Temporal Analysis of Bidirectional Links Between Physical Activity and Sleep Patterns

Social Dynamics Lab

Yishak Tadele Nigatu

Student ID: mat.247815
University of Trento, Italy
yishaktadele.nigatu@studenti.unitn.it

Course Instructors:
Professor Ivano Bison
Professor Giuseppe Alessandro Veltri

***Abstract*—Physical activity and sleep are often described as closely related, yet their day-to-day relations is not clearly defined. While previous cross-sectional studies suggest that more active individuals report better sleep, fewer studies have tested whether activity and sleep influence each other over time.**

In this study, a longitudinal study is performed on a dataset from 204 University of Trento students that combines survey, time-diary response, and motion sensor records. After preprocessing and integration of activity, nutrition, and sleep variables, daily profiles have been created and clustering is applied to identify behavioral groups. Temporal relationships were then examined using mixed-effects models and structural equation modeling.

The results show clear autoregressive patterns within both activity and sleep, but no evidence of significant cross-lagged effects between the two domains. Students who were generally more active did report better sleep quality, but this association was concurrent rather than predictive. These findings suggest that sleep and activity should be considered as parallel lifestyle factors, influenced by broader health behaviors, rather than as direct causal drivers of each other.

I. Introduction

Understanding how physical activity and sleep influence each other is essential for designing effective health interventions, particularly in young adult populations such as university students. Cross-sectional work has long shown robust concurrent associations between greater habitual physical activity and better sleep quality and duration (meta-analytic syntheses and systematic reviews). For example, a meta-analytic review of experimental and observational studies concluded that both acute and regular physical activity improve several sleep outcomes (sleep quality, sleep onset latency and sleep efficiency) [1], [2].

However, whether physical activity and sleep have bidirectional temporal (causal) effects on each other on a day-to-day basis is still unknown. Recent longitudinal and intensive longitudinal (daily/EMA) studies have produced mixed results: work by [3] finds evidence of short-term bidirectional links, whereas [4] report mainly concurrent associations, and [5] only found within-person effects that disappear once stable individual differences are accounted for . These divergent findings suggest that model choice and handling of within- versus between-person variance matter for inference in the topic.

Missingness and preprocessing choices strongly affect longitudinal inference. When longitudinal records are sparse or missing not at random, single imputation or aggressive interpolation can bias estimates of temporal dynamics; principled approaches such as multiple imputation [6] or full-information maximum likelihood (and sensitivity analyses for missing-data mechanisms) are recommended. Moreover, modern longitudinal causal modeling (e.g., random-intercept cross-lagged panel models (RI-CLPM) [7], time-varying vector autoregressive models (mlVAR) [8], and continuous-time SEM) explicitly separate within-person dynamics from stable between-person differences — a crucial distinction when testing bidirectional hypotheses.

Using a multimodal longitudinal dataset collected among University of Trento students (survey, time-diary notifications, and motion sensors), this study aims to (1) characterize joint temporal patterns of physical activity, nutrition, and self-reported sleep quality, (2) identify behavioral clusters via unsupervised learning, and (3) test bidirectional temporal associations between activity and sleep using longitudinal models. Key features are the integration of diary nutrition data with motion sensors and personality/demographic covariates, and the use of mixed-effects and structural equation approaches to probe temporal relations.

II. Methodology and Implementation

A. Study Scope

This study analyzed longitudinal behavioral data collected from students at the University of Trento, Italy, as part of the WeNet – The Internet of Us project. The data integrates three streams: (1) survey responses capturing demographics and Big Five personality traits, (2) time-diary entries of daily activities, food intake, and self-reported sleep quality, and (3) motion sensor records of physical activity.

B. Data Sources and Metadata Description

1) Data Source: The dataset consists four main dimensions:

Physical Activity: Activity data encompasses eight sport categories ranging from low-intensity walking to high-intensity activities. Each activity is coded with corresponding intensity levels (low/moderate/high).

Nutritional Intake: Data on nutrition covers 20 main meal categories and 27 snack categories. To allow quantitative analysis, these categorical records were mapped to nutritional values using the FAO Global Nutrient Conversion Table, and FDA Daily Value on Nutrition and Supplement. This provided estimates of daily intake for energy (kcal), macronutrients (protein, fat, carbohydrates, fiber, alcohol), minerals (e.g., calcium, iron, magnesium, phosphorus, potassium, zinc), and vitamins (A, B, C, retinol, carotene, thiamin, riboflavin). In this way, qualitative food reports were converted into standardized nutritional profiles for each participant.

Individual characteristics: Demographic information and Big Five personality traits (extraversion, agreeableness, conscientiousness, neuroticism, openness)

Sleep Quality: Self-reported sleep quality ratings on a 5-point Likert scale enable assessment of sleep patterns and consistency.

C. Data Preprocessing

The following section will describe in detail the type of data preprocessing performed on the raw data to provide an integrated quality dataset suitable for the study.

1) Participant Selection and Engagement Filtering: The study implemented an engagement-based filtering criterion. The analysis examined participants' notification response rates across three metrics (`tdtot`, `motot`, `sntot` which respectively are "Total TD notification fill in", "Total morning notification fill in" and "Total snack notification fill in") and computed total engagement scores.

Participants were considered for the study given that their total responses exceeded 60% of the maximum observed response count, resulting in a high-engagement cohort of 192 participants.

