

# Analysis and sequencing of the influence factors of the value of review on the low carbon Library Line Based on SPSS data analysis

Mengxin Yi<sup>1,a,\*</sup>

<sup>1</sup> Department of economic management, North China Electric Power University, Baoding 071003, China;

<sup>a</sup> 790769083@qq.com

**Keywords:** Online review sorting, Factor analysis method.

**Abstract:** In order to provide users with the most valuable reviews and help them choose their favorite books in mass book online reviews, we need to develop a ranking method to get more effective reviews from a large number of online reviews. The three children's books Dangdang online review data, statistics on the 9 index will affect the effectiveness of the review, spss19.0 as analysis tools, through descriptive statistical analysis, crosstables analysis, factor analysis and statistical analysis method, nine indexes extracted five factors, and their assignment, calculation according to the comprehensive score, ranking score for comments. The results show that the comments with high scores are informative and can provide more effective information to the user.

## 1. Background introduction

In a low carbon library and other similar online library for student borrowing books, parents often read the book reviews to determine whether their children will read this book. Therefore, regardless of the value of the review, it is important for parents to read a small amount of comments and obtain more complete and real high value information. At present, there are many studies on the rankings of reviews. Comments on applications are emerging, such as the American group, public comments, and so on. However, the research on the online evaluation of books is still a blank. By selecting 3 different children's books Dangdang network, the paper reviews from 300 aspects, completes information collection, data processing, data analysis, systematically analyzes the factors that influence the value of online reviews, and finds out the key points hidden under the link of data and data. For the current study reviewed the existing ranking methods are mostly to guide readers named the target consumer decision, ranking method of reading platform this paper mainly studies the books online review of low carbon library, in order to quickly from online comments by a lot this method to help users to meet the user's personalized comments to obtain higher utility choice you most love the book.

### 2.1 Descriptive statistics

Frequency analysis is used to describe the factors that affect the factors of different kinds of books.

### 2.2 Cross contingency table analysis

Through the collection of sample data, a cross - contingency table is produced to analyze the popularity of different books.

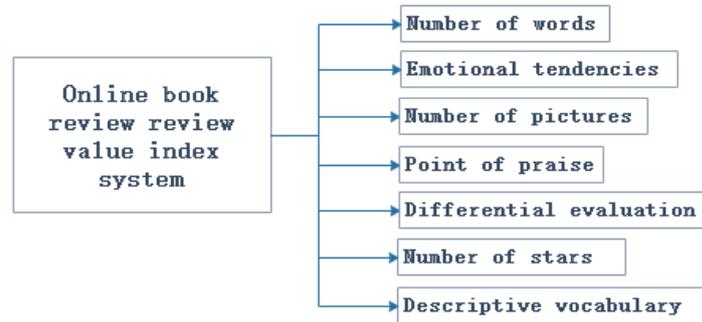
### 2.3 Factor analysis

It's a data reduction technology. By studying the internal dependence of many variables, we explore the basic structure of observation data, and use a few hypothetical variables to represent its basic data structure.

### 2.4 Research model

In the process of data processing, this paper first studied the factors affecting the usefulness of the book review of online reviews, the research about this more, different scholars analysis the influencing factors from different angles, they through a large number of examples, to explore the relationship between various factors and the usefulness of online reviews. For example, Liu and others [1] think that the experience of commentators, the style of commentary and the limitation of comments are useful indicators for consumers' purchase decisions. S., M., Mu-dambi and others [2]

studied the 1587 comments on Amazon website, and found that the depth of reviews, the polarity of reviews and the type of merchandise have an impact on the usefulness of online reviews. Yin Guopeng et al [3] to adopt information theory as the research framework, research found that the comments are related with the length of usefulness, comment star is negatively correlated with extreme usefulness; inverted U relation between reviewers centrality and helpfulness of the review, published historical commentators Never mind between quantity and usefulness. Guo et al. [4] think mobile O2O online reviews the usefulness of the indicators should be critics and readers, review content, review review time for consideration of 4 dimensions, including specific measures: authority, the reader comments reviewers characteristics, review length, product attributes words and emotional characteristics of words, number of pictures and the usefulness of vote and comment timeliness. Comprehensive consideration, the analysis of the characteristics of books about the evaluation, this paper decides to adopt the following indicators:



