

# Project 1: Student Mental Health Analysis

Jadon Chanthavong

9/12/2025

**Dataset link:** <https://www.kaggle.com/code/mdsultanulislamovi/comprehensive-analysis-student-stress-datasets/input?select=StressLevelDataset.csv>

**Introduction to data:** This is a survey of 1100 participants with all being students in a nationwide survey measuring numerous factors such as anxiety level, self-esteem, mental health history, depression, sleep quality, and living conditions.

## **Introduction of problem:**

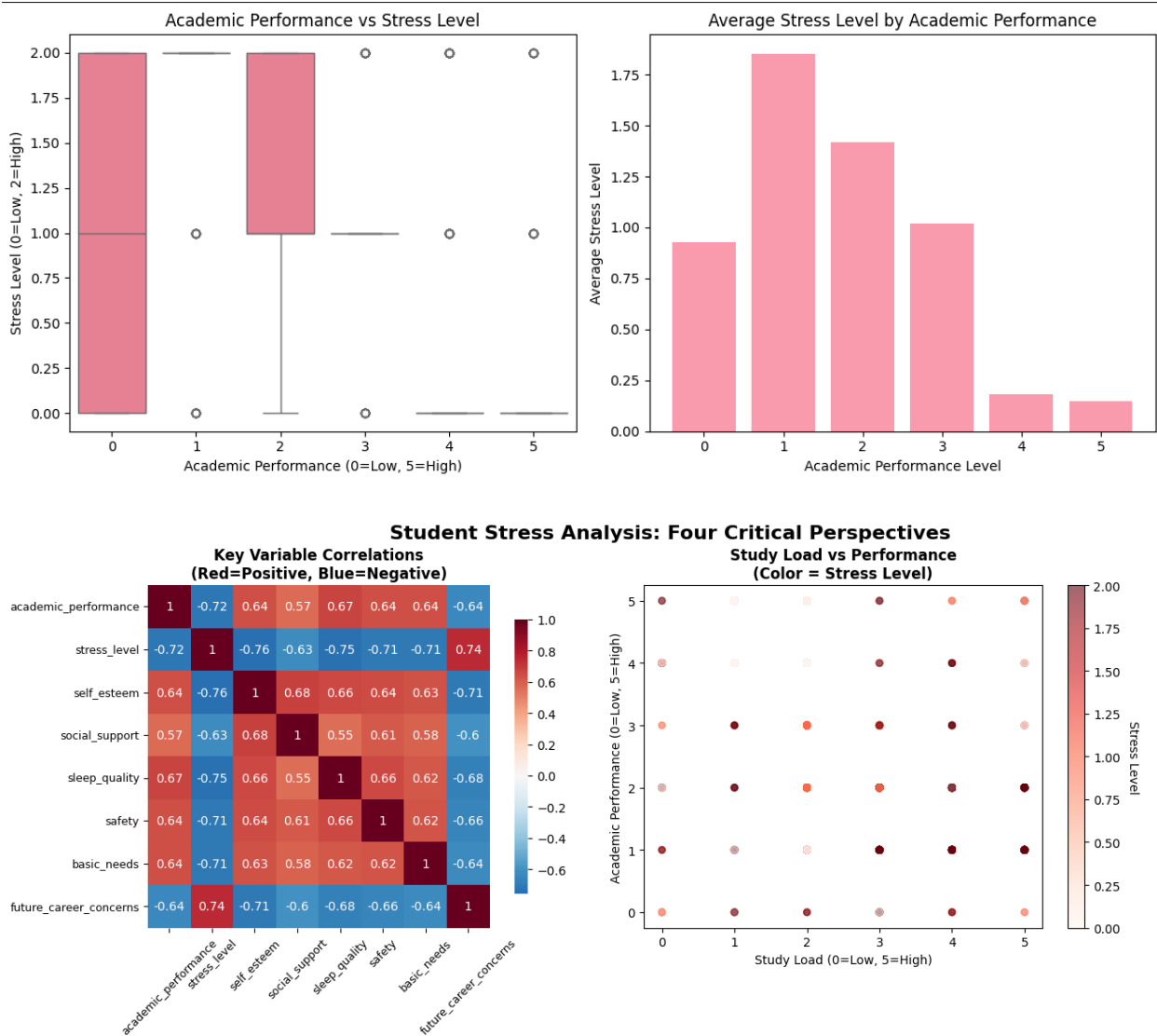
- The questions I want to answer: Are high achievers necessarily more stressed?
- Is there a trade-off between academic success and mental wellbeing, or do some students manage both effectively?
- Does academic pressure create a continuous cycle?
- Do students with higher study loads experience more stress, which then hurts their academic performance, creating even more pressure to do better and it then snowballs?