2) Temporal Alignment and Feature Engineering:

1) Social Time Adjustment: To align recorded timestamps with participants' natural activity and meal patterns, the preprocessing applied a temporal shift of -5 hours to all timestamps, creating a "social datetime" variable:

$$t_{social} = t_{recorded} - 5hours \tag{1}$$

2) Time Period Categorization: Continuous timestamps are categorized into discrete time periods reflecting typical daily routines:
   - Morning: 06:00–09:59
   - Midday: 10:00–11:59
   - Lunch: 12:00–15:59
   - Afternoon: 16:00–17:59
   - Evening: 18:00–19:59
   - Night Time: 20:00–21:59
   - Mid Night: After 22:00

3) Nutritional Data Processing: The transformation of categorical food consumption data to quantitative nutritional values involved several steps:

1) **Food Item Parsing**: Comma-separated food item identifiers were expanded into individual entries.
2) **Nutrient Mapping**: Each food item was mapped to its nutritional profile using the FAO Global Nutrient Conversion Table, yielding values for energy (kcal), macronutrients (protein, fat, carbohydrates, fiber, alcohol), micronutrients (calcium, iron, magnesium, phosphorus, potassium, zinc), and vitamins (A, B, C, retinol, carotene, thiamin, riboflavin).
3) **Aggregation**: Nutrient values were summed for each participant's daily intake.
4) **Missing Value Treatment**: Null nutrient values were replaced with zeros to maintain computational stability.

4) Behavioral Variable Processing:

1) Sleep Quality Standardization: Categorical sleep quality ratings were mapped to a 5-point Likert scale (1=lowest, 5=highest). Missing sleep values were imputed using linear interpolation within each participant's time series to preserve individual sleep patterns while maintaining data completeness.
2) Physical Activity Encoding: Sport activities were processed through:
   a) Categorical encoding of eight activity types
   b) One-hot encoding to create binary indicator variables
   c) Temporal aggregation to capture daily activity patterns

5) Data Integration and Aggregation: The preprocessing pipeline produced two complementary datasets:

1) **Event-level Dataset**: Preserves temporal granularity with individual behavioral events, nutritional intake records, and sleep assessments.
2) **Daily Summary Dataset**: Aggregates data by participant, date, and time period using appropriate summary statistics:
   - Nutrients: mean values to represent typical intake
   - Sports: sum of activity indicators to capture total engagement
   - Sleep: mean values to represent daily well-being

6) Step Count Integration: Motion sensor data underwent timestamp parsing and daily aggregation, with complex timestamp formats handled through custom parsing functions to ensure accurate temporal alignment with other behavioral data. This data is highly sparse, and contains only 128 individuals.

7) Final Dataset Preparation: The final preprocessing stage merged behavioral data with demographic and personality features from the diary dataset. Data type consistency was ensured through standardized ID formatting, and inner joins preserved only participants with complete demographic profiles. The resulting datasets provide comprehensive lifestyle profiles suitable for multi-level behavioral analysis, containing:

- Temporal features (social datetime, time periods, weekday indicators)
- Nutritional profiles (20+ macro and micronutrients)
- Activity patterns (8 sport categories with intensity levels)
- Psychological indicators (sleep quality)
- Demographic characteristics and personality traits

III. Exploratory Data Analysis

A. Data Quality Assessment

Prior to analysis, a comprehensive data quality assessment was conducted to identify missing values and ensure data integrity. Table I presents the missing data patterns across key variables.

TABLE I: Missing Data Assessment

Variable	Missing Count	Missing Percentage
Sleep	208,396	98.46%
Mood	71,638	33.85%
Department	1,114	0.53%
Personality Traits	1,114	0.53%

The analysis revealed substantial missing data for sleep measurements (98.46%), necessitating linear interpolation during preprocessing. Mood data showed moderate missingness (33.85%), while demographic and personality variables demonstrated minimal missing values (0.53%), attributed to incomplete diary responses from a small subset of participants.

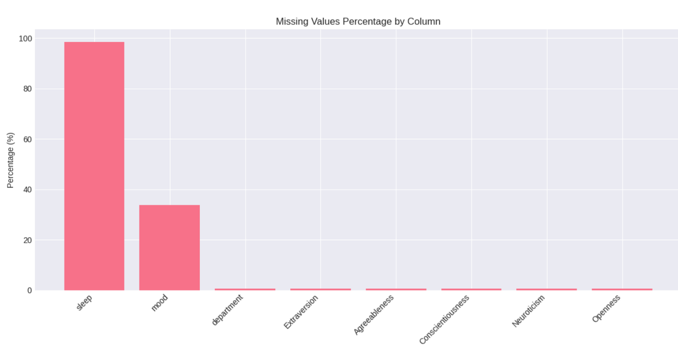

Fig. 1: Missing Values Percentage by Variable

B. Demographic Characteristics

1) Sample Composition: The study sample consisted of 204 Italian university students, providing a culturally homogeneous population for lifestyle analysis. Figure 2 illustrates the distribution across key demographic variables.

2) Academic Distribution: Department representation showed concentration in Engineering and Applied Sciences (47 students) and Social Sciences (40 students), followed by Natural Sciences (27), Law (25), Business/Economics (24), Humanities (17), Medicine and Veterinary (8), and International Relations (2). This STEM-heavy distribution may influence physical activity and dietary patterns.

The cohort distribution spanned from Cohort 17-18 to 31+, with highest representation from Cohort 20 (34 students), followed by Cohorts 22 (27), 19 (27), and 21 (25). Degree distribution revealed 120 Bachelor's students (58.8%) and 70 Master's students (34.3%), reflecting typical university enrollment patterns.

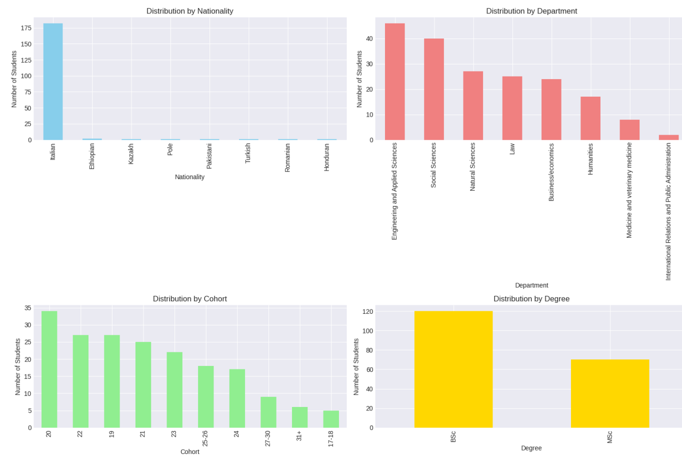

Fig. 2: Demographic Distributions: Nationality, Department, Cohort, and Degree

3) Personality Profile Assessment: Big Five personality trait distributions revealed diverse psychological profiles within the sample (Figure 3):

