Holistic Lifestyle Patterns and Work-Life Balance Among Adults

Potla Manasvi Reddy **Monica Tejaswi Kommareddy**

Abstract

Research Question: To what extent can a combination of lifestyle behaviors such as diet, physical activity, sleep habits, stress management, and social interaction, along with demographic factors like age and gender, help predict overall wellbeing and life satisfaction in adults?

Work-life balance has become an increasingly important concern as working adults face rising levels of stress, burnout, and health decline. This study examines how lifestyle behaviors and demographic factors collectively influence overall work-life balance and wellbeing. Using publicly available survey data, the analysis focuses on identifying modifiable daily habits, such as physical activity, stress-related behaviors, mindfulness practices, and sense of purpose, that contribute most strongly to life satisfaction. Grounded in positive psychology and behavioral science frameworks, the study emphasizes the role of habitual behaviors in shaping wellbeing outcomes. The findings aim to provide data-driven insights that support informed individual decision-making and contribute to broader public health initiatives aligned with the United Nations Sustainable Development Goals related to sustainable and healthy lifestyles.

1 Data Description

Dataset: Lifestyle and Wellbeing Data. **Source**: Kaggle (originally collected via surveys on Authentic-Happiness.com) **Sample Size**: 15,972 adult respondents. **Data Ownership and Collection**: The data was originally collected through voluntary online surveys hosted by Authentic-Happiness.com to assess lifestyle behaviors and perceived wellbeing. The dataset is publicly available, anonymized, and does not contain personally identifiable information. **Ethics and Privacy Considerations**: Because the dataset is public and anonymized, no IRB approval was required.

Outcome Variable (Y):

WORK_LIFE_BALANCE_SCORE a continuous score representing overall perceived work-life balance.

Predictor Variables (X):

Lifestyle factors including physical activity, nutrition, sleep, stress indicators, meditation, social engagement, achievement, purpose, and income sufficiency, along with demographic variables (age, gender, BMI category).

2 Data Preprocessing and Exploratory Data Analysis

Data Cleaning and Preparation: The dataset was inspected for missing values, and no missing data were identified across all variables. The Timestamp variable was removed because it was not analytically relevant to the research question. Categorical variables (GENDER and BMI_RANGE) were converted into numeric dummy variables for modeling.

Exploratory Data Analysis: Summary statistics, histograms, and boxplots were generated to examine distributions, detect potential outliers, and understand variable ranges. A correlation heatmap of numeric variables was created to assess relationships between predictors and the outcome variable. Refer Figure 1 and Figure 2.

Key Findings from EDA: Lifestyle variables such as physical activity, purpose, income sufficiency, and stress indicators showed meaningful associations with work-life balance. After visualization and inspection, the data was deemed sufficient and appropriate to proceed with statistical modeling.

3 Methods and Model Selection

This project used three complementary methods to analyze the relationship between lifestyle behaviors, demographics, and work-life balance. All numeric features were standardized using z-score scaling to ensure comparability across variables with different ranges.

1. **Principal Component Analysis (PCA)**:

   Applied to explore underlying behavioral dimensions and reduce redundancy among correlated lifestyle variables. PCA helped identify key patterns such as purpose-driven behavior, physical activity, and stress-related factors.

2. **Linear Regression**:

   Used to model WORK_LIFE_BALANCE_SCORE as a function of lifestyle and demographic predictors. This method was chosen for its interpretability, allowing clear identification of behaviors that positively or negatively influence work-life balance.

3. **K-Means Clustering**:

   Used to identify natural lifestyle profiles within the population. This unsupervised approach complemented regression by revealing group-level patterns in behavior and wellbeing.

These models were selected because they are interpretable, statistically grounded, and appropriate for addressing the stated research question without unnecessary complexity.

4 Results

4.1 Principal Component Analysis Results:

PCA revealed that the first principal component explained 22.5 percent of the total variance, primarily reflecting purpose-driven and achievement-related behaviors. The first five components together explained approximately 48 percent of the variance, indicating that wellbeing behaviors are multi-dimensional rather than driven by a single factor. Refer Figure 3.

