sustainability

Article

The International City Image of Beijing: A Quantitative Analysis Based on Twitter Texts from 2017–2021

Zhishuo Zhang [1], Manting Luo [1], Ziyu Luo [2] and Huayong Niu [1,*]

[1] International Business School, Beijing Foreign Studies University, Beijing 100089, China
[2] Institute of Art, Communication University of China, Beijing 100024, China
* Correspondence: bfsuniuhy@163.com

Abstract: With the advent of the Internet era, users from numerous countries can express their opinions on social media platforms represented by Twitter. Unearthing people's image perceptions of cities from tweets helps relevant organizations understand the image that cities present on mainstream social media and take targeted measures to shape a good international image, which can enhance international tourists' willingness to travel and strengthen city's tourism competitiveness. This paper collects nearly 130,000 tweets related to "Beijing" ("Peking") from 2017–2021 through web-crawler technology, and uses Term Frequency-Inverse Document Frequency (TF-IDF) keywords statistics, Latent Dirichlet Allocation (LDA) topic mining, and Valence Aware Dictionary and sEntiment Reasoner (VADER) sentiment analysis to further summarize the characteristics of Beijing's international image and propose strategies to communicate its international image. This research aims to tap into the international image of Beijing presented on Twitter, and provide data support for the relevant Chinese and Beijing authorities to develop communication strategies, as well as providing a reference for other cities aiming to manage their international image.

Keywords: international city image; social media; tourism competitiveness; text mining; natural language processing

Citation: Zhang, Z.; Luo, M.; Luo, Z.; Niu, H. The International City Image of Beijing: A Quantitative Analysis Based on Twitter Texts from 2017–2021. *Sustainability* **2022**, *14*, 10675. https://doi.org/10.3390/su141710675

Academic Editor: Yoonjae Nam

Received: 21 July 2022
Accepted: 25 August 2022
Published: 27 August 2022

1. Introduction

With the increase in people's living standards and improvements in transportation conditions, the growing tourism consumption demand have made tourism an important point of growth and a pillar industry to drive the city economy. Therefore, research on the competitiveness of city tourism has attracted extensive attention from scholars [1–5]. The international city image is a comprehensive reflection of its history and culture, geography and humanities, economic level and governance level. People's perception of the city image is a key factor in the choice of tourist city. Only through scientific positioning and dynamic management of its own image can a city stand out in the increasingly fierce competition in the tourist market, and thus enhance its competitiveness. By studying the perceived international image of a city, tourism enterprises and government departments can understand people's concerns and the positive and negative factors that affect people's emotions. This will help in the development of targeted communication strategies and stimulate tourism's vitality.

As the capital of China, Beijing is the political, cultural, international communication, and technology innovation center, with deep historical and cultural deposits. The success of the 2022 Beijing Winter Olympic Games made Beijing the first city to host both the Summer and Winter Olympic Games, which attracted the attention of the world and brought new opportunities to shape the Beijing's international image. By strengthening the research and management of its own international image, Beijing can give full play to its advantages, use major platforms with worldwide influence to shape a good international image, and enhance its international tourism competitiveness.

Twitter is a social platform with wide influence, where users in many countries express their views, comment and have discussions. Unlike the traditional communication mode, the audience is no longer a passive recipient, but a participant and producer of content, and the content mostly derives from users' views and opinions on life, which has a high persuasive power. These features make Twitter an important way to shape and spread an international city image. In addition, compared to other, more travel-oriented social media platforms, Twitter not only shows tourism images of cities, but also non-tourism images of economic trade, sports events, film and entertainment, social issues, etc. These non-tourism images also influence tourists' choices. By mining the content of tweets, we can better understand the current international image of a city and provide a basis for government departments to make decisions and improve the international image in a targeted manner.

This paper selects the international city image of Beijing as the research object, and uses web-crawler technology to obtain tweets containing "Beijing" ("Peking") from 2017 to 2021. The tweets are pre-processed by language recognition, language transcription, word separation, etc. The processed tweets are subjected to Term Frequency-Inverse Document Frequency (TF-IDF) keyword statistics and Latent Dirichlet Allocation (LDA) topic model construction to interpret the keywords and the themes shown in Beijing each year. This paper uses Valence Aware Dictionary and sEntiment Reasoner (VADER) sentiment analysis to further quantify the sentiment values of Beijing city at each timepoint and analyzes the trends in sentiment values from January to December 2017–2021. Finally, we obtain the international image of Beijing, and propose a communication strategy regarding Beijing's international image in response to the analysis results.

The main contributions of this paper are as follows: (1) From the perspective of people's perception of a city, this paper explores the ways in which cities can improve their tourism competitiveness by shaping a good international city image by mining the international city image in mainstream social media, providing a new perspective for such research. (2) As the factors affecting tourists' choices are not only tourism images, this paper examines the international city image, rather than just the tourism image. (3) In terms of research methodology, this paper uses web-crawler technology to capture tweets and quantifies the textual information through natural language-processing techniques. (4) This paper identifies and transcribes the German text in the tweets to further expand the sample size. (5) There is little literature on the international city image of Beijing, and this paper further expands the research in this area.

This paper is structured as follows. The second section sorts the relevant literature in the field. The third section introduces the basic models and methods involved in this study. The fourth section collects nearly 130,000 tweets related to "Beijing" ("Peking") by web-crawlers, and uses natural language-processing techniques to quantify the collected texts. The fifth section further discusses the results of the empirical study, analyzes and summarizes the characteristics of the international city image of Beijing, and then proposes communication strategies for the international city image. The sixth section summarizes the main findings, points out the significance of the and the limitations of the current study, and then provides an outlook on future research directions.

2. Literature Review

2.1. Definition of the International City Image

The term "city image" was first proposed by American urbanist Lynch [6] in 1964, who believed that city image is an impression shared by a certain number of city residents and includes five elements: roads, boundaries, areas, nodes and landmarks. The city image proposed by Lynch was limited to the impressions that city residents had of their places. It was restricted to the perception of the physical form of the city. Later scholars placed more emphasis on the impression of non-residential conditions [7–11], and incorporated city culture and spirit into the concept of "city image", i.e., city image is the public's comprehensive evaluation of the internal and external strengths and future development of a city, reflecting its characteristics and style [12,13]. With the development of modern networks, some scholars have proposed that the construction of city image is a process of external communication, i.e., the city image is formed through the combination of mass media, personal experience and environmental factors [14]. Further, it was suggested that the media is the main tool of city image communication, and the media encodes and decodes comprehensive information regarding the city and acts on public perceptions to eventually form the city image [15–19]. The media has an irreplaceable role in the communication of city image due to its timeliness and comprehensiveness [20]. City image is an objective social existence and a subjective social evaluation, and the construction of a city image is also the process of the city being perceived and re-evaluated by the public [21,22]. According to the different subjects of city image perception, the city image can be further divided into international city image and domestic city image. The international city image is the content presented by the city image in the process of international communication, which is the city image in the eyes of the international public. The international city image relies on international communication, and good international communication ideas are conducive to the establishment of a good international city image [23–26]. Therefore, understanding the current city image of Beijing on international mainstream social media is an essential part of enhancing Beijing's international image.

2.2. International City Image and City Tourism Competitiveness

In recent years, with the booming tourism industry, the relationship between the international city image and the competitiveness of urban tourism has received attention from several scholars. It has been argued that the competitiveness of a tourist destination is expressed as the perception of the tourists [27,28]. Vinyals-Mirabent [29] confirmed through a study that attraction factors such as city architecture and culture are essential features of city tourism competitiveness and help to distinguish a city's image from that of its competitors. Kim and Lee [30] found that the dynamic, static, and concrete dimensions of a city's image positively influence tourists' willingness to return. Some subsequent scholars have also confirmed the view that city image is an important factor influencing the future behavioral intentions of tourists, which directly or indirectly affects their decision-making behavior [31–33]. In addition, some studies have investigated the relationship between city tourism competitiveness and national culture and found that components of city image such as national culture have a significant impact on city tourism competitiveness [34,35]. Destination image was also found to be an antecedent of destination personality, and destination personality directly affects individuals' attitudes toward visiting [36,37]. Existing research on the international city image also suggests that residents' and tourists' perceptions of the city can influence the level of government support for tourism, with more positive perceptions of the city leading to more support for tourism in city-building [38–40]. The international city image can have an impact on the competitiveness of tourism, and the international city image becomes a strategic means of development for the city to improve its tourism competitiveness. Therefore, this paper examines the city's characteristics from the perspective of Beijing's international image, and proposes image-shaping solutions to enhance Beijing's urban tourism competitiveness.

3. Model Methodology

3.1. Word Separation and Word Frequency Statistics

The separation of English words was mainly performed by "word_tokenize()" in the *Natural Language Toolkit (NLTK)* module, a Python-based platform for a natural language processing toolset. To improve the word separation effect, this paper first preprocesses the original text, including abbreviation reduction, *Uniform Resource Locator (URL)* deletion, emoji deletion, "@" "#" symbol deletion, word form reduction, synonym replacement, phrase recognition and deactivation removal, to obtain the word separation results. The word frequency statistics can then be completed using "collections.Counter()".

