Student Mentoring Dataset – Deliverables

Objective: Design a realistic student dataset representing academic, wellness, productivity, and career dimensions.

1. Dataset Overview

The dataset 'students.csv' contains 80 synthetic student records created using realistic academic and behavioral patterns. It is suitable for profiling, exploratory analysis, clustering, and mentoring system design.

2. Data Dictionary

Column Name	Description	Scale / Range
student_id	Unique student identifier	Categorical
age	Student age	18–26
program	Degree / stream	Categorical
semester	Current semester	1–8
gpa	Academic performance	0–10
attendance	Attendance percentage	0–100
assignments_completion	% assignments completed	0–100
stress_level	Self-reported stress	1–10
sleep_hours	Average sleep per night	0–10
mental_wellbeing	Wellbeing score	1–10
productivity_score	Time management & productivity	1–10
distractions	Distraction level	1–10
career_clarity	Career goal clarity	1–10
skill_readiness	Job skill preparedness	1–10
engagement_score	Platform engagement score	0–100

3. Embedded Behavioral Patterns

• High stress → lower productivity and mental wellbeing.
• Low GPA + high engagement → strong attendance and motivation.
• High GPA + low career clarity → moderate skill readiness.
• Sleep, distractions, and productivity show soft correlations rather than perfect relationships.

4. Real-World Reflection

This dataset reflects realistic student behavior by mixing performance, wellbeing, and engagement variables. It captures struggling but motivated students, high-performing but confused students, and stressed low-productivity students. The structure supports mentoring systems, risk detection, and personalized academic guidance.