- **Extraversion**: Approximately normal distribution centered around 50-60, indicating balanced social engagement
- **Agreeableness**: Bimodal distribution with peaks at 80 and 95, suggesting predominantly cooperative tendencies
- **Conscientiousness**: Right-skewed distribution with concentration at 70-80, indicating high self-discipline
- **Neuroticism**: Left-skewed distribution centered at 50-60, suggesting moderate emotional stability
- **Openness**: Concentration around 75-85, indicating high openness to experience

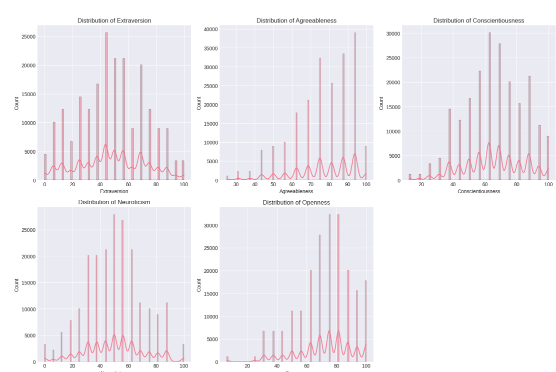

Fig. 3: Big Five Personality Trait Distributions

These personality profiles suggest a sample characterized by high conscientiousness and openness, moderate extraversion and neuroticism, and high agreeableness—traits that may correlate with health behavior engagement and study participation.

C. Nutritional Data Analysis

1) Food Consumption Patterns: Analysis of food consumption revealed water as the most frequently reported item (5,868 records), followed by carbohydrate-rich foods (rice/-pasta: 4,938; bread/cereals: 4,339) and vegetables (4,515). Protein sources showed lower frequency (meat: 1,135; fish: 557), while alcohol consumption was minimal (beer: 201; wine: 250; spirits: 24), consistent with a student population.

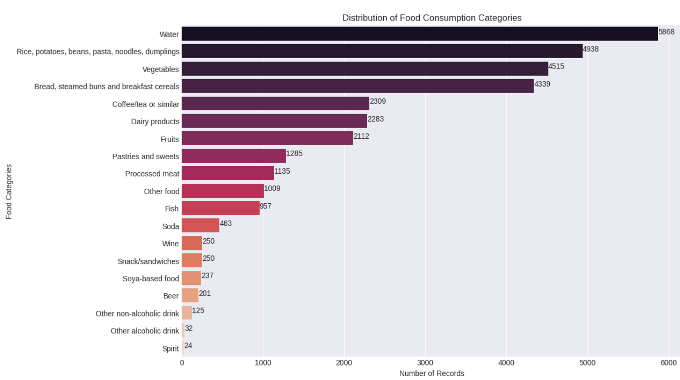

Fig. 4: Distribution of Food Consumption Categories

2) Nutrient Distribution Analysis: Macronutrient correlation analysis revealed strong positive relationships between energy intake and all macronutrients (energy-protein: $r = 0.94$, energy-fat: $r = 0.94$, energy-carbohydrates: $r = 0.94$), indicating proportional consumption patterns. Micronutrient distributions showed log-normal characteristics with substantial right skew, particularly for vitamins A and B complex.

3) Temporal Nutritional Patterns: Daily nutrient intake patterns demonstrated consistent hierarchical relationships throughout the study period. Energy intake showed gradual increase from study initiation (mean: 30 kcal) to completion (mean: 50 kcal), with notable spike around December 3rd coinciding with holiday season. The Nutrient Density Score (NDS) remained relatively stable across clusters after initial adaptation period, with Cluster 2 maintaining highest scores (76-78), Cluster 1 intermediate (73-76), and Cluster 0 lowest (69-74).

4) Circadian Eating Patterns: Temporal analysis of energy intake revealed distinct circadian patterns, with lunch period (12:00-15:00) showing highest consumption (mean: 80-110 kcal), followed by evening (18:00-20:00) and morning (06:00-10:00) peaks. Minimal nighttime consumption (22:00-05:00) indicated adherence to traditional meal patterns. This temporal structure remained consistent across study duration, suggesting stable eating routines within the student population.