3.2. Term Frequency-Inverse Document Frequency

The TF-IDF [41] method consists of two components: Term Frequency (TF) and Inverse Document Frequency (IDF). TF measures the frequency with which words occur in a document and can be calculated by Equation (1).

$$tf_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}} \tag{1}$$

In the context of this paper, we further elaborated the meaning of the symbols represented in the formula, where i is different words, j refers to different tweets, $n_{i,j}$ is the number of times word i appears in tweet j, $\sum_k n_{k,j}$ represents the total number of times all words appear in tweet j.

IDF measures the importance of words, and its mechanism reduces the weight of common words and increases the weight of rare words. The specific calculation formula is shown in Equation (2).

$$idf_i = \log \frac{|D|}{|j : t_i \in d_j| + 1} \tag{2}$$

where $|D|$ is the total number of tweets, $|j : t_i \in d_j|$ represents the number of tweets containing the word t_i. If none of the tweets contain the word t_i, this leads to $|j : t_i \in d_j| = 0$. To avoid the denominator of 0, we use $|j : t_i \in d_j| + 1$.

Ultimately, the value of TF-IDF is obtained by multiplying the resultant value of Equation (1) with the resultant value of (2), as in Equation (3). The more important the word is, the lager its TF-IDF value will be.

$$tf - idf = tf_{i,j} * idf_i \tag{3}$$

3.3. Latent Dirichlet Allocation Topic Model

Word frequency statistics can simply and intuitively extract hot words, but do not consider the semantic association behind the words. Some words may be generated in the same thematic context, such as “It’s very sunny outside and there is no wind” and “I ate a lot of ice cream”, which both reflect the topic of “hot weather”. However, if we only calculate the word frequency, each word in these two sentences is not necessarily a high-frequency word, making it difficult to uncover the topic of the text.

The LDA topic model is a text representation model that takes semantic association into account. The LDA topic model considers that an article has multiple topics and the words in the text have a certain probability of belonging to these topics.

The key to applying the LDA topic model is to determine the optimal number of topics. The effect of topic extraction in the LDA topic model is directly related to the number of potential topics, and the two most common evaluation methods to determine the number of topics are based on coherence [42–44] or perplexity [45]. Coherence refers to the quantitative calculation of whether the semantic association of words under a topic generated by an LDA is closer, and the formula for calculating topic coherence is shown in Equation (4).

$$coherence(T) = \sum_{(v_i, v_j) \in T} p(v_i, v_j) \tag{4}$$

where T is a topic, v_i and v_j are the words within the topic, and $p(v_i, v_j)$ is the scoring function to measure the semantic closeness of the words within the topic.

In practice, the function $p(v_i, v_j)$ is generally used in the UMass algorithm [43], with the formula shown in Equation (5).

$$p(v_i, v_j) = \ln \frac{p(v_i, v_j) + \varepsilon}{p(v_i) p(v_j)} \tag{5}$$

where $p(v_i, v_j)$ denotes the co-occurrence probability of words v_i and v_j, and $p(v_i)$ and $p(v_j)$ denote the occurrence probability of words v_i and words v_j, respectively, and ε is a smaller constant.

Perplexity can be understood as how uncertain the trained model is about the topic to which article d belongs. This degree of uncertainty is perplexity, which is calculated as in Equation (6).

$$perplexity(D_{test}) = \exp\left\{ -\frac{\sum\limits_{d=1}^{M} \log p(w_d)}{\sum\limits_{d=1}^{M} N_d} \right\} \tag{6}$$

In the formula, M represents the number of texts in the test set, N_d denotes the number of words in the dth text, and $p(w_d)$ represents the probability of occurrence of each word in the dth text.

In summary, this paper chose the number of topics under larger coherence and smaller perplexity to complete the construction of the LDA topic model.

3.4. Valence Aware Dictionary and sEntiment Reasoner Sentiment Analysis

Sentiment analysis is a subfield of natural language processing that aims to extract attitudinal dispositions from text. The three most common methods of sentiment analysis are plain Bayesian, Long Short-Term Memory (LSTM), and VADER methods [46–51]. This paper uses the powerful VADER method, which uses a comprehensive, high-quality sentiment vocabulary and complex linguistic rules to generate sentiment scores [52]. In addition to the determination of sentiment words, sentiment intensity is also measured, mainly based on punctuation, degree adverbs, negation, and conjunctions. For example, in the sentence "I really can't recommend it.", "recommend" is the positive sentiment word, while the previous word is the negative word, "can't", so the emotional value of "can't recommend" is -1.11 (-0.74×1.5). The previous word is the degree word, "really". According to the principle that the degree word that is farther away from the emotion word is assigned a lower score, the assignment of really is $0.293 \times 0.95 = 0.27835$. The sentiment score of the whole sentence is $-0.27835 - 0.74 \times 1.5 = -1.38835$. A detailed introduction and code can be found at https://github.com/cjhutto/vaderSentiment (accessed on 10 July 2022).

4. Empirical Research

4.1. Data Collection

This paper used Twitter (https://twitter.com/home, accessed on 5 June 2022) as the original data source for the study, and crawled tweets with "Beijing" or "Peking" between 1 January 2017 and 31 December 2021 through Python language programming. The total number of tweets with "Beijing" or "Peking" between 1 January 2017, and 31 December 2021, was 129,260, and the number of tweets per year is detailed in Table 1.

Table 1. Statistics on the number of tweets from 2017 to 2021.

Year	Number of Tweets
2017	25,806
2018	26,258
2019	26,334
2020	26,113
2021	24,749
Total	129,260

4.2. Data Processing

Firstly, this paper converts the German language in the tweets into English. Since the same two words "Beijing" and "Peking" are also used in German, the authors further used the Language Identification (LangID) module in Python to identify the German language text of each comment. Since the comment content has a small number of words and machine translation can reach a high level of accuracy, this paper used Google Translate (https://translate.google.cn, accessed on 6–7 June 2022) combined with manual proofreading and revision to complete the translation between German and English.

Subsequently, this paper used the NLTK module in Python for word separation, and the collections module in Python for word frequency (number of occurrences), and eliminated words with no real meaning based on the word frequency statistics, such as, "also", "said", "via", etc. In the case of words that are supposed to be phrases but are split into individual words, phrases are written in the dictionary, e.g., "North Korea", "Hong Kong", "New York" etc., were considered. Finally, "Beijing" and "Peking" were excluded because all tweets contained one or the other, and this was of no practical significance for the subsequent study. After pre-processing, this paper performed a second word separation, followed by TF-IDF keywords statistics, and used the LDA topic model for topic-mining and sentiment analysis. The Python code involved in this paper was written and run using PyCharm software.

4.3. Keywords Statistics

In this paper, the TF-IDF value was calculated by the TF-IDF method introduced in Section 3.2, and the keywords were ranked according to the TF-IDF value for each year, while the number of keyword occurrences was counted according to the Word Frequency Statistics method introduced in Section 3.1.

Table 2 shows the top 50 keywords of 2017. "China" has the highest TF-IDF value and the most occurrences; except for the TF-IDF value of 2021, which is in second place, the TF-IDF value and occurrences were all in first place. This shows that Beijing, as the capital of China, is sometimes used as a pronoun for "China". Next, "duck" ranked second in terms of TF-IDF value and number of occurrences. Peking duck is a world-renowned Beijing dish, which has apparently become an iconic food in Beijing and even China in the eyes of friends around the world. In addition, we found many adjectives in the keywords, such as "new", "great", "good", "open", "friend". It is easy to see that the current global image of Beijing is dominated by positive adjectives, which are more related to the continuous strengthening of its international metropolis and expansions of its openness. The words "art", "glass" and "opera" were among the top 50 keywords for 2017, indicating that Beijing's traditional culture and arts are receiving attention from friends around the world. At the same time, many Twitter users expressed their desire to go to Beijing and eat Peking duck, so the keyword "want" ranked seventh in 2017.

According to the top 50 keywords of 2018 (Table 3), it is easy to see that the ranking of "trade" improved, moving from 44th place in 2017 to 7th in 2018, and "tariff", "Washington", "government", "dispute" appeared in the top 50 keywords for 2018. At this point, Beijing was more of a referent for China, and the U.S. government's imposition of tariffs on Chinese products sparked global concern. "duck", "art", "opera", and "glass" remained in the top 50 keywords. The adjectives in the top 50 keywords are also all positive words. The top 50 keywords of 2019 (Table 4) feature 42 keywords from the top 50 keywords of 2018, and generally maintain the same focus as people had in 2018.

The coronavirus epidemic swept the world in 2020, and the 50 keywords of 2020 (Table 5) show that people paid more attention to the epidemic, such as "coronavirus", "virus", "corona", "pandemic", "COVID". The ranking of "duck" dropped, "art", "opera", and "glass" no longer appeared in the top 50 keywords, and the number of positive adjectives decreased.

Table 2. Top 50 keywords of 2017.

