommodations in Los Angeles using reviews from insideairbnb.com. The destination image of Hong Kong has been explored through tourists' comments provided on the Baidu AI platform in a study by Jiang et al. (2021).

The next section will address research methodology and answer research questions. This research will apply qualitative content analysis relying on NVivo 14 software as a research approach. Content analysis is a structured research method utilized to investigate written, spoken, or visual data sources. It involves categorizing the extracted data to uncover patterns, themes, and underlying meanings.

NVivo 14 is qualitative analysis software that is widely used in several research areas. It works with unstructured qualitative data to analyze user sentiments and emotions, whether they are positive, negative, or neutral regarding products, services, or personal experiences. The software focuses on trends in word occurrences and identifies themes within the text. The software can extract data from different sources, including social media platforms, text, surveys, interviews, images, spreadsheets, and videos (Mondal & Samaddar, 2022).

3. RESEARCH METHODOLOGY

TripAdvisor is considered one of the leading platforms in the travel and tourism sector and is considered one of the most widely used tools by travelers, providing valuable insights about a vast range of tourism services like hotels, attractions, restaurants, and activities. Travelers shape their travel decisions based on millions of reviews, ratings, photos, and experiences shared by users from all over the globe. TripAdvisor provides information about destinations and service providers; thus, TripAdvisor is considered a guide for pre-during and post-trip stages. In addition to assisting travelers in making informed travel decisions, this huge repository of information helps businesses get instant feedback about their services and products. This allows them to address negative feedback, appreciate positive comments, and enhance various components of their offerings.

The rating algorithm on TripAdvisor is based on the average ratings of the online reviews. Therefore, businesses must enhance their service quality to get high review ratings, which directly affects the decision-making process of travelers seeking advice from former travelers (Višković, 2022).

TripAdvisor is considered a powerful tool of eWOM, which was defined by Litvin et al. (2008) as any informal communication among consumers via digital platforms addressing the quality, characteristics, or related experiences of goods and services or their providers.

3.1 Data Collection and Preparation

The theoretical framework shed light on the tools used to leverage big data sources to examine user satisfaction levels of products and services on the one hand, and in cultural heritage, specifically on the other hand, as the point of interest of this research. In order to examine the research questions, a qualitative content analysis of visitors' reviews was conducted. TripAdvisor, a highly ranked Consumer-Generated Content (CGC) platform, was used to explore consumer sentiments toward the Giza Pyramids heritage site in Egypt. Content analysis is a method that focuses on specific attributes to systematically recognize and categorize key subjects while assessing their frequency and associated topics. For research purposes, consumer reviews were analyzed over the period from August 2017 to November 2024. A total of 2500 reviews, both in English and non-English languages, were examined. Non-English reviews were translated to facilitate a comprehensive analysis of both negative and positive sentiments. NVivo14, a powerful sentiment analysis software, was used as it provides various options to explore sentiments in addition to numerous visualization tools for the output data, enabling the identification of patterns and relationships in the datasets. The target group for this research was travelers who visited the Giza Pyramids, one of the world's seven wonders, sharing their experiences with other people considering a trip to Egypt. The data preparation involved several stages. Firstly, the reviews were scraped using a TripAdvisor scraper tool into an Excel file. Secondly, the non-English reviews were translated, and the file was imported into NVivo14. The program provides visual results and enables the coding of written, visual, and audio data. Finally, the reviews were coded into parent and child nodes featuring main and sub-themes.

The software allows for the integration of data from different formats into nodes. A node refers to a word that serves as a repository where other words can be organized. All these words relate to a specific subject or research theme. Various materials conveying the same concept can be consolidated into a node. The act of organizing into a node is known as coding. The software enables applying codes to data segments for thematic analysis. It also allows for the retrieval of reference codes, which can then be used to create models and graphs. A variety of different queries can be conducted, such as identifying the frequency of various words across different groups to explore data relationships and patterns. In addition, NVivo 14 combines qualitative and quantitative data for comprehensive analysis (Ríos-Martín et al., 2020).

3.2 Validity of the Data

In order to test the reliability of the content analysis, the coding process was performed by two different coders. Additionally, Cohen's Kappa coefficient (K) was calculated, resulting in a value of 0.80, which indicates good consistency in identifying themes and critical factors. Cohen's Kappa coefficient is a statistical measure used to assess the agreement between two different analyses (Ríos-Martín et al., 2020).

3.3 Findings and Discussion:

The analysis utilized various NVivo14 features, including word frequency analysis, sentiment classification, coding references, thematic exploration, word clouds, and node relationships. Through this comprehensive approach, the study aims to uncover

key trends, patterns, and visitor sentiments regarding the Giza heritage site. The following is a descriptive analysis of the findings.

1. Sentiment Analysis: Distribution of Positive, Neutral, and Negative Reviews

The data was gathered from 2,500 reviews on the TripAdvisor website between August 2017 and November 2024. TripAdvisor's users rated their overall experience on a scale from 1 (terrible) to 5 (excellent). For research purposes, reviews rated as 'excellent and very good' were considered 'positive,' reviews rated as 'average' were considered neutral, and reviews rated as 'poor and terrible' were considered 'negative' (Yetgin et al. 2020). Table 1 shows that 2,189 travelers (87.7%) rated their experience positively, 141 travelers (5.6) rated it as neutral, while 170 travelers (6.7%) rated it negatively.

The majority of reviews (87.7% positive reviews) expressed appreciation for the site's architectural beauty and historical significance. Words such as "stunning," "amazing," "dream," and "breathtaking" frequently appeared in positive reviews. A very small portion of comments was classified as neutral (5.6%), mainly consisting of factual descriptions or general statements regarding the heritage site. These reviews neither praised nor criticized the site but instead provided straightforward observations. A minority of reviews (6.7%) conveyed dissatisfaction, focusing primarily on overcrowding, organization, weather, and merchants pushing to sell their products. Words such as "aggressive," "crowded," and "hot" commonly appeared in this category. Figure 1 and Table 2 show the distribution of languages used in TripAdvisor reviews regarding travelers' visits to the Pyramids of Giza. The majority of reviews were written in English, followed by Spanish and Portuguese.

