BUDT730: Data Models and Decisions
Sleep Helps, Stress Decides:
The Real Driver of Student Success

by

Ioannis Mimigiannis

Vishesh Khare

Kunal Roy Chowdhury

Chia-Hsuan Chang

Summary: This report explores the relationship between student well-being and academic performance by analysing how lifestyle choices, particularly sleep duration, affect stress, depression probability, and CGPA outcomes. Additionally, the study investigates how external factors such as city of residence and degree type influence academic pressure among students. Using a university student dataset sourced from Kaggle, the analysis evaluates the "sweet spot" for sleep duration and identifies broader demographic drivers of stress.

Key findings show that students sleeping 7–8 hours per night achieve the highest average CGPA (7.71/10) and share the lowest reported academic pressure (3.11/5)**.** However, when controlling for all variables through regression modelling, Academic Pressure emerges as the strongest predictor of academic performance ($p < 0.001$), outweighing the direct impact of sleep or study hours. Stress is also shown to strongly correlate with poor mental health outcomes, as the probability of depression rises from 20% at low pressure levels to 86% at maximum pressure.

Finally, the analysis identifies significant differences across demographic segments. Students in large metropolitan cities such as Bangalore and Delhi report substantially higher stress than those in smaller cities. Stress levels also vary across educational stages, with Class 12 students experiencing the highest pressure, while postgraduate students, particularly in M.Ed and M.Tech programs, report the lowest. Although these patterns are statistically significant, they explain only a small portion of the overall variation in stress, suggesting that individual-level and psychological factors also play a central role.

Overall, the evidence shows that student success is shaped less by lifestyle habits alone and more by the broader academic and emotional environment. Universities seeking to improve student performance and well-being should prioritize stress reduction, develop targeted interventions for high-risk groups, and integrate mental-health resources into the academic experience.

Note: While the original scope included additional lifestyle factors (diet, financial stress, and family history of mental illness), the final analysis focused on variables with sufficient data completeness and direct relevance to academic performance and mental health outcomes.

Business Problem 1: Which is the "Sweet Spot" that correlates with the best student outcomes (higher CGPA, lower Academic Pressure).

We conducted statistical analysis to determine the sleep duration that correlates with the highest CGPA and lowest Academic Pressure. We used as values the CGPA (on a scale of 10) and Academic Pressure (on a scale 1-5), grouped by the sleep duration (Less than 5 hours, 5-6 hours, 7-8 hours, More than 8 hours).

Row Labels	Average of CGPA	Average of Academic_Pressure
'7-8 hours'	7,71	3,11
'More than 8 hours'	7,68	3,19
'5-6 hours'	7,65	3,10
'Less than 5 hours'	7,65	3,22
Grand Total	**7,67**	**3,16**

Data Summary:

- **< 5 Hours:** CGPA **7.65** | Pressure **3.22** (Worst Pressure)
- **5-6 Hours:** CGPA **7.65** | Pressure **3.10**
- **7-8 Hours:** CGPA **7.71** | Pressure **3.11** (Best Outcome)
- **> 8 Hours:** CGPA **7.68** | Pressure **3.19**

Key Insights:

1. **(7-8 Hours):** This group achieves the highest average CGPA (7.71) and has one of the lowest Academic Pressure scores (3.11). This confirms the "sweet spot" hypothesis.
2. **<5 Hours:** Students sleeping less than 5 hours report the highest academic pressure (3.22) and the lowest CGPA (7.65), indicating a direct negative impact on both wellbeing and grades.
3. **>8 Hours:** Students sleeping more than 8 hours see a slight drop in CGPA (to 7.68) and a rise in pressure (to 3.19), indicating that balance is required in the sleep duration.

Regression Analysis:

To further examine the relationship of those variables, we conducted a multiple linear regression model, using CGPA as the dependent variable and sleep duration, academic pressure, and study hours as predictors. While descriptive statistics revealed noticeable patterns, particularly the apparent “Sweet Spot” of 7-8 hours of sleep leading to higher academic performance, such visual relationships may be influenced by multiple overlapping factors.

