

Student Sleep & Social media Behavior

By. Owen Larson & Alyson Her

Purpose

Investigate the multifaceted relationship between sleep and social media behavior to accurately predict and interpret how social media and additional underlying factors affects student well-being

- Reshape communication, learning, and relationships
- Influences various areas of life in complex ways

Data:

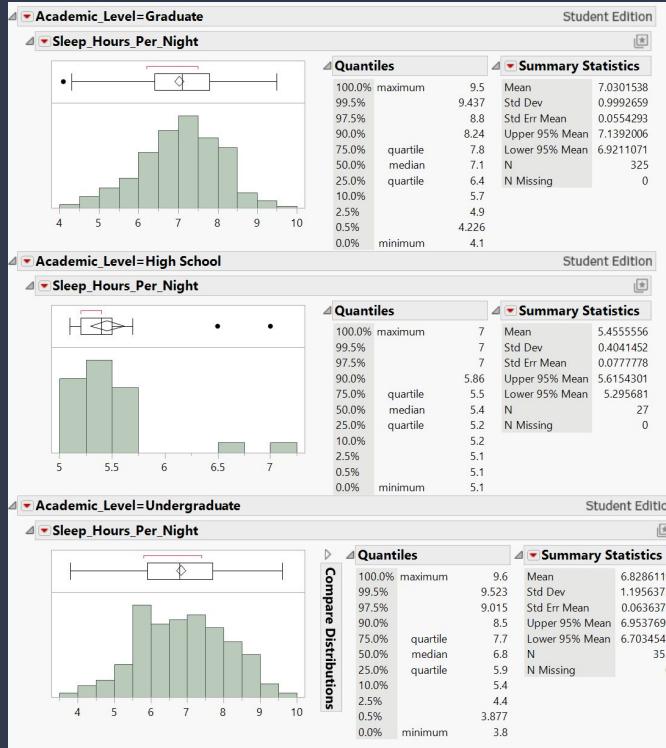
Student_ID	Age	Gender	Academic_Level	Country	Avg_Daily_Usage_Hours	Most_Used_Platform	Affects_Academic_Performance	Sleep_Hours_Per_Night	Mental_Health_Score	Relationship_Status	Conflicts_Over_Social_Media	Addicted_Score
153	153	22	Male	Graduate	4.3	Facebook	Yes	7.3	6	Single	4	8
154	154	21	Female	Graduate	4.8	Instagram	Yes	4.8	5	In Relationship	3	7
155	155	19	Male	Undergraduate	3.5	TikTok	No	8.5	7	Single	2	5
156	156	20	Female	Undergraduate	4.2	Instagram	Yes	7.4	5	In Relationship	4	8
157	157	22	Male	Graduate	4.7	Facebook	Yes	4.7	6	Single	3	7
158	158	21	Female	Graduate	3.4	Instagram	No	8.6	5	In Relationship	2	5
159	159	19	Male	Undergraduate	4.1	TikTok	Yes	7.5	5	Single	4	8
160	160	20	Male	Undergraduate	4.6	Instagram	Yes	4.6	6	In Relationship	3	7
161	161	19	Female	Undergraduate	5.3	Facebook	Yes	6.1	5	Single	3	7
162	162	21	Male	Graduate	4.8	Facebook	No	7.2	5	In Relationship	2	6
163	163	20	Female	Undergraduate	5.5	TikTok	Yes	5.9	6	Single	4	8
164	164	22	Male	Graduate	4.7	Instagram	Yes	6.3	5	In Relationship	3	7
165	165	19	Female	Undergraduate	5.1	Facebook	No	7	7	Single	2	5
166	166	21	Male	Graduate	5.4	TikTok	Yes	6	6	In Relationship	4	8
167	167	20	Female	Undergraduate	4.9	Instagram	Yes	6.4	5	Single	3	7
168	168	22	Male	Graduate	5.2	Facebook	No	7.1	7	In Relationship	2	6
169	169	19	Female	Undergraduate	5.6	TikTok	Yes	5.8	6	Single	4	8
