Article

Exploration of Spa Leisure Consumption Sentiment towards Different Holidays and Different Cities through Online Reviews: Implications for Customer Segmentation

Jianhong Luo *, Shifen Qiu, Xuwei Pan, Ke Yang and Yuanqingqing Tian

School of Economics and Management, Zhejiang Sci-Tech University, Hangzhou 310018, China; Shifen.Qiu@outlook.com (S.Q.); panxw@zstu.edu.cn (X.P.); yangke_keyang@163.com (K.Y.); tyqq947696142@163.com (Y.T.)

* Correspondence: jh_luo@zstu.edu.cn

Abstract: With the improvements in per capita disposable income, and an increase in work-related pressure, demand for leisure consumption such as foot bath spas is constantly increasing. Analysis of leisure consumption sentiment is of great importance for the leisure service industry—to meet customer needs, improve service quality and improve customer relationship management. However, traditional sentiment analysis approaches only aimed to ascertain the overall sentiment of the customer, which is less effective for analyzing customer satisfaction on account of customer size, different customer locations, and different leisure holidays. Sentiment analysis via online reviews can assist different businesses, including foot bath spa services, to better inform the development of customer segmentation strategies and ensure optimal customer relationship management. Hence, the objective of this paper is to explore foot bath spa leisure consumption sentiment towards different holidays and different cities by applying data mining via online reviews, so as to help optimize customer segmentation. A novel general framework and related sentiment analysis methods were proposed and then conducted through a collection of datasets from customers’ textual reviews of foot bath spa merchants in three cities in China on the Meituan social media platform. Findings confirm that the proposed general framework and methods can be used to gain insights into the swing characteristics of sentiment towards different holidays and different cities, to better develop customer segmentation according to the city-holiday emoticon face patterns obtained through sentiment tendency analysis from online social media review data. The study results can help to develop better customer and marketing strategies, thereby creating sustainable competitive advantages, and can be extended to other fields to support sustainable development.

Keywords: sentiment analysis; spa leisure consumption; online reviews; customer segmentation; social media data; customer relationship management

1. Introduction

In recent years, with the improvements in living standards as well as the increase in disposable material wealth, there is an increase in leisure consumption. Furthermore, the increase in work- and life-related stress has led people to take time for leisure activities in order to relieve fatigue and relax. Foot bath spas, including foot baths, tuina, and massage spas, are an excellent way to relax the body and mind, improve sleep quality, and reduce tiredness. Foot bath spas are also a popular spa project among Chinese people, especially those who are under pressure from work and life. As a significant sector of leisure consumption, the spa industry has attracted great attention. Although the spa industry was hit hard during the COVID-19 pandemic, a report by the Global Wellness Institute indicates that the industry is expected to recover fast, with the market hypothesized to grow by 17% annually through 2025 [1]. Faced with such promising growth potential, spa businesses need to seize the opportunity to develop appropriate marketing strategies to promote...
sustainable development. Kauffmann et al. [2] indicate that the long-term sustainability of enterprises depends, to a great extent, on their ability to properly meet consumer demands.

Consumer reviews play an invaluable role for businesses in obtaining consumer feedback to facilitate marketing campaigns [3]. In the past, organizations collected public or customer opinions through surveys, opinion polls, and focus groups [4]. With the increasing use of social media, a large number of people are willing to share their views and opinions online, which has altered customer relationship management and customer behavior analysis [5]. The explosive growth of online reviews provides promising alternative data sources to identify customer needs [6,7]. Thereby, spa businesses can effectively monitor customers through online reviews. However, analyzing the textual content such as online reviews and microblog posts is challenging because they are written in natural language [8]. As one of the common data analytic technologies, sentiment analysis is often used to analyze textual data on the Internet. Extracting customer sentiment from online reviews can help businesses obtain customer attitudes and feedback on services or products.

There are numerous previous studies [9–13] related to sentiment analysis of online reviews. However, most of these studies divided sentiment into only two or three categories (neutral, negative, or positive), which may not reflect emotional intensity well. Moreover, their research focused on the overall or aspect sentiment of the whole period or all the customers, and did not consider dynamic change in sentiment at different times and in different regions. Furthermore, even for the same product, service, or topic, sentiment tendency and focus may be different at different time points. Some researchers [3,14,15] have considered the time attribute; for example, Ibrahim and Wang [16] combined sentiment and time series analyses to identify the critical time points that lead to significant deviations in sentiment trends of five online retailers, so as to provide a deeper understanding of online retail customer behavior. However, few studies consider the effects of special time points, such as festivals and holidays, on customer sentiment and make quantitative analyses. Customers will show different consumption psychology and demand during different leisure holidays. Correctly grasping public festival sentiment is conducive to better cultural development and economic decisions [17].

Therefore, it is necessary to analyze leisure consumption sentiment to better understand customer satisfaction, so as to help leisure businesses segment customers and improve customer relationship management. Thus, the following research questions (RQ) were posed:

RQ1 How can we extract customer leisure consumption sentiment through online review mining?

RQ2 Are there differences in customer sentiment towards different cities and different holidays, and if so, how do we segment customers according to sentiment characteristics?

RQ3 What kind of customer and market management insights can be obtained through leisure consumption sentiment analysis to support sustainable development?

To address the above research questions, from the perspective of foot bath spa businesses, this study proposed a novel general framework and related sentiment analysis methods to explore leisure consumption sentiment towards different holidays and different cities, in which online review data from the Meituan social media platform are utilized to find swing characteristics of sentiment to support developing a new customer segmentation strategy for customers. In particular, we developed a daily sentiment score algorithm for each city to carry out holiday sentiment tendency prediction in order to identify customer segments of foot bath spa businesses so as to provide novel insights for spa merchants to improve customer and market management.

