

**Impact of Social Media Marketing on
the Popularity of
Museum Program in China
— Use Weibo as an Example**

A Thesis Presented to the Faculty of Architecture and Planning
COLUMBIA UNIVERSITY

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in Urban Planning

by

Bingruo Duan

Instructor: Xin Li

Readers: Danil Nagy, Jeffrey Johnson

May, 2015

ACKNOWLEDGEMENTS

I would like to thank Dr. Xin Li for the guidance, patience and understanding throughout the research in the past year. Xin, as my thesis advisor, helped me throughout the process of research. She encouraged me to concentrate on big data research when refining my research scope, which gave me a lot of confidence. I enjoyed the whole research process with Xin in which both of us learned a lot regarding the application of new methodology in the analysis of urban issues.

I would also like to thank Danil Nagy as my reader. Danil gave me huge support on the construction of my methodology related to data mining throughout the whole reserach process. My technical part of methodology could not be completed without Danil's help. The technical guidance he offered is also very helpful for my other research projects. He is also the person who introduced the whole world of data mining and mechine learning to me. I also confirmed my future research direction after his guidance.

I am also very grateful to Jeffrey Johnson as my reader. Jeffrey's research Studio-X on museums in China gave me the inspiration of conducting this research. Jeffrey also provided a lot of helpful guidance on the museum policy and the outline of my research. I enjoyed working with Jeffrey a lot as he is not only an excellent mentor but also a good friend to me.

I also want to send my gratitude to my classmates, families and friends who offered help during my research process.

Table of Contents

ABSTRACT	1
1. INTRODUCTION	2
2. REVIEW OF LITERATURE AND HISTORY	4
2.1 A BRIEF HISTORY OF MUSEUM CONSTRUCTION IN CHINA	4
2.2 MUSEUMS' TRAIL TO GAIN POPULARITY	5
2.3 CURRENT STUDY ON WEIBO DATA AND TRADEOFFS	6
2.4 MUSEUMS' IMPLEMENTATION OF WEIBO	7
3. METHODOLOGY	9
3.1 PILOT STUDY	9
3.2 WEIBO DATA COLLECTION & RECONSTRUCTION	11
4. DATA ANALYSIS	13
4.1 DATA CATEGORIZATION	13
4.1.1 MUSEUM ACCOUNTS' CATEGORIZATION	13
4.1.2 WEIBO POSTS' CATEGORIZATION	17
4.1.3. CATEGORIZATION OF WEIBO USERS WHO CHECKED-IN AT THE MUSEUM	21
4.2 CORRELATION	23
4.2.1 KENDALL'S TAU CORRELATION	23
4.2.2 CASE STUDY	25
5. FINDINGS & CONCLUSION	29
REFERENCE	32

Abstract

After the country's 12th five-year plan, in which cultural industry is defined as one of the pillar industries for the country, China is constructing new museums at an astonishing speed. However, combined with the museum mania is the lack of financial support, collection and most importantly, visitors, among new museums. Thus, how to increase the popularity became a challenge for these museums. With the booming of social media, many museums witnessed the success of social media marketing done by enterprises from many other industries. Therefore, a lot of museums opened their social media account, aiming to gain popularity. By collecting and analyzing data directly from Sina Weibo (Microblog), one of the most popular social media platforms in China, the thesis will explore the effect of Weibo marketing for museums to gain popularity. With the help of API (Application Programming Interface), web scraping and machine learning, it is possible to see directly into the social media data and find out the answer by data-mining methodology.

1. Introduction

In the past 5 years, China has witnessed a boom of museum construction. In May 2011, the Chinese government published the country's 12th five-year plan, in which cultural industry was designated as one of its pillar industries. As a result, on May 7th 2012, the China Ministry of Culture published the Cultural Reform and Development Plan, stating that China is going to have 3500 museums by the end of 2015, of which 2500 will be free to the public, with a series of policy incentives and financial supports (Cultural Reform and Development Plan during the 12th 5-Year Plan, 2012). Incentivized by the general guideline of the central government as well as the Ministry of Culture, the construction of Chinese museums has seen its peak since 2009. According to the annual report of museum of 2013 from the National Bureau of Cultural Heritage of China (NBCHC), there are 4167 museums recorded in the NBCHC's system¹, including 3354 public and 811 private museums, which are 299 more than the statistics of 2012. Moreover, there are 2780 free-entry museums among them. In comparison, the figure is only 2252 in 2009 (Cultural Development Statistical Report 2013, People's Republic of China, 2014). On the other side, the governmental expenditure² on the museum sector also increased dramatically. According to the latest report of National Bureau of Statistics of China, the public expenditure on the museum sector is 14.2 billion Yuan, or 2.3 billion USD, in 2012, which is nearly 3 times of the expenditure of 2007's 4.7 billion Yuan.

Combined with the construction mania all over the country is the pursuit of larger space. The trend happens both in private and public museums as both private developers and governments regard the size of the museum as a way to show their financial power or political achievements. However, many museums do not have enough collections and programs to fill in their huge space. For example, the new National Museum of China, which is renovated in 2010, is more than 200 thousand square meters, which is as big as the Metropolitan Museum of Art in New York City (often called "the MET"), the largest museum in the world by exhibition area. Nevertheless, the number of pieces of collections in the National Museum of China is 1.2 million (Ying, 2010). Although the museum officially declared that it had the largest amount of collections in the country, the figure is still far less than the MET, which has more than 2 million pieces of permanent collections. China's desire for larger museum space is not unique in the capital city. Many less developed provinces and cities also try to build their magnificent museums despite of the fact that a large number of them cannot support the museum on either the financial sector or the program sector. The new Chengdu Contemporary Art Center, designed by the famous female architect Zaha Hadid, is even bigger than the National Museum.

The cases in Beijing and Chengdu are only two cases that are among this great leap of museum construction. The whole industry is booming and it seems that China is enjoying the fulfillment of its cultural dream. However, this is far from what people had expected. After the

¹ The definition of museum in NBCHC's registration system is different from the definition this article is going to adopt. This will be only a reference and the explanation will be presented later.

² Including all levels of governments in China.

construction, many museums remain mostly empty with no or few collections or visitors, and limited financial support. The phenomenon can be seen in both public museums and private museums, which makes a sharp contrast with the huge space and the booming numbers. The issue attracts attention from both inside and outside of China, addressing the over construction problems that lying behind the mania ("Special report - Museums: Mad about museums; China," 2013). Lots of museums nationwide are suffering from lack of professional staffs, operation fund and collections. More importantly, as required by the nationwide free-entry policy, qualified museums³ have to be open to the public for free, regardless of whether its ownership is public or private. As there is no custom of donation in China's museum sector over history, the financial conditions of these museums become even worse.

As the emptiness issue keeps expanding nationwide, museums start to use their own means to solve the problem. One of the means is marketing themselves on online social media. Among different channels of social media, one of the most popular ways for museums to adopt is Weibo, a Chinese localized version of Twitter. Launched in 2009, Weibo has become the largest online social media service provider in China. According to Weibo's 2014 second quarter financial results, there are 156.5 million monthly active users and 69.7 million daily active users on Weibo by the end of June 30, 2014 (Weibo Cooperation("Financial results for the second quarter ended June 30, 2014 of Weibo," 2014). The first museum official account on Weibo is created by Shenzhen Museum on October 2009. Since then, Weibo gradually becomes the marketing tool for museums to brand themselves, publish event notice and interact with their followers as well as Weibo users.

This thesis will focus on the relationship between museum's Weibo marketing strategy and its impact on different types of Weibo users and their check-in frequency, which is used to find out whether museums can use Weibo to boost their popularity. The overall purpose of this thesis is to explore how museums can use Weibo to increase their popularity by looking into the social media data and we are expecting to see a common pattern of data for museums which are successful in Weibo marketing. The main methodology used in the research is data mining, which is to collect the data from Weibo, and machine learning, including supervised learning, unsupervised learning and some basic idea of natural language processing. The analysis will start from the review of Chinese museum history, construction boom, recent trend on big data analysis regarding social media and current situation of museums' Weibo marketing. The following section is the explanation of the methodology, which is the main body introducing the research design of the analysis. Then is the research, which is the analysis of the data. After the analysis, findings and conclusions will be given out.

³ This is an official evaluation program conducted by Ministry of Culture and National Bureau of Cultural Heritage to see if the museum is qualified to operate without entrance ticket. The program starts from 2011.

2. Review of Literature and History

2.1 A Brief History of Museum Construction in China

The development of museum sector has a long history of interaction with the development of China's cultural industry and it is still playing a crucial part in the whole cultural industry today. Thus, in order to look deep into the current museums' marketing strategy, the historic review will focus on two perspectives, which are policy and market, to elaborate the development of China's cultural industry and the museum sector.

From the policy perspective, since the economic reform starting from the late 70s, China's central government started to pay attention to the nation's economic development as the whole country showed a sign of recovery from historical chaos. The cultural industry, which was totally owned and manipulated by the central government in Mao's era for the purpose of ideology control, was open to the market. Combined with the loosening of the censorship on culture, Chinese cultural industries began to regenerate and grow. As a result, the 80s is regarded as the time of resurrection of culture and art sector in China. After 2000, another wave of cultural construction took place in China, coincided with the boom in economy and international events such as Olympic Game and EXPO. At the same time, the museums construction started to accelerate. In December 2005, the CPC Central Committee issued *Several Opinions on Deepening the Reform of the Cultural System*, clarifying the boundary between cultural agencies and cultural industries, offering more space for cultural industries to develop by ordering government to stay away from the operation of cultural market. The policy, which intended to give freedom to cultural enterprises on the level of central government, in turn gave local government more power to control the industry. As a result, the local governments started to use the power to build large-space museums as a showcase of their political achievements (Shan, 2014). The boom is also believed to generate with the macro strategy on the national level such as the nationwide new town construction. Bianchini and Parkinson in 1993 stated that cultural policies in Western Europe, sometimes centered on museums and sometimes other projects, were associated with urban regeneration projects as the case in Sheffield, Bilbao and Ruhr (Bianchini & Parkinson, 1993). The case in China may be slightly different as the construction focuses on the new town instead of the blighted industrial town. Nevertheless, the intention on the government level is the same as to bring the city to life with the cultural industry incentive.

On the market sector, it is believed that the loosen of cultural policy on restricting private museums is due to the rapid emergence and development of mass art and mass social media, based on the experience of Western Europe (Hesmondhalgh & Pratt, 2005). Despite of the incomplete law and the market that is not healthy enough, Chinese art market has witnessed a wave of prosperity between 2009 and 2013 as there are 3 pieces of Chinese art works which are among the 10 most expensive art works in the global art market of 2011 (Liu, Wu, & Zhu, 2012). On the other side, since 2005, with the Internet social media becoming prevalent, the

growth of cultural industry in China has appeared again. On the other hand, benefited from the rapid increase of Chinese economy, the amount of Chinese millionaires and billionaires keeps growing. As a result, many individual collectors want to show their own collections to the public for different purposes. Thus, private museums, especially museums showing contemporary art, are being constructed in recent years.

In summary, throughout the history, policy makers have always been cautious about the ideology of the country, to make sure it won't permit the culture to go too 'far away' from socialist development. Thus the whole post-Mao period is a period that cultural policies are contesting and struggling between the authority and the market (Tong & Hung, 2012). The main contradiction that lies in the whole process of cultural industry development is the contradiction between the control upon ideology and the marketization of cultural industry.

2.2 Museums' Trail to Gain Popularity

Throughout history the question of how to gain popularity has been a trouble for museums all over the world since the museums came into being. With the massive scale of emptiness occurred in Chinese museums, the study on museums' influence and marketing strategy has become more popular than ever before. Among the studies on museum influence, a consensus has been reached that museums can gain the popularity during three phases, the attracting phase, the broadcasting phase and the after-effecting phase (Di, 2013; Hudson, 1987; Kotler, Kotler, & Kotler, 2008). During the attracting phase, which is the phase before people's visit, the core to gain museums' influence is the plan for the museums' construction and market niche. In this phase, the developers and related institutions also have responsibility in the decision of the construction. The market niche of a museum should look back to its geographic, historical and cultural context. Moreover, most of the famous museum clusters were formulated during this period, such as the Smithsonian in the US, the Museum Island in Berlin and the new Jianchuan Museum cluster in Sichuan, China (Di, 2013).