Regression Model:

CGPA = 8.01 − 0.085 × AcademicPressure + 0.056 × Sleep(7–8h) − 0.013 × StudyHours

1. **The Intercept** (8.01)**:** A student who doesn’t sleep 7-8 hours, has 0 Academic Pressure and studies 0 hours would have a predicted CGPA of 8.01.
2. **Academic Pressure** (-0.085) (Highly Significant)
    - **Significance:** The p-value = 0.0003 mean this is highly statistically significant.

- **Impact:** For every 1-point increase in Academic Pressure (on the 1-5 scale), a student's CGPA drops by 0.085 points. In plain English**,** a student moving from "Low Pressure" (1) to "High Pressure" (5) loses roughly 0.34 (0.085 × 4) GPA points, which is the difference between an A- and a B+.

3. **Sleep Duration** (Not Significant)
    - **7-8 hours (+0.056):** Students who sleep 7–8 hours have a CGPA that is 0.056 points higher than those who don't. This aligns with our "Sweet Spot" chart, but the p-value (0.54) means it is not statistically significant.
    - **< 5 hours and > 8 hours:** These coefficients are tiny and not significant.

    While 7-8 hours is visually better, the regression suggests that sleep likely affects CGPA indirectly rather than directly.

4. **Study Hours** (-0.013) (Not Significant)
    - It suggests that for every extra hour of study, CGPA drops slightly. This is likely reverse causality. Students who are struggling might study more to catch up. In addition, it is not significant (p=0.13), so we shouldn't rely too much on it.

Model Fit:

1. **Multiple R-squared:** 0.009 (approx. 1%)
2. **Meaning:** This model explains only 1% of the variation in students' CGPA.

Section Summary:

Predicting human grades is difficult because variables like Intelligence, Major Difficulty, and Attendance are missing from the dataset. While our model identifies Pressure as a significant negative factor, the low R-squared indicates that CGPA is influenced by many other unobserved variables (e.g., innate ability, class difficulty). Our regression analysis revealed that Academic Pressure is the strongest predictor of academic performance ($p < 0.001$).

While descriptive statistics showed a slight "Sweet Spot" at 7-8 hours of sleep, the regression analysis suggests that simply sleeping more does not guarantee higher grades if Academic Pressure remains high.

Therefore, interventions should focus on Mental Health and Stress Reduction rather than just enforcing sleep curfews.

Business Problem 2: Is there a direct correlation between Academic Pressure and the likelihood of Depression? If so, at what pressure level does the risk spike?

We analysed the Academic Pressure (rated 1-5) and Depression columns (Binary: 0=No, 1=Yes) in the pivot table.

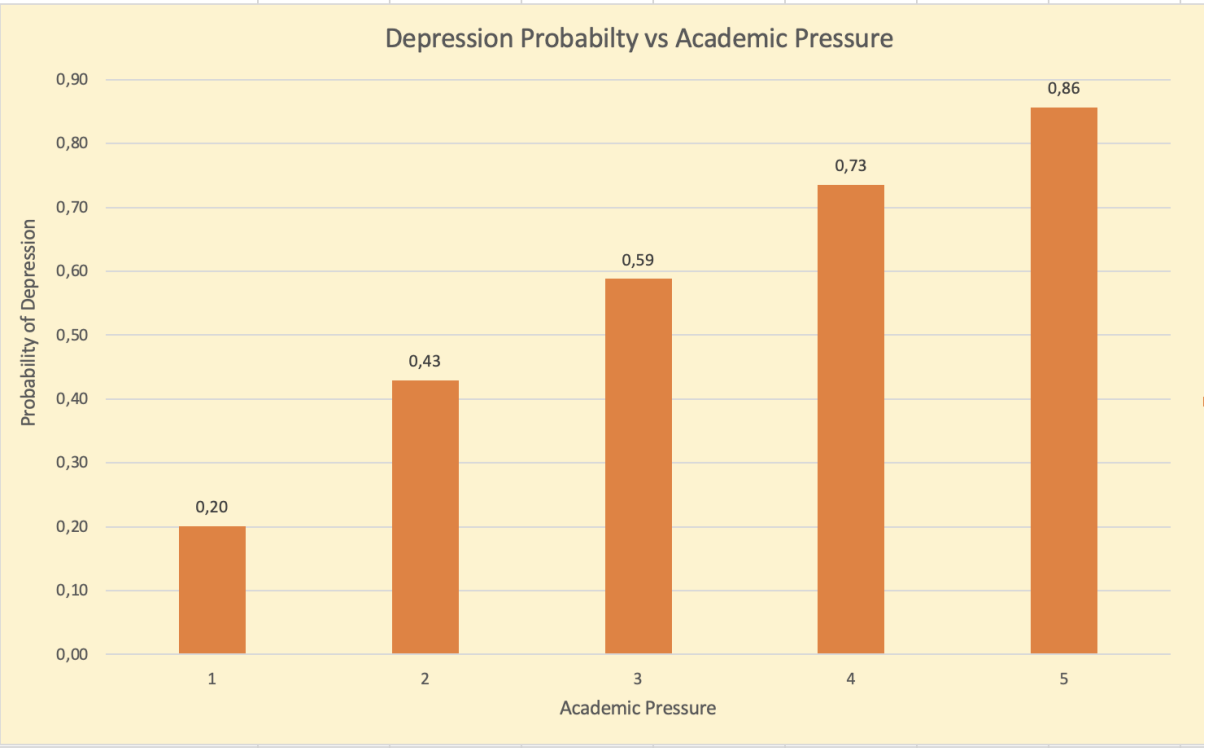
- **Correlation:** A Spearman correlation analysis shows a correlation of 0.45, indicating a moderately strong positive association, meaning that higher academic pressure is linked to higher depression rates.

