

# Data analysis of Higher Education System Based on TOPSIS and Time Series Model

Guangliang Yu<sup>1\*</sup>, Zhuoyang Fu<sup>2</sup>, Yujie Huang<sup>3</sup> and Chun-Te Lee<sup>4</sup>

<sup>1</sup> College of Business and Public Management, Wenzhou-Kean University, Wenzhou, China

<sup>2</sup> College of Science and Technology, Wenzhou-Kean University, Wenzhou, China

<sup>3</sup> College of Business and Public Management, Wenzhou-Kean University, Wenzhou, China

<sup>4</sup> Department of Mathematics, Wenzhou-Kean University, Wenzhou, China

\*Corresponding author. Email: 1098545@wku.edu.cn

## ABSTRACT

Building sustainable and healthy higher education has become a serious problem facing the world. In this article, we have selected 9 representative indicators as our factors in evaluating the higher education system. We developed an evaluation system based on the "Sequence of Preference Techniques Similar to the Ideal Solution" (TOPSIS). We use principal component analysis (PCA) to find the relationship between different feature variables. We apply our evaluation system to the higher education system in the United States and use PCA to analyze it. Then, we made some suggestions on the weaknesses of the PCA results. Finally, we use the seasonal autoregressive integrated moving average (SARIMA) model to simulate the changes after the policy is released. In conclusion, we have built an effective higher education evaluation system, and the results of the evaluation system are positive for the development of the country. In addition, we found that if some relevant policies are introduced to improve the shortcomings of the higher education system, the ranking of the US higher education system will be greatly improved to the 12th place among 50 countries.

**Keywords:** Higher Education, TOPSIS, PCA, Time Series

## 1. INTRODUCTION

### 1.1. Background

On the basis of elementary and secondary education, higher education provides citizens with opportunities to acquire more complex but useful knowledge. It not only has the value of creating new industries, but also has the value of cultivating talents for the further development of the country. In addition, the prosperous higher education attracts a large number of international students, which indirectly enhances the international influence of countries. Looking at the world, developed countries and most developing countries attach great importance to building a healthy and sustainable higher education system. Developing countries are trying to catch up with developed countries in the field of higher education, while developed countries are trying to maintain their advantages in this field. Although different countries have different policies for the doubling of higher education, each country has some similarities. All countries need to adjust, maintain their advantages and make up for their shortcomings. However, it is difficult

to make appropriate adjustments. Appropriate evaluation can help the country find weaknesses in the national higher education system, and then introduce more relevant policies.

### 1.2. Previous Research

The establishment of a sustainable and healthy higher education system has become a common concern. There are many studies related to this topic. According to these studies, some factors are highly related to its health and sustainability. For example, an article [1] surveyed participants in the Portuguese higher education system. The results show that although they understand the concept of sustainable development, different stakeholders are not familiar with the concept of sustainable higher education. The lack of government financial support for higher education and the decline in the number of students in Portuguese universities have created obstacles to the sustainability of higher education. In addition, another article [2] pointed out the importance of institutional service quality and the quantity of educational resources, which are generally accepted by

different stakeholders. In addition, an article [3] analyzes the importance of international students in the higher education system. The number of international students represents the university’s contribution to the “personality” and financial health of the institution. These aspects have profoundly affected the development of the higher education system. We select appropriate factors based on these aspects for further research.

**1.3. Our Task**

Our purpose is to analyze the higher education systems of 50 countries and build an effective and sustainable evaluation system. In addition, this research also tries to find the relationship between different higher education characteristic variables. Then we applied our model to the United States and made some suggestions for the American higher education system. Then we tried to find out whether the weaknesses of the higher education system would change based on the PCA results.

**2. GENERAL ASSUMPTIONS AND SYMBOL DESCRIPTION**

**2.1. General assumption**

1. Selected countries and factors represent the whole situation of world higher education
2. All collected data are valid and precise
3. Ignore private higher education which are not counted in statistical data
4. Data of multiple countries are measured with same standard

**2.2. Symbol description**

**Table 1.** Symbol description

Symbol	Description
$Z_1$	Enrolment in tertiary education per 100,000 inhabitants, both sexes
$Z_2$	Enrolment in tertiary education, all programs, both sexes (number) / Population of the official age
$Z_3$	Expenditure on tertiary as % of government expenditure on education (%)
$Z_4$	Gross enrolment ratio, tertiary, gender parity index (GPI)
$Z_5$	Pupil-teacher ratio in tertiary education (headcount basis)
$Z_6$	Total inbound internationally mobile students, both sexes (number) / Enrolment in tertiary education, all programs, both sexes (number)
$Z_7$	TOP 200(QS200) University / 1000,000
$Z_8$	All staff compensation as % of total expenditure in tertiary public institutions (%)

$Z_9$	Government expenditure per tertiary student as % of GDP per capita (%)
p	Trend autoregression order
d	Trend difference order
q	Trend moving average order
P	Seasonal autoregressive order
D	Seasonal difference order
Q	Seasonal moving average order
s	The number of time steps for a single seasonal period

**3. METHODS, MATERIALS, AND APPLICATIONS**

**3.1. The Technique for Order of Preference by Similarity to Ideal solution (TOPSIS)**

**3.1.1 Basic Theory of TOPSIS Model**

TOPSIS is a commonly used comprehensive evaluation method, which can make full use of the information of the original data, and its results can accurately reflect the gap between the evaluation schemes. Through a certain calculation, the comprehensive distance between the ideal optimal solution and the worst solution of any scheme in the scheme system is evaluated. If a scheme is closer to the ideal optimal solution and farther away from the worst solution, then this kind of scheme is optimal.

