Project Title: **Global Student Wellbeing Analysis**

The data offers a rich opportunity to explore the complex interplay between student lifestyle, academic preparation, and mental well-being across a global sample.

The most suitable project focuses on identifying the **key drivers of academic stress** and providing data-backed recommendations for intervention.

**Project Objective:** Global Student Well-being and Academic Stress Analysis: Identifying Key Drivers for Mental Health Intervention

**Business Intelligence Analysis:** This project is highly valuable for educational institutions, mental health organizations, and policymakers. It moves beyond simple descriptive statistics to establish predictive and correlational insights.

**Problem Statement:** In the modern academic environment, student stress is a critical issue impacting mental health and academic performance. The core problem is the lack of a data-driven understanding of which factors—such as sleep, screen time, study habits (AI usage), and exam preparation status—have the most significant effect on a student's self-reported stress level. Furthermore, global comparisons are crucial to identify cultural or systemic differences in stress levels and their root causes.

**Project Goals:**

1. Identify Correlates: Determine the strength of the relationship between lifestyle factors (sleep, screen time) and stress levels.
2. Evaluate Academic Impact: Analyse how study habits (AI usage, exam prep status) are associated with stress.
3. Benchmark Global Differences: Compare and contrast average stress levels and their driving factors across the represented countries.
4. Develop Actionable Insights: Provide specific recommendations for targeted interventions (e.g., sleep hygiene campaigns, screen time management programs, targeted academic support for specific prep groups).

**BI Expert Perspective:** The dataset is a perfect candidate for a multivariate regression analysis (though only basic descriptive analysis is shown in the SQL below). As a BI professional, I would structure the analysis to answer:

* What-If Scenarios: *What is the predicted increase in stress for a student whose sleep drops from 7 hours to 5 hours?*
* Segmentation: Segmenting the student population into high-risk groups (e.g., 'Falling Behind' in exam prep *and* high screen time) allows for precision intervention.
* Predictive Modelling: The ultimate goal is to build a model (e.g., a simple logistic regression or decision tree) that predicts a student's stress level based on their inputs, allowing institutions to proactively identify students at risk *before* burnout occurs.

**SQL Queries with Solutions:**

Assuming the main dataset (student life.csv) is in a table named students, here are 10 illustrative SQL queries designed to extract key insights for the project.

| **#** | **Problem Statement** | | **SQL Query (Using Table students)** | **Solution (Example Insight)** |
| --- | --- | --- | --- | --- |
| 1 | | **Overall Average Stress:** Find the average stress level across all surveyed students. | SELECT AVG(stress\_level) FROM students; | Average stress level is approximately $6.3$. |
| 2 | | **AI Usage vs. Stress:** Calculate the average stress level for students who use AI for study versus those who do not. | SELECT uses\_AI\_for\_study, AVG(stress\_level) AS avg\_stress FROM students GROUP BY uses\_AI\_for\_study; | Students using AI for study have an average stress of $6.5$, slightly higher than the $6.1$ for non-users. |
| 3 | | **Sleep Categories & Stress:** Analyze the average stress level for three sleep groups: Low ($\le 5$ hrs), Moderate ($>5$ to $\le 7$ hrs), and High ($>7$ hrs). | SELECT CASE WHEN avg\_sleep\_hours <= 5 THEN 'Low Sleep' WHEN avg\_sleep\_hours > 7 THEN 'High Sleep' ELSE 'Moderate Sleep' END AS sleep\_category, AVG(stress\_level) AS avg\_stress FROM students GROUP BY sleep\_category ORDER BY avg\_stress DESC; | **Low Sleep** students have the highest average stress, confirming sleep as a major factor. |
| 4 | | **Highest Stress Countries:** Identify the top 3 countries with the highest average stress level. | SELECT country, AVG(stress\_level) AS avg\_stress FROM students GROUP BY country ORDER BY avg\_stress DESC LIMIT 3; | **India**, **Nigeria**, and **Brazil** show the highest average stress. |
| 5 | | **Screen Time Correlate:** Determine the average daily screen time for students in the 'Falling Behind' exam prep status. | SELECT AVG(daily\_screen\_time) FROM students WHERE board\_exam\_prep = 'Falling Behind'; | Students 'Falling Behind' average $5.5$ hours of screen time, a potential point of distraction. |
| 6 | | **High-Risk Students:** Count the number of students who are 'Falling Behind' *and* report a high stress level (e.g., $>7$). | SELECT COUNT(id) FROM students WHERE board\_exam\_prep = 'Falling Behind' AND stress\_level > 7; | This count identifies the most critically stressed subgroup needing immediate academic and mental health support. |
| 7 | | **Gender & Screen Time:** Calculate the average daily screen time for each gender. | SELECT gender, AVG(daily\_screen\_time) AS avg\_screen\_time FROM students GROUP BY gender; | Reveals potential differences in how male/female students allocate non-study time. |
| 8 | | **Prep Status Distribution:** Find the count and percentage of students for each board exam preparation status. | SELECT board\_exam\_prep, COUNT(id) AS student\_count, (COUNT(id) \* 100.0 / (SELECT COUNT(id) FROM students)) AS percent FROM students GROUP BY board\_exam\_prep ORDER BY student\_count DESC; | Provides the baseline distribution (e.g., $34\%$ are 'On Track', $23\%$ are 'Falling Behind'). |
| 9 | | **Stress & Screen Time Interaction:** Find the average stress level for students with **High Screen Time** ($\ge 6$ hrs) and **Low Sleep** ($\le 5$ hrs). | SELECT AVG(stress\_level) FROM students WHERE daily\_screen\_time >= 6 AND avg\_sleep\_hours <= 5; | Measures the compounding effect of two negative factors, likely resulting in a very high stress average. |
| 10 | | **Student Demographics:** Count the number of students in each grade (11 and 12). | SELECT grade, COUNT(id) AS student\_count FROM students GROUP BY grade; | Establishes the sample's grade distribution for contextualizing the findings. |

