International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 8 Issue 6, Nov-Dec 2024 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Guarding Invisible Wealth - Analysis of Campus Student Health Monitoring Based on Big Data Technology

Liu Yue, Guo Peiyu, Xu Sijia, Li Wenxue

College of Statistics and Data China, Beijing Wuzi University, Beijing, China

ABSTRACT

At present, the field of campus student health detection and analysis is constantly developing with the help of big data technology. On the basis of collecting multi-source data, this paper takes students on campus as the research object for data processing and analysis, and uses data mining and machine learning algorithms to carry out in-depth research. The results show that: (1) Big data technology has been paid more and more attention in campus student health detection and analysis, but it still faces challenges in the application process such as data privacy protection and data security; (2) Campus students' health data mainly come from health examination records, sports activity data and other aspects; (3) Through the construction of student health file system, students' health status can be revealed in multiple dimensions, and targeted management and intervention measures are proposed. The application of big data technology is conducive to the intelligent and personalized campus health management.

KEYWORDS: *physical and mental health; heart rate; bmi; cardiopulmonary function; blood oxygen content*

How to cite this paper: Liu Yue | Guo Peiyu | Xu Sijia | Li Wenxue "Guarding Invisible Wealth - Analysis of Campus Student Health Monitoring Based on Big Data Technology" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-8 | Issue-6, December 2024, pp.663-667, URL: www.ijtsrd.com/papers/ijtsrd72658.pdf

IJTSRD72658

1. INTRODUCTION

In modern times, the health problems of adolescents have received more and more attention. With the change of lifestyle and the influence of social environment, adolescents are faced with many physical and mental health challenges. According to relevant surveys, the adolescent obesity rate is rising year by year, myopia is becoming more common, and psychological pressure is also increasing, such as anxiety, depression and other emotional problems in the adolescent group is not rare. The state attaches great importance to the healthy growth of young people, and has introduced a series of policy measures, such as the Opinions on Strengthening Youth Sports and Enhancing Youth Physical Fitness, aimed at comprehensively improving the health quality of young people. In this context, the vigorous development of big data technology has brought new opportunities and possibilities for adolescent health detection and analysis. Big data technology has shown strong application capabilities in many fields, and it can help disease prediction and precision medicine in the field of medical health. In terms of adolescent health management, big data technology can integrate multi-source data, mine potential information, and achieve accurate assessment and personalized intervention. However, at present, the application of big data technology in campus student health detection and analysis is still in the exploratory stage, facing problems such as uneven data quality and difficult privacy protection. This study focuses on the field of campus student health detection and analysis, aiming to deeply explore how to effectively use big data technology to improve the current situation. Taking campus students as the research object, a comprehensive, scientific and efficient health detection system is constructed by collecting and analyzing various information such as health examination records, sports activity data, dining hall records and daily health monitoring data. On the one hand, big data technology is used to deeply analyze students' physical fitness, nutritional status, mental health, disease prevention and control and other dimensions. On the other hand, a student health file system is constructed to realize dynamic data recording and visual display, which provides convenience for school management, parents' concern and students' self-monitoring. At the same time, personalized health management and intervention measures are proposed in response to the problems found, and the challenges in the application of big data are deeply discussed, and corresponding policy suggestions are put forward to provide strong support for campus health management and help young people grow healthily.

2. Investigation framework

The investigation framework of this study mainly consists of students' basic information dimension and students' health-related dimension. The basic information dimensions of students include demographic characteristics (such as age and gender) and academic-related conditions (such as grade and major). The health-related dimensions of students include lifestyle factors (food intake types and preferences are concerned in terms of diet, work and rest rules are investigated in terms of sleep, exercise frequency and other information are collected in terms of exercise, activity and expression are concerned in terms of social interaction and emotional expression, duration and scene of use of technology are collected and analyzed in terms of health data (speed and heart rate data during exercise) , physical indicators data, cardiovascular health-related data) and health awareness and attitude surveys (perceptions of their own health status, attitudes towards a healthy lifestyle, perceptions and expectations of school health management).

