How Digital and Lifestyle Behaviors Shape Wellness

Kexin Fang, Lihong Gao, Shirley Shen, Shaibah Raiyan

Why This Topic

Motivation

Problem
Screen time often treated as harmful

Gap
Different digital behaviors have different effects

Goal
Identify which behaviors matter for wellness and stress

Data Overview

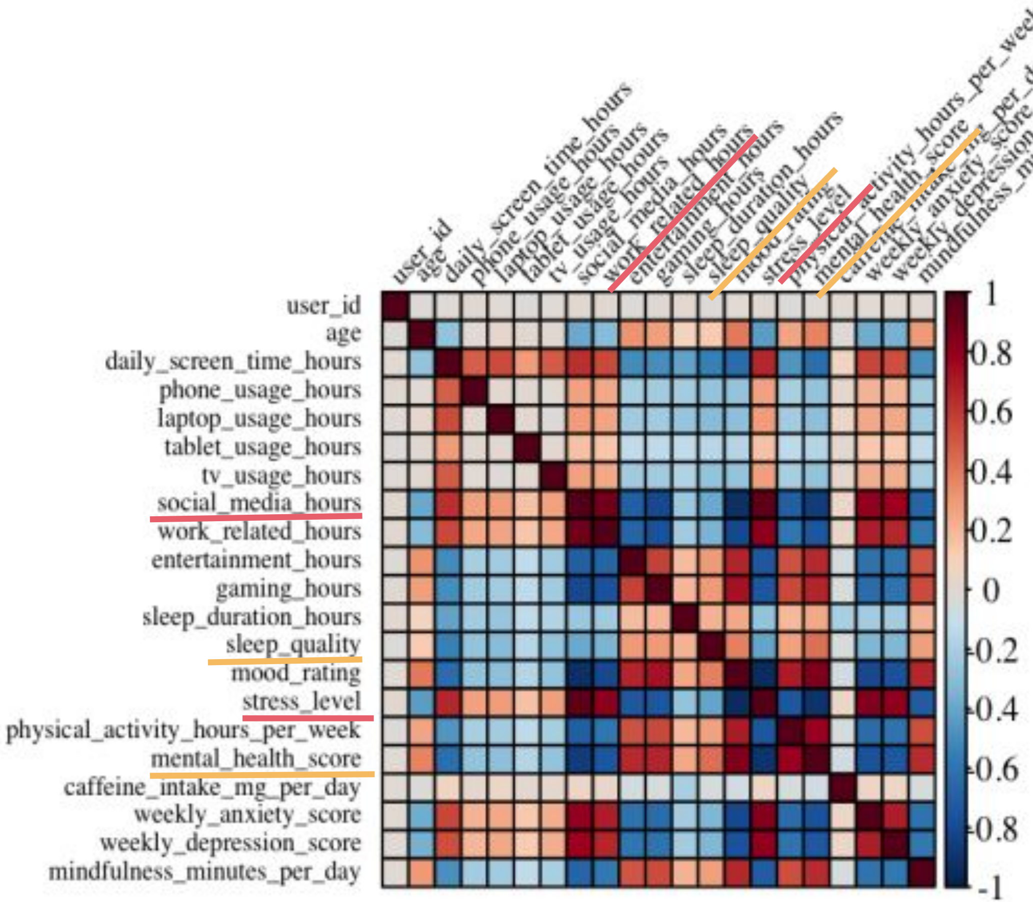
Dataset: 5,000 survey responses

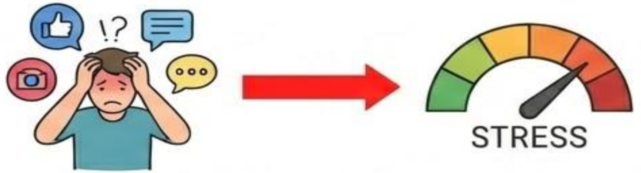
Digital
Social media, gaming, screen hours

Lifestyle
Sleep quality, activity, caffeine

Wellness
Mental health score, stress level

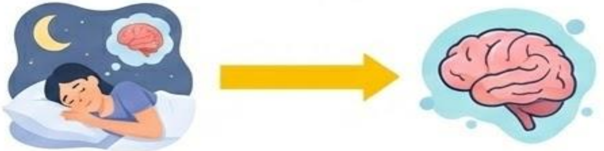
Correlation Matrix





Social media ↔ Stress

Strong Positive Correlation



Sleep quality ↔ Mental health

Strong Positive Correlation

Used to identify key predictors and avoid multicollinearity

Research Question 1: How are digital usage and lifestyle related to mental health scores and stress levels?