The broadcasting phase is the phase when people are directly interacting with the museum. In this phase, the program of the museum is the key to decide whether the museum is attractive to the audiences and will bring more visitors in the future. Several strategies are mentioned to strengthen the positive influence of the museum upon the attractiveness to the audience. The first is the traditional post occupancy evaluation based on audiences' feedback after their visit. According to a survey by Zhang et al. in 5 museums in Hangzhou, Zhejiang, the audiences' most satisfied factors in museums are amount of collections, environment and quietness and layout of exhibition. While the least satisfied factors are guide service, number and scale of the exhibitions and not enough interactions (J. Zhang, Yu, Ma, & Xu, 2008). Thus, during this phase, Kotler et al. and Hudson both stressed the importance of community-oriented program and educational program, which both start from the intention of enhancing the experience of the audiences, as the audiences' experience is the most important element to stress in the broadcasting phase. Another strategy is the stress of visitors' experience

during visit. This theory tries to emphasize the importance of different types of program's attractiveness to different types of visitors (Di, 2013). The key of the efficiency of the broadcast is the character of the content and the way of conveying. Thus, in order to attract different target audience, museum needs different types of exhibition and marketing method to reach out the interest of the target audience, which is also the focus of the analysis of this thesis.

The after-effecting phase is the continuous influence from the museums upon visitors' recognition. In this phase, it is museums' effort to use every possible means to recall memory and experience of the visitors, which is to keep interaction with audiences and bring them back the experience. With the development of modern technology such as social media, this phase has attracted more and more attention and museums' efforts (G. Zhang, 2011; Z. Zhou, 2012).

2.3 Current Study on Weibo Data and Tradeoffs

Weibo, used to be called Sina Weibo, is a twitter-like micro blog service launched in 2009, aiming to serve the Chinese netizen by offering the more localized substitution of twitter service in China. As introduced before, there are 156.5 million monthly active users and 69.7 million daily active users on Weibo by the end of June 2014, which consists of nearly one tenth of the whole Chinese population. There is no cost to open an account and most of the services on Weibo are free. For the reasons above, as a result, Weibo become a popular platform for museums to brand, broadcast and market themselves (G. Zhang, 2011). Meanwhile, because the dataset of Weibo is accessible with its API (Application Programming Interface) service, Weibo is also a perfect dataset for researchers to extract data and conduct further analysis.

Current analysis on Weibo data concentrates on two domains, the Weibo account influence analysis and Weibo broadcast analysis. The Weibo account influence analysis (e.g., Li, J., Chen, Z., & Huang, J, 2012) is used to illustrate the impact of the Weibo account on the entire Weibo platform. In Weibo, an account can be the representative of an individual, a certain interest group or an organization. The analysis mainly focuses on the influence of celebrities and official accounts of institutes and governments. The analysis requires the data of account information and account behavior. Account information is related to follower amount, which is number of followers, as the direct representative of Weibo accounts' popularity. The account behavior contains number of post, repost, reposted, following, mention, mentioned, comment and commented. So far there are three main models to analyze the influence of Weibo account by using these two sets of data. The first is the PageRank (PR) model developed by Larry Page and Sergey Brin. The central idea of the model is to determine the importance of a web page (in our case is the Weibo account) based on the amount of hyperlinks leading to it. The parameter in the model is the number of followers. Thus the model is simple and clear, excluding the zombie followers, which are large amount of empty accounts created by a single IP in a very short time to help the target account look more popular (Weng, Lim, Jiang, & He, 2010). The second model is based on Weibo users' actions. The actions of following, reposting

and mentioning are documented and adopted in this model to test the account's influence. According to Cha, based on the model, the accounts with most followers are more likely to be accounts that post more pictures and public information. The accounts whose post got mostly reposted are commercial elite, Internet service provider and news Weibo. The accounts that are most frequently mentioned by others are celebrities (Cha, Haddadi, Benevenuto, & Gummadi, 2010). The third model combines the first two models together, which is called TU (Twitter User) Rank. The power of the parameters in the model is subjectively decided according to the user pattern and character in Twitter. Moreover, the statistical methods are also popular in the Weibo account influence analysis such as the Spearman's rank correlation coefficient and Kendall's rank correlation coefficient. Based on the TU Rank model, Li et al. built a model that considers more about Chinese Weibo character, which is called WU rank (score) as the biggest difference between Twitter and Weibo is that Weibo account users can reply their followers' comments under their own post (Li, Chen, & Huang, 2012). The analysis of the paper will consider part of this model in analyzing the museum Weibo accounts.

The other popular field of Weibo study is the Weibo communicative model study. This aspect of study focuses on the pattern of a piece of Weibo gets conveyed. Moreover, the study also cares about the role of Weibo in public incidents and as a marketing tool. The boom of social media enables people to share their personal thoughts with the world, changing the Internet platform from a one-direction channel like the traditional media, to a multi-direction platform that everyone can publish information and share with the rest of the world. The current analysis focuses a lot on the influence of Weibo as a platform to share information, especially the information regarding public issues. Comparison with the traditional rumor spreading model (SIR model) is the major analysis tool to examine the information dissemination pattern of Weibo (Qian, Zhang, Zhao, & Zhong, 2012). Moreover, the efficiency of the Weibo information spreading is another field of study regarding the broadcasting issue. According to Jin, Wang and Chen, the broadcasting pattern on Weibo evolved from the traditional media's way of attention-interest-desire-memory-action (AIDMA) to attention-interest-search-action-share (AISAS) (Jin, Wang, & Chen, 2011). Based on this broadcasting pattern, they summarized a model to examine the efficiency of Weibo advertisement. By collecting the data related to the Weibo account that wants to advertise on Weibo, they conclude that the fame of the brand, account's degree of activity and the size of the company are three main factors that decide the efficiency of the advertisement on Weibo. Furthermore, with the analysis on the pattern of communication, the contents and Weibo's uniqueness are also analyzed in these studies. According to Yu et al., in comparison with Twitter, Weibo users are more like to share jokes, pictures and videos and a significantly large percentage of posts are retweets. The trends that are formed are almost entirely due to the repeated retweets of such media content (Yu, Asur, & Huberman, 2011).

However, the analysis on Weibo and the analysis by using Weibo data both have their limitations. The data that directly scraped from Weibo does not have a standard form and is not well structured, which is challenging for the construction of database structure in

computational analysis (X. Zhou, Lu, Li, & Du, 2012). Furthermore, Samuel Arbesman claimed that the model created from big data cannot deal with the changing future because the data is a part of the past (Arbesman, 2012). Also, as the Weibo is main user group ages between 16-50, the data cannot better reflect the thought and opinions of children and elderlies (Wang, Jin, & Cheng, 2013).

2.4 Museums' Implementation of Weibo

Given the importance of Weibo on information broadcasting and marketing, many museums choose to open their verified official Weibo account and interact with Weibo users. Currently, if we type the keyword "museum" in Weibo's search column, we can get more than 14,000 results of Weibo accounts. These accounts include the official account of a museum, museum curators and staffs and also museum media. Besides, not only museums and staffs in China have their Weibo account, lots of museums abroad also have their official Weibo account such as the MET, V&A Museum in London and Art Institute of Chicago. Compared to the traditional marketing method and exhibition notice, Weibo has 5 advantages for museums, according to Zhou's research (Z. Zhou, 2012). First, museums can get feedback from the audience faster through their Weibo comment. Second, Weibo provides more opportunities for museums to interact with Weibo users. Third, museums can use Weibo as an extension of broadcasting and marketing. Fourth, curator can share more of their own thoughts and experience as well as the story behind the exhibition with the Weibo audience. Fifth, museums can get users' data to analyze different users' interest points, which will be covered in the following analysis. However, during the interaction of Weibo, most museums did not fully develop the interactive feature of Weibo when communicating with their audience (Z. Zhou, 2012).

Furthermore, in order to get a good outcome of the Weibo broadcasting and marketing, museums account always choose the following 7 types of Weibo contents, based on Zhang's summary (G. Zhang, 2011). The museums' Weibo contents include information publication, answering questions from the public, initiating public and entertain activities, noticing discount offers, highlighting related news, showing care to the public and starting public topic for discussion. After the categorization, Zhang also evaluated the influence to the public for each kind of Weibo on the repost ratio and interactive ratio. According to his evaluation, initiating public and entertain activities, together with starting public topic for discussion have the highest repost ratio and interactive ratio, followed by answering questions from the public and noticing discount offers. Thus, museums Weibo have to be cautious about the converging of the Weibo contents in the future. Moreover, in order to get rid of the convergence, more attention should be paid on users' analysis side for the museums.

3. Methodology

The method of the research consists of two parts, which are pilot study and data collection & reconstruction. The target region of this research is Beijing, meaning the research will focus on the museums and Weibo users in Beijing municipal region. There are two reasons to choose Beijing as the study region. First is that Beijing is not only the political capital but also the cultural capital of China, according to the strategy established in Beijing Comprehensive Plan 2004-2020. Thus the museum sample is more abundant than elsewhere in China. Also, if the museums in Beijing also have emptiness issue and trouble in marketing by Weibo, then the national situation could be worse and more attention is needed. Second is Beijing has the highest Internet penetration rate¹ all over the country, which is 75.2% by the end of January, 2014, according to the latest report of China Internet Network Information Center. In comparison, the national average Internet penetration rate is 45.8% at the same time ("Statistical Report on Internet Development in China (January 2014)," 2014). A higher Internet penetration rate suggests a higher potential of Weibo users, which will make the analysis sample larger and more meaningful.

The research scope of museum in the thesis not only contains the registered museums in the list of NBCHC's registration system, but also includes private museums, art galleries and exhibition spaces that have hold exhibitions and lectures in the past year. As a result, the amount research object of museums in Beijing municipal area increased from 170 to nearly 500 under this definition. After looking into their Weibo marketing behaviors and the amount of check-ins in these museums, we selected 21 museums as the study objects for the final research.

In this research, we will explore the relationship between two groups of frequencies. The first is the frequency that museum Weibo accounts post on Weibo. The second is the frequency of users checking-in in these museums. By comparing the different outcome of Weibo visitorship, we can have a clear view of the influence of Weibo marketing implementation upon the popularity of museums. More importantly, the research will use both traditional data from statistical yearbooks or related authorities and Social media data directly collected from Weibo dataset. With the two different types of data, the research will use the traditional data as complimentary when testifying the outcome of Weibo marketing.

3.1 Pilot Study

With the huge volume of data that will be extracted from Weibo, it's difficult to categorize the museums, marketing strategies and users without a set of clear criteria. Thus, by going deeply through and analyzing several representative museums and the related users as a pilot study, it is more helpful for the further research as the type of museums, marketing strategies and users are easier to define. Moreover, after examining the sample data and knowing what

¹ The Internet penetration rate is the ratio of population who has access to the Internet to the total population, regardless of the device, explaining the popularity and degree of prevalence of Internet.

the data will be looked like, it is easier for us to collect the data that is related to the research instead of collecting the whole dataset, which will save a lot of time in the data collection process. For the museums that are used for the pilot study, they are either performing well in the Weibo marketing or having their own uniqueness on the Weibo interaction with users.

