

<http://ansinet.com/itj>

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSI*net*

Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

SPSS Impact on the Development of Urban Leisure Sports Principal Component Analysis

Guanqiang Yang

Institute of Physical Education, Shihezi University, Shihezi, 832000, China

Abstract: In the social life of principal component analysis is a statistical method frequently used; it can reduce the computational advantage of dimensionality reduction method. Through statistical software SPSS, the data Alaer, Beitun, Tumushuke, Wujiaqu, Shihezi and other places obtained by Principal Component Analysis (PCA). The results showed that the main factors affecting the recreational sports people with learning time is tight, no space and no partner, followed by the constraints and issues of interest in economic conditions, the last factor is busy. In response to these limitations, the construction of sports facilities should be strengthened to enhance the promotion of sports culture and other specific strategies aimed at promoting physical fitness in order to improve citizens.

Key words: Principal component analysis, city leisure sports, influencing factors

INTRODUCTION

Recreational sports are an integral part of the social sport. Refers to people in their leisure time to improve health, enrich and create interest in life, self-improvement for the purpose of physical exercise activities. To promote health, physical fitness, disease prevention and rehabilitation, improving literacy and spiritual civilization, rich life content and strengthen relationships and promote both the importance and role of human socialization and personality formation.

Beginning of the century is a crucial period of China's overall well-off society, to socialist modernization. Social transition to a modern industrial, technological development, large-scale socialized production scale growing period which will profoundly change the way of social production, social structure, labor structure, class structure and values from a traditional agricultural society to people with to enrich the material and spiritual wealth but also sparked a new understanding of the value of life, lifestyle new thinking-the pursuit of high-quality, healthy and civilized lifestyle. You can reach \$ per capita. Herman Kahn, then China will enter the advanced industrial societies that mass consumer society. At present, China holidays achieve days. Visible, material aspects of life, leisure, institutional level, the cultural level are in place, indicating that China is going into leisure time, the value of pursuing a high quality of life. Leisure Sports is a healthy, scientific and civilized way of life, integrated into people's daily life, the evolution of human leisure center has become a necessity (Li *et al.*, 2010).

Principal component analysis, also known as principal component analysis, aims to use the idea of

dimensionality reduction, the multi-index into a few comprehensive index. In the study of practical problems, in order to comprehensively and systematically analyze the problem, we must consider a number of factors. These are generally referred to as indicators of factors involved in the multivariate statistical analysis, also known as variable. Since each variable in some extent, reflect the different information Research and between certain correlation indicators each other, thus resulting statistics information reflected to some extent overlap. In examining the issue with multivariate statistical methods, variables, too much will increase the amount of computation and complexity analysis of the problem, it is hoped the process of conducting a quantitative analysis of the variables involved in fewer, larger amount of information obtained (Wu, 2008). Over the years, the principal component analysis has gone deep into many aspects of national life and production, such as economics, finance, statistics, including the rapid development of sports and cultural industries (Xia, 2011).

Leisure Sports Research Xianyang City residents aged used questionnaires, literature, logical analysis, mathematical statistics and the status of the sport into the ages of Shenyang City residents' leisure to investigate (Shi *et al.*, 2007). The results show that: The ages of the residents involved in a relatively clear sense of Shenyang City, leisure sports but some people do not have the leisure and fitness awareness; motivation to participate in recreational sports has been a change with respect to the traditional concept of individual motivation is relatively simple; participate Views and inversely proportional to the duration of the basic; leisure and fitness facilities free places charge more places, leisure and fitness programs

has obvious geographical characteristics and gender differences, in order to enhance the participation of family, friendship is the purpose of the majority; leisure Sports Consumption medium-low. There are certain differences. Selection of principal component analysis, a comprehensive analysis of the factors affecting the participation of residents, according to the principle of extracting feature values extracted factors with eigenvalues greater than get the first six factors, the cumulative contribution rate, according to the statistics of Sociology requirements analysis, has been able to reflect the overall amount of information.