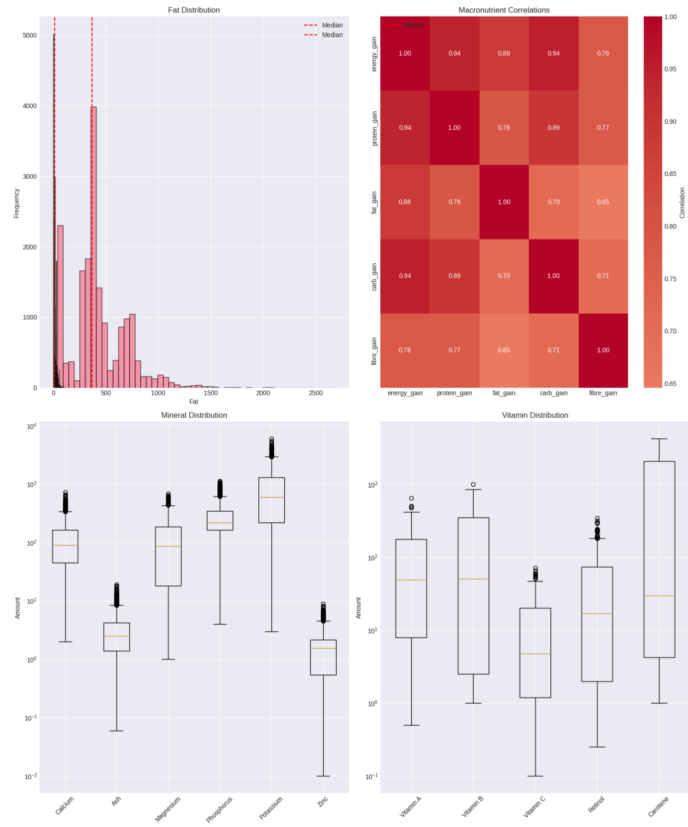

Fig. 5: Macro-Nutrient Correlation and Distribution

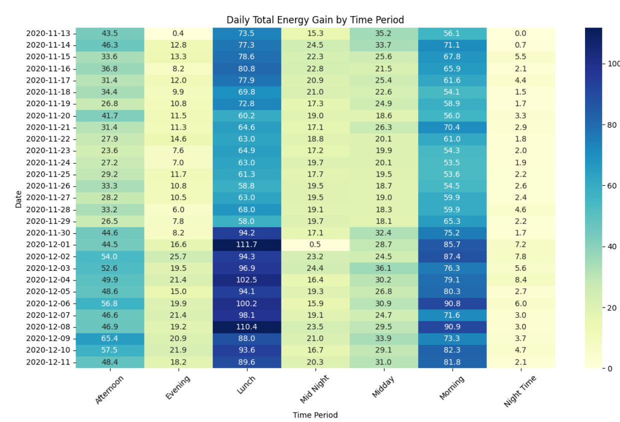

Fig. 6: Energy Gain By Time Classification

D. Physical Activity Analysis

1) Activity Participation Patterns: Physical activity analysis revealed gymnastics/fitness as the most frequently reported activity (approximately 900 participations), followed by other indoor activities (300), walking/trekking (280), and jogging/running (200). Water sports and ball games showed minimal participation, likely due to seasonal and facility constraints.

2) Circadian and Weekly Patterns: Hourly analysis revealed bimodal activity distribution with primary peak at 14:00-15:00 (40+ participations) and secondary peak at 10:00-

11:00, corresponding to typical university break periods. Weekend analysis showed significantly higher activity levels on Sundays (70+ participations) compared to weekdays (20-45 participations), with walking/hiking dominating weekend activities and jogging/running showing consistent weekday preference.

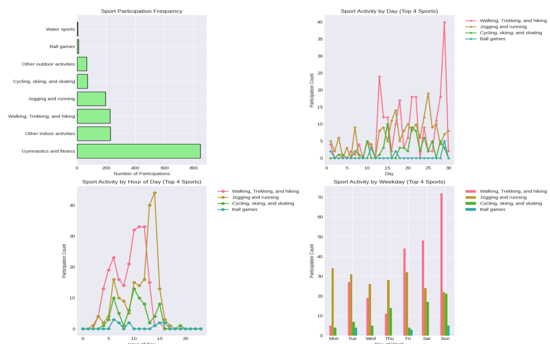

Fig. 7: Sport Participation Frequency Distribution

3) Temporal Activity Patterns: Daily activity trends demonstrated substantial variability across the study period. Walking/hiking showed peak participation in late November (0.0045 daily mean), while gymnastics/fitness maintained relatively stable participation (0.004-0.006 daily mean). Notable spikes in ball games and cycling occurred in early December, suggesting weather-dependent or event-driven participation. Water sports showed isolated peaks, indicating sporadic access or seasonal limitations.

Fig. 8: Sport Activity Patterns by Time of Day and Day of Week

IV. User Profiling

Following data preprocessing, the study developed comprehensive user profiles that capture multidimensional lifestyle patterns. This profiling approach integrates preprocessed physical activity behaviors, nutritional habits, and sleep patterns data to create holistic representations of individual lifestyle characteristics.

A. Physical Activity Profiling

The physical activity profile comprises three key components designed to capture both preferences and patterns:

1) Primary Sport Identification: The analysis identified each participant's dominant physical activity by:

1) Extracting all sport-related variables (prefixed with `sport_`)
2) Computing participation frequency for each activity type
3) Assigning the most frequent activity as the primary sport preference

2) Activity Diversity Quantification: To measure the breadth of physical activity engagement, the study calculated an activity diversity index using Shannon entropy:

$$H = -\sum_{i=1}^{n} p_i \log(p_i) \tag{2}$$

where p_i represents the proportion of activity type i. The index was normalized to a 0–1 scale, with higher values indicating more diverse activity patterns.

3) Intensity Profile Construction: Activities were mapped to three intensity levels based on metabolic equivalents:

- **Low intensity**: Other indoor activities
- **Moderate intensity**: Walking/hiking, gymnastics/fitness, water sports, other outdoor activities
- **High intensity**: Jogging/running, cycling/skiing/skating, ball games

The profiling system computed participation counts for each intensity level, creating a comprehensive activity intensity distribution.

B. Nutritional Profiling

The nutritional profile encompassed multiple dimensions of dietary behavior:

1) Macronutrient Composition: Daily macronutrient intake was aggregated to calculate:

- Average daily consumption of energy (kcal), protein, fat, carbohydrates, and fiber
- Macronutrient distribution ratios relative to total energy intake

2) Meal Timing Analysis: Eating occasions were classified into six temporal categories:

- Breakfast: 05:00–10:00
- Morning snack: 10:00–12:00
- Lunch: 12:00–15:00
- Afternoon snack: 15:00–18:00
- Dinner: 18:00–21:00
- Late snack: 21:00–05:00

The analysis computed meal distribution patterns and eating window duration for each participant.

3) Nutrient Density Score (NDS): A comprehensive nutritional quality metric was developed following established dietary assessment principles:

$$NDS = \frac{1}{n}\sum_{i=1}^{n} \min\left(\frac{N_i/E \times 1000}{DV_i}, 1\right) \times 100 \tag{3}$$

where N_i represents nutrient i intake, E is total energy intake (kcal), and DV_i is the daily value for nutrient i. The score ranges from 0–100, with higher values indicating better nutritional quality.