Table 1: Guest Experience Rating in the Pyramids of Giza on TripAdvisor

	Rating	Number of Reviewers	%
Excellent	5	1850	74
Very good	4	339	13.7
Average	3	141	5.6
Poor	2	48	1.9
Terrible	1	122	4.8
Total		2500	100

Source: NVivo14

Table 2: Number of reviews by language

English	1545	French	73
Spanish	228	Arabic	46
Portuguese	129	Japanese	49
German	82	other	263
Russian	85	Total	2500

Source: NVivo14

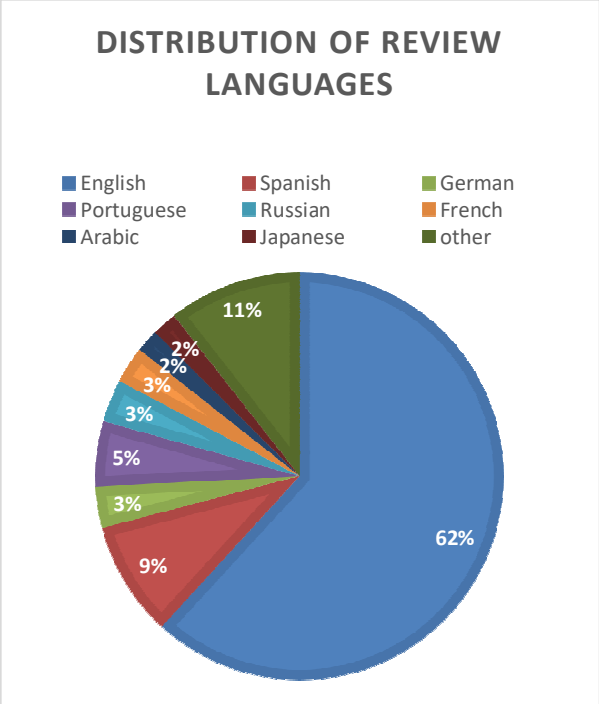

Figure 1: Distribution of reviews by languages.
Source: NVivo14

The analysis revealed that positive comments had 1015 coding references, negative comments had 241, neutral reviews totaled 335, and mixed reviews accounted for 909 (Table 3). The findings highlighted that positive feedback dominates the dataset, with over 1,000 references, indicating an overall favorable perception. Negative feedback is relatively low, appearing much less frequently than positive or mixed reviews. Mixed reviews are substantial (909 references), showing that many responses contained both strengths and weaknesses rather than being purely positive or negative. Neutral reviews are the least frequent (335 references), meaning that fewer comments were entirely neutral or objective. Figure 2 illustrates a *Treemap* that displays data as nested rectangles, where the size of each rectangle represents the relative frequency or proportion of a category.

Table 3: Number of coding references

Codes	Number of coding References
reviews - Neutral	335
reviews - Positive	1015
reviews - Mixed	909
reviews - Negative	241
reviews- Total	2500

Source: NVivo14

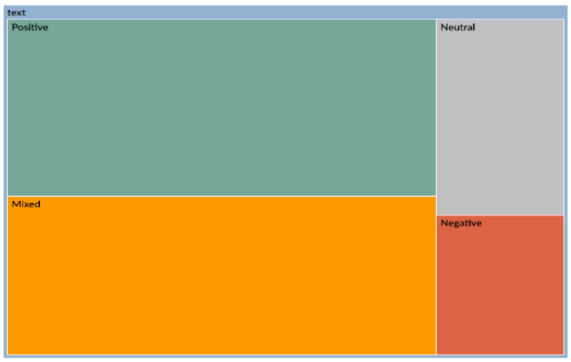

Figure 2: Treemap of frequencies
Source: NVivo14

The analysis revealed several key themes from the reviews, highlighting different aspects of the experiences shared by visitors. Below is a breakdown of the most important themes (Figure 3):

1. **Camel Rides** – Many reviewers mentioned camel rides as a significant part of their experience. This included discussions on the quality of the ride, pricing, the treatment of camels, the expertise of the handlers, and whether the ride enhanced their overall visit. Some visitors may have found camel rides thrilling and authentic, while others might have raised ethical concerns regarding animal welfare.
2. **Tour Guides** – The role of tour guides was another prominent theme in the reviews. This includes aspects such as the knowledge and professionalism of guides, their storytelling abilities, language proficiency, and how well they enhanced the visitor's understanding of the site. Positive reviews likely praised guides for their enthusiasm and expertise, while negative ones might have focused on disorganization, lack of engagement, or high costs.
3. **Tickets** – Discussions related to ticketing included pricing, availability, ease of booking, long queues, or value for money. Some visitors have commented on the administration of the ticketing process and other issues, including hidden fees, delays, and limited accessibility options.

4. **Site** – This theme covers reviews related to the overall management of the heritage site, including its historical significance, cleanliness, infrastructure, crowd management, and preservation efforts. Some visitors have commented on the site's maintenance, accessibility, signage, or whether it met their expectations based on prior research.
5. **Photography** – Many reviews likely mentioned the opportunity for photography. They highlighted the scenic beauty of the site, restrictions on photography, and the presence of professional photographers offering services. Some reviewers expressed concerns about whether the photography policies were too strict or the site was too crowded to capture good shots.
6. **Experience** – This broad theme captures the overall sentiment of the visit, including emotions, expectations versus reality, and whether the trip was memorable or disappointing. Reviews have focused on whether the site lived up to its reputation, whether it was worth the money, or whether it provided a once-in-a-lifetime adventure.
7. **People** – The presence and behavior of other visitors, staff, or locals also emerged as a recurring theme. This includes interactions with friendly or unhelpful staff, the impact of crowds on the experience, cultural exchanges, or potential issues such as aggressive vendors or pushy sales tactics.

Each of these themes provided valuable insights into what visitors appreciate or criticize the most, helping stakeholders improve their experience based on visitors' feedback.

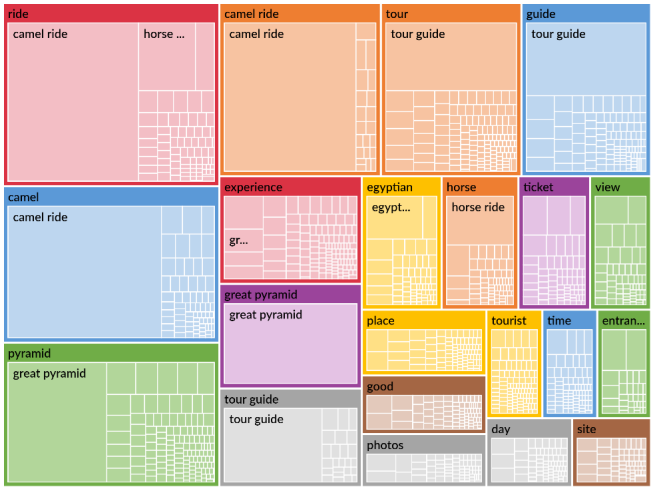

Figure 3: Important Themes from reviews
Source: NVivo14

Table 4: A selection of positive and negative reviews from the travelers

Example of References for Negative Sentiments:	Example of References for Positive sentiments
• Depending on how you travel to Giza, the trip can be pretty exhausting. • Note you can enter the main pyramid if	• A couple of friends and I took Adventures tour of Egypt and we all LOVED it!