170	170	21	Male	Graduate	4.4	Instagram	Yes	6.5	5	In Relationship	3	7
171	171	20	Female	Undergraduate	5	Facebook	No	7.3	7	Single	2	5
172	172	22	Male	Graduate	5.3	TikTok	Yes	5.7	6	In Relationship	4	8
173	173	19	Female	Undergraduate	4.8	Instagram	Yes	6.6	5	Single	3	7
174	174	21	Male	Graduate	5.1	Facebook	No	7.4	7	In Relationship	2	6
175	175	20	Female	Undergraduate	5.7	TikTok	Yes	5.6	6	Single	4	8
176	176	22	Male	Graduate	4.5	Instagram	Yes	6.7	5	In Relationship	3	7
177	177	19	Female	Undergraduate	4.9	Facebook	No	7.5	7	Single	2	5
178	178	21	Male	Graduate	5.2	TikTok	Yes	5.5	6	In Relationship	4	8
179	179	20	Female	Undergraduate	4.7	Instagram	Yes	6.8	5	Single	3	7
180	180	21	Male	Graduate	5	Facebook	No	7.6	7	In Relationship	2	6
181	181	19	Female	Undergraduate	5.8	TikTok	Yes	5.4	6	Single	4	8
182	182	21	Male	Graduate	4.4	Instagram	Yes	6.9	5	In Relationship	3	7
183	183	20	Female	Undergraduate	4.8	Facebook	No	7.7	7	Single	2	5
184	184	22	Male	Graduate	5.1	TikTok	Yes	5.3	6	In Relationship	4	8
185	185	19	Female	Undergraduate	4.6	Instagram	Yes	7	5	Single	3	7
186	186	21	Male	Graduate	4.9	Facebook	No	7.8	7	In Relationship	2	6
187	187	20	Female	Undergraduate	5.9	TikTok	Yes	5.2	6	Single	4	8
188	188	22	Male	Graduate	4.3	Instagram	Yes	7.1	5	In Relationship	3	7
189	189	19	Female	Undergraduate	4.7	Facebook	No	7.9	7	Single	2	5
190	190	17	Male	Graduate	5	TikTok	Yes	5.1	6	In Relationship	4	8
191	191	20	Female	Undergraduate	4.5	Instagram	Yes	7.3	5	Single	3	7
192	192	22	Male	Graduate	4.8	Facebook	No	8	7	In Relationship	2	6
193	193	19	Female	Undergraduate	6	TikTok	Yes	5	6	Single	4	8
194	194	21	Male	Graduate	4.2	Instagram	Yes	7.3	5	In Relationship	3	7
195	195	20	Female	Undergraduate	4.6	Facebook	No	8.1	7	Single	2	5