Our results will assist in identifying the consumption sentiment tendency of customers who consume spa services so as to inform customer segmentation and marketing strategy. In addition, this study provides a new approach for theoretical and practical research of customer segmentation. The results of this research will benefit business managers in the leisure industry, as well as leisure service vendors, to better understand customer
sentiment and satisfaction and evaluative assessments of different customer segments to ensure targeted and efficient marketing strategy.

The remainder of this paper is organized as follows. In Section 2, the related work is introduced. Section 3 illustrates the proposed framework and methods. Section 4 presents the analysis results of foot bath spa sentiment. Section 5 shows the discussion. Finally, the conclusions of this paper are presented in Section 6.

2. Related Work

2.1. Sentiment Analysis in Customer Relationship Management

Recently, there has been a significant rethinking of business management, which is increasingly based on customer-oriented principles [18]. Customer relationship management (CRM) is a technology-based approach designed to improve the management of companies’ interactions with customers [18]. The main goals of customer relationship management are to improve customer satisfaction, reduce customer churn rate, and increase customer quality rate.

In the web 2.0 era, consumers’ willingness to share consumption experiences online, coupled with the application of data mining technologies, offers marketers an unprecedented opportunity to collect market intelligence [19]. Social media and big data have revolutionized customer relationship management by communicating directly with customers [5]. Specifically, big data technologies can be used to efficiently extract useful information from user-generated content in online environments. Sentiment analysis is one of the most data mining techniques employed to extract opinions from the web, which focuses on determining customer emotional polarity and customer attitudes on a specific topic [20].

Sentiment analysis is often used to identify customer needs to improve customer satisfaction. Regarding services, customer satisfaction could be defined as an overall customer attitude toward a service provider or as the cognitive and affective response to a single or a long-term service encounters [21]. For example, Gitto and Mancuso [11] used sentiment analysis to distinguish passenger sentiment polarity for the non-aviation services and aviation services, the results show that passengers are more satisfied with non-aviation services, so they suggested that airlines focus on aviation services to improve customer satisfaction. Customer segmentation is an important way to better understand customers in an era of increasingly fierce competition. Many market segmentation methods classify consumers according to criteria such as socio-economic, demographic, and psychographic factors, but previous studies have shown that these segmentation bases are of limited value in investigating consumer behavior, given that other factors influence customer behavior in service settings, such as customers’ emotions and affect [22]. In this vein, given that emotions differ among different customers, the conceptualization of using emotion as a segmentation variable has received considerable theoretical support; in turn, emotions could be used as the basis for market segmentation [23]. Public sentiment in social media is increasingly seen as an important tool for market understanding and customer segmentation [24]. However, in service industries such as spa services, the use of emotion as a segmentation criteria requires further empirical research.

Therefore, this paper started with the sentiment analysis of online reviews to find sentiment fluctuations, thereby providing valuable implications for customer relationship management and customer segmentation.

2.2. The Basic Methods of Sentiment Analysis

At present, there are mainly two kinds of methods to measure sentiment, including lexicon-based methods and machine learning methods [25].

Sentiment analysis based on machine learning is essentially the task of text categorization [26]. That is to train a sentiment classifier based on a large number of labeled data and then judge sentiment tendency through this classifier. The common classification of sentiment is the dichotomous classification, which means sentiment tendency will be
divided into positive and negative. In 2002, Pang et al. [27] took the lead in employing machine learning algorithms: SVM, Naive Bayes, and maximum entropy classification, to classify movie reviews into positive sentiment or negative sentiment. The results show that SVM performed better. Since then, more and more research [28–30] has been carried out around sentiment classification based on machine learning. Recently, methods based on neural networks [31–33] have been used for sentiment classification. The above methods have achieved good results in sentiment classification, but they cannot reflect the emotional intensity well.

The lexicon-based sentiment analysis method is an unsupervised classification method. It determines the sentiment tendency of the text by calculating text sentiment score with rules [26]. The key to this method is the selection and construction of sentiment dictionaries. Currently, in the sentiment analysis of Chinese texts, the commonly used sentiment dictionaries include the National Taiwan University Semantic Dictionary (NTUSD) [34], the HowNet lexicon [35], and the Chinese Affective Lexicon Ontology [36]. Ding et al. [37] used Chinese Affective Lexicon Ontology to calculate the sentiment scores of tourist attractions through online reviews and then ranked them, and the results showed that the ranking results were highly consistent with the ranking on Ly.com. Wu and Wang [38] constructed an extended sentiment dictionary suitable for shopping areas based on NTUSD and HowNet lexicon. Xu et al. [26] improved sentiment classification performance by adding popular Internet sentiment words, fields, and polysemic sentiment words to the NTUSD and HowNet lexicon. Their experimental results indicated that the dictionary with added field sentiment words showed a higher F1-score. It is obvious that field sentiment words can improve the performance of sentiment analysis. Since there is little research on leisure sentiment analysis with the leisure-related field sentiment dictionary, it is meaningful to explore the sentiment in the leisure field in a fine-grained manner.

3. Materials and Methods

The proposed sentiment analysis framework is shown in Figure 1.

![Figure 1](image-url)
3.1. Data Collection and Preprocessing

3.1.1. Data Collection

Data were collected from Meituan (see www.meituan.com, accessed on 17 November 2021). As one of the most popular life service e-commerce social media platforms in China, it encourages customers to share consumption experiences online, which means that these comments serve the purpose of our research. Three representative cities in China were taken as the research objects: Guangzhou City, Wuhan City, and Shanghai City. We chose them for the following reasons: First, these three cities belong to regions with advanced economies, and the demand for leisure consumption is relatively high; second, the geographical distribution of these three cities is very different. Guangzhou is in South China, Shanghai is in East China, and Wuhan is in Central China. Compared with choosing cities located in the same area, this may better reflect regional differences.