Data Summary:

1. **Pressure Level 1 (Low):** 20% probability of depression.
2. **Pressure Level 2:** 43% probability of depression.
3. **Pressure Level 3 (Moderate):** 59% probability of depression.

4. **Pressure Level 4:** 73% probability of depression.
5. **Pressure Level 5 (High):** 86% probability of depression.

The pivot chart below, illustrates this steep climb:



Key Insights:

1. **(Level 1-2-3):** There is observed a clear upward trend in the Probability of Depression across the first 3 levels of Academic Pressure. The percentage rises from 20% at Level 1 to 43% at Level 2, providing warning signs that even at relatively low-pressure levels, almost 50% of the students report that they experience depression symptoms. Once students report "Moderate" pressure (Level 3), 59% report depression. This suggests that universities should intervene before pressure hits this middle ground, not just at the extremes.

2. **(Level 5):** At the highest-pressure level (Level 5), depression is almost a guarantee (86%), making severe Academic Pressure a strong indicator of mental health risk. This is a powerful statistic to support any recommendation for "academic forgiveness" policies or mandatory wellness days during finals.

3. **Steady escalation pattern:** Depression risk grows steadily as Academic Pressure rises, with each step up in academic pressure associated with a roughly 15–20 percentage point increase in the likelihood of depression, indicating a clear association of higher Academic Pressure and greater Depression Risk.

Regression Analysis of Academic Pressure and Depression:

Academic Pressure=2.39 + factor(Depression)1* 1.29

1. **Coefficient for factor(Depression)1 = 1.29116:** Students who report depression have Academic Pressure scores that are 1.29 points higher than those who do not report

depression. This is a large effect size given the 1–5 scale of the variable. It shows a substantial jump in stress levels when depression is present.

2. **p-value < 2e-16 (highly significant):** Depression variable is highly statistically significant.
3. **R-squared = 0.2072 and Adjusted R-squared = 0.2068**
    - Depression explains around 21% of the variation in Academic Pressure, consisting of a strong effect for a single predictor.
    - Non-depressed students have an average pressure score of 2.40, while depressed students average around 3.69.
    - This indicates that mental health is a major driver of academic stress, far more influential than sleep duration or city.

Section Summary:

The overall model is statistically significant, and depression is a strong predictor of Academic Pressure.

