Academic Journal of Humanities & Social Sciences
ISSN 2616-5783 Vol. 1, Issue 1: 13-22, DOI: 10.25236/AJHSS.040003

Research on the Evaluation of Supporting Ability of "Double-First Class"University Libraries Based on the Principal Component Analysis and Cluster Analysis

Xuhui Ding[*], Yingfei Wang[a]

School of Business Administration, Hohai University, Changzhou 213022, China
[*]Corresponding author e-mail: dingxh@hhu.edu.cn, [a]510558867@qq.com

ABSTRACT. *Based on the data at the end of 2016 and selecting 12 variables such as building area and paper resource purchase fee, this paper uses principal component analysis and cluster analysis to evaluate the support ability and overall strength of 42 universities and colleges under "double-first class" construction. At last, ranking and classifying the university libraries according to the research results. This study has shown that the number of staff members, the purchase fee of electronic resources, paper resources, literature resources, the number of published paper and the number of paper resources are the main factors affecting the construction of university libraries. The gap in support ability among university libraries in the same region, different provinces and cities, and different types are obvious; the imbalance of libraries' own strength leads to their inefficiency; the strength of university libraries does not match the strength of scientific research. It is recommended that universities and colleges should allocate funds reasonably, increase the investment and attention in libraries, and strengthen exchanges and cooperation between university libraries in and between regions.*

KEYWORDS: *double-first class, university libraries, principal component analysis, cluster analysis*

1. Introduction

University libraries are important institutions aimed at providing information and related services for university teaching and scientific research [1]. In the 1970s, Eliot, the President of Harvard University called the library "the heart of the university" [2]. In the report of the 19th National Congress of the Communist Party of China, Comrade Xi Jinping pointed out that it is necessary to speed up the construction of world-class universities and first-class disciplines, strengthen the cultivation of talents and introduction of resources in universities and colleges, and the

ISSN 2616-5783 Vol. 1, Issue 1: 13-22, DOI: 10.25236/AJHSS.040003

construction of world-class universities and first-class disciplines cannot be done without the support of university libraries. Taking 42 “double-first class” universities and colleges for an example, the total area of library buildings is about 240,000 square meters, along with nearly 100 million paper resources. However, the utilization efficiency of libraries sometimes cannot match the development speed of universities and colleges, and there are problems such as low resource utilization rate and loose management, which does not make the library really play its role. There is a significant correlation between the core competitiveness of university libraries with the teaching level and scientific research ability of universities and colleges (Lin Kequan [1], 2014). However, the development level of Chinese university libraries is quite different. Most libraries cannot effectively use their funds, and financial support has not been completely transformed into the library's own strength. There is a large gap in the performance of university libraries in comprehensive, science, engineering, and normal universities and colleges, and the development is uneven (Jiang Yumei [2], 2014). Therefore, it is necessary to conduct customer observation and evaluation on the efficiency of university libraries, so as to provide reasonable suggestions on improving the input-output ratio of libraries, rationally utilizing existing resources, and improving resource utilization efficiency.

2. Research Methods and Indicators Selection

2.1 Research Methods

The principal component analysis is to use the idea of dimensionality reduction to transform multiple indicators into a few comprehensive indicators, each of which can reflect most of the information of the original variables, and the information contained is not repeated. First, the data is standardized to remove the influence of different dimensions on the data values; assuming that the original data X has p observations, n variables, the covariance matrix is calculated by formula (1), as well as the eigenvalue $\lambda_1, \lambda_2, \ldots \lambda_n$, the eigenvector corresponding to the eigenvalue is the eigenvalue and eigenvector of the main component’s load factor aij; finally, the principal component number is determined by the cumulative contribution rate, and the score of observation sample on the principal component Yi is calculated by the formula (2).

$$V_{ij} = \frac{1}{p-1}\sum_{k=1}^{p}(x_{ki} - \overline{x_i})(x_{kj} - \overline{x_j}) \tag{1}$$

$$Y_i = a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n \tag{2}$$

The cluster analysis is based on the basic characteristics of data, to calculate the internal correlation among the data, mainly through the calculation of some statistical indicators, to reflect the degree of similarity between the reaction data. The basic principle of cluster analysis is to calculate the distance between two of all variables, and classify the variables according to the distance. First, classify all the

