

Sentiment Analysis of Chinese Tourists to Marine Tourism Resource in Sabah

QiQi^{1*}, Rosazman Hussin², Md Sayed Uddin³

¹Faculty of Social Sciences and Humanities, Universiti Malaysia Sabah (UMS), 88400 Kota Kinabalu, Sabah, Malaysia.

Email: qiqi6077wzw@gmail.com

²Borneo Institute for Indigenous Studies (BorIIS), Universiti Malaysia Sabah (UMS), 88400 Kota Kinabalu, Sabah, Malaysia.

Email: azzs@ums.edu.my

³Faculty of Social Sciences and Humanities, Universiti Malaysia Sabah (UMS), 88400 Kota Kinabalu, Sabah, Malaysia.

Email: sayed@ums.edu.my

CORRESPONDING AUTHOR (*):

QiQi

(qiqi6077wzw@gmail.com)

KEYWORDS:

Sentiment Analysis

Chinese Tourist

Marine Tourism

CITATION:

Qi, Q., Rosazman Hussin, & Md Sayed Uddin. (2024). Sentiment Analysis of Chinese Tourists to Marine Tourism Resource in Sabah. *Malaysian Journal of Social Sciences and Humanities (MJSSH)*, 9(7), e002937. <https://doi.org/10.47405/mjssh.v9i7.2937>

ABSTRACT

With the tourism industry's recovery, there is an increased demand for outbound Chinese tourists. With Chinese tourists as an important submarket in Sabah, a better understanding of Chinese tourists' perceptions of destination tourism is crucial for development. In this study, the representative Tanjung Aru Beach in Sabah was selected as the study case, and the review content was selected through the travellers' reviews on China's largest travel online platform - Ctrip.com, as the study sample. Through Octopus Collector 8, the text data was analyzed for word frequency, web semantics, and sentiment using the ROST CM6 tool. The results of the study found that most of the Chinese tourists at Tanjung Aru Beach gave satisfactory comments. These positive comments focused mainly on the natural landscape and resources of the beach, the services of the Shangri-La Hotel, the ease of language communication, and the local cuisine. More often than not, Chinese tourists mentioned the beach's sunsets, evening sunsets, flaming clouds, sand, cloud formations, and sea breezes. The study also found that "beach" and "sunset" were the resources that Chinese tourists recognized for the attraction, with the majority of reviews revolving around these two themes. It should be noted that Chinese tourists showed dissatisfaction with the sanitary environment and facilities at Tanjung Aru Beach. The study concludes with recommendations for management.

Contribution/Originality: This is one of the few studies that analyzes the sentiment of Sabah's marine tourism resources from the perspective of Chinese tourists. This study partially compensates for the dearth of research on Chinese tourists in

Malaysian tourism studies.

1. Introduction

The growing number of Internet users and social media users has created a huge amount of data in the form of text, video, images, and audio. According to the [China Internet Network Information Center \(2023\)](#) by the end of 2022, the size of China's Internet users reached 1.067 billion, and the size of online travel booking users reached 423 million. Compared with the number of users using online travel booking in 2021, the number of users increased by 25.61 million ([China Internet Network Information Center, 2023](#)).

Compared to traditional sociological survey methods and traditional questionnaire data, these online data are freely available and save money and human resources ([Kumar et al., 2021](#); [Martin et al., 2018](#); [Frias-Martinez et al., 2012](#)). Text sentiment analysis is a method to identify the emotional attitudes of the study object. It extracts the substance of the reviewer's emotional appraisal of the product or service by evaluating the review content to uncover positive or negative attitudes and views.

Sentiment analysis transforms qualitative data into quantitative data, which can mainly accomplish topic extraction, identification of opinion holders, determination of the scope of reviews, and analysis of positive and negative sentiment. The advantage of text analytics is also in its speed, as it can obtain and process a large number of reviews within a certain period of time. These comments can help other users make consumer decisions, as well as help companies better understand the value of their brands and improve their services ([Dholakia & Zhao, 2010](#)).

With the tourism industry's recovery, China's tourism market momentum and residents' willingness to travel outbound have increased. The [China Tourism Association \(2024\)](#) reports that Chinese travelers surpassed 87 million in 2023 and anticipates exceeding 130 million in 2024. It also states that online platforms are the primary source of information for Chinese tourists on outbound trips, with accommodation, transport, and travel price information being the most popular searches, and that Chinese tourists are increasingly valuing the sense of experience on their trips ([China Tourism Association, 2024](#)).