## **Pre-processing the data:**

Data preprocessing serves as a critical quality assurance step that validates the reliability of analytical findings by identifying potential issues that could compromise analysis validity. Without preprocessing validation, researchers risk building analyses on flawed foundations where missing data patterns could indicate systematic bias, duplicate records could artificially inflate sample sizes, and extreme outliers might represent data entry errors rather than genuine student experiences. The preprocessing assessment of this dataset revealed exceptional cleanliness with no missing values, duplicates, or unrealistic outliers, while confirming that all variables were appropriately scaled integers suitable for correlation analysis and that the balanced distribution of stress levels (33.9% low, 32.5% medium, 33.5% high) could support meaningful comparative analysis across groups. This foundational validation provided confidence that subsequent findings about the relationship between academic performance and

stress reflected genuine patterns rather than data artifacts, enabling reliable conclusions about student wellbeing that could inform educational policy and support systems.

Data visualizations:



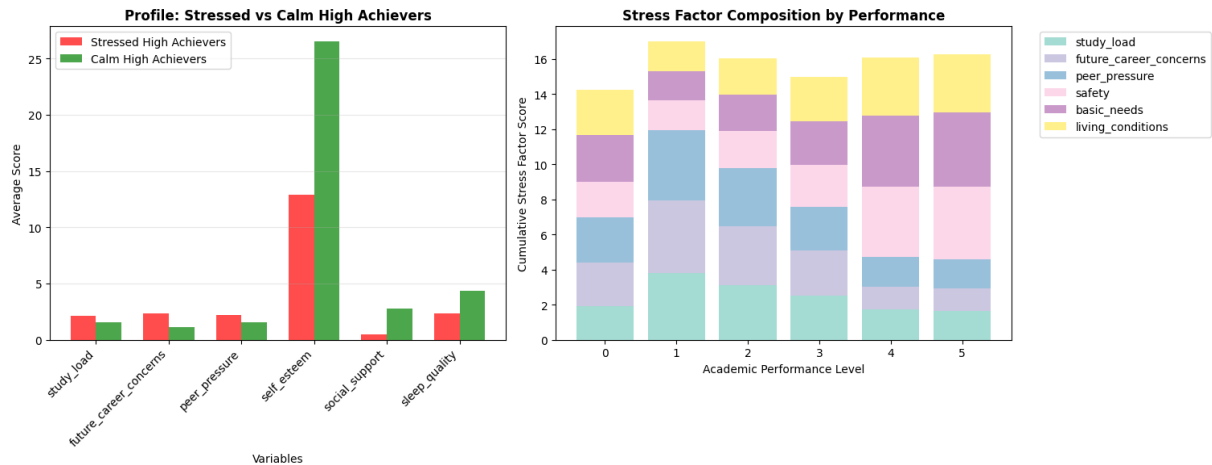
Student Stress Analysis: Four Critical Perspectives

Key Variable Correlations  
(Red=Positive, Blue=Negative)

	academic_performance	stress_level	self_esteem	social_support	sleep_quality	safety	basic_needs	future_career_concerns
academic_performance	1	-0.72	0.64	0.57	0.67	0.64	0.64	-0.64
stress_level	-0.72	1	-0.76	-0.63	-0.75	-0.71	-0.71	0.74
self_esteem	0.64	-0.76	1	0.68	0.66	0.64	0.63	-0.71
social_support	0.57	-0.63	0.68	1	0.55	0.61	0.58	-0.6
sleep_quality	0.67	-0.75	0.66	0.55	1	0.66	0.62	-0.68
safety	0.64	-0.71	0.64	0.61	0.66	1	0.62	-0.66
basic_needs	0.64	-0.71	0.63	0.58	0.62	0.62	1	-0.64
future_career_concerns	-0.64	0.74	-0.71	-0.6	-0.68	-0.66	-0.64	1

