http://dx.doi.org/10.18623/rvd.v23.n2.4198

ANALYSIS OF COLLEGE STUDENTS' PHYSICAL FITNESS TEST DATA AND RESEARCH ON PHYSICAL EDUCATION CURRICULUM REFORM

ANÁLISE DOS DADOS DOS TESTES DE APTIDÃO FÍSICA DE ESTUDANTES UNIVERSITÁRIOS E PESQUISA SOBRE A REFORMA DO CURRÍCULO DE EDUCAÇÃO FÍSICA

Article received on: 8/29/2025
Article accepted on: 11/28/2025

Zhou Runhua*
*Open Education Institute, Mongolian University of Science and Technology, Ulaanbaatar, Mongolia
Orcid: https://orcid.org/0000-0003-2849-4148
zhourunhua1317@163.com

Yanjmaa Battsetseg**
**Department Engineering, School of Technology Darkhan-Uul province, Mongolian University of Science and Technology, Ulaanbaatar, Mongolia
battsetseg.ya@stda.edu.mn

The authors declare that there is no conflict of interest

Abstract
This study collected standardized physical fitness test data from 28,290 college students and conducted an in-depth analysis using a complete data preprocessing pipeline and machine learning algorithms. First, missing values were imputed and outliers were processed, followed by standardization. Principal component analysis (PCA) was then applied for dimensionality reduction, and the K-means clustering algorithm was used, with the optimal number of clusters (K=2) determined by the silhouette coefficient method. The data analysis revealed significant group differences in students' physical fitness, particularly in strength and cardiorespiratory endurance. Based on these findings, a 16-week teaching reform experiment was conducted at Guangzhou Huashang College, incorporating weak test items into the curriculum. The results showed that the pass rate for male students' pull-ups increased by 23.34 percentage points (p<0.01), and the standing long jump improved by 25 percentage points (p<0.01). Female students also demonstrated significant improvements in related indicators. This study validates the effectiveness of data-driven physical education reform and provides actionable pathways for college physical education programs.

Keywords: Physical Fitness. Data Mining. K-Means Clustering. Teaching Reform. Exercise Prescription.

Resumo
Este estudo coletou dados padronizados de testes de aptidão física de 28.290 estudantes universitários e realizou uma análise aprofundada usando um pipeline completo de pré-processamento de dados e algoritmos de aprendizado de máquina. Primeiro, os valores ausentes foram imputados e os outliers foram processados, seguido pela padronização.
Em seguida, a análise de componentes principais (PCA) foi aplicada para redução de dimensionalidade, e o algoritmo de agrupamento K-means foi usado, com o número ideal de agrupamentos (K=2) determinado pelo método do coeficiente de silhueta. A análise dos dados revelou diferenças significativas entre os grupos em relação à aptidão física dos estudantes, particularmente em força e resistência cardiorrespiratória. Com base nessas descobertas, foi realizada uma experiência de reforma do ensino de 16 semanas na Faculdade Huashang de Guangzhou, incorporando itens de teste fracos ao currículo. Os resultados mostraram que a taxa de aprovação dos alunos do sexo masculino nas flexões aumentou 23,34 pontos percentuais (p<0,01) e o salto em distância sem impulso melhorou 25 pontos percentuais (p<0,01). As alunas também demonstraram melhorias significativas nos indicadores relacionados. Este estudo valida a eficácia da reforma da educação física baseada em dados e fornece caminhos viáveis para os programas de educação física universitária.

Palavras-chave: *Condicionamento Físico. Mineração de Dados. Agrupamento K-Means. Reforma do Ensino. Prescrição de Exercícios.*

1 INTRODUCTION

In recent years, with the comprehensive implementation of the "Healthy China 2030" initiative, increasing attention has been paid to college students' physical health. According to the 2022 National Student Physical Health Monitoring Report released by the Ministry of Education, while the overall declining trend in students' physical fitness has been curbed, the excellent/good rate for college students remains at only 23.25%, with a failure rate of 12.44%, indicating a still challenging situation [1].Furthermore, the Ministry of Education's 2014 "Basic Standards for Physical Education in Higher Education Institutions" explicitly requires that "no less than 30% of each physical education class must focus on cardiorespiratory fitness training" [2]. The State Council's 2019 "Opinions on Implementing the Healthy China Initiative" established clear objectives for student physical health, with local authorities subsequently developing corresponding targets and timelines. By 2030, the excellent/good rate in student physical health standards should reach 60% or above, shifting the assessment focus from passing rates to excellence rates, thereby raising the bar for physical education in schools.