Sabah is rich in marine resources, and the uniqueness of its coastal parks attracts many local and foreign tourists ([Mohammed et al., 2018](#)), not to mention Perhentian Kecil Island and Tioman Island, which were ranked as the top 50 beaches in the world by CNN ([Ministry of Tourism and Culture Malaysia, 2015](#)). In the summer of 2024, Chinese travelers have a strong demand for outbound travel, with Asia being their primary destination choice. In particular, Singapore, Malaysia Thailand and other countries with mutual visa exemption with China have become the first choice for Chinese tourists to visit in the summer ([Tongcheng Travel Holdings Limited, 2024](#)). According to [Sabah Tourism Board \(2019\)](#), the number of Chinese tourists is close to half of the total number of inbound tourists. From a supply and demand perspective, it can be concluded that there is a growing demand for outbound tourism from Chinese tourists and Sabah as a tourist destination has been welcoming a large number of Chinese tourists in the past few years.

From the previous statistics, it can be concluded that China is an important submarket for the development of tourism in Sabah. Utilizing the advantages of Sabah's marine tourism resources and understanding Chinese tourists' perception of Sabah's marine resources is particularly important for the development of Sabah's tourism industry. Therefore, Tanjung Aru Beach was chosen as the research site for this study. This is because the site is the most visited and commented site by mainland Chinese tourists when they visit Sabah.

1.1. Research Objectives

This study used the online reviews of Tanjung Aru attraction from Chinese tourism websites as the research data, which were analyzed through word frequency analysis, online semantic analysis and sentiment analysis in order to understand Chinese tourists' perceptions of marine resources in Sabah.

2. Literature Review

With the emergence of large amounts of digitized text on online platforms, researchers have turned their attention to building theoretical framework theories relying on text mining (Kar & Dwivedi, 2020). Researchers have mined unstructured data from online platforms using topic modeling, document classification, natural language processing(NLP),sentiment analysis, language recognition, and web semantics (Mahr et al., 2019; Villarroel & Zhang, 2019). Text mining, also known as text analytics, is an artificial intelligence technique that transforms unstructured data into structured data by using NLP to augment the analysis of machine learning algorithms. Text mining as a data analysis approach has received attention from researchers for its theoretical constructs that utilize unstructured data (Mittal & Agrawal, 2022).

Flores-Ruiz et al. (2021) compared tourist sentiment and behavior before and after the COVID-19 outbreak through social media data analysis. Tweets were mined using the statistical programming language Rstudio for the social network media-Twitter. The data was categorized as collected from August to November 2019 and from August to November 2020 after the COVID-19 outbreak. The software Knime was used to analyze the data, modeling, processing and visualization. The study found that the outbreak had an impact on the behavior of tourists, who were more safety-conscious and tended to travel alone to destinations with shorter and less crowded distances.

Vetitnev et al. (2021) analyzed the emotional attitudes of tourists by mining reviews of resorts in Sochi, Russia, on a travel online platform (TripAdvisor). Python programming language was used for the study. The study found that tourists' overall assessment of the city of Sochi is positive. Tourists were most satisfied with the three attractions rated as alpine ski resorts, cultural centers and embankments. On the contrary, water parks, dolphinariums and aquariums were rated relatively negatively.

Guo et al. (2024) utilized a web crawler, Octopus Collector 8, to collect data from Chinese service retail and commodity retail website, Meituan.com, and a travel online platform, Ctrip.com. The study was conducted on a scenic spot in Zhaoqing City, and the ROST CM6 system was used to process the data (Guo et al., 2024). This study divided the high-frequency words into three first-level themes: "destination factors", "tourist factors" and "objective factors". The following secondary themes were proposed based on the three primary themes, namely, consumption, service, entertainment,

transportation, food, landscape, group, experience, climate, and time. Sentiment analysis was conducted to analyze the positive, neutral and negative percentages based on the secondary themes. It was found that tourists care more about destination factors; management and landscape are important aspects of scenic spots to enhance their competitiveness (Guo et al., 2024).

Ren and Zhao (2024) conducted a study on tourists' perceived characteristics of parks and their influencing factors through text analysis. The study utilized an online review platform, Dianping.com, to obtain comprehensive perceptual characteristics of 10 parks in Guangzhou, China. The data were analyzed by ROST CM6 for high-frequency word extraction, affective tendency and semantic network analysis. It was found that each park was rated differently. Visitors preferred to review the parks in spring and summer; visitors were most satisfied with the parks in spring and winter time; attractions of water bodies, plants, fitness facilities, monumental buildings, play facilities and resting places were more likely to be perceived by visitors; and transportation, admission fees, air quality, internal facilities and activities were the main factors positively rated by visitors.