### 3. The results of the study

#### 3.1 Statistical analysis of sample description

Table 1 Descriptive statistics

Descriptive Statistics						
Title		N	Minimum	Maximum	Mean	Std. Deviation
Chinese history for children	Number of words	300	9	230	85.63	31.421
	Emotional tendencies	300	1	4	3.21	.600
	Number of pictures	300	0	5	.56	1.162
	Point of praise	300	0	154	2.50	11.704
	Differential evaluation	300	0	335	3.22	23.309
	Number of stars	300	8	10	9.98	.152
	Descriptive vocabulary number	300	0	7	2.52	1.394
	A new degree of comment	300	31	1289	198.33	132.796
	Valid N (listwise)	300				
Magic school bus, bridge book	Number of words	300	21	170	65.86	31.617
	Emotional tendencies	300	1	4	3.30	.766
	Number of pictures	300	0	7	.49	1.132
	Point of praise	300	0	63	1.03	5.007
	Differential evaluation	300	0	54	.79	4.356
	Number of stars	300	4	10	9.91	.523
	Descriptive vocabulary number	300	0	5	.90	.898
	A new degree of comment	300	26	710	175.71	117.567
	Valid N (listwise)	300				
Nobel Youth Edition	Number of words	299	21	140	51.16	23.682
	Emotional tendencies	300	1	4	3.05	.558
	Number of pictures	300	0	5	.21	.762
	Point of praise	300	0	95	1.77	8.759
	Differential evaluation	300	0	94	.99	6.041
	Number of stars	300	0	10	9.96	.588
	Descriptive vocabulary number	300	1	13	1.33	.916
	A new degree of comment	299	30	286	175.33	64.609
	Valid N (listwise)	299				

The form lists the collected comments, each book includes three books, including the significant difference between word and emotion tend to average, mode and minimum and maximum. Among them, the effective value of children's Chinese history is 300, the missing value is 0, the average number of words is 85.63 reviews, the pattern is 78, the standard deviation is 31.421, the minimum value is 9, the maximum value is 230, similarly, other data can be read out. All the other two books can be read in the next table. "Emotional expression of love for the child most readers China history", star rating was 10, and 89 year old children reading; emotional expression of love most readers "- the book version of the magic school bus bridge", a score of 10 stars, reading and children at the age of 5; most of the readers "Nobel youth version of" emotional love star a score of 10, and 8 year old children, reading.

**3.2 Cross contingency table analysis**

In order to further understand the popularity of a book in the readers, this paper also uses a cross contingency table to analyze the correlation between books and emotional inclinations.

Table 2 Emotional tendencies \* Title Crosstabulation

Count		Title			Total
		Chinese history for children	Magic school bus, bridge book	Nobel Youth Edition	
Emotional tendencies	Dislike	3	10	1	14
	Generally like	20	26	36	82
	like	188	127	209	524
	Like it very much	89	137	54	280
Total		300	300	300	900

The analysis of the data from the table shows that the three books are generally high in the reader's preference, and are most loved and loved by the readers. Studies show that the reviews and comments on extreme emotions such as love and love are more valuable to readers. Therefore, the comment on book review three is more abundant and convincing from the higher degree of data.

**3.3 Factor analysis**

**3.3.1 Correlation test**

The collected data are measured by KMO and Bartlett ball test. The results are shown as shown in the table.

Table 3 KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.538
Bartlett's Test of Sphericity	Approx. Chi-Square	841.686
	df	28
	Sig.	.000

Result analysis: Kmo shows no good, but Ba can get Bartlett's sphericity test chi square statistical significance is 0, less than 1%, the result shows that the data is correlated, suitable for factor analysis.

**3.3.2 Structural factor variable**

Table 4 Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1.901	23.766	23.766	1.901	23.766	23.766
2	1.300	16.245	40.011	1.300	16.245	40.011
3	1.234	15.420	55.431	1.234	15.420	55.431
4	1.094	13.680	69.111	1.094	13.680	69.111
5	.820	10.245	79.356	.820	10.245	79.356
6	.764	9.548	88.905			
7	.544	6.805	95.709			
8	.343	4.291	100.000			

Extraction Method: Principal Component Analysis.

Can be seen from the table, according to the characteristic values of more than 0.8 and the percentage of more than 79% information interpretation as the standard, extracted a total of 5

common factors, can be found to explain the sub public information related to the 23.766% heat 16.245%, 15.420%,13.680% and 10.245%, finally accumulated can explain the general information of the 79.356%. This result shows that five public factors can better reflect the overall information.

**3.3.3 Scree plot**

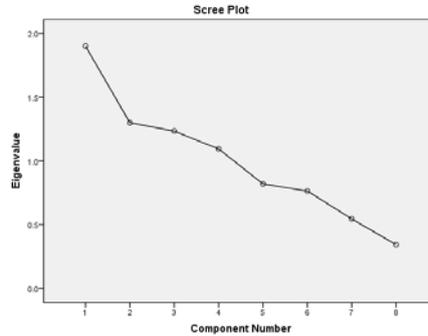


Fig. 1 Scree plot

As you can see from the graph, the components 1, 2, 3, 4, 5 contain most of the information and enter the platform from the beginning of 6. According to the factor analysis feedback component matrix, the meaning and composition of various public factors are analyzed. The results are as follows:

**Table 5 Component Matrix**

	Component				
	1	2	3	4	5
Number of words	.636	.487	-.069	-.230	.003
Emotional tendencies	.061	.311	.613	.279	.643
Number of pictures	.039	-.097	.542	-.644	.036
Point of praise	.790	-.413	.060	.148	.019
Differential evaluation	.741	-.470	.109	.220	-.035
Number of stars	-.104	.300	.387	.644	-.420
Descriptive vocabulary number	.468	.565	.165	-.214	-.359
A new degree of comment	.297	.393	-.606	.140	.312

Extraction Method: Principal Component Analysis.

a. 5 components extracted.

