*sustainability*

MDPI

*Article*

# How Can Big Data Support Smart Scenic Area Management? An Analysis of Travel Blogs on Huashan

**Jun Shao [1], Xuesong Chang [2] and Alastair M. Morrison [3,*]**

[1] School of Landscape Architecture, Beijing Forestry University, Beijing 100083, China; ninashaojun@bjfu.edu.cn

[2] Department of Tourism and Scenic Area Planning & Research, Beijing Tsinghua Tongheng Urban Planning & Design Institute, Beijing 100085, China; changxuesong@thupdi.com

[3] School of Hospitality and Tourism Management, Purdue University, West Lafayette, IN 47907, USA

* Correspondence: alastair@purdue.edu; Tel.: +86-10-6233-8304

Received: 30 October 2017; Accepted: 7 December 2017; Published: 9 December 2017

**Abstract:** Data from travel blogs represent important travel behavior and destination resource information. Moreover, technological innovations and increasing use of social media are providing accessible 'big data' at a low cost. Despite this, there is still limited big data analysis for scenic tourism areas. This research on Huashan (Mount Hua, China) data-mined user-contributed travel logs on the Mafengwo and Ctrip websites. Semantic analysis explored tourist movement patterns and preferences within the scenic area. GIS provided a visual distribution of blogger origins. The relationship between Huashan and adjoining tourism areas revealed a multi-destination pattern of tourist movements. Emotional analysis indicated tourist satisfaction levels, while content analysis explored more deeply into dissatisfying aspects of tourist experiences. The results should provide guidance for scenic areas in destination planning and design.

**Keywords:** travel blogs; big data; smart destination management; scenic areas; Huashan; China; data mining; GIS

## 1. Introduction

Travel behavioral pattern analysis is important for the planning and management of tourism destinations and attractions, allowing managers to more effectively develop strategies, map out travel routes, recommend products and experiences, and manage visitor impacts [1]. Travel blogs on social media are an excellent information source for analyzing tourist movements, activities, preferences, and satisfaction levels [2]. However, these data are not often being applied in scenic area planning in China. There is a tendency to focus just on entry tickets sold, revenue generation, and volumes of tourists, rather than on gathering and analyzing robust data on travel patterns and tourist behavior. Tourism planning in most Chinese scenic areas focuses on the levels of financial investment and GDP increases, but generally ignores tourist services and smart management [3]. The Jiuzhai Valley National Park in Sichuan was the first scenic area with a smart management system in China [1,4,5]. However, despite being an innovator, it was forced to launch a tourist flow forecasting system in 2014 to prevent a repeat of an overcrowding crisis during China's Golden Week in 2013 [6]. Even very well-known attractions in China such as Huashan are lacking basic data, including detailed statistics on the origins of tourists. Smart tourism strategies based on big-data analysis will undoubtedly contribute to solving these information deficiencies.

Huashan (simplified Chinese: 华山) in Shaanxi Province is a popular scenic area and was selected as the case study for this research. Huashan is a mountain situated in Huayin in the Weinan region

of Shaanxi Province, which is 120 km from Xi'an [7]. It is located near the southeast corner of the Ordos Loop of the Yellow River Basin, south of the Wei River Valley, at the eastern end of the Qin Mountains, in southern Shaanxi Province. It is the most western of the Five Great Mountains of China, and has a long history of religious significance. Huashan has five main peaks, the highest being the South Peak at 2155 m (7070 feet). Huashan can be classified as a mature and world-class tourist destination [8,9]. The number of tourists visiting Huashan has been growing steadily; for example, in 2016, total tourist arrivals were 26.2 million and ticket income was RMB 339 million Yuan, representing increases of 5% and 8%, respectively, over 2015 [10]. During the field study at Huashan in May 2016, although it is a leader in Weinan's tourism sector, the Huashan Tourism Group could not provide detailed statistical data on tourists other than tourist arrivals based on ticket counts and information provided by local travel agents and hotels in Weinan. Huashan's smart management system has existed since 2014 [11]; however, the focus is on providing online travel information and generating e-commerce. The system is lacking fine-grained statistics on tourists' behavior, so that it cannot provide data support for intelligent services and further planning. Due to the steep terrain of Huashan, overcrowding during peak attendance periods leads to falls and trampling. Huashan needs to invest great care, time, and effort to ensure the safety of its visitors. As with most of China's state-run scenic areas, Huashan does not feel compelled to significantly enhance visitor services while ticket revenues and profits are high. Most critically, the administration team at Huashan lacks research on tourist satisfaction. The absence of these data constrains intelligent management, marketing, and the sustainable development of Huashan and the surrounding region of Weinan. Paradoxically, due to the popularity of Huashan among tourists, there are many travel blogs on Huashan in Chinese social media, which are waiting to be mined to profile tourist behavior patterns and satisfaction levels. Recognizing the potential for user-generated content, travel blogs uploaded by Huashan tourists were analyzed to document travel movements, site linkages, and satisfaction levels. The principal research questions were:

1. How do Huashan visitors describe their travel experiences in blogs?
2. What sites are visited within the Huashan scenic area?
3. What are the patterns of movement within Huashan and adjoining destinations?
4. Are people satisfied with their experiences at Huashan? If tourists are dissatisfied, what are the reasons?
5. What are the geographic origins of Huashan tourists?
6. What are the monthly distributions of visits, expenditures, and lengths of stay for visitors to Huashan?

Answering these questions by analyzing travel blogs for Huashan is potentially a smart tourism solution leading to more effective scenic area planning and management. It may also contribute ideas and solutions to enhancing the sustainability of the scenic area.

## 2. Literature Review

### *2.1. Travel Blog Data and Tourist Behavior*

Data from travel blogs have five potentially beneficial features (five Vs), which are large scale (volume), content diversity (variety), quickly changing (velocity), authenticity (veracity), and application value (value) [12,13]. Additionally, spatial–temporal data from such blogs is multi-sourced, objective, dynamic, realistic, and fine-grained [1]. The data are free as people publish for personal reasons, such as recording and sharing experiences [14,15]. Blogs contain text, photos, videos, and other data forms. Tourists record some aspects of their actual behavior in blogs, which can be numerous, available at low cost, and contain rich information. Blogs are uploaded on sophisticated platforms for sharing through smartphones, pads, laptops, etc. [16,17]. Many online platforms also contain data tagged by users or by websites using geographic and time references [18]. Consumers are

in control of the text, geographic references, access times, images, and other information, rather than the destinations or attractions [16]. The geographic- (spatial) and time- (temporal) tagged data reveal travel behaviors and site and resource use [19,20]. They constitute the tourists' digital footprints, depicting movement within destinations and surrounding areas [16,21]. In addition, the data can be used to better comprehend tourist preferences for specific products, sites, and experiences, and satisfaction with the management and service quality of attractions [1].

Since travel blogs have rich information for investigating tourist movements and other behaviors, they are being increasingly used in tourism research. However, there is a danger of generalizing results from data that are not representative or have been poorly gathered [22]. Furthermore, as suggested by Hall [23–25] and Shoval [26], qualitative and quantitative analyses should be integrated to explore tourist behavior and traditional approaches should not be disregarded. These data collection approaches include face-to-face interviews [27], survey questionnaires [28–30], secondary data published by government or tourism organizations [27,31,32], onsite observations [14], and data collected by mobile tracking equipment [33–35].

Social media data may reveal tourists' impressions and experiences of the resources and environment of special scenic areas [1]. However, researchers must recognize that these do not reveal actual behavior. Notwithstanding this limitation, the data are less expensive than other sources of mobility big data, such as the Global Positioning System (GPS) logs on mobile phones that are controlled by mobile service providers. Therefore, travel blog data are appropriate in the case of Huashan, especially since it is a Chinese scenic area that has not implemented systems to track and mine tourist behavior data to support smart management.