Lu (2024) conducted a textual analysis of the satisfaction of tourists on online travel platforms, using Python crawler to obtain, Ctrip.com (www.ctrip.com) reviews of the top three attractions in the popular cities for the Labor Day Holiday and National Day Holiday. ROST CM6 was selected as the data analysis tool. Through the analysis of tourists' visit demand for different types of scenic spots, it is found that tourists' demand focuses differently for different scenic spots (Lu, 2024). For scenic spots with natural resources, tourists pay more attention to their scenic services and tourists' experience; for scenic spots with humanistic resources, tourists think that the characteristic buildings, tour guide's explanation service and historical buildings can highlight the core competitiveness of this type of attractions; for scenic spots with development resources, users think that the experience of parent-child interaction and tourists' experience in tourism is the aspect with the greatest demand. Chinese tourists' overall comments on scenic spots, positive comments accounted for 76.75% of the overall comments. Negative comments accounted for 6.76% (Lu, 2024).

Aiming at the perception of cruise brand image and tourists' perception and sentiment analysis of the product, the researchers took the tourists' review data from China's largest online travel platform (Ctrip.com) as a sample, and took the cruise companies as the research object, and carried out text analysis through ROST CM6 (Sun & Ni, 2018).

Through word frequency analysis, semantic network and sentiment analysis, and satisfaction analysis, different brands of cruise ships were analyzed for differences in product image, sentiment expression and satisfaction (Sun & Ni, 2018). It was found that Chinese tourists showed high satisfaction in the attributes of cruise service, accommodation, catering and excursions, but negative evaluations of shore-based viewing. It was also found that there was no significant difference between the core images of different cruise brands (Sun & Ni, 2018).

In terms of tourists' perceptions of cultural heritage ecosystems, Zhang et al. (2024) conducted a content analysis using web data. Travelogues and reports related to the city wall cultural heritage in Nanjing, China were collected using Python web crawler. The word frequency statistics, semantic network analysis and sentiment analysis by ROST CM6 software were used to obtain tourists' ecological environment perception,

emotional perception and overall image perception of Nanjing cultural city wall heritage. It was found that tourists held positive emotions towards the environment of Nanjing city wall cultural heritage; the water environment and plantscape had a particularly significant impact on the image of cultural heritage.

Most of the current studies related to marine tourism resources focus on the relationship between variables and verify the relationship through empirical studies (Su et al., 2021; Papilaya, 2018; Dimitrovski et al., 2021). We can see that most of the studies have used traditional research methods to obtain data. The themes of the studies include a focus on behavioral intentions, such as satisfaction and loyalty. Environmental elements of marine resources are also another focus of researchers (Dimitrovski et al., 2021; Chen et al., 2020).

Understanding tourists' perceptions of destinations and attractions is essential for management and marketing promotional efforts. Su et al. (2021) investigated the relationship between perceived authenticity and perceived value and loyalty of tourists at marine attractions. The study utilized a quantitative survey method and examined the relationship between the variables.

Papilaya (2018) conducted a study on the relationship between marine tourism demand, characteristics and number of visits with the level of perception and willingness to pay for marine tourism destinations. The study was conducted by interviewing tourists from overseas, domestic and local tourists through an interview survey. Data collection was done through questionnaires. The study used Amos to perform structural equation analysis and path analysis on the data. The study found that there is a close relationship between the variables. There was a direct effect of demand for marine tourism, tourist perception and willingness to pay, while the relationship was negative when number of visits was the intervening variable.

Dimitrovski et al. (2021) explored the sustainability of coastal, marine and cruise tourism. The study investigated the perceptions of relevant stakeholders on the future development of coastal and marine tourism through a mixed qualitative and quantitative approach. At the same time, the study established that the relationship between cruise tourists' attitudes and norms and behavioral model mediator model was established through quantitative research.

In order to understand tourists' attitudes and behaviors towards marine resource tourism ecosystems, Mohammed et al. (2018) selected four island parks in Malaysia as the study population and collected data through questionnaires. Structural equation modeling was used to investigate the relationship between destination attractiveness, satisfaction and tourism behavior. The results confirmed that tourists who were satisfied with the coastal parks showed more environmentally responsible attitudes and behaviors.

Chen et al. (2020) used a questionnaire survey to explore the relationship between stakeholder perceptions (locals and tourists, different occupational groups and education level groups) and the suitability of marine ecological parks for tourism development. The study found that the low environmental awareness of the residents of Dongshan Island in Shandong Province was the most significant barrier to the development of ecotourism in the destination. The destination needs to enhance environmental education for local residents.