you are not fearing enclosed spaces • The traffic is horrendous with everyone trying to leave at the same time and a horse cart was hit by our tour bus. • The only thing sometimes camel riders push you to take photos it's annoying but is ok it's not that expensive • We also visited the Egyptian Museum on our own, which was much more affordable • The only downside to my experience was the constant and overly aggressive approaches from people trying to sell souvenirs on the street. • However, if we visit again, we'd plan things more carefully to avoid overspending. • Yet sadly transportation to see around the pyramids was only by camel…	• A lifelong dream come true and so worth it. • They are awe inspiring and took my breath away. • It our Frist time to visit the pyramids we surprised how is the huge is the Khofo pyramid , I recommend everyone to visit it • The Great Pyramids of Egypt are truly amazing! • I was struck by their size and the precision with which they were built. • It's hard not to be impressed by these incredible structures, and visiting them left me with a sense of wonder and admiration for the culture that created them. • What a brilliant place so informative and eye-opening. • The views were absolutely breathtaking. • We captured countless photos of these ancient wonders, each more stunning than the last. • Exploring the rich history and monumental architecture of the pyramids was an incredible journey. • Today was a perfect experience of our first time in Egypt. • The Great Pyramid, in particular, is a marvel, its perfect alignment and towering presence a testament to the genius of ancient Egyptian architects. • Exploring the Pyramids was an unforgettable experience! • Walking around these ancient wonders was surreal, with stunning views and vibrant local culture.

Source: NVivo14

These references highlight different sentiments regarding the experience at the site.
The positive reviews highlighted some insights as follows: (Table 4)

1. **Overwhelming Monuments** – Many visitors described the pyramids as "breathtaking," "awe-inspiring," and "incredible." Their enormous size, history, and architectural precision left a lasting impression.

2. **Dream Experience** – For many, visiting the pyramids was a lifelong dream, and the reality exceeded their expectations, making it a highlight of their travels.
3. **Great Photography Opportunities** – Visitors enjoyed capturing stunning photos of the pyramids and the surrounding landscapes, making it a visually rewarding experience.
4. **Informative and Well-Organized Tours** – Some guests had positive experiences with professional and knowledgeable tour guides who made the history engaging and easy to understand.
5. **Warm Local Interactions** – Visitors appreciated friendly and trustworthy tour operators, with some personally thanking guides for making their trip memorable.
6. **Dreamlike and Cultural Immersion** – Walking around the pyramids, engaging with local culture, and witnessing the historical magnificence firsthand was described as a fantastic and unforgettable journey.

Negative reviews highlighted some shortcomings like:

1. **Travel Difficulties** – Reviews mention the journey to Giza being "exhausting" and due to heavy traffic, making the trip unpleasant.
2. **Animal Welfare Concerns** – Many visitors were disturbed by the treatment of animals, including overworked horses and camels.
3. **Persistent Vendors** – Complaints about demanding merchants, along with pushy camel riders pressuring visitors for photos, made the visit stressful.
4. **Insufficient Tour Experiences** – Some visitors felt their tours lacked detailed explanations**.**
5. **Transportation Issues** – The lack of modern transportation options was a concern, as visitors had to rely on camels and horses.

Figure 4 demonstrates a word cloud generated using NVivo 14, displaying the most frequently used words in the dataset. The size of each word is proportionate to its frequency of use. Stop words such as 'of,' 'at,' 'for,' and 'but' have been excluded from the word cloud, as they aren't relevant to the research. The most prominent words appearing in the visualization included "pyramids," "guide," "experience," "camels," "great," "Giza," and "people." The word cloud provides an intuitive representation of key discussion points, making it easier to identify prevalent themes.

Figure 4: Word cloud for word frequencies
Source: NVivo14

2. Word Frequency Analysis

A word frequency query (Figure 5) was conducted to determine the most commonly used words in the dataset, providing an understanding of the primary topics discussed by visitors. NVivo 14's word frequency query identified recurring words such as "pyramid(s)," "guide," "camel(s)," "great," "experience," "amazing," and "ride(s)." These words reflect major visitors' interests. Specifically, the frequent occurrence of words like "ancient", "impressive," and "wonders" emphasizes the site's perceived cultural value. According to the data obtained, the most frequently used words in the comments were pyramid(s) with 8266 occurrences, camel(s) (2243), guide (1752), visit (1746), great (1668), experience (s) (1518), respectively.

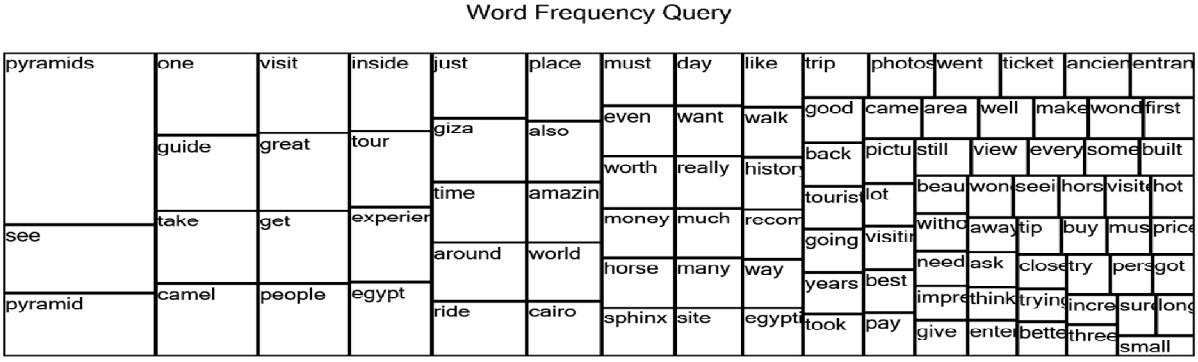

Figure 5: Word frequencies query
Source: NVivo14

3. Coding References, Categorization of Data, and Thematic Analysis

The dataset was systematically coded using NVivo 14's node structure to identify parent and child nodes. Child nodes are essentially sub-codes nested under parent nodes. Coding references were assigned to various themes related to multiple aspects of the heritage site. The primary coding nodes are demonstrated in Table 5. For example, parent node "site construction" was associated with child nodes like "pyramids", "Giza", "Sphinx", "Structure(s)", "Stone(s)", "Khufu", "Tomb(s)", among others.

