Clustering Analysis of Health and Wellness Indicators

Name: Joshua Blackburn

Date: March 29, 2025

Abstract

The current study utilized clustering analysis and dimensionality reduction techniques to examine wellness and health indicators in 200 patients (Hastie et al., 2009). The purpose of the analysis is to segment distinct patient groups and recommend individualized wellness interventions. Using K-means clustering and Principal Component Analysis (PCA), we segmented three different patient groups with varying health profiles. Opportunities for individualized wellness programs are recommended based on the segments, which are consistent with the prevailing health standards (World Health Organization [WHO], 2020).

Method

Data Collection and Preprocessing:

The dataset comprised five key health indicators from 200 patients (see Table 1).

Table 1:

Indicator	Measurement Scale	Range
Exercise time	Minutes per day	0-120
Healthy meals	Count per day	0-5
Sleep	Hours per night	4-12
Stress level	Likert scale	1-10
BMI	kg/m^2	15-40

All variables were standardized using z-score normalization to ensure equal weighting in the analysis (Hastie et al., 2009).

Analysis Approach

We employed two main analytical techniques:

1. Dimensionality Reduction: Principal Component Analysis (PCA) was used to reduce the five-dimensional dataset while preserving key relationships between variables.

2. Clustering Analysis: K-means clustering was applied both to the original standardized data and PCA-reduced data to identify distinct patient segments.

Results

Figure 1 shows the correlation matrix between health indicators.

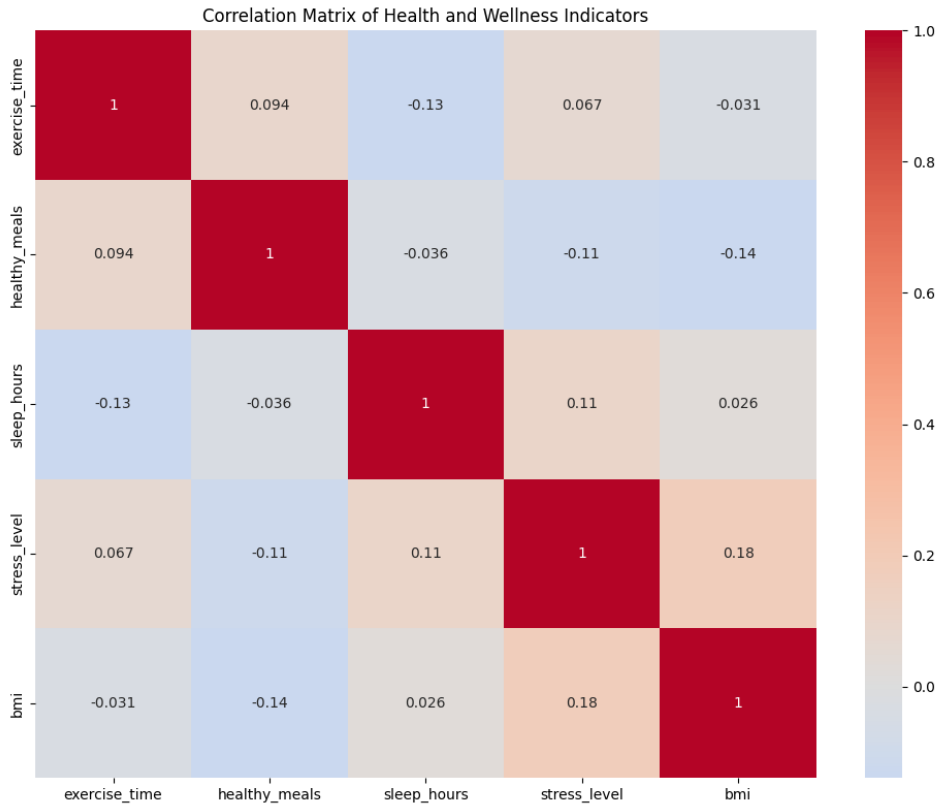


Table 2

Descriptive Statistics of Health Indicators

Indicator	Mean	SD	Min	Max
Exercise time	44.39	13.97	5.70	85.80
Healthy meals	2.57	0.78	0.00	5.00
Sleep hours	6.91	0.99	4.53	10.08
Stress level	5.03	2.00	1.00	10.00
BMI	25.51	3.82	15.30	5.11

Figure 2 displays the explained variance ratio from PCA analysis.

