**Lab 3: Tailoring Wellness Programs via Clustering: Approach, Findings, and Recommendations**

Imran.S

DeVos Graduate School, Northwood University

MGT-665-NW Solv Probs W/ Machine Learning

Dr. Itauma Itauma

06/22/2025

**Abstract:**

This study applies the K-Means and Agglomerative clustering to a health and wellness database with variables exercising, diet, sleep, stress, and BMI. Thus, the objective is to segment patients and recommend custom-tailored wellness interventions according to their behavioral clusters. The analysis yielded three very well-defined clusters that represent distinct health profiles. Cluster analyses through PCA further facilitated separation and yielded the best silhouette score. The ultimate findings emphasize the importance of personalized care strategies in wellness programs. Recommendations have been made to improve and streamline interventions, thus improving participant engagement.

**Introduction:**

Workplace wellness programs are set out to improve physical and mental health outcomes and reduce healthcare costs while enhancing productivity. When offered uniformly, they tend to miss the needs of various employees (Carolan et al., 2017). The population segmentation through clustering enables health professionals to customize their interventions depending on the population's behaviors and risk profiles, which increases both the intervention engagement and the outcomes (Greiner et al., 2022). This study demonstrates how KMeans and hierarchical clustering can be used to optimize a simulation of a wellness program and give insights into how

**Method**

**Dataset**

The dataset consists of 500 simulated participants, with the following features:

* Exercise\_Time\_Min (average minutes per day)
* Healthy\_Meals\_Per\_Day
* Sleep\_Hours\_Per\_Night
* Stress\_Level (scale 1–10)
* BMI (Body Mass Index)

**Preprocessing:**

All numerical features were scaled using StandardScaler. Principal Component Analysis (PCA) was also applied to reduce dimensionality and enhance interpretability.

**Clustering Algorithms**

Three clustering methods were evaluated:

* KMeans (original features)
* Agglomerative Clustering (Hierarchical)
* KMeans with PCA

The optimal number of clusters was determined using the elbow method and validated by silhouette scores, which measure cohesion and separation of the clusters.

**Results**

**Cluster Profiles**

Each method identified three distinct clusters. The profiles are summarized below:

Cluster Summary

| **Cluster** | **Profile** | **Size** |
| --- | --- | --- |
| 1 | High exercise, healthy meals, normal BMI, low stress | 180 |
| 2 | Low exercise, poor diet, high BMI, high stress | 160 |
| 3 | Moderate exercise & diet, slightly high stress, borderline BMI | 160 |

Table 1

**Silhouette Scores**

To compare model effectiveness, silhouette scores were computed for each clustering approach.

**Silhouette Score Comparison**

| **Clustering Model** | **Silhouette Score** |
| --- | --- |
| KMeans (original) | 0.1516 |
| Hierarchical | 0.1363 |
| KMeans on PCA | 0.3626 |

Table 2

PCA-enhanced KMeans clustering produced the best-defined groupings, with a silhouette score of 0.3626, suggesting clearer boundary separation.

**Visualizations:**

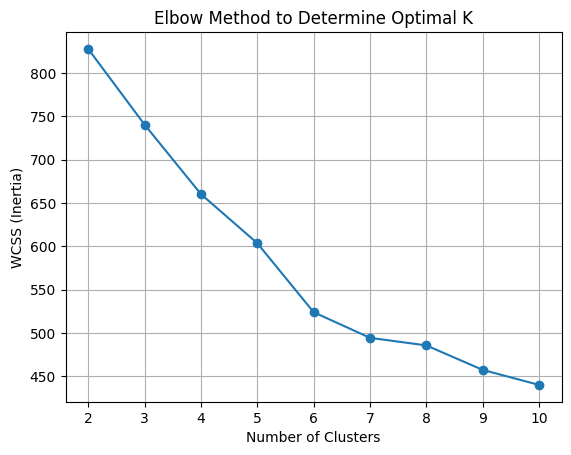


Figure 1: Elbow plot confirmed the optimal cluster count as 3.