The museums chosen for the pilot study are National Museum of China (NMOC), Ullens Center for Contemporary Art (UCCA), Beijing Capital Museum (BCM) and Today Art Museum (TAM). The four museums all have their own official Weibo account. NMOC's Weibo account is the most popular account in museum category as it has more number of followers than any other museum accounts on Weibo, according to Weibo's search suggestion. Besides, NMOC's number of check-in users is also the highest among all the museums in Beijing. UCCA is a prominent museum for contemporary art in the 798 art district, which is a famous case for industrial site's regeneration and the reuse as art studio. It is the most popular private museum by check-in users. Moreover, UCCA's Weibo marketing strategy is not limited in a single account. It has several verified accounts dealing with different issues. For example, the main account focuses more on the exhibition announcement. The store account publishes more discount information. The curator and staffs also have their own account to help with the public relation and marketing. BCM is a famous local museum owned by Beijing municipal government. The Weibo account of BCM received a large amount of check-ins from users. However, the amount of Weibo the account has post is much less than other museum accounts. In comparison, Today Art Museum is another private contemporary art museum, which used to be state-owned and successfully transformed to private-owned during the early 2000s. The museum has the least number of check-in users among the four museums but its Weibo account has more number of posts than the UCCA and BCM.

After deciding the four cases of pilot study, we started to collect the raw data for these four museums and their visitors. First, the collection of museum account data is mainly based on the use of Weibo API. With the API `2/place/poi/show`, the geographic information of a museum is collected, including the exact address, coordinate, total check-in number, total number of check-in users and total number of check-in photos. On the account aspect, with the API `2/users/show`, information of museum accounts' Weibo id, gender, number of followers and friends, number of pieces of Weibo posted, verified status and time of account creation are collected. By using the model of PR Rank and Weibo User Score in the literature review, we can also get the index of PR rank and interaction rate for a museum account.

As the Weibo Company only allow elementary developers to get the latest 2000 pieces of Weibo status that an account has posted, it is hard to get all the Weibo posted by a museum account with API. Thus, the collection of museums' Weibo content is finished with web scraping using Python. In this process, every piece of Weibo and its post time, number of comments, forwards, likes are gathered. Moreover, if this piece of Weibo is not an original Weibo created by the museum account, the code will also extract the original Weibo's content, post time and forward time.

Method	Data Collected
API: 2/place/poi/show	Latitude, Longitude, Total Check-in Amount, Total Number of Check-in Users, Total Number of Check-in Photos
API: 2/users/show	Number of Friends, Number of Followers, Number of Posts, Account Creation Time, Verified Status
Web Scraping	Weibo Content, Number of Comments, Number of Reposts, Number of Likes, Post Time, Repost Time
Derivative Data	Average Posts per Day, Page Rank, Interaction Rate, Original Weibo Rate, Average Comments, Average Reposts

Table 1. Methods Used during Data Collection

For the collection of users' data, we first use the API /2/place/pois/users to get all the check-in users in one museum, and then import the list of users into the API /2/place/user_timeline to see all the check-in activity of every single user.

With the manual examine, combined with literature review, of the data for these three museums, we will come out with a proper categorization method of museum, strategy and user types, which will guide the training of the unsupervised learning model to compare with the traditional dataset. The categorization method will be used in the analysis of the entire dataset. Table 1 shows the methods used in the data collection process.

3.2 Weibo Data Collection & Reconstruction

As mentioned above, the data that is going to be used in the research consists of two parts, the traditional data and the Weibo data. The traditional data, as it appears, is the data that describing the museums' basic information such as ownership and museum type, which can be collected from traditional channel like yearbook and interview. The Weibo data is the data related to museums' Weibo account and Weibo users who checked-in the museum. With the help of API (Application Programming Interface) and web scrapping, most of the Weibo data can be collected. The Weibo data will be divided into three categories, which are museum accounts, museum accounts' Weibo post and check-in users' information. We will focus more on the Weibo data in this section.

Museum account is the museum's Weibo account verified by Weibo as the official representative of the museum on Weibo. There are more than 700 official and individual accounts representing over 500 museums, exhibition centers and galleries in Beijing. However, based on the amount of Weibo visitorship, the museums that with less Weibo visitorship will be eliminated, as the sample is not big enough to analyze. As a result, the size of sample is 21, which is mentioned before. For museum account, the information we need includes the data directly collected from API, such as number of friends, followers and posts. Moreover, we need the derivative data that needs further calculation based on the direct data, such as average posts per day, original posts percentage and PR rank.

Strategy	Program
Announcement	Lecture & Conference
Interaction with Weibo Fans	Education & Community Program
Introduction of Related Knowledge	Moive
Initiation of a Discussion Topic	Live Performance
Sales & Lottery	Exhibition
Others	Others

Table 2. Category of Contents' Strategy and Program

The second part of data is the museum accounts' Weibo posts. This part of data directly comes from the Weibo posts of museum accounts and is gathered with the use of web scrapping. The posts data will mainly serve for two purposes. The first is to categorize the contents of Weibo posts with the help of Natural Language Processing (NLP). The categorization can be conducted in two approaches, first is the categorization into the 6 types of Weibo marketing strategies and 6 types of programs for each piece of post (Table 2). The strategy categorization includes museum accounts' Weibo behavior such as exhibition announcement, interaction with Weibo fans, introduction of the collections in the museum, etc., according to Zhang's research. The program categorization is to see what type of program are introduced by the museum account on Weibo, includes conference, movie, live performance, etc. Both the process need the NLP to process tens of thousands Weibo posts. The second part of posts' data is to find out the most popular type of Weibo post among Weibo users by looking into the data contains Weibo visitorship (check-in amount), number of comments, forwards (retweets) and likes.

The third part of data is the check-in users' information. The data is also directly collected from API, including all the users who have checked-in at one museum in the study list. The data will describe the basic information of users' Weibo account, which is like the museum accounts' information but also with users' gender, originality and other information describing this user.

4. Data Analysis

To analyze the data, what we are going to do consists of two parts, data categorization and correlation, based on our methodology of machine learning.

4.1 Data Categorization

The categorization process aims to classify the massive three groups of data. The three groups of data will be classified by both traditional standard and machine-learning approach.

4.1.1 Museum Accounts' Categorization

For museum accounts, the common sense categorization will describe what type of museum according to the data collected and mentioned above. The first dimension is the ownership and basically, all the 21 museums in the study will fall in either private-owned or public-owned. According to their own websites, 8 out of 21 museums are private-owned museums while

Name	Chinese Name	Alias	Owenship	Content Category
National Museum of China	国家博物馆	NMOC	Public-Owned	Encyclopedic
Today Art Museum	今日美术馆	TAM	Private-Owned	Comtemporany Art
UCCA	尤伦斯当代艺术中心	UCCA	Private-Owned	Comtemporany Art
CAFA	中央美术学院美术馆	CAFA	Public-Owned	Natural, Science & Specified
Beijing Capital Museum	首都博物馆	BCM	Public-Owned	Encyclopedic
Chaoyang Plannning & Art Museum	朝阳规划艺术馆	Chaoyang	Public-Owned	Natural, Science & Specified
National Zoology Museum	国家动物博物馆	NZM	Public-Owned	Natural, Science & Specified
Guanfu Museum	观复博物馆	Guanfu	Private-Owned	Historical
Military Museum of China	军事博物馆	MillitaryM	Public-Owned	Natural, Science & Specified
China Millennium Monument	中华世纪坛	Shijitan	Public-Owned	Relic & Exhibition
Beijing Museum of Natrual History	自然博物馆	BMNatrualHistory	Public-Owned	Natural, Science & Specified
Beijing Exhibition Hall	北京展览馆	BEH	Public-Owned	Relic & Exhibition
Sanyingtang Photography Art Center	三影堂摄影艺术中心	Sanyingtang	Private-Owned	Comtemporany Art
Yanhuang Art Museum	炎黄艺术馆	Yanhuang	Private-Owned	Historical
Songzhuang Art Museum	宋庄美术馆	Songzhuang	Private-Owned	Comtemporany Art
China Movie Museum	中国电影博物馆	CMovieM	Public-Owned	Natural, Science & Specified
Hanmeilin Art Museum	韩美林艺术馆	Hanmeilin	Private-Owned	Comtemporany Art
Guozijian	孔庙和国子监博物馆	Guozijian	Public-Owned	Relic & Exhibition
Beijing Automobile Museum	北京汽车博物馆	Car	Public-Owned	Natural, Science & Specified
798 Art District	798	798	Private-Owned	Comtemporany Art
National Art Museum of China	中国美术馆	NAMOC	Public-Owned	Natural, Science & Specified

Table 3. Museums' Traditional Categorization

the rest 13 are public-owned, controlled by the cultural department of either municipality government (such as Beijing Capital Museum) or the state (National Museums of China).

Another side of the traditional data of common sense is the type of museums based on their exhibition contents. According to Kotler’s way of categorization for museum type, there are five types of museums in our study scope, which are historical, contemporary art, natural, science & specialized, relic & exhibition and encyclopedic. Table 3 shows the categoriation result.

The machine learning categorization for museum accounts will use the data of museums accounts collected from API. By using unsupervised learning, we can have several clusters divided by the degree of influence, popularity and account activeness of the museum Weibo accounts. In order to get a better model of categorization, we adopted four unsupervised clustering models for classifying the museums by their Weibo account behavior. The four models are K-Means, Hierarchical Clustering of Ward linkage criteria, Complete linkage criteria and Average linkage criteria. Two ways of unsupervised learning are adopted.

The first way is comprehensive classification, in which we put all the parameters into one

Type	Parameter
Degree of Online Activeness	number of friends, followers, total Weibo posts, average Weibo status per day, original Weibo rate, PR rank, interaction rate, average comments, average forwards
Degree of Check-in Popularity	number of check-in Weibo, check-in users and check-in photos

Table 4. Input Parameters for Museum's Moedel

unsupervised model. First of all, so as to find out the better classification, we started with looking for the optimum number of clusters. From the function of score in K-Means module we can calculate the average score for each number of cluster. Then we can select the best number of clusters by identifying the turning point if we put all the scores on a line chart (figure 1). The result turned out that 5 clusters are the most suitable for our case. After deciding the number of clusters, we clustered the dataset with 12 dimensions of

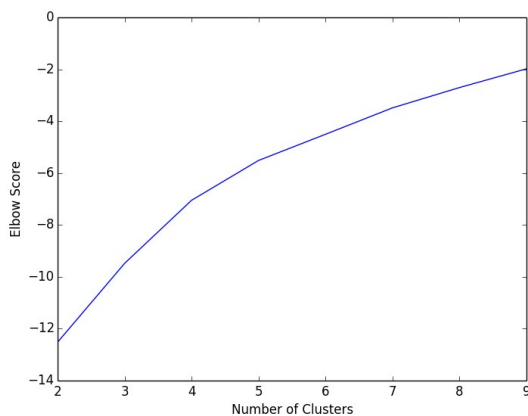


Figure 1. Elbow Score for Museums' Model

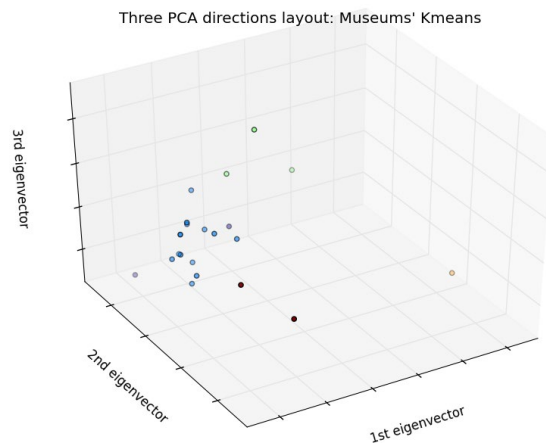


Figure 2. PCA for Museums' Model

data and categorized them into 5 clusters, shown in table 4, which basically cover most figure describing the activeness and popularity of the museums' Weibo account. Because the results of these four unsupervised learning models are almost the same, we choose to show the result of K-Means. The spatial distribution of the result of these 21 museums are shown after the principal component analysis (PCA) in figure 2, which is the three-dimensional projection of the categorical result of these 12 parameters, and the categorical result is shown in figure 3.

As shown in the figure 3, we notice that most of the museum accounts fall into the category A, which can be explained as either inactive on the maintenance of Weibo account or unpopular by looking into the data related to check-ins. From the chart we can see that in these 21 museums, public-owned museums have a relatively higher activeness or popularity, according to the result. Contrary to the common sense and some of the comments on Weibo, public-owned museums tend to have a higher score of overall performance in this chart. Also, the contemporary art museums are not that active and popular in the result, which is also against people's traditional view.