**3.1.2. The analysis of the model**

**Table 2.** Category meaning

Category	Meaning
$x_{ij}$	$i^{th}$ factor’s value for $j^{th}$ country
$w_i$	The weight $i^{th}$ factor
$Z_{ij}$	Elements in $i^{th}$ line and $j^{th}$ column of normalized weighted matrix

Additional assumptions:

1. Assumed that the effect of the influence factors other than the selected ones is negligible.
2. Assume that each factor is not influenced by other selected factors.

**3.1.3 Data preprocessing**

The data in this paper include 50 major countries around the world and 9 factors in 2014 that influence the health and sustainability of higher education systems. These data are extracted and filtered from URL: [Our World in Data](#). The selected countries are evenly

distributed in six continents, and the corresponding factors cover politics, economic and cultural aspects. After selecting corresponding data, we deal with all factors with normalization (1) method

$$x^*_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j} \quad (1 \leq i \leq 9, 1 \leq j \leq 50) \tag{1}$$

$$\bar{x}_j = \frac{1}{9} \sum_{i=1}^9 x_{ij}, s_j = \sqrt{\frac{1}{n} \sum_{i=1}^9 (x_{ij} - \bar{x}_j)^2}$$

Then use medium index maximization method (2) to deal with  $Z_4$ :

$$x'_j = \begin{cases} \frac{2(x_j - m_j)}{M_j - m_j}, m_j \leq x_j \leq \frac{M_j + m_j}{2} \\ \frac{2(M_j - x_j)}{M_j - m_j}, \frac{M_j + m_j}{2} \leq x_j \leq M_j \end{cases} \quad (1 \leq j \leq m, M_j = \max_{1 \leq x \leq 9} \{x_{ij}\}, m_j = \min_{1 \leq x \leq 9} \{x_{ij}\}) \tag{2}$$

For  $Z_5$ , it has a distinctive measurement standard. The lower the decimal is, the worse it behaves. Therefore, we use weighting method to cope with this factor.

**Table 3.** Worldwide Data for TOPSIS method

Parameter	Switzerland	Hong Kong	... (50 countries)
$Z_1$	0.5197	0.6217	...
$Z_2$	0.6337	0.7616	...
$Z_3$	0.6173	0.6841	...
$Z_4$	0.6472	0.7937	...
$Z_5$	0.6514	0.3638	...
$Z_6$	0.4645	0.2672	...
$Z_7$	1	0.8076	...
$Z_8$	0.8093	0.5	...
$Z_9$	0.5115	0.3312	...

3.1.4. Solution of TOPSIS model:

1) Set 9 factors' value for 50 countries as:

$$x_{ij} \quad (i = 1, 2, \dots, 9; j = 1, 2, \dots, 50) \tag{3}$$

2) Find the index weight vector:

To proceed the computation steps, we calculate the weight coefficient by Entropy Method as follows:

2-1) Calculate the characteristic proportion of index:

$$p_{ij} = \frac{x_{ij}}{\sum_{i=1}^{50} x_{ij}} \quad (i = 1, 2, \dots, 50; j = 1, 2, \dots, 9) \tag{4}$$

2-2) Calculate the entropy of index:

$$e_j = -k \sum_{i=1}^{50} p_{ij} \ln(p_{ij}) \quad (j = 1, 2, \dots, 9) \tag{5}$$

2-3) Calculate the difference coefficient of index:

$$g_j = 1 - e_j \quad (j = 1, 2, \dots, 9) \tag{6}$$

If the index difference coefficient is larger, it shows that the index number plays an important role in the model.

2-4) Calculate the weight coefficient of the index:

$$w_j = \frac{g_j}{\sum_{k=1}^m g_k} \quad (j = 1, 2, \dots, 9) \tag{7}$$

$$w = (0.04666, 0.07189, 0.02691, 0.04378, 0.03272, 0.1560, 0.5338, 0.04945, 0.03878)$$

3) Pretreat the decision matrix  $A = X_{ij} 9 \times 50$  to construct normalized decision matrix  $B = (b_{ij})_{9 \times 50}$ :

$$b_{ij} = \frac{a_{ij}}{\sqrt{\sum_{i=1}^9 a_{ij}^2}} \tag{8}$$

4) Construct a normalized weighted matrix:

$$Z = (z_{ij})_{9 \times 50} = w_i b_{ij} \quad (i = 1, 2, \dots, 9, j = 1, 2, \dots, 50) \tag{9}$$

5) Determine the positive ideal point  $S^*(z_1^*, z_2^*, \dots, z_9^*)$  and the negative ideal point  $S^-(z_1^-, z_2^-, \dots, z_9^-)$ :

When the jth indicator is extremely large:

$$\begin{cases} z_j^* = \max_{1 \leq i \leq n} \{z_{ij}\} \\ z_j^- = \min_{1 \leq i \leq n} \{z_{ij}\} \end{cases} \quad (j = 1, 2, \dots, 50) \tag{10}$$

When the jth indicator is extremely small:



Figure 1: The ranking of health and sustainability level:

$$\begin{cases} z_j^* = \min_{1 \leq i \leq n} \{z_{ij}\} \\ z_j^- = \max_{1 \leq i \leq n} \{z_{ij}\} \end{cases} \quad (j = 1, 2, \dots, 50) \quad (11)$$

6). Calculate the distance of each object to both positive ideal and negative ideal point:

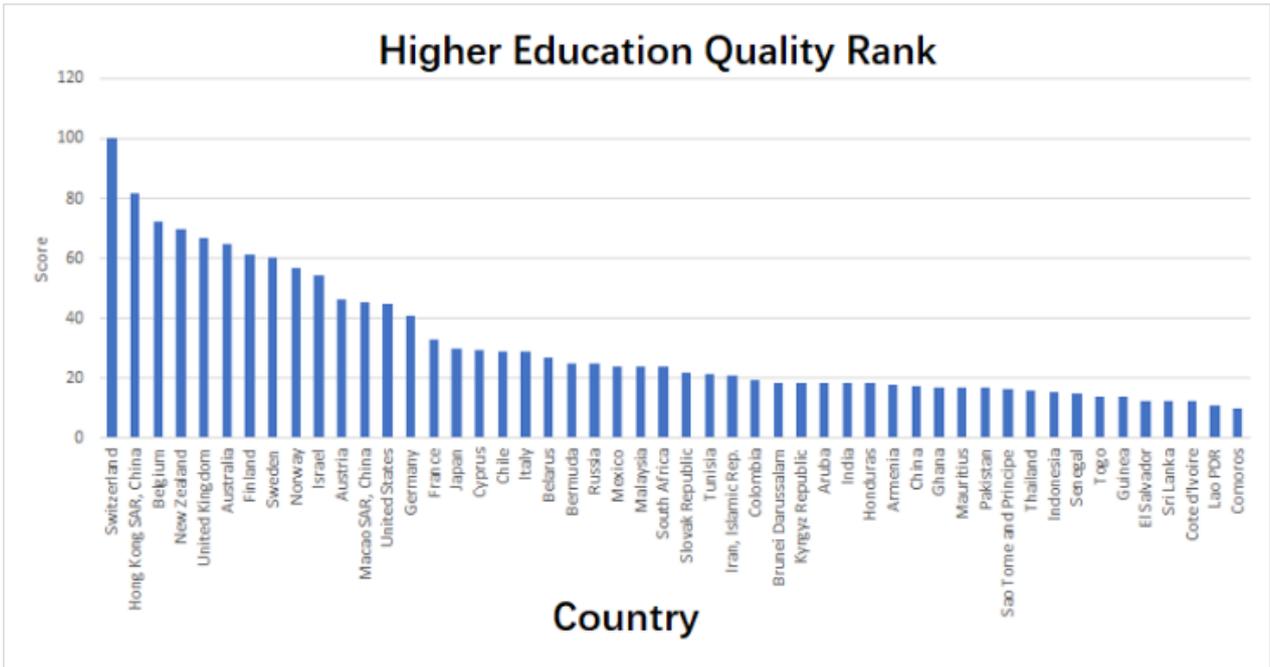


Figure 2: Heat Map of countries' higher education's health and sustainability level

$$\begin{cases} d_i^* = \sqrt{\sum_{j=1}^m (z_{ij} - z_j^*)^2} \\ d_i^- = \sqrt{\sum_{j=1}^m (z_{ij} - z_j^-)^2} \end{cases} \quad (i = 1, 2, \dots, 9) \quad (12)$$

7). Calculate the relative closeness of each country to the ideal point:

$$C_i^+ = \frac{d_i^-}{d_i^* + d_i^-} \quad (i = 1, 2, \dots, 9) \quad (13)$$

8). Rank the objects from large relative closeness to small relative closeness

### 3.1.5 Result of TOPSIS

After the calculation steps of 4.5, we figure out the TOPSIS score of 50 countries. Figure 1 represents the ranking of health and sustainability level based on TOPSIS scores. As a result, we decide to choose an imperfect country- the United States, to conduct our next research.

Heat Map (Figure 2) provides people a visual impression. Excluding pure white countries, the darker the country's color on this map, the better scores it gets.

## 3.2 Principle Component Analysis

### 3.2.1 Basic Theory of PCA model:

The Principal component analysis (PCA) is one of the most important dimensionality reduction methods. Its basic principle is to analyze the interrelationship among the original variables, and resort to the idea of "dimensionality reduction" to transform multiple indicators into a few composite indicators.

### 3.2.2 The analysis of the model

1. Assume that factors are related to each other
2. Assume that components explained more than 85% data characteristics are able to measure the data

Use the same data using in TOPSIS method and do some preprocessing. For  $Z_4$ , we apply the medium index maximization method (2). For  $Z_7$ , we apply the normalization method (1) for the data. For  $Z_5$ , we use the weighting method and create a distinctive measurement standard. Therefore, we use its derivative for PCA model.

**Table 4:** Worldwide Data for PCA model

Parameter	Switzerland	Hong Kong	(50 areas)
$Z_1$	0.035264	0.04218	...
$Z_2$	0.5722	0.6877	...
$Z_3$	0.2637	0.2923	...
$Z_4$	0.6472	0.7937	...
$Z_5$	0.1128	0.06301	...
$Z_6$	0.1710	0.09844	...
$Z_7$	1	0.8076	...
$Z_8$	0.5115	0.3312	...
$Z_9$	0.6734	0.5696	...