ISSN 2616-5783 Vol. 1, Issue 1: 13-22, DOI: 10.25236/AJHSS.040003

samples into one class. For data containing n variables, include n classes. Secondly, calculate the distance between any two of the above n classes. According to the principle that the minimum distance is one class, the two classes with closer distance are clustered into one class, and this is the n-1 class. And so on, the two classes with the smallest merge distance are clustered into one class, until all the data are classified into one class [3]. In this paper, first calculate the Euclidean distance between different variables, as it can be seen from equation (3), and then perform cluster analysis.

$$d(\vec{x_i},\vec{x_j}) = \sqrt{\sum_{k=1}^{p} (x_{ik} - x_{jk})^2} \tag{3}$$

2.2 Data Selection

The supporting ability of university libraries is mainly reflected in the input and output. In the past, most of the research conducted indicators selection from only one aspect, such as scientific research strength, thesis index, service quality, etc. rarely carried out comprehensive research on various indicators. Gao Haitao et al. [4] evaluated universtiy libraries from the perspective of service quality. Liang Ru [5] established a system from funds, staff, facilities investment, information services, teaching and research. In this paper, the input indicators are the area of the building, the number of staff, the total annual funds, etc., the output indicators are the number of published papers, etc. the quantity and quality of library resources are also taken into consideration, as shown in Table 1. The relevant data comes from University Library of the Ministry of Education's fact database and the official website of the university. The research selects the data at the end of 2016, and some missing data will be replaced or smoothed by data from similar years.

The selection of indicators is as follows: ① The building area is one of the basic indicators of library construction, and it is an important place to store a large amount of paper resources and large-capacity storage for storing electronic databases; ② Library seats are the main place for students to learn by themselves. The number of seats should meet the needs of students in colleges and universities; ③ The number of hours of opening in weeks can reflect the length of service provided by the library; ④ The number of staff is directly related to the service and management quality of the university library; ⑤The purchase of literature resources and total annual funds can show the investment of library resources in colleges and universities, whether they pay attention to the resources construction of libraries; ⑥ The more developed the economy of the city, the more attention is paid to education, and the more funds are invested; ⑦ The number of paper published reflects the overall scientific research level and ability of the library, which is closely related to the scientific research level of the university [6]; ⑧ The amount of paper resources and electronic resources can measure the level of document construction [7], which intuitively reflects the achievements of the university in the basic resources construction and

ISSN 2616-5783 Vol. 1, Issue 1: 13-22, DOI: 10.25236/AJHSS.040003

contribution to university research capacity; ⑨ The number of students can estimate the approximate number of visitors to the library and its swipe.

Table 1 Evaluation indicators of supporting ability of "double first-class" university libraries

Code	Indicator name	Code	Indicator name	Code	Indicator name
X1	Building area	X6	Annual total funding	X11	Paper resources number
X2	Number of staff	X7	Library seat number	X12	Chinese & foreign databases
X3	Electronic resource fee	X8	City GDP	X13	Electronic resources number
X4	Paper resource fee	X9	Weekly opening hours	X14	Current student quantity
X5	Document resource fee	X10	Paper published number		

3. Empirical Measurement and Results Analysis

3.1 Results of Principal Component Analysis

Table 2 Description of eigenvalues and contributions of various indicators in in library support

Variable	Eigenvalues	Cumulative contribution rate(%)	Variable	Eigenvalues	Cumulative contribution rate(%)
1	4.835	34.54	8	0.510	90.01
2	2.025	49.00	9	0.416	92.98
3	1.659	60.86	10	0.314	95.23
4	1.178	69.27	11	0.227	96.85
5	0.984	76.30	12	0.194	98.24
6	0.783	81.89	13	0.150	99.30
7	0.628	86.38	14	0.097	100.00

Based on the results of the above principal component analysis, the number of staff members, the purchase fee of electronic resources, paper resources, literature resources, the number of published paper and the number of paper resources have contributed greatly to the research, and the cumulative contribution rate of six indicators is 81.89%, which can better explain the original variables. Other variables are also the evaluation factors, which can give assistance to the main six factors, and help better evaluate the support ability of universities and colleges under “double-first class” construction. According to data, we can get score of every main component. The result of the calculation is divided into two tables, shown in Table 3

ISSN 2616-5783 Vol. 1, Issue 1: 13-22, DOI: 10.25236/AJHSS.040003

and 4.At the same time, according to the scores of different variables in different components, the scores of "double-first class" university libraries can be calculated and sorted. The specific results are shown in Table 5.