- Multi-country online survey
- Sample of students ages 16-25 enrolled in high school, undergraduate, or graduate programs in 2025
- 705 observations in total across 110 countries
- Collected demographic info, measures of social media use, and impact on academic performance, mental health , sleep, and relationships

Data Overview

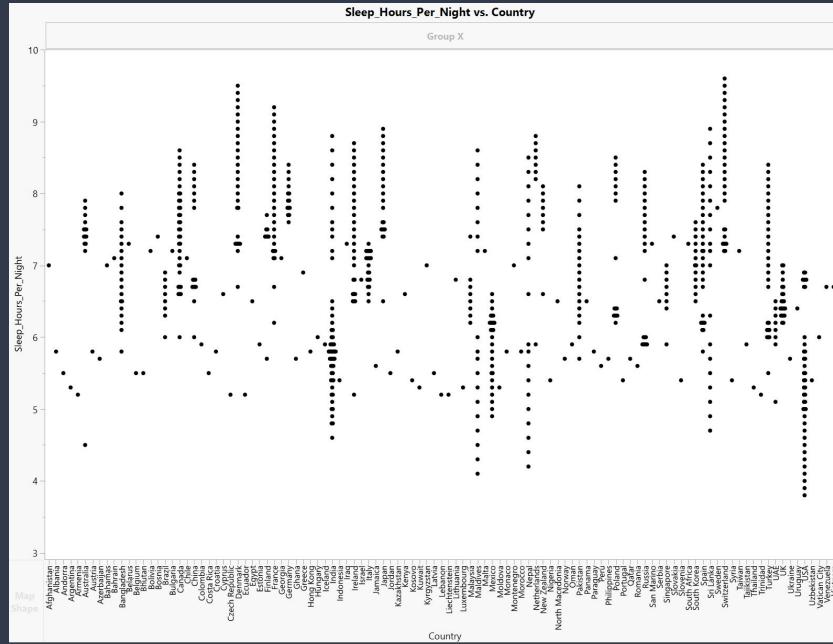
Variable	Description	Type
Sleep_Hours_Per_Night (Y)	Average nightly sleep hours	Continuous
Avg_Daily_Usage_Hours ($x1$)	Average hours per day on social media	Continuous
Addicted_Score($x4$)	Social Media Addiction Score (1 = low to 10 = high)	Continuous
Conflicts_Over_Social_Media ($x3$)	Number of relationship conflicts due to social media	Continuous
Mental_Health_Score ($x2$)	Self-rated mental health (1 = poor to 10 = excellent)	Continuous
Academic_Level	High School / Undergraduate / Graduate	Nominal
Country	Country of residence	Nominal

Data Overview: Sleep vs Academic level



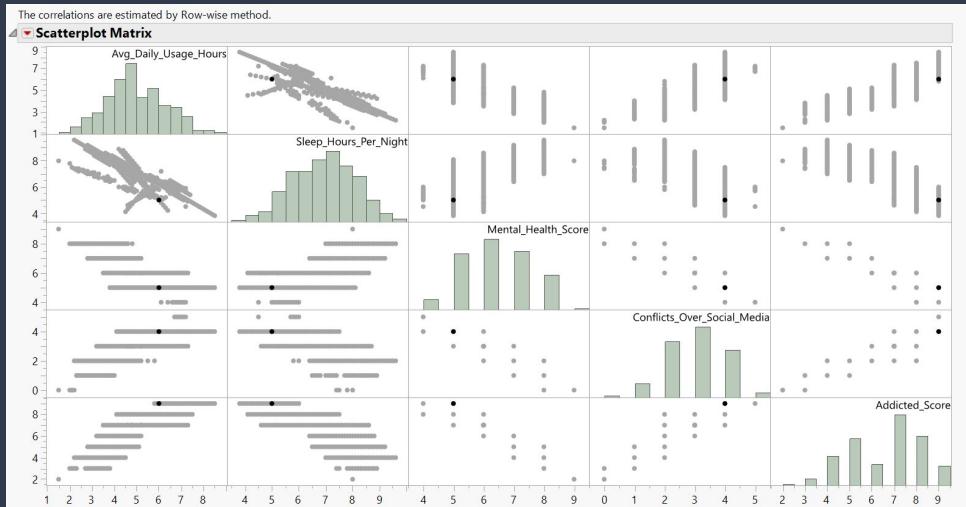
- Sleep vs Graduate:
 - Slightly left skewed
- Sleep vs High school:
 - Extremely right skewed, sleep less than other groups
 - Less observations
- Sleep vs Undergraduate:
 - Slightly right skewed, occasional low hours of sleep

Data Overview: Sleep vs Country



- Sample sizes were not equal for every country
 - African and Caribbean countries
- Reveals similar sleep patterns as academic level
 - Sleep varies but overall matches the median of 6.9 hours of sleep

Multicollinearity Problem



- Sleep: mean= 6.9 hours, max= 9.6 hours, min= 3.8 hours
- Avg Daily Usage Hours: mean= 4.9, max= 8.5 hours, min 1.5, slightly right skewed indicating lower usage hours were more common
- Addicted score: mean= 6.4, max= 9, min=2, left-skewed, occasional high scores
- # of conflicts: mean= 2.8, max= 5, min=0, slightly left-skewed
- Mental Health: mean = 6.2, max=9, min=4, slightly right skewed