We extracted online reviews of the top 10 foot bath spa merchants with the number of reviews ranked on top-ten in each of these three cities from 2014 to 2019, and a total of 80,331 reviews were obtained. Specifically, we collected user information and user-generated content, including username, review text, review date, consumption items, etc. Figure 2 shows an example of online review information on Meituan (we have translated all Chinese into English). Furthermore, the total number of reviews and average rating score of customers (out of 5 points) for each city are shown in Table 1.

![Figure 2. An example of online review information on Meituan.](image)

Table 1. The summary information of foot bath spa merchants in three cities.

| City     | Total Number of Reviews | Average Rating Score |
|----------|-------------------------|----------------------|
| Guangzhou| 12,589                  | 4.31                 |
| Shanghai | 8410                    | 4.23                 |
| Wuhan    | 59,332                  | 4.89                 |

3.1.2. Chinese Word Segmentation

As we all know, in English, words are separated by spaces [37]. Unlike English, there are no obvious separators between Chinese words. Therefore, the first step of Chinese processing task is to identify the sequence of words in the sentence and mark the boundary at the appropriate position [39].

In this research, the Chinese word segmentation module “Jieba” [40] of the R programming language was used to perform text mining and word segmentation processing. Meanwhile, to ensure the accuracy of word segmentation results, we used the Chinese corpus of sougou input method [41] as the word segmentation dictionary and manually added leisure-related custom words to the corpus that often appear in the reviews but not in the corpus. In other words, this research supplemented and extended the word segmentation dictionary in the leisure consumption field and then carried out word segmentation based on this extended dictionary.

3.1.3. Data Preprocessing

In the real world, data are usually incomplete, inconsistent, and easily susceptible to noise interference. Therefore, it is necessary to preprocess the collected data to ensure
the validity of the experimental data. There are many methods for data preprocessing, including data cleaning, data integration, data transformation, and data reduction [37].

For the dataset of customer reviews, we mainly used data cleaning to deal with missing values, duplicate values, and invalid values. First, the review records with less comment text or unrelated with leisure consumption were abandoned. Second, duplicate comments and meaningless templated comments were removed, which hardly contain any emotional information and may even interfere with the results. Then, we can obtain the preprocessed dataset.

3.2. Developing a Daily Sentiment Score Algorithm

To uncover the variation characteristics of leisure consumption sentiment, a daily sentiment score algorithm for each city was proposed to prepare for holiday sentiment exploration.

3.2.1. Extracting Emotion Words

This research utilized the Chinese affective lexicon ontology library developed by Professor Lin Hongfei of Dalian University of Technology [36] to extract emotion words. This resource describes Chinese words or phrases from multiple dimensions, including part of speech, emotion category, emotion intensity, and polarity. It contains 27,466 affective words. We extracted all emotion words from the preprocessed dataset based on this dictionary.

3.2.2. Determining the Weight of Emotion Words

After extracting emotion words, we need to assign weight to them according to their features. The determination of the weight of emotion words can be divided into the following three steps.

1. Setting the Initial Weight Value

The first step is to set the initial weight value of the emotion words. In this paper, we only chose two dimensions, emotion intensity and polarity, to help obtain the initial weight value of emotion words. Emotion intensity refers to the strength of feeling reflected by emotion words, divided into five levels: 1, 3, 5, 7, and 9; among them, “9” represents the highest intensity. Each emotion word also corresponds to a polarity. Polarity has three values: 0, 1, and 2, which represent neutral, positive, and negative, respectively [36]. To make the final sentiment scores easier to distinguish, we reassigned the value to polarity, where “−1” means negative and “+1” means positive. Then, we defined a simple formula to calculate the Initial Weight Value $W_0^i$ of emotion words:

$$W_0^i = I_{word_i} \times P_{word_i}$$

(1)

where $I_{word_i}$, $P_{word_i}$ represent the emotion intensity and the polarity, respectively, of $i$-th word.

2. Computing the Word-Frequency Weight Value

The second step is to calculate the word-frequency weight value of emotion words. The number of repetitive emotion words in reviews is directly related to their significance. Thus, the word frequency needs to be considered in the calculation rules of the sentiment score to make the weight value calculation more reasonable. The calculation formula of Word-Frequency Weight Value $W_1^i$ is as follows:

$$W_1^i = \frac{N_i - N_{min}}{N_{max} - N_{min}}$$

(2)

where $N_i$ represents the number of times of $i$-th emotion word appeared in the preprocessed dataset, $N_{min}$ and $N_{max}$ represent the minimum and maximum number of appearance times of the emotion word in the preprocessed dataset, respectively.
3. Adding Extract Weight value

The third step is to add the extra weight value to emotion words in the leisure field. As mentioned in Section 2.2, field sentiment words can improve the performance of sentiment analysis. When calculating sentiment scores in a certain area through emotion words, higher weight should be given to the emotion words of the field. According to the existing research [26,37], we selected emotion words in the leisure field from the reviews on foot bath spas and then gave Extra Weight Value $W^2_i$ to the emotion words. The specific values are shown as follows:

$$W^2_i = \begin{cases} 0.2, & \text{if word } i \in \text{emotion words of leisure field} \\ 0, & \text{otherwise} \end{cases}$$ (3)

Through the above three steps, we obtain the Final Weight $W_i$ of the emotion words. The formula can be expressed as follows:

$$W_i = W_0^i \times \left( 1 + W^1_i + W^2_i \right)$$ (4)

3.2.3. Calculation of Daily Sentiment Score

To perform a more fine-grained analysis of temporal sentiment, we calculated the daily sentiment score of each city. The daily sentiment score calculation can be divided into the following two steps.

1. Defining Initial Sentiment Score

The Initial Sentiment Score $S^0_j$ in this paper is defined as the ratio of the total sentiment score of positive words in a certain city on a day to the sum of the absolute sentiment score of all emotion words on that day. This value can directly reflect the positive or negative tendency of the leisure consumption sentiment in a city. The calculation formula is as follows:

$$S^0_j = \frac{\sum_{i=1}^{n_1} (x^1_i \times W_i)}{\sum_{i=1}^{n_1+n_2} (x^i_i \times |W_i|)}$$ (5)

where $x^1_i$ represents the frequency of $i$-th positive word, $x^i_i$ represents the frequency of $i$-th emotion word, and $n_1$ and $n_2$ represent the total number of positive words and negative words, respectively, on the $j$-th day of a city.