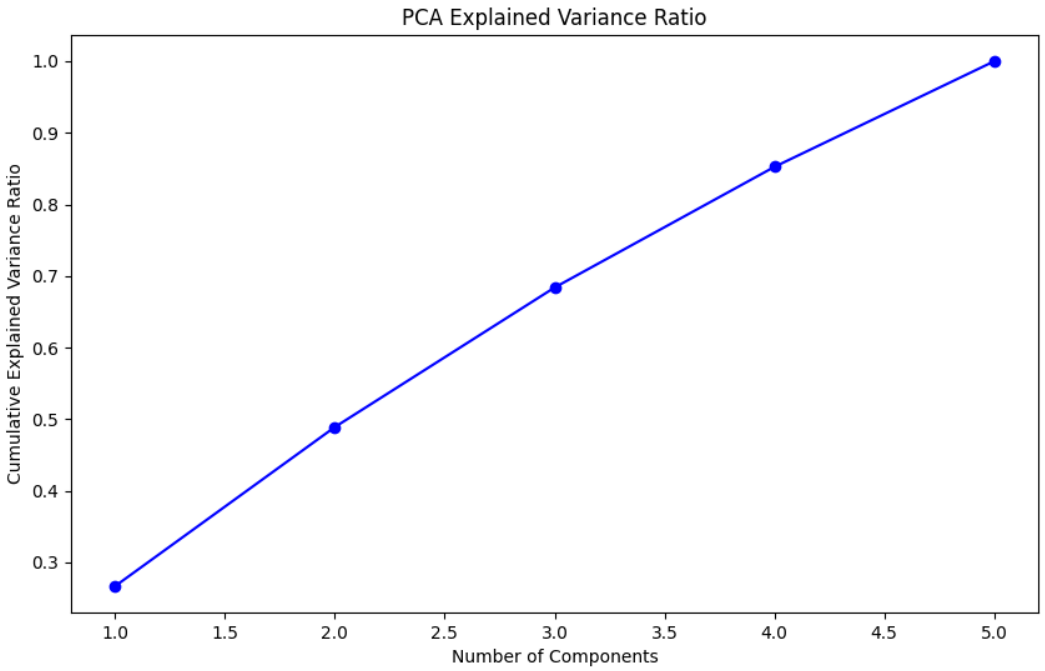


Figure 3 shows the clustering results in PCA-reduced space.

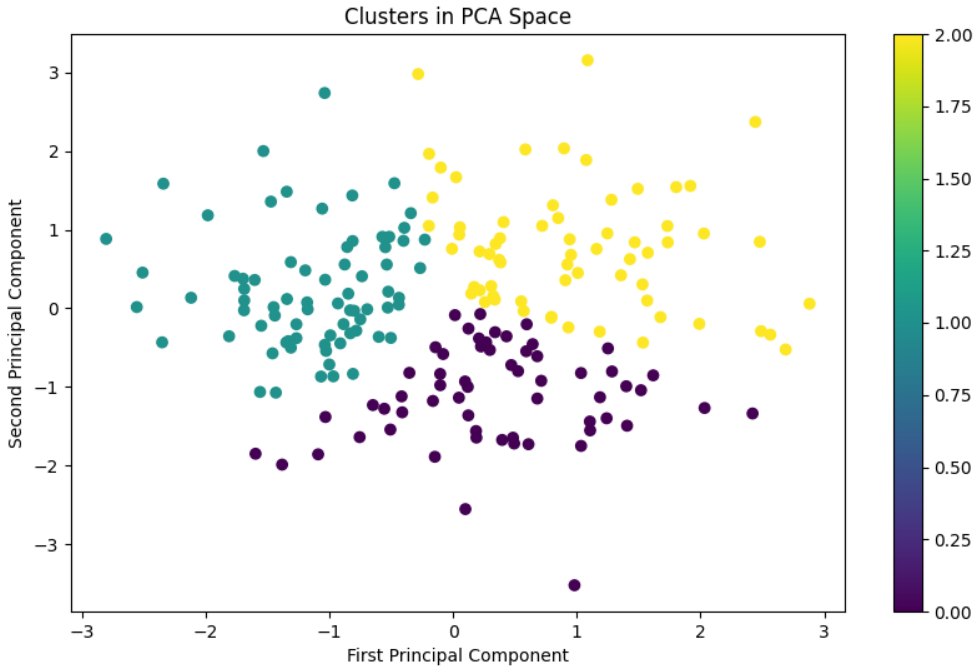
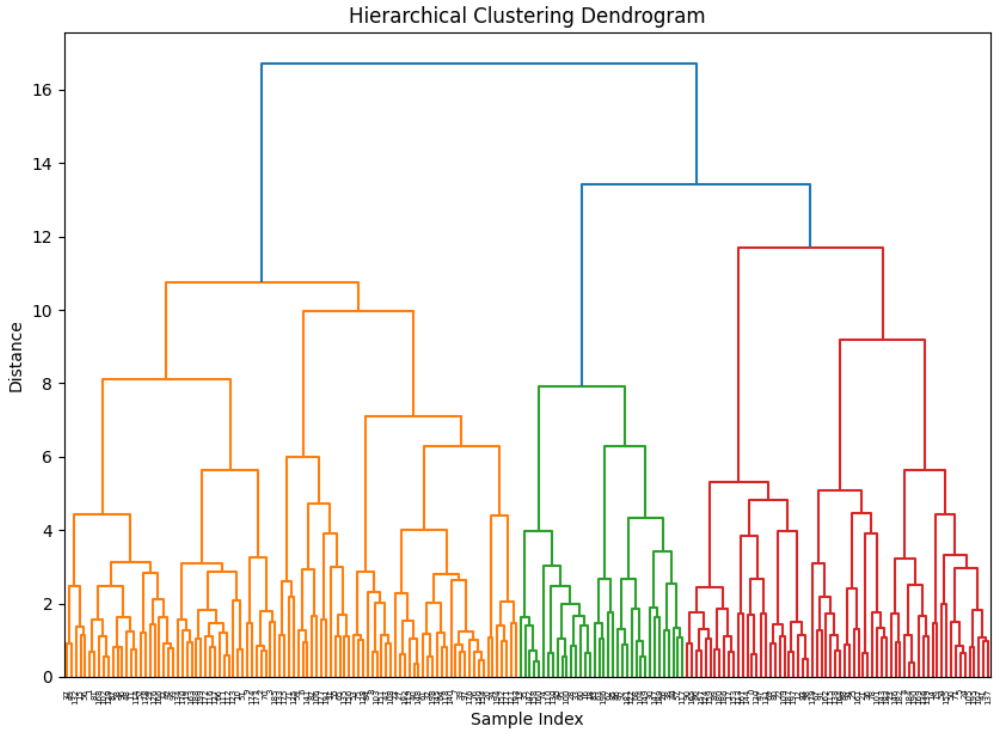