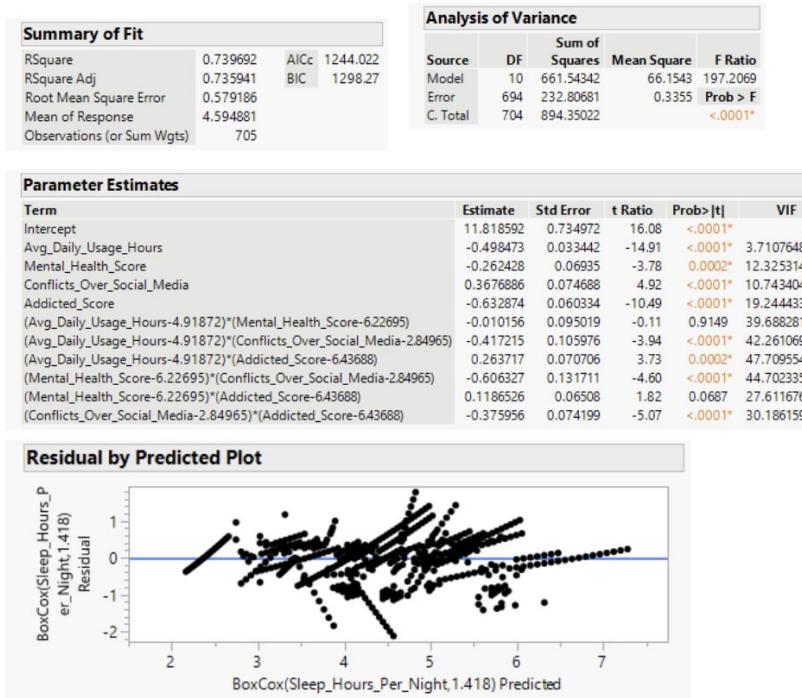
Correlations

	Avg_Daily_Usage_Hours	Sleep_Hours_Per_Night	Mental_Health_Score	Conflicts_Over_Social_Media	Addicted_Score
Avg_Daily_Usage_Hours	1.000	-0.7966	-0.8010	0.8045	0.8324
Sleep_Hours_Per_Night	-0.7966	1.0000	0.7102	-0.6800	-0.7668
Mental_Health_Score	-0.8010	0.7102	1.0000	-0.8934	-0.9451
Conflicts_Over_Social_Media	0.8045	-0.6800	-0.8934	1.0000	0.9336
Addicted_Score	0.8324	-0.7668	-0.9451	0.9336	1.0000

The correlations are estimated by Row-wise method.

- Strong linear correlations indicates multicollinearity
 - Make interpretation difficult

Model 1: Box-Cox with Numerical predictors & 2-way interactions



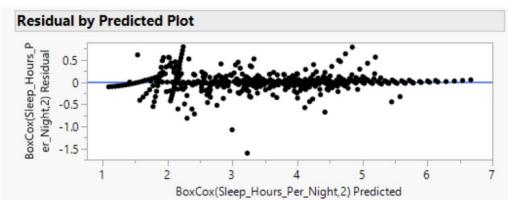
- P-value > 0.05, reject the null, can predict sleep hours
 - Low predicting power, explains 73% of the variance
 - Residual plot shows fanning (heteroscedasticity)
- Hlgh VIF values indicate multicollinearity
- Cannot accurately explain coefficients
- Not a good fit, more predictors we could add to improve it

Model 2: Box-cox of Numerical, Categorical predictors, and 2-way interactions

Summary of Fit	
RSquare	0.974421
RSquare Adj	0.962169
Root Mean Square Error	0.220995
Mean of Response	3.496526
Observations (or Sum Wgts)	705

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Model	228	885.58371	3.88414	79.5300	<.0001*
Error	476	23.24720	0.04884		
C.Total	704	908.83092			