Table 5: Nodes, Related Words, and Number of References

Node	*Related Words and References with the number of occurrences*
Site Constructions	Pyramid(s) (8266), Giza (1480), Sphinx (861), Structure(s) (350), Stone(s) (269), Khufu (250), Tomb(s) (246), Cheops (244), Chamber(s) (235), Monument(s) (226), Plateau (188), Pharoa(s) (188), Khafre (183), King (153), Attraction (137), Sites (143), Construction (142), Temple(s) (108), Menkaure (108), Blocks (102), Queen(s) (76), Sarcophagus (60), Saqqara (59), Memphis (47), Chephren (42), Necropolis (42).
Visitor Experience	Visit(s) (1746), Tour(s) (1710), Experience (1518), Ride(s) (1473), Time (1318), Trip(s) (727), Expectation(s) (330), Dream (245), Travel (239), Excursion (119), Opportunity (107), Journey (102), Chance (71), Mystery (59), Adventure (45), Memories (42).
Location	Place(s) (1430), Area (536), Desert (238), Space(s) (138), Distance (130), Surrounding(s) (116), Destination (44).

Activities at the Site	Camel(s) (2243), Guide (1752), Ride(s) (1473), Horse(s) (1250), Photo(s) (1200), Walk (687), Pictures (495), Carriage(s) (435), Animal(s) (428), Camera (120), Video (52).
Charges	Ticket(s) (943), Money (846), Tip(s) (543), Pay (482), Buy (379), Price (378), Cost(s) (313), Pounds (250), Fee (223), EGP (184), Dollars (96), Euros (75), Bargain (64), Payment (44).
Cultural Importance of the Site	History (715), Ancient (689), Historical (147), Civilization (92), Culture (74), Architecture (70), Heritage 62, Centuries (46), Historic (42), Archaeological (38).
Souvenir Shopping	Vendors (321), Souvenirs (147), Papyrus (82), Merchants (39).
Social Interaction	People (1478), Local(adj.) (308), Locals (275), Egyptians (210), Guys (128), Visitors (96).
Feelings	Feelings (111), Emotion (62).
Security and Safety	Service(s) (189), Security (160), Safe (112), Service (95), Safety (50).

Source: NVivo14

4. Node Analysis and Relationship Mapping

NVivo 14's node analysis tool was utilized to examine the relationships between different concepts in the dataset. The co-occurrence of words within nodes revealed meaningful connections reflecting both positive and negative visitor experiences.

Positive associations: Words like "great, amazing, and breathtaking" would commonly appear alongside terms like *view*, *history*, *guide, place, value, experience*, and *monument*, emphasizing the awe-inspiring nature of the pyramids. Phrases such as *great experience* or *amazing history* would be recurrent. "Majestic" would frequently appear with *history, aura, architecture, and significance.*

Negative associations: Words like "hot" would likely co-occur with *weather*, *sun*, and *heat*, highlighting complaints about the desert climate. Meanwhile, "vendors" might frequently appear together with *pushy*, *overpriced, annoying*, or *scams*, reflecting frustrations with aggressive sellers at the site. "Crowded" would mainly occur with *tourists, lines, busy, chaotic, early,* while "scam" appears frequently with *ticket, camel ride, tour guide, tipping*.

Cluster analysis illustrated these relationships, revealing distinct groupings of terms that reflect different aspects of visitor experiences.

5. Words for Excellence and Satisfaction

Visitors to the Giza Pyramids frequently express deep satisfaction and excitement about their experience, as reflected in their comments. Many of their descriptions include words that convey strong emotions, such as *great*, *amazing*, *wonder*, *worth*, *like*, and *recommend* (Table 7). These expressions highlight the profound impact the site has on travelers, reinforcing the idea that their visit was not only enjoyable but also memorable.

The positive feedback suggests that many guests find the experience to be *breathtaking*, *wonderful*, and *impressive*. This level of enthusiasm is often reflected in a desire to return, as satisfied visitors tend to become repeat travelers, willing to relive what they describe as an *unforgettable* journey. The grandeur of the Pyramids, with their massive scale and historical significance, contributes to a powerful incentive and emotional experience. Many visitors describe their trip as a *dream come true*, emphasizing the fulfillment of their expectations.

In addition to their personal enjoyment, guests frequently express a strong willingness to share their positive experiences with others. The use of the word *recommended* in their reviews indicates that visitors are eager to spread the word about the Pyramids, encouraging friends and family to experience the marvels for themselves. This electronic Word-of-Mouth effect (eWOM) plays a crucial role in attracting more travelers, as personal recommendations often carry more weight than advertisements or promotional content.

Ultimately, the combination of awe-inspiring visuals, emotional fulfillment, and a deep sense of satisfaction ensures that a visit to the Giza Pyramids is not just a trip but a life-enriching experience, one that travelers are excited to relive and share with others.

Table 7: Expressions of Satisfaction in the comments

Word	Count	Word	Count	Word	Count
great	1668	**magnificent**	229	**cool**	119
amazing	1173	**large**	222	**panoramic**	116
wonder(s)	1050	**fantastic**	217	**happy**	114
worth	872	**lifetime**	217	**inspiring**	114
like	743	**top**	206	**unique**	114
recommend	675	**awe**	204	**magical**	113
good	632	**awesome**	193	**safe**	112
wonders	548	**spectacular**	163	**beauty**	108
well	532	**interesting**	158	**organized**	108
best	505	**friendly**	154	**loved**	106
beautiful	469	**enjoyed**	146	**stunning**	105
impressive	439	**fun**	140	**unforgettable**	105
incredible	377	**important**	134	**majestic**	103
wonderful	353	**excellent**	132	**glad**	95
enjoy	312	**breathtaking**	127	**perfect**	94
nice	272	**wow**	127	**fascinating**	91

Source: NVivo14

6. Expressions of dissatisfaction

Visitors to the Giza Pyramids sometimes reported challenges that affect their overall experience. Many of these descriptions include words that convey difficulties and frustrations, such as *hot, hard, difficult, crowded, bad,* and *expensive*. These expressions highlight common concerns among travelers, particularly regarding environmental conditions, logistical challenges, and interactions at the site.

A significant number of visitors mentioned words like *heat (162 mentions)* and *crowds (345 mentions)*, which reveals that the experience may be physically exhausting and overwhelming. Words like *claustrophobic (211 mentions)* and *far (198 mentions)* suggest that some travelers find certain areas confining or feel that reaching the site requires extensive effort.

Issues related to visitor interactions with other people on the site also emerge, with terms like *scammers (168 mentions)*, *pushy (85 mentions)*, and *aggressive (107 mentions)* appearing in their comments, indicating concerns about unwanted requests

to buy or assertive vendors. Additionally, feelings of *disappointment (172 mentions)* and perceptions of the experience being *bad (183 mentions)* or *poor (152 mentions)* suggest that, for some, the visit does not meet expectations. While the Pyramids remain a globally renowned attraction, these negative sentiments underline the importance of the administration's intervention to address common concerns to enhance the overall experience (Table 8).