4.2 Regression Results:

The linear regression model achieved strong predictive performance:

R^2: 0.956

Mean Squared Error: 88.07

Mean Absolute Error: 8.22

This indicates that approximately 95.6 percent of the variability in work-life balance was explained by the predictor variables.

Most influential positive predictors included sufficient income, daily steps, places visited, weekly meditation, and having a clear life vision. The most influential negative predictors were lost vacation time and daily shouting, which represent chronic stress and emotional strain.

4.3 Clustering Results:

K-Means clustering (k = 3) identified three distinct lifestyle profiles:

Cluster 0: Balanced and active individuals with moderate stress and higher wellbeing.

Cluster 1: High-achievement individuals with lower recovery and signs of potential burnout.

Cluster 2: Low-engagement, high-stress individuals with the lowest work-life balance scores.

A PCA visualization showed clear separation between clusters, confirming meaningful behavioral groupings. Refer Figure 4.

5 Conclusions

This project demonstrates that work-life balance is highly predictable from daily lifestyle behaviors and demographic context. Purpose-driven actions, physical activity, mindfulness practices, and financial stability significantly improve wellbeing, while stress-related behaviors substantially reduce it.

The results confirm that adults naturally fall into distinct lifestyle profiles, suggesting that wellbeing interventions may be more effective when tailored to behavioral patterns rather than applied uniformly.

Overall, this analysis provides actionable insights that are relevant to individuals, employers, and public health practitioners seeking to improve wellbeing outcomes.

Github Repository Link

https://github.com/KommareddyMonicaTejaswi/INFO-511-FINAL-PROJECT.git

Future Research

Future work could include inferential statistical testing across demographic groups, validation using cross-validation techniques, and exploration of nonlinear models. Longitudinal data could be used to effectively suggest or investigate potential causal relationships between lifestyle changes and wellbeing outcomes.

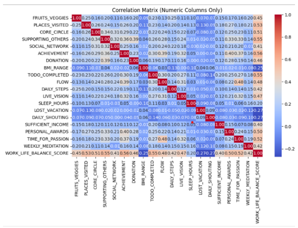

Figure 1: EDA - Corelation Analysis

Generative AI Tool Use Acknowledgement

Generative AI tools were utilized solely to refine and improve the clarity, grammar, and readability of the report's written content. All data analysis, interpretation, results, and code execution were performed, verified, and remain the intellectual product of the authors.

Citations and Acknowledgements

United Nations. (n.d.). *Sustainable Development Goals*. https://sdgs.un.org/goals

Lifestyle and Wellbeing Data [Data set]. Kaggle. Retrieved December 15, 2025, from https://www.kaggle.com/datasets/ydalat/lifestyle-and-wellbeing-data/data

- Authentic-Happiness.com
- 360living.co
- guidebienetre.org

Researcher Bio-Sketches

Potla Manasvi Reddy
Degree Plan: Data Science Graduate Program
Year: First Year, First Semester
Role: Data preprocessing, modeling, PCA, regression analysis, and interpretation.

Monica Tejaswi Kommareddy
Degree Plan: Data Science Graduate Program
Year: Second Year, Second Semester
Role: Exploratory data analysis, clustering analysis, visualization, and results interpretation.

A Appendix

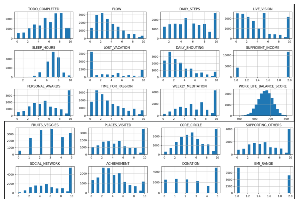

Figure 2: EDA - Histograms

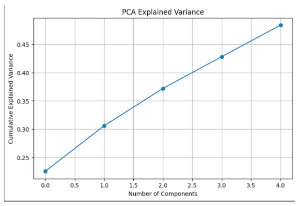

Figure 3: PCA Explained Variance

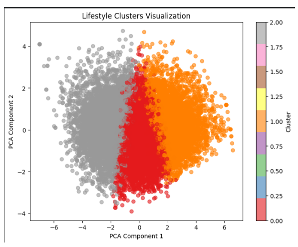

Figure 4: Cluster Visualization