Parameter Estimates					
	Estimate	Std Error	t Ratio	Prob> t	VIF
Academic_Level[High School] Country[Turkey]	Zerod	0	0	.	0
Academic_Level[High School] Country[UAE]	Zerod	0	0	.	0
Academic_Level[High School] Country[UK]	Zerod	0	0	.	0
Academic_Level[High School] Country[USA]	Zerod	0	0	.	0
Academic_Level[High School] Country[Uruguay]	Zerod	0	0	.	0
Academic_Level[High School] Country[Venezuela]	Zerod	0	0	.	0
Academic_Level[High School] Country[Vietnam]	Zerod	0	0	.	0
Academic_Level[Graduate] (Avg_Daily_Usage_Hours-4.91872)	Biased	-0.243685	0.041588	-5.86	<.0001*
Academic_Level[High School] (Avg_Daily_Usage_Hours-4.91872)	Zerod	0	0	.	0
Academic_Level[Graduate] (Mental_Health_Score-5.22695)	Biased	-5.796796	1.170212	-4.95	<.0001*
Academic_Level[High School] (Mental_Health_Score-5.22695)	Zerod	0	0	.	0
Academic_Level[Graduate] (Conflicts_Over_Social_Media-2.84965)	Biased	-5.416897	1.051843	-5.15	<.0001*
Academic_Level[High School] (Conflicts_Over_Social_Media-2.84965)	Zerod	0	0	.	0
Academic_Level[Graduate] (Addicted_Score-6.43688)	Biased	-0.31082	0.219888	-1.41	0.1582
Academic_Level[High School] (Addicted_Score-6.43688)	Zerod	0	0	.	0
Country[Afghanistan] (Avg_Daily_Usage_Hours-4.91872)	Zerod	0	0	.	0
Country[Albania] (Avg_Daily_Usage_Hours-4.91872)	Zerod	0	0	.	0
Country[Andorra] (Avg_Daily_Usage_Hours-4.91872)	Zerod	0	0	.	0
Country[Argentina] (Avg_Daily_Usage_Hours-4.91872)	Zerod	0	0	.	0



- P-value > 0.05, reject the null, can predict sleep hours
 - High predicting power, explains 98% of the variance
 - Residual plot still shows fanning (heteroscedasticity)
- Extremely high or low VIF values indicate multicollinearity
 - VIF=0, is not possible something is wrong with the model or data itself
- Still cannot accurately explain coefficients
- Proven that Academic level and Country are significant predictors but need a new approach for interpretation

Model 3: Weighted final model with key interaction effects

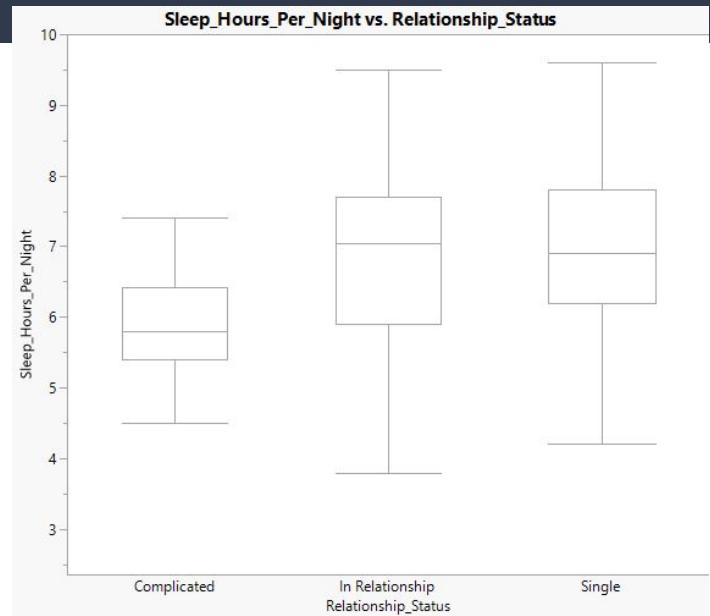
- Refitted using Weighted least Squares
- Additional demographics and categorical variables evaluated using stepwise selection
 - Age
 - Interacted meaningfully with avg daily social media use
 - Academic level
 - Relationship status
 - “complicated” relationships slept less
 - Country
 - Regional groupings to capture cultural differences without overparameterization