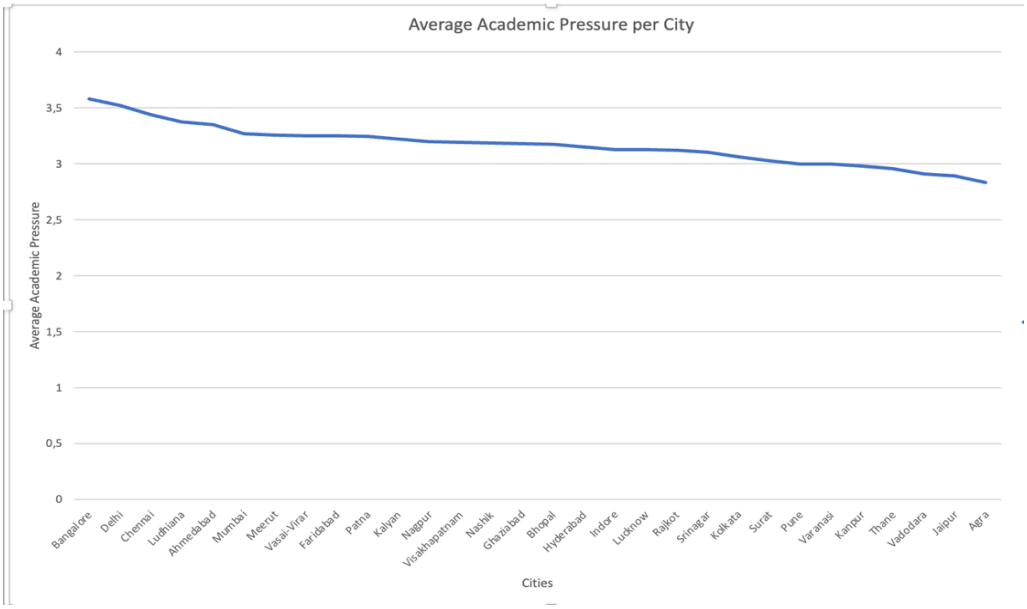
Business Problem 3: How Academic Pressure varies across different Cities and Degrees? What are the most difficult from student perception majors?

1. **City**

    We analysed the average Academic Pressure across 30 cities in a pivot table. The data reveals a clear trend: students in major metropolitan hubs report significantly higher stress levels than those in smaller cities.

    a. **The Stress Capitals:**
        - Bangalore: The most stressful city (Avg Pressure: 3.58/5).
        - Delhi: Close second (Avg Pressure: 3.52/5).
        - Chennai: Third place (Avg Pressure: 3.44/5).

    b. **The "Chill" Cities:**
        - Agra: Lowest stress (Avg Pressure: 2.84/5).
        - Jaipur: Second lowest (Avg Pressure: 2.89/5).
        - Vadodara: Third lowest (Avg Pressure: 2.91/5).



It appears the fast-paced lifestyle of these tech and educational hubs correlates with higher academic pressure.

Regression Models:

a. **The Academic Pressure-City Model Overview:**
   Academic Pressure = 2.835 + 0.517 × Ahmedabad + 0.746 × Bangalore + 0.605 × Chennai + 0.686 × Delhi + 0.541 × Ludhiana + 0.371 × VasaiVirar
   We analysed the average Academic Pressure across 30 cities. The data reveals a clear trend: students in major metropolitan hubs report significantly higher stress levels than those in smaller cities**.** Based on the regression analysis, only Ahmedabad, Bangalore, Chennai, Delhi, Ludhiana, and Vasai-Virar are statistically significant.
b. **The "High Stress" Zones (Statistically Significant):**
   - Bangalore (+0.75, p=0.002): The most stressful city in the dataset. A student in Bangalore is predicted to have pressure 0.75 points higher than one in Agra.
   - Delhi (+0.69, p=0.009): The second most stressful.
   - Chennai (+0.60, p=0.019): Third most stressful.
   - Ahmedabad (+0.52, p=0.029): Significantly stressful.
   - Ludhiana (+0.50, p=0.023): Surprisingly high stress.

   Note: The number in parentheses represents the parameter estimate and the p-value for the variable.
c. **The "Low Stress" Zones (Not Significant):**
   - Cities like Vadodara, Jaipur, and Kanpur have p-values > 0.05. This means their stress levels are statistically indistinguishable from Agra (the low-stress baseline).

Key Insights:

Urbanization is a major driver of student stress. The "Tech Hubs" (Bangalore, Chennai) and the Capital (Delhi) are significantly harder on students than Tier-2 cities.