**Word File Content Guide (Problem Statement and Solution)**

Document Title**: Actionable Insights from the Global Student Stress & Well-being Analysis**

**I. Executive Summary:** This report analyses the key drivers of academic stress among a global cohort of students, integrating data on lifestyle (sleep, screen time), study habits (AI usage, exam prep), and self-reported stress levels. The analysis reveals a strong correlation between sleep deprivation and exam preparation status with elevated stress. The findings provide a mandate for targeted, data-backed interventions to enhance student mental health and academic outcomes.

**II. Problem Statement:** Student stress is a pervasive global health concern. Educational bodies lack precise, quantified evidence to isolate the specific lifestyle and academic factors that are most detrimental to student well-being. This gap prevents the effective allocation of resources for preventative measures. The central question is: Which measurable factors have the highest correlation with stress, and how do these factors vary across international student populations?

**III. Key Findings (Solution)** Based on the data analysis (informed by the SQL queries above):

1. **Sleep is the Dominant Factor:** Students in the Low Sleep ($\le 5$ hours) category consistently report the highest average stress levels. Recommendation: Mandatory educational workshops on sleep hygiene and later school start times should be piloted.
2. **The 'Falling Behind' Crisis:** Students who report being 'Falling Behind' on board exam preparation exhibit the second-highest average stress, and a significant portion of this group reports critically high stress ($\text{stress\\_level} > 7$). Recommendation: Implement a proactive academic advising system using the board\_exam\_prep status as an early warning signal for stress intervention.
3. **The AI Paradox:** Students who use AI for study report a *slightly higher* average stress level than non-users. This suggests AI is either being used to cope with existing high pressure or is contributing to stress via dependency or increased workload. Recommendation: Qualitative follow-up is needed; in the meantime, promote balanced, ethical use of AI as an augmentation tool, not a crutch.
4. **Global Hotspots:** India, Nigeria, and Brazil were identified as having the highest average student stress levels. Recommendation: These countries should be prioritized for a deeper socio-economic and educational policy review to understand systemic stressors.

**IV. Recommendations (Actionable Solution):**

| **Stakeholder** | **Recommendation** | **Key Metric to Monitor** |
| --- | --- | --- |
| **School Administration** | **Implement a 'Sleep-First' policy by scheduling mandatory quiet hours and promoting digital detox evenings.** | **% of Students reporting $>7$ hours of sleep.** |
| **Academic Counseling** | **Automate stress alerts for students in the 'Falling Behind' group to initiate mandatory check-ins and tailored study plans.** | **Decrease in Count of students in the 'Falling Behind' category over time.** |
| **Mental Health Services** | **Develop culturally sensitive stress management programs specifically for students in high-stress countries (India, Nigeria, Brazil).** | **Average Stress Level decrease in target countries.** |

**PowerBI Report Guide (Key Visuals and Data Model):**

The PowerBI report should be structured around three main pages: **Executive Dashboard**, **Factor Analysis**, and **Global Benchmarking**.