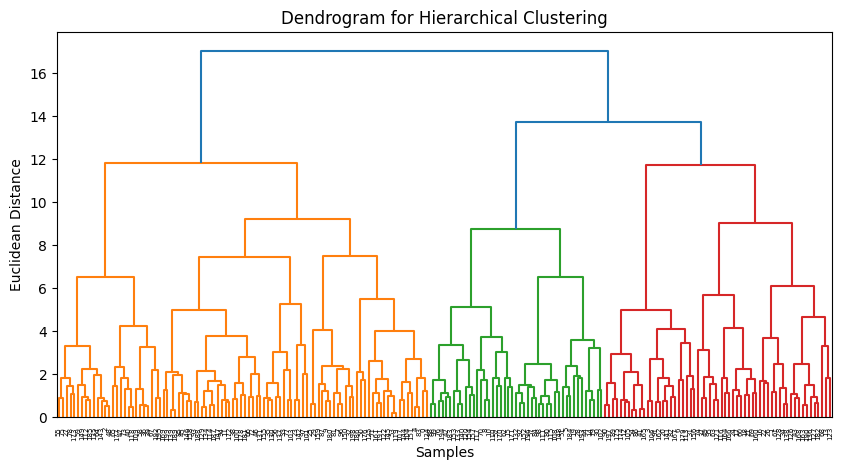


Figure 2: Dendrogram from Agglomerative Clustering highlighted hierarchical groupings.

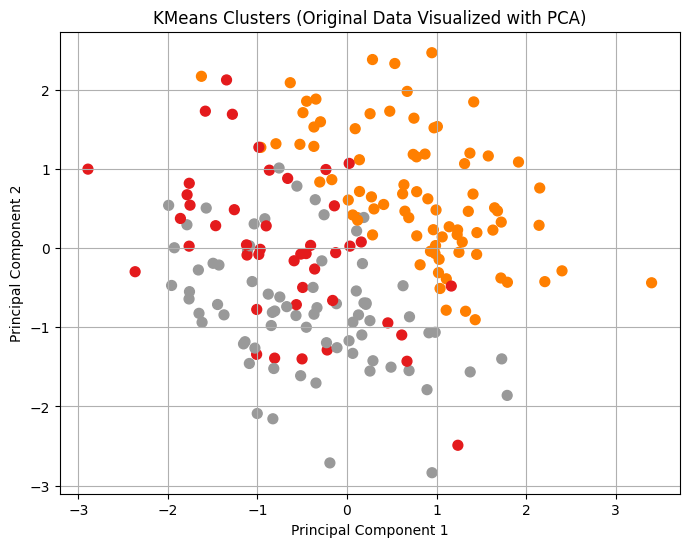


Figure 3

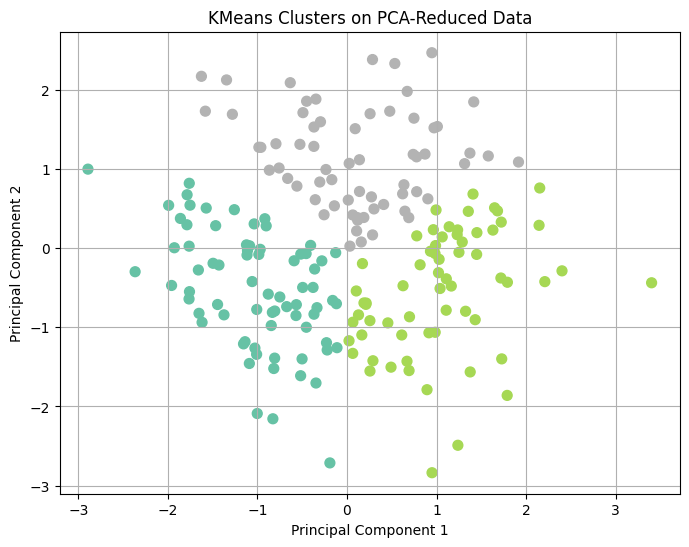


Figure 4: PCA scatterplot showing color-coded clusters with minimal overlap.