### *2.2. Analysis of Tourist-Generated Big Data*

The analysis of blogs and traveler reviews represents a relatively new research method for tourism scholars. Tourist behavioral patterns and behaviors can be observed, recorded, and analyzed [36] through these data. Landmarks, travel routes, and places frequently used by tourists can be identified. Research using data uploaded by tourists on social media is at different geographic levels. At a global level, Hawelka et al. (2014) used Twitter data to explore international tourist travel behavior [37]. According to the nearly one billion tweets in 2012, these researchers investigated the mobility of tourists in different countries, characteristics of tourist flows, radii of rotation, diversity of destinations, and balance of capital inflows and outflows. At a country level, Li and Yang (2017) studied Sina Weibo data to explore travel patterns during China's Golden Week in 2014 [38]. Other studies have focused on urban areas [14,26,34,39] and attractions.

Thanks to the structured comparisons of images published on Flickr, scholars can readily obtain and analyze those photographic data. Some of the urban-scale research focuses on using Flickr's photo and text information [14,39]. As well as exploring the most popular tourist attractions, these studies compare travel patterns and behaviors of tourists from different origins [34]. Researchers also are using big data with geographic labels to analyze tourist preferences and activities [14,40]. For example, Guo et al. (2015) applied compact and sequential mode mining methods to collect and analyze geographical information from blogs on Qunar.com to analyze the interests, tourism activities, and use of specific tourism services [2].

This research analyzes data at the tourist attraction level, specifically for a scenic area. It focuses on tourist movements within and outside of the scenic area, visualizing the results with Geographical Information System (GIS) software.

## 3. Methods

### *3.1. Data Collection*

There is a great quantity of user-generated content on Chinese mainstream online travel websites such as Mafengwo, Baidu Tourism, and Ctrip. The data include not only text and photos in travel

blogs, but also tagged data such as travel dates, travel expenses, lengths of stay, associated destinations, and author residences. These data are gradually becoming easier to obtain with steady improvements of data structure in the travel websites. Several commercial software programs for web information acquisition are available, such as GooSeeker, Enthone, and the Locomotive and the Octopus web crawlers. These programs generally have the advantages of rapid iteration and ease of operation, making UGC acquisition from travel websites more convenient.

The Octopus web crawler tool was used to capture data from the Travel Guide Channels of Mafengwo and Ctrip on 20 May 2016. Two new tasks were created in the capture software and then a complete and clear capture process was established. First, all the lists of travel blogs related to Huashan were obtained by searching the home pages of the Travel Guide Channels using "Huashan" as the keyword. Next, a circular crawl list was created to catch the detailed pages of each travel blog. Then, in each detailed page, different grabbing positions were set up according to the page structure to obtain the corresponding contents, such as title, full text, release time, and tourist behavior. Finally, with the automatic page-turning function of the software, all relevant travel materials were obtained.

A total of 1468 travel blogs (over 840,000 words in Chinese) were captured. Among them, 768 (over 58,000 words) were retrieved from Mafengwo, and 700 (over 265,000 words) were retrieved from Ctrip.

### *3.2. Data Cleaning*

The data were saved in a structured format, that is, importing the basic trip elements, including the blog title, author, and full text as well as tourist behavioral information, including travel dates, travel expenditures, lengths of stay, other destinations visited, and author residences into an Excel file to form a database of travel blogs. Data missing basic information (3977 articles with more than 1,900,000 words), travel website template data extracted by the regular expression function of the software (3341 articles, >2,700,000 words), and advertising text data (1305 articles, >800,000 words) were deleted. The pure-text process was conducted on the full text of each travel blog and contents were sorted by sentences. Duplicate and blank content (>20,000 words), short articles with 10 characters and less or meaningless content such as "I am here!", "This picture is beautiful... "(more than >50,000 words in total) were removed. After sorting and screening, a total of 1080 high-quality Huashan travel blogs (>700,000 words) were obtained. Among them, 549 articles were from Mafengwo (>439,000 words) and 531 were from Ctrip (>265,000 words).

### *3.3. Data Analysis*

To address the first four research questions, content analysis of blogs was conducted. The semantic analysis of blog contents applied ROST CM, NetDraw and other tools for word segmentation and frequency statistics, and semantic structure drawings. A customized lexical pool was created based on the unique vocabulary associated with Huashan, and then integrated with the built-in Chinese word library of ROST CM. For the former, words were included such as "Huashan (华山)," "West Peak (西峰)," "cliff (绝壁)," "sunrise (日出)," "plank walk (栈道)," "Weinan (渭南)," "lamb liver soup (杂肝泡)," and "spicy Chinese food (香椿辣子)." Figure 1 shows the plank walk at the cliff near the South Peak of Huashan. Then, word segmentation processing on the full-text content of the travel blogs was done. Meaningless words were filtered out and word frequency statistics calculated. Using the highest frequency words, the Word Extraction feature was applied to each sentence line of the travel blogs. A co-occurrence matrix was derived by calculating the total frequencies of all the feature words. This matrix was visualized by the topological graph process using NetDraw to represent the semantic structure.

**Figure 1.** The plank walk at Huashan. Source: Shutterstock, Inc. (Nicholas Billington).

The emotional analysis of visitor satisfaction and dissatisfaction assessed inclinations within blog text using an emotional word library. After word segmentation, the text was separated into lines according to the ending punctuation marks such as periods, question marks, exclamations, ellipses, etc. The researchers ensured that each line expressed independent and complete meanings. Next, a dictionary of Chinese commendatory and derogatory terms written by Professor Li Jun of Tsinghua University and sentiment words from the China National Knowledge Infrastructure was selected as the basis for the emotional analysis. Common negative Chinese words were used as negative emotional expressions, and common Chinese adverbs represented the emotional judgments. The words in each line of the travel blogs were compared to those in this emotional lexicon. Also, emotional indications were judged according to multiple negation rules in Chinese language habits. A score was assigned according to the degree of emotion expressed by the adverbs. Positive and negative points indicated positive and negative emotions, respectively; zero point scores were neutral. The higher the absolute score, the greater the degree of emotion being expressed by the tourist.

The data analysis produced descriptive statistics on tourist characteristics. Microsoft Office Excel was used to classify, aggregate, cross-analyze, and visualize charts on structural tag data including tourist origin regions, places visited within Huashan scenic area, travel dates and expenditures, lengths of stay, and visits to adjoining destinations. These results satisfied the requirements for research objectives 5 and 6.

## 4. Results

### *4.1. Content Analysis of Huashan Travel Blogs*

A total of 4561 keywords were found after the word segmentation. Some 224 keywords in the Huashan travel blogs appeared more than 10 times (Appendix A), while the top 35 keywords are shown in Table 1. The semantic network diagram is illustrated in Figure 2. The scenic spots within Huashan that tourists visited most or found of greatest interest were comprised of four peaks (the South, North, East, and West Peaks). The Middle Peaks were less attractive and not included. Figure 2 indicates that East and West Peaks are associated with sunrise views and viewing. Climbing the West Peak was perceived as requiring greater physical exertion than the other peaks. The ropeway (or plank walk) was mentioned most frequently, along with the North and West Peaks. The main attractions within

Huashan were Yuquan Yard, Gold Lock, and Cang Long Ling. The distribution of tourists within Huashan seemed in accordance with natural conditions, and ropeway and tourism product design of the scenic area.

As can be seen from the semantic network diagram (Figure 2), Xi'an (西安) has a great impact on Huashan tourism. The Terracotta Warriors (兵马俑), Huaqing Pool (华清池), Hukou Waterfall (壶口瀑布) on the Yellow River (黄河), and downtown Xi'an were the most frequently visited adjoining attractions for Huashan tourists. Non-local tourists arrived in Huashan mainly by flights (飞机) through Xi'an Xianyang Airport (咸阳机场) or by train (火车).