### 3.2.3 Process of PCA:

1) In our raw data, there are 450 pieces of numbers that are relevant to health and sustainability of the countries. Therefore, the formulas for the mean and variance of each indicator  $Z_i$  are as the following:

$$\bar{x}_i = \frac{1}{50} \sum_{j=1}^{50} x_{ij}, i = 1,2,3,4,5,6,7,8,9 \tag{14}$$

$$s^2_i = \frac{1}{50-1} \sum_{j=1}^{50} (x_{ij} - \bar{x}_i)^2 \tag{15}$$

$i = 1,2,3,4,5,6,7,8,9$

2) To simplify the model, this paper transfers the data into matrix:

$$X = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{19} \\ x_{21} & x_{22} & \dots & x_{29} \\ x_{31} & x_{32} & \ddots & x_{39} \\ \vdots & \vdots & \ddots & \vdots \\ x_{50\ 1} & x_{50\ 2} & \dots & x_{50\ 9} \end{pmatrix}_{50 \times 6} \tag{16}$$

Transfer matrix X into normalized matrix Z:

$$Z_{ij} = \frac{x_{ij} - \bar{x}_{ij}}{s_i}, \quad i = 1,2, \dots, 50; \tag{17}$$

3) Correlation Matrix:

$$R = (r_{ik})_{9 \times 9} \tag{18}$$

$$r_{ik} = \frac{\frac{1}{50-1} \sum_{i=1}^{50} (x_{ij} - \bar{x}_i)(x_{ki} - \bar{x}_k)}{s_i s_k}$$

$$(i = 1,2, \dots, 9; k = 1,2, \dots, 50) \tag{19}$$

R=

$$\begin{pmatrix} 1 & 0.9092 & 0.2881 & 0.6391 & 0.0013 & 0.0834 & 0.3828 & -0.3800 & -0.0367 \\ 0.9092 & 1 & 0.3232 & 0.6306 & 0.1023 & 0.1708 & 0.5002 & -0.3114 & -0.0019 \\ 0.2881 & 0.3232 & 1 & -0.0547 & -0.0233 & 0.3541 & 0.2493 & 0.1992 & -0.2613 \\ 0.6391 & 0.6306 & -0.0547 & 1 & -0.0298 & -0.0272 & 0.2706 & -0.2317 & -0.0969 \\ 0.0013 & 0.1023 & -0.0233 & -0.0298 & 1 & 0.1025 & 0.0779 & -0.0457 & 0.2082 \\ 0.0834 & 0.1708 & 0.3541 & -0.0272 & 0.1025 & 1 & 0.3014 & -0.0073 & 0.2013 \\ 0.3828 & 0.5002 & 0.2493 & 0.2706 & 0.0779 & 0.3014 & 1 & 0.0703 & 0.1005 \\ -0.3800 & -0.3114 & 0.1992 & -0.2317 & -0.0457 & -0.0073 & 0.0703 & 1 & 0.0788 \\ -0.0367 & -0.0019 & -0.2613 & -0.0969 & 0.2082 & 0.2013 & 0.1005 & 0.0788 & 1 \end{pmatrix}_{9 \times 9}$$

4) Eigenvalues and eigenvectors of the R:

$$(1) \sigma_i = \sum_1^9 r_{ij} \quad i = 1,2, \dots, 9$$

$$(2) \sigma = \max\{\sigma_i\}$$

$$U' = \frac{1}{\sigma} (\sigma_1, \sigma_2, \dots, \sigma_9) \tag{20}$$

(3) Calculate  $RU'$  and then the components after transposing can be recorded as  $(b_1, b_2, \dots, b_9)$

$$b = \max\{(b_1, b_2, \dots, b_9)\} \quad (21)$$

$$U'_1 = \frac{1}{b} (b_1, \dots, b_9) \quad (22)$$

(4) Repeat step (3) and iterate until  $U_n \rightarrow U$ . The maximum value of each element in  $R*U$  tends to the first eigenvalue, so the result represents that  $U$  is the first eigenvector

(5) Find the principal components. The corresponding unit vector of  $\lambda_1$  is  $V_1$  ( $V_1 = \frac{U}{|U|}$ ). Each component of  $V_1$  represents the first principal component of  $Z_1, Z_2, \dots, Z_9$ .  $\lambda_1$  represents the contribution of the first principal component  $\tau_1$

$$\tau_1 = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \dots + \lambda_9} = \frac{\lambda_1}{9} \quad (23)$$

Here,  $\lambda_1 + \lambda_2 + \dots + \lambda_9 = \text{tr}R = 9$ .  $\lambda_1$  is the variance of the first principal component.

(6) Calculate the second principal component:

$$A = R - \lambda_1 V_1 V_1' \quad (24)$$

Then  $A$  is still a sixth-order symmetric matrix. Repeat from step (1) to step (5) and obtain  $\lambda_2, V_2$ , and  $\tau_2$ :

$$F_1 = 0.0364Z_1 + 0.6497Z_2 + 0.0383Z_3 + 0.6289Z_4 + 0.0047Z_5 + 0.0292Z_6 + 0.3974Z_7 - 0.1432Z_8 - 0.0133Z_9 \quad (25)$$

$$F_2 = -0.0065Z_1 + 0.0445Z_2 + 0.1058Z_3 - 0.4625Z_4 + 0.0140Z_5 + 0.1053Z_6 + 0.7753 + 0.3532Z_8 + 0.1888Z_9 \quad (26)$$

$$F_3 = -0.0229Z_1 - 0.3962Z_2 - 0.0368Z_3 + 0.5460Z_4 - 0.0242Z_5 - 0.0643Z_6 + 0.0489Z_7 + 0.7297Z_8 - 0.0785Z_9 \quad (27)$$

$$F_4 = -0.0032Z_1 - 0.0535Z_2 - 0.0368Z_3 + 0.1451Z_4 + 0.0371Z_5 + 0.0490Z_6 - 0.1046Z_7 - 0.0333Z_8 + 0.9542Z_9 \quad (28)$$

(7)  $F_1 = (Z_1, Z_2, \dots, Z_9)V_1$  is called the first principal component.