4) Dietary Consistency Assessment: The coefficient of variation (CV) was calculated for NDS and individual nutrients to assess dietary stability:

$$CV = \frac{\sigma}{\mu} \tag{4}$$

Consistency was categorized as:

- Very consistent: CV < 0.2
- Consistent: $0.2 \leq$ CV < 0.4
- Moderate: $0.4 \leq$ CV < 0.6
- Inconsistent: CV ≥ 0.6

C. Sleep Quality Profiling

Sleep profiles were constructed for participants with at least 7 days of sleep data:

1) Linear interpolation addressed missing values while preserving individual patterns
2) Statistical metrics calculated included mean quality, standard deviation, coefficient of variation, and range
3) Quality classifications were assigned:
    - Excellent: mean quality ≥ 4.5
    - Good: $4 \leq$ mean quality < 4.5
    - Fair: $2 \leq$ mean quality < 4
    - Poor: mean quality < 2

D. Profile Integration and Clustering

1) Comprehensive Profile Assembly: Individual component profiles were merged using participant identifiers to create unified user profiles containing:

- Primary sport and activity diversity metrics
- Activity intensity distributions
- Macronutrient intake patterns and ratios
- Meal timing characteristics
- Nutrient density scores and consistency metrics
- Sleep quality indicators

2) Behavioral Clustering: To identify distinct lifestyle patterns, the study employed unsupervised learning techniques:

1) Categorical variables were encoded numerically
2) Multiple clustering algorithms were evaluated: K-means, hierarchical clustering, Gaussian mixture models, and DBSCAN
3) Model performance was assessed using:
    - Silhouette coefficient
    - Calinski-Harabasz index
    - Davies-Bouldin index
4) K-means clustering emerged as the optimal approach based on evaluation metrics

Furthermore, Principal Component Analysis is performed if the clustering clearly captured the user behaviours and the result shows the kmeans clustering performance is somehow better as shown in Fig 10.

V. Cluster Analysis and Behavioral Characterization

Following user profile creation, the study conducted cluster analysis to identify distinct behavioral patterns and validate lifestyle typologies within the study population.

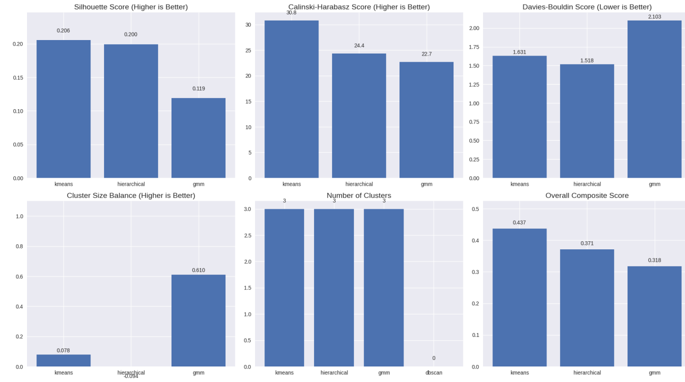

Fig. 9: Clustering Models Evaluation

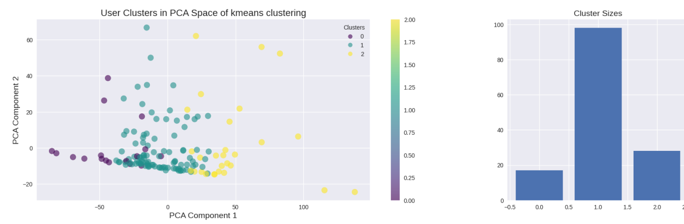

Fig. 10: KMeans Clustering Principal Component Analysis Evaluation

A. Cluster Selection and Validation

1) Clustering Algorithm Evaluation: The analysis evaluated four clustering algorithms: K-means clustering, Hierarchical clustering (Ward's linkage), Gaussian Mixture Models (GMM), and Density-Based Spatial Clustering (DBSCAN). Performance assessment utilized three complementary metrics: Silhouette coefficient (cluster cohesion and separation), Calinski-Harabasz index (between-cluster to within-cluster variance ratio), and Davies-Bouldin index (average similarity between clusters). K-means clustering demonstrated superior performance across all metrics and was selected for final implementation.

B. Physical Activity Characterization

1) Activity Volume and Diversity: The analysis revealed significant differences in activity patterns across clusters ($F = 3.759$, $p = 0.026$. Table II summarizes the activity volume and diversity metrics.

TABLE II: Physical Activity Volume and Diversity Metrics by Cluster

Cluster	Total Activities			Diversity Index		
	Mean	Median	SD	Mean	Median	SD
0	6.9	0.0	14.2	0.034	0.0	0.075
1	7.6	3.0	10.7	0.104	0.0	0.179
2	14.8	7.5	17.2	0.107	0.0	0.171

2) Intensity Distribution: Activity intensity profiles exhibited distinct patterns across clusters, as shown in Table III.

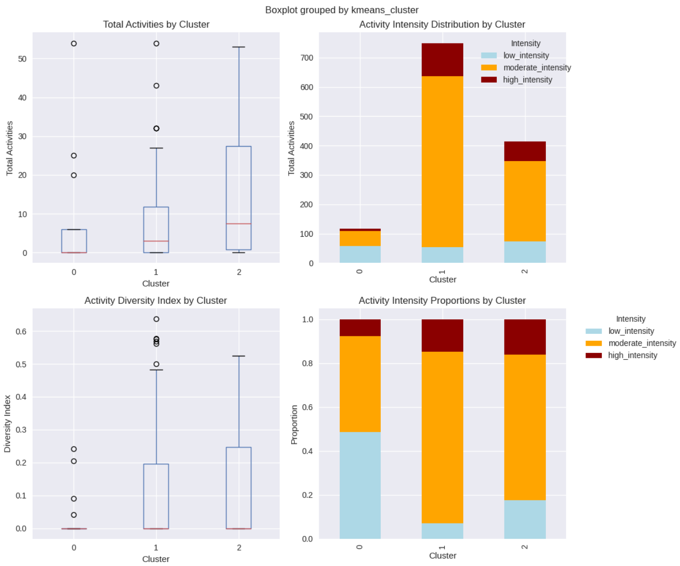

Fig. 11: Activity Distribution by Cluster

TABLE III: Activity Intensity Distribution by Cluster (%)

Cluster	Low Intensity	Moderate Intensity	High Intensity
0	48.7	43.6	7.7
1	7.1	78.1	14.8
2	17.6	66.2	16.2

C. Nutritional Pattern Analysis

1) Mean Nutrition Score Characterization by Cluster: Figure 12 shows daily mean nutrition density scores (NDS) by cluster with shaded confidence bands. Cluster 2 consistently achieves the highest scores (≈76–81) and the narrowest band, indicating stable, higher-quality diets. Cluster 1 remains intermediate (≈73-77), while Cluster 0 is lowest (≈68–75) with the widest band, reflecting greater day-to-day variability. The relative ordering is stable across the period, with only mild increases around early December.