We ran **two Multiple Linear Regression (OLS)** models:

1. **Model 1: Predicts mental health score using lifestyle + digital behavior variables.**

2. **Model 2 : Predicts stress level using different screen-time categories (social media, work-related, etc.).**

Model fit / significance

- **$R^2 = 0.8225$ model explains about 82% of variation in mental health scores.**
- **$p < 2.2e{-}16$. Shows overall model is significant**

```
> summary(lm_model1)

Call:
lm(formula = mental_health_score ~ age + entertainment_hours +
    gaming_hours + sleep_duration_hours + sleep_quality + mood_rating +
    physical_activity_hours_per_week + mindfulness_minutes_per_day,
    data = clean_data)

Residuals:
     Min       1Q   Median       3Q      Max
-21.7925  -3.7467   0.0336   3.8034  22.5391

Coefficients:
                                  Estimate Std. Error t value Pr(>|t|)
(Intercept)                      31.563160   1.142846  27.618  < 2e-16 ***
age                               0.026667   0.004973   5.363 8.56e-08 ***
entertainment_hours               1.506144   0.164858   9.136  < 2e-16 ***
gaming_hours                      0.936562   0.168969   5.543 3.13e-08 ***
sleep_duration_hours              0.381535   0.150177   2.541   0.0111 *
sleep_quality                     2.228866   0.128923  17.288  < 2e-16 ***
mood_rating                       1.952589   0.057625  33.884  < 2e-16 ***
physical_activity_hours_per_week  1.957856   0.048746  40.164  < 2e-16 ***
mindfulness_minutes_per_day       0.129696   0.013373   9.699  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 5.521 on 4991 degrees of freedom
Multiple R-squared:  0.8225,	Adjusted R-squared:  0.8222
F-statistic:  2892 on 8 and 4991 DF,  p-value: < 2.2e-16
```

Model 1: Key Drivers of Mental Health Scores

SLEEP QUALITY

SLEEP QUALITY

Better Sleep Quality
=
Higher Mental
Health Scores

PHYSICAL ACTIVITY (hours/week)

PHYSICAL ACTIVITY
(hours/week)

More Activity is
Strongly Linked to
Higher Mental
Health Scores

RECREATIONAL SCREEN TIME

MODERATE GAMING

May act as **Recovery/Coping,** not harm.

SOCIAL MEDIA

Dominant Driver of Stress in this Model.

Model 2 – Predicts Stress Level

🔴**Biggest stressor: social media hours (largest positive coefficient).**

+1 hour of social media is associated with about **+2.7 higher stress units** (largest coefficient

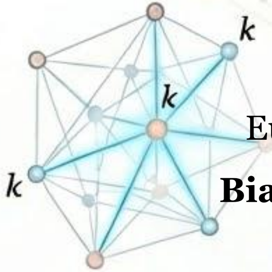
🟡**Work-related use slightly reduces stress mostly depends on the structure and productivity.**

Model fit / significance

- **Adj. R^2 = 0.9256 → explains about 93% of stress variation.**
- **Overall model is significant: $p < 2.2e{-16}$.**

```
> summary(lm_model2)

Call:
lm(formula = stress_level ~ gaming_hours + daily_screen_time_hours +
    social_media_hours + work_related_hours, data = clean_data)

Residuals:
     Min       1Q   Median       3Q      Max
-2.48866 -0.58508 -0.06799  0.53235  2.23906

Coefficients:
                         Estimate Std. Error t value Pr(>|t|)
(Intercept)             -1.398248   0.106185 -13.168  < 2e-16 ***
gaming_hours             0.094403   0.027479   3.436 0.000596 ***
daily_screen_time_hours  0.148974   0.007939  18.766  < 2e-16 ***
social_media_hours       2.696408   0.025958 103.876  < 2e-16 ***
work_related_hours      -0.778843   0.032905 -23.670  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.7949 on 4995 degrees of freedom
Multiple R-squared:  0.9257,	Adjusted R-squared:  0.9256
F-statistic: 1.556e+04 on 4 and 4995 DF,  p-value: < 2.2e-16
```

Can we predict stress level using digital behaviors?