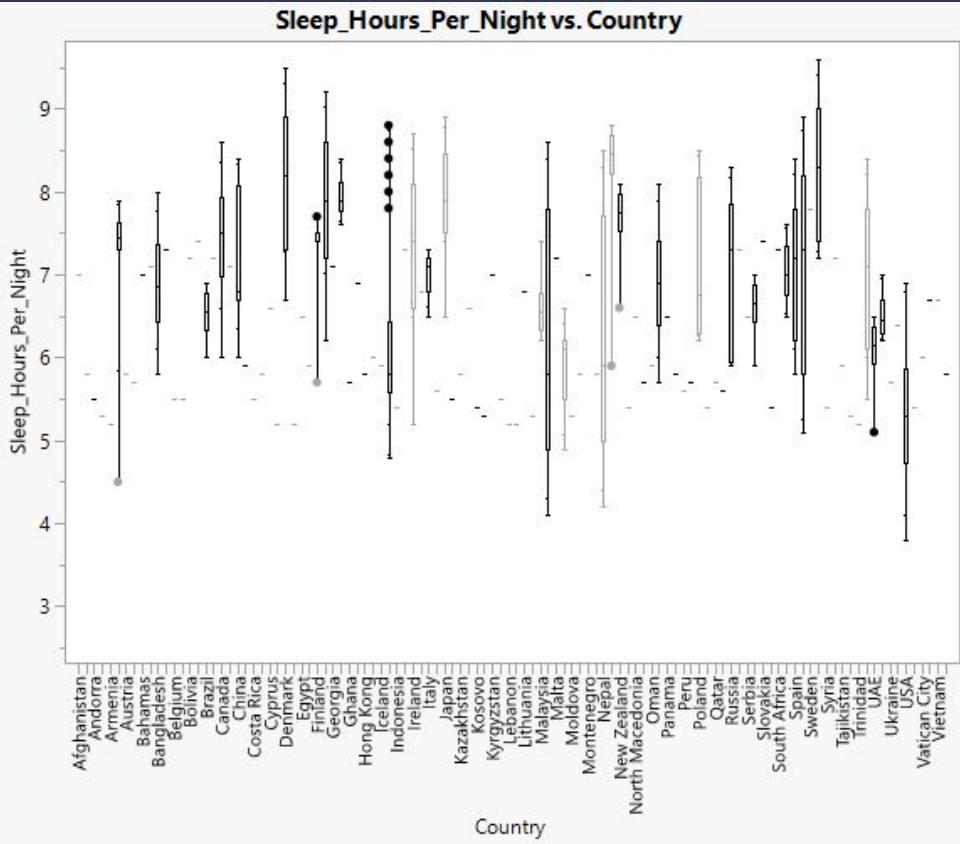
Summary of Fit				Analysis of Variance				
				Source	DF	Sum of Squares	Mean Square	F Ratio
RSquare	0.906018	AICc	-697.088					
RSquare Adj	0.897674	BIC	-439.636					
Root Mean Square Error	0.140044							
Mean of Response	6.709425							
Observations (or Sum Wgts)	104.3459							

	A	B	C	D	E	F
1	Source	1	1	0.9115822	46.4802	<.0001
2	Age	1	1	1.2681454	64.6608	<.0001
3	Avg_Daily_Usage_Hours	1	1	1.0899585	55.5753	<.0001
4	Mental_Health_Score	1	1	2.7598949	140.7228	<.0001
5	Conflicts_Over_Social_Media	1	1	0.1081133	5.5125	0.0192
6	Addicted_Score	1	1	0.8588897	43.7935	<.0001
7	Age*Avg_Daily_Useage_Hours	1	1	1.58564	80.8493	<.0001
8	Mental_Health_Score*Conflicts_Over_Social_Media	1	1	0.0234687	1.1966	0.2744
9	Conflicts_Over_Social_Media*Addicted_Score	1	1	3.5848677	182.7869	<.0001
10	Global Mixed Region (Baseline)	1	1	0.9950239	50.7347	<.0001
11	Global Cluster vs. Southern-Asian Group	1	1	0.0620918	3.166	0.0757