A review of previous studies reveals that in tourism research, studies conducted using online texts are the choice favored by researchers in recent years (Kar & Dwivedi, 2020; Lu, 2024; Sun & Ni, 2018). In the study of marine resource tourism destinations or attractions, studies using traditional research methods still exist (Chen et al., 2020; Dimitrovski et al., 2021; Papilaya, 2018). Therefore, whether the research is conducted using traditional research methods or using online reviews, it is an important research method in tourism research nowadays. Researchers can choose the appropriate research method according to their own research objectives.

3. Research Methods

This study utilizes web crawler means for crawling, by using Octopus Collector 8. Ctrip.com (www.ctrip.com) reviews of Tanjung Aru Beach in Sabah by Chinese tourists was used as the data source website. The statistical information of data from Ctrip.com reveals that Tanjung Aru Beach has the highest number of visits and online reviews from Chinese tourists compared to other tourist attractions with marine resources in Sabah. Tanjung Yalu Beach, which takes its name from the mahogany yellow next to the beach. Tanjung Aru Beach offers the best view of the sunset as the crimson sun slowly dips into the horizon, turning the vast sky a brilliant red. Tanjung Aru Beach is also located next to Shangri-La's Tanjung Aru Resort and Spa and Kinabalu Golf Club.

Ctrip.com, the largest Chinese online tour operator (OTA) in terms of customer visits, provides Chinese tourists with comprehensive information on tourism products, and the reviews on this website are well represented. Octopus Collector8 was utilized to review online reviews of Tanjung Aru Beach, Sabah by Chinese tourists. The captured review data included: user name, overall rating of Tanjung Aru, user-generated content (UGC), user address and time of posting the review. Secondly, to ensure the validity completeness and accuracy of the reviews, the obtained online reviews were processed by removing reviews prior to 2016 as well as duplicates. Finally, 845 valid comment data were obtained.

This study selected the ROST CM6 tool from Wuhan University as a text analysis tool, which is currently the only large-scale free computing platform to assist humanities and social sciences research within China. In this study, the ROST CM6 tool was used to analyze the text for word frequency analysis, semantic network analysis and sentiment analysis. Before textual analysis of the data, the data need to be preprocessed into a desirable format. For this purpose, the data was subjected to word segmentation through ROST CM6 tool.

4. Results

4.1. Word frequency analysis

Word frequency analysis is to count and analyze the frequency of important words appearing in the online comments, the more frequently the words appear indicates that the tourists' attention to or perception is stronger. We can analyze the high-frequency words to understand the evaluation of Chinese tourists' perception of the overall image of Tanjung Aru Beach.

As shown in the [Table 1](#), the first 60 high-frequency words were selected. These high-frequency keywords show the core tourism attributes and resources of the destination, as well as the overall evaluation of the destination resources by tourists. Based on [Table 1](#), it can be found that the keywords for Tanjung Aru Beach are: sunset, beach, evening sun, fire clouds, sand, sky, clouds and sea breeze. These keywords demonstrate Chinese tourists' understanding of Tanjung Aru as a tourist attraction and its main components of tourism resources. It can also be observed that another high-frequency vocabulary is the tourists' evaluation of the attraction and their feelings while traveling, for example, beautiful, top three, worthwhile, best, appreciated, and comfortable.

Table 1: Statistics of high-frequency words in online reviews of Tanjung Aru by Chinese tourists