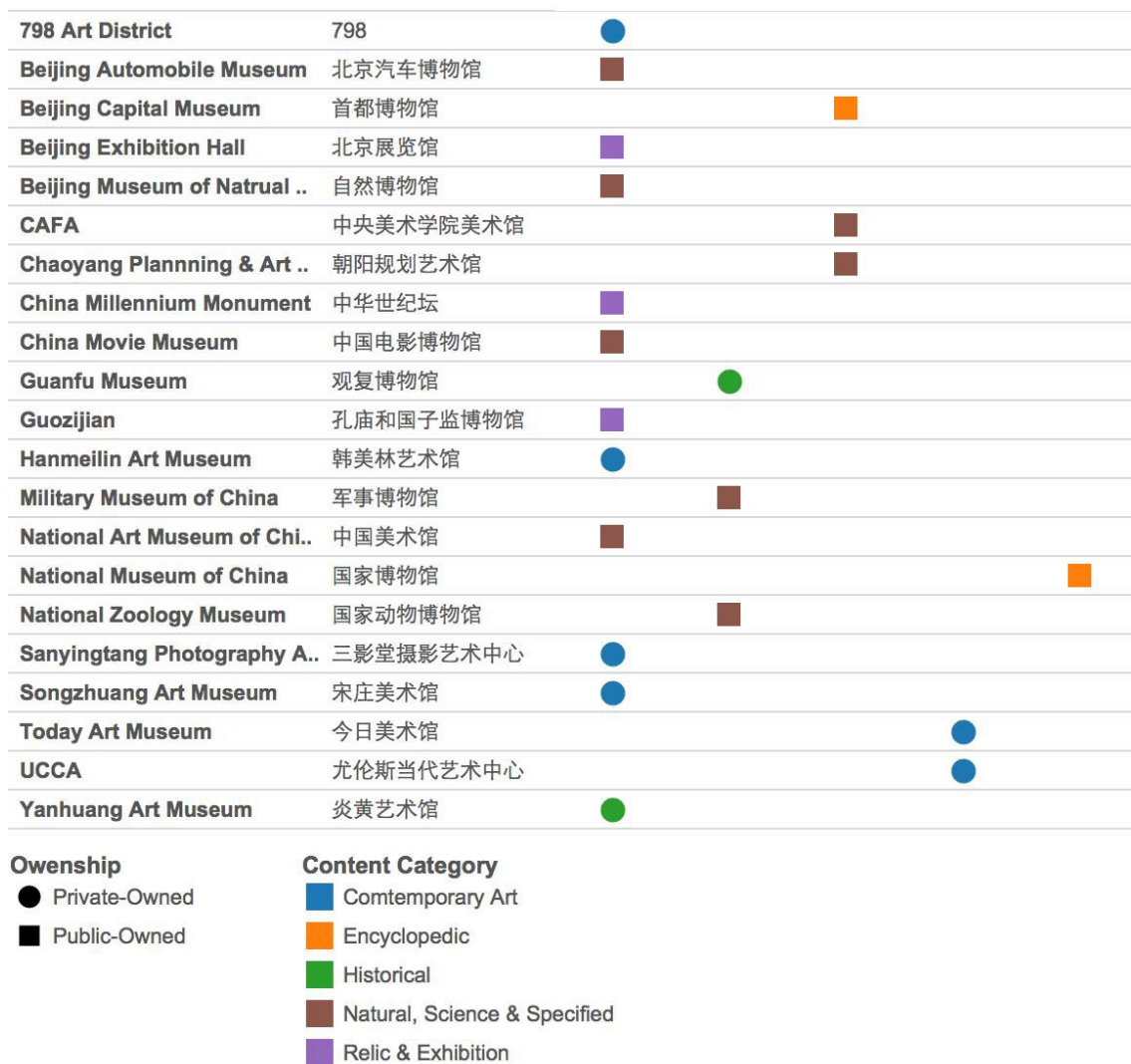


Figure 3. Category Result 1 for Museum Accounts

In order to see the reason for those museums in the same category is inactiveness or unpopularity, we divide the whole dataset into two parts and retrain the unsupervised learning model as the second way of unsupervised learning. The first part is the variables describing the degree of online activeness including number of friends, followers, total Weibo posts, average Weibo status per day, original Weibo rate, PR rank, interaction rate, average comments and average forwards. The second part is describing the degree of offline popularity using number of check-in status, users and photos. Similar to the first way of unsupervised learning, we exhibit the principal component analysis of the result projection in figure 4 & 5. The categorical result is shown in figure 6.

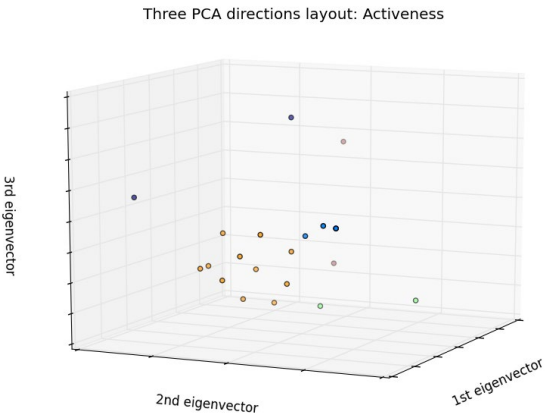


Figure 4. PCA for Activeness Model

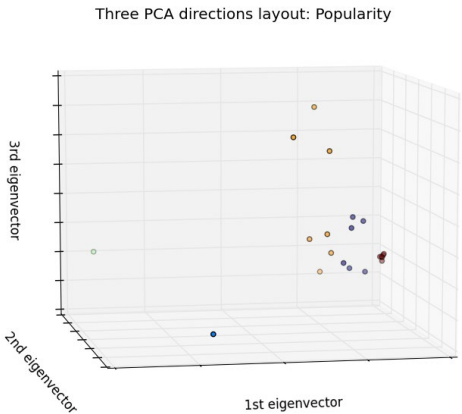


Figure 5. PCA for Popularity Model



Figure 6. Category Result 2 for Museum Accounts

In figure 6, the situation is still similar to the result of figure 3, although the intention is to find out the difference in the museums that fall in the same category. The vertical axis stands for the degree of activeness and the horizontal axis stands for the degree of popularity. Thus, the whole chart is divided into four sections, which are active and popular, active but unpopular, inactive but popular, inactive and unpopular. It is no doubt that the National Museum of China is the museum with the highest popularity and activeness in this list, thus it falls in a cluster itself, no matter it is seen comprehensively or separately by activeness and popularity. For Today Art Museum and UCCA, as the TAM has the triple amount of followers than UCCA and the degree of account activeness is a little bit higher, it also has twice the check-in popularity than UCCA. Also, CAFA (Central Academy of Fine Art Museum) and Beijing Capital Museum both have a higher activeness in Weibo account maintenance, but lacks popularity compared to those popular ones. There are two museums in this list that don't have their own Weibo account, which are 798 and National Art Museum of China. However, there are still lots museums falls into the quadrant 3, as they are both inactive and unpopular. It is obvious that there are several contemporary art museums are neither active nor popular, according to this result. One exception might be the UCCA. UCCA is a successful example in operating and exhibition. The explanation of the less-activeness of UCCA could be that there are several accounts are responsible for the marketing, while we only choose the official account in this analysis. The reason of the less-unpopularity could be UCCA's location. UCCA is in the 798 district and the district contains lots of small art studios and galleries, thus people might select the check-in point as 798 or other galleries instead of UCCA.

However, the overall result of these two approaches shows that although a lot of museums are adopting Weibo marketing, the maintenance of their accounts is not active compared to the active cases such as National Museum of China or Today Art Museum.

4.1.2 Weibo Posts' Categorization

After looking into museums' Weibo account, we will see what they post on their Weibo account. Following the raw data collection of museums' Weibo post, we collect a total of 47083 pieces of posts that are posted by these 19 museum accounts (As 798 and National Art Museum of China do not have their Weibo account), between Jan 1st 2012 to Dec 31st 2014. In these three years, every museum account posts an average of 862 pieces of Weibo per year, or 2.26 pieces per day. If we look into figure 7, CAFA has posted the most number of Weibo posts in these three years, which is 8205, or 7.5 per day. Following CAFA is the National Museum of China, which is 7200. On the other side, there are some museums that establish a much lower post frequency. In figure 7, MilitaryM (China Military Museum) only posted 67 pieces of Weibo post in three years, which is only equivalent to 9 days' post of CAFA. The same situation happens on Guozijian Museum, Guanfu Museum and Hanmeilin Art Museum.

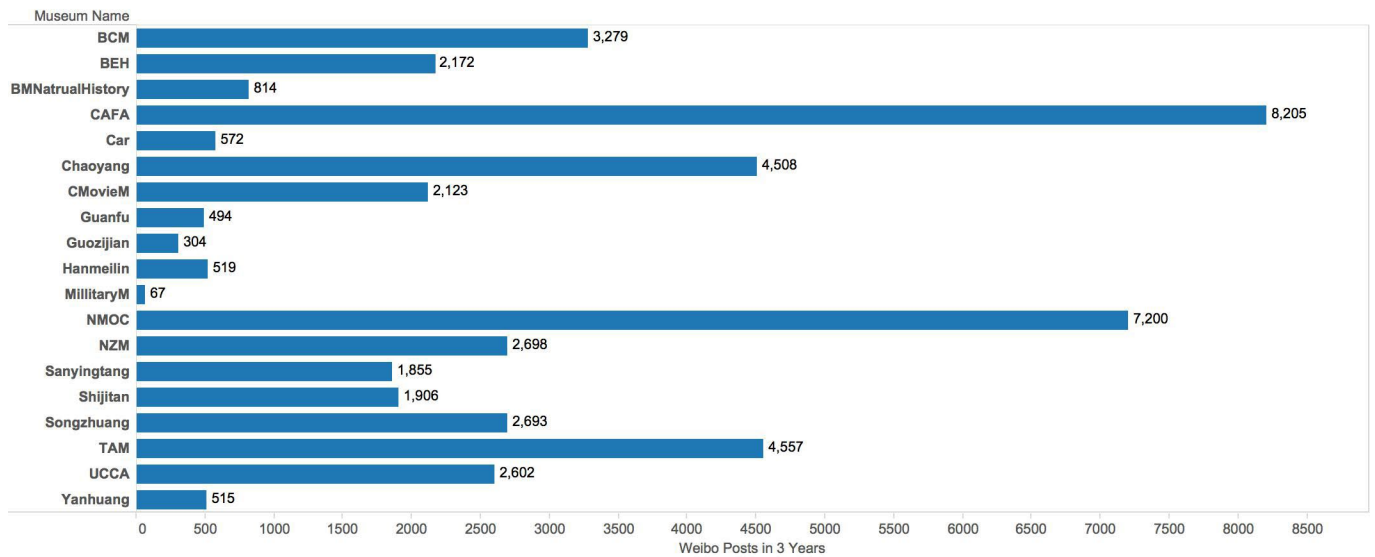


Figure 7. Number of Posts 2012-2014 for 19 Museums

As getting all the Weibo posts of these museum accounts in the three years with web scraping, we start to look into what types of post are made by these museums, which will need the help of Natural Language Processing (NLP).

In order to categorize the post content, we first have to find out the types by looking into the content, which we have completed in the pilot study. There are two perspectives to conduct the categorization, the first is by strategy and the second is by program. For strategy, according to Zhang and combined with the content observation, we list out 6 types of strategy and 6 types of program, which has been shown in table 2. Basically the strategy represents the way of marketing that museum accounts post on their Weibo. While the program is the content they post in it.