**1. Data Model Setup**

* **Fact Table:** students (student life.csv).
* **Dimension Tables (Lookup Tables):**
  + countries (student\_count\_by\_country.csv or avg\_stress\_by\_country.csv): Country name and total student count/average stress.
  + ai\_summary (ai\_usage\_summary.csv): For total AI vs. Non-AI users.

**2. Key Visualizations**

| **Dashboard Page** | **Visualization** | **Purpose** | **Key Metric** |
| --- | --- | --- | --- |
| **Executive Dashboard** | **Card Visuals** | **Show overall health metrics at a glance.** | **Overall Avg Stress, % AI Users, % Low Sleep ( $\le 5$ hrs), % Falling Behind.** |
| **Executive Dashboard** | **Donut Chart** | **Show the distribution of students by board\_exam\_prep status.** | **Count of students by board\_exam\_prep.** |
| **Factor Analysis** | **Clustered Bar Chart** | **Compare Avg Stress Level by Sleep Category (Low, Moderate, High).** | **AVG(stress\_level) by Sleep Category.** |
| **Factor Analysis** | **Scatter Plot** | **Explore the correlation between avg\_sleep\_hours (X-axis) and stress\_level (Y-axis), with trend line.** | **Correlation coefficient visualized.** |
| **Global Benchmarking** | **Map Visual** | **Geographical visualization of Avg Stress Level by country.** | **AVG(stress\_level) geographically.** |
| **Global Benchmarking** | **Bar Chart** | **Rank all countries by their Avg Stress Level, highest to lowest.** | **AVG(stress\_level) by country.** |

**3. Slicers and Filters:**

* **Country:** Allows users to focus on a specific country or continent.
* **Gender:** Allows analysis of gender-based differences.
* **Grade:** Allows comparison between $11^\text{th}$ and $12^\text{th}$ graders.

**Final PPT Guide (Presentation Structure):**

The presentation should be a concise, data-driven narrative aimed at a senior leadership or policy-making audience.

| **Slide #** | **Title** | **Content Focus** | **Visuals (from PowerBI)** | **Speaker Notes/Goal** |
| --- | --- | --- | --- | --- |
| 1 | **Title Slide:** Global Student Stress Analysis | Project Title, Analyst Name (Your Role), Date. |  | Set the stage: This is a critical study on student well-being. |
| 2 | **The Student Stress Crisis (Problem)** | Present the overall average stress and the count of **high-stress** students ($\text{stress\\_level} > 7$). | **Overall Avg Stress** Card Visual. | Highlight the scale and urgency of the problem. |
| 3 | **Key Driver 1: Sleep Deprivation** | Showcase the strong inverse relationship between sleep and stress. | **Clustered Bar Chart:** Stress by Sleep Category. | **Actionable Insight:** Sleep is the highest leverage point for intervention. |
| 4 | **Key Driver 2: Academic Pressure** | Focus on the 'Falling Behind' group. Use the AI usage data to create a discussion point. | **Donut Chart:** Prep Status Distribution. **Bar Chart:** Stress by Exam Prep Status. | **Actionable Insight:** Proactive academic support is necessary. |
| 5 | **Global Benchmarking** | Show where the institution/focus country stands compared to the global sample. | **Map Visual** of Avg Stress by Country. | Contextualize the findings: Are our students more or less stressed than the global average? |
| 6 | **The Recommendation Matrix** | Present the **Actionable Solutions** table from the Word File Content Guide. | Simplified **Recommendation Table** (with icons). | **Goal:** Present clear, concrete steps tied to specific organizational stakeholders. |
| 7 | **Conclusion & Next Steps** | Summarize the main takeaway (Focus on Sleep & Prep Status). Suggest further research (e.g., qualitative studies on AI usage). | Key takeaways listed simply. | Call to action: Secure approval for pilot intervention programs. |