$F_2 = (Z_1, Z_2, \dots, Z_9)V_2$  is called the second principal component.

$F_3 = (Z_1, Z_2, \dots, Z_9)V_3$  is called the third principal component.

$F_4 = (Z_1, Z_2, \dots, Z_9)V_4$  is called the fourth principal component.

$$\frac{\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4}{9} = \tau_1 + \tau_2 + \tau_3 + \tau_4 \geq 85\%$$

As a result, the cumulative variance contribution of the first three principal components can exceed 85% in the vast majority of cases. In practice, 85% is taken as the lower limit of the cumulative contribution of variance.

$$\frac{\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4}{\sum \lambda_i} \geq 85\%$$

The result also indicates that the reliability of the information is more than 85% when the first three principal components are used to illustrate the problem. The first four eigenvalues and eigenvectors in this paper combine information from the original nine indicators, rather than simply drop information from certain indicators.

### 3.2.4 The result of PCA

#### 3.2.4.1. Significance of the First Principal Component $F_1$

According to the previous results, the first principal component value of higher education quality is:

$$F_1 = 0.0364Z_1 + 0.6497Z_2 + 0.0383Z_3 + 0.6289Z_4 + 0.0047Z_5 + 0.0292Z_6 + 0.3974Z_7 - 0.1432Z_8 - 0.0133Z_9$$

In the expression of  $F_1$ , the result can reflect that the coefficients of the second indicator  $Z_2$  and the fourth indicator  $Z_4$  are all positively correlated to  $F_1$ . It can be seen from the above calculation results that  $F_1$  emphasizes the importance of enrolment in tertiary education and gender parity which can influence the health and sustainability of the whole country.

In the expression of  $F_1$ , the result can reflect that the coefficients of the second indicator  $Z_2$  and the fourth indicator  $Z_4$  are all positively correlated to  $F_1$ . It can be seen from the above calculation results that  $F_1$  emphasizes the importance of enrolment in tertiary education and gender parity which can influence the health and sustainability of the whole country.

3.2.4.2. Significance of the First Principal Component

**F<sub>2</sub>**

$$F_2 = -0.0065Z_1 + 0.0445Z_2 + 0.1058Z_3 - 0.4625Z_4 + 0.0140Z_5 + 0.1053Z_6 + 0.7753Z_7 + 0.3532Z_8 + 0.1888Z_9$$

It is seen that F2 has a positive correlation with the seventh indicator Z7. It indicates that the quality of higher education is significantly related to the quality and number of higher education schools in the region. Universities with outstanding academic ability can promote the development and progress of the higher education system in the whole country.

3.2.4.3. Significance of the First Principal Component

**F<sub>3</sub>**

$$F_3 = -0.0229Z_1 - 0.3962Z_2 - 0.0368Z_3 + 0.5460Z_4 - 0.0242Z_5 - 0.0643Z_6 + 0.0489Z_7 + 0.7297Z_8 - 0.0785Z_9$$

According to the above equation, F3 indicates a distinct and positive correlation with the eighth indicator Z8. This correlation illustrates that government financial expenditures targeted at education and the educational investments students received can contribute to the health and sustainability of the higher education system.

3.2.4.4. Significance of the First Principal Component

**F<sub>4</sub>**

$$F_4 = -0.0032Z_1 - 0.0535Z_2 - 0.0368Z_3 + 0.1451Z_4 + 0.0371Z_5 + 0.0490Z_6 - 0.1046Z_7 - 0.0333Z_8 + 0.9542Z_9$$

F4 displays a significant positive correlation with the last factor. This relationship reflects the fact that the proportion of teachers' salaries to education expenditures in higher education in the corresponding countries also affects the health and sustainability of the education system.

**Table 5.** Main components

Component	Explained	Main Relevant factors
F <sub>1</sub>	50.95%	Z <sub>2</sub> , Z <sub>4</sub>
F <sub>2</sub>	19.37%	Z <sub>3</sub> , Z <sub>7</sub> , Z <sub>8</sub>
F <sub>3</sub>	11.19%	Z <sub>4</sub> , Z <sub>8</sub>
F <sub>4</sub>	8.63%	Z <sub>4</sub> , Z <sub>6</sub> , Z <sub>9</sub>

**3.3 Time-series analysis Model**

**3.3.1 introduction of Time Series**

Seasonal Autoregressive Integrated Moving Average (SARIMA) is a prediction model developed from ARIMA, it generalizes the data before a certain point in time and utilizes the generalizations to predict the data after this point. Our data are arranged in time series, which means it is suitable for this model.