Table 8: Expressions of dissatisfaction in the comments

Word	Count	Word	Count	Word	Count
hot	400	**hard**	185	**difficult**	151
crowds	345	**bad**	183	**expensive**	110
avoid	289	**disappointed**	172	**aggressive**	107
claustrophobic	211	**scammers**	168	**traffic**	96
far	198	**heat**	162	**pushy**	85
annoying	186	**poor**	152	**problems**	86

Source: NVivo14

The following graphs (Figure 6) display the trend of positive and negative sentiments over time, spanning the period from 2018 to 2025. The fluctuations vary across the period, indicating changing traveler experiences or external factors affecting sentiment. The graphs show that positive sentiments dominate over time. Negative peak months could be identified to help determine potential causes, such as service issues, weather conditions, or external events. A consistent rise or fall in one sentiment over time may indicate an improvement or decline in service, customer experience, or other related factors. It can also be noticed that sudden spikes in positive sentiments occurred towards the end of 2018 till the end of 2019. During this time, positive sentiment far outweighed negative sentiment. A **decline post-2020 can be noticed due to** the pandemic. In general, positive sentiment appears relatively stable but shows a decline after some peak points.

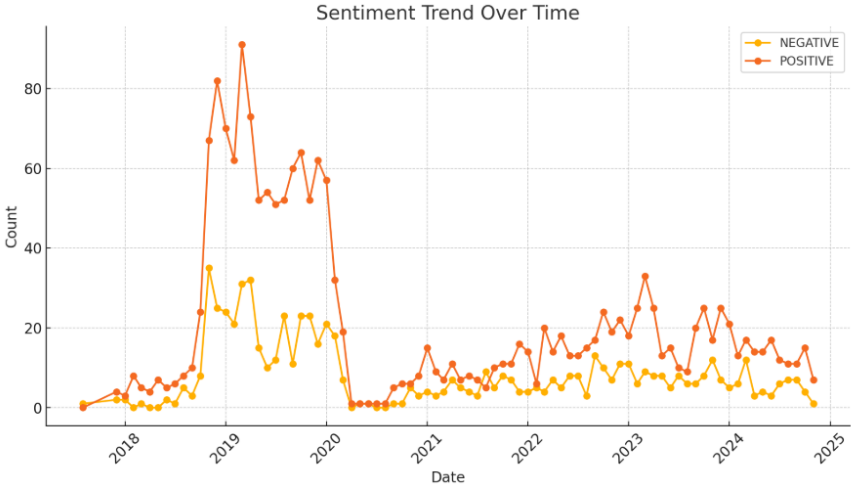

Figure 6: Sentiment trend
Source: NVivo14

7. Pearson's Correlation Coefficient (r)

Pearson's Correlation Coefficient (*r*) is used to examine relationships between variables. The analysis uses the linear correlation coefficient, which ranges between 1 and -1. The Pearson's correlation coefficient (*r*) among variables rating, Triptype, and language indicates minimal linear associations. All correlation coefficients are near zero, with the highest being 0.02068 (between rating and lang). With a sample size of 2,500, statistical testing using the Pearson correlation t-distribution reveals that none of the relationships are statistically significant at the conventional $p < 0.05$ level. Furthermore, the corresponding coefficients of determination (r^2 values) are extremely low, explaining less than 0.05% of the variance in any pair of variables. These results provide strong statistical justification that the examined variables exhibit no meaningful linear relationship and thus are unlikely to contribute substantially to predictive modeling involving one another.

8. Variables Crosstabulation

An analysis of the distribution of reviews in relation to *rating scores* and *trip types* reveals that the majority of ratings fall into the 5-star category, with 1,853 out of a total of 2,500 reviews rating their experience as "excellent." Following this, the 4-star ratings come in significantly lower in number. Each *trip type* predominantly shows 5-star ratings, with the number of 5-star ratings exceeding that of 4-star ratings by a considerable margin. For instance, among couples, 536 out of 716 reviews rated their experience as 5 stars, while only 86 rated it as 4 stars. This data indicates a high level of customer satisfaction among the different *trip types* (Table 9).

Table 9: Trip rating versus trip type Crosstabulation

Count		BUSINESS	COUPLES	FAMILY	FRIENDS	NONE	SOLO	Total
rating	1	4	41	23	14	25	14	121
	2	2	13	8	9	5	10	47
	3	3	40	19	24	39	14	140
	4	10	86	51	60	94	36	339
	5	31	536	315	323	489	162	1853
Total		50	716	416	430	652	236	2500

Source: The researcher

The majority of ratings associated with various languages tend to fall within the range of 4 to 5. This observation suggests that individuals from a diverse range of nationalities, regardless of their cultural backgrounds, generally exhibit positive sentiments regarding their experiences during their visits to the Giza pyramids. This positivity may reflect not only the grandeur and historical significance of the pyramids themselves but also the overall visitor experience, including aspects such as accessibility, hospitality, and the cultural context provided in and around this renowned site. The mean score of *ratings* across *languages* mainly fell between 4 and 5, which indicates a high level of satisfaction with the experience at the Giza Pyramids (Figure 8,9)

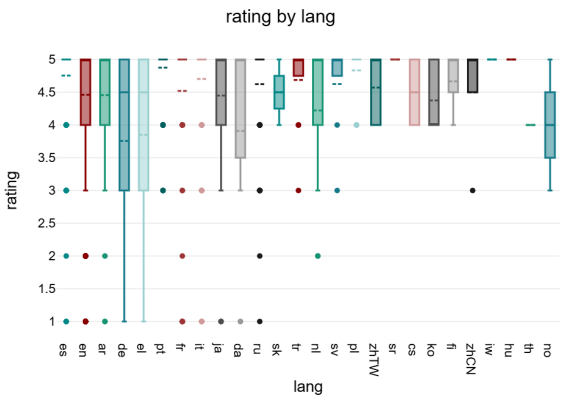

Figure 8: Box Plot of ratings with language
Source: The researcher

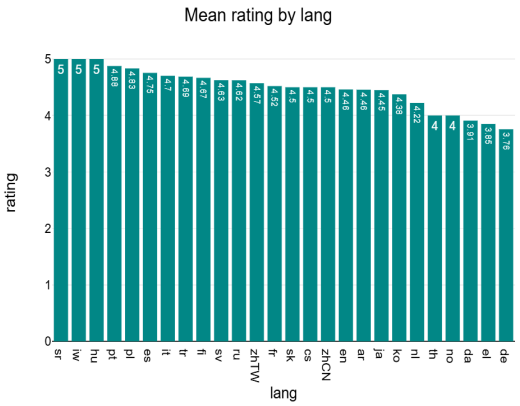

Figure 9: Mean scores of ratings with language
Source: The researcher

4. CONCLUSION AND RECOMMENDATIONS:

This research aimed to investigate the cultural heritage experience at the Giza Pyramids by utilizing big data to analyze user-generated content from TripAdvisor.