Against this backdrop, developing precise physical education reform strategies based on scientific analysis of fitness test data has become a crucial task for higher education institutions.This study innovatively integrates two key dimensions: large-scale fitness data analysis and practical teaching validation. First, we systematically analyze physical test data from 28,290 college students using data mining techniques[3] and machine learning algorithms to identify patterns in physical fitness characteristics. Second, based on these analytical findings, we conduct empirical research at Guangzhou Huashang College to validate the effectiveness of targeted teaching reform measures.This "data analysis-pattern identification-practical validation" research approach provides a novel methodology for reforming physical education in higher education institutions.

2 RESEARCH METHODS

2.1 Data sources and characteristics

This study utilizes data from two independent samples:

1. **Large-scale physical fitness test data**: A total of 28,290 college students, including 9,738 males and 18,552 females. The dataset comprises eight core indicators: height, weight, vital capacity, 50-meter sprint, sit-and-reach test, endurance run (1,000m for males/800m for females), standing long jump, and strength tests (pull-ups for males/sit-ups for females). All data were collected using testing instruments and methods specified in the *National Student Physical Health Standard* (2014 edition).
2. **Teaching experiment data**: Derived from Guangzhou Huashang College, including baseline physical test data from 28,290 students and experimental intervention data from 176 participants (60 males and 116 females). The experiment lasted 16 weeks (March to June 2025).

3 DATA PROCESSING PIPELINE

3.1 Data preprocessing

To ensure data integrity and consistency, strict quality control mechanisms were implemented during data collection, followed by comprehensive cleaning and preprocessing procedures to enhance dataset quality. The following preprocessing steps were applied to college students' physical fitness characteristics [4]:

(1) Missing Value Treatment:

Descriptive statistics and bar charts were used to visualize missing values across variables. The overall missing rate was below 2.5%, satisfying the missing-at-random assumption.Mean imputation was applied for continuous variables, while mode imputation was used for categorical variables. For example, missing male vital capacity values were filled with the group mean of 3663.097ml.

(2) Outlier Detection and Treatment:

Box plots were first employed to identify outliers, defined as data points beyond Q1-1.5IQR or Q3+1.5IQR ranges.Detected outliers (e.g., extreme height values of 194.7cm) were replaced with median values. Post-treatment data distributions were re-examined to ensure processing effectiveness.

Table 1

Example Comparison Before and After Outlier Treatment (Male Height)

Processing Stage	Sample Size	Mean (cm)	Standard Deviation	Minimum	Maximum
Before Treatment	9,738	172.32	6.502	85.4	194.7
After Treatment	9,738	172.31	6.487	150.1	194.2

3.2 Data standardization

To eliminate scale effects, min-max normalization was applied to continuous variables:

$$X_{norm} = (X_i - X_{min}) / (X_{max} - X_{min}) \quad (1)$$

where:

x_i represents the original value,

and $x_{\min}$ and $x_{\max}$ denote the minimum and maximum values of the variable, respectively.

After standardization, all feature values are mapped to the [0,1] interval.

4 ANALYTICAL METHODS

4.1 Descriptive statistics

Calculate the central tendency measures (mean, median) and dispersion measures (standard deviation, range) for each variable, as well as the distribution shape measures (skewness, kurtosis). Taking the male physical fitness test data as an example: pull-ups

(7.35 ± 5.65 times) exhibit a right-skewed distribution (skewness 0.575), and the 1000-meter run (293.79 ± 55.21 seconds) shows a long-tail phenomenon (skewness 1.201).