Study Load vs Performance  
(Color = Stress Level)

Study Load	Academic Performance	Stress Level
0	0, 1, 2, 3, 4, 5	0.00, 0.25, 0.50, 0.75, 1.00, 1.25
1	0, 1, 2, 3, 4, 5	0.00, 0.25, 0.50, 0.75, 1.00, 1.25
2	0, 1, 2, 3, 4, 5	0.00, 0.25, 0.50, 0.75, 1.00, 1.25
3	0, 1, 2, 3, 4, 5	0.00, 0.25, 0.50, 0.75, 1.00, 1.25
4	0, 1, 2, 3, 4, 5	0.00, 0.25, 0.50, 0.75, 1.00, 1.25
5	0, 1, 2, 3, 4, 5	0.00, 0.25, 0.50, 0.75, 1.00, 1.25



-- **Takeaways:** So to answer my question of whether high performing students are more stressed, the answer, to my surprise, is NO. Actually, what we see from the charts is that, generally, higher performing students actually tend to have lower stress levels compared to their lower performing peers who are moderately to extremely stressed. Now, stress can be caused from an extreme multitude of things. It's not just academic stress itself. If it was..that would make this analysis a lot simpler and easier. What does seem to be a pattern though is that high stress typically equates to low academic performance. However, there are some outliers correlating high-performance with high-stress levels. To answer the question of if some students can manage both effectively, that does seem to be the majority case rather than just "some" students.

## The Stress Ecosystem: Beyond Academic Performance

Rather than finding evidence of a simple academic pressure cycle, the data reveals a complex ecosystem where stress emerges from multiple interconnected factors. The correlation analysis shows that future career concerns, not current academic workload, serve as the strongest predictor of student stress. This challenges the assumption that academic pressure itself creates stress.

The most significant finding is the stark difference in stress composition across performance levels. Low-performing students experience what can be described as "survival stress" - simultaneously grappling with peer pressure, basic needs insecurity, safety concerns, and unstable living conditions alongside their academic struggles. This creates a compounding effect

where multiple life stressors reinforce each other, making academic improvement nearly impossible.

High-performing students, conversely, operate in a fundamentally different stress environment. Having largely resolved basic life stability issues, their stress profile is streamlined and future-focused - primarily concerned with maintaining their success rather than achieving basic security. The small minority of stressed high achievers are distinguished not by academic workload, but by significantly lower self-esteem and higher future anxiety despite their current success.

This reveals three key insights:

- Stress is not distributed randomly - it clusters around academic struggle and life instability
- Academic success appears to provide protection from most stress sources, not just academic ones
- The relationship between effort and stress depends entirely on whether that effort produces results - high effort with poor outcomes creates maximum stress, while high effort with good outcomes actually reduces stress

The data suggests that academic performance may be both a cause and consequence of broader life stability, creating either virtuous or vicious cycles depending on a student's starting circumstances'.

## **-- Impact**

### **Potential Positive Impacts:**

- Reducing Achievement Stigma: The findings could help destigmatize academic excellence by showing that high performance doesn't necessarily come with a mental health cost, potentially encouraging students to pursue their academic goals without fear of inevitable stress.
- Early Intervention Frameworks: Identifying that stress clusters around academic struggle and life instability could enable schools to develop predictive models for at-risk students, allowing for proactive support before problems compound.

- Resource Allocation Insights: Understanding that low performers face multiple simultaneous stressors (basic needs, safety, peer pressure) while high performers mainly worry about future concerns could help institutions allocate counseling and support resources more effectively to where they're most needed.

### **Potential Negative Impacts:**

- Oversimplification Risk: These findings might be misinterpreted to suggest that struggling students are simply "not trying hard enough" or that academic pressure is never a concern. This could lead to reduced empathy for students experiencing genuine academic stress.

- Individual Variation Ignored: By focusing on group averages, this analysis might overlook individual high achievers who do experience significant stress, potentially making them feel unseen or invalidated.

- Financial Background Missing: This data provides no data concerning financial situations nor their social class. This would provide even greater, significant insights that help to explain much of the other data in this set. It tells nothing about what resources students have been provided growing up and it leaves significant holes in telling the story of the extent that certain students may have faced in, not only their upbringing, but also currently while in college.

