STUDENTS MENTAL HEALTH REPORT

Tools Used

- Microsoft Excel – Initial data inspection, cleaning and structural validation
- Power Query (Power BI) – Data cleaning, transformation, and validation
- Power BI – Data modeling, DAX-driven calculations, and interactive dashboard development

ANALYSIS OBJECTIVE

Focus: Severity Segmentation & Academic-Level Comparison

Rather than asking "How many students report mental health conditions?", this analysis investigates:

Does psychological burden intensify as students advance academically?

The study was structured around three principles:

1. Cohort-based comparison (Year 1–4 as primary analytical lens)
2. Severity stratification (single vs overlapping conditions)
3. Proportional normalization (within-group prevalence rather than raw counts)

This approach prevents volume distortion and enables accurate cross-cohort comparison.

The objective was not merely descriptive reporting, but risk identification and intervention prioritization.

DATASET OVERVIEW

The raw dataset used was gotten from Kaggle. Here is the link;
https://www.kaggle.com/datasets/shariful07/student-mental-health

After cleaning and normalization, the dataset contains responses from 101 university students across the following variables:

- S/N
- Gender
- Age (18–24)

- Marital Status
- Year of Study (1–4)
- CGPA (Grouped: 0–1.99, 2.00–2.49, 2.50–2.99, 3.00–3.49, 3.50–4.00)
- Anxiety (Yes/No)
- Depression (Yes/No)
- Panic Attack (Yes/No)
- Specialist Treatment
- Any Condition
- Multiple Cases
- Complex Cases

Data Preparation & Validation

To ensure analytical reliability, I:

- Removed incomplete and blank records.
- Standardized categorical values for consistency.
- Validated CGPA group ranges.
- Removed Date/Time column (limited analytical contribution due to narrow window).
- Removed Course variable due to excessive categorical fragmentation and low signal value.
- Structured model for optimized filtering and cross-analysis.

These decisions prioritized analytical clarity and signal strength over dimensional noise.

Severity Engineering

Three derived classifications were created to measure escalation:

1) Any Condition – At least one reported condition
2) Multiple Cases – Two concurrent conditions
3) Complex Cases – All three conditions present

This stratification shifts the analysis from prevalence tracking to burden intensity measurement, enabling prioritization logic rather than uniform treatment of all cases.

Technical Approach

The model was designed around Year of Study as the primary slicer, reinforcing structured cohort comparison.

Key technical elements:

- DAX-based conditional logic for severity classification
- Context-aware prevalence calculations (percentage within academic year)
- Dynamic cross-filtering across demographics and CGPA ranges
- Relational model optimized for interpretability and performance

The deliberate use of proportional analysis prevents enrollment imbalance (Year 1 dominance) from distorting interpretation.

KEY INSIGHTS

1. Enrollment vs Burden Distortion:

Year 1 had the largest population (43 students), producing the highest raw case count.

However, proportional analysis reveals:

Year 1: 58.1% prevalence

Year 2: 73.1%

Year 3: 75.0%

Year 4: 25.0% (small sample caveat)

Insight: Psychological burden does not peak at entry level, it intensifies mid-progression, with Year 3 exhibiting the highest proportional prevalence.

Raw counts alone would have led to incorrect conclusions.

2. Escalation Pattern in Severity

Complex Cases (all three conditions) peak proportionally in Year 3.

This suggests:

- Accumulation of academic and transitional pressure.
- Potential compounding stress exposure over time.
- Escalation risk if early-stage intervention is absent.
- Year 1 reflects volume, but not peak intensity.

3. Academic Performance Lens

CGPA segmentation allows evaluation of performance burden interaction. While performance variation exists across ranges, the analysis indicates that psychological burden is not isolated to low-performing groups, suggesting stress exposure may cut across academic tiers.

This reinforces the need for universal screening rather than performance-based targeting alone.

4. Demographic Cross-Section

Gender, age, and marital status breakdowns reveal variation within cohorts, supporting the argument that mental health strategy should be multi-dimensional rather than being driven by a single variable.

6. Sample Size Integrity

Year 4 displays lower prevalence. However, interpretation is limited due to small sample size (n=8). Maintaining analytical caution prevents overgeneralization.

RECOMMENDATIONS AND CONCLUSION

Strategic Recommendations

1. Front-load prevention – Implement structured screening in Year 1 to prevent escalation into complex cases by Year 3.

2. Mid Progression Intervention Focus – Allocate additional counseling and resilience programs to Year 2 and Year 3 cohorts where proportional burden peaks.

3. Severity based Prioritization – Establish monitoring protocols for students classified under Complex Cases to enable early escalation control.

4. Universal Monitoring Model – Avoid targeting only low CGPA students as psychological burden appears cross-performance.

5. Longitudinal Tracking – Implement follow-up studies to validate progression trends across multiple academic cycles.

Conclusion

This analysis demonstrates that mental health burden is not static across academic stages.

While first-year students contribute the highest volume of reported cases, proportional and severity-adjusted analysis reveals that mid-level academic progression represents the highest risk zone for compounded psychological burden.

By integrating cohort comparison, severity engineering, and proportional normalization, this project moves beyond descriptive dashboarding into structured risk segmentation and intervention modeling.

It demonstrates analytical framing beyond raw metrics, feature engineering for decision support, context-aware DAX modeling, risk-based interpretation, and Institutional decision alignment.

This is not merely a visualization project; it is a cohort-based mental health risk analysis framework.