**3.3.2 Data pre-processing**

We collect 4 corresponding United States data from "Our World in Data" and do some preprocesses. Firstly, we complete a small amount of missing data by averaging two adjacent numbers. Next, we use linear change to simulate the value of four factors over the next five years after policy implementation. Finally, we process time Series into Stationary Time Series by "First-Order Difference method" or "Second-Order Difference method"

**3.3.3 Process of Time-series analysis Model**

1) Construct SARIMA model

$$\begin{cases} \varphi(B)\phi(B^S)(1-B)^d(1-B^S)^D Y_t = c + \vartheta(B)\Theta(B^S)\varepsilon_t, \\ \phi(B^S) = 1 - \beta_1 B^S - \beta_2 (B^S)^2 - \dots - \beta_p (B^S)^p, \\ \Theta(B^S) = 1 - \alpha_1 B^S - \alpha_2 (B^S)^2 - \dots - \alpha_q (B^S)^q, \end{cases} \tag{29}$$

$\phi(B^S)$  Regression coefficient polynomial

$\Theta(B^S)$  Polynomials for moving average coefficients

2) Calculate the optimal SARIMA parameter based on AIC method (Akaike information criterion).

$$AIC = 2k - 2\ln(L) \tag{4} \tag{30}$$

(k is the number of model parameters, L = data likelihood with the given model and k is the number of parameters in the model)

3) The optimal parameters of each time series of the four factors are obtained by AIC method. The following are their augmented Dickey–Fuller test (ADF test) and

their (autocorrelation function) ACF and Partial autocorrelation function (PACF) plot.

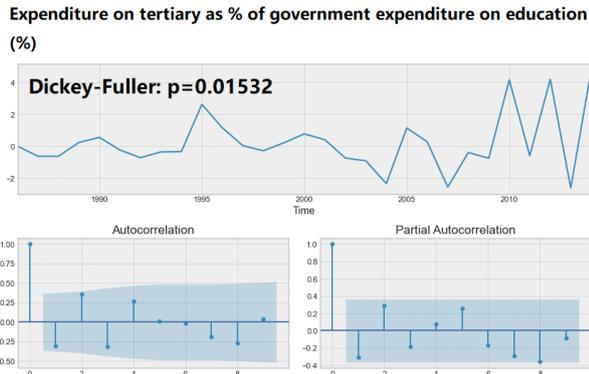


Figure 3: Expenditure on tertiary as % of government expenditure on education.

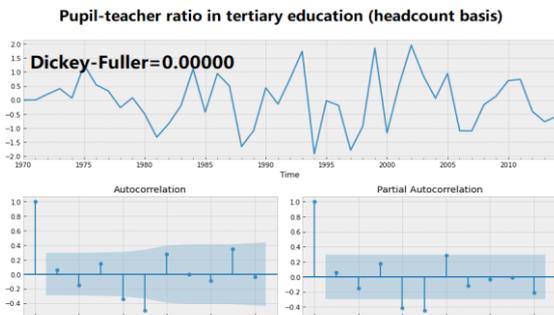


Figure 4: Pupil-teacher ratio in tertiary education

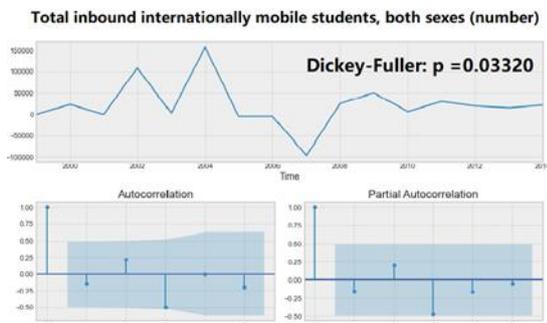


Figure 5: Total inbound internationally mobile students, both sexes

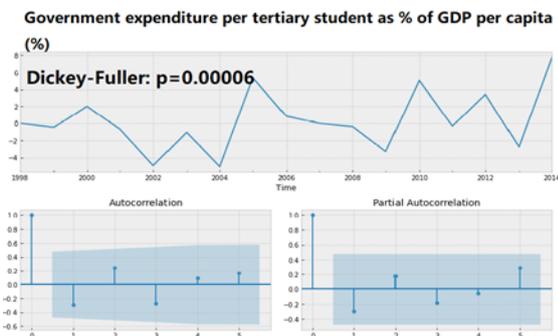


Figure 6: Government expenditure tertiary student as % of GDP per capita

The P value of the ADF test of the residual is almost 0, far less than the significance level 0.05, so the fitting residual is stable and has no autocorrelation.

4) Apply the SARIMA to predict the value of four factors in 2030, and then calculate mean absolute percentage error

$$M = \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \quad (31)$$

$M$  = mean absolute percentage error  
 $n$  = number of times the summation iteration happens  
 $A_t$  = actual value

$F_t$  = forecast value

#### 4. ANALYSIS OF THE UNITED STATES

This article aims to use a modified model based on the PCA and TOPSIS methods to measure the health and sustainability of the United States. Judging from the results of the TOPSIS ranking, the American higher education system has many obvious problems in terms of health and sustainability, which is unexpected. According to the principal component scores, the second and third principal component scores of the United States are both negative. It shows that there is still a certain gap between the US government's expenditure on college students and the most educated countries in the world.

It is well known that world-class universities have a profound influence on regional and national higher education development policies. On the one hand, building more elite universities can effectively promote the sustainable development of higher education. On the other hand, the state's expenditure on higher education is the fundamental factor that promotes the development of the education system and also ensures the fair and reasonable allocation of educational resources.

Taking into account the United States' scores in other factors in the TOPSIS model, we also consider its performance by comparing it with the best countries in the world whose education level is among the highest in the world. Therefore, there are still some problems with the student-teacher ratio in the United States and the acceptance percentage of international students.

The number of faculty and staff in the United States has increased significantly after the expansion of higher education, but the contradiction between the rapidly growing number of college students and the relatively stable number of faculty and staff has become increasingly prominent. As the higher education system develops, it affects the health and sustainability of higher education

Notice that sending and receiving international mobile students has also become an important and

complex issue in higher education. The number of international students in a university affects the operation and international status of the university. Therefore, it has received more and more attention from the government and related institutions. In addition, while adapting to the various requirements of international students, the school has improved and enriched teaching methods and research skills to improve the quality and sustainability of higher education.