No.	Keywords	TF-IDF Value	Occurrences	No.	Keywords	TF-IDF Value	Occurrences
1	China	0.0323	4645	26	art	0.0046	406
2	duck	0.0246	1921	27	video	0.0045	325
3	Chinese	0.0139	1536	28	best	0.0044	308
4	new	0.0122	1205	29	trip	0.0044	299
5	airport	0.0090	548	30	open	0.0044	320
6	Trump	0.0086	748	31	glass	0.0044	406
7	want	0.0081	671	32	Berlin	0.0042	290
8	world	0.0081	754	33	visit	0.0042	322
9	first	0.0071	597	34	Nanjing	0.0042	203
10	great	0.0069	536	35	president	0.0042	345
11	North Korea	0.0067	561	36	car	0.0041	334
12	flight	0.0067	520	37	opera	0.0040	294
13	game	0.0066	484	38	photo	0.0040	271
14	university	0.0066	579	39	New York	0.0040	304
15	night	0.0064	380	40	friend	0.0038	278
16	city	0.0061	458	41	course	0.0038	134
17	train	0.0060	373	42	party	0.0038	298
18	Peking Duk	0.0059	430	43	show	0.0038	281
19	Shanghai	0.0058	449	44	trade	0.0037	281
20	good	0.0056	397	45	home	0.0034	231
21	Moscow	0.0055	412	46	foreign	0.0034	293
22	people	0.0052	457	47	capital	0.0034	246
23	air	0.0051	384	48	Xi	0.0034	264
24	smog	0.0047	317	49	tour	0.0034	222
25	German	0.0047	332	50	congress	0.0034	239

Table 3. Top 50 keywords of 2018.

No.	Keywords	TF-IDF Value	Occurrences	No.	Keywords	TF-IDF Value	Occurrences
1	China	0.0322	6003	26	country	0.0040	456
2	Chinese	0.0163	2467	27	war	0.0049	452
3	duck	0.0243	2209	28	trip	0.0046	447
4	new	0.0132	1686	29	art	0.0045	442
5	university	0.0107	1308	30	right	0.0043	422
6	world	0.0097	1213	31	government	0.0040	418
7	trade	0.0097	1148	32	Washington	0.0047	415
8	first	0.0085	925	33	dispute	0.0043	414
9	Trump	0.0077	865	34	work	0.0041	412
10	great	0.0077	798	35	minister	0.0038	411
11	president	0.0067	789	36	best	0.0047	404
12	people	0.0064	747	37	foreign	0.0038	399
13	want	0.0071	707	38	Xi	0.0037	395
14	Shanghai	0.0065	704	39	national	0.0038	391
15	tariff	0.0062	685	40	team	0.0042	385
16	city	0.0063	673	41	opera	0.0045	379
17	good	0.0069	669	42	company	0.0036	378
18	visit	0.0059	624	43	event	0.0038	374
19	show	0.0062	621	44	glass	0.0038	371
20	airport	0.0070	545	45	open	0.0040	367
21	game	0.0057	540	46	tour	0.0039	367
22	international	0.0049	532	47	car	0.0038	366
23	flight	0.0054	509	48	night	0.0041	365
24	state	0.0048	504	49	Moscow	0.0040	364
25	student	0.0046	492	50	meeting	0.0036	361

Table 4. Top 50 keywords of 2019.

No.	Keywords	TF-IDF Value	Occurrences	No.	Keywords	TF-IDF Value	Occurrences
1	China	0.0322	6199	26	Berlin	0.0049	413
2	duck	0.0196	1634	27	show	0.0049	471
3	Chinese	0.0168	2681	28	flight	0.0048	446
4	new	0.0145	1877	29	student	0.0047	519
5	airport	0.0104	992	30	war	0.0047	457
6	world	0.0104	1331	31	German	0.0046	426
7	trade	0.0103	1233	32	Moscow	0.0044	415
8	university	0.0099	1198	33	Germany	0.0044	432
9	people	0.0085	1068	34	visit	0.0043	450
10	want	0.0083	888	35	photo	0.0043	361
11	first	0.0078	902	36	work	0.0042	417
12	city	0.0072	788	37	tariff	0.0042	435
13	government	0.0068	789	38	dispute	0.0042	391
14	right	0.0066	746	39	team	0.0042	402
15	Trump	0.0065	720	40	video	0.0040	334
16	good	0.0064	620	41	open	0.0040	342
17	great	0.0064	655	42	minister	0.0039	431
18	Shanghai	0.0063	647	43	night	0.0039	302
19	international	0.0062	692	44	deal	0.0039	385
20	president	0.0058	694	45	USA	0.0038	384
21	country	0.0055	644	46	foreign	0.0038	412
22	Washington	0.0052	497	47	game	0.0038	357
23	company	0.0050	522	48	trip	0.0037	342
24	state	0.0050	525	49	business	0.0037	339
25	event	0.0049	436	50	meeting	0.0037	369

Table 5. Top 50 keywords of 2020.

No.	Keywords	TF-IDF Value	Occurrences	No.	Keywords	TF-IDF Value	Occurrences
1	China	0.0349	7558	26	Germany	0.0055	565
2	Chinese	0.0179	3036	27	Washington	0.0054	550
3	new	0.0150	2131	28	national	0.0054	633
4	duck	0.0148	1098	29	good	0.0053	525
5	world	0.0116	1552	30	president	0.0053	588
6	people	0.0109	1508	31	Shanghai	0.0053	541
7	government	0.0089	1147	32	pandemic	0.0051	546
8	want	0.0087	1014	33	Wuhan	0.0051	570
9	coronavirus	0.0086	974	34	Germany	0.0051	520
10	right	0.0081	969	35	Moscow	0.0051	457
11	law	0.0080	985	36	flight	0.0049	497
12	Trump	0.0079	935	37	party	0.0048	494
13	city	0.0079	939	38	million	0.0048	537
14	first	0.0077	918	39	official	0.0046	520
15	corona	0.0076	788	40	global	0.0046	514
16	state	0.0073	857	41	crisis	0.0045	467
17	country	0.0071	906	42	company	0.0044	491
18	security	0.0070	834	43	foreign	0.0044	511
19	virus	0.0066	727	44	EU	0.0044	412
20	Biden	0.0059	496	45	human	0.0044	453
21	case	0.0058	678	46	COVID	0.0043	111
22	outbreak	0.0058	610	47	work	0.0043	450
23	university	0.0057	638	48	medium	0.0043	452
24	show	0.0056	604	49	great	0.0043	407
25	international	0.0056	630	50	life	0.0042	391

In 2021, as the epidemic improved, epidemic-related words no longer appeared in the top 50 keywords (Table 6). “duck” moved to first place, “opera” moved to fifth place, and "brooch" appeared in the top 50 keywords for the first time. Beijing will be the host city of the 24th Winter Olympic Games in 2022, and "Olympics" and "winter" were among the top 10 keywords. At the same time, more positive words, such as "love" and "friend", appeared in the top 50 keywords.

Table 6. Top 50 keywords of 2021.

No.	Keywords	TF-IDF Value	Occurrences	No.	Keywords	TF-IDF Value	Occurrences
1	duck	0.0446	3829	26	great	0.0046	424
2	China	0.0269	4172	27	author	0.0046	237
3	Chinese	0.0180	2634	28	beat	0.0046	236
4	university	0.0143	1748	29	human	0.0045	473
5	opera	0.0134	1441	30	brooch	0.0045	337
6	Olympics	0.0128	1292	31	government	0.0044	472
7	new	0.0111	1246	32	Peking Duk	0.0044	256
8	winter	0.0089	839	33	school	0.0043	455
9	world	0.0081	949	34	encyclopedia	0.0043	196
10	right	0.0080	848	35	national	0.0042	444
11	first	0.0073	778	36	team	0.0041	378
12	people	0.0072	812	37	house	0.0041	304
13	good	0.0072	584	38	official	0.0041	407
14	want	0.0068	514	39	restaurant	0.0039	298
15	game	0.0068	638	40	happy	0.0039	282
16	Biden	0.0063	591	41	international	0.0039	373
17	Olympic	0.0055	511	42	old	0.0038	296
18	city	0.0052	510	43	night	0.0038	291
19	love	0.0051	392	44	photo	0.0038	294
20	food	0.0051	410	45	show	0.0038	359
21	country	0.0050	551	46	eat	0.0037	207
22	site	0.0050	315	47	friend	0.0036	295
23	best	0.0049	358	48	party	0.0035	331
24	life	0.0048	317	49	work	0.0035	336
25	state	0.0047	446	50	dinner	0.0035	231

This paper counted the occurrence of words that appeared in the top 50 keywords in the five-year period from 2017 to 2021, forming Figure 1, with 13 words in total. Among them, "China" and "Chinese" achieved a significant decrease in 2021, despite their increasing number in 2017–2020, which could indicate that Beijing has improved its national image and is more representative of its own city image. "duck" has become a famous dish in Beijing and is known worldwide. This was especially noticeable in 2021, when the epidemic improved, and the number of occurrences of the word "duck" significantly increased. "want" expresses some of the thoughts and aspirations of the global friends regarding Beijing, and "first" was more frequently used in tweets regarding Beijing. The three high-frequency adjectives, "new", "good" and "great", which were all found over the five-year period, show the positive image of Beijing. In 2021, the occurrence of the words "good" and "great" increased, but use of the word "new" decreased compared to the previous year. As Beijing has many famous universities, the word "university" is widely mentioned in tweets. By checking the original tweets, the authors found that many tweeters are studying at universities in Beijing, or hope to come to study at universities in Beijing, and some tweeters have left footprints in university libraries.