Table 2

Descriptive Statistics of Male Physical Test Data (Partial)

Indicator	Mean±SD	Median	Range	Skewness	Kurtosis
Pull-ups (times)	7.35±5.65	7.0	36	0.575	-0.134
1000m Run (s)	293.79±55.21	283.0	358	1.201	4.865

4.2 Correlation analysis

Calculate the Pearson correlation coefficient matrix to analyze the linear relationships among various physical fitness indicators[5]. Visualize the correlation matrix using a heatmap, setting the significance level at α=0.01. The results show a moderate positive correlation between pull-ups and standing long jump ($r=0.413$, $p<0.001$), while weight exhibits a negative correlation with pull-ups ($r=-0.319$, $p<0.001$).

4.3 Principal component analysis (PCA)

(1) Data Suitability Test:The KMO test value was 0.723 (>0.7), and Bartlett's sphericity test yielded $p<0.001$, confirming the suitability for PCA.

(2)Criteria for determining the number of principal components: eigenvalue >1 or cumulative variance contribution rate ≥80%.

(3) Analysis Process:PCA dimensionality reduction was performed on the standardized 8 indicators.The first three principal components accounted for 82.3% of the cumulative variance, with PC1 (comprehensive fitness factor) contributing 45.2%, PC2 (strength-speed factor) 24.7%, and PC3 (flexibility factor) 12.4%.

4.4 K-means clustering analysis

(1) Optimal K Value Determination:The silhouette coefficient method was used to calculate coefficients for K=2-10.Results showed the highest silhouette coefficient (0.412) at K=2, confirming the optimal number of clusters as 2.

(2) Clustering Process:The first three principal components from PCA were used as input features.Parameters: maximum iterations=500, convergence threshold=1e-4, 10 repeated initializations for optimal results.Two cluster center vectors were ultimately obtained (as shown in Table 3).

Table 3

Male Cluster Center Vectors (Partial)

Principal Component	Cluster 0	Cluster 1
PC1	-0.160	0.218
PC2	-0.378	0.516
PC3	-0.058	0.080

5 TEACHING EXPERIMENT DESIGN

5.1 Current status of college physical education curriculum

At Guangzhou Huashang College, compulsory physical education courses are offered to first- and second-year undergraduates, totaling 4 credits across four semesters. The first semester of freshman year consists of mandatory courses in Tai Chi (a traditional Chinese sport) and physical fitness training. The remaining three semesters offer elective courses across 23 sports, including football, basketball, volleyball, aerobics, Tai Chi, outdoor training, badminton, table tennis, and tennis. The curriculum focuses on sports skills combined with sport-specific physical conditioning. Assessment primarily evaluates sports skills and regular performance, emphasizing outcome-based evaluation.

5.2 College physical education curriculum reform experiment

5.2.1 Subjects and methods

Four elective classes at Guangzhou Huashang College, comprising 176 students (60 males, 116 females), participated in a 16-week teaching reform experiment from March to June 2024. The study aimed to validate the effectiveness of data-driven physical education reforms in improving student fitness levels and explore practical implementation methods.Intervention Design Based on Analytical Results:

Experimental Grouping:Experimental group (n=176): Received the reformed physical education curriculum.Control group (n=180): Maintained the traditional teaching model.

Intervention Measures:

Course Content: Each 90-minute session included 60 minutes of specialized skill training followed by 30 minutes of targeted physical conditioning.

Training Program: Design differentiated training plans based on the clustering results. The training focus for Cluster 0 is to maintain strengths and improve weaknesses, with moderate intensity and three sessions per week. The training focus for Cluster 1 is to build foundational strength and aerobic capacity, with low-to-moderate progressive intensity and four sessions per week. (see Table 4).