Leisure sports development impact of Hainan principal component analysis, selected indicators affect leisure sports development in Hainan, the use of statistical analysis of the main components of a comprehensive analysis of indicators will be conducted after the original variable values calculated for each factor score standardization and generation type and calculate the composite score and ranking tables. And use this method to various cities and counties in Hainan Leisure Sports Development comprehensive ranking to identify some areas of leisure sports development is lagging behind and puts forward corresponding countermeasures (Xu *et al.*, 2008).

Several questions about the sports industry statistics of the national economy is the foundation of the national economy statistical classification, as well as the sports industry statistical foundation. On the basis of industrial classification based on theories and methods defined on the meaning of the sports industry, sports industry statistics discussed a range of objects and calculation methods and some information about the academic point of view and are discussed and analyzed.

City Economic Zone in the economic strength of the domain dynamic evaluation-principal component analysis model based on global Economic Zone in the cognitive level of economic development around the city changes the dynamic process of constructing the index system consisting of 18 indicators, combined with 2007-specific data between 2010 global selection of principal component analysis model for 20 cities in the comprehensive economic strength of dynamic evaluation and analysis of their clustering which selected 20 west side of the city is playing a leading role in the economic development of the city, or cooperation is the key cities. The results show that the trend of economic development around the city is good but uneven regional development more obvious characteristics, spatial differences are more significant.

Application of principal component analysis listed in the company's financial analysis states: Principal component analysis is a multivariate statistical method Multiple Attribute Decision Making effective and comprehensive evaluation. Use and comprehensive insights into the company's financial position is very simple and effective. First, a brief introduction to the principles and calculation steps of principal component analysis; then select the appropriate study sample and financial indicators, principal components analysis to build a model using statistical software SPSS calculation, the listed company 's consolidated financial assessment score and then make comprehensive evaluation. And use the listed company's consolidated financial assessment scores, for example were vertical and horizontal comparison compare two applications.

In this study, principal component analysis, statistical software for this study, using data Alaer, Beitun, Tumushuke, Wujiaqu, Shihezi and other places to get. Found that the main factors that affect people's recreational sports learning time is tight, no space and no partners. According to these constraints, indicating that we should strengthen the construction of sports facilities, sports and cultural promotion, strengthening and thus promoting physical fitness in order to improve citizens.

PRINCIPLE AND CALCULATION PROCESS OF PRINCIPAL COMPONENT ANALYSIS

Basic principles of principal component analysis: Principal component analysis is a statistical analysis method that designates the original more variables as a few comprehensive indicators. From a mathematical point of view, this is a dimension reduction technique. A study object is often multi-factor complex system. Too many variables will undoubtedly increase the difficulty and complexity of analyzing problems, use the relationship between the original variables, use fewer new variables to replace the original more variable and make these few variables to retain the reaction information of original more variables as much as possible, so the problem is simplified. Assume there are n samples, each sample has p variables and form a data matrix of order n×p:

$$X = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1p} \\ X_{21} & X_{22} & \dots & X_{2p} \\ \cdot & \cdot & & \cdot \\ X_{n1} & X_{n2} & \dots & X_{no} \end{bmatrix}$$

Take the original variable indicators as x_1, x_2, \dots, x_p , suppose the composite indicators after the dimension reduction process i.e., the new variables are z_1, z_2, \dots, z_m ($m \leq p$), then:

$$\begin{cases} z_1 = l_{11}x_1 + l_{12}x_2 + \dots + l_{1p}x_p \\ z_2 = l_{21}x_1 + l_{22}x_2 + \dots + l_{2p}x_p \\ \dots \\ z_m = l_{m1}x_1 + l_{m2}x_2 + \dots + l_{mp}x_p \end{cases}$$

Determination principles of coefficient l_{ij} : z_1 and z_j ($i \neq j, j = 1, 2, \dots$) are independent of each other; z_1 is the greatest variance of all linear combinations x_1, x_2, \dots, x_p , z_2 is the greatest variance of all linear combinations x_1, x_2, \dots, x_p not relevant with z_1 ; z_m is the greatest variance of all linear combinations x_1, x_2, \dots, x_p not relevant with z_1, z_2, \dots, z_{m-1} . New variable indicators z_1, z_2, \dots, z_{m-1} respectively are the 1, 2, ..., m principle factors of the original variable index x_1, x_2, \dots, x_p .