**5. POLICY RECOMMENDATION**

According to TOPSIS analysis, United States' high education is far away from top countries in government expenditure on tertiary education, Pupil-teacher ratio, and total inbound international students. According to PCA method, government expenditure on tertiary students is another point which is crucial for its higher education performs.

**5.1. Proposes on government expenditure on tertiary education**

For most of American history, the number of the most prestigious elite universities in the United States was small, and their private charters and religious beliefs made them largely independent of the government [5]. On the one hand, the United States intends to cooperate with these famous schools to fund their daily life and campus construction, help them expand the scale of running schools, and improve the quality of higher education. On the other hand, in order to improve the overall level of higher education in the country, the government is more inclined to support public universities with lower levels of education. So the country can cultivate more elites

In our model, we reflect this policy proposes by increasing government expenditure on tertiary education 0.5 percent per year, and this process lasts for 5 years.

**5.2. Proposes on government expenditure on tertiary student**

Financial condition limits American family's higher education. According to American's situation, the middle-class family are required to borrow a large amount of loans to afford their children's higher education intuitions, while lower-class families' students have no chance to enter excellent colleges. Therefore, extra policies are demanded to solve this problem. There are two possible ways to improve this embarrassing situation: one is subsidizing college students; another is offering them loans with low interest rate.

In our model, we reflect this policy proposes by increasing government expenditure on tertiary student 0.4 percent per year, and this process lasts for 5 years

**5.3. Proposes on pupil-teacher ratio**

In recent years, more and more people from different races and countries have received higher education from the United States. The increase in the number of students has increased the teacher-student ratio. Every professor needs to teach more students, thus affecting the quality of higher education. The enrollment of professors is mainly affected by two factors. [6]. One is to formulate wage payment standards. A professor's salary should match his degree level, so it is easy for talents to engage in teaching work. When teachers just started their careers, it is recommended that they have at least US\$50,000. After several years of practice, they should have the opportunity to earn US\$100,000 per year. The low salary of teachers is not the only factor that affects their teaching enthusiasm. When graduates begin their first experience as teachers, they often feel confused and stressed. The training policy solves this problem. New teachers should be guided by old teachers, and gain a set of experience in the real classroom as a teaching assistant, so as to easily adapt to the classroom.

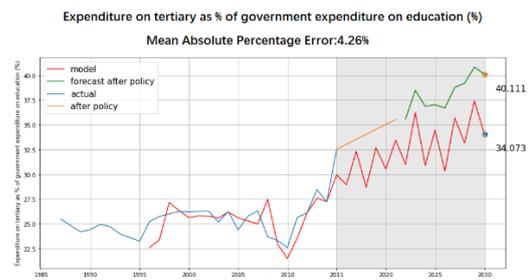
In our model, we reflect this policy proposes by decreasing pupil-teacher ratio 0.2 percent per year, and this process lasts for 6 years.

**5.4. Proposes on Inbound international students**

Although the United States has a large number of inbound international students, the United States still needs more policies to attract and retain them in order to maintain cultural diversity. For example, more and more universities in the United States are adopting STEM policies to encourage international students to stay and work in the United States after graduation. This policy has been helping the United States attract more international students and retain these higher education talents who have made great contributions to the United States. In the near future, this policy should be extended to more international students.

In our model, we reflect this policy proposes by increasing inbound international student number 40k per year, and this process lasts for 5 years.

**6. RESULT**



**Figure 7:** Expenditure on tertiary as % of government expenditure on education.

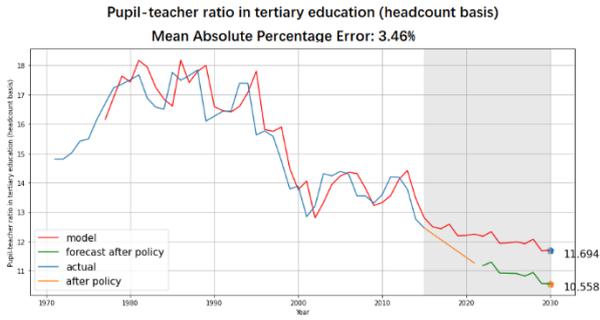


Figure 8: Pupil-teacher ratio in tertiary education

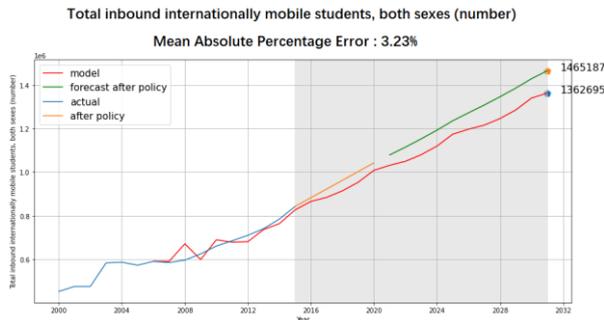


Figure 9: Total inbound internationally mobile students, both sexes

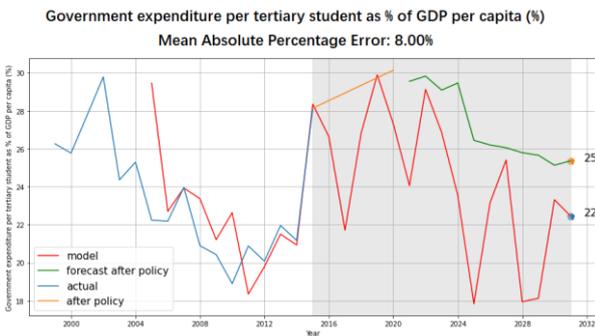


Figure 10: Government expenditure tertiary student as% of GDP per capita

After the implementation of the government's policy of increasing economic investment in higher education, the forecast results show that the growth rate of the indicator "higher education expenditure as a percentage of government education expenditure (%)" is significantly higher than before, and policy adjustments are more stable.