The research results show that ① Among the "double-first class" colleges and universities, the B-class university libraries are slightly behind the A-class universities, and the B-class colleges and universities have weaker support for the libraries, which makes the library develop slowly, such as the total funding of the library of Xinjiang University has a large difference with other colleges and universities, and the number of paper resources and electronic resources is seriously insufficient, which does not meet the demands of teachers and students. ② The top 10 universities are mostly eastern university libraries, mainly in Beijing. Shanghai, Jiangsu Province, these universities have greater support for the library, the library is more balanced in all aspects, and can meet the daily needs of teachers and students. While in the central and western regions, only Wuhan University, Huazhong University of Science and Technology and Sichuan University perform well, other colleges and universities need to strengthen financial support; ③ There is a big gap between different types of universities, normal university libraries (Beijing Normal University and East China Normal University) are under rapid development, the library hardware facilities can meet the needs of teachers and students for scientific research. Meanwhile, their support for library is great with the total annual expenditure ranking in the forefront, other types of colleges and universities have no obvious distribution characteristics; ④ Some universities' ranking is relatively backward, which do not match their research and teaching capabilities, such as Tsinghua University. Although Tsinghua University Library has strong scientific research strength and has published many papers, its number of students is large and the number of seats in the library is small. This may also be related to the lack of scientific selection of indicators.

Table 3 The matrix of socre 1

variable	prin1	prin 2	prin 3	prin 4	prin 5	prin 6	prin 7
1	0.222	-0.379	-0.144	0.123	0.125	-0.196	0.352
2	0.302	-0.280	-0.303	-0.166	-0.009	-0.008	0.330
3	0.319	0.147	-0.036	0.367	0.049	0.496	-0.167
4	0.373	0.142	-0.002	-0.023	0.035	-0.229	-0.401
5	0.327	0.337	-0.259	0.058	0.003	0.159	-0.018
6	0.298	0.273	-0.374	-0.045	-0.201	0.099	0.117
7	-0.034	-0.076	-0.004	0.806	0.361	-0.199	0.041
8	0.113	0.467	0.256	0.055	-0.050	-0.252	0.650
9	-0.058	0.100	-0.236	-0.351	0.844	0.055	0.056
10	0.327	-0.049	0.287	-0.107	-0.038	-0.455	-0.134
11	0.379	-0.149	-0.141	-0.087	0.099	-0.279	-0.278
12	0.286	0.133	0.403	-0.130	0.142	0.219	0.116
13	0.160	-0.164	0.560	-0.059	0.194	0.331	0.038
14	0.220	-0.495	-0.019	-0.003	-0.159	0.282	0.164

Table 4 The matrix of socre 2

variable	prin 8	prin 9	prin 10	prin 11	prin 12	prin 13	prin 14
1	-0.753	0.008	-0.048	-0.102	0.055	0.098	-0.004
2	0.294	-0.044	-0.104	0.393	-0.516	-0.275	0.084
3	-0.807	0.019	-0.099	-0.437	-0.409	-0.017	0.298
4	-0.147	0.042	0.465	0.413	-0.024	0.224	0.405
5	-0.131	0.003	0.029	0.116	0.455	-0.647	-0.161
6	0.118	0.139	-0.384	0.106	0.198	0.614	-0.151
7	0.289	-0.089	-0.051	0.224	0.105	0.066	-0.113
8	0.124	0.226	0.292	-0.120	-0.069	-0.006	0.130
9	0.112	0.0004	-0.001	-0.158	0.110	0.084	0.175
10	0.153	-0.211	-0.538	-0.218	0.185	-0.142	0.332
11	0.163	0.299	0.200	-0.352	-0.170	-0.006	-0.594
12	-0.072	-0.673	0.081	0.123	-0.081	0.161	-0.357
13	-0.041	0.567	-0.215	0.318	0.106	-0.014	-0.199
14	0.345	-0.095	0.376	-0.216	0.456	0.103	0.199

Table 5 Ranking of "double-first class" University Library Support Ability Measurement