KNN — Classification — Decision Tree

KNN

Euclidean distance

Bias VS. Variance

Initially k=5

Classification

High VS. Low Stress

80% Train / 20% Test

Standardized data

Decision Tree

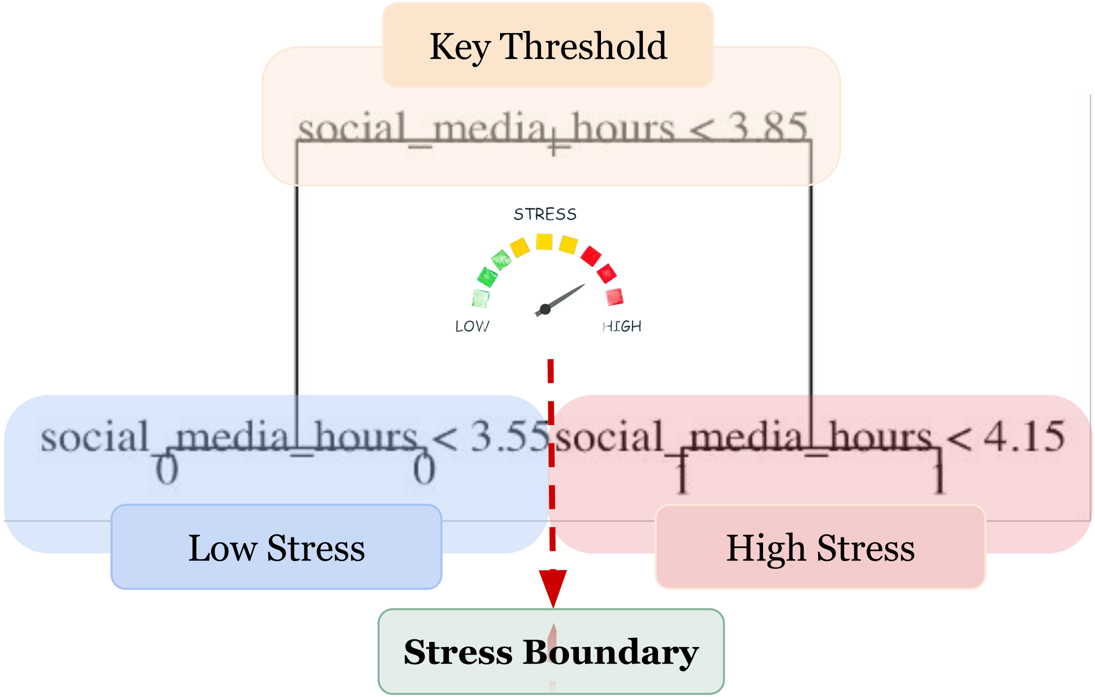
Classification tree

Minimize node impurity

KNN		Decision Tree
92.5%	**Test Accuracy**	96.6%
7.5%	**Error Rate**	3.4%
N/A (Distance-based)	**Strongest Predictive**	*social_media_hours*

Social Media Use Above 3.85 Hours is a Strong Predictor of High Stress

Data reveals a pronounced threshold effect where stress risk sharply increases after 3.85 hours.



From Effects to Personas:
Who Are the Users Behind the Patterns?

What we know

- Digital and lifestyle behaviors show **uneven effects** on wellness
- Social media stands out as a key stress-related behavior

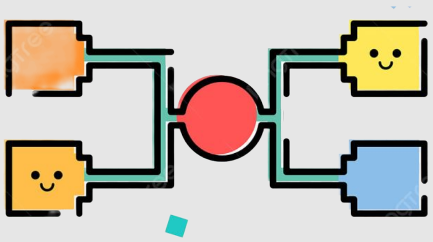
supervised

RQ1 - RQ2

What's missing

Results focus on **variables, not users**

We don't know **how behaviors merge in real life**



What we do next

- **Group** users by **digital usage behavior** patterns
- Identify **natural personas** without predefined outcomes

unsupervised

RQ3

Key Digital Usage Behaviors Used for Clustering

Overall Screen Exposure

- *daily_screen_time_hours*
- Overall level of daily digital engagement

Social Media Intensity

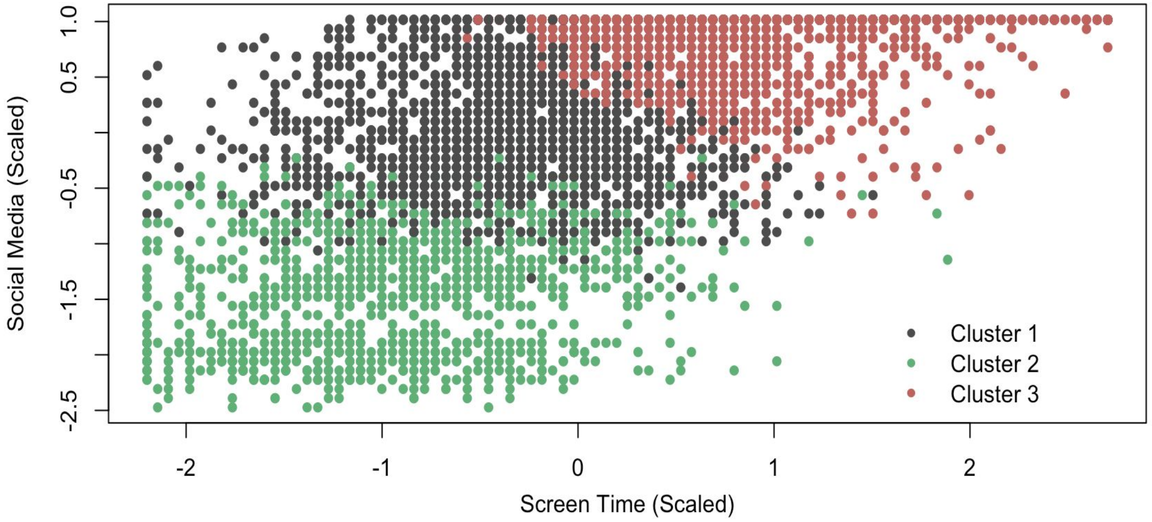
- *social_media_hours*
- Socially driven screen use and online interaction