This qualitative approach was also used Orea-Giner et al. (2022), Baniya et al. (2021), who utilized this technique to evaluate visitor experiences at Cambodian heritage sites, and Liu et al. (2019), who explored the sentiments of Chinese tourists towards Australian destinations.

The research uncovered visitors' sentiments of one of the World's most famous heritage sites, namely, the Giza pyramids in Cairo, Egypt. Content analysis was conducted by using NVivo14, a powerful software for analyzing customer feedback from social media platforms. Several researchers used qualitative content analysis software in their study to get insights from customer online feedback, like Khomsah et al. (2021) and Widyawati et al. (2021),

Destinations and companies must identify key problems derived from user sentiment analysis and try to handle these problems effectively to gain customer satisfaction. This conforms with the research of Amanatidis et al. (2020) and Wallace et al. (2023).

By using an extensive collection of analyses offered by the software, some very important insights into the site in question were revealed. The findings demonstrated that 87.7% rated their experience positively and were satisfied with their visit to the Pyramids of Giza. The majority of comments expressed appreciation for the site's architectural beauty and historical significance. A minority of reviews (6.7%) expressed some aspects of concern, focusing primarily on overcrowding, organization, weather, and pushy vendors.

Expressions of excellence and dissatisfaction were also revealed. Words like "great" recurred (1668), "amazing" (1173), "wonders" (1050), and "recommend" (675). Expressions of dissatisfaction included words like "hot", "crowds", and "annoying". The dominant sentiment for the experience of the Giza Pyramids was a positive one, highlighting the overwhelming grandeur of the place.

The analysis revealed several key themes from the reviews, highlighting different aspects of the experiences shared by visitors like “Camel Rides”, “Tour Guides”, “Tickets”, “Photography “, and “Experience”.

A word frequency analysis was conducted to determine the most commonly used words in the dataset, providing an understanding of the primary topics discussed by visitors. NVivo 14’s word frequency query identified recurring words such as “pyramid(s),” “guide,” “camel(s),” “great,” “experience,” “amazing,” and “ride(s).” Specifically, the frequent occurrence of words like “ancient”, “impressive,” and “wonders” emphasizes the site's perceived cultural value.

NVivo 14’s node analysis tool was utilized to examine the relationships between different concepts in the dataset. The co-occurrence of words within nodes revealed meaningful connections reflecting both positive and negative visitor experiences.

The analysis also revealed a spike in positive sentiments from late 2018 to the end of 2019, where positive feedback significantly outnumbered negative comments. This rise may be attributed to factors, including enhanced visitor services, targeted marketing campaigns, and an increasing global fascination with Egypt’s rich cultural heritage. Despite the fluctuations detected, the overall positive sentiment regarding the Giza Pyramids remains relatively stable over time, indicating a fundamental appreciation from visitors.

The sentiment analysis software and analytical tools contributed to a more in-depth understanding of heritage site experiences as expressed in user-generated online content. The findings answered the research questions as follows:

1. *What are the overall visitor sentiments regarding their experiences at the Giza Pyramids?*

The analysis showed that the overall visitor sentiment is predominantly positive, with many reviews expressing fascination and admiration for the Pyramids and their historical context. Visitors frequently describe their experiences as breathtaking and amazing. However, few reviews reflected negative sentiments, often stemming from issues such as overcrowding, commercialization, and pushy vendors.

2. *What specific aspects of the visit (e.g., guided tours, accessibility, facilities) contribute most to positive or negative sentiments?*

Positive sentiments are closely linked to professionally guided tours that provide in-depth historical insights, enhancing the visitor experience. Accessibility features, such as well-maintained paths and facilities, weren’t one of the great concerns of visitors, which could contribute negatively to the experience. Conversely, negative sentiments often arise from inadequately maintained areas and logistical challenges, such as long wait times and insufficient crowd management.

3. *How do visitors perceive the cultural significance of the Giza Pyramids, and how does this perception affect their overall sentiment?*

The findings indicated that the historical context of the Giza Pyramids is integral to visitor satisfaction. Visitors overwhelmingly perceive the Giza Pyramids as symbolic of Egypt's rich historical and cultural heritage. This perception enhances their emotional connection to the site, often resulting in feelings of admiration and appreciation. Many reviews reflect a desire to understand the historical context of the

Pyramids, which deepens their overall sentiment and promotes a more meaningful engagement with the site.

4. *What suggestions do visitors have for improving their experience at the Giza Pyramids?*

Visitors suggested improvements in crowd management, such as timed entry tickets to reduce congestion. Enhancements to facilities, including shade areas, and the establishment of diverse modes of transportation within the area, were recommended. Furthermore, some visitors expressed a desire for regulations regarding vendors and camel rides to enhance their overall experience. The analysis of visitors' feedback offers valuable insights for site management, enabling those responsible for the site to identify specific areas for improvement. By systematically addressing common concerns and implementing suggested enhancements, site management can create a more satisfying experience that aligns with visitors' expectations and needs.

The research recommends several aspects that officials and industry should take into consideration in order to enhance user experience at the Giza pyramids, and additionally take full advantage of the benefits of Big Data. Some barriers to using big data in cultural heritage, for instance, include issues related to data collection and the invasion of internet users' privacy. It is very important to prioritize responsible use and respect the privacy of users' data, as noted by Wang & Moriarty (2018).

It is also essential for scientists, researchers, historians, and cultural heritage site officials to work collaboratively when implementing big data in cultural heritage to ensure optimal results. Connecting these ends may be challenging, but it is indispensable for success.

Additionally, big data can be effectively used to uncover user patterns and build personas that replicate cultural engagement, enhancing knowledge of cultural heritage visitors and their unique demographics. This information can empower museum officials and those responsible for cultural heritage sites to provide visitors with more customized experiences, as illustrated in the study by De Mauro et al. (2015).

Tourists nowadays use social media platforms to document their travel experiences. They share their photos on Instagram, leave comments on Facebook, and post their reviews concerning travel services on TripAdvisor. A considerable number of travelers rely on these interactions to make decisions about various travel services. This highlights the necessity to trace consumer feedback in order to be able to satisfy their needs. This was emphasized by the work of Surugiu et al. (2022), who analyzed Facebook comments to evaluate visitor opinions on Romanian museum activities, Gitto & Mancuso (2017), who relied on data from the Skytrax blog to examine the quality of Airport service. In addition, Zhang et al. (2022) highlighted the importance of feedback tracking when they explored customer opinions on virtual tourism during COVID-19 from Weibo posts. Furthermore, the studies by Gharzouli et al. (2022), Zhu et al. (2021), and Jiang et al. (2021) evaluated customer satisfaction relying on customer feedback.

