**Students score dataset**

This dataset usually contains columns like:

* Hours Studied
* Previous Scores
* Extracurricular Activities (Yes/No)
* Sleep Hours
* Sample Question Papers Practiced
* Final Exam Score

From this dataset, students can **practice both data analysis and basic AI/ML**. Here are things they can extract or do:

**1. Data Understanding & Cleaning**

* Summarize dataset (mean, median, mode of scores, hours studied, sleep hours).
* Check missing values & data types.
* Create frequency tables (e.g., how many students did extracurricular).

**2. Exploratory Data Analysis (EDA)**

* Distribution of final scores.
* Correlation between hours studied & exam score.
* Do extracurricular activities improve scores?
* Does sleeping more/less affect performance?
* Do students who solve more sample papers score higher?
* Boxplots to see outliers in hours studied or sleep hours.

**3. Visualization Tasks**

* Histograms (scores distribution).
* Scatterplots (Hours studied vs Scores, Sleep hours’ vs Scores).
* Bar chart (Average score by extracurricular activity).
* Correlation heatmap.

**Hands-on practice**

* + Calculate mean study hours, mean score
  + Check correlation between hours studied & scores
  + Plot scatterplot and interpret correlation
  + Do students who study more hours get higher scores?
  + Does attendance play an important role in performance?
  + What’s the correlation between **study hours and marks**?

**Mini Task**

* + Each intern writes 3 insights about the dataset based on descriptive stats + correlation