| Keyword (Chinese characters) | Keyword (in English) | Word frequency | keyword | Keyword (in English) | Word frequency |
|------------------------------|----------------------|----------------|---------|----------------------|----------------|
| 日落 | sunsets | 747 | 海水 | Sea | 42 |
| 海滩 | Beach | 462 | 天空 | Sky | 42 |
| 酒店 | Hotel | 239 | 分钟 | Minutes | 41 |
| 落日 | setting sun | 236 | 景点 | Attractions | 37 |
| 沙滩 | sandbeach | 218 | 三大 | Top three | 37 |
| 香格里拉 | Shangri-La | 206 | 附近 | Nearby | 37 |
| 丹绒亚路 | Tanjung Aru | 202 | 适合 | Suitable | 36 |
| 地方 | Places | 110 | 下午 | Afternoon | 35 |
| 亚庇 | Kota Kinabalu | 105 | 风景 | Scenery | 34 |
| 沙巴 | Sabah | 101 | 享受 | Enjoy | 34 |
| 漂亮 | Beautiful | 88 | 云层 | Clouds | 32 |
| 夕阳 | Declining sun | 86 | 位置 | Location | 32 |
| 景色 | Scenery | 82 | 海风 | Sea Breeze | 31 |
| 美丽 | Beauty | 64 | 旁边 | Next to | 31 |
| 太阳 | sun | 62 | 当地人 | Locals | 31 |
| 市区 | Downtown | 59 | 干净 | cleanly | 30 |
| 天气 | weather | 51 | 运气 | Luck | 30 |
| 免费 | free | 50 | 可惜 | Pity | 30 |
| 火烧云 | flaming clouds | 49 | 拍照 | Take picture | 29 |
| 欣赏 | enjoy | 49 | 公共 | Public | 29 |
| 打车 | take a taxi | 48 | 舒服 | Comfortable | 28 |
| 沙子 | Sandy | 47 | 震撼 | Stunning | 27 |
| 美景 | Fine view | 47 | 环境 | Environment | 27 |
| 观赏 | Sightseeing | 46 | 果汁 | Juice | 27 |
| 最佳 | Best | 46 | 地点 | Location | 25 |

| | | | | | |
|----|--------------|----|----|-----------|----|
| 海边 | Seaside | 44 | 每天 | Everyday | 25 |
| 游客 | Visitors | 44 | 最好 | Best | 25 |
| 时间 | Time | 42 | 下雨 | Rainy | 24 |
| 值得 | Worthwhile | 42 | 机场 | Airport | 24 |
| 晚霞 | evening glow | 42 | 滩边 | Beachside | 23 |

We locate the content of specific Chinese tourists' comments based on the above high-frequency words, and we can find that most of the tourists' comments describe the scenery of Tanjung Aru Beach. For example, *"The sunset is very beautiful, and the sand on the beach is very fine and smooth like a mirror. It is said to be one of the top three spots in the world for sunset viewing. It was a bit cloudy when I went there but it was already beautiful, I hope to see flaming clouds in a while."*, *"One of the most beautiful sunsets in the world, definitely worth going to see once."*, *"A stunning place to watch the sunset, great feeling, memorable and unforgettable."*, *"Came to Kota Kinabalu twice and every time I came here to watch the sunset, it's really beautiful."*, *"The sunset is beautiful, worth going back again and again."*. There is also a part of high frequency words in the reviews that show the infrastructure of the tourist attraction, these words are, Shangri-La, hotel, airport.

Many tourists gave brief descriptions of travel tips in Sabah. Suggestions for transportation and accommodation as well as food were presented in the reviews. For example, *"Tanjung Aru Beach is about 20 minutes away from downtown Kota Kinabalu, so it's quite convenient to take a taxi there"*, *"When traveling to Sabah, you can rent a car at the airport, which costs more than 200 RMB a day"*, *"I stayed at two Shangri-La hotels, and recommended Tanjung Aru. Shangri-La hotels, recommend the Shangri-La Hotel in Tanjung Aru."*, *"The attraction is next to the Shangri-La Hotel. Standing on the beach you can see the airport. Took a taxi from downtown and paid RM15 for the ride after haggling."*, *"There is a food court on the beach with good BBQ and good juices."*

4.2. Semantic network analysis

The semantic network diagram (Figure 1) shows the relationship between the high-frequency words of the content of Chinese tourists' evaluation of Tanjung Aru Beach. It can be found that Chinese tourists consider Tanjung Aru Beach as a popular attraction in Kota Kinabalu, Sabah. It has a beautiful sunset and evening view as well as a clean beach with a comfortable climate.

The semantic web map shows not only the tourism resources of the attraction, but also the recognition and praise from Chinese tourists. Chinese tourists consider "beach" and "sunset" as the unique core tourism resources of Tanjung Aru, and other words around these two words include descriptive words, evaluative words, and words of other tourism resources. We can find that Chinese tourists consider Tanjung Aru, which is not far from downtown Kota Kinabalu, to be a worthwhile destination because of its beautiful beaches and the stunning sunset view in the evening. Moreover, Chinese tourists also recommend staying at Shangri-La Hotel, a hotel near Tanjung Aru. Chinese tourists recommend this hotel because it has the best spot to watch the sunset.

The statistics for negative emotions were: the number of comments with an average degree of negativity was 64; the number of comments with a moderately negative emotion was 25, and the number of comments with a very negative emotion was only two.