Before getting to know the strategy and program for each piece of Weibo posts, first we have to create a dictionary with variables as the training data. We select three museums' Weibo post, which are Beijing Capital Museum, UCCA and National Museum of China, to create the training file, as the style Weibo posts posted by these three museums differs a lot. At the beginning we use a python module Jieba to cut the whole sentence of each piece of Weibo post into Chinese linguistic words. After that we use another function of Jieba to calculate and get the top 300 keywords of all these Weibo posts from three museums. After getting these 300 words, we delete those will not make any sense in our analysis such as 'I', 'Here' and 'That' so as to refine the dictionary and make it more meaningful in the purpose of predicting the content. There are 187 words left in the list. Then we choose 500 pieces of Weibo as the training data. With each piece we will subjectively label them with a strategy and a program based on the judgment of our own. Following the labeling process, we use those 187 words as 187 variables. If the word is in this piece of Weibo, we will give a score of 1, otherwise the score will be 0. In this way, every piece of Weibo is a piece of data with 187 variables. We take 350 of the 500 pieces as the training data and adopt 4 supervised learning models, which are Support Vector Machine, Naïve Bayes, Decision Tree and Random Forest. Following the training data, we use the rest 150 pieces as the validation data to test if the prediction is correct. The result turns

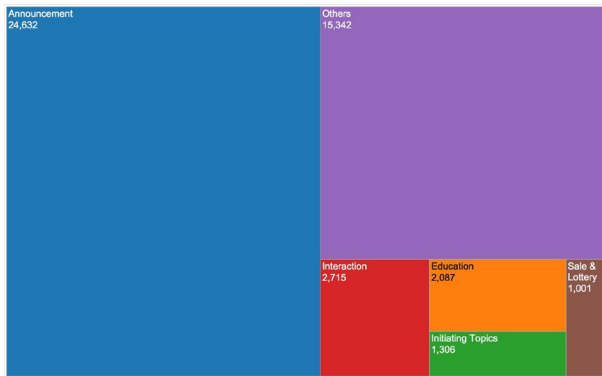


Figure 8. Strategy Overall Distribution

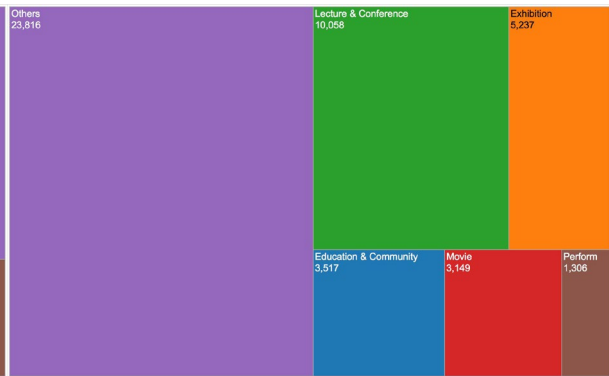


Figure 9. Program Overall Distribution

out the Support Vector Machine model has the highest score of prediction. Thus, we use the model to predict the rest of all the Weibo post we collected for these 19 museums that with Weibo accounts.

With getting the categorization of every piece of Weibo, we will look into the distribution of strategy types and program types regardless of which museum posted them. Figure 8 and 9 shows the result by strategy and program.

In figure 8, we can find out that most popular strategy by these 19 museums is announcement, as we also witness during the pilot study observation and labeling process. Museums tend to use Weibo mainly as a tool of publishing announcement.

The category of others is the 2nd popular strategy. In here, the category of others means those strategies that are not in these 5 categories such as answering the questions asked by Weibo users. On the other hand, due to the limited scale of training data, it is possible that a piece of Weibo will be categorized as others although the piece belongs to another certain category, especially the posts that are used to introduce the collections and related knowledge of the museum as the collections in different museums differ a lot thus it is difficult to label these pieces to the right categories because the keywords are totally different as there are only 187 keywords used as the training parameters.

But we can still see museums' intention of taking the advantage of Weibo' interaction feature, as interaction is the third popular strategy used by these 19 museum accounts. During the pilot study we do find there are lots of pieces that posted in a person's tone, which is a common way for official Weibo to get closer to their audiences. Besides, museum accounts also interact with each other and other organization accounts.

Sales & Lottery is the least popular strategy among these six, as it is not a regular strategy to use, although lottery on Weibo always bring a lot of attentions and followers in a very short time. Initiating a topic & leading discussions is a good way to attract Weibo users and bring the topic to a deeper level. However, it is not that popular among these museums.

If we move onto the result of program, the category of others ranks the top among the six programs. Apart from the same reason of a higher proportion of others as the others in strategy, during the pilot study and the observation of training data, we witness that there are

a lot of Weibo posts that are not directly relevant to the program, especially those introduction of collection and related knowledge.

So if we look into the rest 5 categories in figure 6, 10058 out of the total 47083 pieces of Weibo are talking about lecture and conference, which is nearly double the amount of exhibition. While the amount of movie and education & community are the same. Performance is not a popular option in these museums' programs as dynamic exhibition is not an important part of these museums.

Now we will try to see this result in a different perspective. This time we look into the percentage of each program in each strategy, as is shown in figure 10. It seems that most of the lecture & conference programs are posted as announcement on these museums' Weibo accounts. The exhibition is introduced either with announcement or education & community strategy, which shows these museums' purpose to bring closer of the distance between the local communities and schools. In order to look better into the result, we create a graph without showing the program of others, which is figure 11. By looking at this the content of each type of strategy is clearer. Announcement basically covers most of the programs in the museum, and so does the sale & lottery. While the education & community strategy, which is mentioned before, focus more on the exhibition itself. Initiating topics & leading discussion has the most unique pattern that is the concentration of performance as nearly all the performance programs of these museums are posted as a strategy of discussion. We will look into the detailed result for each museum during the correlation analysis.

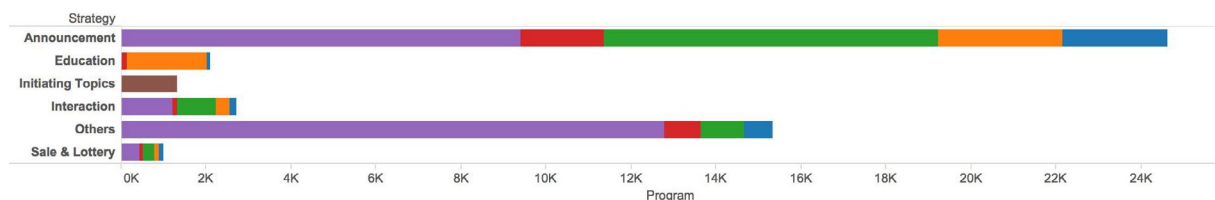


Figure 10. Content Type Overall Combined with Strategy & Program

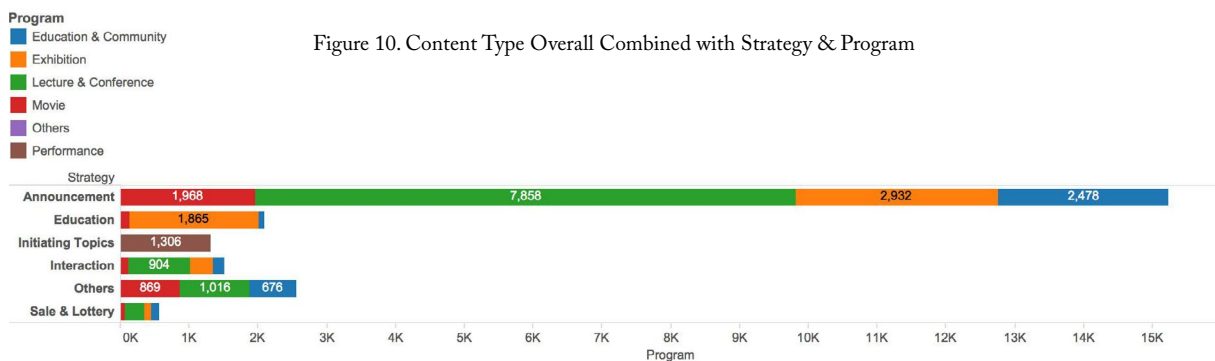


Figure 11. Content Type Overall Combined with Strategy & Program, Not Including Program of Others

4.1.3. Categorization of Weibo Users Who Checked-in at the Museum

By finishing the collection of all Weibo users who has checked-in at these 21 museums, we have a total of 54611 users who has at least one check-in record in the past three years. The basic logic to categorize these users is similar to the categorization of museums' Weibo accounts. This unsupervised learning process will categorize these users based on their number of followers, number of friends, number of Weibo posts and daily average Weibo posts. These four variables describe this user's degree of activeness on Weibo. The same way as we do for the museum accounts, we first try 7 models in order to get a better result, which are K-Means,

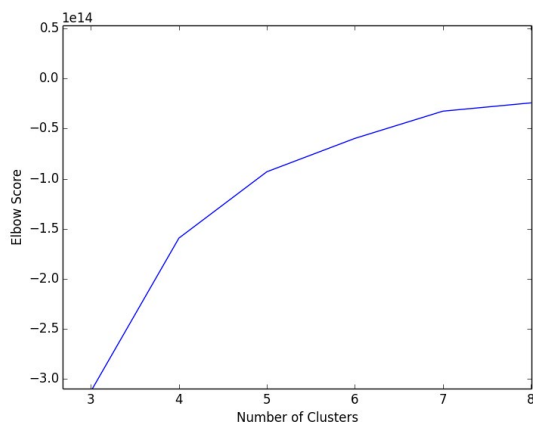


Figure 12. Elbow Score for Users' Model

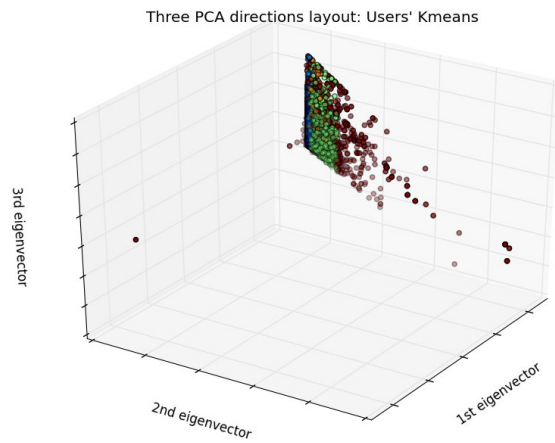


Figure 13. PCA for Users' Model

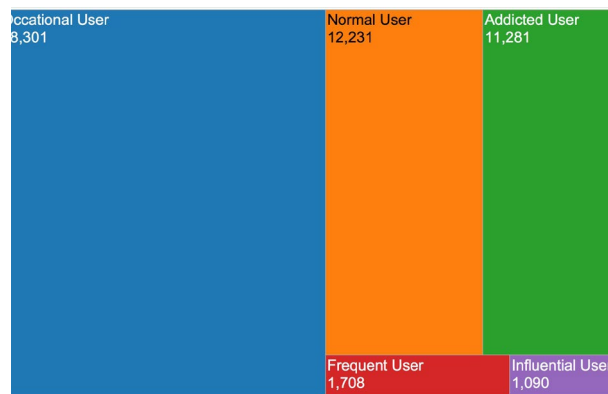


Figure 14. Users' Distribution

User Category	Ave of Followers	Median of Followers	Ave of Friends	Median of Friends	Ave of Posts	Median of Posts
Occasional User	299	204	238	193	555	488
Normal User	1098	603	431	348	1923	1894
Addicted User	1016	757	545	427	6445	4916
Frequent User	14913	12580	641	483	3046	2094
Influential User	273589	63420	1115	899	21556	9592

Table 5. Statistical Description of 5 types of users

Mini Batch K-Means, Spectral Clustering, Affinity Propagation, Hierarchical Clustering of Ward linkage criteria, Complete linkage criteria and Average linkage criteria. After testing the best score for these models, we finally use Mini Batch K-Means as the final model and the number of clusters is set as 5, with the elbow score shown in figure 12. The same as our approach of categorization of museum accounts, we project the result of these four parameters projected in a three-dimension space by principal component analysis, as is shown in figure 13.