3. Research direction

In this study, at the stage of exploratory data analysis, the collected data were cleaned and preprocessed first, and the missing values and outliers were dealt with. Then through a variety of data visualization methods, the indicators of life habits are analyzed. In the correlation analysis of exercise programs, for exercise speed and heart rate, it can be seen from the line chart that the average speed curve has a downward trend, and the speed is different in different stages, reflecting the difference in exercise intensity. Heart rate also showed a downward trend and corresponded with speed, which could reflect the degree of cardiovascular system load at different stages. By analyzing the relationship between speed and heart rate, we can understand the relationship between exercise intensity and cardiovascular health. By comparing the average speed and heart rate under different exercise programs, it is found that each program has different effects on speed and heart rate, and the influence and trend of the overall exercise effect can be evaluated based on this. By analyzing the proportion of different exercise regimens in the dataset, we can understand their relative frequency, popularity, and distribution balance. By observing the scatterplot of velocity and heart rate at different stage levels, we can preliminatively judge that there are linear, non-linear, irrelevant relationships and cluster effects between them, but further statistical analysis and hypothesis testing are needed. As for the visual analysis of BMI, it can be seen from the scatter plot that there is a certain correlation but not a linear relationship between BMI and maximum blood oxygen content, and the maximum blood oxygen content shows a decreasing trend with the increase of BMI, and there are distribution characteristics and possible outliers. To this end, linear relationship test, grouping analysis, outlier detection, correlation analysis, stratification analysis considering gender and age factors, trend analysis, mechanism exploration and health impact analysis were further carried out in order to more comprehensively and deeply understand the relationship between the two, and provide more in-depth insights and guidance for relevant research and clinical practice.

4. Survey questionnaire

4.1. It was found through investigation that the participants in the exercise program could better accept the speed and heart rate

in the exercise program, and none of the exercisers could completely accept the speed and heart rate in the program. In addition, only 5% of the exercisers were less able accept the speed and heart rate in the program. However, being able to accept the speed and heart rate in the exercise program and actually following the program are two concepts. It can be seen that among the five representative exercise programs listed, only 15% of the exercisers actually followed three or more of the programs

4.2. Analysis of the frequency of different exercise programs

Additionally, by analyzing the frequency the exercise programs adopted by the exercisers (not limited to the exercise programs in this study), it was found that the frequency of the exercise programs adopted by theisers was relatively high, with 45.23% of the exercisers adopting the exercise programs 3-6 times in the past six months. inning the investigation results of the mean speed and heart rate under different exercise programs, it was found that the exercise programs in this study still have advantages among many exercise.

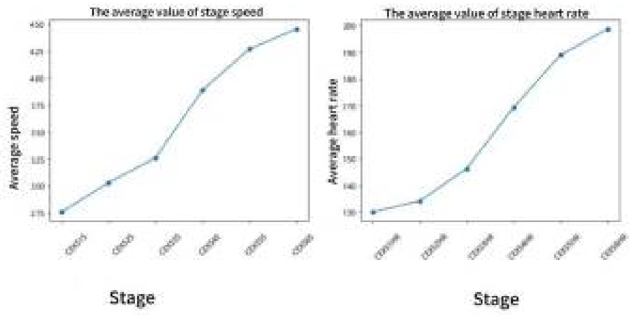

Figure 1. General trend of speed and heart rate in the exercise

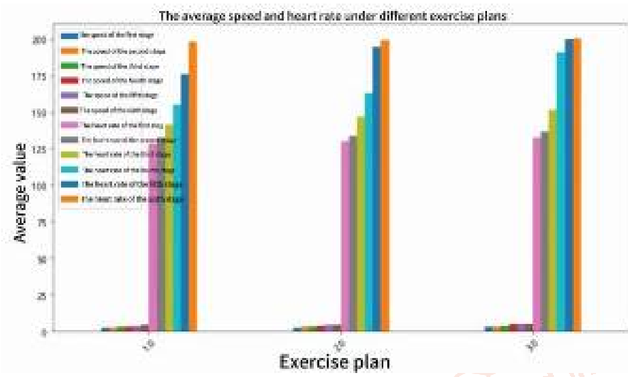

Figure 2. Mean values of speed and heart rate under different exercise protocols

4.3. Proportion of different exercise information sources in the dataset

Through this pie chart, the proportion of different exercise information sources in the dataset can be analyzed thus understanding the relative frequency of each exercise information source.

