The dataset we'll be working with includes a wide range of factors influencing student performance. By analyzing this data, we'll be able to identify key drivers of success and provide insights that could help students, teachers, and policymakers make informed decisions. The table we'll use for this project is called **student\_performance** and includes the following data:

| **Column** | **Definition** | **Data type** |
| --- | --- | --- |
| attendance | Percentage of classes attended | float |
| extracurricular\_activities | Participation in extracurricular activities | varchar (Yes, No) |
| sleep\_hours | Average number of hours of sleep per night | float |
| tutoring\_sessions | Number of tutoring sessions attended per month | integer |
| teacher\_quality | Quality of the teachers | varchar (Low, Medium, High) |
| exam\_score | Final exam score | float |