Positive review content demonstrates the core competencies of the destination. Most of the positive review content is positive feelings about the beauty of the sunsets at Tanjung Aru Beach, and expresses the intention to recommend and revisit the place. For example, *"It's great, great place to watch the sunset"*, *"Not visiting Tanjung Aru Beach is like not visiting Sabah, the sunset is beautiful from any angle at the beach, this sunset is mesmerizing"*, and *"The most beautiful beach in Sabah, the best choice to see the sunset, would like to go again"*, *"The purpose of the trip is to see the sunset, very good"*, *"The moment I saw the sunset, my heart is full of happiness, the power of nature always brings us touching and shocking. The power of nature always brings us touched and shocked"*, *"This is the best place to watch the sunset in Kota Kinabalu, and the kids love it"*, *"The scenery here is very beautiful and recommended"*.

Other tourist features of Sabah are also recommended and evaluated in the positive reviews of Tanjung Aru by Chinese tourists. For example, *"After watching the sunset, you must go to the beach night market to buy the four-color juice for RM15, it's very good, and only this night market sells it"*, *"If you want to watch it better, you can book the Shangri-La Tanjung Aru Hotel, you can watch the sunset while drinking tea"*. *"You can stay at Shangri-La Hotel Tanjung Aru, the service is very good, and the staff is very enthusiastic"*, *"The hotel's supporting facilities are very good, exclusive beach, pool and children's playground, it is very suitable for parent-child travel, and I would like to bring my children here in the future when I have the opportunity"*.

"Stayed two nights at the Shangri-La Hotel because of poor English, fortunately there are Chinese-speaking staff here", *"There are a lot of bubble puters on the beach, you can rent their tools and use them for taking photos"*, *"After watching the sunset, you can also go to the downtown for meat bone tea"*. From the content of the positive reviews, it was found that the uniquely attractive natural landscape, accommodation facilities, good food, service, and ease of language communication are the destination attributes that attract Chinese tourists to visit and generate pleasurable emotions.

Most of the negative comments were about the beach and its facilities, such as *"the beach was very sunny in the afternoon, and the water quality was mediocre"*, *"it rained before I left the house, and the clouds were too thick, so I couldn't see the sunset, and my heart was full of disappointment"*, *"The sea is not clean, the sand is not fine and white, rather disappointing"*, *"the beach has a lot of garbage"*, *"you need to pay to use the toilet, and the sanitary condition is quite bad"*. Fewer tourists said they were not satisfied with the taxi fare, for example, *"a round trip taxi costs RM60, plus meals and drinks, the value for money is not good"*.

5. Conclusion

Based on the evaluation data of Chinese tourists on Tanjung Aru Beach in Sabah from Ctrip.com, word frequency analysis, web semantic analysis, sentiment and satisfaction analysis were conducted. We found that most tourists are satisfied with Tanjung Aru Beach. Chinese tourists showed positive comments on the natural landscape and resources of the beach, the services of Shangri-La Hotel, the ease of language

communication and local cuisine. Chinese tourists mentioned the beach's sunsets, evening sunsets, flaming clouds, sand, clouds and sea breezes more often, and the most frequent descriptions of the beach by Chinese tourists were: beautiful, worthwhile, the best, and comfortable. The study also found that "beach" and "sunset" were the resources that Chinese tourists recognized in the attraction, and most of the evaluations were centered around these two theme words.

It should be noted that Chinese tourists showed dissatisfied comments about the sanitary environment and facilities at Tanjung Aru Beach. Very few tourists also had negative comments because they could not see the beautiful sunset due to the weather. The reasons for these negative comments: on the one hand, it may be due to the fact that Tanjung Aru Beach does have problems with the sanitary environment, such as the lack of timely cleanup of garbage on the beach and the poor sanitary conditions of the beach's toilets.

On the other hand, Chinese tourists are not familiar with the climate of Sabah and do not have a good travel strategy, resulting in a visit during the rainy season in Sabah and fail to meet the expectations of the tour. Another reason for the negative comments may stem from differences in living environment and culture. For example, some tourists mentioned the need to pay to use the toilet, a phenomenon that may not be commonly seen in Chinese tourist destinations in recent years. In particular, Tanjung Aru Beach being a free attraction, Chinese tourists may not be able to accept the need to pay to use the restroom.

Based on the results of the study, the following management recommendations are proposed. First, destination managers should optimize the core resources of the attraction to provide tourists with more viewing and playing options. The results of the study found that Tanjung Aru Beach's sunset, sand, sky, and clouds were mentioned more often by Chinese tourists, which indicates that the attraction's resources are recognized. Destination managers should pay attention to the protection of marine resources and environment to avoid excessive development. Strengthen ecological environment monitoring to prevent environmental pollution. Managers should pay attention to the sanitary environment and try to provide a clean beach environment for tourists.