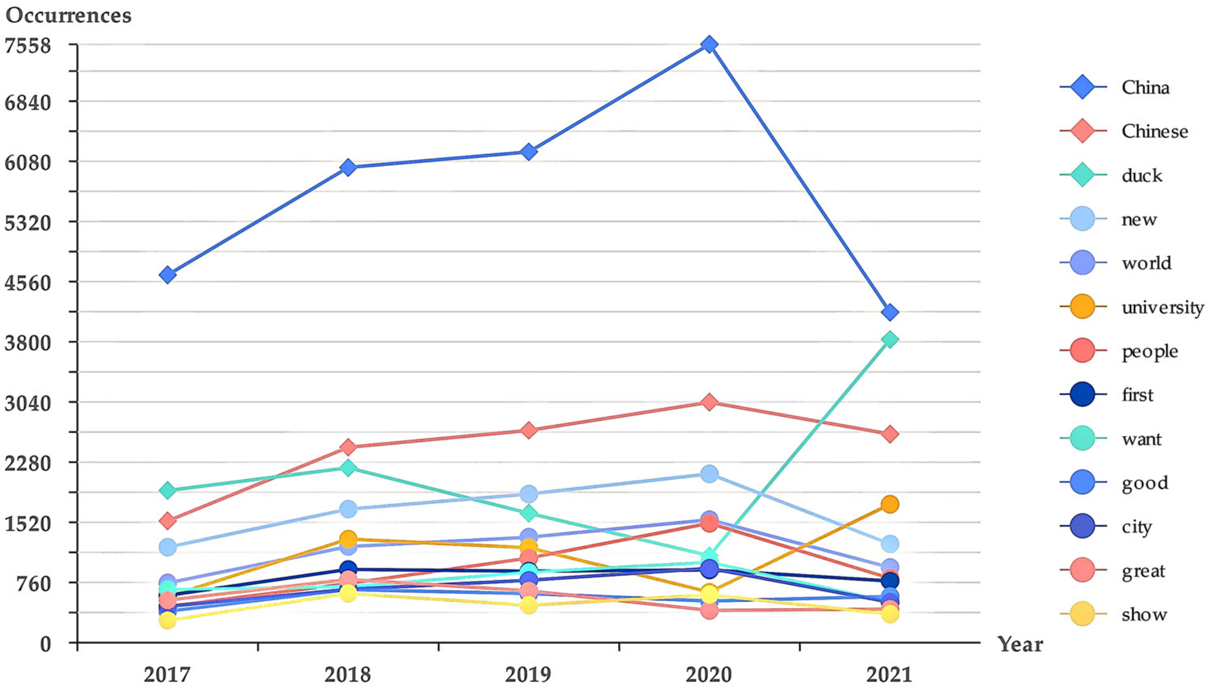


Figure 1. Trends of keywords occurring in all 5 years.

4.4. Topic Mining

Appendix A shows the coherence and perplexity of the LDA topic model for a different number of topics each year from 2017 to 2021, and this paper determined the number of topics for each year based on a larger coherence and smaller perplexity, and finally determined the number of topics for 2017–2021 as 3, 3, 4, 5, and 3, respectively, and Tables 7–11 shows the results of the LDA topic model for 2017–2021.

Table 7. LDA topic model results of 2017.

Topic1 Words	Probability Distribution	Topic2 Words	Probability Distribution	Topic3 Words	Probability Distribution
new	0.0226	duck	0.0838	China	0.0947
university	0.0225	China	0.0341	Chinese	0.0254
friend	0.0194	airport	0.0208	New York	0.0202
first	0.0183	glass	0.0200	Peking Duk	0.0188
flight	0.0155	Trump	0.0185	great	0.0186
Chinese	0.0144	world	0.0174	want	0.0133
air	0.0129	North Korea	0.0169	new	0.0112
good	0.0111	Chinese	0.0160	Barcelona	0.0098
international	0.0109	art	0.0147	Moscow	0.0098
happy	0.0107	city	0.0134	London	0.0098
photo	0.0105	opera	0.0115	trip	0.0095
best	0.0100	Moscow	0.0112	car	0.0091
gold	0.0089	right	0.0105	Shanghai	0.0090
open	0.0088	green	0.0105	old	0.0074
school	0.0080	game	0.0104	video	0.0072

Table 8. LDA topic model results of 2018.

Topic1 Words	Probability Distribution	Topic2 Words	Probability Distribution	Topic3 Words	Probability Distribution
duck	0.0383	China	0.0661	China	0.0204
China	0.0275	trade	0.0217	Shanghai	0.0145
university	0.0196	Trump	0.0169	airport	0.0128
Chinese	0.0167	Chinese	0.0164	museum	0.0127
new	0.0119	president	0.0139	world	0.0114
great	0.0097	new	0.0130	Chinese	0.0107
first	0.0097	tariff	0.0121	want	0.0099
show	0.0094	people	0.0094	right	0.0094
world	0.0088	world	0.0087	New York	0.0082
national	0.0085	war	0.0086	million	0.0080
tour	0.0082	state	0.0084	Germany	0.0079
student	0.0075	visit	0.0084	flight	0.0078
art	0.0063	Washington	0.0083	car	0.0078
best	0.0062	Moscow	0.0081	game	0.0076
opera	0.0055	air	0.0076	largest	0.0075

Table 9. LDA topic model results of 2019.

Topic1 Words	Probability Distribution	Topic2 Words	Probability Distribution	Topic3 Words	Probability Distribution	Topic4 Words	Probability Distribution
China	0.0165	China	0.0270	China	0.0773	duck	0.0599
world	0.0144	Chinese	0.0245	trade	0.0254	airport	0.0334
city	0.0143	university	0.0156	Chinese	0.0153	new	0.0281
event	0.0124	government	0.0099	Trump	0.0141	China	0.0153
Shanghai	0.0114	people	0.0089	want	0.0135	Chinese	0.0152
game	0.0099	school	0.0067	new	0.0114	world	0.0117
great	0.0097	new	0.0063	president	0.0103	open	0.0096
air	0.0090	student	0.0062	war	0.0101	largest	0.0096
fan	0.0082	country	0.0062	Washington	0.0093	international	0.0089
photo	0.0079	police	0.0058	tariff	0.0084	flight	0.0074
night	0.0076	first	0.0057	right	0.0084	glass	0.0073
first	0.0073	good	0.0056	deal	0.0081	car	0.0068
show	0.0073	city	0.0056	USA	0.0078	food	0.0068
video	0.0068	world	0.0055	minister	0.0078	first	0.0066
winter	0.0067	company	0.0052	people	0.0076	film	0.0063

In the results presented in 2017 (Table 7), Topic1 is Beijing school culture, which includes the words "university", "school" and some words describing school or school life. Topic2 is Beijing traditional culture, which includes Beijing traditional food "duck", as well as "art", "glass", Topic3 is closely related cities, and the cities closely related to Beijing are "New York", "Barcelona ", "Moscow", "London", "Shanghai", and many of the tweeters also traveled between these cities and Beijing.

In the results presented in 2018 (Table 8), Topic1 shows the characteristics of the combination of Topic1 and Topic2 themes in 2017, mainly consisting of "duck", "art", "opera" and traditional culture and art, as well as "university" and "student" of campus culture, which can be said to be a wide range of Beijing cultural topics. In 2018, Topic2 is the trade friction, including "trade", "Trump", "tariff", "Washington". Topic3 presents the tourism, including some cities with close ties to Beijing and the words "museum", "airport ", "flight", "car" and other tourist attractions and travel tools.

Table 10. LDA topic model results of 2020.

Topic1 Words	Probability Distribution	Topic2 Words	Probability Distribution	Topic3 Words	Probability Distribution
China	0.0613	duck	0.0420	China	0.0455
Trump	0.0205	Chinese	0.0287	new	0.0280
Chinese	0.0176	China	0.0285	coronavirus	0.0179
want	0.0172	university	0.0116	corona	0.0168
world	0.0148	new	0.0095	case	0.0159
Biden	0.0131	company	0.0084	virus	0.0137
Washington	0.0109	Silk Road	0.0079	outbreak	0.0131
trade	0.0106	good	0.0079	Wuhan	0.0120
president	0.0103	park	0.0078	million	0.0111
war	0.0085	want	0.0071	woman	0.0109
deal	0.0084	great	0.0068	people	0.0106
Moscow	0.0080	restaurant	0.0065	Chinese	0.0105
government	0.0073	love	0.0063	city	0.0094
medium	0.0071	wall	0.0061	pandemic	0.0092
election	0.0065	video	0.0061	first	0.0080
Topic4 words	**Probability distribution**	**Topic5 words**	**Probability distribution**		
China	0.0380	China	0.0181		
law	0.0194	summer	0.0140		
Chinese	0.0181	world	0.0132		
right	0.0176	city	0.0124		
security	0.0151	opera	0.0118		
people	0.0114	first	0.0115		
government	0.0110	palace	0.0111		
national	0.0098	game	0.0104		
human	0.0098	winter	0.0101		
state	0.0092	Shanghai	0.0093		
new	0.0073	Olympics	0.0091		
international	0.0066	photo	0.0084		
freedom	0.0065	flight	0.0084		
country	0.0061	people	0.0081		
HK	0.0054	airport	0.0081		

Table 11. LDA topic model results of 2021.