As can be seen from the above analysis, principal component analysis is to determine the loads $l_{ij} = 1, 2, \dots, m; j = 1, 2, \dots, p$ of the original variables x_j ($j = 1, 2, \dots, p$) on the various principal components z_j ($j = 1, 2, \dots, m$). Can be proved mathematically, they respectively are the corresponding eigenvector to the m large eigenvalue of the correlation matrix.

Calculation steps of principal component analysis:

- Calculating the correlation coefficient matrix:

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1p} \\ r_{21} & r_{22} & \dots & r_{2p} \\ \vdots & \vdots & & \vdots \\ r_{p1} & r_{p2} & \dots & r_{pp} \end{bmatrix} \quad r_{ij} (i, j = 1(2) \dots (p))$$

is the correlation coefficient of the original variables x_i and x_j , $r_{ij} = r_{ji}$, the calculation equation is:

$$r_{ij} = \frac{\sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)}{\sqrt{\sum_{k=1}^n (x_{ki} - \bar{x}_i)^2 \sum_{k=1}^n (x_{kj} - \bar{x}_j)^2}}$$

- Calculating the eigenvalues and eigenvectors:** Solve the characteristic equation $|\lambda I - R| = 0$, commonly use Jacobi (Jacobi) method to calculate the eigenvalues and rank in the order of size $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$; Obtain the eigenvector e_i ($i = 1, 2, \dots, p$) corresponding to the eigenvalue λ_i , require $\|e_i\| = 1$, i.e.:

$$\sum_{j=1}^p e_{ij}^2 = 1$$

where, e_{ij} represents the j -th component of the vector e_i

- Calculate the principal component contribution rate and accumulative contribution rate:** Contribution rate:

$$\frac{\lambda_i}{\sum_{k=1}^p \lambda_k} \quad (i = 1, 2, \dots, p)$$

Cumulative contribution rate:

$$\frac{\sum_{k=1}^i \lambda_k}{\sum_{k=1}^p \lambda_k} \quad (i = 1, 2, \dots, p)$$

Generally we take the eigenvalue when the cumulative contribution rate reaches 85-95%, the 1, 2, ..., m ($m \leq p$) principle factors corresponding to $\lambda_1, \lambda_2, \dots, \lambda_m$.

- Calculate the principal component load:**

$$l_{ij} = p(z_i, x_j) = \sqrt{\lambda_i} e_{ij} (i, j = 1, 2, \dots, p)$$

- The various principal component scores:**

$$Z = \begin{bmatrix} z_{11} & z_{12} & \dots & z_{1m} \\ z_{21} & z_{22} & \dots & z_{2m} \\ \vdots & \vdots & & \vdots \\ z_{n1} & z_{n2} & \dots & z_{nm} \end{bmatrix}$$

IMPACT FACTOR ANALYSIS OF URBAN LEISURE SPORTS DEVELOPMENT BASED ON PRINCIPAL COMPONENT ANALYSIS

Many number of indicators increase complexity of the analysis. The increase of observe index is supposed to make the research process tend to be complete but conversely, to make research results clarity and clear increasing observed indicators blindly can make it unclear. Since in practical work, indexes often have a certain relevance, people want to use fewer indicators to replace the original more indexes but still able to reflect all of the original information and thus generate principal component analysis and factor analysis and other methods (Zhang and Yue, 2013).

Conduct a questionnaire survey on the residents of five city Shihezi, Wujiaqu, Alaer, Beitun and Tumushuke, in which 2317 copies are valid, then select job, economic conditions, learning time is tight, no interest and other eight factors. Use SPSS statistical software to conduct principal component analysis, as shown in Table 1.

The system first output the mean (mean) and standard deviation (Std Dev) of each variable, show a total of five cases of observation units and conduct the analysis, as shown in Table 2.

Then output the correlation coefficient matrix (Correlation Matrix), i.e., the correlation matrix is not a unit matrix, so we consider the factor analysis.