2. Standardizing Sentiment Score

Data standardization is to scale data proportionally to fall within a specific interval, thereby eliminating the adverse effects of data dimensions. Therefore, this study standardized the Initial Sentiment Score $S^0_j$.

By calculating the mean and standard deviation of the initial sentiment score of a certain city over a period of time, we define a formula to calculate the standardized daily sentiment score $S_j$ of each city. The formula is as follows:

$$S_j = \frac{S^0_j - E}{D}$$ (6)

$$E = \frac{\sum_{j=1}^{d} S^0_j}{d}$$ (7)

$$D = \sqrt{\frac{\sum_{j=1}^{d} (S^0_j - E)^2}{d}}$$ (8)

Among them, $d$ represents the total number of days in the sample data, and $E$ and $D$ represent the mean and the standard deviation of the initial sentiment score, respectively, on the $j$-th day of a city.
After the above calculation, we can obtain the daily sentiment scores of each city and then sort them in chronological order to construct the Daily Sentiment Score Time Series (DSTS). Each city corresponds to a DSTS. The formula is as follows:

$$\text{DSTS}_k = (T_j, S_{j,k})$$

(9)

where \( k = 1, 2, \ldots, n \), represents each different city, \( T_j \) is a set of dates, the format of each element is “yyyy-mm-dd”, and \( S_{j,k} \) represents the standardized daily sentiment score of city \( k \) on date \( T_j \).

3.3. Sentiment Variation Exploration towards Different Holidays and Different Cities

In practical research, holidays or some major social events have significant impacts on consumption, such as shopping festivals, and sales data will be greatly improved compared to usual. Traditional time series models such as ARMA cannot accurately predict such data mutation points, so the prediction effects of these models may not be appropriate. However, Prophet can effectively forecast the holiday effects, and has the advantage of not needing to interpolate missing values [42]. Since there are no online reviews on some days, missing values are common in the daily sentiment score data. Therefore, this paper chose the powerful Prophet Model for daily sentiment score prediction analysis.

The time series analysis model Prophet was originally designed to smooth and forecast daily business data and is available as open source software in Python and R [42]. Prophet is an additive model consisting of four components. It uses a Bayesian-based curve fitting method to smooth and predict time series data [43]. The specification of the Prophet Model is similar to a generalized additive model (GAM), including trends, seasonality, holidays, and error terms:

$$y(t) = g(t) + s(t) + h(t) + \epsilon$$

(10)

where \( y(t) \) is the observed value of the time series at time \( t \), \( g(t) \) is the trend function, which represents the non-periodic changes of the time series value, \( s(t) \) represents the periodic changes, \( h(t) \) represents the holiday effects, and \( \epsilon \) represents irregular fluctuations. In the current study, we only described the holiday component \( h(t) \) in detail.

The holiday model \( h(t) \) of Prophet assumes that the effects of holidays are independent, and we can set a window of days around a particular holiday. In other words, it allows us to consider the pre-holiday and post-holiday effects. For example, Christmas Day is December 25th, and we can set parameters of “Lower_window” and “Upper_window” (e.g., Lower_window = −3, Upper_window = 4), which represent three days before Christmas and four days after Christmas, respectively. These eight days will be considered as a holiday period in the analysis.

For each holiday \( i \), let \( D_i \) be the set of past and future dates for that holiday. We added an indicator function to indicate whether time \( t \) is on holiday \( i \), and the value of this function is mainly determined by the parameter \( \kappa_i \); \( \kappa_i \) represents the degree of influence of holiday \( i \) on daily sentiment score. Supposing the number of holidays is \( L \), then we can obtain:

$$h(t) = \sum_{i=1}^{L} \kappa_i 1(i \in D_i)$$

(11)

In order to make the mathematical expression of the holiday component consistent with the seasonality component, the influence degree \( \kappa_i \) of each holiday is combined into a \( 1 \times L \) matrix, and the transpose of this matrix is denoted as \( \kappa \). Then, a \( 1 \times L \) matrix \( Z(t) \), a characteristic matrix about holidays and user-defined time windows, is introduced. Thereby, the holiday model can be expressed as the multiplication of two matrices, as shown in Equation (12).

$$h(t) = Z(t)\kappa$$

(12)
where \( Z(t) = [1(t \in D_1), \ldots, 1(t \in D_L)] , \kappa = [k_1, \ldots, k_L]^T , \kappa \sim \text{Normal} (0, \sigma^2) \). The parameter \( \sigma \) is used to control the flexibility of holiday model, and the default value is 10.

After establishing the Prophet Model, we need to input the DSTS data of each city into the model for fitting and forecasting. An estimated daily sentiment score is the sum of values of the trend, yearly, weekly, and holiday components values at a given time. The variations of daily sentiment score caused by each component can be decomposed separately by the function “plot components”. In this paper, we only focused on the variations of daily sentiment score of each city towards different holidays.

4. Results

4.1. High-Frequency Emotion Words

The word cloud visualization is an excellent way to communicate the findings [44]. Hence, in this paper, a word cloud was used to roughly reflect the positive and negative tendency of leisure consumption sentiment. The polarity of emotion words is the basis for judging whether an emotion word is positive or negative. According to Section 3.2.2, “–1” means negative and “+1” means positive.