On the other hand, destination managers can strengthen the construction of attractions to provide tourists with richer tourism products and improve their experience. The semantic analysis shows that Chinese tourists' comments revolve around "beach" and "sunset", which reflects that Chinese tourists do not recognize other tourism resources of Tanjung Aru attractions. Therefore, the destination can set up a number of photographic cards at the beach to meet the demand of tourists to take pictures and improve the experience of tourists on the attraction. It can also set up some stores where tourist souvenirs can be purchased to meet the needs of tourists to buy specialty products. Destination managers also need to improve the infrastructure and work to improve service capacity. Improve the hygienic state of Tanjung Aru beach toilets.

Ethics Approval and Consent to Participate

All research procedures involving human participants in this study complied with the research organization's ethical standards. Informed consent was obtained from all participants involved in the study.

Acknowledgement

Part of this article was extracted from a doctoral thesis submitted to Universiti Malaysia Sabah, Kota Kinabalu.

Funding

This study received no funding.

Conflict of Interest

The authors reported no conflicts of interest for this work and declare that there is no potential conflict of interest with respect to the research, authorship, or publication of this article.

References

- Chen, F., Lai, M., & Huang, H. (2020). Can marine parks become ecotourism destinations? Evidence from stakeholders' perceptions of suitability. *Ocean & Coastal Management*, 196, 105307. <https://doi.org/10.1016/j.ocecoaman.2020.105307>
- China Internet Network Information Center. (2023). *China Internet development report*. Retrieved from <https://baijiahao.baidu.com/s?id=1794250384458690555&wfr=spider&for=pc>
- China Tourism Association. (2024). *Tourism statistics*. Retrieved from <http://www.ctaweb.org/html/lysjsx/index.html>
- Dholakia, R. R., & Zhao, M. (2010). Effects of online store attributes on customer satisfaction and repurchase intentions. *International Journal of Retail and Distribution Management*, 38(7), 482-496. <https://doi.org/10.1108/09590551011052098>
- Dimitrovski, D., Lemmetyinen, A., Nieminen, L., & Pohjola, T. (2021). Understanding coastal and marine tourism sustainability: A multi-stakeholder analysis. *Journal of Destination Marketing & Management*, 19, 100554. <https://doi.org/10.1016/j.jdmm.2021.100554>
- Flores-Ruiz, D., Elizondo-Salto, A., & Barroso-González, M. D. L. O. (2021). Using social media in tourist sentiment analysis: A case study of Andalusia during the COVID-19 pandemic. *Sustainability*, 13(7), 3836. <https://doi.org/10.3390/su13073836>
- Frias-Martinez, V., Soto, V., Hohwald, H., & Frias-Martinez, E. (2012, September). Characterizing urban landscapes using geolocated tweets. In *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Conference on Social Computing* (pp. 239-248). IEEE. <https://doi.org/10.1109/PASSAT/SocialCom.2012.112>
- Guo, X., Hu, L., Li, M., & Wang, Y. (2024). Construction of a tourist sentiment analysis model from the perspective of online review data mining: A case study of Zhaoqing