After the government's policy to increase the number of teachers in higher education is implemented, the forecast model shows that the policy can significantly reduce the pressure on each teacher, thereby improving their teaching skills.

When the number of international students going to the United States increased due to policies, the annual growth rate of the number of new international students accelerated compared with before the adjustment.

As the government invests in education, the relative amount of national financial aid received by each student in higher education will increase proportionally compared with countries without policy adjustments.

After the implementation of the policy, the TOPSIS score of the United States will be 47.518 in 2030, 2.8 points more than that in 2014. So, the ranking of United States will rise to 12th in 2030, and the gap between the United States and other countries which have health and sustainable higher education system will be further narrowed.

## 7. DISCUSS REAL-WORLD IMPACTS AND CHALLENGES

The implementation of targeted policies will have a real impact on the American higher education system. Increasing the government's financial expenditure on higher education can alleviate the financial pressure and teaching pressure of colleges and universities. More importantly, students can have more opportunities to receive education at top universities, because relaxing the family loan policy can fundamentally prevent students from being financially stressed and unable to afford tuition.

The teacher-student ratio directly affects the health and sustainability of higher education. The increase in the number of teachers in the United States often lags behind the increase in the number of students. Many professors are affected by the pressure of scientific research projects, reducing teaching time. The existing teaching incentive and restraint mechanisms in universities are also not conducive to the expansion of the number of teachers. The fundamental task of solving this problem is to increase the scale of teachers. At the same time, schools need to mobilize qualified teachers to participate in more teaching and research activities. The implementation of this policy will further improve the effectiveness of the allocation and use of teachers, ease the pressure on teachers, and enable the healthy and sustainable development of higher education in the United States.

As the country with the largest number of international students in the world, the United States has had a profound impact on the higher education system by adjusting the policies related to studying abroad. A more open policy for overseas students may reduce the educational resources of domestic students, but it will increase the diversity of the cultural atmosphere within higher education institutions. In addition, the policy will also provide a fairer academic field for international and local students, thereby enhancing the ability of higher education institutions to train high-level graduates.

Notice that the transformation of higher education is obviously difficult to analyze. Although many American universities have gained a good impression abroad in recent years, the academic ability of the United States is

also facing increasingly fierce challenges. Criticism within the American higher education system has also increased. First, the US government's emphasis on higher education has declined, and the government's fiscal deficit has hindered the sound development of schools. The government's large-scale cuts in education manifested itself in funding, research projects, and student loans. The second is that the emphasis on scientific research and postgraduate education has led to a decline in the quality of undergraduate education. In addition, college fees are getting higher and higher, making it impossible for people from poor families to receive higher education. All these factors affect the health and sustainability of the U.S. higher education system.

## 8. CONCLUSION

This article uses the TOPSIS method to develop a model to obtain the health status of any country's higher education system. The bar graph and heat map analysis obtained from the ranking show that some countries with small and compact education systems are the healthiest and most sustainable higher education systems in the world. European countries, especially the Nordic region, can be said to be the most advanced regions for higher education construction.

We select multi-dimensional, easily quantifiable indicators from the factors affecting higher education in a region as principal component analysis factors. The final main components include factors such as enrollment rate and student gender ratio, government education funds, and higher education teacher salaries. According to the analysis of PCA and TOPSIS, there are still many problems in the American higher education system. The United States needs to deal with issues such as government expenditure on higher education, teacher-student ratios, and overseas study policies. This article proposes some targeted policies and verifies their effectiveness through time series analysis. By comparing the actual situation with the predicted results, we also analyzed the actual impact of policy adjustments and the difficulty of policy improvement.

## REFERENCES

- [1] Aleixo, A. M., Leal, S., & Azeiteiro, U. M. (2018). Conceptualization of sustainable higher education institutions, roles, barriers, and challenges for sustainability: An exploratory study in Portugal. *Journal of cleaner production*, 172, 1664-1673.
- [2] Tsinidou, M., Gerogiannis, V., & Fitsilis, P. (2010). Evaluation of the factors that determine quality in higher education: an empirical study. *Quality assurance in Education*.
- [3] Hegarty, N. (2014). Where we are now—The presence and importance of international students to universities in the United States. *Journal of International Students*, 4, 223-235.
- [4] Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6), 716-723.
- [5] Educational-Policy-United-States. (2019). The Basics of Educational Policy, The Pressure for Reform in American Education, Defining Policy Read more: United States Educational Policy - The Basics of Educational Policy, The Pressure for Reform in American Education: <https://education.stateuniversity.com/pages/1937/Educational-Policy-United-States.html>
- [6] Ronald Thorpe, "Residency: Can it transform teaching the way it did medicine?", *Kappan* 96 (1) (2014): 36–40, available at [http://www.nbpts.org/wpcontent/uploads/ron\\_thorpe\\_pdk.pdf](http://www.nbpts.org/wpcontent/uploads/ron_thorpe_pdk.pdf).