**Table 1.** Occurrence of keywords in Huashan travel blogs.

| Number | Keyword | Occurrence Number | Number | Keyword | Occurrence Number |
|---|---|---|---|---|---|
| 1 | Huashan 华山 | 580 | 19 | Accommodation 住宿 | 78 |
| 2 | Xi'an 西安 | 464 | 20 | Downhill 下山 | 74 |
| 3 | North Peak 北峰 | 125 | 21 | Snack 小吃 | 74 |
| 4 | Train 火车 | 103 | 22 | Route 路线 | 74 |
| 5 | West Peak 西峰 | 95 | 23 | Yan Pagoda 雁塔 | 72 |
| 6 | Wall 城墙 | 132 | 24 | Terracotta Warriors 兵马俑 | 71 |
| 7 | Cableway 索道 | 131 | 25 | Shanxi 陕西 | 69 |
| 8 | East Peak 东峰 | 122 | 26 | South Peak 南峰 | 69 |
| 9 | History 历史 | 108 | 27 | Bell Tower 钟楼 | 68 |
| 10 | Yuquan Yard 玉泉院 | 107 | 28 | Climbing 登山 | 66 |
| 11 | Huis 回民 | 105 | 29 | Delicacy 美食 | 65 |
| 12 | Train station 火车站 | 104 | 30 | Square 广场 | 65 |
| 13 | Airport 机场 | 102 | 31 | Plank walk 栈道 | 62 |
| 14 | Hotel 酒店 | 101 | 32 | Metro 地铁 | 60 |
| 15 | Sunrise 日出 | 93 | 33 | Culture 文化 | 57 |
| 16 | Tourist 游客 | 92 | 34 | Rest 休息 | 57 |
| 17 | Admission ticket 门票 | 84 | 35 | Chang'an 长安 | 57 |
| 18 | Drum Tower 鼓楼 | 78 | ... | ... | ... |

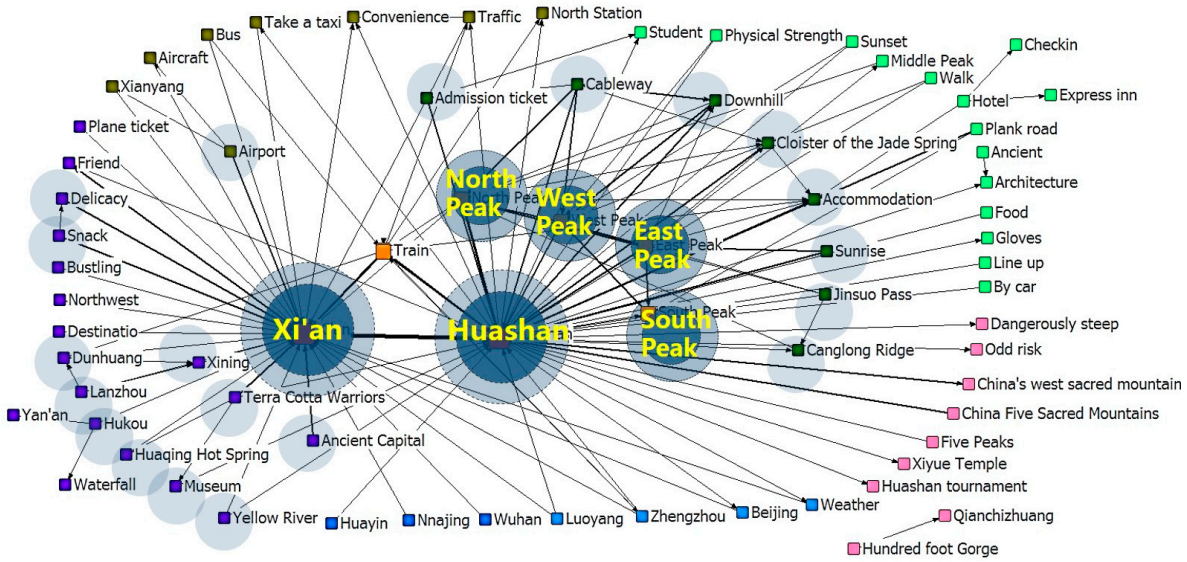


**Figure 2.** Semantic network of Huashan travel blog key words.

## *4.2. Travel Pattern Analysis*

The co-occurrence numbers of the place names appearing in Huashan blogs were calculated to represent the association degree between destinations. Then the association degree of Huashan and its adjoining destinations were obtained; only the elements whose co-occurrence numbers were greater than 10 were listed in Table 2.

**Table 2.** Association degree of destinations visited by Huashan tourists.

| No. | Destination | Destination | Co-Occurrence | No. | Destination | Destination | Co-Occurrence |
|---|---|---|---|---|---|---|---|
| 1 | Huashan 华山 | Xi'an 西安 | 353 | 10 | Huashan 华山 | Qinghai Lake 青海湖 | 10 |
| 2 | Huashan 华山 | Terracotta Warriors 兵马俑 | 50 | 11 | Xi'an 西安 | Terracotta Warriors 兵马俑 | 62 |
| 3 | Huashan 华山 | Dunhuang 敦煌 | 22 | 12 | Huaqing Pool 华清池 | Terracotta Warriors 兵马俑 | 30 |
| 4 | Huashan 华山 | Huaqing Pool 华清池 | 22 | 13 | Xi'an 西安 | Xianyang 咸阳 | 29 |
| 5 | Huashan 华山 | Luoyang 洛阳 | 21 | 14 | Xi'an 西安 | Dunhuang 敦煌 | 26 |
| 6 | Huashan 华山 | Huayin 华阴 | 19 | 15 | Xi'an 西安 | Hukou Waterfall 壶口瀑布 | 25 |
| 7 | Huashan 华山 | Hukou Waterfall 壶口瀑布 | 16 | 16 | Yan'an 延安 | Hukou Waterfall 壶口瀑布 | 23 |
| 8 | Huashan 华山 | Yan'an 延安 | 13 | 17 | Xi'an 西安 | Huaqing Pool 华清池 | 22 |
| 9 | Huashan 华山 | Xianyang 咸阳 | 11 | 18 | Xi'an 西安 | Luoyang 洛阳 | 11 |

Using a batch acquisition tool named xGeocoding, the coordinate information of these destinations, the latitude and longitude data, was derived. Taking the correlation degree between two destinations as input variables into GIS, a map showing the association degrees between Huashan and its adjoining destinations was obtained. Results illustrated at Figure 3 were exported from GIS.

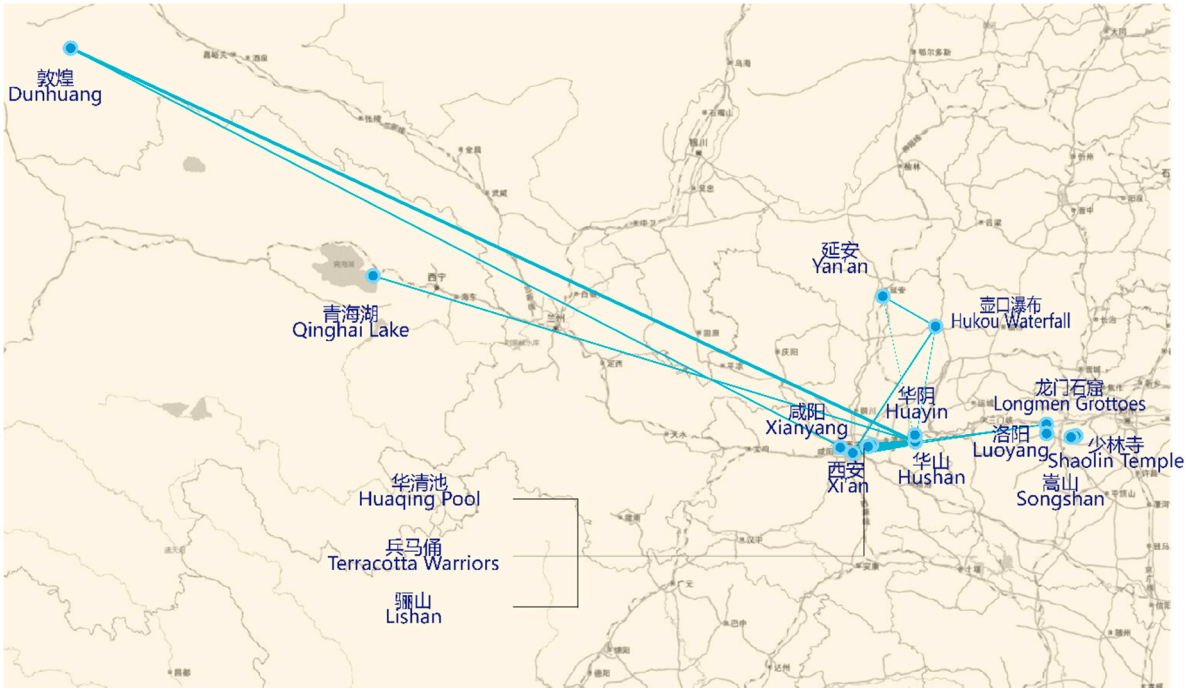


**Figure 3.** Multi-destination choices of Huashan tourists.

Only 17.4% of visitors had Huashan as their sole destination. The attractions jointly visited with Huashan included Xi'an, Terracotta Warriors, Huaqing Pool, Lishan, Xianyang, Huayin, Yan'an, Hukou Waterfall, Huangdi Mausoleum, Lishan, and some other popular places in Shaanxi Province. Visitors also had other joint destinations with Huashan, such as Luoyang, Songshan Shaolin Temple, Longmen Grottoes, Qinghai Lake, and other top attractions in neighboring provinces. However, Xi'an and Huashan was the most popular multi-destination itinerary found in the travel blogs.