- Qixingyan Scenic Area. *Science and Technology Innovation and Application*, 13, 9-13. <https://doi.org/10.19981/j.CN23-1581/G3.2024.13.003>
- Kar, A. K., & Dwivedi, Y. K. (2020). Theory building with big data-driven research—Moving away from the “What” towards the “Why.” *International Journal of Information Management*, 54, 102205. <https://doi.org/10.1016/j.ijinfomgt.2020.102205>
- Kumar, S., Kar, A. K., & Ilavarasan, P. V. (2021). Applications of text mining in services management: A systematic literature review. *International Journal of Information Management Data Insights*, 1(1), 100008. <https://doi.org/10.1016/j.ijinfomgt.2021.100008>
- Lu, Y. (2024). A study on tourist satisfaction of online OTA platforms based on text analysis: A case study of Ctrip.com. *Journal of Xingtai College*, 2, 83-92. Retrieved from https://kns.cnki.net/kcms2/article/abstract?v=cp4X_KT6XPjDmBHvdUiusKSzrSJ70wYSFUIfTSwfHWQeMzB83yM09JlqPy_7BPi0_F28zghiEReOuhp9-1wIo9p8St-UEgH3gjTYSVMY0zlyOc082NNsYjQH_uMFbZfXCVmaouhcJjg5xxoUUczixup_X_ORBnJ0VFV45lsoMbuTqNpg0LLj_IXFbcLMg2n90xJc25MJQHlXjVSGyEoG6wg_3Vud&uniplatform=NZKPT&language=CHS
- Mahr, D., Stead, S., & Odekerken-Schröder, G. (2019). Making sense of customer service experiences: A text mining review. *Journal of Services Marketing*, 33(1), 88-103. <https://doi.org/10.1108/JSM-10-2018-0295>
- Martín, C. A., Torres, J. M., Aguilar, R. M., & Diaz, S. (2018). Using deep learning to predict sentiments: A case study in tourism. *Complexity*, 2018, Article 7408431. <https://doi.org/10.1155/2018/7408431>
- Ministry of Tourism and Culture Malaysia. (2015, September 21). Achievements and recognitions. Retrieved September 5, 2015, from <http://www.motac.gov.my/en/profile/achievements-and-recognitions>
- Mittal, D., & Agrawal, S. R. (2022). Determining banking service attributes from online reviews: Text mining and sentiment analysis. *International Journal of Bank Marketing*, 40(3), 558-577. <https://doi.org/10.1108/IJBM-08-2021-0380>
- Mohammed, N. A. D. Z. I. R. A. H., Aziz, Y. A., Awang, K. W., & Samdin, Z. A. I. T. O. N. (2018). Influence of destination attractiveness on place satisfaction and environmentally responsible behavior in marine parks. *Journal of Sustainability Science and Management*, 5, 52-67. Retrieved from <https://jssm.umt.edu.my/wp-content/uploads/sites/51/2020/05/5.15.1pdf.pdf>
- Papilaya, R. L. (2018, February). Effect of tourist characteristic, marine tourism demand, and number of visits to the value perceptions and willingness to pay to environmental marine tourism in Ambon City. *IOP Conference Series: Earth and Environmental Science*, 116(1), 012097. <https://doi.org/10.1088/1755-1315/116/1/012097>
- Ren, X., & Zhao, G. (2024). Visitor perceptions of comprehensive parks in the city center of Guangzhou based on social media comments. *Gardening*, 6, 141-150. <https://doi.org/10.12193/j.laing.2024.06.0141.016>
- Sabah Tourism Board. (2019). *Tourism statistics*. Retrieved February 21, 2020, from <https://www.sabahtourism.com/statistics/>
- Su, Y., Xu, J., Sotiriadis, M., & Shen, S. (2021). Authenticity, perceived value and loyalty in marine tourism destinations: The case of Zhoushan, Zhejiang Province, China. *Sustainability*, 13(7), 3716. <https://www.mdpi.com/2071-1050/13/7/3716>
- Sun, X., & Ni, R. (2018). Product awareness, emotional expression, and brand image perception of Chinese cruise tourists: A content analysis based on online reviews. *Geographical Research*, 6, 1159-1180. <https://doi.org/10.11821/dlyj201806008>

- Tongcheng Travel Holdings Limited. (2024). *Summer travel consumer trend report*. Retrieved June 10, 2024, from <http://www.ec100.cn/detail--6640430.html>
- Vetitnev, A., Chigarev, D., & Enushevskaya, V. (2021). Sentiment analysis as a tool for assessing the negative impact of tourism on a destination. In *E3S Web of Conferences*, 311, 06008. EDP Sciences. <https://doi.org/10.1051/e3sconf/202131106008>
- Villarroel Ordenes, F., & Zhang, S. (2019). From words to pixels: Text and image mining methods for service research. *Journal of Service Management*, 30(5), 593-620. <https://doi.org/10.1108/JOSM-08-2019-0254>
- Zhang, W., Tang, S., Chen, L., & Xu, M. (2024). A study on visitor perceptions of cultural heritage ecological environment and its impact on heritage tourism image based on network text analysis: A case study of the Ming City Wall in Nanjing. *Ecological Economy*, 6, 135-140. Retrieved from https://kns.cnki.net/kcms2/article/abstract?v=cp4X_KT6XN9DZQZoW6snP8Fpr4E80Iq2sGIMIDR9CVnHIC_IcluAwTXL5UMZwvV6iREKSrI3OZmkTFrRm1WAsXy05hd1EBi6zoCp02zQN8HajZpwhahD68WgwcghL1xBFvJdCSfVocCxl2SuphfTryw2EWurcj43sOQBG4uWufpLDGtnFJPoFGvS4Ct7ZEAKZINR50KtX2XghTslSa5MMGqKTkXmWv&uniplatform=NZKPT&language=CHS