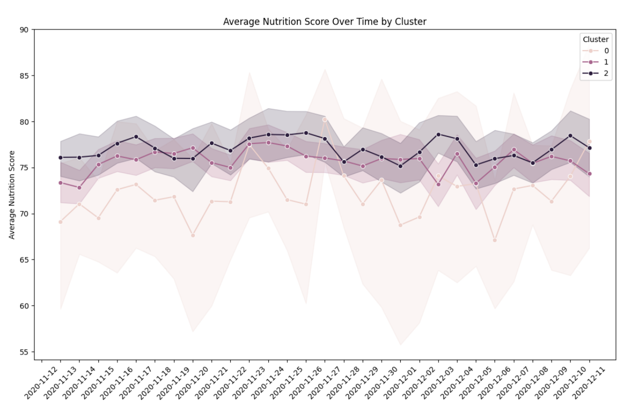

Fig. 12: Average Nutrition Score Over Time by Cluster

2) Diet Quality and Energy Intake: Statistical Nutrient Density Score (NDS) analysis revealed significant inter-cluster differences ($F = 3.749$, $p = 0.026$). Table IV presents the nutritional quality and energy intake patterns.

TABLE IV: Nutritional Quality and Energy Intake Metrics by Cluster

Cluster	NDS		NDS CV	Energy (kcal)		
	Mean	SD		Mean	Median	SD
0	72.7	7.87	0.864	20.0	19.0	7.4
1	75.7	4.81	0.414	35.0	35.0	8.1
2	77.0	4.42	0.237	62.0	58.0	13.3

D. Temporal Behavior Patterns

Time-of-day analysis revealed distinct temporal signatures across clusters, with Cluster 2 demonstrating significantly higher activity counts across all time periods, particularly during morning, lunch, and midnight hours. This pattern suggests more engaged behavioral reporting throughout the day.

E. Cluster Profiles Summary

TABLE V: Summary of Cluster Characteristics

Characteristic	Cluster 0 Low Engage	Cluster 1 Moderate	Cluster 2 High Engage
Activity Level	Minimal (6.9)	Moderate (7.6)	Highest (14.8)
Intensity Profile	Low-dom. (48.7%)	Mod.-dom. (78.1%)	Balanced
NDS Score	72.7	75.7	77.0
Energy Intake	20 kcal	35 kcal	62 kcal
Diet Consistency	Inconsist. (CV=0.864)	Moderate (CV=0.414)	Consistent (CV=0.237)
Macronutrient	Carb-dom.	Balanced	Fat-enriched

F. Longitudinal Behavioral Trends

1) Physical Activity Temporal Patterns: Time series analysis of daily sports activities revealed distinct temporal patterns across clusters (Figure 13). Notable observations include:

- **Walking & Hiking**: Cluster 2 showed sustained higher engagement throughout the study period, while Clusters 0 and 1 demonstrated minimal and declining participation after initial weeks.
- **High-Intensity Activities**: Jogging/running and cycling showed sporadic engagement across all clusters, with Cluster 2 maintaining slightly higher consistency.
- **Team Sports**: Ball games exhibited pronounced peaks in Cluster 1 during mid-December, suggesting seasonal or event-driven participation.
- **Indoor Activities**: Gymnastics/fitness showed gradual increase across all clusters from November to December, with Cluster 2 maintaining highest levels.

2) Nutritional Intake Temporal Dynamics: Longitudinal analysis of nutrient intake patterns revealed consistent hierarchical differences between clusters maintained throughout the study period (Figure 14):

Key temporal trends observed:

- **Consistent Hierarchy**: Cluster 2 > Cluster 1 > Cluster 0 maintained across all nutrients throughout the study

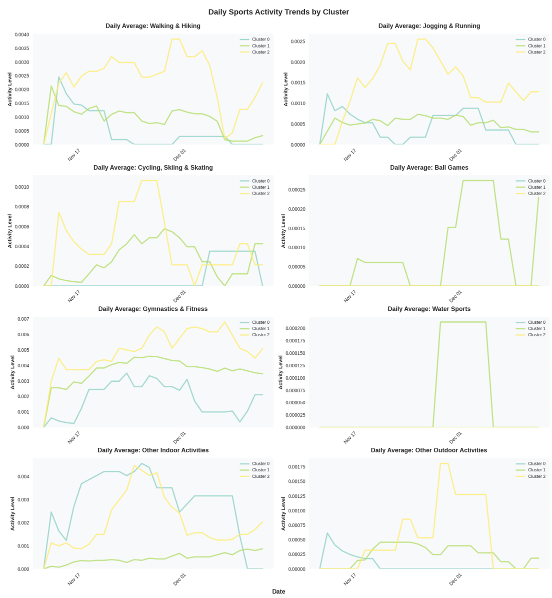

Fig. 13: Daily Sports Activity Trends by Cluster (7-day rolling average)