Import Table 1 into SPSS and conduct factor analysis. The basic purpose of factor analysis is to use some factors to describe the link between many indicators or factors, namely group several more closely related variables in the same class, each class variable becomes a factor, use several less factors to reflect most information of the original data and obtain the correlation matrix as in Table 3.

From Table 3 we can see that the correlation coefficient of total number of sportsman and economic conditions, learning time is tight and no venues is above 0.8, indicating that the four are closely linked and the correlation is very strong.

Extraction principle of PCA is the first x principal components that the corresponding eigenvalues are greater than 1. Eigenvalues to some extent can be seen as the index to present the size of component influence degree; If the eigenvalue is less than 1, indicating that the explanatory degree of the principle component is not greater than the mean explanatory degree of introducing an original variable. Thus we often take the eigenvalues greater than 1 as the inclusion criteria. By Table 4, extract three principle components, namely $x = 3$.

As can be seen from Table 5, the total number of sportsman, learning time is tight, no venues and no partner have a higher load in the first principal component. The first principal component basically reflects the information of these indicators. Economic conditions and no interest have a higher load in the second principal component, indicating that the second principal component basically reflect the information of the two indicators the economic conditions and no interest. The third principle component can basically reflect busy work this basic information.

Can be drawn from Table 4, the eigenvalues of the correlation matrix are $\lambda_1 = 0.478$, $\lambda_2 = 0.558$, $\lambda_3 = 1.362$, $\lambda_4 = 0.3$.

Table 1: Five impact factors table of urban leisure sports

City	N	a1	a2	a3	a4	a5	a6	a7	a8
Shihezi	467	49	49	44	47	125	57	42	29
Wujiaqu	592	50	37	105	51	152	94	24	57
Alaer	474	66	42	57	57	99	60	30	30
Beitun	446	28	60	68	84	84	36	28	30
Tumushuke	484	18	24	42	54	134	72	48	54

Sign convention: N: The total number of sportsman; a1: Busy with work; a2: Family burden; a3: Economic conditions; a4: Lack of skills; a5: Learning time is tight; a6: No venues; a7: No interest; a8: No partner

Table 2: Description statistical table

	Mean value	Standard deviation	Analysis N
N	492.60	57.287	5
a1	42.20	19.110	5
a2	42.40	13.428	5
a3	63.20	25.626	5
a4	58.60	14.673	5
a5	118.80	27.271	5
a6	63.80	21.288	5
a7	34.40	10.139	5
a8	40.00	14.195	5

Table 3: Factor analysis correlation matrix

Index	N	a1	a2	a3	a4	a5	a6
N	1.000	0.235	-0.442	0.804	-0.468	0.800	0.917
a1	0.235	1.000	0.184	0.252	-0.379	-0.021	0.181
a2	-0.442	0.184	1.000	0.091	0.630	-0.698	-0.740
a3	0.804	0.252	0.091	1.000	0.089	0.323	0.508
a4	-0.468	-0.379	0.630	0.089	1.000	-0.779	-0.711
a5	0.800	-0.021	-0.698	0.323	-0.779	1.000	0.916
a6	0.917	0.181	-0.740	0.508	-0.711	0.916	1.000
a7	-0.415	-0.480	-0.468	-0.839	-0.382	0.197	-0.069
a8	0.769	-0.347	-0.779	0.439	-0.365	0.810	0.843

Table 4: Explanatory total variance

Components	Initial eigenvalues			Extract square and load		
	Total	Variance (%)	Accumulation (%)	Total	Variance (%)	Accumulation (%)
1	4.780	53.110	53.110	4.780	53.110	53.110
2	2.558	28.425	81.535	2.558	28.425	81.535
3	1.362	15.128	96.663	1.362	15.128	96.663
4	0.300	3.337	100.000			
5	3.145E-16	3.494E-15	100.000			
6	6.961E-17	7.735E-16	100.000			
7	-1.416E-17	-1.574E-16	100.000			
8	-2.265E-16	-2.517E-15	100.000			
9	-3.453E-16	-3.837E-15	100.000			