With getting these 5 groups of people we rearrange the sequence of category based on their activeness. The result is shown in figure 14 and the statistical description of these 5 groups is shown in table 5. As we can see, occasional users make up for more than 50 percent of the total users, followed by normal users, addicted users, frequent users and influential users. For these 5 groups, the amount of users in the group drops with the increase of average followers, friends and posts, which suggests that most of the users on Weibo platform are occasional users and the active and influential users are only very small portion of the overall users. In the category result, influential users only take up 2% of the total users. One interesting finding about the result is that the major difference between the normal user and addicted user. These two groups have similar number of followers and friends. However, addicted users post their Weibo nearly 5 times more frequent than normal users, and they are even more frequent than frequent users. If we put these users in a three-year timeline, which is figure 15, we can notice that the temporal behavior patterns of these 5 groups of users are similar. They share the same trend in increase and decrease with time. When we look separately into the check-in behavior of Beijing and non-Beijing users, there is a different scenario of their behavior shifting with time can be observed in figure 16.

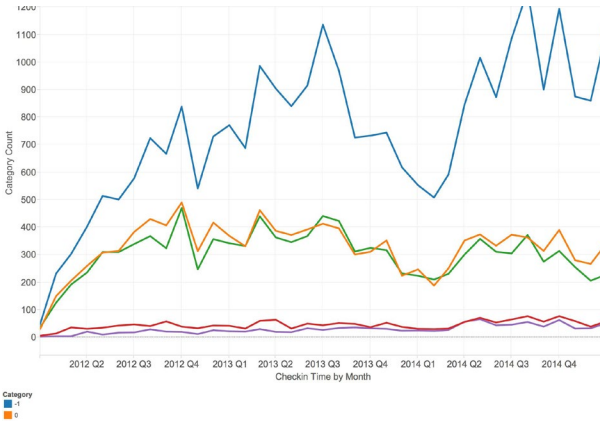


Figure 15. Check-in Frequency by User Category

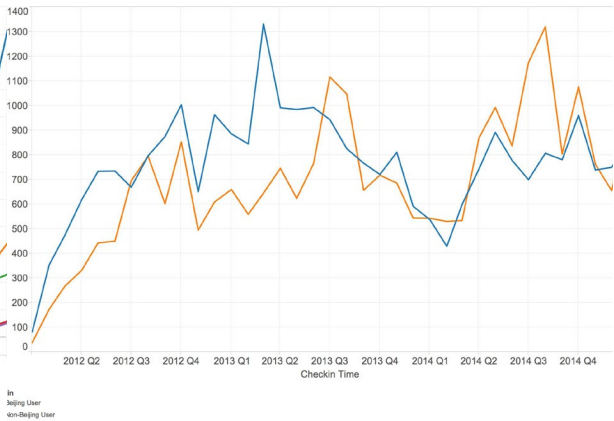


Figure 16. Check-in Frequency by User Origin

4.2 Correlation

4.2.1 Kendall's Tau Correlation

In order to test our origin hypothesis regarding the Weibo marketing efficiency, the most direct way is build a correlation model between the frequency of museum accounts posting Weibo and the frequency of Weibo users checking-in at the museum. If the posting frequency is higher and so is the checking-in frequency, then we can say that the Weibo marketing starts to be effective. The main model we use here to test the correlation is Kendall's tau. Kendall's tau is a coefficient between -1 and 1. When the coefficient is 1 the two sets of data are highly correlated and share the same trend while -1 represents that the two sets of data are highly correlated but have the opposite trend.

Then we calculate the frequency of posting and checking-in for each of 21 museums in 3 ways, which are daily, weekly and monthly. Moreover, for the daily and weekly frequency calculation, we adopt two ways. First is calculating the correlation between the two frequencies on the same day. The second is to calculate the correlation based on the day of posting frequency and its next day of checking-in frequency. The same approach has been done when dealing with the weekly frequency. Thus, each museum gets five results in the correlation calculation, which are daily, next-daily, weekly, next-weekly and month frequency correlations.

Museum Name	Type					Museum Name	Type				
	daily	nextday	weekly	nextweek	montly		daily	nextday	weekly	nextweek	montly
BCM	0.1422	0.0860	-0.0652	-0.0444	-0.1791	BCM	0.0363	0.0000	0.2270	0.4119	0.1244
BEH	0.0718	0.1174	0.1396	0.0416	0.2210	BEH	0.0004	0.0000	0.0097	0.4420	0.0579
BMNatrualHistory	0.1054	0.1509	0.1023	0.1060	0.2052	BMNatrualHistory	0.0000	0.0000	0.0581	0.0503	0.0783
CAFA	0.0987	0.1040	0.0231	-0.0130	-0.0143	CAFA	0.1543	0.8431	0.6688	0.8099	0.9020
Car	0.0238	0.0602	0.0223	-0.1093	0.1835	Car	0.2378	0.0029	0.6797	0.0435	0.1153
Chaoyang	0.0888	0.0077	0.1403	-0.0583	-0.0069	Chaoyang	0.0000	0.7025	0.0093	0.2813	0.9527
CMovieM	0.0320	0.0233	-0.0313	0.0283	0.0870	CMovieM	0.1125	0.2491	0.5620	0.6014	0.4555
Guanfu	0.0084	-0.0457	-0.0107	-0.0047	0.1440	Guanfu	0.6789	0.0235	0.8428	0.9308	0.2167
Guozijian	0.0702	0.0694	0.3009	0.2167	-0.1952	Guozijian	0.0005	0.0006	0.0000	0.0001	0.0940
Hanmeilin	-0.0177	-0.0406	-0.0247	0.0131	-0.0871	Hanmeilin	0.3789	0.0445	0.6473	0.8092	0.4548
MillitaryM	0.0034	0.0036	-0.0125	0.0263	-0.0018	MillitaryM	0.8662	0.8596	0.8161	0.6274	0.9876
NMOC	-0.0504	0.0061	-0.0116	-0.0969	0.0669	NMOC	0.0124	0.7624	0.8303	0.0736	0.5660
NZM	-0.0460	0.0044	-0.0399	-0.0689	0.0527	NZM	0.0227	0.8268	0.4599	0.2034	0.6508
Sanyingtang	0.1440	0.0454	-0.1126	-0.0122	0.1632	Sanyingtang	0.0000	0.0244	0.0370	0.8211	0.1615
Shijitan	-0.1945	-0.0811	0.0226	-0.0298	0.0625	Shijitan	0.0000	0.0001	0.6748	0.5821	0.5920
Songzhuang	0.0543	0.0872	-0.0347	-0.0339	-0.0016	Songzhuang	0.0071	0.0000	0.5200	0.5317	0.9890
TAM	-0.0492	-0.0063	-0.1085	-0.0652	-0.0558	TAM	0.0148	0.7557	0.0444	0.2289	0.6322
UCCA	0.1676	0.1015	0.1012	0.0325	0.0704	UCCA	0.0008	0.0004	0.9823	0.5489	0.5458
Yanhuang	0.0067	0.0201	0.0172	-0.0422	0.1798	Yanhuang	0.7385	0.3196	0.7499	0.4363	0.1229

Figure 17. Kendall's tau between Museum's Posting Frequency and User's Checking-in Frequency

Figure 18. P-Value for the Correlation

Figure 17 and figure 18 show the overall result of Kendall's tau and p-value for all these 19 museums. From figure 17, we find there are several museums having a relatively higher Kendall's tau regardless of the type of frequency, such as Beijing Museum of Natural History and UCCA. If we compare the Kendall's tau by the type of frequency, we will notice that daily and next-day coefficient seem have a higher Kendall's tau compared to other types, which indicates a higher potential of correlation between the posting frequency and checking-in frequency. However, from this table we also notice that there are a lot of museums which have a negative correlation between the posting frequency and checking-in frequency, as their Kendall's tau is less than 0.

Then we move onto the p-value table to decide the confidence level of the result. As in the result of Kendall's tau the monthly frequency has a relatively higher score, similarly, the p-values of all the monthly frequency are larger than 0.05, which is in a lower confidence level. On the other side, the monthly frequency tends to have a higher confidence level, which suggests that the result of the daily and the next-day frequencies are more reliable. This result indicates that the Weibo marketing strategy acts more like an instant notice, as Weibo users are more likely to check-in right after a museum posting its on-going programs.

By looking through the coefficient result we pick out 4 museums to look into their Weibo posts and users who checked-in in these 4 museums, as they have a higher Kendall's tau and confidence level. These 5 museums are Beijing Capital Museum (BCM), Beijing Museum of Natural and History (BMNH), Songzhuang Art Center (Songzhuang) and Ullens Center for Contemporary Art (UCCA).

4.2.2 Case Study

In this section we will look into the composition of the Weibo posts from these four museums' Weibo accounts.

In figure 19, Beijing Capital Museum's Weibo posts mostly focus on their exhibitions. Beijing Capital Museum is a publicly owned encyclopedic museum. The Weibo account of Beijing Capital Museum falls in the category of active and popular. They used a lot of announcements to let Weibo users know their exhibition. At the same time, they introduced a lot of collections and related knowledge of their exhibitions, which can be seen in their education strategy. Also half of their interaction contents are related to the exhibition. To sum up, Beijing Capital Museum tried to make the best use of Weibo to advertise their exhibition programs by introducing their collections and interaction with their followers by talking about related knowledge.

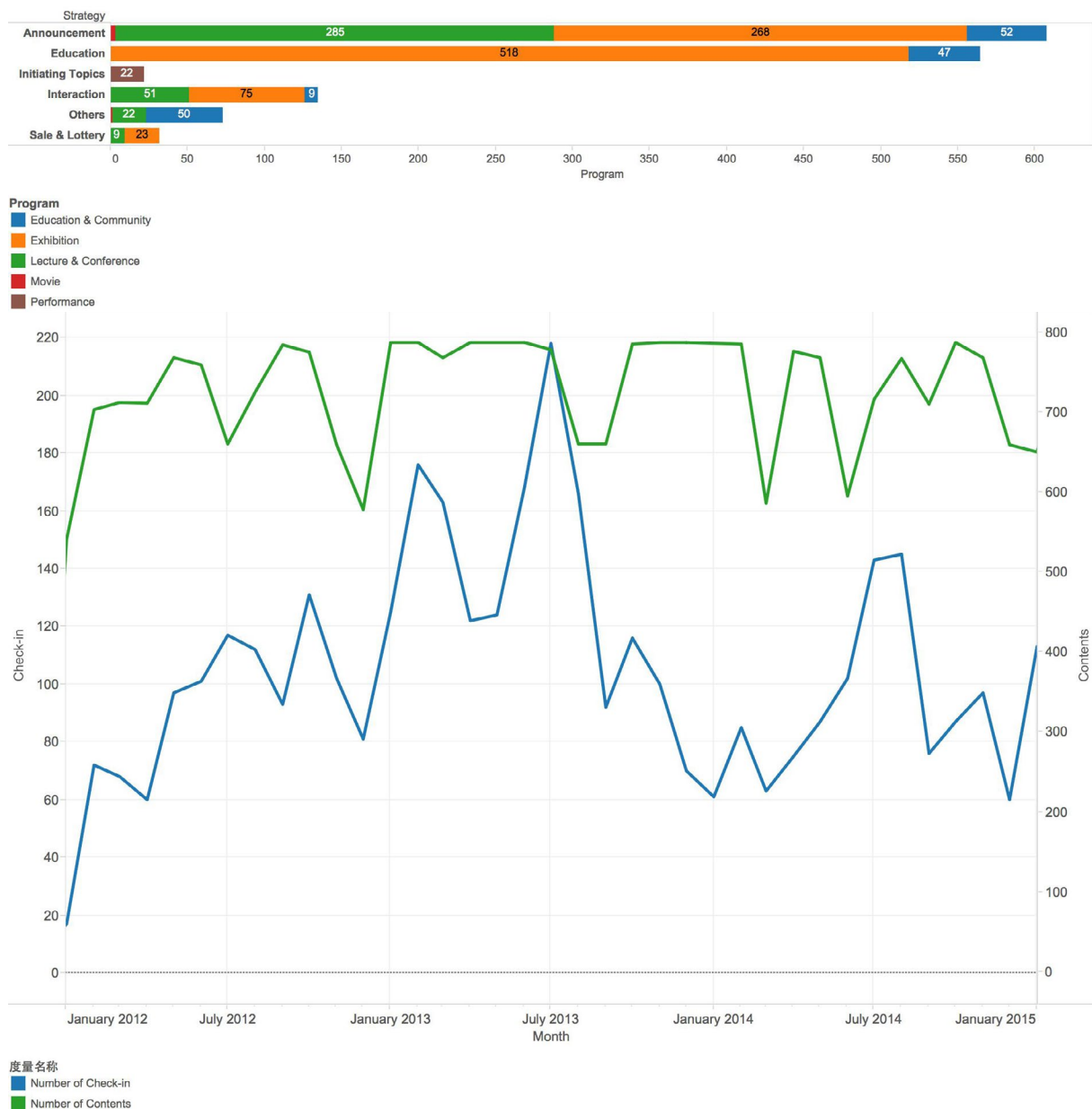


Figure 19. BCM's Weibo Content and Check-in & Post Frequency

According to figure 20, the Beijing Museum of Natural and History used Weibo in a more dynamic way. The BMNH is a public owned specific museum concentrating on natural science. Its account falls in the category of active but unpopular. Their focus is on the conference and lectures when using announcement strategy. At the same time, they cared a lot about the program related to education and communities. When looking into their Weibo account, we find lots of the content are related to family program, which explains there are lots of education & community programs in BMNH's Weibo.