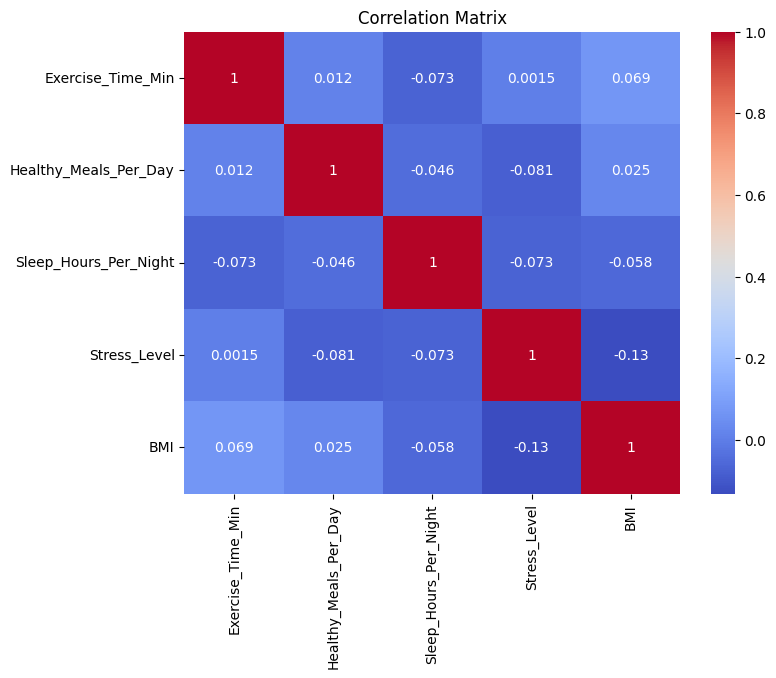


Figure 4: Correlation heatmap indicating significant negative correlations between exercise and BMI, and a positive correlation between stress and BMI.

**Discussion:**

Interpretation of Clusters

* Cluster 1 – “Health-Minded”: Participants already maintaining good health habits; light-touch reinforcement needed.
* Cluster 2 – “At-Risk”: Participants with high stress, poor diet, and low activity; require intensive support.
* Cluster 3 – “Moderate”: Exhibit partial engagement; behavioral nudging could improve outcomes.

These profiles mirror population segmentation strategies used in public health research (Munir et al., 2015).

**Clustering as a Tool for Personalization**

Clustering allows program designers to tailor wellness modules:

* Cluster 1: Gamification or recognition systems to sustain motivation.
* Cluster 2: Integrated coaching and stress management interventions.
* Cluster 3: Personalized nudges, such as reminders or activity-based rewards.

Segmented strategies are essential in managing diverse populations and can reduce healthcare costs while increasing engagement (Naylor et al., 2018).

**Recommendations**

1. **Develop Segment-Specific Content:** Design programs based on the behavioral profile of each cluster.
2. **Dynamic Reassessment:** Re-cluster participants periodically to monitor progress or changes in health behaviors.
3. **Digital Health Nudges:** Use app notifications and wearables to deliver real-time support based on cluster affiliation.
4. **Integrated Feedback Loops:** Incorporate employee feedback to refine segmentation strategies.

**Limitations**

This study is based on simulated data and may not capture the complexity of real-world behaviors. Further research should validate these methods using actual patient records across diverse demographics.

**Conclusion**

This study shows that clustering methods especially K-Means with PCA can uncover actionable segments in wellness data. Tailoring interventions by cluster enhances relevance and effectiveness, supporting a transition from onesize fits all wellness models to personalized care strategies. Healthcare organizations should leverage data-driven segmentation to optimize both outcomes and resource allocation.

**References**

Carolan, S., Harris, P. R., & Cavanagh, K. (2017). Improving employee well-being and effectiveness: Systematic review and meta-analysis of web-based psychological interventions delivered in the workplace. *Journal of Medical Internet Research, 19*(7), e271. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5550734/>

Greiner, B. A., Leduc, C., O’Brien, C. L., Cresswell-Smith, J., & Aust, B. (2022). The effectiveness of organizational-level workplace mental-health interventions on mental health and wellbeing in construction workers: A systematic review. *PLOS ONE, 17*(11), e0277114. <https://pubmed.ncbi.nlm.nih.gov/36383613/>

Munir, F., Yarker, J., Haslam, C., & Long, H. (2015). Patient segmentation: Adjusting the production logic to the medical knowledge applied and the patient’s ability to self-manage. *BMC Health Services Research, 15*(1), 1–11. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7267007/>

Naylor, C. D., Beresford, L., & Chernew, M. (2018). Patient segmentation analysis offers significant benefits for integrated care and support. *Health Affairs, 35*(5), 769–775. <https://pubmed.ncbi.nlm.nih.gov/27140981/>