Table 4

Differentiated Training Program Based on Clustering Results

Group Type	Training Focus	Intensity Control	Weekly Frequency
Cluster 0	Maintain strengths, improve weaknesses	Moderate intensity	3 times
Cluster 1	Foundational strength and aerobic capacity	Low-to-moderate progressive intensity	4 times

Evaluation Metrics: Primary outcome: Changes in pass rates of physicaltest.items.Secondary outcome: Improvement magnitude in individual test scores

5.3 Data-driven reform pathways for college physical education based on student fitness test analysis

5.3.1 incorporating fitness test items into curriculum content

Guided by the "health-first" principle outlined in the National College Physical Education Guidelines and Higher Education Physical Work Standards, we integrated the three lowest-performing test items into regular instruction. Each 90-minute weekly session comprised 60 minutes of sport-specific skill development followed by 30 minutes of targeted conditioning: pull-ups, standing long jump, and 1000m run for males; standing long jump and 800m run for females. Instructors demonstrated proper techniques, safety protocols, and training principles to facilitate independent practice.

5.3.2 Making fitness tests part of final assessments

Reform the evaluation system by integrating process assessment with outcome assessment. The limited scope of assessment content, singular evaluation formats, and lack of process-oriented evaluation have failed to effectively motivate students' engagement in physical conditioning exercises [6].The reformed evaluation system combines process (40%) and outcome (60%) assessments, with 30% dedicated to fitness test performance - meeting the 30% minimum requirement for cardiorespiratory assessments per national standards. An additional 10% process evaluation tracks conditioning progress, creating incentives for consistent participation.

5.3.3 Personalized instruction based on fitness data

Using stratified teaching methods, students were grouped into three tiers (excellent/passing/failing) according to test results. Customized programs adjusted exercise difficulty and objectives for each tier, with instructors modifying:Exercise selection (assisted vs. standard pull-ups),intensity progression (low-medium vs. medium-high),achievement benchmarks.

5.3.4 Integrating classroom and extracurricular training

Extracurricular physical exercise serves as an effective approach to enhancing students' health, playing an irreplaceable role in fostering adolescents' moral character, intellectual development, aesthetic literacy, and healthy lifestyles[7]. It is challenging to achieve the goals of improving students' physical health and strengthening their constitution solely through the two weekly hours of physical education classes. Therefore, extracurricular physical exercise is essential as an effective supplement.Physical education classes primarily teach sports skills and methods for physical fitness training, requiring students to complete "physical education assignments" during their extracurricular exercise time. The content of these assignments can either fully replicate the activities conducted in class or be adjusted in difficulty based on individual circumstances. Each "physical education assignment" session lasts 30 minutes, with a frequency of 2–3 times per week, serving a function similar to that of an exercise prescription.

The content of extracurricular physical exercise is not limited to the top three items with the highest failure rates in student fitness tests. After completing their "physical education assignments," students can access their test results through the school’s fitness testing online system and arrange appropriate practice content based on their performance. Through the process of completing "physical education assignments," students gradually transform the acquired sports knowledge, skills, and training methods into lifelong physical fitness capabilities.

5.3.5 Embedding ideological education in fitness training

Arranging classroom physical education instruction based on the ranking of student fitness test failure rates inherently poses a challenge for students who fail to meet the standards. Teachers should explain the intrinsic connection between physical fitness and sports skills, clarify the fundamental patterns of physical fitness development, and help students establish a correct understanding of physical education, thereby alleviating concerns that their physical fitness cannot be improved. This approach encourages students to adopt a positive attitude toward their current situation, to push their limits through physical fitness exercises, and to enhance their overall health.

During the process of physical fitness training, students will naturally encounter various difficulties and setbacks, including temporary instability or regression in performance, as well as physical discomfort such as soreness and pain caused by lactic acid buildup during exercise. Teachers should actively cultivate students' courage to overcome challenges and push their boundaries, foster a sense of teamwork characterized by mutual encouragement and support, and integrate ideological and political education into the teaching of physical fitness within the physical education curriculum.

6 RESULTS AND ANALYSIS

6.1 Data analysis findings

Multiple factors—both subjective and objective—may influence test outcomes [8]. While this study's conclusions are based on pre-/post-intervention comparisons, other variables (e.g., ambient temperature, lifestyle habits, physical condition) could have contributed to observed improvements. Through triangulating test results with classroom observations, we attribute the enhanced pass rates primarily to the pedagogical reforms.

