**🎓 2. Student Performance Dataset**

**Goal:** Predict final grade or identify at-risk students.

**📊 EDA Questions**

1. What is the distribution of final grades (G3)? Is it skewed?
2. How do parental education levels affect student grades?
3. Does study time correlate with better final grades?
4. Are students who take extra educational support performing better?
5. Use boxplots to compare grade distribution across schools (school) or gender.
6. Are alcohol consumption (Dalc, Walc) and absences negatively impacting performance?
7. How many students are failing (e.g., G3 < 10)? What are their common characteristics?

**🛠️ Feature Engineering Questions**

1. Create a total grade feature: G\_total = G1 + G2 + G3. Is it more predictive?
2. Bin G3 into categories: fail, pass, excellent.
3. Encode binary yes/no features like schoolsup, famsup, paid into 0/1.
4. Engineer a risk\_score = Dalc \* Walc \* absences. Does it highlight at-risk students?
5. Combine studytime and failures to create a resilience feature.
6. Create new interaction features: studytime \* absences, failures / studytime.