Topic1 Words	Probability Distribution	Topic2 Words	Probability Distribution	Topic3 Words	Probability Distribution
duck	0.0951	Olympics	0.0328	China	0.0393
opera	0.0438	China	0.0274	university	0.0276
Chinese	0.0215	winter	0.0212	Chinese	0.0212
food	0.0105	game	0.0160	new	0.0139
glass	0.0096	right	0.0146	school	0.0119
defend	0.0091	Olympic	0.0141	people	0.0111
art	0.0090	Chinese	0.0105	world	0.0080
love	0.0088	growing	0.0094	great	0.0066
green	0.0078	first	0.0089	quite	0.0065
restaurant	0.0076	human	0.0088	Chan	0.0062
good	0.0075	world	0.0079	national	0.0060
vintage	0.0075	state	0.0071	city	0.0056
want	0.0071	team	0.0070	student	0.0056
brooch	0.0070	summer	0.0067	house	0.0053
old	0.0068	president	0.0066	friend	0.0049

In the results presented in 2019 (Table 9), Topic1 contains words such as “event”, “game”, “winter” and other words closely related to winter sports. After reviewing the content of the original tweets, these words were found more often found in tweets expressing expectations for the Beijing Winter Olympics. Topic2 is the campus culture, including "university", "school" and "student" and "student". Topic3 presents the same topic as Topic2 in 2018: the trade friction. Topic4 is mainly composed of the words "duck ", "glass", "food", "film", and other words that form the cultural topic.

In the results presented in 2020 (Table 10), in addition to some words representing trade friction, words such as "Biden" and "election" also appear in Topic1, so it can be said that Topic1 focuses more on the Sino–US relations. Topic2 has words such as "duck" and "restaurant", making the topic more focused on the food culture. Topic3 focuses on the epidemic, including the words "coronavirus", "virus", "corona", "outbreak ", "Wuhan", "pandemic" and other related words. Topic4 focuses on the violent impact of Hong Kong society, mainly involving "law ", "security", "HK" and other words. Topic5 is Beijing's traditional culture and sports, including "opera ", "palace", "game", "winter", "Olympics".

In the results presented in 2021 (Table 11), Topic1 focuses on Beijing's food culture and traditional arts culture, with the main categories being "duck", "opera", "food ", "glass", "art", "restaurant", "vintage", "brooch". Topic2 is the Winter Olympics, mainly including "Olympics", "winter ", "game", "Olympic" and other words. Topic3 is campus culture, including "university ", "school", "student".

4.5. Sentiment Analysis

The adjectives in the keyword statistics, such as "new", "great", "good", "open", "international", and "friend", help us to understand people's emotional perception of Beijing. Most of them are adjectives that express positive emotional tendencies.

Through the VADER sentiment analysis method, this paper further quantified the sentiment values of Beijing city at various timepoints. Figure 2 depicts the trends in sentiment value from January to December 2017–2021, and found two timepoints showed significant decreases in August 2019 and April 2020. When the image of Beijing as the capital of China was somewhat affected by the violent social shock in Hong Kong in August 2019, and the global outbreak of the coronavirus outbreak in April 2020, some global friends changed their emotional perception of China and the city of Beijing within a short period of time. The emotional perception of a city is closely related to major events.

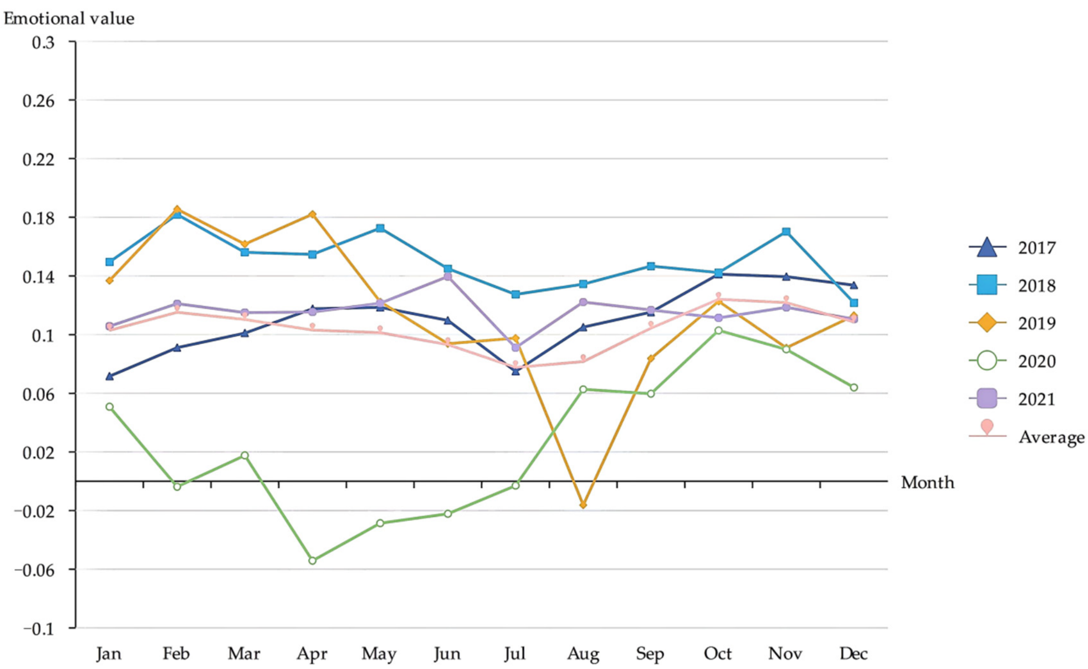


Figure 2. Trends in sentiment value from January to December 2017–2021.

Compared to the mean value of 5-year sentiment, the majority of months in 2017, 2018, 2019 and 2021 showed sentiment values above the mean, with the exception of 2019 and 2020, where the sentiment trend line fluctuated somewhat, and 2017, 2018 and 2021 show relatively stable monthly changes. The 5-year sentiment mean line has mean values below 0.1000 in June, July and August, and above 0.1000 in all other months, and above the mean values in October and November. After further calculation, the annual sentiment values for 2017–2021 are all above 0, showing a positive sentiment tendency, with the annual sentiment values for 2017, 2018, 2019 and 2021 being above 0.1000.

5. Further Discussion

5.1. Analysis of the Characteristics of the International Image of Beijing City

Beijing presents a kind of ancient capital charm. The city has a long history and is a humanistic city where cultural preservation and development go hand in hand. "palace", "museum", and "wall" appear in keywords and the LDA topic model.

Secondly, tweets about Beijing reveal a certain cultural atmosphere. On the one hand, this is a traditional culture, with words such as "opera", "art", "glass", "vintage", "brooch" and so on becoming keywords. On the other hand, the campus culture, with words such as "university", "school", "student", and the many famous universities in Beijing, make the city's campus culture strong.

Third, "duck" was mentioned many times in tweets, and appeared in the top 10 keywords every year. Peking duck has become a famous Beijing specialty and a food card of Beijing.

Fourthly, Beijing, as the capital of China, has many opportunities to host major international events, which gives it a huge advantage in international communications. As the host city of the 2022 Winter Olympic Games, Beijing is highly anticipated worldwide.

Fifthly, through the adjectives included in keywords, we can describe the international image of Beijing city more clearly. Words such as "new", "great", "good", "open", "international" and "friend" show that Beijing is gradually becoming a modern and civilized international city, and that the friendliness and openness of Beijing attracts people from all over the world.

Sixthly, Beijing, as an international metropolis, has close ties with other cities, "New York", "Barcelona", "Moscow", "London", "Shanghai" appear in the high-frequency vocabulary many times, and tweeters repeatedly say that they travel between these cities and Beijing.

Finally, the paper also notes that the international city image of Beijing is vulnerable to national events, especially controversial events that can cause some people around the world to change their emotional disposition toward Beijing within a short period of time.

5.2. Strategy for Beijing International Image Communication

First, the city image is part of the national image, and the city image of Beijing is highly linked to its national and governmental image. Therefore, the communication strategy for Beijing should be an important part of a national policy, rather than a propaganda piece or a stand-alone project. The communication strategy regarding city image should focus on the creation of a clearer national consensus about Beijing's positioning, making the nation understand that a country's reputation is the property of its people, and encouraging and practicing a culture of creativity across government, culture, business, investment, education, industry, and other areas. Only by standing as a nation and encouraging creative culture, so that new ideas continue to emerge in every field, can we continue to correct and counteract existing stereotypes, reaching a situation where the image of the nation and the image of the city develop together and complement each other.

Among the adjectives in keywords, words such as "new" and "first", which are always associated with innovation, are often associated with positive content, making people feel positively about Beijing. Therefore, it is important to pay close attention to new concepts and achievements in various areas when promoting tourism in Beijing. It is also important to note that major national events that take place in Beijing will bring more direct tourism opportunities. The word "Olympics" jumped to the sixth keyword in 2021, and the Winter Olympics provided the city with the new title of the first city in the world to host both the Summer and Winter Olympic Games, bringing the new concept of "Green Olympics" and a new market of "skiing fever". Tourism organizations and practitioners should make full use of these new opportunities to innovate and promote tourism in Beijing and build a good international image of the city.

Secondly, the focus should be on promoting Beijing's image with regard to its culture, entertainment and appearance. According to data from the Beijing Municipal Cultural Heritage Bureau [53], Beijing has 138 national key cultural relics in the protection list, including the Summer Palace, the Great Wall, and the Palace Museum. It also has 144 national intangible cultural heritage representative projects [54], including the Beijing Opera, Kunqu Opera and others. There are also many modern landmarks, such as National Grand Theatre, the Water Cube and others. However, according to keywords, statistics and the results of LDA topic model, these rich and colorful cultural resources are far from being fully explored and displayed, and there is less relevant discussion. Media in China can create sub-accounts with different topic contents according to audience interests, make full use of pictures and videos to show more intuitive contents, explore the filmable resources of Beijing culture, and promote city tourism through storytelling.