### *4.3. Satisfaction or Dissatisfaction with Huashan Trips*

Table 3 shows the emotional analysis results. There were 11,686 positive (56.6%), 6126 neutral (29.7%), and 2833 negative evaluation items (13.7%). The significant level of negative comments should be of concern to the Huashan Management Committee.

**Table 3.** Emotional analysis of travel blogs.

| Category | Absolute Value > 20 | | 10< Absolute Value < 20 | | Absolute Value < 10 | |
|---|---|---|---|---|---|---|
| | Number | Proportion | Number | Proportion | Number | Proportion |
| Positive | 1157 | 5.60% | 2653 | 12.85% | 7876 | 38.15% |
| Negative | 22 | 0.11% | 375 | 1.82% | 2362 | 11.44% |

A thick data analysis of the unsatisfactory evaluations was conducted, selecting the top 100 blogs with the lowest scores for negative emotions, and reflecting the most unsatisfactory experiences with Huashan visits. Content analysis on each unsatisfactory evaluation was carried out, including in-depth analysis of different tourism experience elements, infrastructure, and services. Dissatisfaction with Huashan was focused on service facilities, congestion and garbage at scenic spots, and self-driving navigation difficulties (Table 4).

**Table 4.** Dissatisfaction evaluation of Huashan visits.

| Issues | Existing Problems | Travel Blogs with Translation in English |
|---|---|---|
| Accommodation | Poor conditions, difficulty in booking at the height of the tourist season | • Accommodation conditions are rudimentary; no bathing was possible because there is a lack of water in the mountains.<br>住宿条件简陋，因山上缺水，无洗澡条件。 |
| | | • Go inside a hundred meters, you can see the hotel; too many young people, sound insulation is not good.<br>往里走一百米就是这个酒店，年轻人较多，隔音不好。 |
| | | • Accommodation during National Day must be booked half a month in advance, or it is difficult to choose a cheap hotel or youth hostel.<br>国庆住宿必须提前半个月预定，不然很难选到便宜的酒店或青年旅社。 |
| | | • I called the police; the hotel name is Jinxin Hotel.<br>我就打报警，看了一下宾馆的名字是金鑫宾馆。 |
| Catering | Management confusion, high prices, bad service attitude | • A chicken is so expensive, WTF; in the off-season, you can choose to kill fresh chickens or have the chickens from last night, but now you can only have the latter.<br>一只鸡卖这么贵，太坑爹了，而且淡季的话可以选择昨夜杀的鸡和现杀的鸡，现在是旺季只有隔夜杀的鸡了。 |
| | | • At night, I went to the platform below the north peak, which was super-crowded. There was not a private shop selling bottled water, only the official one, very very expensive.<br>晚上，我到了北峰下面平台，人山人海。平台没有私人开得卖水小店，都是官方的售价，死贵死贵的。 |
| | | • I was cold and hungry, I had no choices and had to eat the most expensive instant noodles, which I have never had before.<br>我又冷又饿，没办法吃了自己史上最贵的一桶方便面。 |
| | | • Food is expensive but not delicious; even frozen dumplings are better than that.<br>饭菜又贵又不好吃，还不如速冻饺子呢。 |
| | | • The boss is very snobbish, you are not allowed to sit without ordering; Gansu boys had breakfast there, only porridge, bread, pickles, but it was so expensive.<br>老板很势力，不吃饭不许坐，甘肃男生在那里吃了早餐只有粥、馒头、咸菜，还卖那么贵。 |
| | | • I was very hungry in the mountains, so I ate a bowl of instant noodles. I had the waitress bring me some boiled water, which I had to pay for that.<br>在山上饿的不行，吃了碗泡面。我让阿姨给我加点白开水，还得另外收费。 |

**Table 4.** *Cont.*

| Issues | Existing Problems | Travel Blogs with Translation in English |
| --- | --- | --- |
| Toilets | Small quantity, heavy smell | • The hillside is full of people; it's very difficult to find a toilet.<br>山坡上全是人，上个厕所何其困难。 |
| | | • Looking around for a toilet, no mood for sightseeing.<br>四处寻找厕所，无心游览了。 |
| | | • There is a small temple on the middle peak, but there is nothing inside. There's a public toilet in the back and it smells bad.<br>中峰上有个小庙，什么东西都没有，后面是个公共厕所，发出难闻的异味。 |
| Passenger ropeway | Long wait time | • There are too many people waiting for the ropeway, mostly occupied by tour groups.<br>坐索道的人太多了，都是旅行团，带团去的能等死。 |
| | | • After waiting for the ropeway for nearly 40 min, when getting down I felt sick.<br>索道等候近四十分钟，下行时觉得恶心想吐。 |
| Capacity | Full of tourists | • It so happened that it rained during May Day holiday, it took nearly three hours to buy tickets, and I was wet, cold, and sad.<br>五一刚好赶上下雨，排队买票花了近三个小时，被雨淋湿湿的，更冷的难过。 |
| | | • Huashan has too many drawbacks. It's a tragedy that I encountered too many tourists.<br>华山的弊端太多，碰到游客一多真的会悲剧。 |
| | | • The pedestrians in front are not moving, so the traffic jams; I had to tilt my head stupidly just looking at the sea of clouds.<br>前面的行人也不动，交通堵车，只能侧头傻呆呆地望着云海。 |
| Sanitary | Litter everywhere | • It is chaotic; there is garbage on both sides of the tourist queue.<br>秩序有些混乱，队伍两边到处是随手丢弃的垃圾。 |
| | | • I saw very annoying things on Sun-watching Platform; there is so much garbage left behind by tourists.<br>观日台看到了让人很气愤的事情，有很多游客留下的垃圾。 |
| Transportation | Illegal taxis, rip off, old railway station | • These minibus drivers do not compete with the illegal taxi drivers. But when you get on the minibus, the driver said the illegal taxi drivers were actually the hotel's - confederates, who would tell tourists to have a rest at hotels first before climbing the mountain at night. They would then take you to - tourist traps that would rip you off.<br>这些小巴司机也不会主动跟黑车司机抢客，但是上了小巴，小巴司机就说了，这些黑车司机竟是旅馆的托，骗着游客说，晚上登山，现在可以去旅馆休息休息。然后，就把你拉去黑旅馆继续宰你 |
| | | • The train station is a bit old, which is reasonable. But the road to the tourist center is also old, full of potholes. Fortunately, I have a seat. Halfway there, the driver was afraid of being checked, so he asked the standing tourists to squat; there were many large bags, so this was very embarrassing.<br>火车站有点老旧情有可原，去游客中心的路居然一样老旧，坑坑洼洼。好在还有个座位，中途怕检查，司机让站着的游客蹲下，都是大包小包的也够难为情的。 |
| | | • In the train station I did not want to stay longer than it takes to take a few photos, so I left quickly. There were always black car drivers around, as well as pimping, very annoying.<br>在火车站不愿多做停留拍几张照片，赶紧离开，时时刻刻都有黑车司机，还有拉皮条的来骚扰烦死人了。 |
| | | • High-speed rail tickets from Huashan North to Xi'an North are quite difficult to buy; this time we had a loss.<br>从华山北到西安北的高铁票很难买，这次就吃了亏。 |
| | | • Rushing out of the trap of the black car drivers surrounded, you will see only a few regular taxis<br>穿过黑车司机的包围，会看到不多的几辆正规出租车。 |
| | | • Do not believe any taxi drivers who say they will charge you legally; when detouring, they won't tell you.<br>不要相信任何说打表的出租车司机，绕路的时候，他们是不会告诉你的。 |
| | | • It took us four hours to reach the foot of Huashan; we stopped frequently.<br>用了四个小时到达华山脚下，中途经常停下。 |

### *4.4. Regions of Origin of Huashan Tourists*

Table 5 shows the blog authors' residence cities, retrieved from tagged data in blog websites. Using GIS software, the visual pattern of origin distribution is shown in Figure 4. Huashan tourists were mainly from cities outside of Shaanxi Province. The first tier of origins according to volumes of tourists were Beijing, Shanghai, and Guangzhou; the surrounding region composed by Zhengzhou, Taiyuan, etc.; Southwest China consisting of Chengdu, Chongqing, etc. The next tier was comprised of Central China with Wuhan, Changsha, etc.; Northeast China including Shenyang, Dalian, etc.; and other parts of China. Spatial distance determined the distribution of Huashan tourist markets, which is influenced by economic development levels and convenience of transportation. The closer, the higher the level of economic development, and the more convenient the transportation, the greater the volumes of tourists are to Huashan.