The value in the upper table is the correlation coefficient between the common factor and the original variable. The greater the absolute value, the closer the relationship is. Is the first common factor and the point of praise (0.790), the differential evaluation (0.741) is higher; Second the common factor has a high correlation with the number of words (0.487) and descriptive vocabulary number (0.565); third common factors and the number of pictures (0.542), a new review of comment (-0.606) there is high; fourth common factors and the number of stars (0.644) have a higher correlation; fifth common factors and emotional tendency (0.643) the relationship between.

The first factor named "other user evaluation" (F1), the second factor named comment integrity (F2), the third factor named "comments attraction" (F3), the fourth factor named "Star" (F4), the fifth factor named "emotion" (F5).

**3.3.4 Calculation factor score**

**Table 6 Component Score Coefficient Matrix**

	Component				
	1	2	3	4	5
Number of words	.334	.375	-.056	-.210	.003
Emotional tendencies	.032	.239	.497	.254	.785
Number of pictures	.020	-.074	.439	-.588	.044
Point of praise	.415	-.318	.049	.135	.023
Differential evaluation	.390	-.362	.088	.201	-.043
Number of stars	-.055	.231	.314	.588	-.513
Descriptive vocabulary number	.246	.435	.134	-.195	-.438
A new degree of comment	.156	.303	-.491	.128	.381

Extraction Method: Principal Component Analysis.

Component Scores.

Judging from the highest score coefficient of every row, which index has the highest correlation with which component, for example, the correlation coefficient between the first line word number and the second component is the highest, reaching 0.375, which has the highest correlation. Similarly, can be summed up the following information: the first component and comments like the number of points, there was a correlation between a high rating number, number second and Book attributes; composition of words, words, attribute Book comment number of pictures and comments the new has a high degree of correlation; and third component words, and comments, new comments there is a strong correlation between the number of pictures; and the fourth component and commentary star, book attributes word number, there is a strong correlation between the number of pictures; and the fifth components with a review of the emotional tendency and number of pictures, there is a strong correlation between the degree of new comments.

### **3.3.5 Assignment**

The comment points like the number, number, number of books on attribute F1 represents the element as other users to evaluate the influence of other users, but the evaluation is one-sided, so with 15% information interpretation; Review words, words, pictures, books attribute number comment on the new degree as the composition of F2 represents the comment integrity influence, commented on the value of great help, so with 30% information interpretation; The number of reviews, the number of new reviews, and the number of pictures as components, F3, represents the attraction influence of the reviews, and has a role in user's attention to content, so it has 20% content interpretation. The number of comments, book words, star attribute picture as the composition of F4 represents the comment star influence, but the user rating has great randomness, so it has 15% information interpretation; The emotional tendencies, the number of pictures and the novelty of reviews are the elements of F5. It represents the emotional impact of the reviews, and it can show readers' emotional inclination to books, so it has 20% information interpretation.

### **3.3.6 Calculating comprehensive score**

All the components above can be obtained, that is, the comprehensive score of each dimension in the index system:

$$F=0.15*F1+0.3*F2+0.2*F3+0.15*F4+0.2*F5$$

### **3.3.6 Sorting and comparison**

According to the results of the sorting, the order order of the original website of the three books is quite different from the sort order of the new analysis, especially the former reviews. But for rankings section, the Dan comments ranked overall difference, such as the original 1-100 comments than the original 101-200 comments, the new ranking overall than before.

According to statistics, compared with the first book, the great comments F1 data difference, F4, F5 data is relatively small difference; the second book, the great comments F4 data difference, F2 data is relatively small difference; the third book, the great review F1 and F4 data difference, F5 data difference is small. F2 data has a great impact on the new ranking, while F1 and F4 data have less impact on the new ranking.

Between the data, the difference between the original ranking data is larger, so the new ranking has a larger span.

## **4. Conclusions**

The results of data analysis show that the ranking results of comprehensive scores based on factors are very different from those of the original network. By reading the new comments in the new ranking, we can find that they are rich in content and often have pictures of books. The description of books is more precise, allowing people to quickly understand various attributes of a book, and then determine whether the book is suitable for children to read. It can be seen that this sort of sorting method is more objective for the evaluation of the value of book reviews.

## References

- [1] Liu Y, Huang X, An A, et al. Modeling and predicting the helpfulness of online reviews[A]. Data Mining, 2008. ICDM'08. Eighth IEEE International Conference on[C]. Pisa: IEEE, 2008: 443-452.
- [2] Mudambi S M, Schuff D. What makes a helpful online review? A study of customer reviews on Amazon.com [J]. MIS Quarterly, 2010, 34( 1) : 185—200.
- [3] Yin Guopeng, Liu Wenwen, Zhu Shan. Online community commentary online usefulness impact model research -- Based on information adoption and social network perspective [J]. Library and information work, 2012, 56 (16): 140 - 147.
- [4] Guo Guoqing, Chen Kai. Research on the factors influencing the credibility of consumer online reviews [J]. Contemporary economic management, 2010, 10 (32): 123.