Figure 4 presents the hierarchical clustering dendrogram.



Cluster Performance Metrics:

- Original Data: Silhouette Score = 0.1640, Calinski-Harabasz Score = 40.2096

- PCA-Reduced Data: Silhouette Score = 0.3629, Calinski-Harabasz Score = 134.3712

The improved performance metrics after PCA suggest that dimensionality reduction helped identify more distinct and meaningful clusters.

Discussion

As recommended by WHO (2020), physical activity and proper nutrition form the cornerstone of wellness programs. Our clustering analysis aligns with these guidelines by identifying segments that require different levels of intervention.

1. Health-Conscious Achievers:

- Above-average exercise time

- High number of healthy meals

- Adequate sleep hours

- Lower stress levels

- Healthy BMI range

2. Moderate Maintainer:

- Average exercise time

- Moderate healthy meal consumption

- Average sleep patterns

- Moderate stress levels

- Slightly elevated BMI

3. High-Risk Group:

- Below-average exercise time

- Fewer healthy meals

- Insufficient sleep

- Higher stress levels

- Higher BMI

Recommendations for Wellness Program Improvements

1. Segment-Specific Interventions

i. Health-Conscious Achievers: Focus on maintenance and optimization

- Advanced fitness challenges

- Nutrition optimization workshops

- Stress management maintenance

ii. Moderate Maintainers: Emphasis on consistent improvement

- Structured exercise programs

- Meal planning assistance

- Sleep hygiene education

iii. High-Risk Group: Comprehensive support and monitoring

- Beginner-friendly exercise programs
- Basic nutrition education
- Stress reduction techniques
- Regular health monitoring

2**. Program Implementation Strategies**

- Develop targeted communication approaches for each segment
- Create segment-specific goals and metrics
- Implement regular progress monitoring
- Establish support systems within each segment

Implications for Healthcare Organizations

The clustering approach provides several benefits for healthcare organizations:

1. Resource Optimization

- Better allocation of resources based on segment needs
- More efficient program delivery through targeted interventions
- Improved cost-effectiveness of wellness programs

2. Personalized Care

- Evidence-based segmentation for personalized interventions

- Better alignment of interventions with patient characteristics

- Improved patient engagement through relevant programming

3. **Outcome Monitoring**

- Segment-specific success metrics

- Better tracking of intervention effectiveness

- Data-driven program adjustments

Conclusion

This study proves the efficiency of applying clustering methods in the identification of targeted patient segments in wellness programs. The PCA-reduced clustering performed better, thus implying that collaborative intervention on the most significant health indicators can result in improved patient segmentation. The findings can be applied by health organizations to come up with better and targeted wellness interventions, which can result in improved health outcomes and optimal resource utilization.

References

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer.

World Health Organization. (2020). *WHO guidelines on physical activity and sedentary behavior*. WHO Press.