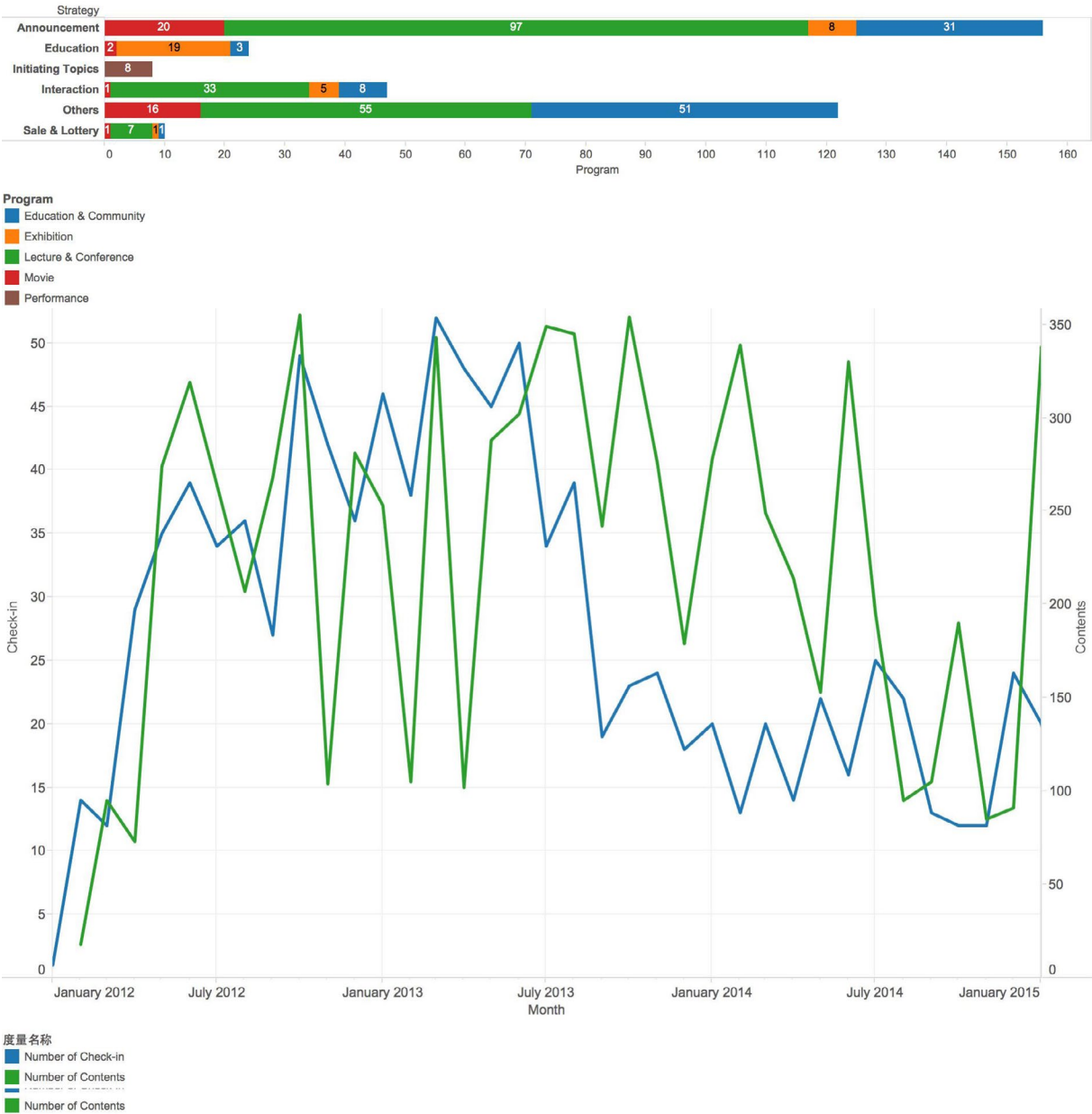


Figure 20. BMNH's Weibo Content and Check-in & Post Frequency

Songzhuang Art Museum is a public owned contemporary museum. Its account belongs to active and unpopular. In figure 21, the main program that published on the Weibo account is conference and education, which turns out to be the news and related information about interactive contemporary art, as it is the concentration for the Weibo account's coverage. Also Songzhuang provided a lot of contemporary art movies as their program published on Weibo. Compared to other museums, Songzhuang tends to have a higher proportion in the strategy of initiating topics.

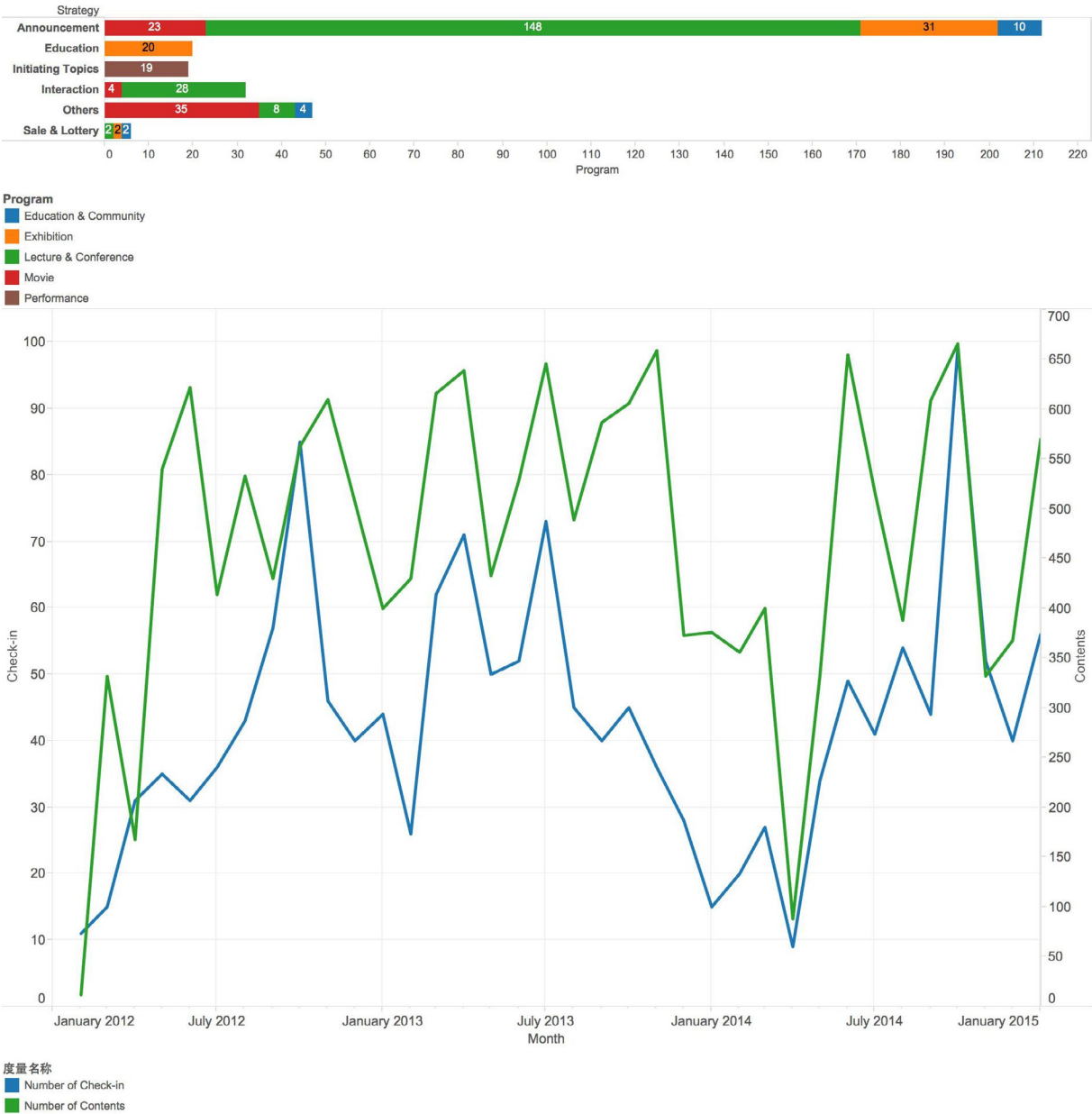


Figure 21. Songzhuang's Weibo Content and Check-in & Post Frequency

Ullens Center for Contemporary Art is a private contemporary museum and its account is active and popular. Similar to Beijing Capital Museum, as shown in figure 22, UCCA uses their Weibo account as a tool of announcement. The main content of their announcement is conference & lectures and movies. UCCA initiates the most of the discussion among these four museums, which is hard to notice because of the large amount of announcement posts.