**Table 5.** Cities of origin of Huashan tourists.

| Number | Origin | Proportion | Number | Origin | Proportion |
|---|---|---|---|---|---|
| 1 | Beijing 北京 | 16% | 16 | Shenyang 沈阳 | 2% |
| 2 | Shanghai 上海 | 12% | 17 | Qingdao 青岛 | 1% |
| 3 | Guangzhou 广州 | 9% | 18 | Luoyang 洛阳 | 1% |
| 4 | Zhengzhou 郑州 | 6% | 19 | Shijiazhuang 石家庄 | 1% |
| 5 | Xi'an 西安 | 5% | 20 | Hangzhou 杭州 | 1% |
| 6 | Chengdu 成都 | 4% | 21 | Wuxi 无锡 | 1% |
| 7 | Chongqing 重庆 | 3% | 22 | Xingtai 邢台 | 1% |
| 8 | Taiyuan 太原 | 3% | 23 | Ningbo 宁波 | 1% |
| 9 | Weinan 渭南 | 3% | 24 | Lanzhou 兰州 | 1% |
| 10 | Nanjing 南京 | 3% | 25 | Haerbin 哈尔滨 | 1% |
| 11 | Changsha 长沙 | 2% | 26 | Xianyang 咸阳 | 1% |
| 12 | Wuhan 武汉 | 2% | 27 | Suzhou 苏州 | 1% |
| 13 | Dalian 大连 | 2% | 28 | Yuncheng 运城 | 1% |
| 14 | Tianjin 天津 | 2% | 29 | Xuzhou 徐州 | 1% |
| 15 | Jinan 济南 | 2% | 30 | …… | …… |


**Figure 4.** Distribution of origins of Huashan tourists.

Figures 5–7 provide the travel date, expenditure, and length of stay results respectively. These statistics were calculated using the tagged data in Huashan travel blogs, which was structured by the blog websites. Most people visited Huashan in the months of April to October. This seasonal pattern for Huashan tourism is not as limited as for many other destinations in Northern China. The per capita travel expenditure of Huashan tourists was in the range of RMB 1000 to 3000 Yuan, which is slightly higher than for other scenic areas in China. The length of stay was three to five days. It bears noting that the length of stay of tourists who only visited Huashan was not as long as for people who also visited the Terracotta Warriors, Huaqing Pool, Yellow River (Hukou Waterfall), and other attractions in downtown Xi'an. This implies that Huashan needs to add to and upgrade its tourism products to increase length of stay and expenditure, and to enhance the holding power of the surrounding Weinan region.

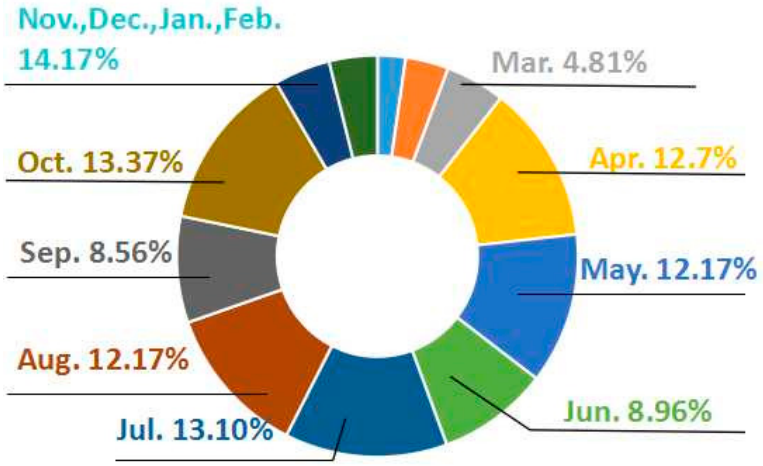


**Figure 5.** Monthly distribution of visits to Huashan.

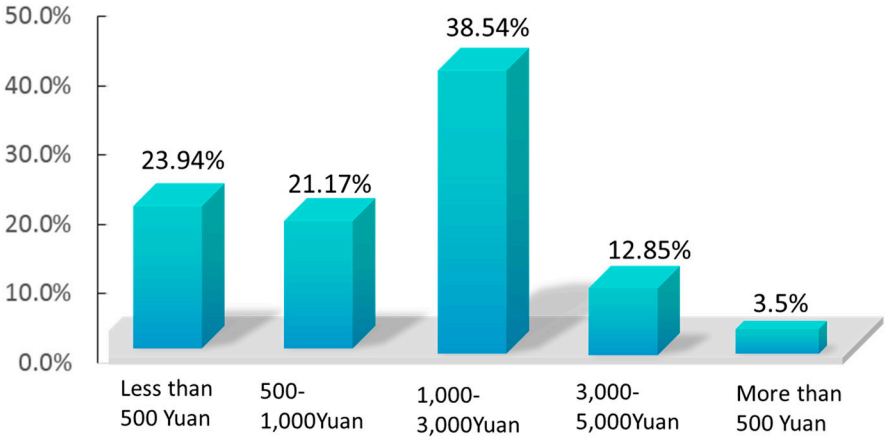


**Figure 6.** Per capita expenditures of Huashan tourists.

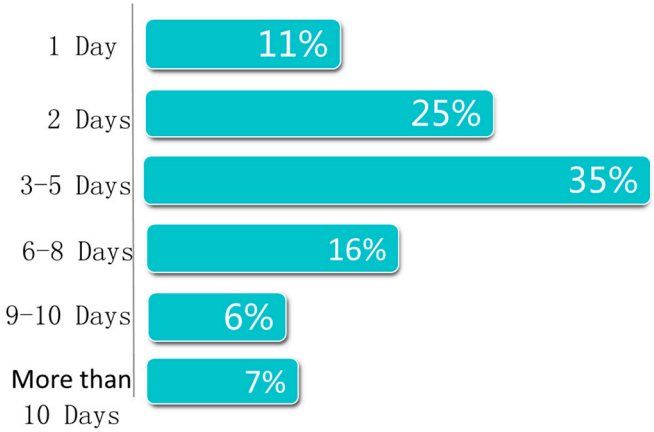


**Figure 7.** Lengths of stay of Huashan tourists.

## 5. Conclusions, Management Implications, and Research Limitations

This research is among the first to data-mine scenic area travel blogs by incorporating semantic analysis along with GIS visualizations. It demonstrates the value of these user-generated contents for market and satisfaction analysis of scenic area attractions. It is an exploratory analysis on travel blog data about scenic area attractions and there is considerable scope for future studies. Suggestions include analyzing the photographic content of travel blogs; conducting preference analyses among different tourist market segments; and cross-validation analysis with data from traditional research methods.

The results show that the tourist experience at Huashan is based on climbing and especially associated with the iconic "plank walk." Xi'an and Huashan are linked as destinations in the minds and actions of tourists. Specifically, downtown Xi'an, Terracotta Warriors, Huaqing Pool, and Yan'an are often grouped with Huashan in multi-destination trips.