Moreover, keywords statistics show that international attention to Beijing mostly occurs at the abstract level (representatives of the state and government, economic and social issues, etc.), ignoring its unique feeling as a physical space. In addition to the virtual image presented on the Internet, the physical image should be paid equal attention. The government should continue to improve and create urban humanistic landmarks, enrich the cultural and recreational life of residents, and improve transportation, roads and the ecological environment. According to the concepts of absorption and immersion in tourism management, people need to obtain immersion experiences through visual and auditory sensory perceptions, which could help them to form a deep and complex city image [55]. In short, the influence of Beijing's culture and urban landscape on its overall image should be continuously expanded, which will help to create a more positive and diverse image of Beijing and allow for people to discover its diverse charms.

Finally, openness to the outside world needs to be maintained and encouraged, and the tourism organizations in Beijing should cooperate with unofficial accounts. The shaping of an international city image is a cross-regional, cross-cultural and cross-linguistic external communication. Ordinary communication only needs to convert the original information into a message that is acceptable to the general public, while external communication needs to cross multiple barriers, such as language differences and cultural differences, which means that opinion leaders play an important role in shaping the international city image. Compared with general audiences, opinion leaders are more frequently exposed to media, and they influence audiences with lower levels of media exposure, knowledge and interest by providing information and conveying their views.

With the popularity of Internet applications, the role played by online opinion-leaders is receiving more and more attention. There are two types of information sources that can play the role of opinion leaders. One is the national media of foreign publics. The other is foreign publics with direct or indirect contact with Chinese cities. During the epidemic, the international tourism market was hit hard, which also made it difficult for the foreign public to gain direct knowledge of Chinese cities. There is a Chinese idiom that originates from a historical story: "Seeing is believing". In the post-epidemic era, the government should establish a more scientific and efficient public health policy, promote international tourism recovery and development, organize study-abroad activities between universities, and establish international friendship cities, ultimately achieving image improvement and renewal.

In addition, the role of local citizens in spreading the city image should not be ignored. According to the Beijing Urban Master Plan (2016–2035) released by the Beijing government in 2017 [56], the blueprint for the future development of Beijing is to make Beijing an international, first-class, harmonious and livable capital. When residents feel that the city is livable, they will naturally be proud to promote their city to the outside world, ultimately creating an attractive urban atmosphere in Beijing. The official media and related institutions should actively cooperate with domestic and foreign unofficial media and cooperate with the agenda-setting function of traditional media, i.e., traditional media influences the focus of public attention and the perception of the social environment by giving prominence to various topics, and continuously carrying out data tracking and analysis.

5.3. Discussion

This paper searched papers with the topic “city image” through the Web of Science (https://www.webofscience.com/wos/woscc/basic-search, accessed on 12 August 2022). After browsing and sorting, we found that most studies focused on the city image in the country, but fewer studies studied the international image of cities. In addition, most studies on the international image of cities were conducted in a qualitative way, without sufficient data support.

Subsequently, this paper set the topic as “city image”, “Beijing” and “Twitter”, and found only six papers related to this topic. Among them, two papers used social media data to map human activities, and one paper proposed a new paradigm for urban residential sensitivity to heatwave risks based on social media Big Data. All three papers were in the field of remote sensing. In addition, one paper examined how the Burberry brand has become a major trendsetter in social media marketing. One paper examined how the Beijing Museum uses social media tools as a marketing strategy to increase its visibility among the public. Another paper proposed a fine-grained spatiotemporal dynamic pattern analysis approach and applied this to a Flickr dataset from Beijing: the paper shows that the method can also be applied to Twitter data.

In summary, international city image studies lack case studies; notably, there are few studies on the international image of Beijing. The analysis is mostly qualitative, and there are few studies on the international image of cities based on international social media platforms using text-analysis techniques. This paper further expands the research in related fields and enriches the research results. It also makes some contributions in terms of methodology, theory and practical application, including the following three aspects:

1. From the methodological perspective, this paper further mines the international image of Beijing based on Twitter comments, using keyword extraction through the TF-IDF method, topic analysis through the LDA model and sentiment analysis through the VADER method. The combination of the three methods allowed for a more comprehensive analysis of Twitter users’ perception of Beijing’s city image. A social media platform text-mining framework was formed for city-image research.
2. From the theoretical perspective, this paper extended the research related to the international image of cities and city tourism competitiveness, emphasizing that tourism competitiveness is enhanced through the city image, which is not limited to the city tourism image but includes a more comprehensive international image of cities, including economic trade, sports events, film and entertainment, and social issues.
3. From the perspective of practical application, Beijing is currently in the promotional period of international communication center construction, and better shaping the international image of the city plays a crucial role in the future development of Beijing. This study provides powerful data support and communication strategies to allow for relevant departments to make decisions, which, in turn, helps Beijing to establish a good international image.

6. Conclusions

This paper collected tweets about “Beijing” (“Peking”) on Twitter from 2017 to 2021 through web-crawler technology, and analyzed the international city image presented through TF-IDF keywords statistics, LDA topic mining, and VADER sentiment analysis.

Based on the analysis results, this paper proposed a strategy to communicate the international city image, further stimulate tourism's vitality and enhance the city's competitiveness in terms of tourism.

According to the research analysis, the international city image of Beijing in Twitter presents these characteristics: (1) Beijing has some cultural connotations, but, as an ancient capital with a long history, there are still many cultural resources that are not fully displayed. (2) Beijing, as the host city of the 2022 Winter Olympic Games, is attracting the attention of people around the world. (3) Beijing, as an international metropolis, has close ties with other cities. Many keywords also show that Beijing is gradually becoming a modern and civilized international city, and its friendliness and openness attract people from all over the world. (4) As Beijing is the capital of China, the world people's emotional tendency towards the city is easily influenced by the national events in China.

Through an analysis of the international city image of Beijing and the presented characteristics, the corresponding strategies were further proposed based on communication, including the following steps: (1) encouraging creative culture and making use of major international activities and events to build a good image; (2) deeply exploring and promoting the content of city image focusing on culture, entertainment and city style, so that people can discover the diverse charm of Beijing; (3) insisting on and encouraging openness to the outside world and actively developing cooperation with domestic and foreign unofficial media.

The main significance of this study is as follows. (1) It further enriches the research perspective of urban tourism competitiveness. The international city image of Beijing was analyzed through user comments on the mainstream social media platform Twitter. Based on the analysis, image enhancement strategies to attract more tourists and improve urban tourism competitiveness were proposed. (2) The international city image was measured based on the TF-IDF method, LDA method and VADER method, which enriched the image perception research theory and method, and provided references for relevant research. (3) Taking Beijing as a case study site, powerful data support and communication strategies were provided to support governmental decision making, as well as references for other cities manage their international city image.

However, we should note the limitations of the current study further research directions. (1) The research in this paper mainly uses text data, and the original data types could be further expanded in the future to include pictures, videos, audios, etc. Furthermore, multiple data types could be studied from multiple perspectives. (2) Due to the limitation of the languages learned by the authors, it was not possible to proofread the English results obtained by the machine translation of tweets in other languages. From the perspective of data rigor, the tweets collected in this paper were mainly in English and German. The research results mainly reflect the countries and regions that use these languages, but do not reflect the perceptions of other countries and regions that speak other languages. Therefore, there is a need to expand the multilingual tweets in the future to present a more comprehensive international perception of Beijing's city image. (3) In addition, the authors believe that the prediction of keyword statistics, LDA topics, and sentimental analysis value based on machine learning methods can be further explored in the future, and the Empirical Mode Decomposition (EMD) could be applied to reduce noise in non-stationary and non-linear data. (4) Finally, the international image of Beijing as presented on a wider range of social media platforms should be a future research direction.

Author Contributions: Conceptualization, H.N. and Z.Z.; methodology, H.N. and Z.Z.; writing—original draft preparation, Z.Z., M.L. and Z.L.; writing—review and editing, Z.Z., M.L., Z.L. and H.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Key Project of Beijing Social Science Foundation for Decision-making and Consulting (No. 21JCB044).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data in this study can be obtained by contacting the author: bfsuniuhy@163.com.