University	Score	Rank	University	Score	Rank
Zhongshan University	8791372	1	Jilin University	2687930	22
Fudan University	7951467	2	Northwestern Polytechnical University	2636715	23
Wuhan University	7762838	3	National University of Defense Technology	2505276	24
Beijing University	6624861	4	Renmin University of China	2474998	25
Zhejiang University	6439945	5	Nankai University	2433219	26
Beijing Normal University	5504325	6	Ocean University of China	2430029	27
East China Normal University	54286345	7	Jilin University	24085055	28
Sichuan University	53825935	8	Harbin Institute of Technology	2269509	29
Huazhong University of Science and Technology	50144395	9	Zhengzhou University	21541547	30
Nanjing University	48167745	10	China Agricultural University	2147475	31
Southeast University	4399047	11	Northwest A&F University	2146426	32
Xi'an Jiaotong University	4235625	12	Lanzhou University	2082711	33
Shanghai Jiao Tong University	4095997	13	University of Science and Technology of China	1959878	34
Tongji University	3607840	14	Hunan University	19343288	35
Northeastern University	3456607	15	Yunnan University	1700347	36
Central South University	3330853	16	Shandong University	1651697	37
Northwestern Polytechnical University	3090688	17	Beijing Institute of Technology	1557905	38
Xiamen University	3042478	18	Dalian University of Technology	1358392	39
University Of Electronic Science And Technology Of China	2967051	19	Minzu University of China	1234691	40
Chongqing University	2952488	20	Tsinghua University	616628	41
Beihang University	2921313	21	Xinjiang University	117	42

ISSN 2616-5783 Vol. 1, Issue 1: 13-22, DOI: 10.25236/AJHSS.040003

3.2 Results of cluster analysis

According to the results of cluster analysis, it could be seen from Table 6 that the "double-first class" universities and colleges can be divided into four categories by the library supporting ability: ① Class I. The utilization efficiency of these 10 university libraries is in the forefront of all the double-first class universities and colleges. It provides a large amount of financial support for the library's infrastructure and resource purchase. Facility development and resource development along with each other, while meeting the needs of teachers and students, it also enhances its scientific research strength. Most of these university libraries have an area of 50,000-60,000 square meters, with an annual total funding of 40 million yuan, an average of 5,000 seats, an average of 5.5 million paper resources, and more than 300 Chinese and foreign databases. The human resources, material resources and financial resources invested by these schools are well transformed into teaching, scientific research, and library resources construction. And their utilization efficiency is the leader of the country. ② Class II, the utilization efficiency of these 30 university libraries is similar to each other. Universities and colleges have also invested a large amount of funds to support the development of libraries, but they have not paid attention to the simultaneous development of facilities construction and resource construction. Their resource input, teaching and research output are at a medium level, and there is a mismatch between these two. The purchase fee of literature resources is more than 14 to 30 million yuan, the total annual expenditure is more than 25 to 32 million yuan, the number of seats is about 3,000, and the amount of paper resources is between 2 million and 5 million. ③ Class III and Class IV are special universities and colleges with an average area of 30,000 square meters. The purchase fee of electronic resources and paper resources are about 5 million yuan, the purchase fee of literature is about 27 million yuan. And the total annual expenditure is about 26 million yuan, basically meet the daily needs of teachers and students. The resource input is mainly transformed into the infrastructure construction, which has not been completely transformed into the teaching & research needs [8], and the library efficiency needs to be improved.

ISSN 2616-5783 Vol. 1, Issue 1: 13-22, DOI: 10.25236/AJHSS.040003

Table 6 Classification of Supporting Capabilities of "Double-First Class" University Libraries

Category	Name of university and college
Class I	Peking University, Beijing Normal University, Fudan University, East China Normal University, Shanghai Jiaotong University, Huazhong University of Science and Technology, Zhejiang University, Sichuan University, Wuhan University, Zhongshan University
Class II	Renmin University of China, Tianjin University, Harbin Institute of Technology, University of Science and Technology of China, Northwest A&F University, Northwestern Polytechnical University, Northeastern University, Hunan University, South China University of Technology, Chongqing University, Lanzhou University, Yunnan University, Ocean University of China, Central South University Nanjing University, Southeast University, Xi'an Jiaotong University, Tongji University, Beijing University of Aeronautics, China Agricultural University, Nankai University, University of Electronic Science and Technology, Beijing Institute of Technology, Dalian University of Technology, Zhengzhou University, Tsinghua University, Xinjiang University, Jilin University, Shandong University, Xiamen University
Class III	Minzu University of China
Class IV	National University of Defense Technology