- Causal Assumption Errors: The correlational data might be misinterpreted as proving that academic success causes reduced stress, when the relationship could be reversed (low stress enables better performance) or driven by unmeasured factors like socioeconomic status. This could lead to misguided interventions that focus on boosting grades rather than addressing underlying stressors.

### **-- Visualization Chart Stories:**

### **Visualization #1: Box Plot of Academic Performance vs. Stress Level:**

Figure 1 below looks at the distribution of stress levels across different academic performance levels using a box plot visualization. This chart examines the relationship between academic achievement and student stress by showing the median, quartiles, and outliers for stress levels within each performance category. After analyzing this plot, several conclusions emerged including that higher performing students (levels 4-5) consistently show lower median stress levels, while lower performing students (levels 0-2) display higher stress with greater variability. This addresses the core research question about whether high achievers are more stressed and clearly demonstrates that academic success correlates with reduced rather than increased stress levels. The visualization reveals the inverse relationship between performance and stress across the entire 1,100-student dataset.

### **Visualization #2: Bar Chart of Average Stress Level by Academic Performance:**

Figure 2 is a bar chart displaying the average stress level for each academic performance level, which answers the question of the specific stress patterns associated with different levels of academic achievement. A bar chart effectively shows the magnitude of differences between categorical groups by displaying mean values. After examining this plot, the conclusions drawn include that average stress decreases consistently as academic performance increases, with the highest stress levels occurring among the lowest performers and the lowest stress levels among top achievers. This graph displays a clear downward trend that challenges common assumptions about academic pressure and provides quantitative evidence for the protective effect of academic success against stress.

### **Visualization #3: Correlation Heatmap of Key Variables:**

Figure 3 below is a correlation heatmap showing the relationships between stress level and key psychological, academic, and life circumstance variables including self-esteem, social support, sleep quality, safety, basic needs, and future career concerns. This visualization uses color coding to display correlation strengths and directions across the 1,100-student dataset. After analyzing

this heatmap, conclusions include that future career concerns show the strongest positive correlation with stress (0.74), while self-esteem (-0.76) and sleep quality (-0.75) serve as the strongest protective factors. The academic performance correlation with stress (-0.72) is significant but positioned within a broader ecosystem of psychological and life factors, revealing that stress operates through multiple interconnected pathways rather than academic pressure alone.

#### **Visualization #4: Scatter Plot of Study Load vs. Performance by Stress Level:**

Figure 4 below is a scatter plot examining the relationship between study effort and academic outcomes, with stress level indicated by color intensity. This visualization displays each of the 1,100 students as individual points to show the distribution of effort-reward combinations and their associated stress levels. After examining this plot, key findings include that students with high study loads but low performance (bottom-right quadrant) show the highest stress levels, while students achieving high performance regardless of effort level display consistently low stress. This addresses the research question about academic pressure cycles by revealing that stress emerges from ineffective effort rather than high workload, and demonstrates that successful students maintain low stress even when working intensively.

#### **Visualization #5: Profile Comparison of Stressed vs. Calm High Achievers:**

Figure 5 is a comparative bar chart analyzing the 5.4% of high-performing students who experience high stress against the 88.9% who remain calm, examining variables including study load, future career concerns, peer pressure, self-esteem, social support, and sleep quality. This visualization answers the question of what differentiates successful students who struggle with stress from those who don't. After analyzing this comparison, conclusions include that stressed high achievers are characterized by significantly higher future career concerns and lower self-esteem despite their academic success, while calm achievers demonstrate superior psychological resilience through higher confidence and better support systems. This reveals that sustainable high performance requires psychological resources beyond academic capability alone.

### **Visualization #6: Stacked Bar Chart of Stress Factor Composition by Performance:**

Figure 6 below is a stacked bar chart showing how different stress factors (study load, future career concerns, peer pressure, safety, basic needs, living conditions) contribute to the total stress burden across academic performance levels. This visualization displays the cumulative stress factor scores for each performance level while breaking down the composition of stressors affecting different student groups. After examining this chart, conclusions include that low-performing students face "survival stress" with multiple simultaneous challenges including peer pressure, basic needs insecurity, and safety concerns, while high-performing students experience streamlined stress focused primarily on future career concerns. This demonstrates that academic struggle comes with compounding life stressors, while academic success provides protection from most stress sources except future-oriented anxiety.