After word segmentation and emotion words extraction, we counted the word frequency of emotion words and then ranked them. Table 2 shows the examples of negative and positive words in the top 50 emotion words based on word frequency ranking. Due to some words having the same word frequency, their rankings are the same. For instance, “优雅 (grace)” and “支持 (support)” have the same word frequency of 388, so they are both ranked 23rd. We obtained a total of 56 words and then drew the emotion word cloud, as shown in Figure 3. In the figure, black indicates negative words while white indicates positive words. Most high-frequency words are positive, and the proportion of negative words is relatively small, which means customers in these three cities are generally satisfied with their spa experience during the data collection period, so their sentiment tendencies are overall active. Online ratings are often regarded as proxies of customer satisfaction [45, 46], so this result is also in line with the average rating score in Table 1, where the average rating score is above 4.2.

### Table 2. Examples of positive and negative words in the top 50 emotion words.

| Polarity  | Examples (Order by Weight)                                      | Number |
|----------|------------------------------------------------------------------|--------|
| Negative | 不好 (bad) · 不行 (useless) · 不怎么样 (not up to much) · 随便 (casual) · 无语 (speechless) · 敷衍 (perfunctory) · 稀烂 (very poor) ... | 8      |
| Positive | 不错 (good) · 值得 (worth) · 热情 (enthusiasm) · 满意 (satisfaction) · 朋友 (friends) · 喜欢 (like) · 好评 (acclaim) ... | 48     |

Figure 3. Emotion word cloud.
4.2. The Variation in Leisure Consumption Sentiment towards Different Holidays and Different Cities

Based on the daily sentiment score algorithm proposed in Section 3.2, we obtained the daily sentiment scores of each city and constructed three DSTS, pointing to Guangzhou, Shanghai, and Wuhan, respectively. Then, we analyzed the time series data through Prophet Model and decomposed the holiday component.

At first, New Year’s Day, Spring Festival, Valentine’s Day, Tomb-Sweeping Day, Christmas Day, and other holidays were chosen for analysis. The time window was set according to the statutory leave of festival or festival time, and the pre-holiday and post-holiday was set to 10 days. After analyzing and filtering, four holidays with significant emotional fluctuation were left: Spring Festival, Dragon Boat Festival, National Day, and Christmas Day. Figure 4a–c show the characteristics of consumption sentiment variations in four festivals in three different cities: The y-axis is the predicted value of daily sentiment score caused by holiday effects, and the x-axis represents the different holidays. The negative daily sentiment score indicates that consumption sentiment is negative; otherwise, it is the opposite. The greater the absolute value of y, the more significant the sentiment holiday effects. For a more intuitive display, gray color blocks were added to holidays with the most apparent sentiment fluctuations. From Figure 4, we can easily observe different sentiment fluctuations for each holiday and each city.

![Figure 4](image_url)

**Figure 4.** The variations in consumption sentiment in different holidays and different cities.

There are many significant characteristics. Firstly, from the perspective of festivals, consumption sentiment shows negative fluctuations during the Spring Festival and Dragon Boat Festival, but the negative fluctuations during the Spring Festival are even greater. This may be related to the lack of massage technicians during the Spring Festival and
the lack of attention to festival customs. Consumption sentiment on National Day and Christmas shows positive fluctuations. This may be because the National Day holiday is sufficient, giving spa customers more choices, and as China gradually moves towards internationalization, spa merchants are paying more and more attention to Christmas.

Secondly, from the perspective of cities, consumption sentiment in Guangzhou and Wuhan shows obvious negative fluctuations during the Spring Festival. This may be related to the shortage of employees caused by many migrant workers returning to their hometowns on vacation. In Shanghai, consumption sentiment shows significant positive fluctuations during the Christmas period. As an international metropolis, Shanghai’s spa businesses may pay more attention to Christmas than other cities.

### 4.3. Customer Segmentation Based on Consumption Sentiment Tendency Prediction

To show the difference more intuitively, based on the above line chart, we innovatively adopted emoticon face graphs to show the consumption sentiment tendencies of different holidays and different cities, as shown in Figure 5.

![Figure 5. The emoticon face graph of consumption sentiment tendencies.](image)

The emoticon face is generated and marked according to the predicted effect values of consumption sentiment. Three steps can be conducted: First, compute the mean of the predicted effect values of each holiday in each city; secondly, select the maximum and minimum values of the mean and then divide the interval; and thirdly, match the emoticon face for each interval. The detailed description is shown in Table 3.

| ID | Emoticon Face | Meaning | Interval | Sentiment Tendency |
|----|---------------|---------|----------|--------------------|
| S1 | 😊             | Very happy | [0.4,0.5] | Highly positive    |
| S2 | 😊             | Slightly happy | [0.1,0.4] | Slightly positive  |
| S3 | 😊             | Neutral | [−0.05,0.1] | Neutral           |
| S4 | 😊             | Slightly sad | [−0.1,−0.05] | Slightly negative |
| S5 | 😊             | Very sad | [−0.3,−0.1] | Highly negative   |

According to the statistical data of the Local Bureau of Statistics, the per capita disposable income of urban residents in Guangzhou [47], Shanghai [48], and Wuhan [49] were CNY 65,052, CNY 76,437, and CNY 51,706, respectively in 2019, which were higher than the national average [50] of CNY 42,359. It means that the residents of these three cities have the financial ability to engage in leisure activities. However, from Figure 5, we can observe that in these three cities, the proportion of non-smiling faces (10/12, 10 and 12 represent the number of non-smiling faces and the total number of emoticons in Figure 5, respectively) is significantly higher than that of smiling faces (2/12). This shows that the customers are not very satisfied with the spa services they received during the holidays.
The high average rating score in Table 1 and the majority of positive words shown in Figure 3 informed good customer satisfaction. However, comparing the emoticon face graphs of consumption sentiment tendencies in Figure 5, it is implied that customer satisfaction is varied in different holidays and different cities. The overall customer sentiment obtained by traditional methods cannot effectively obtain the fine-grained differences in customer satisfaction, resulting in insufficient understanding of customers and the market.