Extraction Method: Principal Component Analysis

Table 5: Components matrix

Index	Components		
	1	2	3
N	0.904	0.418	0.057
a1	0.084	0.518	-0.836
a2	-0.757	0.526	-0.042
a3	0.481	0.837	0.254
a4	-0.702	0.334	0.615
a5	0.946	-0.169	-0.036
a6	0.994	0.061	-0.068
a7	-0.006	-0.991	-0.063
a8	0.880	-0.111	0.452

Table 6: Comprehensive principal component value

Influencing factor	F1	F2	F3
N	15605.16	-5881.95	6335.88
a1	19964.53	-5197.36	4506.00
a2	10015.51	3467.83	5132.78
a3	14049.7	1160.94	9863.21
a4	21778.77	-5587.41	7724.23
a5	22960.75	3202.46	6240.97
a6	14746.51	-2511.37	4438.96
a7	11271.67	-213.01	4892.59
a8	9819.83	-3990.12	3979.48

Each column value in the “matrix component” table is divided by the square root of eigenvalues and then obtains the eigenvector corresponding to each eigenvalue. Thus we can obtain the expression of the first, second and third principal component (suppose each factor is X_1, X_2, \dots, X_3).

$$F1 = 0.7527 \times X1 + 0.5095 \times X2 + 0.7274 \times X3 + 0.635 \times X4 + 1.3053 \times X5 - 1.37 \times X6 - 2.4318 \times X7 + 6.72 \times X8$$

$$F2 = 0.2982 \times X2 - 0.5534 \times X2 + 0.6501 \times X3 - 1.1387 \times X4 + 2.0468 \times X5 + 0.6784 \times X6 + 2.33 \times X7 + 3.4864 \times X8$$

$$F3 = 0.1324 \times X1 + 0.5828 \times X2 + 0.8923 \times X3 + 0.3804 \times X4 - 0.3959 \times X5 + 2.7855 \times X6 + 1.718 \times X7 + 1.8228 \times X8$$

Based on principal component comprehensive model the comprehensive principal component values can be calculated and rank in the order of comprehensive principal component values, you can conduct a comprehensive evaluation comparison of various factors, the results are shown in Table 6. For the results of comprehensive principle component (evaluation) values, we can use the actual results experience and the original data to do clustering analysis and can precede test, for the disputed results, we can use the raw data to discriminate the disputes for discriminate analysis to resolve the dispute.

CONCLUSION

This study, by statistical software SPSS, used the data Which obtained from Alaer, Beitun, Tumushuke and other places, conducted Principal component analysis, found that the main factors that affect people proceed leisure sports Include that learning time is tight (a5), no venue (a6) and no partner (a8), followed by the constraints of economic conditions (a3) and interest issues (a7), the last impact factor is The busying work (a1). In response to these limitations, we should strengthen the construction of sports facilities; strengthen sports and cultural promotion, in order to promote the atmosphere and interest of national leisure fitness sports.

ACKNOWLEDGMENTS

Project name: This 2012 Shihezi University youth engagement Fund projects “under the background of building a well-off society study on urban development of leisure sports in production and construction Corps of Xinjiang” research Project number (RWSK12-Y18).

REFERENCES

- Li, S., J. Wang and X. Lei, 2010. Leisure sports research of 24-38 year-old resident in Xianyang City. Sports Cult. Guide, 7: 22-25.
- Shi, Z., B. Sun, Y. Tian, C. Sun and Y. Jia, 2007. Investigation and analysis of current situation of residents' leisure sport in five cities in China. J. Wuhan Inst. Phys. Educ., 4: 84-90.
- Wu, J., 2008. Research on the leisure sport consumption actuality of the town resident in region of Zhujiang delta. J. Alaer Sport Univ., 2: 70-73.
- Xia, R., 2011. Effects of Hainan Leisure Sports development principal component analysis. Beijing Agric., 33: 15-19.
- Xu, Z.X., X.R. Du and J. Xu, 2008. Study of gender difference of the leisure sports in China. J. Guangzhou Sport Univ., 1: 1-5.
- Zhang, B. and H. Yue, 2013. Bio-mechanical mathematical model analysis for race walking technique. Int. J. Applied Math. Stat., 44: 469-476.