6.2 Fitness characteristics distribution

The distribution characteristics of male physical fitness indicators are as follows: pull-ups exhibit a right-skewed distribution (skewness 0.575), with most students clustered in the lower score range (median 7 repetitions, 25th percentile only 3 repetitions). The 1000-meter run also shows a right-skewed distribution (skewness 1.201), with a long-tail phenomenon indicating that some students' performance lags significantly behind.

The distribution characteristics of female physical fitness indicators are as follows: sit-ups approximately follow a normal distribution (skewness -0.086). The 800-meter run exhibits a right-skewed distribution (skewness 1.269), with relatively high variability in performance (SD = 43.255 s).

6.3 Correlation analysis (Figure 2)

Pull-ups and standing long jump show a moderate positive correlation (r=0.413, p<0.001); weight and pull-ups exhibit a negative correlation (r=-0.319, p<0.001); the 1000-meter run has the most significant negative correlation with the total score (r=-0.585, p<0.001).

6.4 Cluster analysis

Cluster 0 (Higher-fitness group, 58.2%):Pull-ups: 10.45±4.86;1000m run: 277.24±44.89s;Vital capacity: 3621.22±683.24ml

Cluster 1 (Lower-fitness group, 41.8%):Pull-ups: 3.11±3.49;1000m run: 316.42±59.80s;Vital capacity: 3720.40±709.99ml

All inter-group differences were significant (p<0.001, independent t-tests).

6.5 Intervention outcomes

After 16 weeks, the experimental group showed:

Pull-ups: The experimental group improved from 5.2±3.1 times to 8.7±4.3 times (+3.5 times, p<0.001); the control group improved from 5.3±3.0 times to 6.1±3.5 times (+0.8 times, p=0.112).

Standing long jump: The experimental group improved from 198±21 cm to 215±19 cm (+17 cm, p<0.001); the control group improved from 197±20 cm to 203±18 cm (+6 cm, p=0.023).

Table 5

Evaluation of Teaching Experiment Outcomes

Indicator	Group	Pre-test	Post-test	Improvement	p-value
Pull-ups (count)	Experimental	5.2±3.1	8.7±4.3	+3.5	<0.001
	Control	5.3±3.0	6.1±3.5	+0.8	0.112
Standing long jump (cm)	Experimental	198±21	215±19	+17	<0.001
	Control	197±20	203±18	+6	0.023

Key intervention effects:

1.**Strength**: Male pull-ups pass rate increased by 23.34%, with large effect size (Cohen's d=0.82)

2.**Endurance**: 1000m run pass rate improved by 11.66%, with average time reduction of 18.5 seconds

3.**Lower-body power**: Standing long jump showed significant gains, with experimental group outperforming controls ($p<0.01$)

7 DISCUSSION

7.1 Value of data analysis methods

This study employed a comprehensive data analysis pipeline with the following advantages:

1.**Data Quality Control**: Rigorous handling of missing values and outliers ensured result reliability. For example, post-treatment height data showed reduced standard deviation (from 6.502 to 6.487), yielding more rational distributions.

2.**Effective Feature Engineering**: PCA retained 82.3% information while reducing features from 8 to 3, enhancing clustering efficiency and interpretability.

3.**Actionable Clustering Results**: The K=2 classification (determined via silhouette coefficient) facilitated practical implementation, with clear inter-group differences enabling targeted interventions.

7.2 Implications for teaching reform

The data-informed reforms demonstrated:

1. **Precision Intervention Necessity**: The lower-fitness group (Cluster 1) showed greater improvement (e.g., +4.2 vs. +2.1 pull-ups for Cluster 0, $p<0.05$), validating personalized training.
2. **Curricular-Extracurricular Integration**: "Exercise prescriptions" (2-3 sessions/week, 30 minutes/session) were pivotal for sustained gains.
3. **Assessment-Driven Engagement**: Incorporating fitness tests into evaluations (30% weight) significantly boosted participation, aligning with prior research [8].