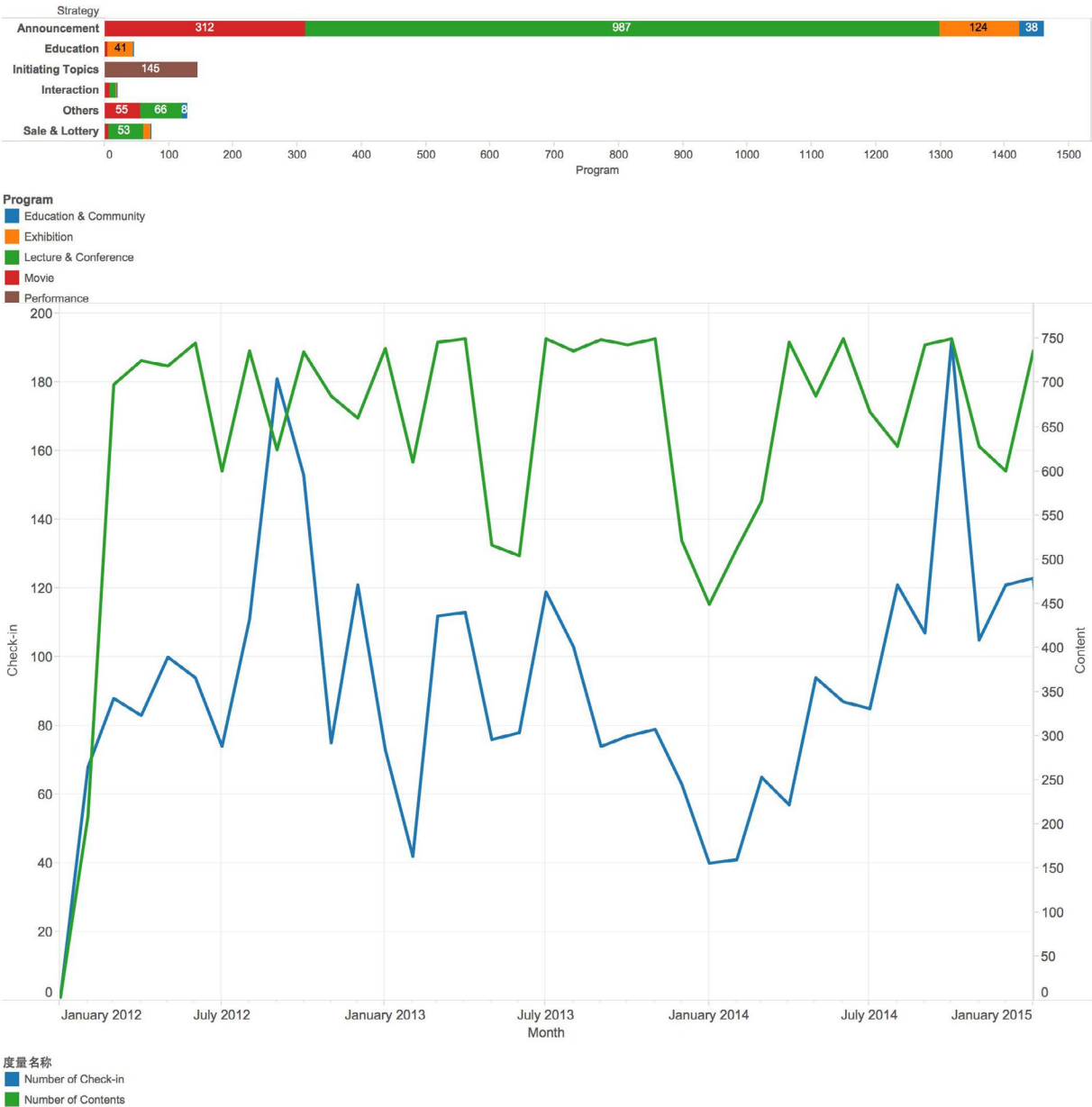


Figure 22. UCCA's Weibo Content and Check-in & Post Frequency

5. Findings & Conclusion

After all of the analysis, we will give some findings concluded from the data analysis, and give some suggestions to museums for their Weibo account maintenance. Also, we will look back to our original hypothesis to conclude the how can Weibo marketing help to attract Weibo users for museums.

The four case studies are the museums that outperformed in the Kendall's tau correlation analysis which show a higher coefficient between the posting frequency and checking-in frequency. By looking into their details as well as the overall dataset we come up with several characters of the successful museum Weibo maintenance.

First, if we look into the ownership of museums in the category result of museum Weibo accounts, we will find that publicly owned museums tend to behave either extremely active or inactive in regard of maintaining their Weibo accounts. On contrary, the behavior style of privately owned is more diverse. This result coincides with the fact that public museums either have assigned staff for Weibo marketing or no staff and no Weibo marketing at all. Also the result matches the Weibo maintenance style of private museums as in most cases they do not have a specific staff for Weibo marketing however the account will be managed by several people, which suggests the diverse Weibo behavior.

Second, from the category result of museums' Weibo accounts we witness that being active on Weibo cannot ensure the popularity of the museum. Museums such as CAFA post their Weibo even more frequently than National Museum of China. However, their category result of popularity falls into the category of the least popular group. On the other side, keeping active on Weibo is necessary for museums as we see only two museums falls into the category of inactive but popular. The fact tells us that there should be a strategy for museums to maintain their Weibo account in order to attract more visitors. Moreover, all the four of our successful cases are located in the middle part of the category result, which indicates a moderate degree of activeness and popularity. This also tells us that the content and quality of Weibo is more important than the degree of activeness. In the frequency's correlation perspective, we also find that the posting frequency will not ensure the increase of user check-ins in comparison of the posting strategies. In our study sample we have several museum Weibo accounts that have posted a large amount of Weibo such National Museum of China, Today Art Museum and CAFA. However, their coefficient results turn out not optimistic although some of them did attract a lot of Weibo users to check-in.

Third, a successful museum Weibo has to have a posting style, which is concentrating on one topic and trying to interact with Weibo users. For Beijing Capital Museum, it focuses on the introduction of exhibition by posting related collections and knowledge to the audience, which can bring more curiosity to Weibo users. Also users will enjoy the new knowledge when there are some interesting facts in it. By posting interesting knowledge, the BCM also increase the degree of interaction with its audiences as they can have more discussion based on the

knowledge and other contents posted by the BCM. Beijing Museum of Natural History uses a different strategy. It focuses on the introduction of natural science knowledge to children and teenagers. When comes to the interaction, it posts a lot of family programs in order to attract both parents and children to the museum. Songzhuang Art Museum's concentration is interactive contemporary art and visual art. The method of increase interaction of Songzhuang is to initiate a lot of discussion with Weibo users. As mentioned before, Songzhuang's Weibo accounts' proportion of initiating discussions is the highest among the four cases. On the other hand, UCCA's Weibo account covers a large portion of salons and news related to contemporary art. Meanwhile, UCCA also tries to initiate more discussions with its audience in a similar way to Songzhuang.

Fourth, by looking into the checking-in frequency in a user-categorical perspective and regional perspective, it is obvious that different categories of users, regardless of occasional user or influential user, tend to have a similar temporal check-in frequency pattern. However, if we look into the temporal check-in frequency pattern of users from Beijing or outside Beijing, the check-in pattern of these two groups of users will be much more different. In the overall users' result, we see that in the most of time check-ins of Beijing users outnumber the check-ins of non-Beijing users. However, there are some specific seasons and time periods every year that non-Beijing users will be outnumber the amount of Beijing users. This indicates that marketing strategy aiming on user of different degree of activeness might not be as effective as the strategy on user from different regions. Museums could aim the target audience as not only in Beijing but also people from other regions.

Fifth, based on the p-value result of the Kendall's tau correlation, it is obvious that the daily and next-daily relationships between posting frequency and check-in frequency tend to be more significant. The result implies that Weibo marketing, to some degree, is a short-term strategy. A piece of Weibo tends to be effective in a relatively shorter time. After a few days the effect of this piece of Weibo will expire. Thus, museums should target their strategy on a short-term basis, preparing a more effective plan to attract people instantly.

However, this research does have its limitations both theoretically and technically. The sample for the research contains 21 museums, which is only a small portion of all the museums in Beijing. Thus, an alternative research method could be applied to examine the whole dataset. Moreover, the machine learning models, especially the supervised learning model when categorizing the museum content, need to be refined. Because the sample data contains more than 45,000 pieces of Weibo, the size of training size, which is 500, is still small compared to the size of the sample data. This could be one of the reasons that cause so much type of others in the prediction result both in strategy and program. Furthermore, the traditional data used in the research is not enough. In the further research, the traditional data will play a more important role as the ultimate intention of the research is to draw the conclusion based on the comparison of traditional data and the social media data.

To sum up, we believe that Weibo marketing, if handled properly, is a tool to help museums

to attract more people, especially people who are Weibo users. For most of the museums, we find that the Kendall's tau is positive although the coefficient is close to 0, which suggests that the relationship between posting and checking-in is not very strong. But we do find out these four successful cases and notice that the Weibo marketing tool is more suitable for attracting people in a shorter term. The daily coefficient and next-day coefficient is higher and more significant than the weekly and monthly coefficient. As a result, it is museums' way of maintenance and posting strategy that will have an important impact on the effect of Weibo marketing. In our study, some museums suddenly stopped their maintenance of Weibo account and it is no doubt the check-in amount shows a decline after that. The tool of Weibo is used by different types of museums and the posting style is dramatically different. However, no matter the museum is owned by the state or the museum focus more on contemporary art, they can always find their own way to gain more attention with the help of Weibo.

Reference:

- Arbesman, S. (2012). *The half-life of facts: Why everything we know has an expiration date*: Penguin.
- Bianchini, F., & Parkinson, M. (1993). *Cultural policy and urban regeneration: the West European experience*: Manchester University Press.
- Cha, M., Haddadi, H., Benevenuto, F., & Gummadi, P. K. (2010). Measuring User Influence in Twitter: The Million Follower Fallacy. *ICWSM*, 10, 10-17.
- . Cultural Development Statistical Report 2013, People's Republic of China. (2014). (000014348/2014-000023).
- . Cultural Reform and Development Plan during the 12th 5-Year Plan. (2012). (000014348/2012-00320). Beijing: Retrieved from http://59.252.212.6/auto255/201205/t20120510_28451.html.
- Di, L. (2013). The Public Influence of Museums. *Southeast Culture*, 3, 017.
- Financial results for the second quarter ended June 30, 2014 of Weibo. (2014). from <http://ir.weibo.com/phoenix.zhtml?c=253076&p=irol-newsArticle&ID=1958713&highlight=>
- Hesmondhalgh, D., & Pratt, A. C. (2005). Cultural industries and cultural policy. *International journal of cultural policy*, 11(1), 1-13.
- Hudson, K. (1987). *Museums of influence*: Cambridge University Press Cambridge.
- Jin, Y., Wang, R., & Chen, X. (2011). A Short-term Interactive Model on the Relationship between Enterprise
- Microblogging Marketing Effect and Followers' Number. *Journal of Management Science*, 24(4), 71-83.
- Kotler, N. G., Kotler, P., & Kotler, W. I. (2008). *Museum marketing and strategy: designing missions, building audiences, generating revenue and resources*: John Wiley & Sons.
- Li, J., Chen, Z., & Huang, J. (2012). Micro-blog Impact Evaluation Study. *Information Network Security*(3), 10-13.
- Liu, X., Wu, M., & Zhu, X. (2012). Ecology Report of Chinese Art Market. *Art Market*(6), 64-69.
- Qian, Y., Zhang, N., Zhao, L., & Zhong, Y. (2012). The Spread of Public Sentiment on Micro-blogging under Emergencies. *Journal of the China Society for Scientific and Technical Information*, 31(12), 1299-1304.
- Shan, S.-l. (2014). Chinese cultural policy and the cultural industries. *City, Culture and Society*, 5(3), 115-121. doi: <http://dx.doi.org/10.1016/j.ccs.2014.07.004>

Special report - Museums: Mad about museums; China. (2013, 2013 Dec 21). *The Economist*, 409, 8-n/a.

. Statistical Report on Internet Development in China (January 2014). (2014). Beijing: China Internet Network Information Center.

Tong, Q. S., & Hung, R. Y. Y. (2012). Cultural policy between the state and the market: regulation, creativity and contradiction. *International journal of cultural policy*, 18(3), 265-278.

Wang, Y., Jin, X., & Cheng, X. (2013). Network Big Data: Present and Future. *Chinese Journal of Computers*, 36(6), 1125-1138.

Weng, J., Lim, E.-P., Jiang, J., & He, Q. (2010). TwitterRank: finding topic-sensitive influential twitterers. Paper presented at the Proceedings of the third ACM international conference on Web search and data mining, New York, New York, USA. http://delivery.acm.org/10.1145/1720000/1718520/p261-weng.pdf?ip=129.236.227.152&id=1718520&acc=ACTIVE SERVICE&key=7777116298C9657D%2ECCAFA7F43E96773E%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35&CFID=591087593&CFTOKEN=64523944&__acm__=1414440936_cd75481bf2712b584f2d5ccd463c07c2

Ying, N. (2010). The National Museum of China is Officially Open to the Public with more than 1.2 Million Pieces of Collections. 2010, from <http://www.chinanews.com/cul/2012/03-01/3711416.shtml>

Yu, L., Asur, S., & Huberman, B. A. (2011). What trends in Chinese social media. arXiv preprint arXiv:1107.3522.

Zhang, G. (2011). The survey on the content of the museum's microblog and its applied value: with the ferences to sina microblog. *Shanghai Science & Technology Museum*, 3(3), 46-53.

Zhang, J., Yu, Y., Ma, L., & Xu, D. (2008). The Study on the Development Strategy of the Museum Industry in Hangzhou City According to the Customer-satisfaction. *Journal of Tourism College of Zhejiang*, 1, 010.

Zhou, X., Lu, J., Li, C., & Du, X. (2012). The Challenge of Big Data from a Perspective of Data Management. *Communication of the China Computer Federation*, 8(9), 16-20.

Zhou, Z. (2012). The application of Weibo during the museum development. *Chinese Museums*, 1, 012.