Therefore, it is necessary to segment customers to provide more targeted services. Therefore, we proposed the segmentation rules based on the city-holiday emoticon face patterns, where each city-holiday emoticon face represents a customer group with a kind of sentiment tendency shown in Table 3 during a city’s holiday. As a result, we can segment spa customers into five segments according to the city-holiday emoticon face patterns, as shown in Figure 6. For instance, Segment S1 represents customers whose consumption sentiment shows a highly positive tendency.

![Customer segmentation based on city-holiday emoticon face patterns.](image)

**Figure 6.** Customer segmentation based on city-holiday emoticon face patterns.

From the above results, it can be seen that the novel general framework and related sentiment analysis methods proposed in this paper can effectively understand the fine-grained consumption sentiment, assist in optimizing customer segmentation, and enable leisure service vendors to formulate more precise marketing strategies. For example, for Segment S5, spa merchants in Guangzhou and Wuhan should pay attention to the negative tendency of customer sentiment during the Spring Festival, ensure that there are sufficiently skilled massage technicians on duty, and provide festive services, such as gifting customers with New Year’s Zodiac dolls. For Segment S2, it is suitable for spa merchants in Shanghai to pay attention to the festivals with national attributes while improving the service level during the Christmas period.

5. **Discussion**

5.1. **Research Implications**

This research has made significant theoretical, methodological, and practical contributions. From a theoretical viewpoint, this paper presents an effort to improve customer segmentation and customer relationship management by applying sentiment analysis via online reviews. Our findings reveal that customers’ consumption sentiment exists significant differences towards different cities and different holidays. As shown in the study of Calvo-Porral and Lévy-Mangin [22], the emotion-based segmentation is meaningful, and the service customers cannot be perceived as a homogenous group. In this study, the exploration of customer sentiments confirmed that customers could experience different emotions for spa products or services at different times and in different regions, and presented that emotions experienced during service use can be used to develop service customer segmentation. The findings will help to identify the customer satisfaction through the city-holiday emoticon face patterns and provide information for customer segmentation.
and marketing strategies. The results of this research provide a new customer segmentation strategy for leisure businesses to assist in offering appropriate services for each identified segment, improve customer satisfaction, and promote the reasonable allocation of marketing resources.

From a methodological viewpoint, this research proposed a general framework and related sentiment analysis methods for sentiment exploration towards different cities and different holidays. Instead of judging sentiment orientation of the whole period for all the customers, sentiment analysis in this paper was held by computing daily sentiment score, which is not only at a more fine-grained level (showing the dynamic changes of customer sentiment during holidays) but also gives the intensity degree of emotion (specific values). In addition, the proposed sentiment analysis methods provide a way to obtain the city-holiday customer consumption emotion and understand customer satisfaction and also provide a way for customer segmentation based on sentiment analysis. As López-Chau [51] pointed out, sentiment analysis helps organizations discover sentiment and improve customer relationship management.

From a practical viewpoint, this research will benefit business managers in the leisure industry as well as leisure services vendors to better understand customer sentiment and satisfaction and evaluate different customer segments to ensure targeted and efficient marketing strategy. First, we discovered four festivals where customer sentiment fluctuates significantly: the Spring Festival, the Dragon Boat Festival, National Day, and Christmas. Compared with other festivals, these four festivals deserve more attention from leisure businesses. Second, we found that the swing characteristics of customer sentiment in different cities and different holidays are different. The variations in consumption sentiment during holidays are related to the quality of service provided by merchants at different times and in different regions. In addition, we segmented customers according to their city-holiday emoticon face patterns and divided them into five categories so that businesses can provide more targeted services to different groups. Segmentation allows marketers to better tailor their marketing efforts to various customer subsets [52]. Thereby, segmentation is vital for spa companies because, as far as we know, many foot bath spa merchants in China exist in the form of chain stores. Market differences should be taken seriously, and their marketing strategy needs to be adjusted in light of time and local conditions.

5.2. Limitations and Future Work

It should be pointed out that this paper still has some limitations. First of all, we only collected online reviews of foot bath spa merchants in three major cities in China, and discovered the swing characteristics of foot bath spa leisure consumption sentiment towards different holidays and different cities. Further studies need to consider more cities and merchants to give more suggestions on creating competitive advantages and extend to other fields, such as the catering industry. Secondly, this study segmented customers according to sentiment tendencies. Future work can consider more attributes to achieve finer customer segmentation, so as to better meet the diverse demands of customers. Thirdly, this research mainly focused on the spa consumption experiences of Meituan users through the online reviews. Future studies could collect opinions of other spa customers through interviews, focus groups, or surveys. Finally, despite the COVID-19 pandemic affecting leisure businesses hard in recent years, the industry is expected to recover fast, and the proposed framework and methods in this paper are also applicable to sentiment analysis and customer segmentation for this context.

6. Conclusions

This research demonstrates how leisure businesses can more effectively understand the leisure consumption sentiment and customer segmentation. A novel general framework and related sentiment analysis methods were proposed and then analyzed using a collection of datasets from customers’ textual reviews of foot bath spa merchants in three cities in China. The results confirm that the proposed general framework and methods can be
used to gain insights into the swing characteristics of sentiment towards different cities and different holidays. Accordingly, five customer segments were established via the city-holiday emoticon face patterns obtained through sentiment tendency analysis. The results demonstrate that the proposed sentiment analysis techniques are effective in spa customer segmentation by using reviews data from social media platforms. The findings can help to provide implications for spa merchants, even leisure businesses, develop differentiated marketing strategies for segments and improve customer relationship management.

The proposed approaches and the exploration results of leisure consumption sentiment towards different cities and different holidays provide novel insights for businesses to understand customer satisfaction and conduct customer segmentation according to customer sentiment. It can be extended to other fields or online platforms as well.