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

Appendix A

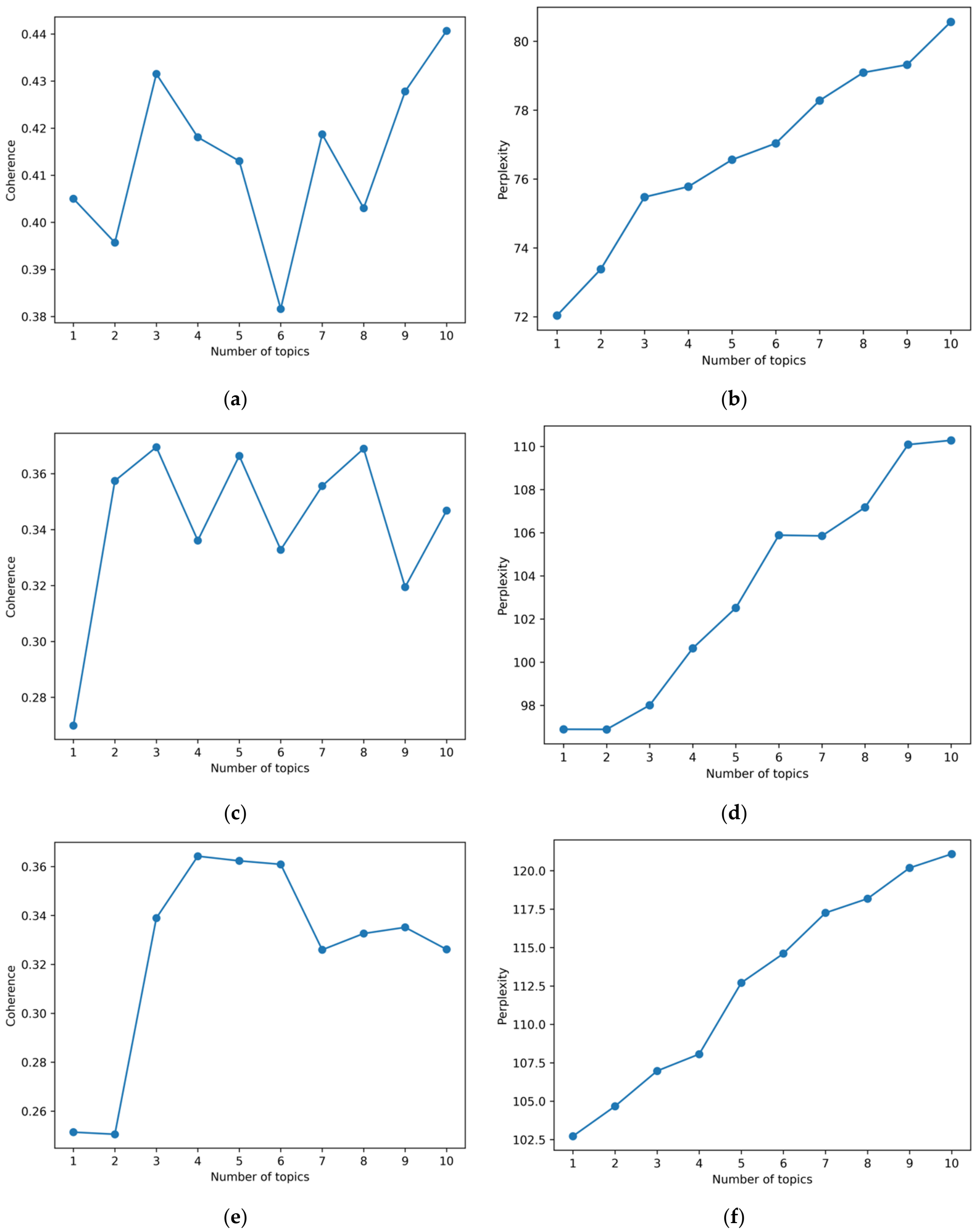
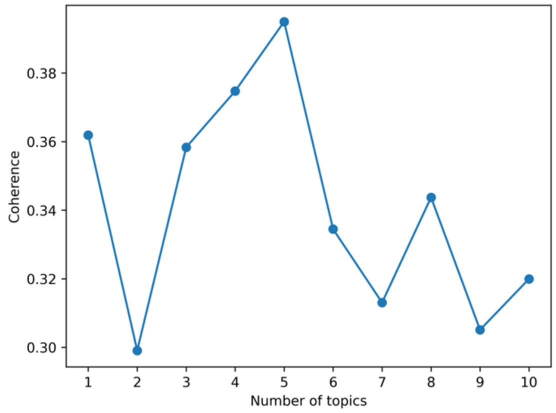
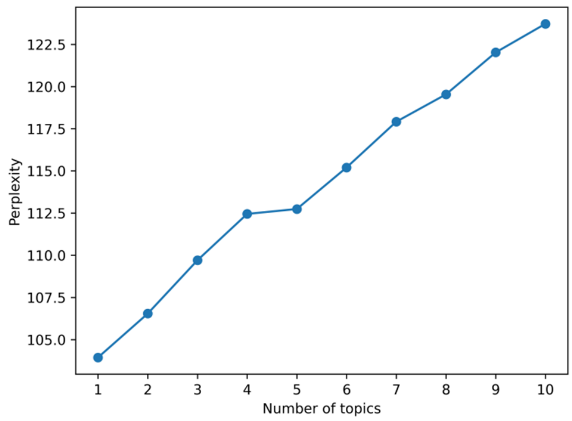


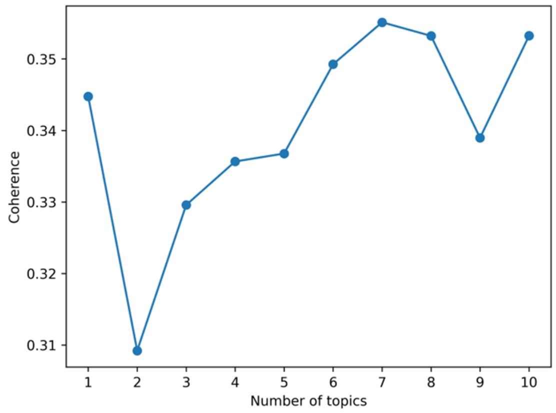
Figure A1. *Cont.*



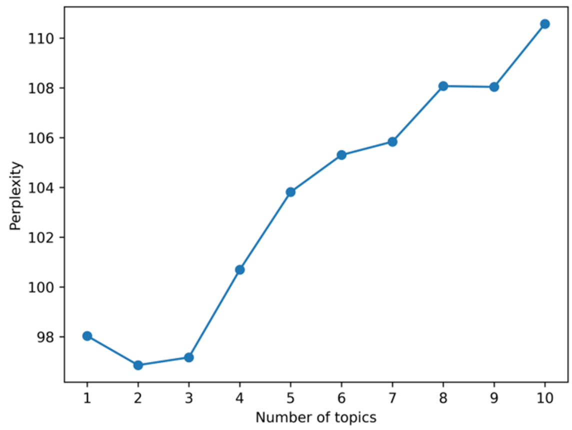
(**g**)



(**h**)



(**i**)



(**j**)

Figure A1. (**a**) Coherence of the 2017 LDA Topic Model; (**b**) Perplexity of the 2017 LDA Topic Model; (**c**) Coherence of the 2018 LDA Topic Model; (**d**) Perplexity of the 2018 LDA Topic Model; (**e**) Coherence of the 2019 LDA Topic Model; (**f**) Perplexity of the 2019 LDA Topic Model; (**g**) Coherence of the 2020 LDA Topic Model; (**h**) Perplexity of the 2020 LDA Topic Model; (**i**) Coherence of the 2021 LDA Topic Model; (**j**) Perplexity of the 2021 LDA Topic Model.

References

1. Boivin, M.; Tanguay, G.A. Analysis of the determinants of urban tourism attractiveness: The case of Québec City and Bordeaux. *J. Destin. Mark. Manag.* **2019**, *11*, 67–79. [CrossRef]
2. Rodríguez-Díaz, B.; Pulido-Fernández, J.I. Sustainability as a key factor in tourism competitiveness: A global analysis. *Sustainability* **2019**, *12*, 51. [CrossRef]
3. Khanzadeh, H.; Rahnavard, F.; Bamdad, N.; Mahmoudzadeh, S.M. Explanation Model of Tourism Competitiveness in Iran Touristic Cities. *New Mark. Res. J.* **2021**, *11*, 45–68.
4. Jia, Z.; Jiao, Y.; Zhang, W.; Chen, Z. Rural tourism competitiveness and development mode, a case study from Chinese township scale using integrated multi-source data. *Sustainability* **2022**, *14*, 4147. [CrossRef]
5. Sirkis, G.; Regalado-Pezúa, O.; Carvache-Franco, O.; Carvache-Franco, W. The Determining Factors of Attractiveness in Urban Tourism: A Study in Mexico City, Buenos Aires, Bogota, and Lima. *Sustainability* **2022**, *14*, 6900. [CrossRef]
6. Lynch, K. *The Image of the City*, 1st ed.; The MIT Press: Cambridge, MA, USA, 1964; pp. 46–90.
7. Hunt, J.D. Image as a Factor in Tourism Development. *J. Travel Res.* **1975**, *13*, 1–7. [CrossRef]
8. Crompton, J.L. An assessment of the image of Mexico as a vacation destination and the influence of geographical location upon the image. *J. Travel Res.* **1979**, *17*, 18–23. [CrossRef]
9. Embacher, J.; Buttle, F. A Repertory Grid Analysis of Austria's Image as A Summer Vacation Destination. *J. Travel Res.* **1989**, *27*, 3–7. [CrossRef]
10. Chon, K.S. The role of destination image in tourism: A review and discussion. *Tourist Rev.* **1990**, *45*, 2–9. [CrossRef]
11. Baloglu, S.; McCleary, K.W. A model of destination image formation. *Ann. Tourism Res.* **1999**, *26*, 868–897. [CrossRef]