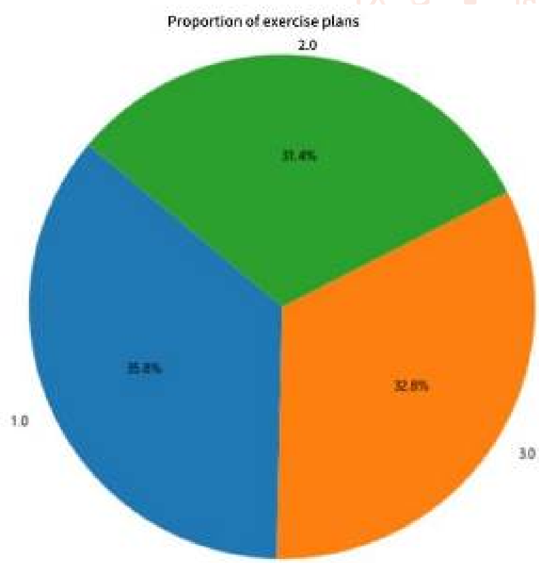

Figure 3 The proportion of different exercise plans in the dataset

The specific analysis is as follows:

1. Proportion of exercise plans:

The size of each sector represents the proportion of the exercise plan in the dataset. The part on the far left is blue, occupying one quarter of the entire pie chart. The middle part is green, occupying two fif of the pie chart, which is the largest proportion. The part on the far right is orange, also occupying two fifths of the pie chart. This allows to understand the relative importance of each exercise plan in the dataset.

2. Most popular exercise plan:

Green usually symbolizes vitality and health, representing exercise plans as aerobics and cardio, which are widely popular and have significant benefits to physical health.

3. Balance of exercise plans:

The sizes of the sectors not significantly different, indicating that the distribution of exercise information sources in the dataset is relatively balanced.

5. Updates on speed and heart rate at different stage

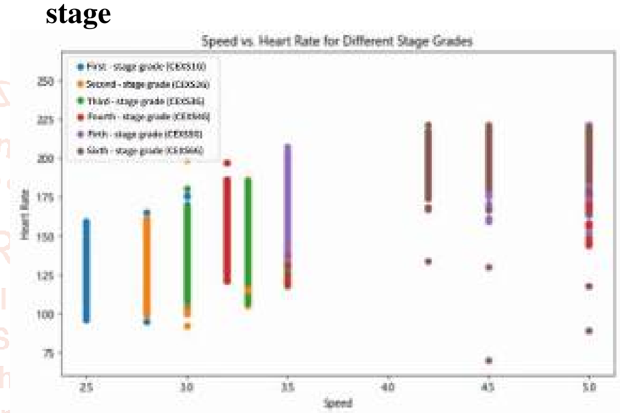

Figure 4. Speed and heart rate at different stage levels

This chart shows the relationship between speed and heart rate at stage levels.

1. The first stage level (CEXS1G, blue)

had a heart rate of about 150 at a speed of 25. As the speed increases to around 30, the heart rate slightly rises to about 160. When the speed further increases to 5, the heart rate reaches about 175. At a speed of 40, the heart rate reaches about 190. Finally, a speed of 45, the heart rate is around 200.

2. The second stage level (CEXS2G, orange)

a heart rate of about 125 at a speed of around 25. As the speed increases to around 30, the heart rate rises to 150. When the speed further increases to 35, the heart rate reaches about 175. At a speed of 40 the heart rate reaches about 190. Finally, at a speed of 45, the heart rate is around 200.

3. The third stage level (CEXS3G, green)

has a heart rate of about 100 at a speed of around 25. As speed increases to around 30, the heart rate rises to about 125. When the speed further increases to 35, the heart rate reaches 150. At a speed of 40, the heart rate reaches about 175. Finally, at a speed of 45 the heart rate is around 190.

4. The fourth stage level (CEXS4G, red)
has a heart rate of about 5 at a speed of around 25. As the speed increases to around 30, the heart rate rises to about 100. When the further increases to 35, the heart rate reaches about 125. At a speed of 40, the heart rate reaches about 10. Finally, at a speed of 45, the heart rate is around 175.

5. Fifth Stage Level (CEXS5G, Purple)
At a speed of around 25, the heart is approximately 50. As the speed increases to around 30, the heart rate rises to about 75. When the speed further increases to35, the heart rate reaches around 100. At a speed of 40, the heart rate is about 125. Finally, the speed reaches 45, the heart rate is approximately 150.

6. Sixth Stage Level (CEXS6G, Brown)
At speed of around 25, the heart rate is approximately 25. As the speed increases to around 30, the heart rate rises to about50. When the speed further increases to 35, the heart rate reaches around 75. At a speed of 40, the heart is about 100. Finally, when the speed reaches 45, the heart rate is approximately 125.

Scatter plots can help this visually observe the relationship between speed and heart rate at different stage levels.

It can be observed that: as the speed increases, the heart rate at each stage level an upward trend.

At the same speed, the increase in heart rate is different for different stage levels. Higher stage levels (such as CEXS1G have significantly higher heart rates at the same speed compared to lower stage levels (such as CEXS6G).