2. Degree

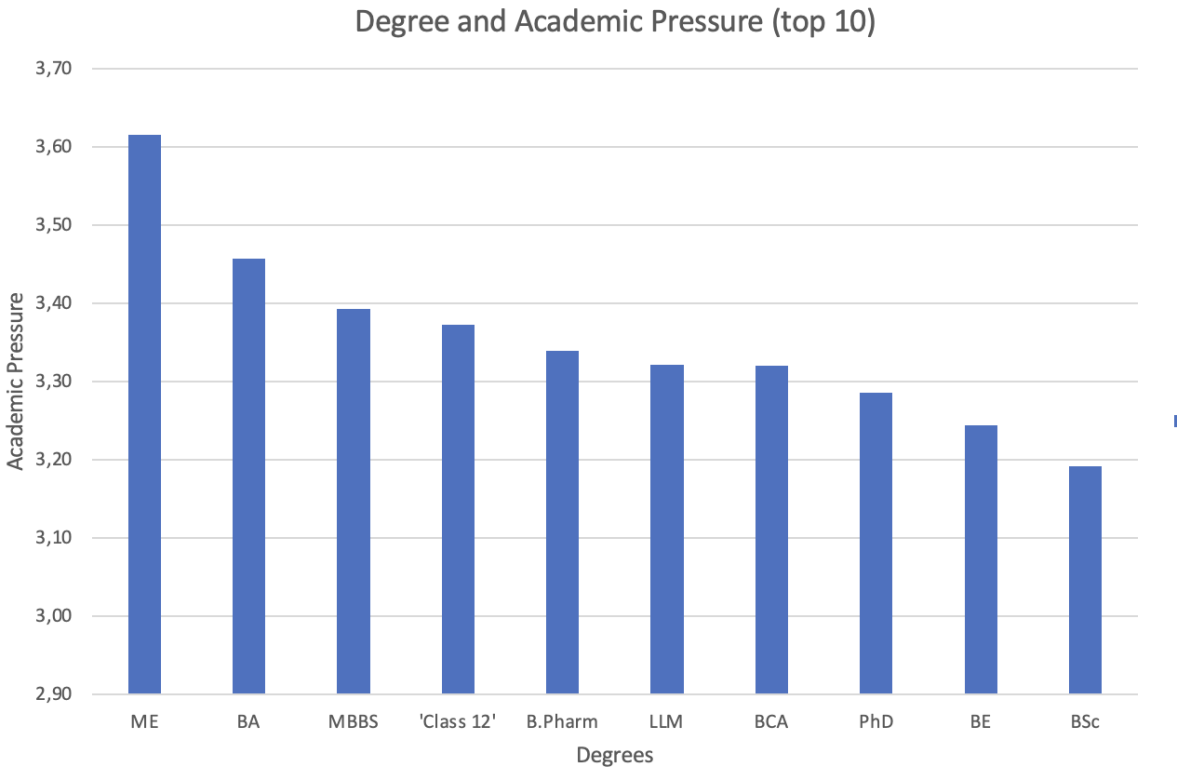
Since almost 100% of the dataset labelled their profession as "Student," we analysed the Highest Qualification column to understand how stress varies by degree type.

Visualization: This bar

a. **The "High Pressure" Degrees:**
   - ME (Master of Engineering): Ranked #1 (Avg Pressure: 3.62/5).
   - BA (Bachelor of Arts): Surprisingly ranked #2 (Avg Pressure: 3.46/5).
   - MBBS (Medical): Ranked #3 (Avg Pressure: 3.39/5), which is expected given the rigor of medical school.
   - Class 12: High stress (Avg Pressure: 3.37/5), likely due to the intense pressure of college entrance exams.
b. **The "Low Pressure" Degrees:**
   - M.Ed (Master of Education): One of the least stressful (Avg Pressure: 2.84/5).

- M.Tech: Interestingly, while ME was the highest, M.Tech is among the lowest (Avg Pressure: 2.86/5). This might suggest differences in curriculum intensity or job market security between the two engineering masters.
- M.Com & MA: Both report lower-than-average stress.

Visualization: This bar chart highlights the top 10 degrees with the highest reported academic pressure.