The multi-destination tendency of Huashan tourists underlines the potential for cooperative marketing by the Huashan Management Committee along with the neighboring provinces of Henan, Qinghai, and Gansu. The other closest sites and attractions within Weinan did not appear in multi-destination patterns, which suggests that Huashan is overshadowing neighbors through its much greater destination image and market popularity. The Huashan Management Committee must, therefore, strengthen its role as a tourism development agent for the Weinan region. Greater attention must be focused on regional tourism development and marketing, integrating the tourism resources in Eastern Shaanxi and along the Yellow River.

There is a significant level of dissatisfaction with the facilities, services, and operational management of Huashan, which requires immediate attention. Overcrowding and littering are already serious issues, and will worsen as tourist numbers continue to increase. The sustainability of the Huashan experience is under threat. Visitor monitoring and management are insufficient at the current time; however, smart data-gathering and analysis such as demonstrated in this research can help point to solutions that will improve resource and experience sustainability.

Many attraction administration teams in China still have a narrow "ticket revenue and GDP" mindset and need to broaden their perspectives to operate more professionally as destination managers while assuring the sustainability of precious natural and cultural resources. The Huashan Management Committee should gather and use contemporary information sources, including smartphone 'footprint' data, to obtain real-time, spatial data on tourist and personnel movements within the scenic area that impact on the natural resources and environment, traffic flows and convenience of navigation, and visitor safety, experiences, and enjoyment. Managers should be accessing real-time data from big-data centers and cloud computing platforms, as well as analyzing tourist preferences and requirements.

As Gretzel et al. (2015) claimed, the lifeblood of smart tourism is big data, and the final purpose of smart tourism planning is extracting intelligence from big data [41]. Smart scenic area management will be assisted by technological approaches to gathering, analyzing, and interpreting big data [41,42], along with taking care of the human side by providing the types and quality of experiences that visitors are seeking [3,42]. This research verified that the results from travel blog data could help reveal tourists' opinions on services offered [42] at the area level, although there is a risk of bias by under-representing Huashan visitors who do not post online. Through the development of a smart scenic area system, the administration will be able to monitor tourist flow distribution, traffic conditions, and service facility use in real time. Timely diversion measures can be adopted to ensure the safety, comfort, and enjoyment of tourists. Moreover, service and facility quality must be continuously evaluated and improved based on visitor survey results and observation on usage of facilities and service encounters. Capacity measurement of most popular sites needs immediate attention as overcrowding is spoiling the tourist experience at Huashan.

It is recognized that this is only one example of a famous scenic area in China and the results may not be generalizable to other countries, let alone to other similar destinations in China. However, the research and its analysis can be helpful to protected area managers for smart destination management and promoting sustainability. The combination of qualitative and quantitative techniques applied to a scenic area using traveler blogs is rather unique. It has the potential of providing protected area managers with visitor monitoring and management data that can enhance resource sustainability and visitor satisfaction.

There are some limitations to this research that must be recognized. The research data were all from social media sources and there is a danger that they may be biased in under-representing Huashan visitors who do not post online. Additionally, all tourists were treated alike, and differences in demographics, travel group composition, and arrangements (e.g., independent vs. group tours) were not investigated. It is very important to stress that big data processing methods should be combined with other approaches, rather than being considered an independent method.

**Acknowledgments:** The authors would are grateful for the suggestions from anonymous reviewers.

**Author Contributions:** Jun Shao prepared the literature review and supervised the assembly of the initial manuscript. Xuesong Chang conducted the data collection and prepared the data analysis. Alastair M. Morrison provided the overall guidance on destination management and sustainability aspects, and was the editor of the final manuscript.

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

## Appendix

**Table A1.** Keywords in Huashan travel blogs that appear more than 10 times.

| No. | Word | Occurrence Number | No. | Word | Occurrence Number | No. | Word | Occurrence Number | No. | Word | Occurrence Number |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Huashan 华山 | 580 | 57 | Walk 步行 | 41 | 113 | Cost 费用 | 26 | 169 | On foot 走路 | 17 |
| 2 | Xi'an 西安 | 464 | 58 | Gold Lock 金锁关 | 41 | 114 | Photograph 拍照 | 26 | 170 | Graduation 毕业 | 17 |
| 3 | North Peak 北峰 | 125 | 59 | Take taxi 打车 | 39 | 115 | Mutton 羊肉 | 24 | 171 | Steep 陡峭 | 15 |
| 4 | Train 火车 | 103 | 60 | Plane ticket 机票 | 39 | 116 | Reserve 预定 | 24 | 172 | Economic 省钱 | 15 |
| 5 | West Peak 西峰 | 95 | 61 | Tour line 路线 | 39 | 117 | Vacation 假期 | 24 | 173 | Spectacular 壮观 | 15 |
| 6 | Wall 城墙 | 132 | 62 | Student 学生 | 38 | 118 | Hiking 徒步 | 24 | 174 | Direct 直达 | 15 |
| 7 | Cableway 索道 | 131 | 63 | Urban 市区 | 38 | 119 | Yummy 好吃 | 24 | 175 | Alone 独自 | 15 |
| 8 | East Peak 东峰 | 122 | 64 | North Station 北站 | 38 | 120 | Xining 西宁 | 24 | 176 | Luggage 行李 | 14 |
| 9 | History 历史 | 108 | 65 | Beijing 北京 | 38 | 121 | Mogao Grottoes 莫高窟 | 24 | 177 | Plan 规划 | 14 |
| 10 | Yuquan Yard 玉泉院 | 107 | 66 | Taste 味道 | 38 | 122 | Tang Paradise 芙蓉园 | 23 | 178 | Entrance 入口 | 14 |
| 11 | Huis 回民 | 105 | 67 | China's West Mountain 西岳 | 36 | 123 | Camera 相机 | 23 | 179 | Lanzhou 兰州 | 14 |
| 12 | Train station 火车站 | 104 | 68 | Experience 体验 | 36 | 124 | Taxi 出租车 | 23 | 180 | Private Cabs 黑车 | 14 |
| 13 | Airport 机场 | 102 | 69 | Luoyang 洛阳 | 36 | 125 | Destination 目的地 | 23 | 181 | Cloud Peak 云峰 | 12 |
| 14 | Hotel 酒店 | 101 | 70 | Convenient 方便 | 35 | 126 | Ancient 古代 | 23 | 182 | Lishan 骊山 | 12 |
| 15 | Sunrise 日出 | 93 | 71 | Steps 台阶 | 33 | 127 | Train Tickets 火车票 | 23 | 183 | Express Inn 快捷酒店 | 12 |
| 16 | Tourist 游客 | 92 | 72 | Scenery 风景 | 33 | 128 | Worry 担心 | 23 | 184 | Tourist 游人 | 12 |
| 17 | Admission ticket 门票 | 84 | 73 | Fountain 喷泉 | 33 | 129 | By car 坐车 | 23 | 185 | Expectation 期待 | 12 |
| 18 | Drum Tower 鼓楼 | 78 | 74 | Knapsack 背包 | 33 | 130 | Park 公园 | 23 | 186 | Whole Course 全程 | 12 |
| 19 | Accommodation 住宿 | 78 | 75 | Museum 博物馆 | 33 | 131 | Guide 导游 | 21 | 187 | Clothes 衣服 | 12 |
| 20 | Downhill 下山 | 74 | 76 | Cheap 便宜 | 32 | 132 | Freedom 自由 | 21 | 188 | Guangzhou 广州 | 12 |
| 21 | Snack 小吃 | 74 | 77 | Aircraft 飞机 | 32 | 133 | Hotel 宾馆 | 21 | 189 | East Gate 东门 | 12 |
| 22 | Route 路线 | 74 | 78 | The Forest of Steles 碑林 | 32 | 134 | Huayin 华阴 | 21 | 190 | Natural 自然 | 12 |
| 23 | Yan Pagoda 雁塔 | 72 | 79 | China's Five Sacred Mountains 五岳 | 30 | 135 | Wuhan 武汉 | 21 | 191 | Unfortunately 可惜 | 12 |
| 24 | Terracotta Warriors 兵马俑 | 71 | 80 | Huaqing Pool 华清池 | 30 | 136 | Many People 人多 | 21 | 192 | Challenge 挑战 | 12 |
| 25 | Shanxi 陕西 | 69 | 81 | Perform 表演 | 30 | 137 | Chengdu 成都 | 21 | 193 | Leave 离开 | 12 |
| 26 | South Peak 南峰 | 69 | 82 | Legend 传说 | 30 | 138 | Regret 遗憾 | 21 | 194 | Unique 唯一 | 12 |
| 27 | Bell tower 钟楼 | 68 | 83 | Hukou Waterfall 壶口瀑布 | 30 | 139 | Journey 旅途 | 20 | 195 | Yaozifanshen 鹞子翻身 | 11 |
| 28 | Climbing 登山 | 66 | 84 | Cold Rice Noodles 凉皮 | 30 | 140 | Xianyang 咸阳 | 20 | 196 | Shuttle Bus 班车 | 11 |
| 29 | Delicacy 美食 | 65 | 85 | Driver 司机 | 30 | 141 | Tent 帐篷 | 20 | 197 | Huangshan 黄山 | 11 |
| 30 | Square 广场 | 65 | 86 | Security 安全 | 30 | 142 | Nanjing 南京 | 20 | 198 | Expenditure 花费 | 11 |
| 31 | Plank walk 栈道 | 62 | 87 | Middle Peak 中峰 | 29 | 143 | Tianjin 天津 | 20 | 199 | Impression 印象 | 11 |
| 32 | Metro 地铁 | 60 | 88 | Map 地图 | 29 | 144 | Glove 手套 | 20 | 200 | Shock 震撼 | 11 |
| 33 | Culture 文化 | 57 | 89 | Sunset 日落 | 29 | 145 | Gate 山门 | 20 | 201 | Yuntai 云台 | 11 |
| 34 | Rest 休息 | 57 | 90 | Line up 排队 | 29 | 146 | Memorial Gateway 牌坊 | 20 | 202 | Imagine 想象 | 11 |
| 35 | Changan 长安 | 57 | 91 | Side 旁边 | 29 | 147 | Shaanxi 陕西省 | 20 | 203 | Setting sun 夕阳 | 11 |
| 36 | Online 网上 | 57 | 92 | Baidu 百度 | 29 | 148 | Smoothly 顺利 | 20 | 204 | Ticket Office 售票处 | 11 |
| 37 | Architecture 建筑 | 57 | 93 | Music 音乐 | 29 | 149 | Environment 环境 | 18 | 205 | Sell ticket 售票 | 11 |
| 38 | Story 故事 | 57 | 94 | Love 爱情 | 29 | 150 | Happy 开心 | 18 | 206 | Early morning 清晨 | 11 |
| 39 | Traffic 交通 | 56 | 95 | Chartered 包车 | 29 | 151 | Beauty 漂亮 | 17 | 207 | Shanxi opera 秦腔 | 11 |
| 40 | Friend 朋友 | 56 | 96 | Youth 青年 | 29 | 152 | Check in 入住 | 17 | 208 | Strange 陌生 | 11 |