This indicates that at higher stage levels, the speed has a greater impact on the heart rate, reflecting the differences in exercise intensity and individual adaptability at different stage levels.

6. The relationship between BMI and blood oxygen content

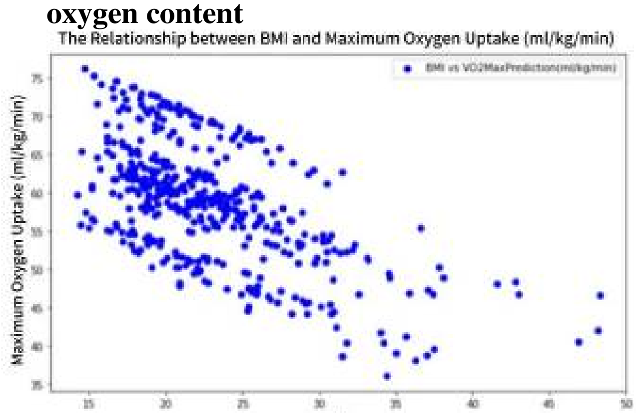

Figure 5.The relationship between BMI and maximum blood oxygen capacity observed

from the scatter plot shows a certain degree of correlation between BMI and maximum blood oxygen capacity but it is not a linear relationship.

Specifically:

1. Overall trend:
As BMI increases, the maximum blood oxygen capacity seems to show a downward trend,. e., individuals with higher BMI generally have lower maximum blood oxygen capacity.

2. Distribution characteristics:
The scatter plot shows a certain degree of scatter, i.e there is some variation in maximum blood oxygen capacity at the same BMI value. This may be due to the influence of other factors (such as activity level, health, etc.).

3. Outliers:
There may be some outliers in the scatter plot that deviate from the main trend, which may need further investigation and.

7. Build a model
In this study, a linear regression model was established after reading data from Excel files. The mean square error of the model is 0.20, which indicates that the average difference between the predicted value and the actual value is small. The R-square score is 0.79, which is close to 1, indicating that the model fits the data well. The model coefficient shows that when the independent variable increases by 1 unit, the dependent variable is expected to increase by about 0.01475311 units, and the intercept is -0.012440168990315636. In practical application, the interpretation should consider the actual value range of the independent variable. According to the classification report, the model performs very well in all categories, with precision, recall and f1-score all being 1.0. However, UserWarning appears, indicating that there may be problems in cross-validation under certain circumstances because the sample number of the minimum category is less than the fold number. However, since the current accuracy of 1.0 May not be a problem, if encountered in general, consider reducing the fold number or using stratified sampling to ensure category representation. The cross-validation scores ranged from 0.99365079 to 1.0, reconfirming that the model performed well on the test set. By calculating the Pearson correlation coefficient between the two variables as -0.61, it means that there is a moderate negative correlation between BMI and the maximum blood oxygen content, that is, with the increase of BMI, the maximum blood oxygen content tends to decrease, and the relationship is relatively strong but not completely linear, which may be affected by other

factors, and further consideration should be given to the interaction and nonlinear relationship. Based on the results of exploratory data analysis, this study selects multiple linear regression, Logistic regression and other methods to establish a model to analyze the relationship between lifestyle habits and physical and mental health. Considering the mutual influence of lifestyle factors, variable screening and interaction item analysis are carried out to ensure the accuracy and interpretability of the model. At the same time, the robustness and generalization ability of the model are verified by cross-validation and other methods.

8. Conclusions and Suggestions

Through the in-depth analysis of campus students' health-related data and the study of teenagers' living habits and other aspects, the following conclusions are drawn: The physical and mental health of adolescents is affected by a variety of living habits. Bad living habits such as unreasonable diet, lack of sleep, lack of exercise, lack of social interaction and over-reliance on technology products have many negative effects on their physical and mental development, and good living habits help promote healthy growth. Big data technology has some effectiveness in the analysis of campus student health detection, which can build a health detection system and reveal the health status of students, but it faces challenges such as data privacy and security, and there is room for improvement of statistical models. At the same time, there are some problems in the development path of youth sports, such as students' weak willingness to exercise, lack of data thinking in evaluation and insufficient visual supervision. Based on the above conclusions, the following suggestions are put forward: Families, schools and society should make joint efforts to help young people develop healthy lifestyle habits by carrying out health education courses, establishing work and rest schedules, enriching physical activities, developing social skills and guiding rational use of scientific and technological products. In the application of big data technology, it is necessary to strengthen data security management and sharing mechanism construction, improve data literacy of teachers and students, and continue to optimize the health detection system. For the development of youth sports, intelligent exercise mode should be built, the evaluation system should be improved by using data thinking, a visual supervision platform should be established, and its development should be promoted by policy and system guarantee, so as to improve the physical and mental health level of teenagers, help their healthy growth, narrow the gap with the ideal health state, and lay a solid foundation for the future development of teenagers. It is hoped that with the joint attention and efforts of all parties, adolescent health issues will be better improved and solved, and relevant fields will continue to develop and progress.