Regions, Academic level, and relationships

- Region (country clusters)
 - Statistically merged using stepwise (lowest BIC) to capture behavioral patterns
 - (graph on next slide)
- Relationship status
 - “Complicated” status showed association with lower sleep than single/in relationship
- Academic level
 - Small main effect, moderate role through interactions.





Global Mixed Region (Baseline)

Global Cluster vs. Southern-Asian Group

Western Global vs. Eastern Eurasian Cluster

Globalized vs. Mixed Developing Regions

Low-Engagement vs. High-Engagement Regions

Low-Connectivity vs. High-Connectivity Regions

U.S. vs. Developing & Transitional Regions

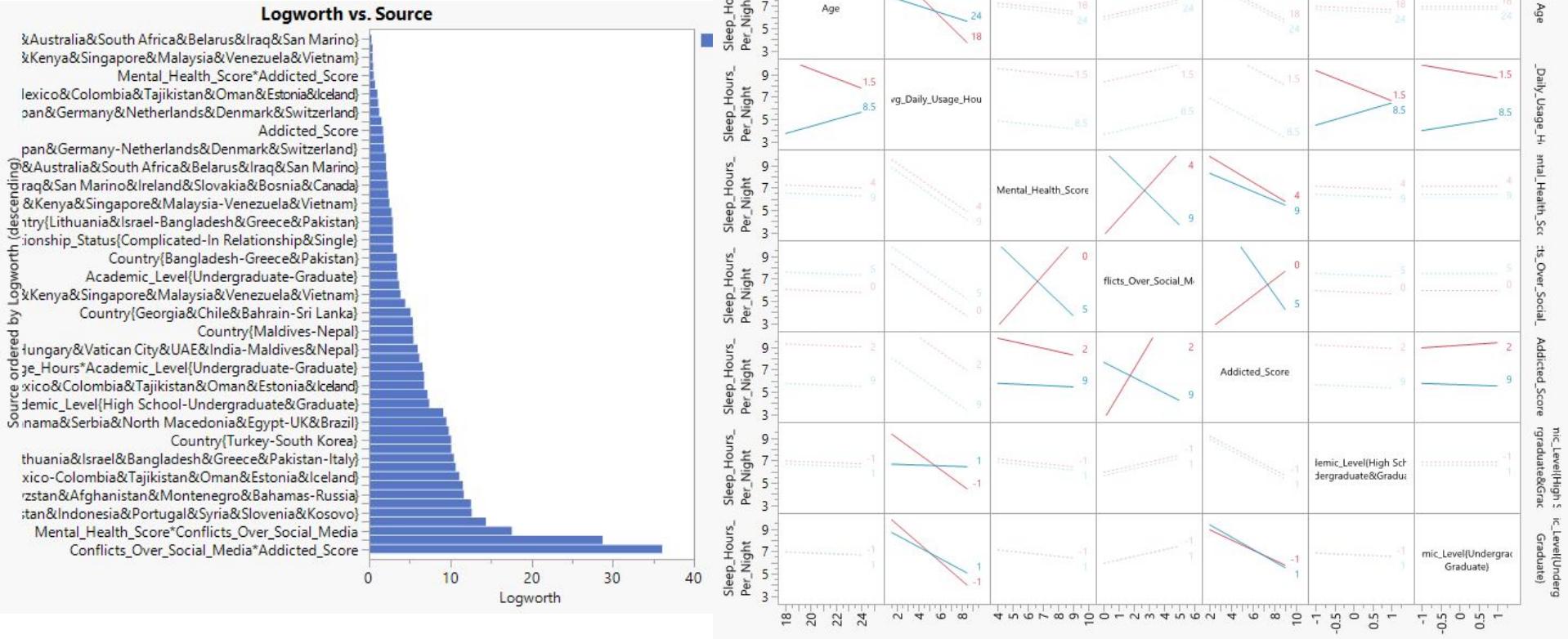
Transitional Eurasia vs. Developed Global Regions