TABLE VI: Average Daily Nutrient Intake Progression by Study Phase

Nutrient	**Early Phase (Nov)**			**Late Phase (Dec)**		
	C0	C1	C2	C0	C1	C2
Energy (kcal)	18	32	55	20	40	80
Protein (g)	0.5	1.0	1.5	0.6	1.2	2.5
Fat (g)	0.5	0.9	1.5	0.6	1.3	2.8
Carbs (g)	2.5	4.0	6.0	2.8	5.0	9.0
Fiber (g)	0.3	0.6	1.0	0.4	0.8	3.0

- **Temporal Stability**: After initial stabilization period (first week), intake patterns remained relatively constant within clusters
- **Alcohol Consumption**: Notable spike in Cluster 0 during mid-December, suggesting event-related consumption
- **Hydration Patterns**: Water intake showed parallel trends across clusters with gradual increase over time

3) Micronutrient Temporal Patterns: Vitamin and mineral intake analysis revealed proportional scaling with overall dietary intake (Figure **??**):

- **Fat-Soluble Vitamins**: Vitamins A and retinol showed strongest differentiation between clusters, correlating with fat intake patterns
- **Water-Soluble Vitamins**: B vitamins and vitamin C maintained stable ratios across clusters despite absolute differences
- **Minerals**: Calcium, iron, and zinc intake tracked closely with protein consumption patterns
- **Temporal Consistency**: Micronutrient profiles remained stable after initial two-week adaptation period

These longitudinal patterns reinforce the engagement gradient hypothesis, with Cluster 2 demonstrating sustained higher reporting across all behavioral domains throughout the study duration.

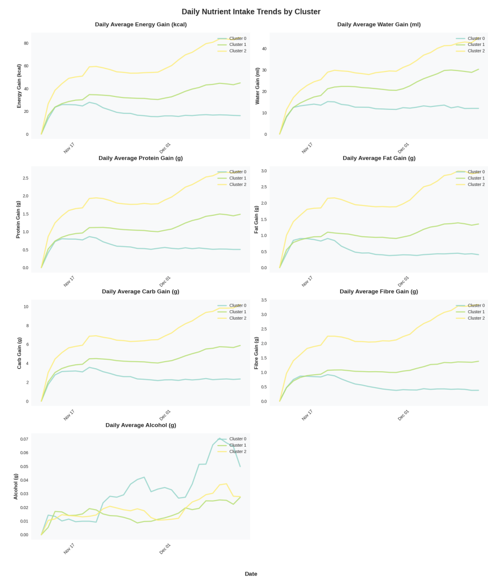

Fig. 14: Daily Nutrient Intake Trends by Cluster (7-day rolling average)

G. Interpretation

The cluster analysis reveals three distinct health behavior profiles with clear hierarchical relationships. Cluster 2 represents a "good health profile" characterized by highest physical activity levels (14.8 activities), superior nutritional quality (NDS = 77.0), and most consistent dietary patterns (CV = 0.237). Cluster 1 demonstrates a "moderate health profile" with balanced activity engagement (7.6 activities), moderate nutritional quality (NDS = 75.7), and intermediate consistency. Cluster 0 exhibits a "poor health profile" with minimal physical activity (6.9 activities), lowest nutritional quality (NDS = 72.7), and highly inconsistent dietary patterns (CV = 0.864).

The longitudinal analysis confirms these profiles remain stable throughout the study period, with consistent hierarchical relationships (Cluster 2 $>$ Cluster 1 $>$ Cluster 0) maintained across all behavioral domains. This temporal stability suggests cluster membership represents persistent lifestyle patterns rather than transient states.

VI. Statistical Analysis and Results

This section presents comprehensive analyses examining the relationships between physical activity patterns, nutritional behaviors, and sleep quality within the identified behavioral clusters.

A. Mixed-Effects Modeling of Activity-Sleep Dynamics

1) Activity Score Predictors: Mixed linear model analysis (n=3,074 observations, 106 participants) revealed significant predictors of daily activity scores. High-intensity activity participation showed positive association ($\beta = 0.151$, SE = 0.045, $p = 0.001$), while activity diversity demonstrated stronger effects ($\beta = 2.194$, SE = 0.813, $p = 0.007$). The interaction between cluster membership and lagged sleep showed no significant effects across clusters, suggesting sleep-activity relationships are consistent across behavioral phenotypes.

2) Sleep Quality Predictors: Sleep quality modeling revealed strong autoregressive effects ($\beta = 0.416$, SE = 0.018, $p < 0.001$), indicating previous night's sleep strongly predicts subsequent sleep quality. Lagged activity effects were non-significant across all cluster groups, suggesting no direct temporal influence of physical activity on next-day sleep quality.

B. Structural Equation Modeling of Bidirectional Relationships

Cross-lagged panel analysis using structural equation modeling examined bidirectional relationships between activity and sleep across two time waves. The model revealed:

- Negative autoregressive path for activity ($\beta = -0.614$, SE = 0.208, $p = 0.003$), suggesting activity levels fluctuate rather than persist
- Strong positive autoregression for sleep ($\beta = 0.423$, SE = 3.631, $p = 0.907$), though with high standard error
- No significant cross-lagged effects between domains, indicating independence of sleep-activity cycles

Model fit indices indicated adequate fit with random intercept variances of 2.358 (SE = 0.225) for activity and 0.488 (SE = 0.629) for sleep.

C. Between-Group Comparisons by Activity Level

1) Activity-Based Sleep Quality Differences: Participants were categorized into activity groups based on total activity levels. Table VII presents sleep quality distributions across groups.

TABLE VII: Sleep Quality by Activity Level Groups

Activity Level	n	Mean	SD	Median	Min	Max
Low	52	3.640	0.563	3.665	2.190	4.973
Moderate	60	3.921	0.644	3.801	1.912	5.000
High	38	4.502	0.449	4.637	3.276	5.000

Kruskal-Wallis test revealed significant differences ($H = 39.673$, $p < 0.001$). Post-hoc pairwise comparisons using Mann-Whitney U tests showed all group differences were significant: Low vs. Moderate ($U = 1157.0$, $p = 0.019$), Low vs. High ($U = 232.0$, $p < 0.001$), and Moderate vs. High ($U = 540.0$, $p < 0.001$).