**Table A1.** *Cont.*

| No. | Word | Occurrence Number | No. | Word | Occurrence Number | No. | Word | Occurrence Number | No. | Word | Occurrence Number |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 41 | Canglong Ridge 苍龙岭 | 56 | 97 | Northwest 西北 | 27 | 153 | Food 食物 | 17 | 209 | Beautiful 美丽 | 11 |
| 42 | Since ancient 自古 | 54 | 98 | Classmate 同学 | 27 | 154 | Thousands of Years 千年 | 17 | 210 | Hotel 旅馆 | 11 |
| 43 | Physical Strength 体力 | 54 | 99 | Qinghai Lake 青海湖 | 27 | 155 | Dayan Pagoda 大雁塔 | 17 | 211 | Lotus 莲花 | 11 |
| 44 | Mountaintop 山顶 | 53 | 100 | Ticket Price 票价 | 27 | 156 | Street 街道 | 17 | 212 | Nervous 紧张 | 11 |
| 45 | Xiyue Temple 岳庙 | 53 | 101 | Problem 问题 | 27 | 157 | Taiyuan 太原 | 17 | 213 | Explain 讲解 | 11 |
| 46 | Ancient City 古城 | 53 | 102 | Taishan 泰山 | 27 | 158 | Dinner 晚饭 | 17 | 214 | Memory 回忆 | 11 |
| 47 | Onhill 山上 | 53 | 103 | Huashan Road 华山路 | 27 | 159 | The Great Wall 长城 | 17 | 215 | Reasonable 合理 | 11 |
| 48 | Bus 大巴 | 51 | 104 | Restaurant 饭店 | 27 | 160 | Comfortable 舒服 | 17 | 216 | Hanzhoung 汉中 | 11 |
| 49 | Museum 博物馆 | 51 | 105 | Visit 游览 | 27 | 161 | Real 真实 | 17 | 217 | Altitude 海拔 | 11 |
| 50 | Weather 天气 | 50 | 106 | University 大学 | 26 | 162 | Kaifeng 开封 | 17 | 218 | Country 国家 | 11 |
| 51 | Cable Car 缆车 | 44 | 107 | Yan'an 延安 | 26 | 163 | Hard Seat 硬座 | 17 | 219 | Thank 感谢 | 11 |
| 52 | Transit 公交 | 42 | 108 | Xi'an Downtown 西安市 | 26 | 164 | Famous 有名 | 17 | 220 | Bustling 繁华 | 11 |
| 53 | Zhengzhou 郑州 | 42 | 109 | Bus 公交车 | 26 | 165 | Dangerously Steep 险峻 | 17 | 221 | Metro Station 地铁站 | 11 |
| 54 | Bell and Drum Tower 钟鼓楼 | 42 | 110 | Yellow River 黄河 | 26 | 166 | Desert 沙漠 | 17 | 222 | Road 道路 | 11 |
| 55 | Dunhuang 敦煌 | 42 | 111 | Qianchi Zhuang 千尺幢 | 26 | 167 | Hundred Foot Gorge 百尺峡 | 17 | 223 | Miss 错过 | 11 |
| 56 | Ancient Capital 古都 | 41 | 112 | Ruins 遗址 | 26 | 168 | One way 单程 | 17 | 224 | White clouds 白云 | 11 |