Mexico vs. Smaller Global Regions

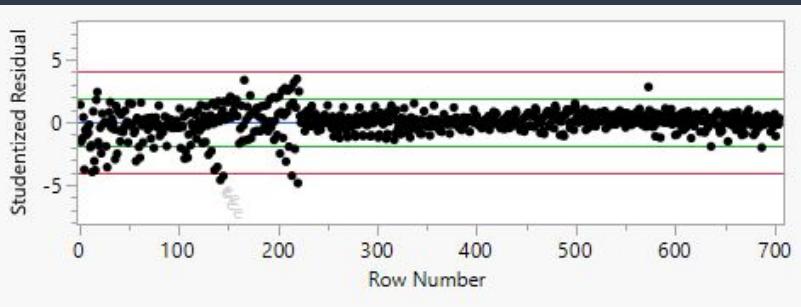
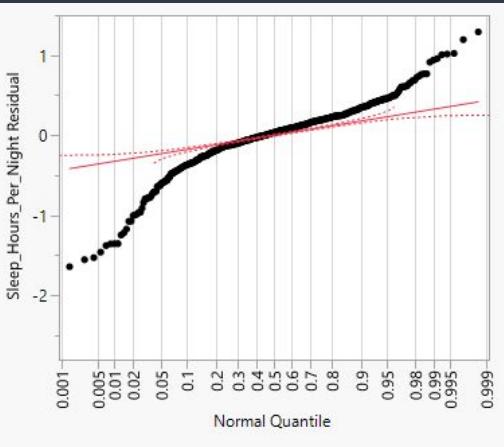
South Asia & Central Europe vs. Urbanized Global Regions

Central Europe vs. South Asia

Key Effects and interactions



Model 3 ~ residuals and validation



- Studentized residual plot
 - Suggest heavier tails and some influential outliers
- QQ Plot
 - Mild tail deviations, consistent with residual distribution
 - Some outliers persist, but residual behaves predictively
 - Residuals roughly centered but show mild tails: non-normality from extreme cases
- Studentized residuals
 - $> \pm 3$ indicate a handful of influential observations
- Cook's D ≈ 0.06 (max)
 - shows limited leverage: model robust

Train Test Validation

Training

Summary of Fit			
RSquare	0.919574	AICc	-510.461
RSquare Adj	0.908851	BIC	-294.954
Root Mean Square Error	0.128854		
Mean of Response	6.70295		
Observations (or Sum Wgts)	68.81338		

Validation

Summary of Fit			
RSquare	0.90726	AICc	-134.171
RSquare Adj	0.883343	BIC	15.12856
Root Mean Square Error	0.156777		
Mean of Response	6.721965		
Observations (or Sum Wgts)	35.53248		

- Small difference between these metrics indicates Model 3 maintains strong predictive accuracy when applied to unseen observations
- Confirms final weighted model is stable and reliable

Conclusion

- Stepwise data-driven modeling identified key behavioral and demographic predictors of sleep duration.
- Average daily usage, social media conflict, and addiction scores were the strongest and most consistent predictors of reduced sleep.
- Mental health amplified these effects, while age moderated them—older participants were less affected by heavy use.
- Regional clusters captured broader cultural similarities in online behavior and sleep habits.
- Weighted Least Squares ($R^2 = 0.90$) corrected heteroscedasticity and improved model fit.

Key Takeaways

- Social and emotional factors surrounding media use were far more influential than usage time alone.
- Conflicts and addiction interactions most strongly predicted sleep loss.
- “Complicated” relationships and regional context further shaped outcomes.
- Findings highlight that social media’s impact on sleep is both behavioral and emotional, varying across cultures and demographics.