7.3 Limitations and future directions

7.3.1 Limitations

1. Limited generalizability to sports-specialized institutions.
2. Short-term (16-week) intervention period; long-term effects require tracking.
3. Incomplete control of lifestyle confounders (diet, sleep).

7.3.2 Future work

1. Longitudinal databases to analyze fitness trends.
2. AI-powered recommendation systems for personalized training.
3. Multimodal data integration (e.g., wearables) for holistic assessment models.

7.3.3 Key innovations included

Cluster-specific programming: Differentiated protocols for distinct fitness profiles

Dual-focus pedagogy: Simultaneous skill development and metabolic conditioning

Progress-based motivation: Real-time feedback through standardized testing

These findings advance evidence-based PE reform, demonstrating how data analytics can bridge research and practice to optimize student outcomes.

8 CONCLUSIONS AND RECOMMENDATIONS

8.1 Key findings

1.Significant heterogeneity exists in college students' physical fitness. PCA and K-means clustering effectively identified distinct groups ($p<0.001$ for strength/endurance differences).

2.Data-driven interventions significantly improved fitness levels: 18.34% average pass-rate increase post-16-week intervention (effect sizes $d=0.61-0.82$).

3.The "data analysis → classification → targeted intervention" framework demonstrated scalability and practical value for PE reform.

8.2 Implementation strategies

1. **Smart Fitness Monitoring Systems:** Institutional fitness databases with regular updates/analytics;Automated visualization platforms generating student/class profiles.
2. **Curriculum Optimization**:Differentiated lesson plans tailored to fitness clusters;≥30% class time dedicated to weakness areas (e.g., strength training).
3. **Assessment Enhancements**:Incorporate progress metrics (≥40% process evaluation);"Most Improved" awards to motivate struggling students.
4. **Teacher Development**:Data literacy training for evidence-based instruction;Collaborative lesson planning to share best practices.

This three-tiered approach—technological infrastructure, pedagogical innovation, and professional development—provides a replicable model for data-informed physical education modernization.

REFERENCES

[1]. National Student Physical Health Standard (2014 Revision) [Z]. Ministry of Education: JTY[2014]No.5, 2017,7.

[2]. Basic Standards for Physical Education in Higher Education Institutions [Z]. Ministry of Education: JTY[2014]No.4, 2014,6.

[3] Zhang Xueqin, Jiang Fan, Xi Benyu. Design and Application Research of Student Physical Health Testing Platform Based on Data Mining [J]. Electronic Design Engineering, 2022,30(13):87-90+95. DOI:10.14022/j.issn1674-6236.2022.13.018.

[4] Huang Rong, Han Zhezhe. Prediction of College Students' Physical Test Scores Based on Ensemble Learning Model [J]. Fujian Sports Science and Technology, 2023,42(06):112-118.

[5] Fan Linli. Analysis and Prediction of College Physical Test Scores Based on Machine Learning [D]. Dalian Polytechnic University, 2020. DOI:10.26992/d.cnki.gdlqc.2020.000106.

[6]. Zhou Zuhui. Research on Student Physical Health from Environmental Factors Perspective [J]. Youth Sports, 2016,(8):98-99

[7]. Yang Wenxuan, Zhang Xiqian, Deng Xinghua. School Physical Education [M]. Beijing: Higher Education Press, 2016,238
[8]. Zhang Zongguo. Analysis of Subjective and Objective Factors Influencing National Student Physical Health Standard Test Results [J]. Sports Science, 2009,29(9):86-91

Authors' Contribution

All authors contributed equally to the development of this article.

Data availability

All datasets relevant to this study's findings are fully available within the article.

How to cite this article (APA)

Runhua, Z., & Battsetseg, Y. (2026). ANALYSIS OF COLLEGE STUDENTS' PHYSICAL FITNESS TEST DATA AND RESEARCH ON PHYSICAL EDUCATION CURRICULUM REFORM. *Veredas Do Direito*, 23(2), 234198. https://doi.org/10.18623/rvd.v23.n2.4198