12. Hong, S.; Kim, J.; Jang, H.; Lee, S. The roles of categorization, affective image and constraints on destination choice: An application of the NMNL model. *Tourism Manag.* **2006**, *27*, 750–761. [CrossRef]
13. Chen, X.; Li, J.; Han, W.; Liu, S. Urban Tourism Destination Image Perception Based on LDA Integrating Social Network and Emotion Analysis: The Example of Wuhan. *Sustainability* **2021**, *14*, 12. [CrossRef]
14. Foot, J.M. From boomtown to bribesville; the images of the city. Milan, 1980–1997. *Urban Hist.* **1999**, *26*, 393–412. [CrossRef]
15. Al-ghamdi, S.A.; Al-Harigi, F. Rethinking image of the city in the information age. *Procedia Comput. Sci.* **2015**, *65*, 734–743. [CrossRef]
16. Martí, P.; Serrano-Estrada, L.; Nolasco-Cirugeda, A. Social media data: Challenges, opportunities and limitations in urban studies. *Comput. Environ. Urban* **2019**, *74*, 161–174. [CrossRef]
17. Priporas, C.V.; Stylos, N.; Kamenidou, I.E. City image, city brand personality and generation Z residents' life satisfaction under economic crisis: Predictors of city-related social media engagement. *J. Bus. Res.* **2020**, *119*, 453–463. [CrossRef]
18. Lin, M.S.; Liang, Y.; Xue, J.X.; Pan, B.; Schroeder, A. Destination image through social media analytics and survey method. *Int. J. Contemp. Hosp. Manag.* **2021**, *33*, 2219–2238. [CrossRef]
19. Pan, X.; Rasouli, S.; Timmermans, H. Investigating tourist destination choice: Effect of destination image from social network members. *Tourism Manag.* **2021**, *83*, 104217. [CrossRef]
20. Huang, J.; Obracht-Prondzynska, H.; Kamrowska-Zaluska, D.; Sun, Y.; Li, L. The image of the City on social media: A comparative study using "Big Data" and "Small Data" methods in the Tri-City Region in Poland. *Landscape Urban Plan.* **2021**, *206*, 103977. [CrossRef]
21. Luque-Martinez, T.; Del Barrio-García, S.; Ibanez-Zapata, J.A.; Molina, M.Á.R. Modeling a city's image: The case of Granada. *Cities* **2007**, *24*, 335–352. [CrossRef]
22. Adamus-Matuszyńska, A.; Michnik, J.; Polok, G. A systemic approach to city image building. The case of Katowice city. *Sustainability* **2019**, *11*, 4470. [CrossRef]
23. Avraham, E. Media strategies for improving an unfavorable city image. *Cities* **2004**, *21*, 471–479. [CrossRef]
24. Alvarez, M.D. Creative cities and cultural spaces: New perspectives for city tourism. *Int. J. Cult. Tour. Hosp. Res.* **2010**, *4*, 5–171. [CrossRef]
25. Das, D.K. Engendering creative city image by using information communication technology in developing countries. *Urban Plan.* **2016**, *1*, 1–12. [CrossRef]
26. Lu, X. The Current Situation and Strategy of the International Communication of Shaanxi City Image. *J. Front. Soc. Sci. Technol.* **2021**, *1*, 69–73.
27. Novais, M.A.; Ruhanen, L.; Arcodia, C. Destination competitiveness: A phenomenographic study. *Tourism Manag.* **2018**, *64*, 324–334. [CrossRef]
28. Fernández, J.A.S.; Azevedo, P.S.; Martin, J.M.M.; Martin, J.A.R. Determinants of tourism destination competitiveness in the countries most visited by international tourists: Proposal of a synthetic index. *Tour. Manag. Perspect.* **2020**, *33*, 100582. [CrossRef]
29. Vinyals-Mirabent, S. European urban destinations' attractors at the frontier between competitiveness and a unique destination image. A benchmark study of communication practices. *J. Destin. Mark. Manag.* **2019**, *12*, 37–45. [CrossRef]
30. Kim, H.B.; Lee, S. Impacts of city personality and image on revisit intention. *Int. J. Tourism Cities* **2015**, *1*, 50–69. [CrossRef]
31. Bigne, J.E.; Sanchez, M.I.; Sanchez, J. Tourism image, evaluation variables and after purchase behaviour: Inter-relationship. *Tourism Manag.* **2001**, *22*, 607–616. [CrossRef]
32. Bongkoo, L.; Choongki, L.; Jaeseok, L. Dynamic nature of destination image and influence of tourist overall satisfaction on image modification. *J. Travel Res.* **2014**, *53*, 239–251.
33. Chen, C.F.; Tsai, D.C. How destination image and evaluative factors affect behavioral intentions. *Tourism Manag.* **2007**, *28*, 1115–1122. [CrossRef]
34. Nazarian, A.; Atkinson, P.; Foroudi, P. Influence of national culture and balanced organizational culture on the hotel industry's performance. *Int. J. Hosp. Manag.* **2017**, *63*, 22–32. [CrossRef]
35. Kumar, S.; Dhir, A. Associations between travel and tourism competitiveness and culture. *J. Destin. Mark. Manag.* **2020**, *18*, 100501. [CrossRef]
36. Souiden, N.; Ladhari, R.; Chiadmi, N.E. Destination personality and destination image. *J. Hosp. Tour. Manag.* **2017**, *32*, 54–70. [CrossRef]
37. Afshardoost, M.; Eshaghi, M.S. Destination image and tourist behavioural intentions: A meta-analysis. *Tourism Manag.* **2020**, *81*, 104154. [CrossRef]
38. Ramkissoon, H.; Nunkoo, R. City image and perceived tourism impact: Evidence from Port Louis, Mauritius. *Int. J. Hosp. Tour. Admin.* **2011**, *12*, 123–143. [CrossRef]
39. Stylidis, D.; Biran, A.; Sit, J.; Szivas, E.M. Residents' support for tourism development: The role of residents' place image and perceived tourism impacts. *Tourism Manag.* **2014**, *45*, 260–274. [CrossRef]
40. Zuo, B.; Gursoy, D.; Wall, G. Residents' support for red tourism in China: The moderating effect of central government. *Ann. Tourism Res.* **2017**, *64*, 51–63. [CrossRef]
41. Jones, K.S. A statistical interpretation of term specificity and its application in retrieval. *J. Doc.* **1972**, *28*, 11–20. [CrossRef]
42. Newman, D.; Bonilla, E.V.; Buntine, W. Improving Topic Coherence with Regularized Topic Models. In Proceedings of the 25th Annual Conference on Neural Information Processing Systems, Granada, Spain, 12–14 December 2011.

43. Stevens, K.; Kegelmeyer, P.; Andrzejewski, D.; Buttler, D. Exploring topic coherence over many models and many topics. In Proceedings of the Empirical Methods in Natural Language Processing 2012, Jeju Island, Korea, 12–14 July 2012.
44. Röder, M.; Both, A.; Hinneburg, A. Exploring the space of topic coherence measures. In Proceedings of the WSDM 2015: Eighth ACM International Conference on Web Search and Data Mining, Shanghai, China, 2–6 February 2015.
45. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent Dirichlet Allocation. *J. Mach. Learn. Res.* **2013**, *3*, 993–1022.
46. Kumar, S.; Yadava, M.; Roy, P.P. Fusion of EEG response and sentiment analysis of products review to predict customer satisfaction. *Inform. Fusion* **2019**, *52*, 41–52. [CrossRef]
47. Kumar, S.; Gahalawat, M.; Roy, P.P.; Dogra, D.P.; Kim, B.G. Exploring impact of age and gender on sentiment analysis using machine learning. *Electronics* **2020**, *9*, 374. [CrossRef]
48. Dang, N.C.; Moreno-García, M.N.; De la Prieta, F. Sentiment analysis based on deep learning: A comparative study. *Electronics* **2020**, *9*, 483. [CrossRef]
49. Mujahid, M.; Lee, E.; Rustam, F.; Washington, P.B.; Ullah, S.; Reshi, A.A.; Ashraf, I. Sentiment analysis and topic modeling on tweets about online education during COVID-19. *Appl. Sci.* **2021**, *11*, 8438. [CrossRef]
50. Nemes, L.; Kiss, A. Information Extraction and Named Entity Recognition Supported Social Media Sentiment Analysis during the COVID-19 Pandemic. *Appl. Sci.* **2021**, *11*, 11017. [CrossRef]
51. Kamal, S.; Sharma, S.; Kumar, V.; Alshazly, H.; Hussein, H.S.; Martinetz, T. Trading Stocks Based on Financial News Using Attention Mechanism. *Mathematics* **2022**, *10*, 2001. [CrossRef]
52. Hutto, C.J.; Gilbert, E.E. VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. In Proceedings of the Eighth International Conference on Weblogs and Social Media (ICWSM-14), Ann Arbor, MI, USA, 1–4 June 2014.
53. Beijing Municipal Cultural Heritage Bureau. Available online: http://wwj.beijing.gov.cn/bjww/wwjzzcslm/1737418/1738088/1742737/index.html (accessed on 17 July 2022).
54. China Intangible Cultural Heritage Network & China Intangible Cultural Heritage Digital Museum. Available online: https://www.ihchina.cn/news2_details/24607.html (accessed on 17 July 2022).
55. Pine, B.J.; Gilmore, J.H. *The Experience Economy: Work is Theater and Every Business A Stage*, 1st ed.; Harvard Business School Press: Boston, UK, 1998; pp. 56–58.
56. The People's Government of Beijing Municipality. Available online: http://www.beijing.gov.cn/gongkai/guihua/wngh/cqgh/201907/t20190701_100008.html (accessed on 17 July 2022).