Entertainment-Oriented Usage

- *gaming_hours*
- Entertainment-focused and recreational screen use

These dimensions describe how users allocate their digital time.
Different combinations of these behaviors naturally form distinct usage patterns.

Beyond Screen Time: Three Digital Usage Personas



KEY TAKEAWAY:

Digital engagement is defined by **modality** (Social vs. Gaming) rather than just **intensity**, distinguishing specific 'Gamer' niches from 'Social-driven' usage.

Cluster 1: Moderate Users

- **The Average Profile**
- Balanced usage; no extreme behaviors.

Daily Screen Time	Social Media	Gaming
4.3h	3.3h	1.5h

Cluster 2: Gamer Niche

- **Low Engagement, Specific Interest**
- Lowest total screen time, but highest gaming

Daily Screen Time	Social Media	Gaming
3.4h	1.6h	**2.5h**

Cluster 3: Heavy Social Users

- **The High-Risk Group**
- Driven by intense Social Media
- Correlates with high stress levels.

Daily Screen Time	Social Media	Gaming
6.6h	**4.3h**	1.1h

Key Findings: The Tipping Point

🔴 3.85 Hours / Day

🚨 Social Media

Strongest predictor of high stress

↓

High stress predictor

🎮 Gaming

Higher mental health scores

💤 Sleep quality

↓

More important than duration

Conclusion: Relationships to Patterns

Key Implications

🧠 **Core insight**
Digital behavior type > total screen time

🎮 **Main risk driver**
Social media — the strongest stress indicator

🎯 **Actionable takeaway**
Focus on Digital Hygiene to reduce toxic scrolling + allow restorative use

Limitations & Future Direction

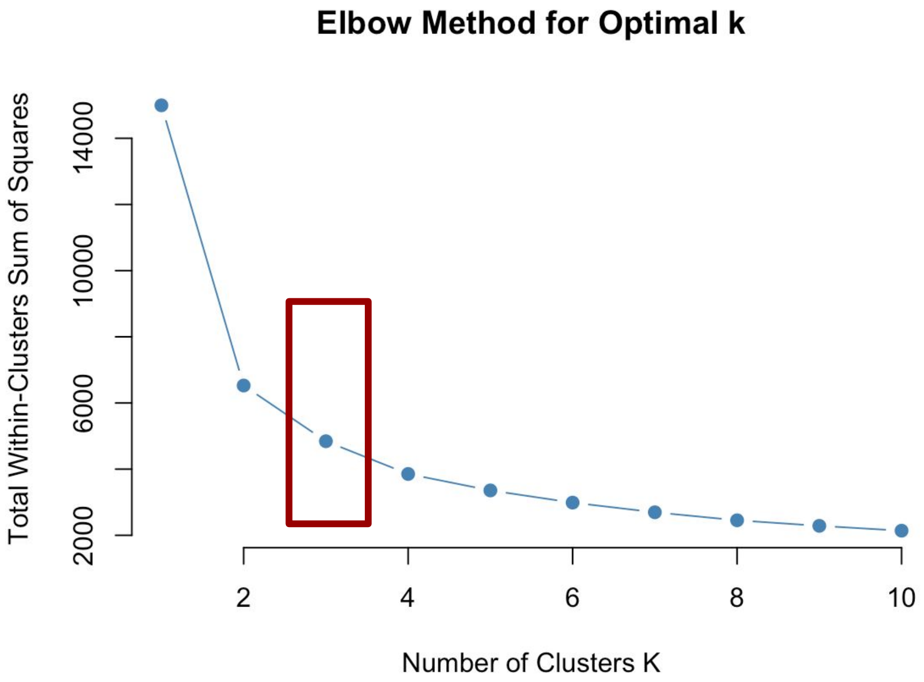
⚠️ **Scope**
Results show associations, not causal effects

⚠️ **Data limitation**
Findings rely on self-reported behavior

🚀 **Next step**
Validate results using objective device logs and APIs

Appendix

RQ3 - Clustering



Appendix

RQ3 - Clustering

Cluster	Daily Screen Time (h)	Social Media (h)	Gaming (h)
Cluster 1	4.3	3.3	1.5
Cluster 2	3.4	1.6	**2.5**
Cluster 3	**6.6**	**4.3**	1.1