Author Contributions: Conceptualization, J.L. and X.P.; methodology, J.L.; formal analysis, J.L. and K.Y.; writing—original draft, S.Q. and Y.T.; writing—review and editing, J.L. and S.Q.; project administration, funding acquisition, J.L. and X.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Zhejiang Provincial Natural Science Foundation of China (Grant No.LZ22G010001; No.LZ18G010001), and the Fundamental Research Funds of Zhejiang Sci-Tech University (Grant No.2021Q075).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used to support the findings of this study are available from the corresponding author upon request.

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

References

1. GWI. Global Wellness Economy: Looking Beyond COVID: 2021. Available online: https://globalwellnessinstitute.org/industry-research/the-global-wellness-economy-looking-beyond-covid/ (accessed on 17 November 2021).

2. Kauffmann, E.; Peral, J.; Gil, D.; Ferrández, A.; Sellers, R.; Mora, H. Managing Marketing Decision-Making with Sentiment Analysis: An Evaluation of the Main Product Features Using Text Data Mining. *Sustainability* 2019, 11, 4235. [CrossRef]

3. Ye, W.-J.; Lee, A.J.T. Mining sentiment tendencies and summaries from consumer reviews. *Inf. Syst. E-Bus. Manag.* 2021, 19, 107–135. [CrossRef]

4. Liu, B. Sentiment Analysis and Opinion Mining. *Synth. Lect. Hum. Lang. Technol.* 2012, 5, 1–12. [CrossRef]

5. Ahani, A.; Nilashi, M.; Ibrahim, O.; Sanzogni, L.; Weaven, S. Market segmentation and travel choice prediction in Spa hotels through TripAdvisor’s online reviews. *Int. J. Hosp. Manag.* 2019, 80, 52–77. [CrossRef]

6. Timoshenko, A.; Hauser, J.R. Identifying Customer Needs from User-Generated Content. *Mark. Sci.* 2019, 38, 1–20. [CrossRef]

7. Vriens, M.; Chen, S.; Vidden, C. Mapping brand similarities: Comparing consumer online comments versus survey data. *Int. J. Mark. Res.* 2019, 61, 130–139. [CrossRef]

8. Luo, J.; Pan, X.; Wang, S.; Huang, Y. Identifying target audience on enterprise social network. *Ind. Manag. Data Syst.* 2019, 119, 111–128. [CrossRef]

9. Miyoshi, T.; Nakagami, Y. Sentiment Classification of Customer Reviews on Electric Products. In Proceedings of the 2007 IEEE International Conference on Systems, Man and Cybernetics, Montreal, QC, Canada, 7–10 October 2007; pp. 2028–2033.

10. Nagamma, P.; Pruthvi, H.R.; Nisha, K.K.; Shwetha, N.H. An improved sentiment analysis of online movie reviews based on clustering for box-office prediction. In Proceedings of the Communication Automation International Conference on Computing, Noida, India, 15–16 May 2015; pp. 933–937.

11. Gitto, S.; Mancuso, P. Improving airport services using sentiment analysis of the websites. *Tour. Manag. Perspect.* 2017, 22, 132–136. [CrossRef]

12. Rasool, G.; Pathania, A. Reading between the lines: Untwining online user-generated content using sentiment analysis. *J. Res. Interact. Mark.* 2021, 15, 401–418. [CrossRef]

13. Ren, G.; Hong, T. Investigating Online Destination Images Using a Topic-Based Sentiment Analysis Approach. *Sustainability* 2017, 9, 1765. [CrossRef]

14. AL-Sharuee, M.T.; Liu, F.; Pratama, M. Sentiment analysis: Dynamic and temporal clustering of product reviews. *Appl. Intell.* 2021, 51, 51–70. [CrossRef]