It is fundamental for the Giza Heritage site officials to closely monitor social media platforms and gather insights from online reviews, as noted in the study by Jangid et al. (2021). The researchers here categorized users based on their sentiments. They identified loyal fans, unsatisfied users, silent followers, and cheerleaders. They noted

that every category should be addressed differently to effectively engage them and turn them into advocates for the destination.

Lastly, the analysis showed that the Giza Pyramids remain a symbol of historical magnificence and a top global tourist attraction, with overwhelmingly positive sentiment from visitors despite certain persistent challenges. The findings highlighted the importance of visitor-centered site management, where cultural heritage is not only preserved but experienced in a way that meets modern expectations of comfort, safety, and engagement. This dynamic interchange of visitor sentiment and site management presents an ongoing opportunity to enrich the cultural and historical engagement that the Giza Pyramids offer to the world.

5. FUTURE RESEARCH

Future research may address topics linking Big Data with Cultural Heritage like analyzing visitor satisfaction using cross-media big data sources, leveraging big data for long-term digital preservation of cultural heritage, exploring predictive analytics to forecast heritage site usage, deterioration, or tourism trends, and using big data to document and preserve intangible heritage.

6. LIST OF REFERENCES:

Amanatidis, D.; Mylona, I.; Mamalis, S.; Kamenidou, I.E. (2020). Social media for cultural communication: A critical investigation of museums' Instagram practices. *J. Tour. Herit. Serv. Mark. (JTHSM),* Vol. 6, pp. 38–44.

Amato, F.; Moscato, V.; Picariello, A.; Colace, F.; Santo, M.D.; Schreiber, F.A.; Tanca L. (2017). Big data meets digital cultural heritage: Design and implementation of scrabs, a smart context-aware browsing assistant for cultural environments. *J. Comput. Cult. Herit*., 10, pp.1–23.

Antoniou, A. (2017). Social Network Profiling for Cultural Heritage: Combining Data from Direct and Indirect Approaches. *Soc. Netw. Anal. Min*., 7, pp.1–11.

Atabay, L., & Çizel, B. (2020). Comparative content analysis of hotel reviews by mass tourism destination. *Journal of Tourism and Services, 11*(21), pp.147–166.

Baniya, R., Dogru-Dastan, H., & Thapa, B. (2021). Visitors'experience at Angkor Wat, Cambodia: Evidence from sentiment and topic analysis. *Journal of Heritage Tourism*, 16(6), pp.632–645.

Burns.L, Moore.M, Speno.L &Talley-McRae.H. (2010). *Heritage tourism Handbook: A How-to-Guide for GeorGiA.* Georgia Department of Natural Resources & Tourism Division.

Camacho, D., Panizo-LLedot, A., Bello-Orgaz, G., Gonzalez-Pardo, A., & Cambria, E. (2020). The four dimensions of social network analysis: An overview of research methods, applications, and software tools. *Information Fusion*, *63*, pp.88–120.

Centobelli, P., & Ndou, V. (2019). Managing customer knowledge through the use of big data analytics in tourism research. *Current Issues in Tourism*, 22(15), pp.1862-1882.

Chittiprolu, V., Samala, N., & Bellamkonda, R. S. (2021). Heritage hotels and customer experience: a text mining analysis of online reviews. *International Journal of Culture, Tourism and Hospitality Research*, *15, 2,* pp. 131–156

De Mauro, A., Greco, M., & Grimaldi, M. (2015, February). What is big data? A consensual definition and a review of key research topics. In AIP conference proceedings (Vol. 1644, No. 1, pp. 97-104). American Institute of Physics.

Deligiannis, K.; Raftopoulou, P.; Tryfonopoulos, C.; Platis, N.; Vassilakis, C. (2020). Hydria: An Online Data Lake for Multi-Faceted Analytics in the Cultural Heritage Domain. *Big Data Cogn. Comput*., 4, 7.

Drivas, I.C.; Sakas, D.P.; Giannakopoulos, G.A.; Kyriaki-Manessi, D. (2020). Big Data Analytics for Search Engine Optimization. *Big Data Cogn. Comput*., 4, 5.

Escandon-Barbosa, D., & Salas-Paramo, J. (2021). Tourism Amidst COVID-19: consumer experience in luxury hotels booked through digital platforms. *Tourism Recreation Research*, 1–6.

Farisi, A. A.; Sibaroni, Y. and Al Faraby, S.(2019). Sentiment analysis on hotel reviews using Multinomial Naïve Bayes classifier. Journal of Physics: Conference Series, 1192, 012024.

Gharzouli, M., Hamama, A. K., & Khattabi, Z. (2022). Topic-based sentiment analysis of hotel reviews. *Current Issues in Tourism*, 25(9), pp.1368–1375.

Gitto, S., & Mancuso, P. (2017). Improving airport services using sentiment analysis of the websites. *Tourism Management Perspectives*, 22, pp.132–136.

Innocente, C., Ulrich, L., Moos, S., & Vezzetti, E. (2023). A framework study on the use of immersive XR technologies in the cultural heritage domain. *Journal of Cultural Heritage, 62*, pp.268–283.

Jangid, A., Mishra, A., & Sharma, A. (2021). Effective marketing strategies for the tourism industry using sentiment analysis. *IOSR Journal of Business and Management (IOSR-JBM), 23*(6, Ser. VIII), pp.59–66.

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

Julia, C. & Junshean, K. (2015). Fun in The Philippines: Automatic Identification and Sentiment Analysis of Tourism-related Tweets. *IEEE 15th International Conference on Data Mining Workshops,* pp. 660-667.

Khomsah, S., Ramadhani, R. D., & Wijayanto, S. (2021). Big data analytics to analyze sentiment, emotions, and perceptions of travelers (Case study: Tourism destination in Purwokerto, Indonesia). *Jurnal E-KOMTEK (Elektro-Komputer-Teknik), 5*(2), pp.284–297.

Konstantakis, M.; Alexandridis, G.; Caridakis, G. (2020). A Personalized Heritage-Oriented Recommender System Based on Extended Cultural Tourist Typologies. *Big Data Cogn. Comput*., 4, 12.

Kontogiannis, S.; Kokkonis, G.; Kazanidis, I.; Dossis, M.; Valsamidis, S. (2020). Cultural IoT Framework Focusing on Interactive and Personalized Museum Sightseeing. In Towards Cognitive IoT Networks; Springer: Cham, Switzerland, pp. 151–181.

Kosmopoulos, D.; Styliaras, G. (2018). A survey on developing personalized content services in museums. *Pervasive Mob. Comput*., 47, pp.54–77.

Lama, R. (2024). Importance of social media platforms in tourism industry. In *Futuristic trends in management* (Vol. 3, Book 26). IIP Series.