## References


1. Dang, A.; Zhang, D.; Ma, Q. The Intelligent Scenic Spot Management and Service Discussion in the Big Data Era. *J. Hum. Settl. Forum West China* **2016**, *31*, 8–13. (In Chinese)
2. Guo, L.; Li, Z.; Sun, W. Understanding Travel Destination from Structured Tourism Blogs. In Proceedings of the 2015 Wuhan International Conference on e-Business, Wuhan, China, 19–21 June 2015; pp. 144–151.
3. Zhang, L. Smart Tourism Development Needs Intelligent Thinking. Available online: http://www.pinchain.com/article/33689 (accessed on 24 May 2016).
4. Shao, Z.; Zhang, X.; Ma, J. Smart Management of Jiuzhai Valley Scenic Area Based on Internet of Things. *Geogr. Inf. World* **2010**, *5*, 12–16. (In Chinese)
5. Deng, G.; Shao, Z. Intelligent Management and Service of Jiuzhai Valley Scenic Area Based on Video Cruise. *Comput. Eng. Des.* **2011**, *11*, 3920–3924. (In Chinese)
6. Zhao, M. Jiuzhai Valley Launched A Big Data Mining System to Prevent a Repeat of the Over-Crowded Crisis. Available online: http://www.thepaper.cn/newsDetail_forward_1269459 (accessed on 24 May 2016).
7. China Central TV. First Light in the Morning. Episode 1 of Chinese Civilization. Available online: http://shaoer.cntv.cn/children/C19222/classpage/video/20100301/101570.shtml (accessed on 26 May 2017).
8. China Today. Huashan: The Top One Steepest and Wondrous Mountain in China. Available online: http://www.chinatoday.com.cn/chinese/mlzg/mlsh/201608/t20160831_800066194.html (accessed on 26 May 2017).
9. China Internet News Center. Mount Huashan—The Root of Chinese. Available online: http://www.china.com.cn/culture/zhuanti/shz09/2010-04/09/content_19780491.htm (accessed on 26 May 2017).
10. Huashan Administration Committee. Mount Huashan Won the Most Popular Tourist Attraction in 2017. Available online: http://www.weinan.gov.cn/news/bmdt/hgw/590986.htm (accessed on 18 October 2017).
11. Weinan Daily News. Smart Tourism Plugs Wings for Great Huashan's Soar. Available online: http://www.weinan.gov.cn/cul/lldt/560297.htm (accessed on 18 October 2017).
12. Dang, A.; Yuan, M.; Shen, Z. Rational Planning and Urban and Rural Governance and Based on Smart City and Big Data. *Constr. Sci. Technol.* **2015**, *5*, 64–66. (In Chinese)
13. Bian, F.; Du, J.; Meng, X. The Demand, Application and Challenge of Big Data Processing. *J. Geomat.* **2016**, *4*, 1–4. (In Chinese)
14. Vu, H.Q.; Leung, R.; Rong, J. Exploring Park Visitors' Activities in Hong Kong Using Geotagged Photos. In *Information and Communication Technologies in Tourism 2016*; Springer International Publishing: Gewerbestrasse, Switzerland, 2016; pp. 183–196.
15. O'Neill, E.; Kostakos, V.; Kindberg, T. Instrumenting the City: Developing Methods for Observing and Understanding the Digital Cityscape. In Proceedings of the International Conference on Ubiquitous Computing, Orange County, CA, USA, 17–21 September 2006; Springer: Heidelberg, Germany, 2006; pp. 315–332.
16. Zheng, Y.; Zha, Z.; Chua, T. Mining Travel Patterns from Geotagged Photos. *ACM Trans. Intell. Syst. Technol.* **2012**, *3*, 1–18. [CrossRef]
17. Zheng, Y.; Zhang, L.; Xie, X. Mining Interesting Locations and Travel Sequences from GPS Trajectories. In Proceedings of the International Conference on World Wide Web, Madrid, Spain, 20–24 April 2009; ACM: New York, NY, USA, 2009; pp. 791–800.
18. Wong, E.; Law, R.; Li, G. Reviewing Geotagging Research in Tourism. In *Information and Communication Technologies in Tourism 2017*; Springer International Publishing: Gewerbestrasse, Switzerland, 2017; pp. 43–57.
19. Henderson, J.; Milligan, K.; Niccols, A. Event-Based Analysis of People's Activities and Behavior Using Flickr and Panoramio Geotagged Photo Collections. In Proceedings of the International Conference Information Visualisation, London, UK, 26–29 July 2010; IEEE Computer Society: Washington, DC, USA, 2010; pp. 289–296.
20. Kisilevich, S.; Mansmann, F.; Keim, D. P-DBSCAN: A Density Based Clustering Algorithm for Exploration and Analysis of Attractive Areas Using Collections of Geo-Tagged Photos. In Proceedings of the International Conference and Exhibition on Computing for Geospatial Research & Application, Washington, DC, USA, 21–23 June 2010; ACM: New York, NY, USA, 2010; pp. 1–4.

21. Jiang, K.; Wang, P.; Yu, N. Context Rank: Personalized Tourism Recommendation by Exploiting Context Information of Geotagged Web Photos. In Proceedings of the Sixth International Conference on Image and Graphics, Hefei, China, 12–15 August 2011; IEEE Computer Society: Washington, DC, USA, 2011; pp. 931–937.
22. Xiang, Z.; Du, Q.; Ma, Y.; Fan, W. A Comparative Analysis of Major Online Review Platforms: Implications for Social Media Analytics in Hospitality and Tourism. *Tour. Manag.* **2017**, *58*, 51–65. [CrossRef]
23. Hall, C.M. *Tourism: Rethinking the Social Science of Mobility*; Prentice Hall: Harlow, UK, 2005.
24. Hall, C.M. Time, space, tourism and social physics. *Tour. Recreat. Res.* **2005**, *30*, 93–98. [CrossRef]
25. Hal, C.M.; Page, S.J. Progress in Tourism Management: From the Geography of Tourism to Geographies of Tourism—A Review. *Tour. Manag.* **2009**, *30*, 3–16. [CrossRef]
26. Shoval, N.; McKercher, B.; Ng, E. Hotel location and tourist activity in cities. *Ann. Tour. Res.* **2011**, *38*, 1594–1612. [CrossRef]
27. Forer, P.; Simmons, D. Serial experiences: Monitoring, Modelling and Visualising the Free Independent Traveller in New Zealand at Multiple Scales with GIS. In *Monitoring and Management of Visitor Flows in Recreational and Protected Areas*; Arnberger, A., Brandenburg, C., Muhar, A., Eds.; Institute of Landscape Architecture and Landscape Management, Bodenkultur University: Vienna, Austria, 2002; pp. 173–180.
28. Lau, G.; McKercher, B. Understanding Tourist Movement Patterns in a Destination: A GIS Approach. *Tour. Hosp. Res.* **2006**, *7*, 39–49. [CrossRef]
29. McKercher, B.; Lau, G. Movement Patterns of Tourists within a Destination. *Tour. Geogr.* **2008**, *10*, 355–374. [CrossRef]
30. Xia, J.; Zeephongsekul, P.; Arrowsmith, C. Modelling Spatio-temporal Movement of Tourists Using Finite Markov Chains. *Math. Comput. Simul.* **2009**, *79*, 1544–1553. [CrossRef]
31. Li, X.; Meng, F.; Uysal, M. Spatial Pattern of Tourist Flows Among the Asia-Pacific Countries: An Examination over a Decade. *Asia Pac. J. Tour. Res.* **2008**, *13*, 229–243. [CrossRef]
32. Wu, C.; Carson, D. Spatial and Temporal Tourist Dispersal Analysis in Multiple Destination Travel. *J. Travel Res.* **2008**, *46*, 311–317. [CrossRef]
33. Asakura, Y.; Iryo, T. Analysis of tourist Behaviour Based on the Tracking Data Collected Using a Mobile Communication Instrument. *Transp. Res. Part A* **2007**, *41*, 684–690. [CrossRef]
34. Vu, H.; Gang, L.; Law, R. Exploring the Travel Behaviors of Inbound Tourists to Hong Kong Using Geotagged Photos. *Tour. Manag.* **2015**, *46*, 222–232. [CrossRef]
35. Zheng, Y.; Zha, Z.; Chua, T. Research and Applications on Georeferenced Multimedia: A Survey. *Multimed. Tools Appl.* **2011**, *51*, 77–98. [CrossRef]
36. Girardin, F.; Calabrese, F.; Fiore, F. Digital Footprinting: Uncovering Tourists with User-Generated Content. *IEEE Pervasive Comput.* **2008**, *7*, 36–43. [CrossRef]
37. Hawelka, B.; Sitko, I.; Beinat, E. Geo-located Twitter as Proxy for Global Mobility Patterns. *Cartogr. Geogr. Inf. Sci.* **2014**, *41*, 260. [CrossRef] [PubMed]
38. Li, D.; Yang, Y. GIS Monitoring of Traveler Flows Based on Big Data. In *Analytics in Smart Tourism Design*; Springer International Publishing: Gewerbestrasse, Switzerland, 2017; pp. 111–126.
39. Kádár, B.; Gede, M. Where Do Tourists Go? Visualizing and Analysing the Spatial Distribution of Geotagged Photography. *Cartogr. Int. J. Geogr. Inf. Geovis.* **2013**, *48*, 78–88. [CrossRef]
40. Sun, Y.; Fan, H.; Helbich, M.; Zipf, A. Analyzing Human Activities through Volunteered Geographic Information: Using Flickr to Analyze Spatial and Temporal Pattern of Tourist Accommodation. In *Progress in Location-Based Services. Lecture Notes in Geoinformation and Cartography*; Krisp, J., Ed.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 57–69.
41. Gretzel, U.; Sigala, M.; Xiang, Z.; Koo, C. Smart tourism: Foundations and developments. *Electron. Mark.* **2015**, *25*, 179–188. [CrossRef]
42. Buhalis, D.; Amaranggana, A. Smart Tourism Destinations Enhancing Tourism Experience through Personalisation of Services. In *Information and Communication Technologies in Tourism*; Tussyadiah, I., Inversini, A., Eds.; Springer: Berlin/Heidelberg, Germany, 2015; pp. 377–389.