Regression Models:

The Academic Pressure-Degree Model Overview

Model: Academic Pressure = 3.372 − 0.333·BArch − 0.390·BCom − 0.515·MCom − 0.536·MEd − 0.513·MTech − 0.461·MA − 0.440·MCA − 0.407·MSc

Based on the regression analysis, only B.Arch, B.Com, M.Com, M.Ed, M.Tech, MA, MCA, and MSc are statistically significant. While descriptive statistics show that ME and BA appear among the most stressful degrees, the regression analysis indicates that these differences are not statistically significant once we control for the baseline group. Therefore, Class 12 clearly emerges as the most stressful academic stage.

1. **Master's degrees:**
    - **M.Com (-0.515)**: Students pursuing Master of Commerce report substantially lower stress levels.
    - **M.Ed (-0.536)**: Master of Education students experience the largest reduction in stress.
    - **M.Tech (-0.513)**: Despite being in a technical/engineering pathway, Master's-level students feel significantly less pressure than high school students.

- **MA (-0.461):** Master of Arts students show a clear decrease in reported academic pressure.
- **MCA (-0.440)**: Master of Computer Applications students also report meaningfully lower stress.
- **MSc (-0.407)** – Master of Science students show a consistent downward shift in academic pressure.

2. **Undergrad Degrees:**
    - **B.Com** (-0.390) & **B.Arch** (-0.333): Less stressful than Class 12.

Note: The number in parentheses represents the parameter estimate for the variable.

Key Insights:

"Class 12" students (baseline group) experience the highest pressure—likely due to entrance exams. Stress declines as students enter college, and decreases further in Master's programs.

Model Limit:

Adjusted $R^2 \approx 0.006$, meaning cities and degree types explain only a small fraction of stress. Most academic stress is driven by personal and psychological factors.

Section Summary:

While the specific cities and degrees are statistically significant, they don't explain the whole picture. Personal factors (family history, financial stress) likely make up the rest**.** Academic pressure varies significantly by both geographic location and educational level. Students in major metropolitan hubs such as Bangalore and Delhi report substantially higher stress than those in smaller cities like Agra, indicating that urban academic environments intensify competition and psychological strain. Stress levels also differ across qualifications: Class 12 students represent the peak stress group, likely due to competitive entrance exams, whereas postgraduate students (especially M.Ed and M.Tech**)** report much lower pressure. Regression results confirm these patterns, identifying urban location and early educational stages as key stress amplifiers.

Conclusion

Across all analyses, student well-being emerges as a decisive factor shaping academic performance. Although sleeping 7–8 hours is associated with the most favourable outcomes in descriptive statistics, our regression results make clear that academic pressure, rather than lifestyle habits alone, is the dominant force influencing CGPA. High pressure not only suppresses academic performance but also dramatically increases the likelihood of depression, with risk levels escalating steadily across the pressure spectrum and reaching critical severity at the top end.

The data further demonstrates that stress is not experienced uniformly. Urban academic environments and early educational stages (particularly Class 12) act as key amplifiers of pressure, whereas postgraduate programs, especially M.Ed and M.Tech, are associated with significantly lower stress levels.

Overall, the evidence suggests that improving student outcomes requires more than promoting healthy habits such as adequate sleep. Institutions must focus on mitigating excessive academic pressure, particularly in competitive urban settings and during transitional stages such as Class 12. Strategic initiatives, such as workload redesign, supportive academic policies, accessible counselling services, and structured mental wellness programs, are not only beneficial for student health but are also essential for sustaining academic success and enhancing institutional reputation.

Limitations

This study has several limitations that should be considered when interpreting the results. First, the data are self-reported, which means that some variables, such as sleep duration, academic pressure, and depression, may not be measured perfectly and could be affected by personal bias or inaccurate reporting. Second, the dataset is cross-sectional, capturing student responses at only one point in time. Additionally, several important factors that could influence CGPA and stress levels are not included in the dataset, such as course difficulty, attendance, family background or individual personality traits. Also, Some lifestyle variables (diet, financial stress, and family history of mental illness) were not included in the final models due to incomplete data and limited relevance to the core research questions. Finally, because the sample mainly reflects students studying in India, the results may not fully generalize to students in other countries or educational systems.

AI usage: ChatGPT was used exclusively to support the writing process by organizing the report's structure (proposing business questions, the regression analysis presentation format etc.). All statistical calculations, data analysis and data transformations, pivot tables, visualizations, and regression analyses were performed independently using Excel and R.