Li, H., Hu, M., Li, G., (2020). Forecasting tourism demand with multisource big data. *Annals of Tourism Research*, 83, pp.1-13. 102912.
Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. *Tourism Management*, 68, pp.301-323.
Litvin, W. S., Goldsmith, E. R., & Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism Management. *Tourism Management*, 29, 3, pp. 458–468.
Liu, Y., Huang, K., Bao, J., & Chen, K. (2019). Listen to the voices from home: An analysis of Chinese tourists' sentiments regarding Australian destinations. *Tourism Management*, 71, pp.337–347.
Lu, J. (2022). Personalized Recommendation Algorithm of Smart Tourism Based on Cross-Media Big Data and Neural Network. *Computational intelligence and neuroscience*, *2022*, 9566766.
Lu, Y., & Zheng, Q. (2021). Twitter public sentiment dynamics on cruise tourism during the COVID-19 pandemic. *Current Issues in Tourism*, 24 (7), pp.892–898.
Mariani, M. (2020). Big data and analytics in tourism and hospitality: a perspective article. *Tourism Review*, 75 (1), pp.299-303.
McKercher, B.; Du Cros, H. (2003). Testing a Cultural Tourism Typology. *Int. J. Tour. Res.*, 5, pp.45–58.
Mondal, S., & Samaddar, K. (2022). Future of sharing economy and its resilience post pandemic: A study on Indian travel and tourism industry. *Management of Environmental Quality: An International Journal, 33*(6), pp.1591–1610.
Orea-Giner, A., Fuentes-Moraleda, L., Villace´-Molinero, T., Mun˜ oz-Mazo´ n, A., & Calero-Sanz, J. (2022). Does the implementation of robots in hotels influence the overall tripadvisor rating? A text mining analysis from the industry 5.0 approach. *Tourism Management,* 93, 104586.
Park, E., Kang, J., Choi, D., & Han, J. (2020). Understanding customers' hotel revisiting behaviour: A sentiment analysis of online feedback reviews. *Current Issues in Tourism*, 23(5), pp.605–611.
Poulopoulos, V., & Wallace, M. (2022). Social media analytics as a tool for cultural spaces—The case of Twitter trending topics. *Big Data and Cognitive Computing, 6*(2).
Ríos-Martín, M. Á., Folgado-Fernández, J. A., Palos-Sánchez, P. R., & Castejón-Jiménez, P. (2020). The impact of the environmental quality of online feedback and satisfaction when exploring the critical factors for luxury hotels. *Sustainability, 12*(1), 299.
Salas-Olmedo, M. H., Moya-Gómez, B., García-Palomares, J. C., & Gutiérrez, J. (2018). Tourists' digital footprint in cities: Comparing Big Data sources. *Tourism Management, 66*, pp.13–25.
Solano-Barliza, A., Arregocés-Julio, I., Aarón-Gonzalvez, M., Zamora-Musa, R., De-la-Hoz-Franco, E., Escorcia-Gutiérrez, J., & Acosta-Colla, M. (2024). Recommender systems applied to the tourism industry: A literature review. *Cogent Business & Management, 11*(1), 2367088.
Spiliotopoulos, D.; Margaris, D.; Vassilakis, C. (2020). Data-Assisted Persona Construction Using Social Media Data. *Big Data Cogn. Comput.*, 4, 21.

Statista, (2024a). *Number of visits to travel booking sites worldwide 2014*. [online] Available at: https://www.statista.com/statistics/459983/number-of-visits-to-travel-booking-sites-worldwide/. [Accessed 16 November 2024]

Statista, (2024b). *TripAdvisor: Total website visits worldwide 2024*, *Statista*. Available at: https://www.statista.com/statistics/1215473/total-visits-to-tripadvisor-website/ [Accessed: 16 November 2024].

Surugiu, C., Tudorache, D. M., Asxtefa˘noaiei, R. M., & Surugiu, M. R. (2021). Bucharest cultural heritage through the eyes of social media users. *Anatolia*, 32(4), pp.576–587.

Teles da Mota, V., & Pickering, C. (2020). Using social media to assess nature-based tourism: Current research and future trends. *Journal of Outdoor Recreation and Tourism*, 30, 100295.

TripAdvisor reports second quarter 2023 financial results, 2023. Available at: https://ir.tripadvisor.com/static-files/d6752890-85a0-4764-a90e-f6213480c7e1. [Accessed 16 November 2024].

TripAdvisor reports second quarter 2024 financial results (2024). Available at: https://ir.tripadvisor.com/static-files/9dc2a668-7789-4445-b7c7-61d82930b5d9: (Accessed 18 November 2024).

UNESCO (2024). *What is meant by "cultural heritage"?* Unesco.org. [online] Available at: https://uis.unesco.org/en/glossary-term/cultural-heritage [Accessed September 2024]

Vassilakis, C.; Antoniou, A.; Lepouras, G.; Poulopoulos, V.; Wallace, M.; Bampatzia, S.; Bourlakos, I. (2017). Stimulation of reflection and discussion in museum visits through the use of social media. *Soc. Netw. Anal. Min*., 7, 40.

Višković, K., Rašan, D., & Prevolšek, D. (2022). Content analysis of TripAdvisor online reviews: The case of Valamar Riviera hotels in Dubrovnik. In *Tourism & Hospitality Industry 2022: Congress proceedings* , pp. 307–319.

Wallace, M., Poulopoulos, V., Antoniou, A., & López-Nores, M. (2023). An overview of big data analytics for cultural heritage. *Big Data and Cognitive Computing, 7*(1), 14.

Wang, S.J., Moriarty, P. (2018). Barriers to the Implementation of Big Data. In: Big Data for Urban Sustainability. Springer, Cham.

Widyawati, R. S., Irawan, H., & Ghina, A. (2021). Content analysis of tourist opinion based on tourism quality (TOURQUAL) by text mining online reviews: The case of Borobudur. In *Proceedings of the 1st International Conference on Sustainable Management and Innovation (ICoSMI 2020), 14–16 September 2020, Bogor, West Java, Indonesia*.

Yetgin, D., Keküç, M. & Şayin, M. (2020). Regulation on Inspection and Control of Food Safety and Quality. *Journal of Tourism and Gastronomy Studies*, 8 (3), pp. 2090-2109.

Zhang, S. N., Li, Y. Q., Ruan, W. Q., & Liu, C. H. (2022). Would you enjoy virtual travel? The characteristics and causes of virtual tourists' sentiment under the influence of the COVID-19 pandemic. *Tourism Management*, 88, 104429.

Zhang, X., Yang, Y., Zhang, Y., & Zhang, Z. (2020). Designing tourist experiences amidst air pollution: A spatial analytical approach using social media. *Annals of Tourism Research,* 84, 102999.

Zhu, J., Cheng, M., & Li, Z. (2021). Peer-to-peer accommodation experience and guest actual recommendations: A novel mixed-method approach. *Tourism Management Perspectives,* 38, 100816.