	Specified training plan (CEXPROT)	First - stage speed (CEXS1S)	First - stage level (CEXS1G)	Time spent in the stage (CEXS1TIM)
0	2.0	2.8	0.0	120.0
1	2.0	2.8	0.0	120.0
2	1.0	2.5	0.0	120.0
3	2.0	2.8	0.0	120.0
4	2.0	2.8	0.0	120.0

	Heart rate in the first - stage (CEXS1HR)	Second- stage speed (CEXS2S)	Second - stage level (CEXS2G)	Time spent in the stage (CEXS2TIM)
0	140.0	3.0	0.0	120.0
1	125.0	3.0	0.0	120.0
2	101.0	2.8	0.0	120.0
3	129.0	3.0	0.0	120.0
4	125.0	3.0	0.0	120.0

	Heart rate in the second - stage (CEXS2HR)	Third stage speed (CEXS3S)		Time spent in the stage (CEXS4TIM)	Heart rate in the fourth - stage (CEXS4HR)
0	144.0	3.3	...	120.0	189.0
1	135.0	3.3	...	120.0	167.0
2	105.0	3.0	...	120.0	136.0
3	132.0	3.3	...	120.0	160.0
4	127.0	3.3	...	120.0	159.0

	Fifth- stage speed (CEXS5S)	Fifth - stage level (CEXS5G)	Time spent in the stage (CEXS5TIM)	Heart rate in the fifth - stage (CEXS5HR)	Sixth - stage speed (CEXS6S)
0	4.5	7.5	85.0	210.0	NaN
1	4.5	7.5	120.0	185.0	4.5
2	3.5	7.5	120.0	162.0	4.2
3	4.5	7.5	120.0	196.0	4.5
4	4.5	7.5	120.0	201.0	4.5

	Sixth - stage level (CEXS6G)	Time spent in the stage (CEXS6TIM)	Heart rate in the sixth - stage (CEXS6HR)
0	NaN	NaN	NaN
1	10.0	33.0	NaN
2	10.0	118.0	203.0
3	10.0	120.0	203.0
4	10.0	120.0	208.0

[5 rows x 25 columns]

Explained variance ratio: [0.7294759 0.14457021]

	Body part status code (BMDSTATS)	Weight (BMXWT)	Height	BMI	BMI comment	Upper arm length	Arm circumference	Waist circumference	Calf circumference
0	1	38.6	131.6	22.3	4	27.7	25.4	71.9	32.3
1	1	89.7	172.0	19.8	2	38.4	28.0	79.4	35.3
2	3	92.5	167.1	33.1	4	35.9	37.9	96.4	46.8
3	1	12.4	90.2	15.2	2	18.3	15.1	46.8	19.4
4	1	66.4	160.1	25.9	4	34.2	29.5	90.0	36.9

	PC1	PC2
0	0.837409	1.209006
1	2.054399	-1.932101
2	6.451439	1.783303
3	-3.896997	1.015552
4	3.366372	0.315506

Figure 6 and 7

References

[1] Qi Donggui, Liu Rong, Wu Xiaoxi, Pang Jing, Deng Jiangbo, Wang Aiguo Investigation on the sleep quality of college students and its influencing factors [J]. Modern Preventive Medicine, 2007(5):875-77.

[2] Hu Dayi, Han Yaling, Ning Guang, Ma Changsheng. Prevention and rehabilitation of cardiovascular diseases [J]. Chinese of Cardiology, 2020,48(12):1000-1038.

[3] Liu Changrong. vascular diseases are prevalent in winter, and these young people are at risk [J]. China Youth Daily, 2024.

[4] Zhang Fang Lancet Public Health: Body weight and risk of comorbid cardiovascular diseases [J]. MedSci Original Translation.

[5] BMI and overall health: Scientific and practical guidance [J]. Baidu Literature.

[6] National Health Statistics Center: NNYFS examination data, 2022 (3)EB/OL]. NNYFS Examination Data.

[7] Relationship among physical literacy, mental health, and resilience in college students [EB/OL]. YAXDAY.