15. Tsytarsaru, M.; Palpanas, T.; Castellanos, M. Dynamics of news events and social media reaction. In Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA, 24 August 2014; pp. 901–910.
16. Ibrahim, N.F.; Wang, X. Decoding the sentiment dynamics of online retailing customers: Time series analysis of social media. *Comput. Hum. Behav.* 2019, 96, 32–45. [CrossRef]
17. Yuan, H.; Song, Y.; Hu, J.; Ma, Y. Design of Festival Sentiment Classifier Based on Social Network. *Comput. Intell. Neurosci.* 2020, 2020, 8824009. [CrossRef] [PubMed]
18. Capuano, N.; Greco, L.; Ritrovato, P.; Vento, M. Sentiment analysis for customer relationship management: An incremental learning approach. *Appl. Intel. 2021*, 51, 3339–3352. [CrossRef]
19. Ramboças, M.; Pacheco, B.G. Online Sentiment Analysis in Marketing Research: A Review. *J. Res. Interact. Mark.* 2018, 12, 146–163. [CrossRef]
20. Wang, L.; Wan, Y. Sentiment Classification of Documents Based on Latent Semantic Analysis. In Proceedings of the Sentiment Classification of Documents Based on Latent Semantic Analysis, Wuhan, China, 18–19 June 2011; pp. 356–361.
21. Levesque, T.; McDougall, G.H.G. Determinants of Customer Satisfaction in Retail Banking. *Inter. J. Bank Mark.* 1996, 14, 12–20. [CrossRef]
22. Calvo-Porral, C.; Lévy-Mangin, J.-P. An Emotion-Based Segmentation of Bank Service Customers. *Inter. J. Bank Mark.* 2020, 38, 1441–1463. [CrossRef]
23. Barrena, R.; Sánchez, M. Using Emotional Benefits as a Differentiation Strategy in Saturated Markets. *Psychol. Mark.* 2009, 26, 1002–1030. [CrossRef]
24. Sharef, N.; Mat Zin, H.; Nadali, S. Overview and Future Opportunities of Sentiment Analysis Approaches for Big Data. *J. Comput. Sci.* 2016, 12, 153–168. [CrossRef]
25. Gonçalves, P.; Araújo, M.; Benevenuto, F.; Cha, M. Comparing and combining sentiment analysis methods. In Proceedings of the First ACM Conference on Online Social Networks, New York, NY, USA, 7–8 October 2013; pp. 27–38.
26. Xu, G.; Yu, Z.; Yao, H.; Li, F.; Meng, Y.; Wu, X. Chinese Text Sentiment Analysis Based on Extended Sentiment Dictionary. *IEEE Access 2019*, 7, 43749–43762. [CrossRef]
27. Pang, B.; Lee, L.; Vaithyanathan, S. Thumbs up? Sentiment Classification Using Machine Learning Techniques. In Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing, EMNLP 2002, Philadelphia, PA, USA, 6–7 July 2002; pp. 79–86.
28. Chaovilat, P.; Zhou, L. Movie Review Mining: A Comparison between Supervised and Unsupervised Classification Approaches. In Proceedings of the 38th Hawaii International Conference on System Sciences, Big Island, HI, USA, 6 January 2005; p. 112c.
29. Hasan, A.; Moin, S.; Karim, A.; Shamshirband, S. Machine Learning-Based Sentiment Analysis for Twitter Accounts. *Math. Comput. Appl.* 2018, 23, 11. [CrossRef]
30. Maipradit, R.; Hata, H.; Matsumoto, K. Sentiment Classification Using N-Gram Inverse Document Frequency and Automated Machine Learning. *IEEE Softw.* 2019, 36, 65–70. [CrossRef]
31. Poria, S.; Cambria, E.; Gelbukh, A. Aspect extraction for opinion mining with a deep convolutional neural network. *Knowl.-Based Syst.* 2016, 108, 42–49. [CrossRef]
32. Lin, X.-M.; Ho, C.-H.; Xia, L.-T.; Zhao, R.-Y. Sentiment analysis of low-carbon travel APP user comments based on deep learning. *Sustain. Energy Technol. Assess.* 2021, 44, 101014. [CrossRef]
33. Behera, R.K.; Jena, M.; Rath, S.K.; Misra, S. Co-LSTM: Convolutional LSTM model for sentiment analysis in social big data. *Inf. Process. Manag.* 2021, 58, 102435. [CrossRef]
34. Ku, L.-W.; Liang, T.; Chen, H.-H. Tagging Heterogeneous Evaluation Corpora for Opinionated Tasks. In Proceedings of the 5th International Conference on Language Resources and Evaluation, Genoa, Italy, 22–28 May 2006; pp. 667–670.
35. Dong, Z.; Dong, Q. HowNet—A hybrid language and knowledge resource. In Proceedings of the 2003 International Conference on Natural Language Processing and Knowledge Engineering, Beijing, China, 19–26 October 2003; pp. 820–824.
36. Xu, L.; Lin, H.; Pan, Y.; Ren, H.; Chen, J. Constructing the Affective Lexicon Ontology. *J. China Soc. Sci. Tech. Inf.* 2008, 27, 180–185.
37. Ding, Y.; Li, B.; Zhao, Y.; Cheng, C. Scoring tourist attractions based on sentiment lexicon. In Proceedings of the 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chongqing, China, 25 March 2017; pp. 1990–1993.
38. Wu, X.; Wang, L. Investigation on sentiment of reviews with shoppingfield dictionary construction. *Comput. Technol. Dev.* 2017, 27, 194–199.
39. Xue, N. Chinese Word Segmentation as Character Tagging. *Comput. Linguist. Chin. Lang. Process.* 2003, 8, 29–48.
40. jieba. Available online: https://github.com/firstfaconjieba (accessed on 15 August 2021).
41. Sougou Input Method. Available online: https://pinyin.sogou.com/dict/ (accessed on 15 August 2021).
42. Taylor, S.J.; Letham, B. Forecasting at Scale. *Appl. Comput. Intell. Soft Comput.* 2016, 2016, 2385429. [CrossRef]
43. Radojevic, T.; Stanisic, N.; Stanić, N. Ensuring positive feedback: Factors that influence customer satisfaction in the contemporary hospitality industry. *Tour. Manag.* 2015, 51, 13–21. [CrossRef]
46. Nilashi, M.; Ahmadi, H.; Arji, G.; Alsalem, K.O.; Samad, S.; Ghabban, F.; Alzahrani, A.O.; Ahani, A.; Alarood, A.A. Big social data and customer decision making in vegetarian restaurants: A combined machine learning method. *J. Retail. Consum. Serv.* 2021, 62, 102630. [CrossRef]

47. Guangzhou Statistical Yearbook 2020. Available online: http://112.94.72.17/portal/queryInfo/statisticsYearbook/index (accessed on 8 October 2021).

48. Per Capita Disposable Income and Consumption Expenditure of Residents in Shanghai City in 2020. Available online: http://tjj.sh.gov.cn/ydsj71/20210122/caafdd75af224d29a266ee483e5aafe5.html (accessed on 8 October 2021).

49. Wuhan Statistical Yearbook 2020. Available online: http://tjj.wuhan.gov.cn/tjfw/tjnj/202102/t20210202_1624450.shtml (accessed on 8 October 2021).

50. China Statistical Yearbook 2020. Available online: http://www.stats.gov.cn/tjjs/njdsj/2020/indexch.htm (accessed on 8 October 2021).

51. López-Chau, A.; Valle-Cruz, D.; Sandoval-Almazán, R. Sentiment Analysis of Twitter Data through Machine Learning Techniques. In *Software Engineering in the Era of Cloud Computing*; Ramachandran, M., Mahmood, Z., Eds.; Computer Communications and Networks; Springer International Publishing: Cham, Switzerland, 2020; pp. 185–209, ISBN 978-3-030-33624-0.

52. An, J.; Kwak, H.; Jung, S.; Salminen, J.; Jansen, B.J. Customer segmentation using online platforms: Isolating behavioral and demographic segments for persona creation via aggregated user data. *Soc. Netw. Anal. Min.* 2018, 8, 54. [CrossRef]