D. Multiple Regression Analysis of Sleep Quality Determinants

Four hierarchical regression models examined sleep quality predictors with increasing complexity:

TABLE VIII: Model Comparison for Sleep Quality Prediction

Model	R^2	Adj. R^2	AIC	BIC
Basic (Activity only)	0.306	0.288	258.7	273.9
+ Nutrition	0.316	0.293	258.6	276.8
Full (+ Demographics)	0.334	0.272	270.4	312.9
+ Interactions	0.324	0.266	270.7	310.1

VII. Discussion

The findings from this longitudinal study provide important insights into the temporal dynamics between physical activity and sleep quality among university students. Several key patterns emerged that warrant further discussion.

A. Independence of Sleep-Activity Cycles

The unexpected finding is the absence of significant cross-lagged effects between physical activity and sleep quality domains. Despite theoretical expectations and previous cross-sectional evidence suggesting bidirectional relationships, our structural equation modeling revealed that these behavioral domains operate as independent systems over time. The negative autoregressive coefficient for activity ($\beta = -0.614$) suggests day-to-day fluctuation rather than stability, while sleep quality showed strong positive autoregression ($\beta = 0.416$), indicating persistent patterns.

B. Activity Intensity and Diversity Matter

While temporal causality was not established, the analysis revealed that activity characteristics significantly predict daily activity engagement. Activity diversity showed particularly strong effects ($\beta = 2.194$, $p = 0.007$), suggesting that varied physical activity repertoires may be more important than volume alone. This finding aligns with recent literature emphasizing the benefits of diverse movement patterns over repetitive single-activity focus.

C. Concurrent Association vs. Temporal Causation

The significant differences in sleep quality across activity level groups (Low: 3.64, Moderate: 3.92, High: 4.50) demonstrate clear concurrent associations. However, the absence of lagged effects indicates these relationships are contemporaneous rather than predictive. This distinction has important implications for intervention design, suggesting that sleep and activity improvements may need to be targeted simultaneously rather than expecting cascade effects.

VIII. Conclusion

This longitudinal investigation of 204 Italian university students revealed that while physical activity levels and sleep quality show robust concurrent associations, they do not exhibit significant temporal or causal relationships. The absence of cross-lagged effects between these domains challenges prevailing assumptions about bidirectional sleep-activity interactions and suggests these health behaviors operate as independent systems with potentially shared underlying determinants.

The study's key contributions include: (1) demonstration of domain independence through rigorous temporal analysis, (2) identification of activity diversity as a stronger predictor than volume, and (3) evidence for contemporaneous rather than lagged associations between sleep and activity. These findings suggest that effective health interventions should target both domains simultaneously rather than expecting improvements in one to naturally enhance the other.

Future research should investigate potential mediating factors that may explain the concurrent associations observed, employ objective measurement tools to validate self-reported patterns, and extend the observation period to capture longer-term dynamics. Understanding why sleep and activity co-occur without causal linkage may reveal novel intervention targets for promoting holistic student wellness.

IX. Limitations

This research has some important limitations. The main issue is the large amount of missing data on sleep quality and time diary response, which led to the use of interpolation. This makes the results on temporal dynamics less reliable. Another limitation is that both sleep and activity were mostly self-reported, while sensor data were available and objective, only a smaller group of participants had that data.

In addition, the study period was relatively short, which limits comprehensive studies from long-term and seasonal patterns. Finally, the nutritional data were based on self-reported food categories that were converted into nutrient values, which may not fully capture real intake.

References

[1] B. A. Dolezal, E. V. Neufeld, D. M. Boland, J. L. Martin, and C. B. Cooper, "Interrelationship between Sleep and Exercise: A Systematic Review," *Advances in Preventive Medicine*, vol. 2017, no. 1, p. 1364387, 2017, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1155/2017/1364387. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1155/2017/1364387

[2] M. A. Kredlow, M. C. Capozzoli, B. A. Hearon, A. W. Calkins, and M. W. Otto, "The effects of physical activity on sleep: a meta-analytic review," *Journal of Behavioral Medicine*, vol. 38, no. 3, pp. 427–449, Jun. 2015.

[3] S. Lee, C. G. Park, and K. Saban, "Temporal Dynamics Between Sleep and Physical Activity After Cardiac Surgery: A Cross-Lagged Panel Model," *Western Journal of Nursing Research*, vol. 47, no. 9, pp. 810–817, Sep. 2025, publisher: SAGE Publications Inc. [Online]. Available: https://doi.org/10.1177/01939459251344565

[4] J. Marquez, M. Panayiotou, R. Farzinnia, Q. Cheng, and N. Humphrey, "Longitudinal relationships across sleep, physical activity, and mental wellbeing in early-to-mid-adolescence: a developmental cascades investigation," *Quality of Life Research*, vol. 34, no. 5, pp. 1233–1250, 2025. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC12064606/

[5] E. L. Hamaker, R. M. Kuiper, and R. P. P. P. Grasman, "A critique of the cross-lagged panel model," *Psychological Methods*, vol. 20, no. 1, pp. 102–116, Mar. 2015.

[6] I. R. White, P. Royston, and A. M. Wood, "Multiple imputation using chained equations: Issues and guidance for practice," *Statistics in Medicine*, vol. 30, no. 4, pp. 377–399, Feb. 2011.

[7] S. Alismail, C. P. Tribby, J.-A. Yang, D. D. Sears, N. Letellier, T. Benmarhnia, and M. M. Jankowska, "Daily sleep and physical activity from accelerometry in adults: Temporal associations and lag effects," *Sleep Health*, vol. 11, no. 2, pp. 234–240, Apr. 2025.

[8] A. Jover Martínez, L. H. J. M. Lemmens, E. I. Fried, and A. Roefs, "Developing a transdiagnostic Ecological Momentary Assessment protocol for psychopathology," *International Journal of Methods in Psychiatric Research*, vol. 33, no. 3, p. e2028, Jul. 2024. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11258473/