4. Conclusions and Recommendations

This paper uses principal component analysis and cluster analysis method to evaluate and compare the library support ability of 42 “double-first class” universities and colleges. The results show that ① The main factors affecting the library support capacity are the number of staff members, the purchase fee of electronic resources, paper resources, literature resources, the number of published paper and the number of paper resources; ② The gap in the utilization efficiency of university libraries in different regions is large, showing a stepped distribution in the eastern, central and western regions; ③ Different types of university libraries has obvious utilization efficiency gap. The performance of the library support ability of the normal colleges is excellent. ④ The unbalanced strength of the university libraries’ own leads to the inefficiency of the library, and the strength of some university libraries is not matched with the scientific research strength. ⑤ The main reasons that restrict the efficiency of the university libraries is whether the capital investment has a higher return and whether the libraries' own construction is matched with education and scientific research strength.

The author thinks that the support ability of "double-first class" university libraries should be improved from the following aspects: ① Reasonable allocation of library construction funds. It should be clear about its own shortcomings, starting

from the main driving factors of efficiency, increase financial support, balance development, improve the efficiency of the library itself and meet the demands of its own universities. All universities and colleges can optimize the proportion of capital investment according to their own shortcomings, appropriately increase the opening time of the library, increase the number of databases, and meet the needs of teachers and students; ② Increase the capital investment and attention of university libraries. The country should increase financial support and guarantee the supply of educational resources in the central and western regions, pay attention to the enhancement of scientific research capabilities of libraries, and appropriately increase the recruitment of highly educated and highly qualified librarians; ③ Strengthen exchanges and cooperation between university libraries in and between regions. University libraries with higher efficiency in the east should exchange with lower efficient university libraries in the central and western regions or in the same region, to achieve resource sharing and common progress. Carry out one-on-one assistance and cooperation, the eastern university libraries can be docked with the western universities with less development to achieve common progress; ④ Strengthen data statistics work and open it to the public. University libraries should strengthen their own data statistics work, find out the deficiencies in library efficiency construction through data analysis and improve them, so as to improve library efficiency and give teachers and students more help.

Acknowledgments

This research was financially supported by the China Postdoctoral Science Foundation Project (Grant NO.2017M621622) and supported by "the Fundamental Research Funds for the Central Universities" (Grant NO. 2018B23614).

References

[1] Lin Kequan, Kuang Wanling, Gao Bo. Study on Correlation between Universities's Research Ability and its Library's Core Competence: Taking China's 211 Project Universities for Example. Library and Information Service, Vol. 58 (2011), No.24, p.5-10.

[2] Jiang Yumei, Tian Jingmei, Li Xinyun. Research on Performance Evaluation of University Libraries Based on CRITIC-TOPSIS Methods, Library Tribune, Vol.2018, No.3, p.1-9.

[3] Qi Yanhong, Li Qianqing.A Comprehensive Evaluation on Major University Libraries in China, Journal of Modern Information, Vol.30 (2010), No.7, p.103-106.

[4] Gao Haitao, Xu Kaiying, Li Han. Research on the Evaluation System of University Libraries Based on Super Efficiency DEA, Library and Information Service, Vol.58 (2014), No.5, p.17-21.

[5] Liang Ru, Li Jianxia, Liu Ying, Liu Yun, Sun Xiaoxing. Evaluation Study on Digital Resources Comprehensive Service Ability in University Libraries, Journal of Academic Libraries, Vol.2015, No.2, p.38-46.

[6] Zhang Guochen. Research on the Correlation Between the Amount of Digital Resource with Teaching and Researching in University Library--Take Domestic Finance and Economics Universities for Example, Library and Information Service, Vol. 55 (2011), No.7, p.90-93.
[7] Zhao Yinghong. Assessment and Countermeasure of University Library Competitiveness Based on Empirical Research of 52 Universities, Journal of Academic Libraries, Vol. 2011, No.4, p.38-44.
[8] Sun Zhijing, Tian jingmei, Li Xinyun. Research on Performance Evaluation of University Library Construction to Academic Discipline Development in the "Input—Output" Perspective —— Based on Universities under the Direct Supervision of Ministry of Education, Library Tribune, Vol. 2016, No.5, p.63-71.